

# Sommario

Prefazione		1
Dedica e riconoscimenti		4
Capitolo I.	<b>Cosmologie &amp; Cosmologia</b>	6
	1. L'uso delle scritte	7
	2. Pitagora da Samo	9
	3. L'eresia astrologica	11
	4. Dall'astronomia alla cosmologia	11
	5. Edwin Hubble	12
	6. Il flusso di Hubble	14
	7. La cosmologia e la fisica	18
	8. Antiche dispute, nuovi veleni	20
Capitolo II.	<b>Nel cielo le tracce dell'origine</b>	21
	1. Astronomia e cosmologia	21
	2. Il flusso di Hubble	22
	3. Una finestra sull'Universo	24
	4. Galassie, gruppi e ammassi	26
	5. Le scale cosmologiche	30
	6. Le distanze cosmiche	31
	7. Verso l'omogeneità	32
	8. La varianza di massa	35
	9. Campioni e cataloghi	36
	10. La relazione log N-log S	43
	11. Funzione a due punti e varianza di massa	44
	12. Funzioni a $n$ punti	50
	13. Lo spettro delle fluttuazioni	51
	14. I vuoti cosmici e la funzione di probabilità dei vuoti	54
	15. Funzioni di correlazione degli ammassi di galassie e segregazione in luminosità	57
	16. Uso di campioni angolari. Relazioni di scala	58
	17. Conclusioni	59

<i>Appendice A1.</i>	<i>Smussamento e filtri</i>	61
	1. I filtri a top-hat e Gaussiano	61
	2. Convenzione di somma sugli indici continui	62
<i>Appendice A2.</i>	<i>Distribuzioni stocastiche nel discreto e nel continuo</i>	64
	1. Processi e loro realizzazioni. Valor medio, varianza, momenti d'ordine $n$	64
	2. Distribuzioni Gaussiane	66
	3. Distribuzioni in uno spazio funzionale	67
	4. Valore di aspettazione, momenti, funzionale di partizione	68
	5. Varianza di massa	68
	6. Funzione di correlazione a 2 punti	69
	7. Funzione di correlazione a 2 punti e varianza di massa	69
	8. Funzione a 2 punti per processi Gaussiani	70
<i>Appendice A3.</i>	<i>Le funzioni a <math>n</math>-punti</i>	73
	1. La funzione a 4 punti	73
	2. Espressioni alternative delle funzioni a 3 e 4 punti	73
	3. Espressioni del modello gerarchico	75
	4. Funzioni a $n$ punti per processi Gaussiani e non Gaussiani	76
	5. Funzioni di correlazione come coefficienti dello sviluppo del funzionale di partizione	77
<i>Appendice A4.</i>	<i>Lo spettro delle fluttuazioni</i>	78
	1. Lo spettro	78
	2. Espressioni per la varianza di massa e la funzione a 2 punti	78
<i>Appendice A5.</i>	<i>Il bias</i>	81
<i>Appendice A6.</i>	<i>Profondità dei campioni e equazione di Limber</i>	84
Capitolo III.	<b>Le equazioni di Friedmann</b>	86
	1. Metrica e geometria	86
	2. Dalla geometria alla cinematica	91
	3. Il parametro di decelerazione	93
	4. Dalla cinematica alla dinamica	94
	5. Le equazioni di Friedmann	94
	6. Densità critica e parametro di densità	96
	7. Evoluzione del parametro di densità	98
	8. Il valore della curvatura	100
	9. La scala di Planck	100
<i>Appendice A1.</i>	<i>Dalle equazioni di Einstein alle equazioni di Friedmann</i>	103

<i>Appendice A2.</i>	<i>Altre forme delle equazioni di Friedmann</i>	106
<i>Appendice A3.</i>	<i>Espressioni della metrica FRW nei casi <math>\kappa = \pm 1</math></i>	107
Capitolo IV.	<b>La sostanza cosmica e l'espansione dell'universo</b>	108
	1. Le sorgenti dell'espansione	108
	2. Distribuzioni di particelle	109
	3. Entropia e numero nell'espansione cosmica	112
	4. Produzione di entropia nelle transizioni di fase	113
	5. Integrazione delle equazioni di Friedmann per $w$ costante	115
	6. Integrazione delle equazioni di Friedmann in presenza di piú componenti	120
	7. Orizzonti	121
	8. Cosmografia classica	123
<i>Appendice A1.</i>	<i>Distribuzioni statistiche: densità di energia e pressione</i>	126
<i>Appendice A2.</i>	<i>Entropia e altre funzioni termodinamiche</i>	128
	1. Transizioni di fase. Ordine della transizione	128
	2. Conservazione dell'entropia nell'espansione cosmologica e generazione di entropia nelle transizioni di fase	129
<i>Appendice A3.</i>	<i>Evoluzione del cosmo in una transizione di fase a temperatura <math>T_c</math></i>	130
<i>Appendice A4.</i>	<i>Dipendenza della densità dal fattore di scala</i>	131
<i>Appendice A5.</i>	<i>Evoluzione del fattore di scala in modelli a 2 componenti</i>	132
<i>Appendice A6.</i>	<i>Orizzonti</i>	134
Capitolo V.	<b>Un modello per l'universo</b>	135
	1. Modello e realtà	135
	2. DM: perché	136
	3. Le ere cosmiche	142
	4. DE: perché	144
	5. $\Lambda$ CDM	151
	6. Conclusioni	152
<i>Appendice A1.</i>	<i>Dipendenza del redshift della distanza e del tempo di emissione di una sorgente</i>	154
Capitolo VI.	<b>L'epoca della nucleosintesi</b>	155
	1. Elio nelle stelle	155
	2. Processi dovuti a forze deboli nell'era leptonica	158

	3. Annichilazione degli elettroni e temperatura del fondo neutrico	162
	4. Il «collo di bottiglia» del Deuterio	165
	5. L'elio e gli altri nuclidi	168
	6. Modifiche alla nucleosintesi primordiale	171
<i>Appendice A1.</i>	<i>Evoluzione cosmologica delle distribuzioni nello spazio delle fasi</i>	173
<i>Appendice A2.</i>	<i>Integrali della distribuzione di Bose</i>	174
<i>Appendice A3.</i>	<i>Equazione di Saha</i>	175
Capitolo VII.	<b>Evoluzione lineare delle disomogeneità</b>	176
	1. Lo scenario dell'instabilità gravitazionale	176
	2. Il problema di Jeans	178
	3. Instabilità di Jeans e crescita delle fluttuazioni	179
	4. Oltre il modello di fluido	182
<i>Appendice A1.</i>	<i>Trattazione analitica</i>	185
	1. Le equazioni di Jeans	185
	2. Evoluzione di perturbazioni rispetto un fondo in espansione	186
	3. Scala di Jeans e soluzioni di Lifshitz	189
	4. Caso relativistico. La gauge gravitazionale	190
	5. Il problema di Jeans relativistico	193
	6. Modi di fluttuazione relativistici	193
Capitolo VIII.	<b>Evoluzione lineare delle disomogeneità</b>	195
	1. Scale e disomogeneità	195
	2. La designazione delle scale	196
	3. La scala dell'orizzonte	197
	4. Interazione barioni-fotoni nelle ere intermedie	201
	5. La ricombinazione	202
	6. Natura delle fluttuazioni di densità	206
	7. Spettri delle fluttuazioni	209
	8. Instabilità di Jeans fino alla ricombinazione	210
	9. Dinamica delle fluttuazioni tra orizzonte e ricombinazione. Componente oscura e plasma	211
	10. Oscillazioni acustiche nei barioni: BAO	218
	11. Soppressione delle onde soniche di piccola scala	219

<i>Appendice A1</i>	<i>La ricombinazione dell'idrogeno</i>	221
<i>Appendice A2</i>	<i>La banda dell'ultimo scattering</i>	222
<i>Appendice A3</i>	<i>CDM e fluido barioni-fotoni accoppiati gravitazionalmente</i>	223
	1. Equazioni di Einstein linearizzate	223
	2. Il problema di Jeans nei modelli con CDM	224
	3. Il problema dinamico	224
	4. Fluttuazioni nella CDM a $z > z_{eq}$ . Effetto di stagnazione	225
	5. Fluttuazioni a $z < z_{eq}$	227
Capitolo IX.	<b>Evoluzione non lineare</b>	229
	1. Spettro primordiale e spettro trasferito	229
	2. Dalla ricombinazione alla non-linearità	231
	3. Evoluzione post-lineare	233
	4. Espressioni di Press & Schechter e generalizzazioni	240
	5. Espressioni di Press & Schechter e dati cosmologici	242
	6. L'approssimazione di Zel'dovich	245
	7. Simulazioni N-body	248
	8. Simulazioni idrodinamiche	251
	9. Scopo delle simulazioni	252
	10. Alcuni risultati delle simulazioni: il profilo NFW	258
<i>Appendice A1</i>	<i>Fluttuazioni nella DM e nei barioni dopo la ricombinazione</i>	262
<i>Appendice A2</i>	<i>Evoluzione di una sovradensità sferica in una cosmologia SCDM</i>	263
<i>Appendice A3</i>	<i>Funzione di massa di Press &amp; Schechter</i>	265
<i>Appendice A4</i>	<i>L'approssimazione di Zel'dovich</i>	267
Capitolo X.	<b>Cosmologia e forze elementari</b>	269
	1. Il dogma di gauge	269
	2. Simmetrie discrete	271
	3. Numero barionico e teorema di Sacharov	273
	4. Particelle e campi	274
	5. Le algebre di Lie	279
	6. Algebre di Lie e interazione dei campi vettoriali	282
	7. La natura delle forze	285
	8. Il modello elettrodebole	292
	9. Il gruppo di rinormalizzazione	294
	10. Supersimmetrie in cosmologia	297
	11. Transizione quark-adroni	301
	12. Transizione quark-adroni in cosmologia	304

	13. La fisica alla scala $T_{GUT}$	307
	14. La $CP$ forte e l'assione	309
<i>Appendice A1</i>	<i>Il teorema di Noether</i>	311
	1. Invarianze e conservazioni nei sistemi continui	312
<i>Appendice A2</i>	<i>La rappresentazione spinoriale</i>	316
<i>Appendice A3</i>	<i>Trasformazioni di gauge abeliane e non abeliane</i>	319
<i>Appendice A4</i>	<i>Il modello standard delle interazioni elettrodeboli</i>	321
<i>Appendice A5</i>	<i>La rottura spontanea delle simmetrie</i>	324
	1. Il modello di Goldstone	324
	2. Il modello di Higgs in teoria di gauge abeliana	327
	3. La rottura della simmetria $SU(2) \otimes U(1)$	327
<i>Appendice A6</i>	<i>Formazione dei residui di particelle fredde</i>	329
<b>Capitolo XI.</b>	<b>L'universo inflazionario</b>	331
	1. Il laboratorio cosmologico	331
	2. Perché l'inflazione	332
	3. La dinamica inflazionaria	335
	4. Leggi di espansione e orizzonti	339
	5. Diluizione	343
	6. Piattezza e parametro di densità	345
	7. Le fasi dell'inflazione e i modelli d'inflazione	347
	8. Il reheating	355
	9. Fluttuazioni e onde gravitazionali	360
	10. Da fluttuazioni quantistiche alle distribuzioni classiche osservate	362
<i>Appendice A1</i>	<i>Cinematica dell'inflazione</i>	364
	1. Espansione esponenziale	365
<i>Appendice A2</i>	<i>Dinamica inflazionaria</i>	366
<i>Appendice A3</i>	<i>Modelli di inflazione</i>	368
	1. Inflazione con campo «grande». Inflazione caotica	368
	2. Inflazione da campo «piccolo»	369
	3. Inflazione ibrida	370
<i>Appendice A4</i>	<i>Il reheating</i>	372
	1. Reheating perturbativo	372
	2. Un modello con reheating non perturbativo	373

<i>Appendice A5</i>	<i>L'origine delle fluttuazioni</i>	376
	1. Metrica e equazioni dinamiche	376
	2. Campi classici	377
	3. Quantizzazione	378
	4. Funzione spettrale	380
Capitolo XII.	<b>L'energia oscura</b>	381
	1. Espansione accelerata oggi: i problemi	381
	2. DE dinamica	384
	3. Evoluzione della DE dinamica	387
	4. Modelli cosmologici con DE dinamica: confronto coi dati	391
	5. Modelli cosmologici con DE dinamica: problemi concettuali aperti	393
<i>Appendice A1</i>	<i>Soluzioni traccianti con potenziali a legge di potenza</i>	395
Capitolo XIII.	<b>Il fondo cosmico microonde</b>	397
	1. Introduzione	397
	2. Lo spettro del CMB	397
	3. L'effetto SZ	399
	4. Anisotropie e polarizzazione: definizioni	399
	5. Mappe del cielo MW	403
	6. Dalle fluttuazioni di densità primordiali alle anisotropie del CMB	405
	7. Forma dello spettro angolare e sua dipendenza dai parametri cosmologici	408
	8. Conclusioni	416
<i>Appendice A1</i>	<i>La dinamica della distribuzione di fotoni a partire dall'equazione di Boltzmann</i>	417
	1. Equazioni del moto per barioni, DM e neutrini	419
	2. Equazioni per il campo gravitazionale	419
<i>Appendice A2</i>	<i>Calcolo delle anisotropie di temperatura</i>	420
Capitolo XIV.	<b>Il tempo e l'entropia</b>	422
	1. Topologia e metrica del tempo	422
	2. Il dominio quantistico	424
	3. La microfisica e il tempo	427
	4. Sistemi chiusi e aperti	428
	5. L'informazione	429
	6. L'abbandono degli equilibri	430
	7. Il ruolo della massa	432

Capitolo XV.	<b>Cosmologia &amp; Cosmologie</b>	436
	1. Cosmologia scientifica e cosmologie tradizionali	436
	2. Il nesso soggettivo	438
	3. Il nesso oggettivo	446
	4. Dal nesso oggettivo alla cosmologia scientifica	452
Riferimenti bibliografici		455
Indice analitico		463
Appendice tavole a colori		471

## Prefazione

La cosmologia descrive l'origine dell'Universo, cioè di tutto ciò che possiamo sperimentare e osservare, coi sensi e con gli apparati forniti da ogni tecnologia.

Questo libro è scritto, in primo luogo, per gli studenti dei corsi di cosmologia nelle lauree in Fisica e in Astronomia. Difatti, i dati astronomici e fisici sono l'ovvio punto di partenza dell'indagine cosmologica. Ma c'è molto di più: la cosmologia è divenuta strumento per esplorare ambiti anche assai lontani dall'astrofisica, e da essa ci aspettiamo «ricadute» persino a livello tecnologico.

Penso peraltro sia un'esperienza abbastanza comune, quella di trovarsi tra le mani un testo di paleontologia, di filosofia medievale, o di altre discipline universitarie, e di trovarsi «invischiati» nella sua lettura. È anche chiaro perché la stessa cosa difficilmente succede con un testo di cosmologia, come per molti altri capitoli della fisica. L'ostacolo è costituito dal linguaggio matematico.

Si tratta di un ostacolo a due facce: gli sviluppi concettuali della cosmologia moderna si esprimono adeguatamente solo tramite matematiche avanzate, molto al di là delle conoscenze acquisibili anche in un buon liceo; ma c'è un secondo aspetto. Un noto divulgatore anglosassone diceva che, per spiegare un concetto, meglio è aggiungere una pagina a un libro, che una formula: a ogni formula che compare al centro di una pagina – diceva – il numero degli aspiranti lettori si dimezza. Questo è il secondo aspetto: l'uso dello strumento matematico è ostico anche quando la matematica è facile.

In questo libro abbiamo tentato un'operazione precisa: fornire un testo che non si curasse all'aforisma del divulgatore anglosassone; dove la matematica occorre, essa viene usata. In prima battuta, tuttavia, la matematica è solo quella che non può essere ignorata da un buono studente di liceo, o da un valido cultore di altre discipline tecnico-scientifiche. Certamente questo sarebbe limitativo, per degli studenti universitari. A ciò abbiamo ovviato con una serie di appendici, dove queste limitazioni cadono e si affrontano i temi con gli strumenti analitici adeguati.

Senza appendici, ci piacerebbe definire questo testo “amatoriale”. Con le appendici, è nostra ambizione portare il lettore ai confini della ricerca.

La storia dell'astronomia è ricca del contributo di ricercatori non-professionali. Non occorre aver letto Tomasi di Lampedusa, per sapere quanto diffuso fosse, tra nobili e clerici, l'uso di cannocchiali e telescopi, tra il Diciottesimo e il Diciannovesimo secolo.

novesimo secolo. Legare il nome del proprio casato alla scoperta di una nuova *cometa*, o di un *asteroide*, era un vanto per le case nobiliari. Le sonde interplanetarie degli ultimi decenni hanno portato a un *revival* dell'interesse per l'astronomia planetaria. Ma il discorso è più vasto e le molte riviste semi-divulgative di astronomia, in Europa e in America, testimoniano ampiamente di questo diffuso interesse amatoriale.

Questa operazione, in realtà, dovrebbe avere anche un risultato speculare. Prese a sé, le appendici sono compendi necessariamente compatti su molti argomenti, alle soglie della ricerca. Molti dei temi trattati si trovano anche in altri testi recenti. Ma spesso, a fini didattici, il discorso si fa opportunamente articolato; nelle appendici, invece, abbiamo mirato a un'esposizione essenziale, asciutta, ma comunque auto-consistente. Su altri temi trattati esistono validi articoli di rivista, a cui spesso ci siamo ispirati. Anche in questo caso, il nostro scopo è stato quello di estrarne l'essenziale, porlo in ordine logico, lasciando ogni argomentazione «descrittiva» al testo del capitolo.

Da entrambi i punti di vista, dunque, ci siamo prefissi degli obiettivi impegnativi. Speriamo di essere riusciti, almeno in parte, nello scopo. Ma l'assieme dei due punti di vista, a nostro avviso, potrebbe avere una terza valenza. Tra le conoscenze medie degli operatori culturali, da un lato, e la moderna cosmologia scientifica, dall'altro, esiste una discontinuità, un «gradone». A esso qui cerchiamo di sostituire due «gradini», mantenendo la completezza della dottrina e enucleando la contiguità tra le conoscenze accessibili e la scienza più avanzata. Se siamo riusciti, almeno in parte, in tutto ciò, sarebbe un grosso successo.

A nostro modo di vedere, le conoscenze scientifiche dovrebbero divenire una base comune di tutto il contesto culturale. Non è infrequente, oggi, sentire validi operatori culturali che denunciano la propria ostilità alle matematiche e, con ciò, si precludono la strada alla conoscenza scientifica. Ma la talora ostentata denuncia della propria insensibilità alla forma più raffinata di arte che esista, la matematica, ci sembra faccia il paio con l'altrettanto ostentata avversione, da parte di tanti nobili o principi della storia, rispetto ai segni alfabetici, alla lettura e alla scrittura. È pur vero che Carlo-magno a mala pena sapeva firmare, e comunque utilizzava il sigillo imperiale. Ma chi prenderebbe oggi esempio da lui, nell'istruzione dei suoi figli?

In attesa che le conoscenze scientifiche si allarghino, comunque, non possiamo non ricordare quanto le cosmologie delle varie civiltà della Storia siano state organiche agli sviluppi culturali e alla stessa costruzione dell'ordine sociale. Il trapasso dalle cosmologie «culturali» alla moderna cosmologia scientifica è stato complesso, e dedichiamo il primo capitolo di questo testo alla sua analisi. È pur vero che anche oggi esiste un flusso di notizie, dall'ambito scientifico all'ambito della più vasta cultura, e che esso ha spesso estese ricadute. Basti pensare, per andare alle origini della cosmologia moderna, quanto abbia pesato il *Big-bang*, non solo nelle dispute tra scienziasti e creazionisti, ma anche nel discorso giornalistico, per divenire un'allusione semantica, influenzando persino la moda o la borsa valori.

Negli anni recenti si è spesso assistito a un «debordare» di conoscenza o semantica cosmologica, verso il mondo culturale. Spesso le allusioni, dai «buchi neri» alle «stelle fredde», risultavano persino fastidiose per chi nella scienza operava. Ma ciò potrebbe anche essere irrilevante, se non risultasse chiaro, almeno agli operatori di scienza, quanto simili superficialità siano contrassegno del muro che Ottocento e Novecento hanno elevato, tra cultura e cultura scientifica.

Vorremmo dunque contribuire, almeno un poco, al ristabilimento di un flusso senza distorsioni. Da nostra parte, oltre che nel primo, anche negli ultimi due capitoli tentiamo di entrare nel campo interdisciplinare. Sappiamo quanto certe operazioni siano delicate e chiediamo venia al lettore più competente di noi in campo non scientifico.

Voglia egli leggere quei testi come un'elaborazione di tematiche, su cui altri è forse più competente, da parte di chi ha una sensibilità culturale diversa, maturata in decenni di ricerca fisico-matematica.

## Dedica e riconoscimenti

Dedico questo libro alla memoria di **Francesco Lucchin**. L'idea iniziale di scriverlo era maturata nelle nostre discussioni. Avevamo anche cominciato a scriverlo assieme, ma un destino impietoso ha voluto allontanare da noi uno dei protagonisti della ricerca cosmologica internazionale e certamente un uomo di grande spessore culturale e morale.

Ma a Francesco debbo comunque molto di più dell'idea del libro. Gli debbo molte delle mie competenze astrofisiche e cosmologiche, ma soprattutto gli debbo la spinta a approfondirle, in un tempo in cui il contesto della ricerca fisica in Italia vedeva molti di questi argomenti con malcelato fastidio.

Lavoravamo assieme nell'ultima stanza a destra del corridoio dei «teorici», al secondo piano del Dipartimento di Fisica G. Galilei, a Padova. D'inverno il riscaldamento era insufficiente e il sole estivo arroventava la stanza. All'inizio non avevamo un computer, per i calcoli numerici si sottoponevano pacchi di schede al centro di calcolo, al di là del cortile. Poi venne un computer, fuori dalla porta dello studio, in condominio con altri ricercatori e, nel primo pomeriggio, lo si utilizzava anche per le battaglie di *Star wars*, abbattendo le navi dei *Klingons* o finendo sbriciolati a *warp* 7.

Per diversi anni rinunciavo alle vacanze estive e passavo luglio e agosto all'Institute of Astronomy, subito fuori dall'abitato di Cambridge. Mi pareva di accedere a un mondo di infinito sapere, illustravo il lavoro di quell'anno a **Martin Rees** e agli altri amici che lavoravano in quel luogo, venivo da loro indirizzato a discutere con chi aveva le competenze più specifiche, ritornavo alla fine a Padova con un pieno di idee e di entusiasmo, per un nuovo inverno di lavoro. Ma spesso era Francesco a porre un ordine in quel bagaglio di progetti; mi pareva a volte che egli impoverisse il mio slancio, ma a distanza di anni debbo riconoscere molte delle sue ragioni e, certo, senza di lui difficilmente avrei sortito molti dei risultati di allora.

Non è forse qui il luogo per continuare a raccontare la storia di anni tanto diversi da quelli attuali. Quando le speranze non avevano limite e ci pareva che i successi dei giovani fossero comunque destinati a meritata ricompensa. Un tempo ben diverso dall'attuale; oggi i migliori dei nostri allievi trovano facilmente collocazione nelle grandi istituzioni internazionali, ma di lì non rientreranno più. Quando qualche risorsa diviene disponibile, di norma è troppo tardi, la loro vita familiare e scientifica è ormai assestata e continuerà altrove.

A dispetto dell'arroganza di una parte dell'*establishment* di allora, a dispetto di molte frustrazioni causate dall'incompetenza dei clerici, abbiamo certo vissuto, Francesco e io, una delle stagioni più vivaci dell'avanzamento della scienza in questo paese. Voglio solo sperare che la sensazione di essere alla vigilia dell'eclisse, l'ultima generazione *avant le deluge*, sia smentita dai fatti.

Ma alla scrittura di questo libro hanno dato una mano in molti. *In primis* debbo ringraziare **Alfonso Cavaliere**, che per un po' ha condiviso con me questo progetto e che ha avuto anche la capacità di fare un passo indietro, quando altrimenti il completamento del libro sarebbe stato troppo rimandato. Debbo poi ringraziare **Sabino Matarrese**, l'abitatore attuale dell'ultima stanza a destra in fondo al corridoio a Padova, in lontani anni allievo di Francesco e mio, oggi certamente uno dei migliori cosmologi europei. I suoi consigli e le sue correzioni, specialmente per il capitolo sull'inflazione, mi sono stati utilissimi. Ulteriori espliciti riconoscimenti vanno a **Giuseppe Murante**, le cui critiche su diversi capitoli sono state estremamente costruttive, e a **Giuseppe (Peppo) Gavazzi**, senza il cui contributo molte delle immagini riprodotte in questo testo non sarebbero state disponibili. Colgo poi l'occasione per ringraziare molti altri, tra cui **Andrea Macciò**, **Loris Colombo**, **Roberto Mainini**, alcuni dei miei allievi negli anni più recenti. Anche senza avere tra le mani il mio lavoro, vi hanno contribuito con le loro idee, o aiutandomi a impadronirmi di algoritmi o tecniche, o fornendomi aggiornamenti bibliografici.

Non posso concludere questa lista di ringraziamenti, dalla quale ho certo indebitamente escluso qualche altro amico, senza ricordare **Nicolò Dallaporta**, recentemente scomparso, cui debbo non solo l'iniziazione al mondo della fisica, ma anche un importante contributo dialettico al mio modo di vedere la scienza e il suo ruolo nella società e nella cultura.

L'ultimo e più importante dei ringraziamenti va infine a mia moglie **Bruna**, soprattutto per la pazienza che ha avuto, in tante serate, week-end, giorni di vacanza, quando il mio lavoro continuava e potevo dedicarle solo gli scampoli della mia attenzione.

## CAPITOLO I

# Cosmologie & Cosmologia

Tutte le culture, da sempre, hanno utilizzato le loro tecniche più avanzate per affrontare la sfida della cosmologia, cioè per esprimere le loro valutazioni sull'origine del cosmo. Ciò è, quanto meno, una conferma dell'enorme impatto di questo tema, anche sul piano sociale. Inoltre, sin dall'emergere della memoria storica, è ovunque registrata la relazione tra astronomia e cosmologia, la coscienza nell'uomo di essere figlio di quegli stessi cieli che riempiono di ammirazione il suo animo nelle notti serene: «Due cose riempiono il mio animo di ammirazione sempre nuova e crescente, la volta stellata sopra di me, e la legge morale dentro di me» (Immanuel Kant, conclusioni della *Critica della Ragion Pratica*).

Sarebbe un errore non rendersi conto, da un lato, della contiguità tra l'ammirazione antica e la cosmologia moderna; dall'altro, della gradualità del trapasso dalla cosmologia *culturale*, frutto avanzato di speculazione filosofica e sociologica, alla cosmologia *scientifica*, figlia del dato sperimentale. Come sarebbe un errore non rendersi conto di quanto radicale, alla fine, sia stata la transizione, poiché la nostra cultura è andata ben oltre l'espressione di valutazioni, pervenendo a una struttura cosmologica ricavata cogentemente dai dati.

In questo capitolo vorremmo ripercorrere brevemente le tappe del trapasso; nostro scopo non è la completezza storica e l'*excursus* tra le cosmologie del passato è esclusivamente mirato a comprendere il meccanismo del trapasso stesso e se ancora ci sia qualcosa di a-scientifico nella cosmologia di oggi.

A questo proposito vorremmo subito asserire con forza che la struttura portante della cosmologia moderna ha pieni connotati di scientificità: *la cosmologia è oggi una branca dell'astrofisica e della fisica*. Le sue metodologie di base sono le stesse dell'astronomia e della fisica fondamentale. C'è di più: il paradigma cosmologico sta diventando il tramite della *verifica sperimentale* della fisica di altissima energia. Ma, proprio per questo, è importante anche provare a distinguere tra le strutture portanti ormai assestate, le induzioni dettate dal procedere della ricerca e conformi al metodo scientifico, e talune speculazioni a-scientifiche rivestite del linguaggio della scienza.

Del resto non è che queste ultime non abbiano una loro importanza. Per oltre vent'anni, dal 1929 al tardo «dopoguerra», tutta la cosmologia autodichiaratasi

«scientifica» ebbe simili connotati e fu anche da quelle speculazioni che nacque la cosmologia come vera scienza.

## 1. L'uso delle scritture

Non ci rendiamo spesso conto, oggi, di cosa fosse la *scrittura* ai suoi inizi. Essa richiedeva strumenti altamente dispendiosi, disponibili solo alle classi sociali cui le istituzioni permettevano di sfruttare il lavoro delle moltitudini contadine. Non era concepibile un addestramento generalizzato a scrittura e lettura; del resto, a capo degli imperi stavano spesso principi cui non s'era nemmeno posta la questione di imparare a scrivere.

Tutto ciò serve a comprendere la carica ideologica e la rilevanza sociale di disporre di «scritture» concernenti l'origine del cosmo. La tecnologia allora più avanzata era messa a disposizione del quesito più difficile.

È in questo stadio della storia umana che nasce, quasi contemporaneamente in diverse culture, il *corto-circuito* astrologico. Sembra di sentire il dotto che si rivolge al giovane adepto, che lo ha interrogato sulla meraviglia dei cieli punteggiati di stelle, nella notte di Babilonia: «Non solo i cieli sono meravigliosi, ma, conoscendo la chiave di lettura, in essi è iscritto tutto il futuro dell'Uomo». Una frase che colpisce il giovane, ma il giovane non comprende; noi invece sappiamo quale mistificazione il sapiente sacerdote aveva in mente; e infine una frase che contiene molta verità, ma in un senso completamente diverso da quello significato dal dotto.

Nell'avvicinarci quindi alla descrizione delle origini, contenuta nelle scritture delle culture più antiche, dobbiamo renderci conto dello sforzo anche tecnologico che la loro stesura ha implicato. Uno sforzo non inferiore a quello che impegna uno scienziato moderno, che scrive le sue equazioni, esprimendo così in modo altamente sintetico l'ortodossia di concetti logici altrimenti troppo difficili da enunciare; allo stesso modo in cui la tradizione orale si dimostrava ormai inadeguata a stabilire nel tempo l'ortodossia dei sottili distinguo, di cui sono costellate le scritture. Possiamo anche immaginare un futuro, neppure troppo remoto se la specie umana riuscirà a superare la sfida del controllo tecnologico dell'ambiente, in cui le conoscenze matematiche saranno altrettanto diffuse quanto è oggi la scrittura, e la cosmologia scientifica sarà altrettanto accessibile a tutti quanto lo sono oggi le scritture antiche.

Una chiave di lettura delle scritture antiche è l'ispezione di come vengano tra loro bilanciati gli spazi destinati all'*Essere* con quelli destinati al *Divenire*. Ci sta a cuore attirare l'attenzione del lettore su questi due archetipi, perché essi sono destinati a ripresentarsi, in forme più o meno evolute, in tutte le costruzioni cosmologiche, fino ai nostri giorni.

La cosmologia vedica, contenuta nelle più antiche scritture che ci siano note, è ampiamente concentrata sugli aspetti dell'Essere. Sulla cosmologia vedica torneremo ancora, da altra prospettiva, nel capitolo conclusivo di questo testo. L'Induismo amplia il concetto di «Essere» nel *Brahman*, entità suprema primordiale personale e im-

personale, che include tutto l'Essere ma anche il *non-Essere*, la negazione che può essere definita solo in presenza del reale. Nei primordi a-temporali, il Brahman si «polarizza» nelle entità maschile e femminile, di *Purusha* e *Prakrti*. *Prakrti*, a priori, è una miscela indifferenziata dei tre *Guna* (cfr. latino *genus*): *Sattwa*, *Rajas*, *Tamas*; la luce, l'azione, la gravità. Le personificazioni dei *Guna* saranno poi le divinità della *Trimurti* indù: *Brahma* (da non confondere col Brahman), *Visnù* e *Siva*.

Il lettore che ha avuto la pazienza di seguire questa piccola parte della descrizione vedica delle articolazioni dell'Essere – che si dilunga ancora molto, includendo per esempio nel *SatChitAnanda* un'immagine strettamente legata alla *Trinità* cristiana – comprenderà che il suo scopo è anzitutto il confronto con lo spazio che il libro della *Genesi*, alla base della cultura giudaico-cristiana, riserva all'articolazione dell'Essere. Esso è così breve da poter essere totalmente riportato a seguire: «All'inizio Iddio creò il Cielo e la Terra, la Terra era informe e deserta e le tenebre regnavano sull'abisso. Ma sulle acque spirava un Vento Divino».

A ciò segue la parte più celebrata della *Creazione*, quella che avviene tramite la *Parola*, il *Verbo*: «E Dio disse "sia la Luce"...». È la parte destinata a far risaltare la centralità dell'uomo sulla scena cosmica, in una successione di eventi che allora inevitabilmente ricalca – veramente con pochissimi errori – le scoperte moderne di cosmologia, geologia e paleontologia. Ma è anche la parte in cui domina la sequenza temporale segnata dai giorni; tutto ciò avviene ormai nel tempo, è «divenire».

L'opposta focalizzazione delle due diverse culture si riflette sulla concezione stessa del tempo. Nella cosmologia induista si sussegue una sequenza indefinita di *eoni* o *brahmānda* (uova di Brahma), che durano 100 anni di Brahma, calcolabili in  $3.11 \times 10^{14}$  anni, 20 000 volte il tempo trascorso dal *Big-bang*. Tutti gli esseri viventi, «dal Dio Brahma dai quattro volti, fino allo stelo d'erba», cessano di esistere alla fine di ogni eone. Dopo 100 anni di Brahma di assenza, il cosmo rinascerà, con nuovi mondi e nuovi dei. Al *tempo ciclico* della cosmologia vedica si contrappone il *tempo lineare* della cosmologia giudaica, in cui domina l'*uomo*, essere di caratura divina, fatto «a Sua immagine e simiglianza», cui viene comandato di assoggettare il mondo e le sue creature. Ancor più pronunciato diviene il vettore temporale con la *Divina Incarnazione*, che segna uno spartiacque nella storia.

L'unità e univocità della sequenza temporale è del resto parte di tutto l'afflato monistico della cultura giudaica, mirabilmente enfatizzato nella preghiera mosaica: «Shemà Israel, Adonai Elohinu, Adonai Echad» (Ascolta Israele, l'Eterno è il Signore, l'Eterno è Uno). L'unità dell'Essere, espressa nel Brahman, ma poi oltremodo articolata nella cultura vedica, non è mai in discussione nelle dottrine monoteistiche, anche quando appare l'articolazione trinitaria. Questo afflato all'unità è veramente iscritto nella nostra cultura e sarà bene sottolinearlo anche in vista delle dinamiche più recenti, destinate a sbocciare nella cosmologia come scienza.

Su un terreno in qualche modo intermedio si pone la tradizione *olimpica*, destinata a fondersi con la cultura giudaica per originare la cultura occidentale. Le sue scritture hanno un senso mistico decisamente minore. Ma, nella *Teogonia* esiodea, rintracciamo

facilmente lo spazio dedicato all'articolazione dell'Essere che si conclude con la nascita di *χρόνος*, il dio che divora i suoi figli e da cui verrà salvato Giove, per essere posto a capo dell'Olimpo classico.

La modesta profondità esoterica della religione olimpica si riflette nella quasi assoluta assenza di riflessi non formali sulla cultura occidentale. Tuttavia è proprio nel corso dello sviluppo della cultura ellenica che la storia del pensiero cosmologico compie un primo significativo passo avanti, essenziale a focalizzare il processo che porta alla cosmologia moderna, di cui tratteremo nel prossimo paragrafo.

## 2. Pitagora da Samo

Il quinto secolo avanti Cristo fu tempo di rivoluzioni ideologiche su tutto il pianeta. È il secolo del *Buddha* in India, di *Confucio* e *Lao-Tse* in Cina, è il secolo in cui parlò *Zaratustra*. Nel mondo ellenico è il secolo di Pitagora.

I dati sulla vita di Pitagora sono stati registrati sei secoli dopo la sua morte. Ciò ha reso gli storici dubbiosi sulla sua stessa esistenza, un po' come per il cieco vate Omero. Della vita e delle idee di Pitagora ci parla *Aristotele*; ma due autori della tarda romanità, Giamblico e Porfirio, hanno anche dedicato interi libri a questo pensatore, *Vita Pitagorae* e *De vita pitagorica*. Ai nostri fini va benissimo accettare l'esistenza del Pitagora che emerge da questi dati, perché ciò che ha interesse è la memoria storica che da lui si propaga.

Pitagora nacque dunque a Samo, dove si trovava un antico santuario della dea Era. Ancor giovanissimo emigrò in Egitto, dove viveva una vasta colonia greca dedita ai commerci. Ma il suo interesse andava a altro; cercò difatti di essere ammesso alle scuole che preparavano i futuri sacerdoti della religione egizia e, dopo alcuni fallimenti, venne iniziato nel santuario di Diospoli.

Ma l'orologio della storia stava segnando il tempo della crisi. Erodoto ci racconta come i persiani guidati da Cambise schiacciarono le difese egizie nel Sinai e dilagarono nel paese del Nilo. Qui distrussero e depredarono, portando con sé enormi ricchezze in oro, manufatti e schiavi, nonché i sacerdoti dei templi, giustamente ritenuti depositari del sapere più avanzato dell'epoca. Così, dopo essere vissuto nell'ambito delle caste sacerdotali egizie, a Pitagora capitò la ventura di essere iniziato anche al sacerdozio persiano e di apprendere la cultura dei *magi*.

Partito dalla Grecia poco più che quindicenne, Pitagora vi rientra a quarant'anni d'età, con un bagaglio articolato di sapere, smitizzato dal confronto tra le tradizioni. Al centro della dottrina di cui egli va a farsi promotore non ci sono più i miti del divino, ma le valenze scientifiche dell'esoterismo che, in Egitto come in Persia o in Mesopotamia, erano state volte al controllo sociale.

Per Pitagora, gli dei verranno dopo: al centro dell'Universo sta la *bellezza matematica*, vero veicolo al soprannaturale. Il dominio della casta sacerdotale è sostituito da quello dei *μαθηματικοί* e, su questa base, si fonda un nuovo sistema politico, sperimentato per breve tempo a Krotona. Ma l'esercizio del potere politico è secondario rispetto alla ricerca del sapere. Ai pitagorici sono accreditate molte acquisizioni di matematica, come la

definizione del numero dei solidi regolari convessi, vari teoremi sulle successioni, la scoperta dei numeri irrazionali. Se, al centro del mondo pitagorico sta dunque la *scienza del numero*, la maggiore acquisizione che gli va forse accreditata è quella della *bellezza del numero*, che esso legò alla *bellezza musicale*, fissando le proporzioni tra le lunghezze delle corde e i toni, definendo l'accordo di ottava, di quarta, di quinta, e così via.

Pitagora avrebbe forse potuto essere il riformatore della religione olimpica, colui che la riportava ai fulgori delle radici indoeuropee, parallele ai *Veda*, da cui probabilmente proveniva. Egli invece sceglie una strada laica, emblema dell'empiricità dell'ambiente sociale da cui proveniva, ma in cui risaltano i veri legami che avvicinano l'umanità all'assoluto.

Ebbene, se questa ci appare la base del pensiero pitagorico, varrà subito la pena di seguirne gli effetti nella cultura dei secoli che seguirono.

Alcuni di essi sono gravemente deleteri. La parte più deteriora dei cascami pitagorici venne inserita nelle ideologie di alcune scuole *neoplatoniche*, ai tempi di Plotino e Proclo. Si tratta dei circoli del tardo impero in cui si cercava di resistere al dilagare delle sette mitraiche e giudaiche. Si tratta degli stessi circoli che probabilmente curarono le trascrizioni dei testi di Ermete Trismegisto, quelle stesse trascrizioni che giunsero poi nella Firenze medicea per essere tradotte dal chierico Giovanni Ficino e divenire la base della scuola magica rinascimentale.

Un legame forse mai adeguatamente esplorato connette poi il pitagoreismo alla *Cabbalàh* ebraica, dallo *Zohàr* a Isaac Luria, in quella che è probabilmente la più profonda elaborazione esoterica fiorita nella cultura occidentale.

Ma la storia della scuola pitagorica si presta soprattutto a mostrare come le punte più avanzate del pensiero scientifico siano poi strumento di elaborazione cosmologica. Ai numeri vengono associate allora le divinità tradizionali, il cui ruolo viene rielaborato alla luce degli esoterismi dell'oriente, di modo che i nomi sono quelli degli dei olimpici, ma il ruolo si approssima alle connotazioni vediche. L'*Uno*, emblema dell'unità dell'essere, ritrova il nome di *ὀύρανος* (il cielo) e reca con sé le fattezze del Brahman: esso è l'assolutezza dell'Essere, destinato a articolarsi nelle molteplicità senza tuttavia essere corrotto, esso viene anche visto come *fuoco primordiale*, rispetto cui vanno misurati i moti altrui. Il *Due* è *γῆα*, la madre Terra, sostanza primordiale. Il *Tre*, similmente, si riconduce alla divinità solare, al dio *απόλλων*, al ruolo maschile del Purusha.

Al di là dei dettagli di queste associazioni, su alcuni dei quali varrà comunque la pena di ritornare, ciò che emerge con chiarezza è l'uso dello strumento scientifico più avanzato, delle nuove scoperte matematiche, per rivisitare e rinvigorire l'exoterismo olimpico, per l'occasione coniugato agli esoterismi orientali. Tra cui fa decisamente spicco il paradigma astrologico, che promana dal mondo egizio e dalla mezzaluna fertile, e fa la sua trionfale entrata nella cultura occidentale.

Eppure, mentre non possiamo che annotare l'origine di un filone foriero di tanti fuorviamenti, ci sorprende la precognizione di dati fisici allora radicalmente sconosciuti. Sulla base delle astrazioni pitagoriche e non di dati, per i quali si sarebbero dovute attendere le puntigliose osservazioni secentesche di Tycho Brahe, due secoli più tardi

Aristarco da Samo, rinunciando all'Uno inconoscibile, definì per primo il sistema eliocentrico. Al suo tempo Tolomeo aveva dalla sua parte tutte le ragioni della scienza. Dalla parte di Aristarco troviamo solo speculazioni e metafisica. Se poi accreditiamo ad Aristarco l'eliocentrismo, per quale ragione non attribuire ai pitagorici la scoperta del CMB? Non è forse il *fuoco primordiale*, ancora presente nel fondo cosmico a microonde (CMB), a definire il sistema di riferimento assoluto, quello rispetto cui tutti i moti vengono definiti?

### 3. L'eresia astrologica

Varrà la pena di spendere poche parole sul corto-circuito astrologico, vista la sua diffusione quasi in tutte le culture tradizionali. Accanto all'astrologia egizia e babilonese, i cui ultimi cascami sono le rubriche oggi diffuse sui *mass media*, ci fu un'astrologia cinese, un'astrologia indiana, e le astrologie delle civiltà precolombiane.

Nel medioevo europeo il paradigma astrologico ricoprì un ruolo incerto; esso rifulse prepotentemente con la riscoperta dei testi di Ermete Trismegisto, che lo comprendevano e, in certo modo, lo superavano, istituendo la magia. I testi ermetici testimoniano dell'esistenza di tre mondi: terreno, celeste e divino. Il mondo celeste influenza il mondo terreno e tali influenze sono decrittate dall'astrologia. Ma anche dal mondo terreno si riesce a propagare un'influenza sul mondo celeste, che poi si ripercuoterà sul futuro degli uomini. Questa è la strada della magia, che il dotto rinascimentale attua tramite amuleti e talismani, tramite lo studio delle pietre e dell'influenza dei colori. Dal Ficino, traduttore dei testi del Trismegisto, si passa a Giovanni Pico della Mirandola, che compie una mala lettura di testi cabalistici e pretende la possibilità di giungere alle porte del mondo divino, con le *magie angeliche*. Per giungere infine alla tragedia di Giordano Bruno, che pretende di portare le gerarchie cattoliche e il Papa a rifondare il cristianesimo su basi magiche.

A fronte di questa infezione planetaria va sottolineato il ruolo della tradizione giudaica, unica a mantenersi pienamente immune dall'eresia astrologica. Cancellando il corto-circuito astrologico dall'ortodossia del pensiero, il cristianesimo apre la strada allo sviluppo scientifico. Se Isaac Newton resta a lungo in bilico tra pensiero scientifico e pensiero magico, per cadere poi nella trappola alchemica, i chierici Martin Mercenne (1588-1648) e Pierre Gassendi (1592-1655) coniugano il meccanicismo postcartesiano alla visione cristiana del mondo.

Non c'è dubbio che il ruolo della chiesa romana, rispetto alla scoperta della strada che dalla scienza porta alle tecnologie, sia quanto meno ambiguo, a partire dalla vicenda ultranota di Galileo Galilei. Ma non c'è neppure dubbio che sono le radici del filone giudaico-cristiano-islamico a porre correttamente il ruolo dell'uomo rispetto al cosmo e a proporre correttamente il suo ruolo nel rapporto con la natura.

### 4. Dall'astronomia alla cosmologia

Dagli albori della storia e fino all'Ottocento, il problema cosmologico riguarda stelle e pianeti. A tutt'oggi, riferendosi allo studio della nascita del sistema solare, si usa il

termine *cosmogonia*, di etimo prossimo a quello che riguarda la nascita dell'Universo intero.

Ancora all'inizio del Novecento mancavano i dati sperimentali che permettessero di iniziare uno studio cosmologico, come mancava una base teorica per inserirveli, ove fossero stati noti. Quando Albert Einstein si pone il problema cosmologico, il cosmo cui egli fa riferimento è quello popolato di stelle, la nostra galassia, la *Via Lattea*. C'erano sì stati alcuni pensatori che avevano proposto, ancora nel Settecento, che talune nebulose potessero essere altri universi-isola, entità simili alla Via Lattea, di cui fa parte il Sole. Lo aveva pensato e scritto anche Immanuel Kant, con riferimento a alcune osservazioni effettuate coi modesti apparati ottici dell'epoca.

Ma, sul finire dell'Ottocento, due elementi poderosi avevano persuaso gli astronomi «professionisti» a non persistere in questa «eresia». Anzitutto la scoperta dell'esplosione di *Supernovae* nelle nebulose. Una supernova può avere una luminosità  $\sim 10^9 L_{\odot}$ , oltre un miliardo di volte la luminosità del Sole. Gli astronomi giustamente si chiedevano: «Com'è possibile che una sola stella pareggi in luminosità una nebulosa, se essa contiene tante stelle quante ne contiene la Via Lattea?». Il secondo elemento, più sottile ma ancor più potente, era di natura statistica. Si osservava un'anticorrelazione tra il piano della Via Lattea e la distribuzione delle nebulose. Oggi sappiamo che la scarsità di galassie osservabili sul piano galattico è dovuto alla copiosità delle *polveri* sul piano stesso. Ma allora, giustamente, i professionisti si chiedevano: «Come sarebbe possibile che la distribuzione di sistemi non appartenenti alla Via Lattea risenta della geometria della Via Lattea stessa?».

Per questo va riconosciuto a Edwin Hubble il merito di aver sdoganato le galassie, riconosciute come gli abitanti di un Universo su cui si può finalmente speculare in termini cosmologici.

## 5. Edwin Hubble

La memorabile 33<sup>a</sup> Riunione della American Astronomical Society, iniziata alla fine del 1924, tenne una sessione anche in data 1 gennaio 1925. A essa non partecipò Edwin P. Hubble, e altri lesse la sua comunicazione sul risultato di un programma di osservazioni rese possibili dal nuovo telescopio da 100 pollici di Mount Wilson.

Concettualmente semplici ma cruciali, esse dimostravano che M31 in Andromeda era un sistema stellare esterno alla Via Lattea e a essa simile. Hubble aveva identificato in M31 delle stelle *cefeidi* che usava come indicatori di distanza. Le cefeidi sono stelle variabili sulla scala dei giorni, per cui si conosceva allora una relazione empirica tra periodo delle variazioni ( $P$ ) e luminosità intrinseca ( $L$ ). Varrà la pena di rimarcare che si trattava di una relazione errata, ma non di ordini di grandezza.

Grazie alla fotometria svolta in collaborazione con Humason, Hubble aveva ricavato i periodi  $P$  delle cefeidi in Andromeda, deducendone le luminosità  $L$ . Confrontando poi  $L$  con la luminosità apparente  $l$  di ciascuna cefeide, tramite la legge

$$(I.1) \quad l = L/4\pi d^2,$$

è facile ricavare la distanza  $d$ . Hubble valutò così la distanza di M31 in 250 kpc, a fronte di un raggio della Via Lattea di circa 10 kpc; il parsec (pc) è l'unità di distanza più comunemente usata in astronomia e vale 3.26 anni-luce, circa  $3.09 \times 10^{18}$  cm; kpc e Mpc sono, rispettivamente 1000 pc e 1 milione di parsec. Nonostante la distanza di M31 sia 3 volte maggiore, come si vede usando i dati corretti sulle cefeidi, già i numeri di Hubble collocavano definitivamente M31 al di fuori della Via Lattea.

Questa scoperta segna la data di nascita dell'astronomia extragalattica. Il primo dato di questo nuovo capitolo riguarda allora, curiosamente, la nostra stessa galassia, la Via Lattea (*Milky Way*), per cui si misura una luminosità totale  $L_{mw} \simeq 10^{10} L_{\odot}$  e si stima una massa  $M_{mw}$  tra  $10^{11}$  e  $10^{12} M_{\odot}$  (questo ampio intervallo non è dovuto a difficoltà sperimentali, ma alla progressiva scoperta di nuove componenti oscure; ne riparleremo diffusamente più avanti). Le unità  $L_{\odot}$  e  $M_{\odot}$  sono la luminosità intrinseca e la massa del Sole, rispettivamente  $4 \times 10^{33}$  erg/sec e  $2 \times 10^{33}$  grammi.

Ma la vera astronomia extragalattica riguarda ciò che è esterno alla Via Lattea. In ogni direzione, a profondità via via crescenti, si osservano altre galassie simili alla Via Lattea stessa. La distanza media galassia-galassia,  $\lambda_{gg}$ , è di 3-5 Mpc. Anche qui non ci troviamo di fronte a un'incertezza sperimentale, ma a diverse possibili scelte sulla massa minima degli oggetti cui attribuire la qualifica di "galassia"; uno spostamento anche modesto del limite inferiore causa un significativo spostamento nel valore di  $\lambda_{gg}$ , dal momento che il numero di oggetti cresce rapidamente al calare della scala.

Non deve piuttosto sorprendere che la distanza di M31 dalla Via Lattea sia inferiore al Mpc e che le *Nubi di Magellano* siano ancora più vicine. Il fatto è che le galassie tendono a presentarsi in gruppi, legati gravitazionalmente; alcuni sono semplici sistemi binari, ma altri sono più consistenti, come il Gruppo Locale in cui si trovano la Via Lattea e M31, fino a giungere a *ammassi* di migliaia di galassie, come Abell 1689 (vedi Fig. I.1). In sostanza, tutto ci appare come un sistema di scatole cinesi: il gas forma le stelle, le stelle formano le galassie, le galassie formano gruppi e ammassi. Sin qui i sistemi sono gravitazionalmente legati, seppure in condizioni dinamiche assai difforni, per cui si tratta di associazioni stabili o stazionarie.

Su scale maggiori riscontriamo associazioni che possono essere transienti, come i *superammassi*, su scale fino a 10 Mpc, e tutta una rete cosmica di filamenti e superfici. Tra i filamenti risaltano i grandi *vuoti*, di raggio fino a 30-40 Mpc.

Al crescere della scala di massa, cala il contrasto di densità tra sistema cosmico e spazi circostanti. Si va da un contrasto di densità di venti ordini di grandezza per le stelle, per passare a  $\sim 1 : 10^7$  per le galassie,  $\sim 1 : 100$  per gli ammassi,  $\sim 1 : 5$  per i superammassi. È dunque oggi legittimo pensare che, su scale ancor maggiori, si approssimi l'omogeneità.

Ciò concorda coi risultati sull'Universo profondo forniti dai radiotelescopi e coi dati sul fondo cosmico a microonde (CMB: *cosmic microwave background*) rilevato per la prima volta da Arnold A. Penzias & Robert W. Wilson nel 1965, di cui parleremo nel prossimo paragrafo, che testimoniano di una *isotropia* superiore a  $1 : 10^5$ .

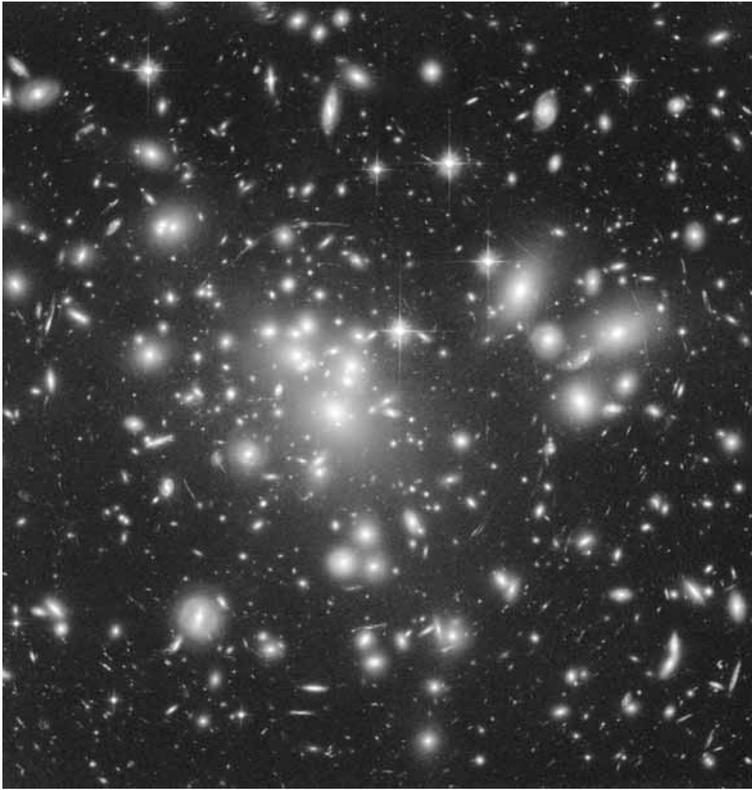


FIGURA I.1

L'ammasso di galassie Abell 1689. Si notino gli archi dovuti al *lensing* gravitazionale forte, mentre sono anche chiaramente riconoscibili sistematiche distorsioni nella forma delle galassie di *background*, da attribuire a *lensing* debole. (si veda l'appendice TAVOLE A COLORI).

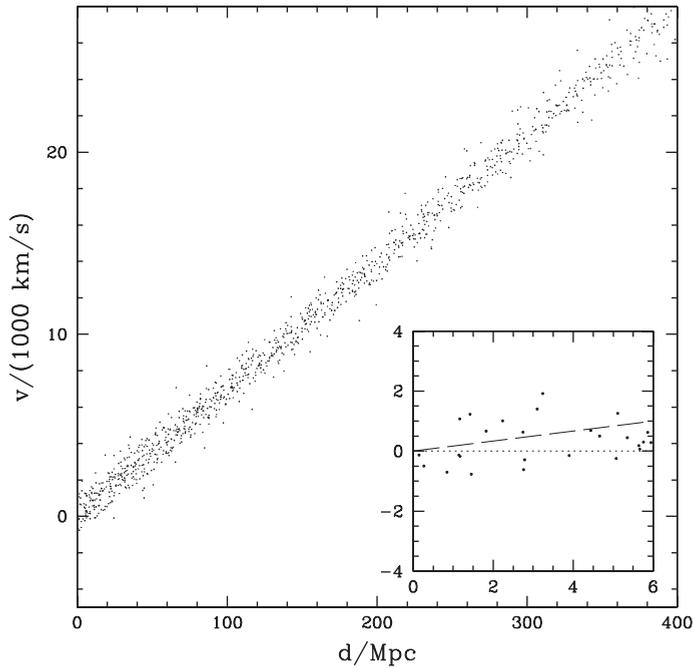
Ma questi dati sono il risultato di un lavoro lungo e sofisticato, durato alcuni decenni, che prende sì le mosse dalla comunicazione di Hubble del 1925, ma che richiede il lavoro di centinaia di ricercatori. Nel prossimo capitolo discuteremo in maggior dettaglio tutto questo quadro. Qui vogliamo ritornare al contributo di Edwin Hubble cui si accredita la scoperta dell'espansione dell'Universo, il moto d'insieme delle galassie, che va appunto sotto il nome di *flusso di Hubble*.

## 6. Il flusso di Hubble

Fu Hubble a suggerire, nel 1929, che valesse la celebre *legge di Hubble*

$$(I.2) \quad v = H_0 d,$$

( $d$  : distanza delle galassie osservate;  $v$  : velocità di allontanamento), fornendo per  $H_0$  un valore  $\sim 500$  (km/s)/Mpc: la velocità di allontanamento cresce linearmente con la

**FIGURA I.2**

Relazione tra velocità relativa alla Terra e distanza, per un campione di galassie entro  $\sim 400$  Mpc. Nel riquadro sono magnificate velocità e distanze delle galassie entro 6 Mpc.

distanza, di 500 km/s per ogni Mpc in più. Ciò non significa che l'osservatore terrestre occupi un sito privilegiato. Le distanze spaziali crescono come crescono le distanze sulla superficie di un pallone sferico che viene gonfiato; qualsiasi osservatore «bidimensionale», sulla superficie del pallone vedrebbe gli altri punti della superficie allontanarsi, con velocità proporzionale alla distanza.

La legge (I.2) è corretta, il valore di  $H_0$  è errato. Ma, in sé, l'acquisizione concettuale sarebbe stata formidabile, se fosse stato vero che i dati permettevano di ottenerla. Essa si basava sulla misura dei *redshift* delle righe spettrali di una ventina di galassie, entro 5-6 Mpc. Più in là, il telescopio di Mount Wilson e le tecniche spettroscopiche dell'epoca non permettevano di giungere.

In figura I.2 è rappresentata la relazione tra velocità di galassie osservabili entro  $\sim 400$  Mpc e le loro distanze. Da un campione di questa profondità, la legge di Hubble risalta chiaramente, al di sopra del «rumore» dovuto alle velocità peculiari delle singole galassie. (La figura, in realtà, è ottenuta tramite una simulazione basata sui dati fisici reali. Per ottenerla direttamente dai dati, dovremmo avere usato un indicatore di distanza, indipendente dal redshift, fino a 400 Mpc di distanza, su un campione di decine di migliaia di galassie; in futuro lo si potrà fare.)

Se però restringiamo l'esame ai dati entro 6 Mpc, ci troviamo nella situazione descritta nel riquadro, ove non c'è neppure una maggioranza statisticamente significativa

di galassie in allontanamento. La linea tratteggiata indica la *legge di Hubble* con  $H_0 \simeq 170$  (km/s)/Mpc, corrispondenti ai 500 (km/s)/Mpc stimati da Hubble, rinormalizzati correggendo il suo errore sulle stelle cefeidi. Un'occhiata al grafico potrebbe far pensare che la stima possa essere corretta e Hubble pensò indubbiamente che lo fosse. Ma a queste distanze i dati sono dominati dalle *velocità peculiari*, il «flusso di Hubble» non si può vedere, ciò che Hubble misurò è poco più di un *fluke* statistico.

Vale la pena di chiedersi se Hubble fosse in qualche modo predisposto a «vedere» quel segnale. Illuminante a questo proposito è il suo articolo sui *Proc. Nat. Ac. Sci.* (15, 169, 1929), dove Hubble anzitutto afferma esistere «... a roughly linear relation between velocities and distances...», aggiungendo poi: «The outstanding feature ... is ... the possibility that numerical data may be introduced into discussions of the general curvature of space». Il riferimento alla Relatività Generale è chiarissimo, anche se il lavoro di Alexandr Friedmann, pubblicato sul *Z. Phys.* nel 1922, non viene citato.

Si tratterebbe quindi di un fatto che dobbiamo dire quasi normale, uno sperimentale che cerca di rintracciare nei suoi dati una previsione teorica. Lascia forse perplessi la mancata citazione.

Dai telescopi da 4 metri, nel dopoguerra, l'analisi del flusso di Hubble fu spinta oltre i 10 Mpc, grazie al lavoro dei gruppi guidati da deVaucouleurs e da Sandage & Tammann. Il valore di  $H_0$ , al finire degli anni Cinquanta, era sceso sotto i 100 (km/s)/Mpc, con Sandage & Tammann che lo portavano fino a 50-55 (km/s)/Mpc. Questa fu la vera scoperta del flusso di Hubble. Sino a allora tutta la comunità scientifica credette a un risultato «pre-maturo» e molte speculazioni su di esso non avrebbero avuto ragione di esistere.

Se una parte importante della «scoperta» di Hubble va allora attribuita a Friedmann e agli altri relativisti della prima ora, ci dobbiamo chiedere su quali dati sperimentali/osservativi si basasse la costruzione teorica dei relativisti stessi. Il punto è che, in realtà, con l'avvento della relatività generale, si poteva impostare compiutamente quel problema cosmologico che persino Isaac Newton aveva tentato di affrontare: «Come evolve una distribuzione isotropa e omogenea di materia autogravitante». A questa domanda Friedmann, Lemaitre, Robertson e Walker stavano dando la risposta corretta.

Ciò che non risultava invece chiaro era se il cosmo potesse, in qualche modo, essere trattato come un continuo isotropo e omogeneo. All'epoca di Friedmann, le distanze accessibili erano intragalattiche. Nella galassia i contrasti di densità sono enormi e la geometria non è certo isotropa, si è ben lontani da un «continuo isotropo e omogeneo». Ma anche 3 anni più tardi, con la scoperta delle galassie, diveniva accessibile una scala non superiore ai 6-8 Mpc e oggi sappiamo che la fluttuazione quadratica media di densità, su una tale scala, è dell'ordine dell'unità. Siamo cioè ancora in un regime ampiamente non-lineare. È vero che si constata una progressiva diminuzione dei contrasti di densità al crescere della scala, ma nulla garantisce che ciò porti a un quadro di generale isotropia e omogeneità. Un contro-esempio, su cui si discusse molto negli anni Ottanta, fu che la distribuzione della materia cosmica fosse un *frattale*. Nessun dato, all'epoca di Hubble, poteva escluderlo. Se la distribuzione è un frattale, o un *multifrattale*, non esiste una scala di isotropia e omogeneità.

I lavori dei relativisti meritano quindi il massimo di considerazione come soluzione di importanti e difficili esercizi di fisica-matematica. Si trattava di problemi che potevano anche avere risvolti sperimentali, ma la cui aderenza alla realtà fisica era tutta da verificare.

Molta letteratura indica negli anni Venti la data di partenza della cosmologia scientifica. Si tratta certamente della data in cui viene certificata la natura extragalattica delle galassie, mattone su cui si costruirà la cosmologia; per questa scoperta, Edwin Hubble merita tutti i riconoscimenti che gli furono poi tributati, fino a aver denominato col suo nome il primo telescopio spaziale. Per quanto concerne la comprensione dell'origine e della dinamica del cosmo, in quegli anni si erano fatte importanti speculazioni fisico-matematiche, selezionando il problema da risolvere, utilizzando la relatività generale, sulla base di un afflato ideologico all'unità. Probabilmente motivata da quelle speculazioni, era poi nata una possibile interpretazione (in realtà prematura) di dati osservativi.

In conclusione: in quell'epoca la cosmologia non era ancora uscita dal paradigma antico; l'uso delle tecniche più avanzate per esprimere valutazioni in ordine all'origine del cosmo.

Di tutto ciò, il mondo scientifico contemporaneo non era consapevole. Il flusso di Hubble si presentava come l'unico dato che facesse uscire la cosmologia dalla minorità prescientifica. Tra i nemici ideologici di un'interpretazione «semplicistica» dell'espansione si annoverano Hermann Bondi, Thomas Gold e Fred Hoyle. Fu quest'ultimo, in particolare, a coniare la denominazione *Big-bang* e nessun termine nato a scopo di irrisione ebbe mai maggior fortuna.

Il conflitto ideologico s'incentrava sull'esistenza (o meno) di un'origine del tempo e dello spazio: facendo scorrere all'indietro il film dell'espansione, tutto tendeva a concentrarsi in un unico punto: ma non uno degli infiniti punti di uno spazio. Non c'era altro spazio che il punto stesso, esso *era* l'Universo. Non ci voleva molto per estrapolare un atto creativo come matrice della realtà fisica nel suo complesso e molti caddero in quella che è oggettivamente una trappola, perché implica un'assolutizzazione immotivata della topologia e della metrica del tempo, anche al di là dell'entrata nel regime della gravità quantistica.

Della trappola non si accorsero neppure i fautori dei modelli di *Stato Stazionario*, la cui tesi veniva rfrasata nei termini di un *Principio Copernicano Assoluto*, come ora ricorderemo.

La cosmologia si basa indubbiamente sull'assunto che la posizione occupata dall'osservatore terrestre sia del tutto generica, su scale maggiori alle scale interne al Sistema Solare. È la generalizzazione alle grandi scale del principio copernicano, che degrada la Terra a sito generico, a dispetto della sua rilevanza «teologica»: la Terra non è il centro dell'Universo fisico.

Se dunque la posizione spaziale dell'osservatore è generica, perché il tempo dell'osservatore deve essere peculiare? Per salvare la genericità dell'epoca attuale, negando l'evoluzione legata all'espansione, occorre che creazione di materia avvenga con

continuità, sia pure in quantità sperimentalmente impercettibili. Essa coagulerà poi in nuove galassie, destinate a occupare i vuoti causati dall'espansione di Hubble. In versioni successive, meno note, la creazione di materia avviene in modo esplosivo, in tanti piccoli *bang*, dando luogo a *apparenti* ammassi di galassie che, come certificavano i dati (trascurando la materia oscura), erano sistemi in espansione.

Di quest'ultima tesi si fece propugnatore, in particolare, l'astrofisico armeno Ambartsumian, e comparvero anche articoli che collocavano la creazione di materia negli stessi nuclei galattici; citiamo a memoria il titolo di un articolo dell'epoca: «*About an explosive origin of galactic spiral arms*». Le braccia delle spirali galattiche emesse dal nucleo, dove la materia aveva origine.

La disputa sopravvisse alla definizione statistica del flusso di Hubble e alla fissazione di valori realistici per  $H_0$ . Ma ormai il suo tempo era agli sgoccioli. Già i dati sulle *radiogalassie* apparivano in contrasto con il principio copernicano assoluto. Infine, nel 1965, Penzias e Wilson notarono per caso, tramite un radiometro a microonde costruito per la Bell Telephone, una radiazione di intensità indipendente dall'orientazione cosmica dello strumento, debole ma costante. Si trattava del CMB, atteso in modelli di *Big-bang*, non comprensibile in modelli di *Stato stazionario*.

La cosmologia stava diventando scienza.

## 7. La cosmologia e la fisica

È in questo periodo che, dato per acquisito il paradigma del *Big-bang*, si inizia uno studio sistematico dell'evoluzione del cosmo a temperature via via crescenti, utilizzando la fisica nota, in contesti primordiali.

Della scienza costruita allora si occupa ampiamente questo libro; si tratta di lavoro la cui originalità sta nello scoprire che le stesse forze, viste in azione nel laboratorio, agivano quando il contesto cosmico era radicalmente difforme, privo dell'articolazione necessaria a ospitare qualunque sorta di osservatore.

In parte si tratta di lavoro che prosegue speculazioni già iniziate prima della scoperta del CMB, in un tempo in cui alcuni fisici speculavano in varie direzioni, quasi un *divertissement* a proporre curiose aspettative. In effetti, oltre all'alternativa tra *steady-state* e *Big-bang*, c'era stata la disputa se il *Big-bang* fosse stato *caldo* o *freddo*. Si trattava in sostanza di fissare il rapporto primordiale tra numero barionico e entropia. Molto notevoli, in questa fase, furono le speculazioni di George Gamow, un fautore del *Big-bang* caldo. In sostanza egli costruì un valido schema della BBN (*Big-bang nucleosynthesis*, la nucleosintesi primordiale). Dai dati sull'abbondanza di  ${}^4\text{He}$  nelle stelle, Gamow arrivò a predire l'esistenza di un *fondo cosmico* di radiazione dallo spettro termico, per il quale prevedette una temperatura di 4-5 K.

I suoi risultati, talora contestati dalle più accreditate riviste scientifiche, andarono anche in testi divulgativi. Ciò di cui Gamow non si preoccupò fu di accertarsi se una tale radiazione cosmica sarebbe stata visibile da terra o avrebbe avuto altre conseguenze osservabili. Ancor prima della rilevazione di Penzias e Wilson, c'erano dati

astronomici che mostravano che l'equilibrio tra l'occupazione dei livelli atomici e molecolari, negli spazi interstellari, era alterato rispetto a quello di temperatura zero. Ma dovette giungere la scoperta dei due tecnici della *Bell-telephone* perché si passasse dalle speculazioni alla scienza.

Fino agli anni Ottanta, dunque, una generazione di cosmologi lavorò all'uso di *fisica nota* nel contesto abnorme del cosmo primordiale, con notevoli risultati.

La percezione definitiva che la cosmologia era divenuta in tutto una branca della scienza fisica la si ebbe comunque all'inizio degli anni Ottanta, quando da Kazanas e Guth giunsero le prime proposte dell'*Universo inflazionario*. Un paio d'anni dopo parve che un cerchio si chiudesse, quando Andrei Linde e Albrecht & Steinhardt posero alla base della dinamica inflazionaria il *campo di Higgs* responsabile per la rottura della simmetria *GUT*, tra forze forti e elettrodeboli.

Anche se tali proposizioni sono state successivamente superate, il laboratorio cosmologico è ormai divenuto la palestra nella quale le idee più avanzate di fisica fondamentale vengono elaborate per essere portate al confronto coi dati.

Ma le ricadute della cosmologia sulla fisica fondamentale vanno ben oltre, specialmente con la scoperta delle componenti oscure del cosmo. Quasi tutta la fenomenologia delle interazioni fondamentali rientra oggi nel cosiddetto «modello standard», basato sul «dogma di gauge» e sul gruppo di simmetria  $U_Y(1) \otimes SU_L(2) \otimes SU_c(3)$ . Oltre un migliaio di dati microscopici rientrano nelle previsioni del modello. I fisici però concordano nell'affermare che il modello standard è... brutto. Essenzialmente perché esso dipende da oltre 20 costanti, *a priori* arbitrarie, mentre arbitrario e stranamente articolato è anche il gruppo di simmetria. Perciò è fondamentale la ricerca di «fisica oltre il modello standard», che certamente esiste, ma si potrebbe esplorare sperimentalmente, con certezza, solo a energie molto superiori a quelle raggiungibili con gli acceleratori.

Va anche notato che le sole evidenze di rottura del modello standard vengono dal mondo dell'astrofisica. Esse riguardano in primo luogo i neutrini prodotti nell'alta atmosfera dai *raggi cosmici* e i *neutrini solari*.

Ma i due elementi fondamentali che testimoniano dell'esistenza di una fisica articolata e complessa oltre il modello standard sono l'esistenza della *materia oscura* e dell'*energia oscura*.

La scoperta delle componenti oscure è probabilmente il più grande risultato della ricerca cosmologica, dopo il flusso di Hubble e il CMB. La loro importanza risulterà subito chiara al lettore, sapendo che la materia ordinaria, quella di cui si compone il corpo umano, i pianeti e le stelle, usualmente definita *materia barionica*, è meno del 5% dei contenuti globali dell'Universo. Un altro 20-22% della sostanza cosmica è *materia oscura*, mentre il 70-75% è *energia oscura*.

Cosmologia e fisica fondamentale convergono oggi nei programmi di ricerca, mirando alla comprensione della natura di queste componenti, fatte di particelle o di campi che non rientrano nel *modello standard* e ciò fa apprezzare quanto lontana sia ormai la cosmologia dalle sue stesse origini culturali.

La novità principale, rispetto alle cosmologie antiche, non è dunque l'uso del linguaggio matematico a elaborare archetipi antichi o a risolvere significativi problemi fisico-matematici, ma il confronto con l'esperimento e i suoi dati. Sono i dati sul CMB a porre la parola fine sulle elucubrazioni dello stato stazionario. Sono i dati sulle radiosorgenti, i QSO e gli AGN, a rendere cogente l'evoluitività dell'Universo. Sono i dati sull'ampiezza delle anisotropie del CMB a portare alla scoperta della materia oscura. Il procedimento della cosmologia è ormai totalmente quello della fisica.

A fronte dei dubbi antichi, per cui furono edificate le cosmologie tradizionali, e che ancora permeano la dialettica classica, tra Essere e Divenire, la cosmologia scientifica ormai prospetta la definitiva risposta che il divenire del mondo è la base stessa del suo essere.

## 8. Antiche dispute, nuovi veleni

In sostanza, di fronte all'alternativa tra *tempo ciclico*, con gli eterni ritorni, e un *tempo lineare* – una semiretta, forse un segmento – le evidenze sembrano a favore della seconda scelta. Alla fine del libro torneremo comunque su questo punto, svolgendo una disamina articolata sul concetto di tempo, nelle sue accezioni *topologica* e *metrica*. Essa ci confermerà che una durata finita nel tempo è legata alla scelta della metrica, mentre il discorso topologico può essere estremamente articolato.

Andare più oltre, richiede di fare ipotesi *ad-hoc*, sia sul contorno cosmologico che sul contenuto fisico. È un'industria cui alcuni si sono dedicati e che, in sostanza, ritorna a vecchi schemi: quali sono le ipotesi *metriche* e *fisiche* che debbo fare, perché la storia del cosmo sia stata di un certo tipo?

Opportune ipotesi permettono di ricostruire un cosmo multieonico, in cui il *Big-bang* risulti da una struttura formatasi in un pre-Universo. Come se determinate singolarità metriche, i *buchi neri* evolutisi in presenza di opportuna fisica fondamentale, possano a loro volta essere il punto di partenza di nuovi *bang*, in una catena infinita.

Allo stato attuale della ricerca cosmologica queste formulazioni sono possibili ma i loro risultati non vanno intesi come acquisizioni scientificamente definite. Esse passano sotto il nome di *metacosmologia* e, di per sé, potrebbero avere una validità analoga alle speculazioni di Gamow, quando prevedeva il CMB sulla base dei dati su  $^4\text{He}$ . Ciò che attendiamo quindi, da queste speculazioni, è la previsione di un qualche fenomeno osservabile.

Il pericolo insito in questo lavoro è che esso venga posto sullo stesso piano della ricerca di fisica cosmologica. Già si sono visti su riviste di divulgazione scientifica rappresentazioni di quadri *metacosmologici*, senza che fosse rimarcata la loro incompletezza metodologica. Queste rappresentazioni servono talora a confortare visuali «cosmologiche» in cui la cosmologia non è intesa come scienza moderna, ma nella sua antica accezione, rivestendole del linguaggio più avanzato che la nostra cultura è in grado di produrre. Col rischio di mettere a repentaglio, assieme alla scelta culturale, la validità dello strumento d'indagine; non solo la cosmologia, dunque, ma lo strumento scientifico della conoscenza.

## CAPITOLO II

# Nel cielo le tracce dell'origine

## Gli abitatori del cielo e la scala di omogeneità

### 1. Astronomia e cosmologia

Una delle caratteristiche più pregnanti dell'animo umano è *Ulisse*, la spinta all'esplorazione. In un racconto di fantascienza degli anni Cinquanta, gli esploratori di un pianeta, in un remoto sistema solare della Via Lattea, vi scoprivano le *piante pensanti*. Esse condividevano con l'animo umano il desiderio di sapere e anzi, con la sola *osservazione* unita a enormi capacità logiche e computazionali, avevano ricostruito la fisica della materia e la storia del cosmo, dalle sue origini. Ma, anche per questo, giudicavano risibile l'impulso umano a esplorare.

Lo scienziato, invece, è Ulisse. Esplora l'intimità della materia vivente o inanimata, gode delle sorprese che si presentano dietro le svolte del suo cammino. Quando un esperimento sta finendo, ha già in mente il successivo, per avanzare nelle terre che il primo ha lasciate inesplorate.

Dove tradizionalmente Ulisse ha dovuto segnare il passo è di fronte al Cielo. Dal mito di Icaro ai favolosi tentativi di Leonardo, l'angoscia di essere schiacciati alla superficie del pianeta è stata faticosamente rimossa e il *celeste* si è confuso col *divino*. Con la scoperta del motore a scoppio e l'ardimento dei primi aviatori (Gabriele D'Annunzio: «Volare necesse, vivere non necesse») parve che il vaso di Pandora fosse finalmente scoperchiato. Purtroppo, al pacifico raid su Vienna, fecero seguito il bombardamento di Dresda e i lanci delle V2. Ma, finalmente, il sogno umano di esplorare altre terre parve sul punto di concretizzarsi nei giorni dell'Apollo 11.

Eppure, più di questo ardire verso i cieli, suona condanna il limite fisico alla velocità degli spostamenti. In un tempo uguale a quello trascorso dagli albori della civiltà, muovendosi alla velocità della luce, la nave degli esploratori potrebbe raggiungere una distanza  $\sim 1.5$  kpc, poco più di 1/10 del raggio della Via Lattea.

Il volo interplanetario ha mutato il quadro dell'astronomia del Sistema Solare; che si fa sempre meno dalla specola dello Schiapparelli e sempre più con le sonde della NASA. La futuribile scoperta di tecnologie capaci di avvicinare il limite della velocità della luce, potrebbe mutare il quadro dell'astronomia stellare. Ma, di fronte all'a-

stronomia che conta ai fini cosmologici, Ulisse è definitivamente frustrato. L'astronomo cosmologo, come l'astronomo ottocentesco, è condannato a restare «pianta pensante», seppure il suo *lebensraum* si estenda ormai ben oltre i confini del pianeta e gran parte dell'input per le sue capacità logiche e computazionali provenga da dati non più censurati dalla bidimensionalità degli spazi accessibili e dall'opacità dei gas atmosferici.

In questo capitolo ripercorreremo i legami tra la nuova astronomia e la cosmologia. Entreremo talora anche negli aspetti tecnici della scienza astronomica, quando ciò è reso necessario a spiegare la dialettica del discorso cosmologico. Altrimenti, dall'astronomia mutueremo nozioni acquisite, dando loro un ordine strumentale al nostro discorso. La cosmologia è principalmente basata su nozioni astrofisiche, ma esiste e progredisce solo grazie agli sviluppi della fisica fondamentale.

Si tratta di un campo in cui, viste le proporzioni degli spazi intergalattici, molto al di là di ogni concetto di esplorabilità, potrebbe sembrare che Ulisse abbia ceduto il passo alle «piante pensanti». Ma, se ciò è vero per l'astronomo extragalattico, non è più vero per lo scienziato cosmologo che, partito dalle nozioni astrofisiche di base, compie ormai gran parte della sua esplorazione nella fisica delle altissime energie.

## 2. Il flusso di Hubble

Nel capitolo precedente abbiamo dibattuto il graduale trapasso dalla cosmologia come fatto culturale alla cosmologia come scienza. Non c'è dubbio che il primo elemento si cui quest'ultima di basa sia la legge

$$(II.1) \quad v \simeq H_o d ,$$

che fissa un sostanziale rapporto di *proporzionalità* tra velocità di allontanamento  $v$  e distanza  $d$  delle galassie osservate. Si parametrizza la costante di proporzionalità ponendo

$$(II.2) \quad H_o = 100 h \text{ (km/s)/Mpc}$$

e  $h$  viene denominato *parametro di Hubble adimensionale*. I dati più recenti danno

$$(II.3) \quad h \simeq 0.72 \pm 0.03 .$$

A questo valore si arriva convolvendo i risultati di diverse tecniche, per esempio usando i dati sullo spettro delle anisotropie del CMB (la radiazione di fondo a 2.73 K), ma il modo principe resta comunque l'utilizzo di dati profondi sul moto delle galassie, di modo che le velocità di allontanamento effettivamente osservate

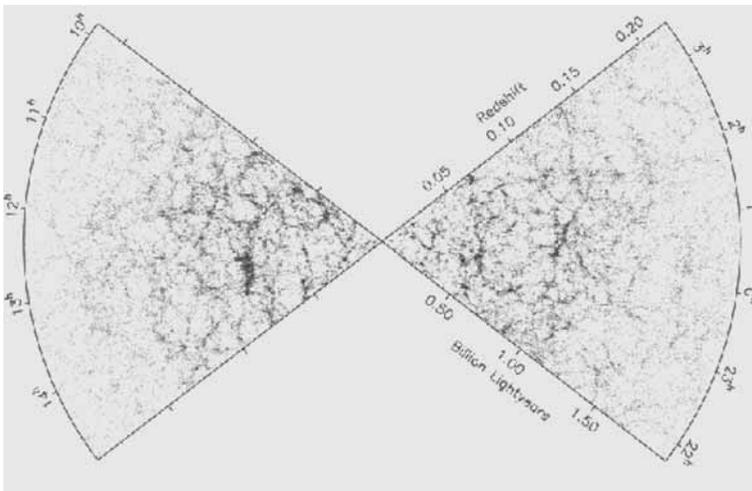
$$(II.4) \quad v = H_o d + v_p \cos \theta ,$$

siano poco inquinate dal rumore dovuto alle *velocità peculiari*  $v_p$  (di cui conta la proiezione nella direzione di vista,  $\theta$  è l'angolo tra  $v_p$  e tale direzione). La velocità peculiare della Via Lattea, per esempio, è circa un millesimo della velocità della luce  $c$ ,

quindi  $\sim 300$  km/s; si tratta di un valore abbastanza tipico e si può notare che  $H_0 d$  raggiunge un valore  $\sim 300$  km/s a una distanza  $d_h \sim (3/h)$  Mpc. Per  $h \simeq 0.71$  sarà  $d_h \simeq 5$  Mpc, sicché un segnale significativo sul *flusso di Hubble* si avrà oltre i 10 Mpc, mentre non ci sarà nessun segnale sotto i 5 Mpc. A conferma dell'impossibilità di confrontare l'eq. (II.1) coi dati di Hubble, sta il valore allora fornito per  $H_0$ : sia pure rinormalizzandolo in modo da scontare l'errore sulla distanza delle stelle cefeidi, che Hubble e Humason usarono come indicatori di distanza, si ottiene  $\sim 170$  (km/s)/Mpc. Se quel valore fosse vero, l'età del cosmo sarebbe di 6 miliardi di anni; cosa assurda, visto che in esso vi sono sicuramente stelle più vecchie di 10 miliardi di anni.

Nel primo capitolo si è poi anche visto che, anche se l'induzione di Hubble era in qualche modo fondata sui modelli relativistici del cosmo, resta il fatto che l'ipotesi di omogeneità e isotropia su grande scala su cui essi si basano, a quel tempo, era del tutto strumentale.

A distanza di ottant'anni, disponiamo di campioni statisticamente ben selezionati sul tipo di quello rappresentato in figura II.1, che includono galassie fino a distanze anche maggiori di 1 Gigaparsec, mille volte più lontane di quelle viste da Hubble. Oltre a confermare la coerenza del flusso di Hubble, l'analisi di questi campioni permette di confermare che le ipotesi su cui si basavano i modelli relativistici. *In primis*, è confermato che, salendo di scala, ci si approssima sempre più a un regime di isotropia e omogeneità. Accertare questo secondo punto non è stato un lavoro da poco e questo capitolo è principalmente dedicato alle tecniche con cui si è gradualmente giunti a mostrare la ragionevolezza di tale assunto.



**FIGURA II.1**

Il campione di galassie denominato 2dF. Ogni punto corrisponde alla collocazione di una singola galassia. Nel suo insieme il campione comprende circa 200 000 galassie. Sono indicate in figura una coordinata celeste, il *redshift* e la distanza in miliardi di anni-luce

### 3. Una finestra sull'Universo

Un punto fondamentale, senza cui non sarebbe possibile «fare cosmologia», è la constatazione che il cosmo è praticamente vuoto, la sua trasparenza si prospetta ottimale. (Ci si può chiedere quale sarebbe stata la storia della scienza se la Terra fosse stata immersa in una *nebula* di polveri, come ce ne sono molte nella Via Lattea. Si tratta di *nebulae* di densità assai modesta, sicché l'esistenza di uno schermo davanti al cosmo non sarebbe stata facilmente rilevata.)

Mediando sulle scale cosmiche, un cubo di lato pari alla distanza Terra-Luna contiene in media la massa di una grossa mela, ma non basta: il 96% della mela è fatto di sostanze diverse dalla materia *barionica* di cui si compongono Sole, Terra, pianeti e, in definitiva, il nostro corpo. Tali sostanze aliene sono la *materia oscura* e l'*energia oscura*, sulle cui proprietà discuteremo a lungo.

Accanto a esse ci sono le *radiazioni*, elettromagnetica e neutrinica. Il contributo dei loro quanti, *fotoni* e *neutrini*, al budget energetico è oggi trascurabile. La radiazione elettromagnetica è invece essenziale nel portarci informazione diretta sulla distribuzione nel cosmo della materia barionica. La distribuzione delle altre componenti, invece, può solo essere ricostruita tramite argomentazioni astrofisiche. Infatti *solo la materia barionica emette o assorbe fotoni*.

Si pensi che un fotone, emesso 13 miliardi di anni or sono, sulla *banda dell'ultimo scattering*, ha 9 probabilità su 10 di giungere indisturbato sino a noi. Il 99% della probabilità di essere assorbito nel cammino si colloca poi a un tempo preciso, quando il cosmo aveva 1/10 dell'età attuale.

Se la finestra sull'Universo è ben aperta, il compito della *nebula* viene assolto in parte dell'atmosfera terrestre. Ma anche l'atmosfera dell'attuale era geologica è generosa: essa lascia passare, anzitutto, la radiazione della *banda ottica* tra  $\sim 4$  e  $7 \times 10^{14}$  Hz, ragion per cui l'evoluzione ha forgiato occhi sensibili ai fotoni della luce... visibile. Sulle altre bande, si va da un assorbimento totale alle alte frequenze, a partire dall'ultravioletto (UV), oltre il  $10^{15}$  Hz, a una trasparenza simile a quella ottica su ampie bande delle frequenze radio ( $10^4$ - $10^8$  Hz). Nella zona intermedia andiamo da un assorbimento fortemente variabile a seconda della frequenza nell'infrarosso (IR), a una elevata trasparenza alle microonde (MW), attorno ai  $10^{10}$  Hz, che permise a Penzias & Wilson di scoprire il CMB (*cosmic microwave background*).

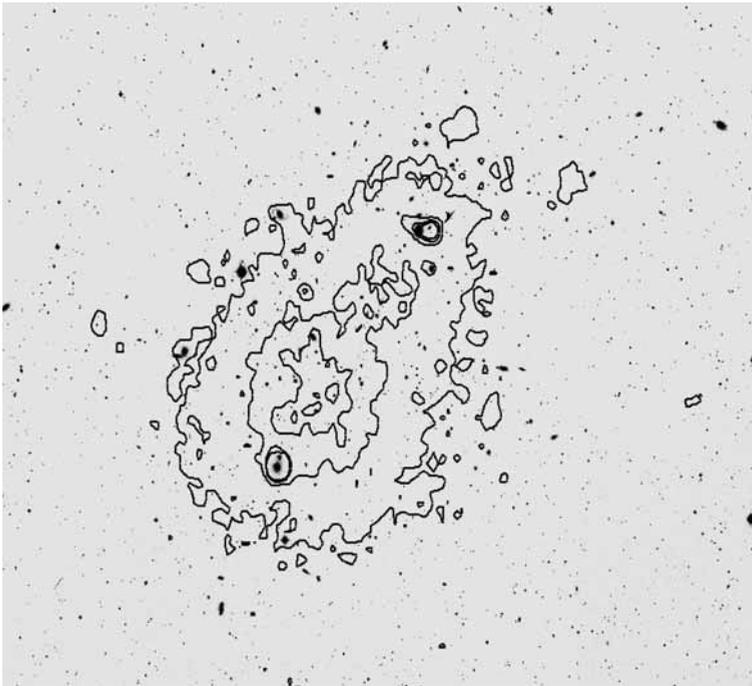
Alla finestra ottica lavorano i telescopi. A molti di essi possono essere agganciati apparati sensibili all'IR. Fin dalla fine degli anni Cinquanta, sono stati costruiti i *radiotelescopi*, che scoprirono rapidamente un enorme numero di *radiosorgenti*. La distribuzione delle loro luminosità radio (la curva Log N-Log S) fu essenziale nello stabilire che l'«Universo è evolutivo», diverso oggi dal passato. Su questa curva torneremo più avanti in questo stesso capitolo.

Per l'osservazione nelle altre zone spettrali, nell'UV, nella banda X e nella banda  $\gamma$  dovette invece attendere il lancio di sonde extra-atmosferiche. Ma il loro lancio sfiora

ormai l'assoluta normalità e anche l'opacità atmosferica è ormai quasi solo un problema da relegare nella storia dell'astronomia

L'astronomia X è oggi un vasto capitolo della scienza astrofisica. Le sorgenti X, galattiche e extragalattiche hanno la stessa importanza, almeno, delle sorgenti ottiche. Alcuni oggetti cosmici, già noti prima dell'astronomia X, sono stati compresi nella loro reale essenza solo tramite i dati X. Primi tra questi gli *ammassi di galassie* di cui, di fatto, le galassie non costituiscono più del 20% della massa barionica. Il grosso dei barioni di un ammasso si trova invece nella forma di *gas caldo* ma trasparente, che emette copiosamente in X per *bremstrahlung*. In figura II.2 riportiamo l'immagine dell'ammasso Abell 1367 nel *visibile* e in *banda X*. Si è scoperta anche una radiazione di fondo, in banda X, che oggi si interpreta come sintesi di sorgenti profonde non risolte.

Va qui aggiunto che un ulteriore meccanismo di rivelazione degli ammassi sta diventando ormai operativo. Esso si basa sul cosiddetto effetto SZ (*Sunyaev-Zel'dovich*), su cui daremo qualche ulteriore dettaglio nel capitolo sul CMB. In sostanza, il gas caldo dell'ammasso urta una minuscola frazione dei fotoni del CMB che attraversano l'ammasso. Ma i fotoni del CMB sono moltissimi, più di 400 per  $\text{cm}^3$ , e



**FIGURA II.2**

L'ammasso Abell 1367. Sono qui sovrapposte le immagini ottica (le macchie scure rappresentano le galassie) e le isofote dell'emissione in banda X, dovuta a gas a temperatura  $\sim 5 \cdot 10^7$  K.

anche una piccolissima frazione di essi, se spostata su un'altra banda, diviene osservabile.

L'astronomia  $\gamma$  ha avuto invece origini sconcertanti. Esplosioni di raggi  $\gamma$  vennero scoperte da satelliti spia, americani e sovietici, costruiti per monitorare i test nucleari. La scoperta, di cui i militari non capirono significato e portata, rimase a lungo classificata. Il cielo  $\gamma$  è caratterizzato da una straordinaria variabilità; non ci sono importanti sorgenti fisse di raggi  $\gamma$ , ma rapidi *flares*, di cui si è disputata a lungo l'origine galattica o extra-galattica.

Oggi sappiamo che i  *$\gamma$ -ray-burst* (GRB) traggono origine da esplosioni avvenute nel passato cosmologico, sconosciute nel mondo d'oggi. Lo studio delle sorgenti  $\gamma$  si rivelerà probabilmente fondamentale per la conoscenza delle prime popolazioni di oggetti cosmici, ma i suoi dati non hanno sinora trovato applicazione alla comprensione del cosmo nel suo complesso.

Abbiamo sin qui lasciato in disparte la radiazione neutrinica. I neutrini hanno una piccolissima sezione d'urto. A differenza dei fotoni (almeno alcuni dei) neutrini hanno una massa, sicché il loro contributo al budget energetico potrebbe essere confrontabile al contributo della materia visibile nelle galassie. In ogni  $\text{cm}^3$  ci sono infatti varie centinaia di neutrini, che però non possono essere percepiti tramite alcun apparato oggi esistente o progettabile. Essi appartengono a un fondo neutrino in tutto analogo al CMB fotonico.

Neutrini più energetici possono essere invece rivelati grazie alle trasmutazioni che essi inducono tra determinati elementi chimici; tra le reazioni utilizzate è importante quella che trasforma un nucleo di *Gallio* in un nucleo di *Germanio*. Tra le sorgenti neutriniche a alta intensità così rivelate c'è il Sole; il flusso di *neutrini solari* è stato misurato, tuttavia, con apparati inadeguati a fissare la direzione del segnale. Il passaggio a apparati direzionali, oggi in costruzione, segna la transizione a veri e propri *telescopi neutrinici*.

Tra le sorgenti cosmiche di neutrini vanno annoverate le *Supernovae*. Neutrini provenienti dalla SN 1987a sono stati misurati da vari apparati, con risultati importanti per la fisica delle interazioni fondamentali e l'astrofisica.

Un'ulteriore sorgente di dati sul cosmo è costituita dai *raggi cosmici*, particelle dotate di massa che giungono dallo spazio. La composizione dei raggi cosmici ha rilevanza cosmologica solo indiretta. Alcuni dati basati sui raggi cosmici verranno citati nel testo.

#### 4. Galassie, gruppi e ammassi

Un astronomo del primo Ottocento non avrebbe avuto dubbi. Oggetto primario dell'indagine astronomica erano i pianeti e le stelle. C'erano poi i pianetini e le *stelle comete*, fenomeni «importantissimi», perché transienti. Alla fine dell'Ottocento Messier compilò un catalogo delle *Nebulose* allora note, per evitare ai cacciatori di comete di scambiare una nebulosa nota per una nuova cometa. Molte delle nebulose di Messier sono galassie.

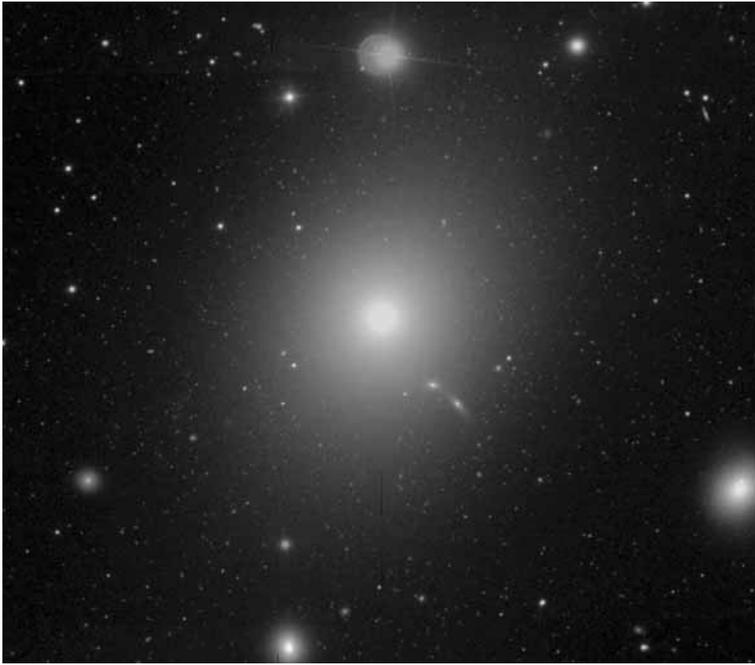
**FIGURA II.3**

Immagine di M87, l'ellittica gigante al centro dell'ammasso della Vergine. Ai bordi destro, superiore e inferiore dell'immagine, nonché in basso presso il bordo sinistro, sono visibili le immagini di altre ellittiche nane, satelliti di M87; le immagini galattiche prossime a M87, invece, sono dovute a oggetti di *background*. I globi rossi nascono dalla saturazione di immagini stellari. (si veda l'appendice TAVOLE A COLORI).

Dopo che Hubble ebbe collocato le galassie al di fuori della Via Lattea, grazie all'uso di specchi di dimensioni crescenti e tecniche fotografiche sempre più avanzate, gli astronomi hanno ottenuto immagini di galassie, come quelle nelle figure II.3 e II.4, che colpiscono il senso estetico e quasi divengono in sé una ragione dell'indagine astronomica. La massa delle galassie va da  $10^8$  a  $10^{12}M_{\odot}$  (masse solari), ma si considerano galassie anche oggetti cosmici minori, di  $10^6$ - $10^7M_{\odot}$ , o enormi ellittiche al centro di ammassi, contenenti oltre  $10^{13}M_{\odot}$ .

Lasciando in disparte il senso estetico, è importante sottolineare invece che le galassie sono gli oggetti «più piccoli» la cui genesi può essere direttamente legata al campo di tenui fluttuazioni primordiali, sviluppatasi per autogravità. Il processo è dapprima lentissimo, ma subisce una progressiva accelerazione quando il contrasto di densità, tra l'interno di una fluttuazione e la media cosmica, avvicina l'unità.

In via del tutto generale è legittimo chiedersi se tutta la dinamica dello sviluppo di fluttuazioni primordiali possa essere ricondotta alla sola interazione gravitazionale. In questo contesto gli effetti delle altre interazioni verranno denominati *dissipativi*, in



FIGURA II.4

Immagine della galassia a spirale M100. Si distinguono, oltre ai bracci brillanti di luce stellare, anche dei tenui bracci color arancio, dovuti a assorbimento da polveri. (si veda l'appendice TAVOLE A COLORI).

quanto alterano il bilancio energetico delle forze di gravità. Ebbene, la scala galattica, così cospicua dal punto di vista osservativo, marca il confine tra la dinamica puramente gravitazionale sulle grandi scale e l'intervento delle forze dissipative, nello stabilire le configurazioni osservate, sulle scale minori.

Si tratta comunque di un confine permeabile, dipendente dalla geometria dei sistemi in formazione e dalla loro eventuale immersione in sistemi di dimensioni maggiori; di qui, l'oggettiva variabilità delle masse delle galassie osservate.

Va altresì notato che le forze dissipative sono all'opera anche al di sopra della scala galattica. Abbiamo già citato l'emissione X, da parte del gas caldo negli ammassi di galassie, per *bremstrahlung*, cioè a causa dell'accelerazione delle cariche elettroniche nel campo di altre cariche. Si tratta di processo che produce una potenza

$$(II.5) \quad W \propto \rho^2 R^3,$$

in quanto la probabilità degli incontri è proporzionale al quadrato della densità  $\rho$  (delle cariche) e l'emissione totale cresce col volume  $R^3$ . Esso è talmente cospicuo da costituire il principale canale di osservazione di gruppi ad ammassi, senza però influire in modo determinante sulla loro dinamica.

### Il teorema del viriale

Per meglio delineare la dialettica tra forze gravitazionali e dissipative, seguiamo la *crescita non-lineare* delle fluttuazioni, quando i materiali vengono riscaldati dal lavoro delle forze di gravità e la loro pressione aumenta fino a bilanciare la spinta gravitazionale. In assenza di dissipazione, il sistema evolve verso una configurazione *quasi-stazionaria* di *equilibrio viriale*.

Il teorema del viriale stabilisce che un sistema di molte particelle raggiunge una configurazione stazionaria se le sue energie cinetica e potenziale gravitazionale soddisfano la relazione:

$$(II.6) \quad 2E_{cinetica} + E_{potenziale} = 0.$$

Notiamo che un pianeta che percorre un'orbita circolare di raggio  $R$  a velocità  $v$  ha accelerazione centripeta  $v^2/R$  pari all'accelerazione di gravità dovuta al Sole,  $GM_{\odot}/R^2$ , sicché per esso risulta

$$(II.7) \quad v^2 - GM_{\odot}/R = 0.$$

Il primo termine, moltiplicato per la massa  $m$  del pianeta è pari al doppio della sua energia cinetica; il secondo termine, sempre moltiplicato per  $m$ , ne dà invece l'energia potenziale gravitazionale: la condizione di equilibrio dinamico su un'orbita circolare ricalca quindi la condizione di equilibrio viriale.

Raggiunto l'equilibrio viriale e in assenza di meccanismi capaci di drenare energia, il contrasto di densità non cresce più; se crescesse sarebbe come se l'orbita del pianeta non fosse stabile e il pianeta potesse cadere verso il Sole.

Si tratta comunque di un equilibrio dinamico; come il pianeta continua a muoversi, così le «particelle» in equilibrio viriale percorrono traiettorie articolate, soggette alla mutua interazione.

Gli *ammassi di galassie* sono il più cospicuo esempio di sistema in equilibrio viriale; le loro dimensioni e massa tendono a congelarsi, anche se le forme possono mutare su lunghe scale di tempo, molto al di là di quelle osservative.

### Distribuzione spaziale delle galassie

Per un pianeta in orbita attorno al Sole non ci sono meccanismi che producano oggi significative perdite di energia. Invece, un sistema che perda energia con la legge (II.5), vede mutare il suo raggio e le sue caratteristiche morfologiche sulla scala di tempo  $\tau_{em} = E_k/W$  ( $E_k$  è la sua energia cinetica,  $W$  la potenza emessa). Questo tempo cresce con le dimensioni del sistema, in modo  $\propto R^2$ , in quanto  $E_k \propto \rho^2 R^5$  e  $W \propto \rho^2 R^3$ .

La seconda legge di proporzionalità è già in eq. (II.5). Nello studiare la dipendenza dell'energia dalla scala  $R$  possiamo basarci sulla dipendenza da  $R$  dell'energia potenziale  $E_p$ , dato che il sistema è in equilibrio viriale, sicché  $E_k = -E_p/2$ .

Essendo allora

$$(II.8) \quad E_p \simeq GM^2/R \quad \text{e} \quad M \simeq \rho R^3,$$

risulta  $E_k \propto \rho^2 R^5$ . Nel rapporto  $E_k/W$ , pertanto, le dipendenze da  $\rho$  si elidono e rimane solo una dipendenza da  $R$ .

La scala  $R_g$  per cui  $\tau_{em} = E_k/W$  raggiunge l'età del cosmo segna il limite oltre cui le forze dissipative non influiscono più sulla dinamica, vale a dire la scala galattica. La conclusione è che, per definizione, *nel ricostruire la genesi della distribuzione delle galassie, possiamo sostanzialmente trascurare ogni azione non-gravitazionale.*

### 5. Le scale cosmologiche

Quando studiamo la distribuzione della materia su scala super-galattica, ci dobbiamo dunque preoccupare delle sole forze gravitazionali. Ma fin dove potremo utilizzare la gravità newtoniana? esistono scale su cui gli effetti relativistici sono imprescindibili, sicché diviene imperativo l'uso della Relatività Generale?

Il quadro del cosmo descritto in figura II.5, sintesi di secoli di dati astrofisici, ci

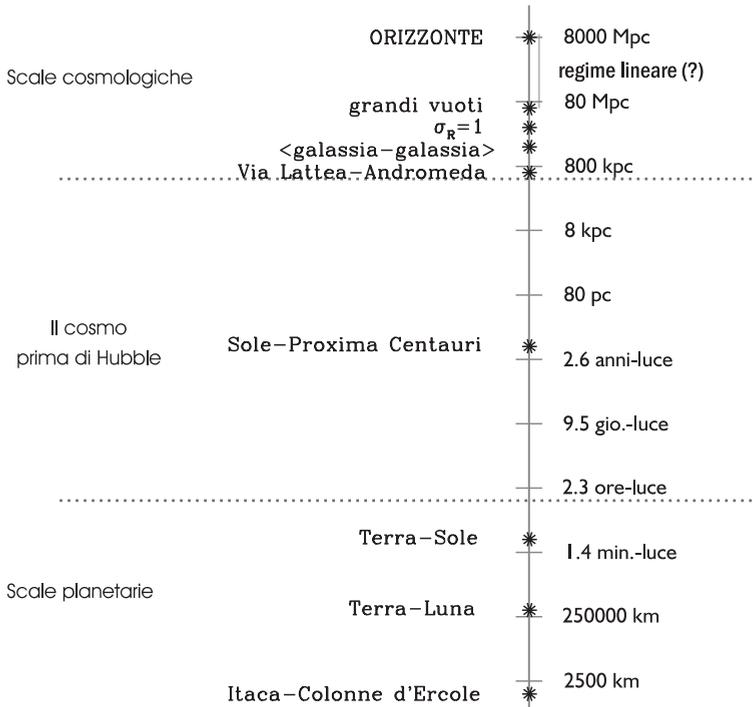


FIGURA II.5

Le scale cosmiche. A ogni tacca sull'asse verticale le distanze crescono di un fattore 100. Il grafico è realizzato assumendo  $H_0 \simeq 70$  km/s/Mpc. (si veda l'appendice TAVOLE A COLORI).

mostra la situazione cui dobbiamo fare riferimento. Sull'asse verticale, dal basso all'alto, sono indicate scale via via crescenti; a ogni tacca la scala cresce di un fattore 100. Ci sono due grossi iati, prima tra le scale planetarie e le scale stellari – dalle ore-luce agli anni-luce – poi dalle scale intragalattiche alle distanze tra galassie – da pochi kiloparsec al Megaparsec. Il grafico assume un parametro di Hubble  $H_0 = 70$  km/s/Mpc. (Il parsec, pari a 3.26 anni-luce, è l'unità naturale per le distanze esterne al Sistema Solare, come spieghiamo nel prossimo paragrafo.)

Le distanze indicate da punti stellati, sulle scale planetarie e per il cosmo pre-Hubble, hanno un significato immediato. È anche chiaro cosa siano la distanza tra Via Lattea e Andromeda e la distanza media tra galassie.

Con  $\sigma_8 = 1$ , invece, ci si riferisce alla scala su cui la fluttuazione quadratica media di densità (detta *varianza di massa*) ha valore unitario. Sotto tale scala la dinamica è non-lineare; ne discuteremo ancora diffusamente. Tuttavia, anche quando le fluttuazioni, in media, sono inferiori all'unità, ci saranno siti dove si osservano eccessi o carenze «anomali» di densità. Ivi sono i grandi ammassi o i grandi vuoti; il diametro di questi ultimi raggiunge i 40-50 Mpc.

Infine, attorno a 8000 Mpc, si raggiunge l'orizzonte cosmologico, la massima distanza da cui ci possono giungere segnali. L'esistenza di questa scala, che non esclude la realtà di distanze ulteriori, è un'acquisizione relativistica; le distanze accessibili nel cosmo hanno un limite, non si prolungano all'infinito, l'osservabilità si riduce ai contenuti di una bolla di un infinito meta-Universo.

La linea violetta, in alto a destra, indica le scale che stanno tuttora evolvendo in «regime lineare». Su quelle scale ci sono fluttuazioni di densità, ma la differenza tra densità media e densità locale va progressivamente attenuandosi, al crescere della scala, finché diviene fisicamente sensato considerare il cosmo come un continuo isotropo e omogeneo. Alla fine della lunga scala si recupera il modello che i relativisti «della prima ora» avevano trattato come problema fisico-matematico. Ricordiamo comunque che, quando compare l'articolo di Friedmann sul *Zei. P.*, l'astronomia aveva sondato scale fino a 20 kpc. Anche 3 anni dopo, quando Hubble colloca le galassie al di fuori della Via Lattea, le distanze raggiunte dall'ispezione astronomica non superavano i 6 Mpc. Il regime lineare non era ancora stato raggiunto, tanto meno una scala su cui valesse omogeneità e isotropia.

## 6. Le distanze cosmiche

La costruzione di un grafico delle scale cosmiche, come quello esibito nel precedente paragrafo, richiede una buona capacità di misurare le distanze tra corpi celesti. Il primo gradino nella storia della misura delle distanze extraterrestri venne compiuto, nel 1672, da Giandomenico Cassini, astronomo bolognese trapiantato in Francia, realizzando una triangolazione sul pianeta Marte. Alla stessa ora siderale dello stesso giorno, Cassini e il suo assistente Richer, rispettivamente da Parigi e dalla Guaiana, misurarono la posizione angolare di Marte nel cielo. Nota la distanza Parigi-Caienna e

misurati i due angoli, se ne dedusse la distanza di Marte e, con essa, tutte le scale interne al Sistema Solare. A quel punto divenne noto, sia pure con un errore dell'ordine del 10%, anche il diametro dell'orbita terrestre.

Grazie a Cassini, una simile triangolazione poté allora essere tentata per misurare la distanza delle stelle, usando il diametro dell'orbita terrestre (parallassi) come lato noto del triangolo. Nel 1838, Friedrich Wilhelm Bessel misurò in questo modo la distanza della stella *61 Cygni* e, poco dopo, anche le distanze di  $\alpha$ -*Centauri* e *Vega* furono misurate alla stessa maniera. Lo spostamento angolare di queste stelle, tra i due estremi dell'orbita terrestre, è di poco inferiore al secondo d'arco.

Per ovviare all'incertezza sul valore della parallasse, allora ancora significativa, le distanze delle stelle vennero definite in parallassi-al-secondo (*parsec*, pc): per esempio, lo spostamento di 0.1 secondi d'arco, tra un estremo e l'altro dell'orbita terrestre, significa una distanza di 10 pc.

L'incertezza attuale sulla lunghezza del parallasse è legata alle perturbazioni planetarie, non a problemi di misura. Ma il numero delle distanze calcolate prima di raggiungere l'attuale precisione fu molto elevato e la conversione a altre unità di distanza è inutile. Così il parsec resta l'unità di distanza cosmica.

Va segnalato anche il limite intrinseco che questa procedura ebbe fino agli anni Novanta. Esso è dovuto alla rifrazione nell'atmosfera terrestre. Indipendentemente dalle dimensioni dello specchio del telescopio, lo sfavillio di una stella impedisce di fissarne la posizione al di là di 1/30 di secondo d'arco. Non si possono quindi misurare da terra distanze di stelle più lontane di 30 pc.

Questo limite «storico» è oggi superato. Il satellite astrometrico *Hipparcos*, lanciato nel 1989, è giunto a fissare la posizione angolare di oltre 100 000 stelle con una precisione  $\sim 10^{-2}$  secondi d'arco, misurando quindi distanze fino a  $\sim 0.1$  kpc. La sonda *Gaia*, che l'ESA lancerà verso il punto lagrangiano L2 nel 2012, arriverà a una precisione 200 volte maggiore,  $\sim 10^{-5}$  *arcsec.*, rendendo possibile la misura della distanza di stelle oltre i 10 kpc, in pratica qualsiasi stella della Via Lattea.

Ciò permetterà di tarare in modo definitivo gli indicatori di distanza, come le stelle cefeidi, che si rilevano ormai anche in galassie oltre i 10 Mpc. La celebre *scala a pioli* delle distanze cosmiche perde così molti pioli, e si accederà facilmente alle distanze su cui domina il *flusso di Hubble*, per cui il resto dello schema segue in modo semplice.

Su questa base possiamo affermare che l'esistenza e l'isotropia del *flusso di Hubble*, già accertata nei dati correnti, potrà essere implementata da sempre più precise misure del parametro  $H_0$ .

## 7. Verso l'omogeneità

Sulla stessa base possiamo ora dibattere come si giunge all'acquisizione del secondo dato cosmologico di base, cioè che, su grande scala, il cosmo è isotropo e omogeneo.

Il cosmo è certamente disomogeneo sulle piccole scale, ma ciò che interessa è il livello di disomogeneità a partire dalla scala galattica: *Le galassie sono gli abitanti del*

*cosmo della cosmologia*, il cosmo dei Mpc,  $10^{15}$  volte sopra le distanze percorribili sulla Terra,  $10^{30}$  volte sopra le dimensioni di una molecola.

Il microcosmo molecolare ci aiuta però a capire come il livello di disomogeneità dipenda dalla scala. Basta pensare all'omogeneità in un liquido su scale  $> \sim 10^{-1}-10^{-2}$  cm, a confronto delle disomogeneità molecolari, su scale  $\sim 10^{-5}-10^{-6}$  cm. Nulla, nella melmosa fluidità di un olio, per esempio, segnala le enormi discontinuità su scale minori di appena 4-5 ordini di grandezza.

L'idea intuitiva che l'ampiezza delle disomogeneità dipende dalla scala va però posta su basi quantitative. Il dato statistico da costruire, a questo fine, è l'andamento della *varianza di massa*  $\sigma_R$  in funzione della scala  $R$ .

Tracciamo peraltro sin d'ora lo schema dell'indagine da svolgere.

- 1)  $\sigma_R$ , opportunamente definita, permetterà di stabilire la deviazione quadratica media dalla linearità, in funzione della scala  $R$ .
- 2) La misura di  $\sigma_R$  può essere ricondotta alla misura di  $\zeta(R)$ , la *funzione di correlazione* a due punti.
- 3) La «funzione a due punti»  $\zeta(r)$  può essere direttamente correlata allo *spettro delle fluttuazioni*  $P(k)$  e alla *funzione spettrale*  $\Delta^2(k)$ .
- 4) Il quadro fenomenologico che risulta dall'analisi di questi quattro elementi, [ $\sigma_R$ ,  $\zeta(r)$ ,  $P(k)$ ,  $\Delta^2(k)$ ], permette effettivamente di designare le scale su cui si raggiunge la linearità e le scale su cui un'ipotesi di isotropia e omogeneità è ragionevole.

È anche il caso di segnalare subito le ragioni per cui l'ordinaria problematica statistica si complica in un contesto cosmologico.

La distribuzione osservata della materia è una realizzazione di un processo stocastico, la cui conoscenza deve essere avvicinata dall'analisi statistica dei dati. A questo fine occorre compiere un *gran numero* di misure tra loro indipendenti, mediando poi sui risultati. È lo stesso procedimento con cui si misura, per esempio, la *vita media* di una particella, osservando molte tracce che essa ha lasciato. All' $i$ -esima traccia corrisponderà una vita  $\tau_i$ , la vita media si ha mediando gli  $N$  valori misurati  $\tau_i$ :

$$(II.9) \quad \langle \tau \rangle = \frac{1}{N} \sum_{i=1}^N \tau_i.$$

Se le  $N$  misure non vengono considerate sufficienti, in questo caso si può continuare a realizzarne, finché si ottiene un numero di misure adeguato.

Allo stesso modo, se si vuole conoscere la *varianza di massa*  $\sigma_R$ , bisognerà osservare  $N$  volumi di dimensioni  $R$ . La definizione esatta di  $\sigma_R$  e la tecnica da seguire, saranno oggetto della seconda parte di questo paragrafo; supporremo qui che non ci siano difficoltà tecniche legate a ciascuna misura.

Proponiamoci allora, per esempio, di cercare la varianza di massa sulla scala di  $R = 10 h^{-1}$  Mpc. Dovremo avere dati riguardanti un volume  $\mathcal{V} \gg R^3$ , in cui vi sia un

numero adeguato di *realizzazioni* su cui effettuare la misura. Per esempio, se  $\mathcal{V} \sim (100 h^{-1} \text{Mpc})^3$ , avremo a disposizione  $\sim 10^3$  realizzazioni.

È importante che si tratti di realizzazioni *dello stesso processo*. Se, per esempio, abbiamo evidenza di un'evoluzione della varianza col *redshift*, dovremo operare su volumi aventi, quanto meno, lo stesso *redshift* medio.

Per questa ragione potrebbe essere necessario considerare volumi tutti alla stessa distanza dall'osservatore. Uno *shell* di profondità  $R$ , a distanza  $d$  dall'osservatore ha volume

$$(II.10) \quad \mathcal{V} \simeq 4\pi d^2 R$$

e in esso si trovano

$$(II.11) \quad N \simeq \mathcal{V}/R^3 = 4\pi(d/R)^2$$

realizzazioni indipendenti. Se vogliamo allora  $10^3$  realizzazioni indipendenti, dovrà essere

$$(II.12) \quad d > R(1000/4\pi)^{1/2} \sim 9R.$$

Una distanza  $\sim 90 h^{-1} \text{Mpc}$  non è proibitiva per l'osservazione.

L'esempio deve tuttavia mettere a fuoco l'esistenza di un limite oggettivo alla riproducibilità del dato osservativo. Al crescere della scala le difficoltà aumentano. Se il nostro obiettivo fosse di confrontare 1000 volumi di lato  $\sim 50 h^{-1} \text{Mpc}$ , per esempio, dovremmo andare a distanze  $\sim 450 h^{-1} \text{Mpc} \sim 1.6$  miliardi di anni luce. Si tratta ancora di distanze praticabili, ma che osserviamo nelle condizioni in cui si trovavano 1.6 miliardi di anni orsono, circa 1/8 della vita del cosmo. È legittimo ritenere che le condizioni di oggi siano rappresentate dallo stato del cosmo allora?

Come mostrato in figura II.5,  $50 h^{-1} \sim 80 \text{Mpc}$  sono l'inizio del regime lineare. Di fatto il raggiungimento di tale regime si può considerare certificabile da un numero adeguato di misure indipendenti.

Viceversa, il raggiungimento di una scala in cui le fluttuazioni lineari divengono effettivamente trascurabili, di modo da poterci considerare in regime di isotropia e omogeneità, presume veramente un'esplorazione nel passato. L'acquisizione dell'esistenza di una simile scala non potrà quindi essere un dato direttamente conseguito, si deve passare attraverso una conoscenza delle dinamiche dell'evoluzione, da assodare con statistiche adeguate, in modo da rifletterne i risultati su una simile scala.

Il problema qui considerato raggiunge il suo «parossismo» quando ci si propongono misure su scale confrontabili con l'orizzonte. Omogeneità e isotropia sono raggiunte su scale decisamente inferiori, ma le anisotropie di un fondo cosmico, per esempio del CMB, su scale confrontabili con l'intero angolo solido di  $4\pi$ , sono misurabili su un'unica realizzazione o su un numero modestissimo di realizzazioni. In questo caso si dice che le misure sono soggette a *varianza cosmica* e le incertezze intrinseche nel confronto dei dati, sulle grandi scale angolari, coi modelli cosmologici dovranno essere attentamente dibattute.

## 8. La varianza di massa

Proviamo allora a trasporre su scale cosmiche i passi che faremmo su scale accessibili in laboratorio. Partiamo da un «grande» volume  $\mathcal{V}$ , che contiene un elevato numero  $\mathcal{N}$  di galassie e tale che la sua scala lineare  $\mathcal{R} = \mathcal{V}^{1/3}$  sia molto maggiore della scala  $R$  su cui vogliamo fare misure. In generale le galassie hanno masse diverse, ma, con ottimi risultati operativi, conviene attribuire a ciascuna galassia una massa pari alla *massa galattica media*  $m_g$ .

Partiamo allora dalla definizione di alcune quantità *medie*:

La densità in numero di galassie in $\mathcal{V}$ :	$n = \mathcal{N}/\mathcal{V}$
La densità in massa in $\mathcal{V}$ :	$\rho = m_g n$
Il volume di una sfera di raggio $R$ :	$V_R = (4\pi/3)R^3$
Il numero medio di galassie nella sfera di raggio $R$ :	$N_R = V_R n$
La massa media in una sfera di raggio $R$ :	$M_R = V_R \rho = N_R m_g$

Usando il *filtro a top-hat* definiamo poi alcune quantità *dependenti dal sito*. Tale filtro è definito coerentemente in appendice A1, in modo tale da poterlo anche sostituire con filtri diversi, di cui non faremo però uso nel testo del capitolo.

A questo proposito conveniamo di indicare  $V_R(\mathbf{r})$  il *volume di una sfera di raggio  $R$  centrata in  $\mathbf{r}$* . Il valore del volume non cambia con  $\mathbf{r}$ , ma ciascun  $V_R(\mathbf{r})$  avrà diversi contenuti. Definiamo allora:

Il numero di galassie in $V_R(\mathbf{r})$ :	$N_R(\mathbf{r})$
La massa in $V_R(\mathbf{r})$ :	$M_R(\mathbf{r}) = m_g N_R(\mathbf{r})$
La densità in numero in $V_R(\mathbf{r})$ :	$n_R(\mathbf{r}) = N_R(\mathbf{r})/V_R$
La densità in massa in $V_R(\mathbf{r})$ :	$\rho_R(\mathbf{r}) = m_g n_R(\mathbf{r})$

In genere, i rapporti  $n_R(\mathbf{r}) = N_R(\mathbf{r})/V_R$  hanno valori diversi da  $n$  e dipendenti dalla collocazione  $\mathbf{r}$  del volume. Mediando, come è espresso dalla notazione

$$(II.13) \quad \lim_{v \rightarrow \infty} \frac{1}{v} \sum_{i=1}^v n_R(\mathbf{r}_i) = n,$$

si deve ottenere  $n$ . Qui troviamo per la prima volta l'operazione di media; operativamente essa si ridurrà all'operazione

$$\langle n_R(\mathbf{r}) \rangle \equiv \frac{1}{v} \sum_{i=1}^v n_R(\mathbf{r}_i)$$

con  $v$  finito; ma il campione di volumi su cui si media dovrebbe essere abbastanza grande da renderlo «rappresentativo» dell'intero Universo. Il fatto che sia  $\langle n_R(\mathbf{r}) \rangle = n$  è riesprimibile tramite le relazioni

$$(II.14) \quad \langle N_R(\mathbf{r}) \rangle = N_R \quad \text{ovvero} \quad \langle N_R(\mathbf{r}) - N_R \rangle = 0,$$

che si ottengono moltiplicando per  $V_R$ . Nel modo di effettuare la media  $\langle \dots \rangle$  e nella stessa possibilità concettuale di effettuarla sta gran parte delle difficoltà. Le discuteremo nell'ultimo sottoparagrafo.

La seconda relazione (II.14) ci dice che le deviazioni in più e in meno, rispetto alla media, si compensano per definizione. Se invece quadriamo prima di mediare, otteniamo la quantità

$$(II.15) \quad \sigma_R^2 = \frac{\langle [N_R(\mathbf{r}) - N_R]^2 \rangle}{N_R^2},$$

che riceve contributi positivi sia dalle deviazioni in più che da quelle in meno, ed è la *varianza di massa* su scala  $R$ .

In eq. (II.15) possiamo dividere o moltiplicare numeratore e denominatore, a secondo membro, per diverse quantità, ottenendo altre utili espressioni per  $\sigma_R^2$ .

$$(II.16) \quad \begin{array}{ll} 1) \text{ Dividendo per } V_R^2 : & \sigma_R^2 = \langle [n_R(\mathbf{r}) - n]^2 \rangle / n_g^2 \\ 2) \text{ Moltiplicando per } m_g^2 : & \sigma_R^2 = \langle [M_R(\mathbf{r}) - M_R]^2 \rangle / M_R^2 \\ 3) \text{ Dividendo quest'ultima per } V_R^2 : & \sigma_R^2 = \frac{\langle [\rho_R(\mathbf{r}) - \rho]^2 \rangle}{\rho^2} = \left\langle \left( \frac{\delta \rho_R(\mathbf{r})}{\rho} \right)^2 \right\rangle \end{array}$$

Quest'ultima espressione mette in evidenza che la varianza di massa coincide con la media quadratica della *fluttuazione di densità*

$$(II.17) \quad \varepsilon_R(\mathbf{r}) = \delta \rho_R(\mathbf{r}) / \rho$$

smussata sulla scala  $R$ . L'eq. (II.16) assume quindi la forma compatta

$$(II.18) \quad \sigma_R^2 = \langle \varepsilon_R^2(\mathbf{r}) \rangle$$

che mostra come la varianza di massa su scala  $R$  sia la media quadratica delle fluttuazioni di densità smussate sulla medesima scala.

Sin qui non si è tenuto conto dell'eventuale difficoltà a reperire un volume  $\mathcal{V}$  abbastanza grande per effettuare le medie richieste.

## 9. Campioni e cataloghi

Questo punto va dibattuto a due livelli. Concettualmente è importante sapere se una determinata misura è statisticamente fattibile. Operativamente, c'è un limite più immediato con cui occorre confrontarsi, la quantità di materiale osservativo effettivamente disponibile.

Per quanto concerne le galassie, esso è raccolto in *cataloghi*. Un catalogo di galassie contiene tutte le galassie che (i) *soddisfano determinati requisiti* e di cui (ii) *sono stati misurati alcuni parametri*. Tra i requisiti ci sarà l'appartenenza a una

determinata zona di cielo e, nella quasi totalità dei casi, anche il superamento di un determinato livello di luminosità apparente. Tra i parametri misurati ci sono, di norma, le coordinate angolari sulla volta celeste.

### ***Le magnitudini***

Al momento di trattare i limiti in luminosità o di valutare la luminosità dei singoli oggetti, occorre conoscere le unità in uso.

I moderni apparati che scrutano il cielo non perdono neanche uno dei fotoni utilizzabili. Non sempre è stato così. Nel Seicento, Tycho Brahe compilò una mappa dei moti planetari senza disporre nemmeno di un cannocchiale e i suoi dati furono tuttavia così precisi da permettere a Keplero di ricavarne le sue leggi. Ma, ancora nella prima parte del Ventesimo secolo, l'abilità dell'osservatore giocava un ruolo dominante nella stima delle luminosità apparenti degli oggetti cosmici. Per esempio, molte galassie sono catalogate secondo la *magnitudine di Zwicky* e solo strumenti raffinati sono in grado di battere le misure di Fritz Zwicky, che peraltro disponeva di uno strumento oggettivamente straordinario, nella sua come in ogni altra epoca, il suo occhio infallibile.

Il ruolo giocato dall'occhio, nella definizione della scala su cui misurare le luci celesti, influisce sul modo in cui ancora oggi ne esprimiamo l'intensità. Orecchio e occhio umani hanno una *sensibilità logaritmica*. Se la differenza d'intensità tra due sorgenti **A** e **B** ci sembra pari alla differenza d'intensità tra **B** e **C**, le misure fisiche ci dicono che tra le rispettive intensità vale la relazione

$$(II.19) \quad S_A/S_B \simeq S_B/S_C$$

(e non  $S_A - S_B \simeq S_B - S_C$ ). Non deve quindi sorprendere che la «tradizione astronomica», nata prima dell'introduzione dei moderni apparati fisici, abbia ricondotto le intensità  $S$  a quantità, dette *magnitudini*, che si ricollegano alle intensità energetiche tramite la relazione

$$(II.20) \quad m = -\frac{5}{2} \log_{10} S + K_1,$$

dove  $K_1$  è una costante opportuna (vedi sotto). Prendendo il logaritmo decimale della relazione (II.19) è infatti facile vedere che risulta

$$(II.21) \quad m_A - m_B \simeq m_B - m_C.$$

Il mantenimento delle magnitudini, come moderna unità, è dovuto alla costante necessità di non disperdere il patrimonio di informazione gradualmente accumulato e accumulantesi in queste unità.

La relazione (II.20) vale tra luminosità apparenti o *flussi* e quindi le  $m$  sono *magnitudini apparenti*. In stretta analogia con le magnitudini apparenti si usano anche le

*magnitudini assolute*

$$(II.22) \quad M = -\frac{5}{2} \log_{10} L + K_2,$$

legate alla luminosità intrinseca  $L$  delle sorgenti.

Il valore delle costanti  $K_1$  e  $K_2$  è fissato attribuendo anzitutto al Sole la magnitudine intrinseca  $M_{\odot} = 5.5$  e ciò fissa  $K_2$ . Vale poi la relazione

$$(II.23) \quad S = L/4\pi r^2$$

tra flusso e luminosità intrinseca; prendendone il logaritmo avremo

$$(II.24) \quad \log_{10} S = \log_{10} L - 2 \log_{10} r + \text{costante},$$

in cui la costante dipende dalla scelta di un'unità per  $r$ . Tale unità di misura e la costante  $K_1$  in eq. (II.20), si fissano traducendo la relazione (II.24) nella forma

$$(II.25) \quad m = M + 5 \log_{10}(r/\text{Mpc}) + 25.$$

Per esempio, una stella come il Sole, sita a 1 pc ( $10^{-6}$ Mpc!) dall'osservatore terrestre, apparirebbe di magnitudine

$$m = 5.5 + 5 \log_{10} 10^{-6} + 25 = 0.5,$$

assai prossima alla magnitudine apparente di *Proxima Centauri*. Viceversa, una stella simile al Sole in una galassia del Gruppo Locale, sita a 1 Mpc dall'osservatore, ha magnitudine apparente

$$m = 5.5 + 25 = 30.5,$$

ma stelle della trentesima magnitudine non sono facilmente osservabili!

Questi esempi mettono in rilievo un fatto insito nel segno negativo delle definizioni (II.20) e (II.22), che oggetti di apparenza più fioca sono di magnitudine più elevata:  $m$  ( $M$ ) cresce al calare di  $S$  ( $L$ ).

### **Cataloghi angolari**

Usando le magnitudini a definire le luminosità consideriamo ora i vari tipi di catalogo che sono (e sono stati) disponibili per l'indagine cosmologica.

Storicamente, il primo catalogo di galassie con proprietà definite, tale quindi da potere essere utilizzato a fini statistici, compilato da Zwicky, conteneva circa 3750 galassie. Si trattava di tutte le galassie collocate entro un determinato angolo di cielo ( $b^I > +40^\circ$ ,  $\delta(1950) > 0$ ), con magnitudine  $m < 15$ . Di ciascuna galassia era fornita la posizione in coordinate galattiche e la magnitudine apparente  $m$ .

Un secondo catalogo fu successivamente compilato da Shane e Wirtanen su dati del telescopio Lick e catalogava un milione di galassie fino alla 18<sup>a</sup> magnitudine, senza fornire le luminosità individuali e dandone le coordinate angolari con un certo tipo di approssimazione, che esula dai nostri scopi discutere.

Questi *cataloghi angolari* non offrivano nessuna informazione sulla effettiva distanza delle sorgenti galattiche e sulla loro reale luminosità intrinseca.

### **Cataloghi con redshift**

Tra i dati che rendono più significativi i cataloghi, quello più importante è il *redshift*, che fornisce una stima della distanza. Infatti, se una galassia si allontana a velocità  $v_g (\ll c)$ , il *redshift* della luce da essa emessa sarà  $z_g = v_g/c$ . Per stimare la sua distanza  $d_g$ , da  $v_g$  va sottratto il contributo della sua *velocità peculiare*  $v_{p,g}$ , cioè del suo moto rispetto al *flusso di Hubble*. Come abbiamo discusso in paragrafo 1, la velocità osservata è infatti

$$(II.26) \quad v_g = H d_g + \mathbf{v}_{p,g} \cdot \mathbf{n}_g,$$

e, come abbiamo visto, velocità peculiari  $\sim 10^{-3}c$  sono comuni. Qui il versore  $\mathbf{n}_g$  definisce la posizione della galassia sulla volta celeste.

$z_g$  diviene quindi un ragionevole misuratore di distanza solo quando  $d_g > \sim 10 h^{-1} \text{Mpc}$  e la distanza fornita da  $z_g$  sarà comunque imprecisa di  $\sim 10 h^{-1} \text{Mpc}$ . L'errore è significativo per galassie «vicine», meno importante per galassie «lontane», ma l'uso di indicatori di distanza più sofisticati è assai più laborioso.

Peraltro, anche per ricavare il *redshift* di una galassia, occorrono dettagliate analisi spettrali. Solo all'inizio degli anni Ottanta divenne disponibile il primo *catalogo con redshift*, che forniva il *redshift* di galassie di Zwicky con magnitudine  $m < 14.5$  (*CfA1 redshift sample*). Il catalogo venne poi esteso a comprendere tutte le galassie fino a  $m = 15.5$  (*CfA2 redshift sample*).

A causa del basso limite in luminosità, essi comprendevano galassie relativamente vicine. Se la distanza è stimata sulla base del *redshift*, ne nasce il cosiddetto effetto delle *dita di Dio*: ogni gruppetto di galassie nel catalogo appare dilatato nella direzione della linea di vista, a causa della componente lungo  $\mathbf{n}$  delle velocità peculiari  $\mathbf{v}_p$ , che invece non hanno alcun effetto in direzioni ortogonali a  $\mathbf{n}$ .

### **Indicatori di distanza indipendenti dal redshift e cataloghi profondi**

Come suddetto, l'uso di indicatori di distanza indipendenti dal *redshift* è laborioso e solo piccoli campioni sono attualmente disponibili, in cui ciascuna galassia sia catalogata con distanza e *redshift* stimati in modo indipendente.

Piuttosto che in questa direzione, grossi sforzi sono stati fatti per avere campioni profondi, estesi a magnitudini elevate, che annoverino molte galassie fornendone coordinate angolari, magnitudine  $m$  e *redshift*  $z$ . La compilazione di questi cataloghi è stata automatizzata, in modo che non occorre più che l'astronomo valuti personalmente il *redshift* di ogni singola galassia. Esempi di questi cataloghi sono il campione 2dF (vedi Fig. II.1) e la *SLOAN digital sky survey*, tuttora in fase di compilazione.

### Dai cataloghi ai campioni

Tutti i cataloghi sinora prodotti sono limitati in magnitudine apparente  $m$ . Si tratta di un valido criterio osservativo cui non corrisponde alcuna proprietà intrinseca delle galassie. Usando  $m$  e la distanza  $d$ , si può dedurre la magnitudine intrinseca  $M = m - 5 \log(d/\text{Mpc}) - 25$  e ricavare insiemi di galassie con luminosità superiore a un determinato limite. (Di fatto, come precedentemente rimarcato,  $d$  è spesso noto con qualche approssimazione.)

In questo modo si realizzano dei *campioni*, che contengono solo una parte delle galassie del catalogo, tali da soddisfare requisiti più stringenti e adatti a determinate indagini statistiche. Ciascun catalogo costituisce di per sé un campione, mentre non è vero l'opposto.

La figura II.6 è tracciata nell'ipotesi ideale di velocità peculiari trascurabili. Essa mostra come un catalogo selezionato in funzione di  $m$  (in figura  $m = 17$ ) contenga galassie al di sotto di una magnitudine intrinseca  $M$  che varia con la distanza. Il limite

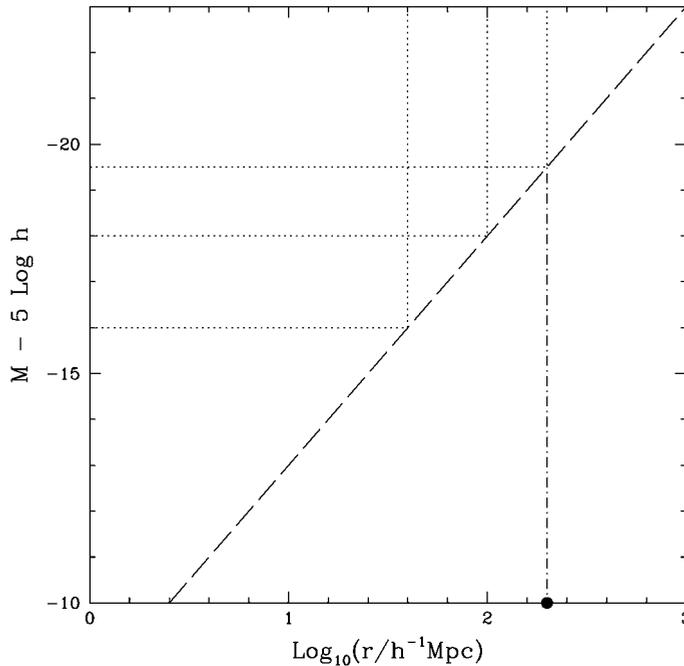


FIGURA II.6

Magnitudine intrinseca delle galassie contenute in un ipotetico catalogo limitato alla  $m = 17$ . La linea obliqua col tratteggio lungo indica la minima magnitudine intrinseca delle galassie comprese nel catalogo a ciascuna distanza. Le linee punteggiate si riferiscono al contenuto di ipotetici campioni ricavati dal catalogo, che annoverano solo galassie fino a  $M - 5 \log_{10} h = -16, -18, -19.5$ , rispettivamente. Dal grafico si vede che ciascuno di questi campioni si trova entro una diversa profondità. Il punto a  $200 h^{-1} \text{Mpc}$ , sull'asse delle ascisse, indica la *profondità* del catalogo (vedi testo).

è indicato dalla linea tratteggiata lunga. Dal catalogo si possono desumere vari campioni limitati in magnitudine assoluta ( $M - 5 \log h = -16, -18, -19.5$ ). Ciascuno di tali campioni contiene galassie entro una distanza definita (rispettivamente  $rh/\text{Mpc} = 10^{1.6}, 100, 10^{2.3}$ ). Si tratta dunque di campioni limitati sia in luminosità assoluta che in volume.

La magnitudine assoluta  $M^* = -19.5 + 5 \log h$  ha un significato specifico, che discuteremo nel prossimo sottoparagrafo. La distanza corrispondente  $D^* = 10^{2.3} h^{-1} \text{ Mpc} \simeq 200 h^{-1} \text{ Mpc}$  sarà definita *profondità* del catalogo.

### **Funzione di luminosità delle galassie**

Tramite i campioni si cerca di ricavare, in primo luogo, come sia distribuita la luminosità delle galassie. Di tale distribuzione si suole dare l'espressione

$$(II.27) \quad \psi(L/L^*) = \psi^*(L/L^*)^{-\alpha} e^{-L/L^*},$$

detta *funzione di luminosità* di Schechter, che è una sintesi di effetti complessi e concomitanti, discussi nei testi di fisica delle galassie. Un problema è costituito dal fatto che la sintesi di questi effetti vari in funzione dell'epoca cosmica, dal momento che l'espressione (II.27) è spesso usata in campioni estesi a una discreta profondità.

Con buona approssimazione, riferendosi a una popolazione di galassie a  $z \sim 0$ , nell'espressione (II.27) si possono usare i valori

$$(II.28) \quad \psi^* = 10^{-2} (h^{-1} \text{ Mpc})^{-3}, \quad \alpha \simeq 1.1 - 1.2, \quad L^* \simeq 10^{10} h^{-2} L_{\odot},$$

delle costanti ivi contenute. A  $L^*$  corrisponde la magnitudine

$$(II.29) \quad M^* = -19.5 + 5 \log h,$$

citata nel precedente sottoparagrafo.

In figura II.7 è mostrata la funzione di Schechter con  $\alpha = 1.2$ . L'area al di sotto della curva rappresenta il numero di galassie. Per esempio, l'area del trapezoide è il numero di galassie con luminosità compresa tra  $10^{9.8}$  e  $10^{10.2} M_{\odot}$ .

Formalmente tale area è l'integrale della funzione di Schechter e, in particolare, la funzione di luminosità *integrata*

$$(II.30) \quad \psi(> L) = \int_L^{\infty} \frac{dL}{L^*} \psi\left(\frac{L}{L^*}\right) = \psi^* \int_{L/L^*}^{\infty} dx x^{-\alpha} e^{-x}$$

ci dice quante galassie ci sono con luminosità maggiore di  $L$ .

Le caratteristiche fenomenologiche che l'espressione (II.27) vuole mettere in rilievo sono le seguenti:

- 1) A bassi valori di  $L$  il numero di galassie osservate cala al crescere di  $L$  con una legge di potenza di esponente  $\alpha$ .

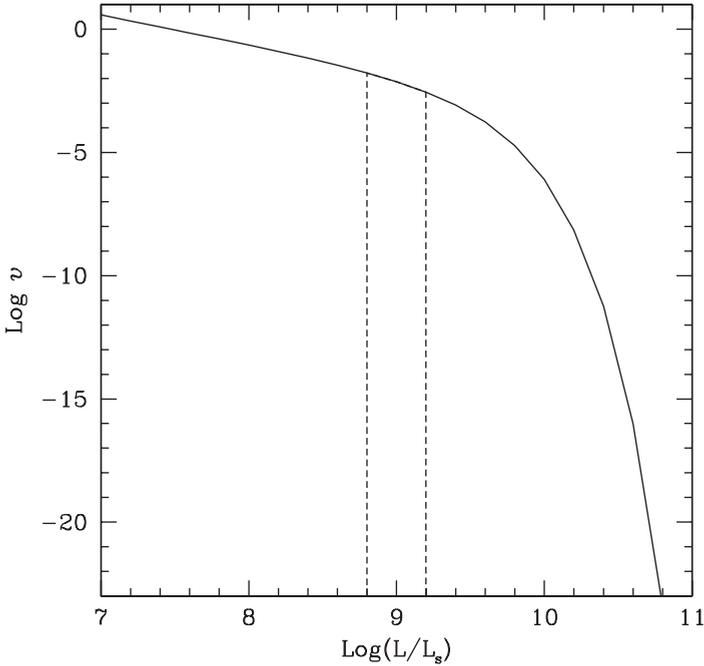


FIGURA II.7

Funzione di luminosità delle galassie. L'area del trapezoide dai contorni tratteggiati è il numero di galassie con luminosità compresa tra i valori indicati dalle ascisse. (si veda l'appendice TAVOLE A COLORI).

- 2) Giunti a una luminosità  $\sim L^*$  (magnitudine  $M^*$ ), la legge esponenziale subisce un brusco taglio esponenziale; il numero delle galassie con  $L > L^*$  è quindi assai modesto.
- 3) Il numero di galassie per unità di volume con luminosità  $> L$ ,  $n_g(> L)$ , è dato dal prodotto di  $\psi^*$  per una funzione  $F(L/L^*, \alpha)$  ( $= \int_{L/L^*}^{\infty} dx x^{-\alpha} e^{-x}$ ) che esprime una blanda dipendenza dal rapporto  $L/L^*$  e da  $\alpha$ , e ha valori prossimi all'unità. Il coefficiente  $\psi^*$ , espresso in unità  $(h^{-1} \text{ Mpc})^3$ , fornisce quindi direttamente l'ordine di grandezza della densità in numero di galassie. La distanza media tra galassie risulta allora  $\sim 100^{1/3} h^{-1} \text{ Mpc} \sim 4\text{-}5 h^{-1} \text{ Mpc}$ .

### ***Profondità di un campione***

I cataloghi di galassie sono compilati includendo tutti gli oggetti con  $m < m_l$  (magnitudine limite). La *profondità* ( $D^*$ ) di un campione è la distanza a cui deve trovarsi galassia di magnitudine intrinseca  $M^*$  per essere osservata con magnitudine apparente  $m_l$  (vedi ancora Fig. II.6). Dalla relazione (II.25) ricaviamo allora

$$(II.31) \quad m_l = M^* - 5 \log h + 5 \log (D^*/h^{-1} \text{ Mpc}) + 25$$

e quindi

$$(II.32) \quad \frac{D^*}{h^{-1} \text{ Mpc}} = 10^{\frac{m_l - M^* + 5 \log h}{5} - 5} = 10^{\frac{m_l}{5} - 1.1}$$

Il campione considerato in figura, con  $m_l = 17$ , ha profondità  $D^* = 10^{2.3} h^{-1} \simeq 200 h^{-1} \text{ Mpc}$ , come indicato nella figura e nella sua didascalia. Il catalogo di Zwicky che comprende galassie fino alla 15<sup>a</sup> magnitudine ha profondità  $D^* = 10^{1.9} h^{-1} \text{ Mpc} \simeq 80 h^{-1} \text{ Mpc}$ .

## 10. La relazione log N-log S

Ci siamo sin qui riferiti a galassie osservate in banda ottica. Abbiamo poi utilizzato la relazione tra luminosità intrinseca e flusso,

$$(II.33) \quad S = L/4\pi r^2 \quad \text{ovvero} \quad \log_{10} S = \log_{10} L - 2 \log_{10} r + \text{cost.},$$

corretta in assenza di ogni sorta di attenuazione della luce e *se lo spazio è euclideo*, nel suo assieme e localmente.

In questo paragrafo mostreremo che esiste un modo semplice per esplorare la geometria globale del cosmo e l'evoluzione dei suoi «abitanti», senza misurare *redshift*. Tale tecnica, utilizzata per le *radiogalassie* scoperte dai primi radiotelescopi operanti negli anni Sessanta, venne usata a provare che il cosmo non può essere simultaneamente (i) euclideo e (ii) non-evolutivo.

Si seleziona una particolare famiglia di oggetti, determinandone la luminosità apparente  $S$ ; ciò permette di calcolare il numero  $N(> S)$  di oggetti con luminosità  $> S$  (in un determinato angolo solido). Verificheremo ora che, se vale l'eq. (II.33), deve essere

$$(II.34) \quad \log N(> S) = -1.5 \log S (+ \text{cost.}).$$

Si tratta di una relazione semplice e apparentemente facile da verificare. Se è violata, uno dei punti (i)-(ii) è falso (o entrambi).

Se tutti gli oggetti cosmici avessero uguale luminosità  $L_1$ , gli oggetti con luminosità apparente  $> S$  sarebbero tutti e soli quelli situati entro una distanza  $R_{lim}$  data dall'equazione

$$(II.35) \quad S = \frac{L_1}{4\pi R_{lim}^2} \quad \text{e, quindi:} \quad R_{lim} = \left( \frac{L_1}{4\pi S} \right)^{\frac{1}{2}}.$$

Usando l'espressione di  $R_{lim}$  si vede allora che se ne osserverebbero in tutto

$$(II.36) \quad N_1(> S) = n_{L_1} \frac{4\pi}{3} R_{lim}^3 = n_{L_1} L_1^{\frac{3}{2}} \frac{1}{3(4\pi)^{1/2}} S^{-\frac{3}{2}},$$

se  $n_{L_1}$  è la loro densità in numero. Nel cosmo ci sono galassie con luminosità distribuita

su un continuo di valori. Ma, per semplicità, immaginiamo che ci sia solo un insieme discreto di luminosità possibili:  $L_1, L_2, \dots, L_n, \dots$  ecc. Il numero totale di galassie con luminosità  $>S$ , sarà allora dato dalla somma

$$(II.37) \quad N_{tot}(>S) = N_1(>S) + N_2(>S) + \dots + N_n(>S) + \dots = \\ = \left[ n_{L_1} L_1^{\frac{3}{2}} + n_{L_2} L_2^{\frac{3}{2}} + \dots + n_{L_n} L_n^{\frac{3}{2}} + \dots \right] \frac{1}{3(4\pi)^{1/2}} S^{-\frac{3}{2}}.$$

Qualunque siano quindi i valori delle densità  $n_{L_i}$ , purché indipendenti dalla distanza, avremo allora

$$(II.38) \quad N_{tot}(>S) = K S^{-\frac{3}{2}}$$

in cui  $K$  è una costante. Il logaritmo dell'eq. (II.38) ridà l'eq. (II.34).

Questa relazione può essere riespressa in termini di magnitudini. La densità angolare di oggetti cosmici al di sotto di una determinata magnitudine limite  $m$  ha l'espressione

$$(II.39) \quad n_{\Omega}(<m) = F \times 10^{3m/5};$$

la dipendenza da  $m$  segue subito da quella da  $S$  e dalla relazione tra  $m$  e  $S$ . L'espressione del coefficiente  $F$  è data in appendice A6, assieme alla prova dell'espressione

$$(II.40) \quad n_{\Omega}(m)\delta m = \delta m F \times 10^{3m/5} \times 0.6 \ln 10$$

che dà la densità angolare di oggetti con magnitudine in un intervallino  $\delta m$  attorno a  $m$ .

Confrontate coi dati delle galassie ordinarie, queste equazioni testimoniano una evoluzione complessivamente modesta e la compatibilità della geometria dello spazio con una metrica euclidea, fino a  $m \sim 16-17$ .

Invece, una volta scoperte le radiogalassie, oggetti situati a distanze assai maggiori, venne notata una distribuzione del loro *flusso* con pendenza  $\neq 1.5$ . Su questa base venne proposta la conclusione che *il cosmo è evolutivo* e/o ha una geometria non euclidea.

Una dettagliata analisi degli errori, su queste quantità integrali, rende oggi questionabili le prime evidenze. Attualmente, tuttavia, disponendo di dati assai più dettagliati, tale conclusione è confermata, tramite l'uso di relazioni «differenziali» del tipo (II.40) e una accurata trattazione degli errori osservativi.

## 11. Funzione a due punti e varianza di massa

La ricerca della scala di omogeneità, concettualmente, richiede la misurazione della varianza di massa in funzione della scala. Relativamente più facile da misurare è la *funzione di correlazione a 2 punti*, la cui importanza abbiamo segnalato sin dall'inizio di questa discussione, da cui si può dedurre analiticamente l'andamento della varianza.

Le funzioni di correlazione, in generale, sono uno strumento statistico molto in voga, in vari rami della fisica, quando si vogliono ricavare le caratteristiche di distribuzioni stocastiche.

### **La funzione a due punti: definizione**

Sul finire degli anni Sessanta gli astronomi erano giunti a compilare i primi *campioni di galassie*, su cui operare misure delle funzioni di correlazione. La prima misura su un catalogo fu operata da due giapponesi, Totsuji & Kihara. Il loro lavoro venne tuttavia ignorato dalla comunità scientifica, finché la questione venne riaffrontata indipendentemente da Jim Peebles e collaboratori.

Non seguiremo qui il cammino storico, ma supporremo di disporre di un campione ideale di  $\mathcal{N}$  galassie con luminosità  $L > L_*$ , situate in punti di coordinate note, entro un volume  $\mathcal{V}$  assai ampio e ben delimitato. La stima della densità in numero di galassie, data dal campione, è  $n = \mathcal{N}/\mathcal{V}$ .

La funzione (di correlazione) a 2 punti ci deve dire come viene influenzata la probabilità di trovare una galassia in un determinato sito di  $\mathcal{V}$ , dalla presenza di un'altra galassia in un sito vicino.

Per operare con rigore quantitativo, fissiamo *a caso*, all'interno di  $\mathcal{V}$ , un volume  $\delta V$  «piccolo», ma comunque in grado di contenere una galassia. La probabilità che esso contenga una galassia, ovvero il *valore di aspettazione* del numero di galassie in  $\delta V$ , sarà

$$(II.41) \quad \langle \delta N \rangle = n \delta V .$$

Consideriamo poi, in  $\mathcal{V}$ , due volumetti  $\delta V_1$  e  $\delta V_2$  siti a distanza  $r_{12}$ , ma altrimenti scelti completamente a caso. La probabilità congiunta di trovarvi due galassie, ovvero il valore di aspettazione del prodotto dei numeri di galassie nei due volumetti, sarà

$$(II.42) \quad \langle \delta N_1 \delta N_2 \rangle = n^2 \delta V_1 \delta V_2 [1 + \xi(r_{12})] .$$

Notiamo che, se fosse  $\xi(r_{12}) = 0$ , avremmo  $\langle \delta N_1 \delta N_2 \rangle = \langle \delta N_1 \rangle \langle \delta N_2 \rangle$ : il valore di aspettazione del prodotto dei numeri coinciderebbe col prodotto dei valori di aspettazione dei numeri. Ciò significherebbe che la presenza di una galassia in un sito non influenza la presenza di un'altra galassia in un sito vicino. Viceversa, ove questa influenza sussista, l'assunzione di isotropia dello spazio implica che essa possa dipendere solo dalla mutua distanza dei siti ( $r_{12}$ ).

Ciò che  $\xi(r)$  misura è il *clustering* delle galassie. Se  $\xi(r)$  è positivo (negativo), la presenza di una galassia in un punto accresce (diminuisce) la probabilità di trovarne altre nei punti vicini.

Con opportune tecniche la  $\xi(r)$  è stata stimata usando i cataloghi di galassie. Il risultato delle misure, con discreta approssimazione, è una legge di potenza:

$$(II.43) \quad \xi(r) = \left( \frac{r_0}{r} \right)^\gamma .$$

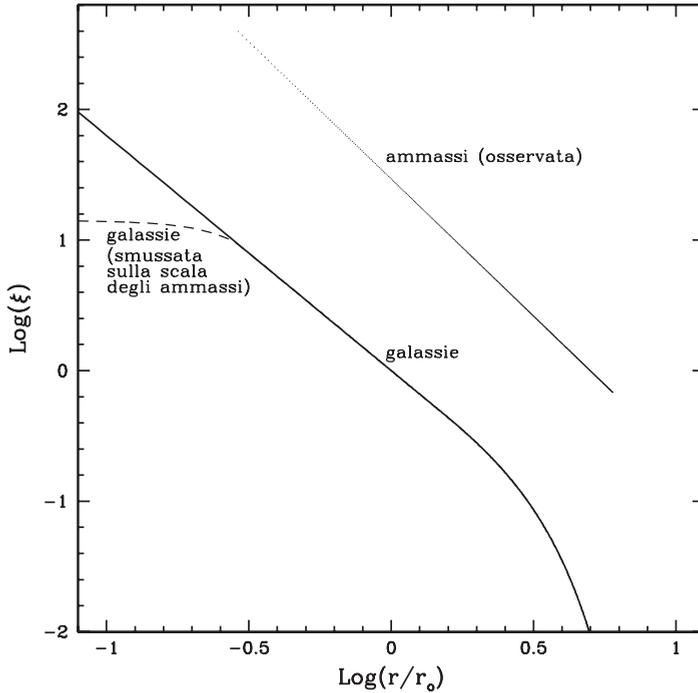


FIGURA II.8

Funzioni di correlazione a 2 punti. La linea compatta contrassegnata «galassie» corrisponde all'espressione (II.43). La linea tratteggiata dà la funzione a 2 punti delle galassie smussata sulla scala degli ammassi. La linea più alta, contrassegnata «ammassi», dà la funzione a 2 punti osservata per gli ammassi ricchi (vedi testo, più avanti).

$r_0$  è detta *lunghezza di correlazione*. Il suo valore si aggira tra  $5.5$  e  $6 h^{-1} \text{Mpc}$ . L'esponente, invece, ha valore  $\gamma = 1.77 \pm 0.03$ . Questa espressione è verificata per  $r$  compreso tra  $\sim 0.05 h^{-1} \text{Mpc}$  e  $\sim (1.5-2) r_0$  e l'andamento di

$$\log \xi = C - \gamma \log(r) \quad \text{con} \quad C = \gamma \log(r_0).$$

è mostrato in figura II.8, assieme a dati concernenti gli ammassi di galassie, discussi più avanti. La scoperta di questa espressione, negli anni Settanta, fece scalpore. Gli astronomi pensavano che, sopra della scala galattica, ci fossero «scale tipiche» per il *clustering*. In particolare, agli *ammassi di galassie* veniva attribuito un raggio  $\simeq 1.5 h^{-1} \text{Mpc}$ , il *raggio di Abell*. L'espressione (II.43) e la presenza apparente di scale tipiche si conciliano grazie al fatto che la distribuzione delle galassie è *non-Gaussiana*, come illustreremo tra breve.

Una gran parte del mondo astronomico, però, rimase a lungo scettica di fronte all'assenza di un qualsiasi segno della scala degli ammassi. Misure recenti mettono oggi in evidenza qualche deviazione dalla legge di potenza, sebbene in misura assai minore di quanto gli astronomi dell'epoca si attendessero.

L'andamento con  $r$  della funzione di correlazione a 2 punti è oggi sostanzialmente compreso, anche nelle sue deviazioni da una singola legge di potenza, sulla base delle conoscenze sulla distribuzione della materia in aloni con profilo NFW, di raggio e concentrazione variabile. Tutto ciò verrà discusso nel capitolo VII.

In figura II.8, oltre all'andamento della funzione a 2 punti da una scala di  $0.1 h^{-1}\text{Mpc}$  a oltre  $2r_o$ , riportiamo il risultato del suo smussamento alla scala  $R = 1.5 h^{-1}\text{Mpc}$  (vedi appendice A1), nonché l'andamento della funzione di correlazione a 2 punti per gli ammassi di galassie (vedi sotto).

### Relazione tra funzione a due punti e varianza di massa

In appendice A2 si prova un'importante relazione tra funzione di correlazione a 2 punti e varianza di massa:

$$(II.44) \quad \sigma_R^2 = G(\gamma)\zeta(R),$$

in cui  $G(\gamma)$  è una funzione nota del solo esponente  $\gamma$ , esprimibile tramite un integrale multiplo. In figura II.9 è rappresentato l'andamento della funzione, sostanzialmente piatto nell'intervallo rilevante, con  $G(1.8) \simeq 2$ . La scala  $R_1$ , per cui  $\sigma_{R_1} = 1$ , sarà allora tale che  $r_o/R_1 \simeq (1/2)^{1/1.8} \simeq 0.68$ , per cui

$$(II.45) \quad R_1 \simeq 1.47 \times r_o \simeq 8 h^{-1}\text{Mpc}, \quad \text{ovvero} \quad \sigma_{8 h^{-1}\text{Mpc}} \simeq 1.$$

La *varianza di massa* su  $8 h^{-1}\text{Mpc}$  si denomina brevemente  $\sigma_8$ .

Se la *distribuzione delle galassie traccia la distribuzione di materia*,  $\sigma_8$  è prossimo all'unità.

Più in generale, se l'andamento di  $\zeta(R)$  è noto fino a un determinato  $R > r_o$ , ne

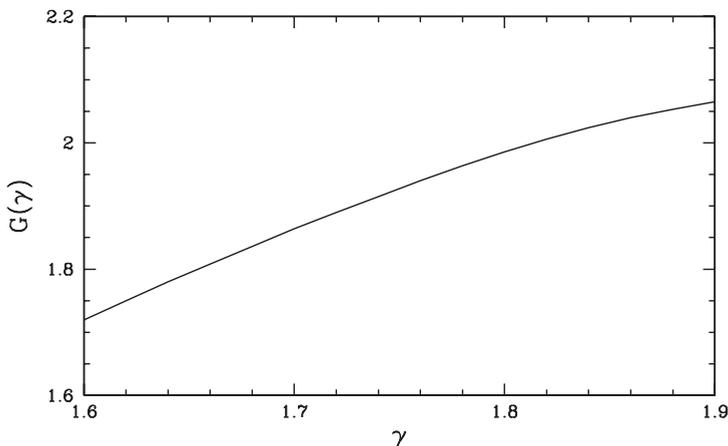


FIGURA II.9  
Andamento di  $G(\gamma)$ .

possiamo dedurre il valore di  $\sigma_R$  fino a tale  $R$ . Se, per esempio, la pendenza  $\gamma$  fosse costante fino a  $R = 3r_o$ , ne dedurremmo un valore  $\sigma_{25} \sim 0.3$ . Una maggiore pendenza  $\gamma$  implica un  $\sigma_{25}$  minore.

A una tale scala siamo quindi in regime mediamente lineare. Ciò non esclude però che, in qualche sito, siano presenti fluttuazioni eccezionali. Per esempio, una fluttuazione 3-4 volte sopra la media sarebbe sostanzialmente già entrata nel regime di non-linearità. Se la fluttuazione è positiva essa potrebbe dare luogo a un grosso ammasso di galassie, di massa  $\sim 5 \times 10^{15} \Omega_m h^{-1} M_\odot$ . Se negativa, la fluttuazione potrebbe dare invece luogo a un enorme vuoto dal diametro  $\sim 50 h^{-1} \text{Mpc}$ .

Strutture non virializzate, ma con contrasti di densità ormai significativi, sono invece ampiamente osservate su questa scala. Esse corrispondono a fluttuazioni 1.5-2 volte sopra la media e vengono definite *super-cluster*. Sebbene non si tratti di strutture già legate, la loro dinamica non è lineare.

Su scale superiori di un ulteriore fattore 2, non si riscontra nessuna struttura significativamente distante dal regime lineare.

### ***Il bias***

Bisogna inoltre considerare la possibilità che la distribuzione della massa non ricalchi quella delle galassie. In particolare, si sono valutate due ipotesi: (i) le galassie (o altri oggetti cosmici) si situano nei picchi di densità; (ii) tra le distribuzioni delle galassie (o di altri oggetti) e della materia c'è un fattore moltiplicativo. In parte, le due ipotesi sono collegate tra loro.

L'ipotesi (i) appare fisicamente ben fondata: gli oggetti di scala  $R$  si formano nei primi siti dove  $\epsilon_R(\mathbf{r})$  avvicina l'unità.

Consideriamo per esempio una scala  $R$  per la quale, a  $z = 0$ , risulti  $\sigma_R < 1$ . Le fluttuazioni positive di ampiezza media, su scala  $R$ , non hanno formato oggetti. Questa è la situazione, per esempio, sulla scala degli ammassi di galassie, che però sono in numero consistente. Ciò non è affatto un paradosso e anzi permetterà di valutare il valore di  $\sigma_R$  sulla scala degli ammassi, a condizione che la distribuzione delle ampiezze di fluttuazione sia inizialmente Gaussiana (vedi cap. VII).

Consideriamo poi una scala minore, su cui oggi sia  $\sigma_R > 1$ . In questa seconda condizione, si può pensare che i primi oggetti si siano formati dov'erano i picchi di densità, ma poi, gradualmente, se ne sono andati formando ovunque, arrivando finalmente a ricalcare la distribuzione della materia. Alternativamente, i primi oggetti formati potrebbero «drenare» materiali cosmici dai loro dintorni, alterando significativamente la distribuzione primordiale. Nelle zone «drenate» non si formeranno più altri oggetti, e la collocazione degli oggetti osservabili ricalcherebbe comunque quella dei picchi primordiali.

L'alternativa tra queste due ipotesi non si risolve sul piano fisico-matematico; occorre visitare in dettaglio fisica e fenomenologia, scala per scala.

È comunque possibile, per via analitica, calcolare la densità (in numero) di siti in cui

$\varepsilon_R$  supera di  $v$  volte  $\sigma_R$ . In appendice A5 si dimostra che il numero di oggetti derivanti da fluttuazioni di ampiezza superiore a  $v\sigma_R$  è

$$(II.46) \quad \langle n_{v,R} \rangle = 3(2\pi)^{-3/2} R^{-3} v^{-1} e^{-\frac{v^2}{2}},$$

mentre la loro funzione a 2 punti ha l'espressione

$$(II.47) \quad \xi^{(v,R)}(r) = \exp\left(\frac{v^2}{\sigma_R^2} \xi^{(R)}(r)\right) - 1.$$

Se l'argomento dell'esponenziale è  $\ll 1$ , risulta

$$(II.48) \quad \xi^{(v,R)}(r) \simeq \frac{v^2}{\sigma_R^2} \xi^{(R)}(r)$$

e tutto avviene come se ci fosse semplicemente un fattore moltiplicativo

$$(II.49) \quad b = v/\sigma_R,$$

detto *fattore di bias*, tra campo di fluttuazioni della materia e campo di fluttuazioni definito dagli oggetti osservati. Per esempio, se le fluttuazioni di densità materia e galassie,  $\varepsilon_m(\mathbf{r})$  e  $\varepsilon_g(\mathbf{r})$ , sono tali che risulta

$$(II.50) \quad \varepsilon_g(\mathbf{r}) = b \varepsilon_m(\mathbf{r}),$$

allora sarebbe

$$(II.51) \quad \xi_g(\mathbf{r}_1, \mathbf{r}_2) = \langle \varepsilon_g(\mathbf{r}_1) \varepsilon_g(\mathbf{r}_2) \rangle = b^2 \langle \varepsilon_m(\mathbf{r}_1) \varepsilon_m(\mathbf{r}_2) \rangle = b^2 \xi_m(\mathbf{r}_1, \mathbf{r}_2),$$

conformemente a eq. (II.48)-(II.49) e, analogamente,

$$(II.52) \quad \sigma_g^2(R) = b^2 \sigma_m^2(R).$$

Di fatto, l'argomento dell'esponenziale in eq. (II.47) coincide con la funzione a 2 punti delle galassie, o degli altri oggetti osservati, e assume anche valori  $\gg 1$ .

Ciononostante, l'assunto di un *bias lineare*, conforme alla relazione (II.50), si è dimostrato sinora sufficiente alla comprensione dei dati. Tra le due formulazioni (i) e (ii) considerate all'inizio di questo sottoparagrafo, si segue dunque l'ipotesi (ii), sottintendendo che essa è una buona approssimazione per l'ipotesi (i).

Come già sottolineato, per le galassie, anche sulla base dei dati osservativi più recenti, risulta  $\sigma_g(8 h^{-1} \text{Mpc}) \simeq 1$ . L'ampiezza della varianza, sulla scala di  $8 h^{-1} \text{Mpc}$ , viene anche semplicemente denominata  $\sigma_8$ . Per la materia, dunque

$$(II.53) \quad \sigma_8 = b^{-1}$$

e, quando si assegna un valore per  $\sigma_8$ , grazie a un determinato insieme di osservazioni o sulla base di opportuni assunti teorici, si intende di norma riferirsi a quello della materia, che è l'inverso del *fattore di bias delle galassie*.

## 12. Funzioni a $n$ punti

Nel paragrafo precedente si è visto che la probabilità di trovare una galassia in un volume  $\delta V_1$  è diversa se prendiamo un  $\delta V_1$  generico o se prendiamo un  $\delta V_1$  a distanza  $r$  da un volume  $\delta V_2$  in cui si situa un'altra galassia.

Supponiamo ora di considerare 3 volumi  $\delta V_i$  ( $i = 1, 2, 3$ ). La probabilità congiunta di trovare una galassia in ciascuno dei 3 volumi sarà data da

$$(II.54) \quad \langle \delta N_1 \delta N_2 \delta N_3 \rangle = n^3 \delta V_1 \delta V_2 \delta V_3 [1 + \zeta(r_{12}) + \zeta(r_{23}) + \zeta(r_{31}) + \zeta^{(3)}(r_{12}, r_{23}, r_{31})].$$

Tra gli addendi in parentesi quadra a secondo membro, l'unità esprime la probabilità complessiva che ci sarebbe se le tre probabilità fossero tra loro scorrelate. I 3 addendi in cui compare la funzione a 2 punti, invece, esprimono la correlazione esistente tra le probabilità di trovare simultaneamente una galassia in ciascuna delle 3 coppie di punti. Se, oltre a ciò, esiste un sovrappiù di probabilità in ciascun punto, dovuta non alla semplice presenza di galassie negli altri due punti, ma alla loro *simultanea* presenza, esso viene espresso dalla funzione (di correlazione) a 3 punti  $\zeta^{(3)}(r_{12}, r_{23}, r_{31})$ .

La ricetta si può estendere a numeri crescenti di punti, dove però si trovano crescenti complicazioni topologiche e di altra natura, come si vede già nell'esempio della funzione a 4 punti trattata in appendice A3.

I dati osservativi mostrano che, per la funzione a 3 punti si può dare una forma precisa, denominata *espressione del modello gerarchico*:

$$(II.55) \quad \zeta_{123}^{(3)} = Q [\zeta_{12}\zeta_{23} + \zeta_{23}\zeta_{31} + \zeta_{31}\zeta_{12}]$$

con  $Q \sim 1.2 \pm 0.1$ , tendenzialmente maggiore dell'unità, ma compatibile con essa. In questa espressione, la funzione a 3 punti appare quindi come una combinazione di prodotti di 2 funzioni a 2 punti, completamente simmetrica nei 3 punti considerati.

Vi sono analoghe espressioni del modello gerarchico che valgono per le funzioni a  $n$  punti ( $n > 3$ ; vedi appendice A3). La ragione della validità di queste espressioni non è del tutto chiarita.

Si noti che, anche in presenza di *clustering* delle galassie, le funzioni a  $n$  punti, con  $n > 2$ , potrebbero essere tutte nulle. Il fatto che esse non lo siano esprime la *non-Gaussianità* della presente distribuzione delle galassie nel cosmo.

In appendice A2 si discute il significato di *distribuzione Gaussiana*, nel caso in esame. In effetti, se abbiamo a che fare con una singola variabile stocastica  $\varepsilon$ , l'affermazione che essa è distribuita in modo Gaussiano ha un significato semplice: la probabilità di ciascun valore  $\varepsilon$  è ottenibile dalla *curva a campana*

$$(II.56) \quad P(\varepsilon) = \frac{h}{\sqrt{2\pi}} e^{-h^2 \varepsilon^2 / 2}.$$

Questo concetto si può estendere alla distribuzione su  $n$  o su un'infinità discreta di variabili  $\varepsilon_i$ . Nel caso fisico esso deve essere ulteriormente esteso fino alla distribuzione

su un'infinità continua di variabili  $\varepsilon(\mathbf{x})$ , in cui la variabile  $\mathbf{x}$  viene a assumere lo stesso ruolo dell'indice  $i$ .

In presenza di non-Gaussianità si dimostra che saranno non-nulle alcune o tutte le funzioni a  $n$  punti. Sulle espressioni delle funzioni a  $n$  punti con  $n > 3$  vedi appendice A3.

### 13. Lo spettro delle fluttuazioni

Il campo di fluttuazioni  $\varepsilon(\mathbf{x})$  descrive la distribuzione della materia nello spazio. Ricordiamo, per esempio, che la varianza di massa su una determinata scala,  $\sigma_R$ , è il valore quadratico medio di  $\varepsilon_R(\mathbf{x})$ , cioè del campo di fluttuazioni smussato sulla scala  $R$ .

Di questo campo è essenziale determinare le *componenti armoniche*, come si ricercano le singole note musicali contenute in un suono, e la relazione col suono andrà molto avanti, quando scopriremo *onde di compressione*, del tutto analoghe alle *onde soniche*, nella materia primordiale ionizzata, a temperature superiori ai 3000 K, alcune centinaia di migliaia d'anni dopo il *Big-bang*.

L'operazione matematica che fornisce le componenti armoniche è la *trasformata di Fourier* e l'importanza di lavorare su queste componenti è legata alla *linearità* delle equazioni dinamiche, quando le fluttuazioni sono piccole, che permette di studiare separatamente l'evoluzione su ogni componente, grazie al *principio di sovrapposizione*.

Anche senza entrare nei dettagli della trasformata di Fourier (per cui si veda l'appendice A4), basta rendersi conto che il campo di fluttuazioni può essere scritto nella forma

$$(II.57) \quad \varepsilon(\mathbf{x}) = \sum_{\mathbf{k}} \delta(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}},$$

dove ciascuna componente armonica ha numero d'onda  $\mathbf{k}$  e ampiezza

$$(II.58) \quad \delta(\mathbf{k}) = |\delta(\mathbf{k})| \exp(i\theta_{\mathbf{k}}).$$

Come per un suono, la relazione tra numero e lunghezza d'onda si scrive

$$(II.59) \quad \mathbf{k} = \mathbf{n} \frac{2\pi}{\lambda}$$

dove  $\mathbf{n}$  è il versore di  $\mathbf{k}$ . Nelle fasi in cui la singola onda si propaga, con velocità  $c_s$ , la sua frequenza sarà  $\nu = c_s/\lambda$  e il suo periodo  $P = 1/\nu$ .

Tramite le ampiezze  $\delta(\mathbf{k})$  si costruisce lo *spettro* del campo di fluttuazioni  $\varepsilon(\mathbf{x})$ . Mentre conviene che  $\delta(\mathbf{k})$  sia una quantità complessa, lo spettro

$$(II.60) \quad P(k) = \langle |\delta(\mathbf{k})|^2 \rangle$$

è una quantità reale, in cui non entrano le fasi  $\theta_{\mathbf{k}}$ . La procedura di media indicata dal simbolo  $\langle \dots \rangle$  va effettuata tra diverse realizzazioni del processo che distribuisce le

disomogeneità cosmiche. Si cerca di ottenere questo risultato mediando su ampi volumi nello spazio. Va anche detto che, concettualmente, questa media è ancora più problematica di altre analoghe. La bontà del risultato è peraltro *certificata* dall'indipendenza della stima di  $P(\mathbf{k})$  dalla direzione  $\mathbf{n}$  del numero d'onda  $\mathbf{k}$ . Così non fosse, ci sarebbero direzioni privilegiate nello spazio e le fluttuazioni sarebbero distribuite in modo anisotropo.

Lo spettro non dipende dalle fasi  $\theta_k$ , ma un osservabile che da esse dipendesse dovrà comunque essere anch'esso funzione di  $k$  e non di  $\mathbf{k}$ .

In appendice A4 si dimostra che, se vale l'espansione (II.57) per  $\varepsilon(\mathbf{k})$ , valgono le espansioni

$$(II.61) \quad \begin{aligned} \xi(r) &= \frac{1}{2\pi^2} \sum_k \frac{\sin(kr)}{kr} k^3 P(k) \\ \sigma_R^2 &= \frac{1}{2\pi^2} \sum_k \tilde{W}^2(kR) k^3 P(k) \end{aligned}$$

per la funzione a due punti e la varianza di massa; quest'ultima con

$$\tilde{W}(kR) = \frac{3}{(kR)^3} (\sin kR - kR \cos kR).$$

In figura II.10 si confrontano i coefficienti di  $k^3 P(k)$  nei due casi, mostrando che tendono a 1 per  $kR, kr \rightarrow 0$ , mentre i contributi dei diversi valori di  $k$  sono pesati in modo molto simile, almeno fino a  $kR, kr \sim \pi$ , quando il loro valore supera 0.2-0.3. Ciò conferma il risultato espresso in eq. (II.44).

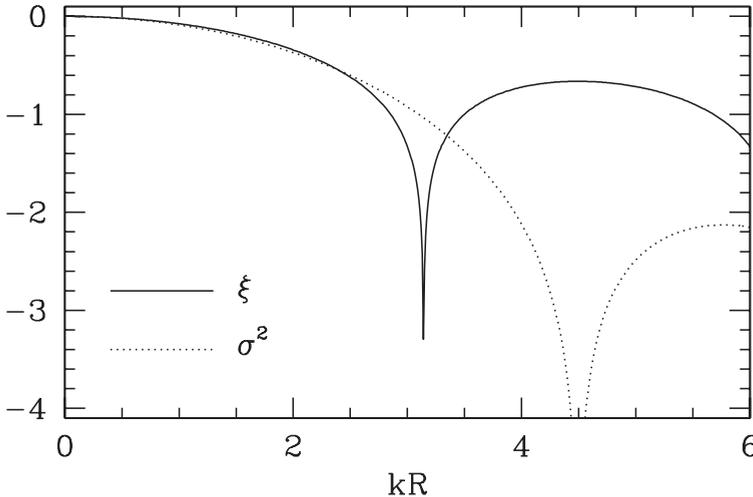


FIGURA II.10

Andamento con  $kR$  del logaritmo in base 10 dei coefficienti di  $k^3 P(k)$  nelle espressioni (II.61).

Accanto allo spettro  $P(k)$  si usa la *funzione spettrale*

$$(II.62) \quad \Delta^2(k) = \frac{1}{2\pi^2} k^3 P(k)$$

che permette di scrivere semplicemente

$$(II.63) \quad \zeta(r) = \sum_k \Delta^2(k) \frac{\sin(kr)}{kr}, \quad \sigma_R^2 = \sum_k \Delta^2(k) \tilde{W}(kR)$$

Se assumiamo una legge di potenza per lo spettro,  $P(k) = A k^{n_{\text{eff}}}$ , le espressioni (II.61) danno

$$\zeta(r) = \frac{A}{2\pi^2} r^{-(n_{\text{eff}}+3)} \sum_{kr} \frac{\sin(kr)}{kr} (kr)^{n_{\text{eff}}+3}.$$

Al limite in cui le componenti armoniche sono distribuite su tutti i valori di  $k$ , la sommatoria diviene l'integrale definito  $\int_0^\infty dx x \sin x$ , che non dipende da  $r$  (vedi appendice A4). Se è  $\zeta(r) \propto r^{-\gamma}$ , ne segue che

$$(II.64) \quad n_{\text{eff}} = 3 - \gamma.$$

Considerazioni analoghe per  $\sigma_R^2$  portano alla stessa conclusione.

Le eq. (II.61) mostrano che la varianza di massa e la *funzione a 2 punti non dipendono dalla distribuzione delle fasi*, in quanto non ne dipende lo spettro. Se la distribuzione delle fasi è casuale, pertanto, tutta la statistica risulta determinata dallo spettro o dalla funzione a due punti.

Poiché l'annullarsi delle funzioni a  $n$  punti (con  $n > 2$ ) indica che la distribuzione è Gaussiana, se ne conclude che, in questo caso, nessuna informazione può essere contenuta nella distribuzione delle fasi  $\theta_k$  sicché, *in una distribuzione Gaussiana le fasi sono distribuite in modo casuale*.

La distribuzione osservata delle galassie è quindi non-Gaussiana.

Al di là degli aspetti formali, è importante sottolineare che la non-Gaussianità osservata nella distribuzione della materia su scala galattica ha un contributo essenziale dalla dinamica non-lineare delle fluttuazioni, quando la loro ampiezza si avvicina all'unità. Un ulteriore contributo alla non-Gaussianità nella distribuzione delle galassie può provenire dal *bias*, ove la distribuzione delle galassie e della materia siano diverse. Le fluttuazioni primordiali, invece, si suppongono Gaussiane. Ciò è confermato dalle osservazioni sul CMB, che stabiliscono limiti stringenti alla non-Gaussianità. La (non-)Gaussianità delle fluttuazioni primordiali è importante in relazione alle teorie che ne studiano l'origine, a partire dai possibili modelli di *inflazione*.

Le informazioni sulla distribuzione delle disomogeneità, su scale  $\gg r_o \simeq 5-6 h^{-1} \text{Mpc}$ , sono espresse convenientemente in termini di spettro. C'è una rozza corrispondenza tra valori di  $r$  e di  $k$ , per cui il grosso dei contributi a  $\sigma_R^2$  proviene da valori  $k \sim 2\pi/R$ , come se,

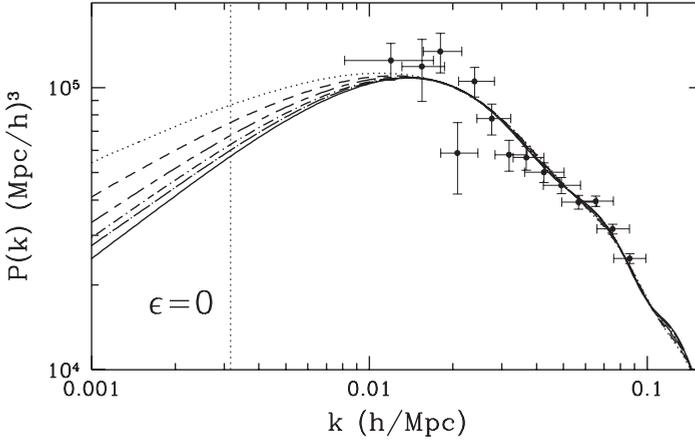


FIGURA II.11

Dati sullo spettro delle fluttuazioni provenienti dalla SLOAN *Digital Sky Survey*, con le loro barre d'errore, confrontati con le previsioni di un insieme di modelli.

da eq. (II.63), si potesse estrapolare

$$(II.65) \quad \sigma_R^2 \sim \Delta^2 \left( \frac{2\pi}{R} \right) = 4\pi P \left( \frac{2\pi}{R} \right) R^{-3}$$

A scale  $\gg r_o$  corrispondono quindi  $k \ll 1 h\text{Mpc}^{-1}$ .

In figura II.11 sono esibiti dati recenti sullo spettro a  $k \ll 1 h\text{Mpc}^{-1}$ , confrontati con le previsioni di determinati modelli cosmologici. L'andamento di  $P(k)$  si suppone quindi comunque crescente a piccoli  $k$  (dato confermato dalle osservazioni del CMB, qui non riportate) raggiunge poi un massimo, per declinare a valori maggiori. Come si vede in figura, i dati si estendono ormai fino a  $k \sim 0.01$ , corrispondenti dunque a scale  $\sim 600 h^{-1}\text{Mpc}$ , quasi 1/10 dell'orizzonte. Su queste scale, lo spettro è risalito di un fattore  $\sim 20$  rispetto ai valori che ha per  $k \sim 2\pi/r_o$ , mentre il coefficiente  $R^{-3}$  è calato di 3 ordini di grandezza. Complessivamente abbiamo un decremento della varianza per un fattore  $\sim 50$ . Sulla scala indicata l'ampiezza quadratica media delle fluttuazioni è quindi  $\sim 10^{-2}$ ; sono dunque disponibili dati che mostrano fluttuazioni nella distribuzione delle galassie in regime ampiamente lineare, in condizioni ormai assai prossime all'omogeneità primeva.

#### 14. I vuoti cosmici e la funzione di probabilità dei vuoti

Nel campione di galassie 2dF, in figura II.1, pur esibito in proiezione bidimensionale, sono evidenti ampie zone di «vuoto», specialmente a  $z > 0.1$ , dove il campione comprende solo galassie di alta luminosità intrinseca.

Una misura statistica dei «vuoti» deve essere comunque riferita alle galassie con luminosità intrinseca superiore a un determinato limite  $\bar{L}$ .

Se si immagina di distribuire casualmente  $N^3$  particelle in un cubo di lato  $R$ , sicché la loro distanza media è  $\delta = R/N$ , nel cubo ci saranno abbondanti sfere vuote di diametro  $\delta$ , ma anche qualche sfera vuota di diametro  $d > \delta$ ; il loro numero andrà comunque rapidamente scemando al crescere di  $d/\delta$ . Con opportune tecniche statistiche tale probabilità può essere stimata e l'espressione di tale stima è data in eq. (II.67), qui sotto.

Una distribuzione del tutto casuale di punti si dice *Poissoniana*. Anche per una distribuzione *Gaussiana* la probabilità di trovare sfere vuote può essere stimata (vedi sotto).

Quando passiamo a esaminare la distribuzione reale delle galassie, dobbiamo immaginare di gettare, nel volume  $\mathcal{V}$  coperto da un campione osservativo, sfere di raggio  $R$  (e volume  $V_R$ ), con centri distribuiti a caso, «in modo Poissoniano», con tutte le cautele necessitate dal fatto che, in genere,  $\mathcal{V}$  non è un cubo. Si misura così la probabilità  $P_0(R)$  che esistano sfere di raggio  $R$  non contenenti galassie di luminosità  $> \bar{L}$ . Essa si denomina *funzione di probabilità dei vuoti*.

Esiste una relazione tra  $P_0(R)$  e le funzioni di correlazione a  $n$  punti:

$$(II.66) \quad P_0(R) = \exp \left[ -\bar{N}_R + \sum_{n=2}^{\infty} \frac{(-\bar{N}_R)^n}{n!} \bar{\xi}_R^{(n)} \right].$$

In questa relazione  $\bar{N}_R$  è il numero medio di galassie in  $V_R$ ;  $\bar{\xi}_R^{(n)}$  sono medie delle funzioni a  $n$  punti, ricavate lasciando variare gli  $n$  punti entro  $V_R$  e mediando tra i valori così ottenuti. La somma si estende da 2 a  $\infty$ , sicché la funzione di probabilità dei vuoti riceve contributi da tutte le funzioni di correlazione, a partire da quella a 2 punti.

Per distribuzioni Gaussiane, dunque, la  $P_0$  è calcolabile a partire dalla  $\zeta$ . Nel caso di distribuzioni non-Gaussiane, il calcolo si può effettuare sulla base di un'ipotesi concernente le espressioni nelle funzioni a  $n$  punti, che non saranno mai tutte osservativamente note. Assumendo le espressioni del modello gerarchico, per esempio, restano da fissare delle costanti simili a  $Q$ , per ogni  $n$ . Pertanto, pur essendo correlata alle funzioni a  $n$  punti, la funzione di probabilità dei vuoti dà risultati indipendenti.

Nel caso di una distribuzione Poissoniana, essendo nulle tutte le funzioni di correlazione, sarà

$$(II.67) \quad P_{o,Poisson}(R) = \exp(-\bar{N}_R);$$

ciò che ha rilievo statistico è dunque la discrepanza tra  $P_0(R)$  e  $P_{o,Poisson}(R)$ . La figura II.12 mostra una misura della funzione di probabilità dei vuoti per un campione di galassie, confrontata con la funzione di probabilità dei vuoti per un campione Poissoniano con numero di punti pari a quello delle galassie. La discrepanza, in figura II.12, deve essere confrontata con gli errori di misura, che dipendono dal campione

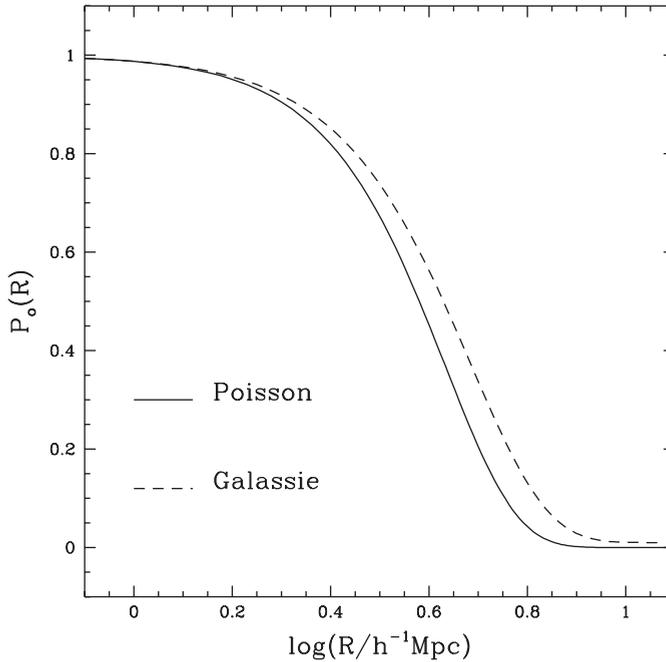


FIGURA II.12

Funzione di probabilità dei vuoti, per un campione di  $N$  galassie, confrontata con la stessa funzione, per un insieme di  $N$  punti gettati a caso.

considerato. In genere  $P_0(R)$  è ben determinata e permette di discriminare non solo tra dati e previsione Poissoniana, come in figura, ma anche tra dati e previsione Gaussiana con funzione a 2 punti data dall'espressione (II.43), confermando quindi la non-Gaussianità della distribuzione delle galassie.

Tra i dati che figura II.12 evidenzia, c'è il valore di  $P_0$  a  $R \simeq 10 h^{-1} \text{Mpc}$ . Nel caso Poissoniano, esso è praticamente nullo, mentre la differenza da zero, per il campione di galassie, non è trascurabile, specie se si tiene conto che l'ordinata, in figura, non è logaritmica. In effetti, esistono vuoti con raggi anche  $> 20 h^{-1} \text{Mpc}$ . Essi sono dunque il frutto di una «congiura» delle fasi tra le armoniche in cui si decompone  $\epsilon(\mathbf{x})$ , ciò che è possibile – appunto – perché la distribuzione di galassie è non-Gaussiana e le sue fasi non sono distribuite a caso.

Una conseguenza dell'esistenza dei grandi vuoti è che, volendo compilare un campione statistico di galassie, si debbono esplorare scale 4-5 superiori a quelle dei vuoti più grandi, cioè fino a  $D \sim 200 h^{-1} \text{Mpc}$  ( $\sim 300 \text{Mpc}$  se  $h \sim 0.7$ ). Come già sottolineato, la velocità di espansione del cosmo a questa distanza è  $H \times D = 100 h \cdot 200 h^{-1} \text{km/s} \sim 20\,000 \text{km/s} \sim 0.07 c$ ; essa causa un *redshift*  $z_D \sim 0.07-0.1$ .

Ciò non è senza conseguenze, in quanto gli effetti evolutivi a questo *redshift* cominciano a non essere trascurabili.

## 15. Funzioni di correlazione degli ammassi di galassie e segregazione in luminosità

Le galassie costituiscono l'oggetto più cospicuo osservabile in banda ottica. Gli ammassi (*cluster*) di galassie, invece, sono gli oggetti più cospicui in banda X, come testimoniarono per primi i dati dei satelliti *Uhuru* e *Rosat*. Inoltre essi costituiscono di norma entità ben delineate anche nel visibile, dove si presentano come associazioni di un gran numero di galassie entro raggi  $R_c \sim 1.5 h^{-1} \text{Mpc}$ , e dove sono stati inizialmente scoperti.

La prima misura significativa della funzione di correlazione a due punti dei cluster,  $\xi_c(r)$ , fu basata su dati ottici. A priori ci si poteva attendere che  $\xi_c(r)$  fosse ottenibile dalla funzione a due punti delle galassie, smussandola su una scala  $\sim R_c$ . Essa avrebbe allora avuto l'andamento rappresentato in figura II.8 dalla linea tratteggiata.

I dati erano consistenti con una legge di potenza e valori di  $\gamma$  compatibili con quelli delle galassie, o leggermente superiori. Viceversa, il valore trovato per la lunghezza di correlazione era nettamente superiore. Nel caso di ammassi «ricchi», selezionati in banda ottica da Abell, si ottenne una curva del tipo mostrato in figura II.8, simile a quella delle galassie, ma con  $r_0$  circa 5 volte superiore.

Sulla base di dati più recenti si può sintetizzare un quadro d'insieme, che comprende galassie, gruppi e ammassi, per i quali la funzione a 2 punti ha comunque la forma

$$(II.68) \quad \xi(r) = \left(\frac{r_0}{r}\right)^\gamma$$

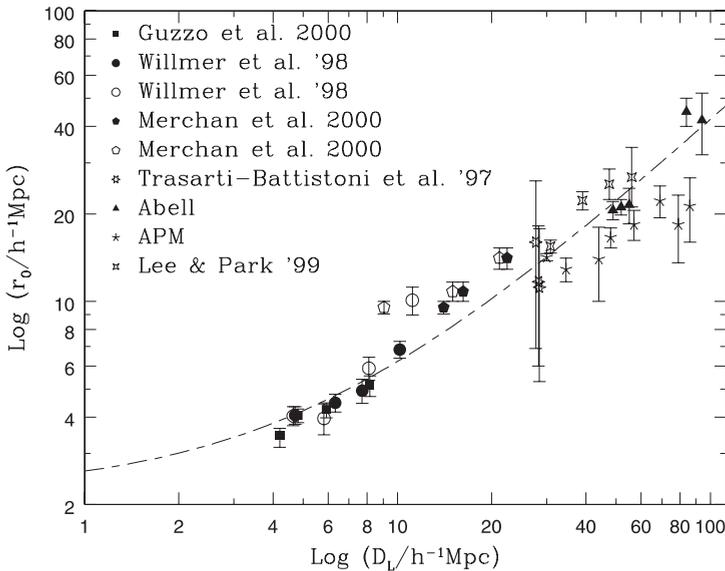


FIGURA II.13

Lunghezza di correlazione  $r_0$ , in funzione della separazione media tra gli oggetti considerati  $D_L$ , che sono galassie, gruppi di galassie, ammassi di galassie.

con  $r_o$  variabile a seconda della densità in numero  $n = D^{-3}$  degli oggetti considerati. Via via che si escludono gli oggetti più fiochi e si considerano solo quelli con luminosità  $> L$ , di densità (in numero)  $n(> L)$ , va crescendo la loro distanza media  $D_L [\simeq n^{-1/3}(> L)]$ .

In figura II.13 si riportano stime di  $r_o$  effettuate su campioni di galassie con  $D_L$  crescente. Il grafico mostra una modesta *segregazione in luminosità*:  $r_o$  cresce da  $\sim 3.5$  e  $\sim 8 h^{-1}$ Mpc quando  $D_L$  va da  $\sim 4$  e  $\sim 10 h^{-1}$ Mpc.

Nello stesso grafico si considerano anche *gruppi di galassie*, sistemi legati non particolarmente cospicui, ma ben individuati sia in ottico che in  $X$ , con masse tra  $\sim 10^{12}$  e  $10^{13.5} h^{-1} M_\odot$  e distanze medie  $D_L$  tra  $\sim 10$  e  $20 h^{-1}$ Mpc.

Nel grafico, si considerano infine anche *ammassi di galassie*, a vario limite di luminosità, con  $D_L$  fino a  $\sim 100 h^{-1}$ Mpc.

I dati sono ancora soggetti a ampie incertezze statistiche e la curva

$$(II.69) \quad r_o(D_L) = 0.4 D_L + 2.2,$$

rappresenta un tentativo di interpolazione, senza una precisa motivazione fisica.

La variabilità della distanza di correlazione mostra che nessuna famiglia di oggetti, *a priori*, può essere preferita per misurare la lunghezza di correlazione della materia che è, presumibilmente,  $\leq$  di ciascuno degli  $r_o$  in eq. (II.69). Ciò conferma l'esistenza del *bias*, per cui gli oggetti più luminosi nascono da fluttuazioni di ampiezza maggiore.

## 16. Uso di campioni angolari. Relazioni di scala

Abbiamo sin qui fornito stime sulle funzioni di correlazione senza approfondire come esse vengano ricavate dai campioni. Le tecniche sono oggi assai sviluppate e sono disponibili programmi numerici «di libreria».

Vale la pena però di ricordare che i primi campioni davano solo le coordinate angolari delle galassie. Depurandoli dagli effetti di proiezione, si riuscì tuttavia a dedurre gli andamenti delle funzioni a 2, 3 e 4 punti. In appendice A6 diamo informazioni sull'*equazione di Limber* e le sue generalizzazioni, che si usano a questo scopo. Esse non hanno solo rilevanza storica: sinché *tutti* i *redshift* delle galassie osservabili non saranno misurati, saranno comunque disponibili campioni angolari più ricchi dei campioni con *redshift*. Inoltre, lo studio dei campioni angolari mette in evidenza delle relazioni di scala (*scaling*) analoghe alla  $\log N$ - $\log S$ , che permettono di evidenziare e studiare effetti evolutivi o violazioni della geometria euclidea, ancora a prescindere da misure di *redshift*.

In questo paragrafo forniremo quindi la definizione della funzione di correlazione a 2 punti angolare e esplicheremo le succitate relazioni di scala.

Consideriamo un campione angolare di galassie con magnitudine  $< m_l$ ; detti  $\delta N_1$  e  $\delta N_2$  i numeri di galassie del campione che si trovino negli angoli solidi  $\delta \Omega_1$  e  $\delta \Omega_2$ , siti a

distanza angolare  $\theta_{12}$ , si consideri il valore di aspettazione

$$(II.70) \quad \langle \delta N_1 \delta N_2 \rangle = n_{\Omega}^2(< m_l)[1 + w_{m_l}(\theta_{12})] \delta \Omega_1 \delta \Omega_2,$$

che esprime la probabilità congiunta di osservare una galassia in entrambi gli angoli solidi;  $n_{\Omega}(< m_l)$  è la densità angolare di galassie con magnitudine  $< m_l$ .

La funzione angolare a 2 punti,  $w_{m_l}(\theta)$ , dipende dalla magnitudine limite  $m_l$  o, equivalentemente, dalla *profondità*  $D^*$  del campione, per cui la indicheremo anche con  $w_{D^*}$ .

L'equazione di Limber dimostra che, se la funzione spaziale a 2 punti ha la forma  $\zeta(r) = (r_o/r)^\gamma$ , la funzione angolare è

$$(II.71) \quad w_{D^*}(\theta) = A(\gamma) \left( \frac{r_o}{D^*} \right)^\gamma \theta^{1-\gamma}$$

( $D^*$ : profondità del campione). La pendenza della funzione angolare a 2 punti è  $\beta = \gamma - 1$ , attutita di un'unità rispetto a quella spaziale. Storicamente, proprio dal valore di  $\beta \simeq 0.7-0.8$ , fu rilevata per la prima volta la pendenza della funzione spaziale a 2 punti.

Se  $A(\gamma)$  è indipendente dalla profondità del campione di galassie, dell'equazione (II.71) è facile vedere che vale la relazione di scala

$$(II.72) \quad w_{D_1^*}(\theta) = \frac{D_2^*}{D_1^*} w_{D_2^*} \left( \theta \frac{D_1^*}{D_2^*} \right).$$

In realtà, il coefficiente  $A(\gamma)$ , oltre che da  $\gamma$ , dipende dalla funzione di luminosità delle galassie; la sua espressione è fornita in appendice A6 (eq. II.57 e successive). L'evoluzione della funzione di luminosità implica perciò violazioni delle relazioni di *scaling* (eq. II.71 e II.72), che sarebbero esatte solo se la funzione di luminosità non evolvesse e  $A(\gamma)$ , quindi, non dipendesse dalla profondità del campione. Altre violazioni dello *scaling*, oltre che dall'eventuale limitatezza del campione, possono trarre origine da violazioni della geometria euclidea e dalla correzione  $-\kappa$ . Dati opportuni assunti geometrici e evolutivi, tali violazioni sono prevedibili e possono essere messe a raffronto coi dati.

Abbiamo quindi un secondo modo, dopo quello definito dalla curva  $\log N$ - $\log S$ , di studiare l'evoluzione degli oggetti cosmici e la geometria dell'Universo.

## 17. Conclusioni

In questo capitolo, dall'intimità del nostro microcosmo planetario, abbiamo compiuto un'ispezione del cosmo. I suoi abitanti si sono rivelati essere galassie, gruppi e ammassi; abbiamo quindi approntato le tecniche per studiarne la distribuzione e, in particolare, per individuare la scala di omogeneità. Abbiamo quindi definito la varianza di massa, ci siamo imbattuti nella statistica dei vuoti, per concentrarci a lungo sulle funzioni di correlazione delle galassie.

Questa statistica e in particolare la funzione a due punti  $\xi(r)$  si sono rivelate assai importanti; nel regime non-lineare, risulta  $\xi(r) \propto r^{-\gamma}$ , ma  $\xi(r)$  si valuta fino a scale non ancora entrate nella non-linearità. Le funzioni di correlazione di ordine più elevato, invece, misurano quanto le attuali disomogenità si siano allontanate dal regime Gaussiano.

C'è dunque un ricco messaggio statistico nella distribuzione delle galassie nel cielo. In parte esso è determinato dalle leggi della crescita delle disomogeneità per auto-gravità, ma in parte esso reca la memoria della distribuzione primordiale delle tenui disomogeneità «iniziali». In questa distribuzione sono criptati i meccanismi fisici che presiedono all'origine stessa delle disomogeneità.

L'attenzione alla distribuzione delle galassie, prima di introdurre gli stessi modelli cosmologici, è ben motivata: proprio di lì l'avventura cosmologica si è dipartita, quando Edwin Hubble ebbe l'ispirazione della proporzionalità tra velocità di allontanamento e distanza,  $v = Hd$ .

La legge di Hubble, *a posteriori*, si è dimostrata una realtà inconfutabile. I modelli cosmologici, proposti quando non si sapeva nemmeno se esistesse una scala di omogeneità, si dimostreranno aderenti ai dati. Ma la lunghezza stessa di questo capitolo, in cui molti temi sono appena sfiorati per essere approfonditi quando vengono sviluppate le dinamiche che presiedono ai dati, richiama l'attenzione del lettore sulla distanza tra realtà sperimentale pre-cosmologica e teoria.

# Appendice A1

## Smussamento e filtri

### 1. I filtri a top-hat e Gaussiano

In un contesto cosmologico, la densità  $\rho(\mathbf{x})$  o le fluttuazioni di densità  $\varepsilon(\mathbf{x}) = \delta\rho(\mathbf{x})/\rho$  vengono «ripulite» dalle oscillazioni sulle scale piccole, tipicamente sub-galattiche. La ricetta più semplice per tale smussamento consiste nell'attribuire a ciascun sito  $\mathbf{x}$  la densità media in una sfera di raggio  $R$  centrata in  $\mathbf{x}$ :  $\rho_R(\mathbf{x})$ . Analogamente, per le fluttuazioni di densità si pone

$$(A1.1) \quad \varepsilon_R(\mathbf{x}) = \frac{1}{V_R} \int_{V_R(\mathbf{x})} d^3u \varepsilon(\mathbf{u}),$$

dove  $V_R(\mathbf{x})$  è la sfera di raggio  $R$  centrata in  $\mathbf{x}$ ; l'eq. (A1.1) si può riscrivere

$$(A1.2) \quad \varepsilon_R(\mathbf{x}) = \int_0^\infty du u^2 \int_{4\pi} d\Omega W_R(|\mathbf{x} - \mathbf{u}|) \varepsilon(\mathbf{u}).$$

In questa espressione compare la **funzione di smussamento o filtro**

$$(A1.3) \quad W_R(y) = 1/V_R \text{ per } y < R, \quad W_R(y) = 0 \text{ altrove,}$$

chiaramente normalizzata a 1, che attribuisce ugual peso a tutti i valori del campo  $\varepsilon(\mathbf{u})$  per  $|\mathbf{u} - \mathbf{r}| < R$ . La brusca discontinuità di  $W_R$  a  $y = 1$  può talora causare effetti spuri. Si cerca di evitarli usando altri filtri, che pesino  $\varepsilon(\mathbf{u})$  in modo decrescente al crescere della distanza  $|\mathbf{u} - \mathbf{r}|$  senza mai bruscamente annullarsi. Il filtro più comune usato a questo scopo è il **filtro Gaussiano**

$$(A1.4) \quad W_R(y) = (2\pi R^2)^{-3/2} \exp[-(1/2)(y/R)^2],$$

prodotto di distribuzioni Gaussiane normalizzate, di varianza  $R$ , lungo i 3 assi.

Si noti che i filtri, essendo invarianti di scala, hanno comunque la forma

$$(A1.5) \quad W_R(y) = (1/V_R)w(y/R);$$

il coefficiente  $1/R^3$  assicura l'invarianza dell'integrale di volume.

Lo smussamento agisce sulle componenti di Fourier  $\delta(\mathbf{k})$  di  $\varepsilon(\mathbf{x})$ , deprimendo quelle di lunghezza d'onda  $2\pi/k < R$ .

Si può infatti mostrare (vedi sotto) che le componenti di  $\varepsilon_R(\mathbf{x})$  sono

$$(A1.6) \quad \delta_R(\mathbf{k}) = \delta(\mathbf{k})\tilde{W}(kR), \quad \text{dove} \quad \tilde{W}(kR) = \int_0^\infty d^3s W_R(s) e^{i\mathbf{k}\cdot\mathbf{s}}.$$

Grazie all'eq. (A1.5) e ponendo  $\mathbf{s} = \mathbf{u}R$ , si vede poi che

$$\begin{aligned}\tilde{W}(kR) &= V_R^{-1} \int_0^\infty d^3s w(s/R) e^{i\mathbf{k}\cdot\mathbf{s}} = V_R^{-1} R^3 \int_0^\infty d^3u w(u) e^{i\mathbf{k}\cdot\mathbf{u}R} = \\ &= \frac{3}{4\pi} \int_0^\infty du u^2 w(u) 2\pi \int_{-1}^1 d \cos \theta e^{ikRu \cos \theta} = \frac{3}{kR} \int_0^\infty du u w(u) \sin(kRu)\end{aligned}$$

è funzione esclusivamente del prodotto  $kR$ .

In particolare, per i filtri top-hat e Gaussiano, risulta rispettivamente

$$(A1.7) \quad \tilde{W}^{(t.h.)}(x) = \frac{3}{x^3} (\sin x - x \cos x), \quad \tilde{W}^{(G)}(x) = e^{-x^2/2}.$$

Difatti, per il filtro a top-hat, risulta

$$\tilde{W}^{(t.h.)}(x) = \frac{3}{x} \int_0^1 du u \sin(xu) = \frac{3}{x^3} \int_0^x dv v \sin v = \frac{3}{x^3} (\sin x - x \cos x)$$

L'analogia verifica per il filtro Gaussiano, per cui  $w(u) = (2\pi)^{-3/2} \exp(-u^2/2)$ , è quasi immediata se si tiene conto della relazione

$$\int_0^\infty du u e^{-p^2 u^2} \sin xu = x\sqrt{\pi} p^{-3} \exp(-x^2/4p^2)$$

(Gradshteyn & Ryzhik; 3.952.1) in cui va posto  $p = 1/\sqrt{2}$ .

Entrambe le  $\tilde{W}$  tendono a 1 per  $x \rightarrow 0$  (scale  $\gg R$ ). Per  $x \rightarrow \infty$ , il filtro Gaussiano dà un taglio esponenziale, mentre  $\tilde{W}^{(t.h.)}(x) \propto x^{-2}$ . Per un spettro che, asintoticamente, abbia comportamento  $\delta(k) \propto k^\alpha$ , con  $\alpha > 2$ , il taglio sulle piccole scale operato da un filtro a top-hat è quindi insufficiente. Per il filtro Gaussiano, il rischio c'è solo se lo spettro che diverge esponenzialmente. L'uso del filtro Gaussiano richiede però un'integrazione nello spazio estesa fino  $\infty$ .

Veniamo infine alla prova di eq. (A1.7) con (A1.6). Si parte da un cambio di variabile in eq. (A1.2), in cui si inserisce la trasformata di  $\varepsilon(\mathbf{x})$ :

$$\begin{aligned}\varepsilon_R(\mathbf{x}) &= \int_0^\infty d^3s W_R(s) \varepsilon(\mathbf{x} + \mathbf{s}) = \int_0^\infty d^3s W_R(s) (2\pi)^{-\frac{3}{2}} \int_0^\infty d^3k \delta(\mathbf{k}) e^{i\mathbf{k}\cdot(\mathbf{x}+\mathbf{s})} = \\ &= (2\pi)^{-\frac{3}{2}} \int_0^\infty d^3k \delta(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} \int_0^\infty d^3s W_R(s) e^{i\mathbf{k}\cdot\mathbf{s}} = (2\pi)^{-\frac{3}{2}} \int_0^\infty d^3k \delta_R(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}\end{aligned}$$

avendo posto  $\delta_R(\mathbf{k}) = \delta(\mathbf{k})\tilde{W}(kR)$  con  $\tilde{W}(kR) = \int_0^\infty d^3s W_R(s) e^{i\mathbf{k}\cdot\mathbf{s}}$ . Quindi, le componenti di  $\varepsilon(\mathbf{x})$  sono quelle definite in eq. (A1.6).

## 2. Convenzione di somma sugli indici continui

Un'espressione come (A1.2) si presta a illustrare un modo compatto di esprimere talune integrazioni definite. In tale caso, il dominio di integrazione non è equivoco e conviene considerare

$W_R(|\mathbf{x} - \mathbf{u}|)$  come una «matrice continua» di indici  $\mathbf{x}$  e  $\mathbf{u}$ , scrivendo

$$\boldsymbol{\varepsilon}_x^{(R)} = W_{xu}^{(R)} \boldsymbol{\varepsilon}_u .$$

Qui l'integrazione sull'indice continuo ripetuto è sottintesa, in analogia a quanto si fa per le somme sugli indici discreti.

In generale si opera la convenzione che le notazioni

$$(A1.8) \quad A_x = W_{xy} B_y, \quad A(x) = \int_{\mathcal{D}_n} d^n y W(x, y) B(y)$$

sono tra loro equivalenti, qualunque sia la dimensionalità  $n$  del dominio  $\mathcal{D}_n$ , purché non ci siano equivoci sulla sua definizione.

Questo formalismo si dimostrerà estremamente efficace a compattare molte espressioni nel prossimo paragrafo e in quelli successivi, agevolando la comprensione del significato delle espressioni analitiche.

# Appendice A2

## Distribuzioni stocastiche nel discreto e nel continuo

### 1. Processi e loro realizzazioni.

#### Valor medio, varianza, momenti d'ordine $n$

Consideriamo una variabile  $\varepsilon$  e un **processo** stocastico che distribuisce valori  $\varepsilon_r$  di tale variabile. Un esempio è la «misura» di una grandezza fisica tramite un apparato sperimentale, sia che la grandezza fisica abbia un valore vero (lunghezza di un tavolo) sia che la grandezza fisica possa intrinsecamente variare (tempo  $\tau$  che intercorre tra creazione e decadimento di una particella).

Sulla base di  $N$  valori si definisce il **valor medio** di  $\varepsilon$ ,

$$(A2.1) \quad \langle \varepsilon \rangle = N^{-1} \sum_{r=1}^N \varepsilon_r$$

e la **varianza** rispetto al valor medio

$$(A2.2) \quad \sigma_N^2 = N^{-1} \sum_{r=1}^N [\varepsilon_r - \langle \varepsilon \rangle]^2.$$

$\sigma_N$ , in genere, è la «distanza tipica» da  $\langle \varepsilon \rangle$  cui si situano i valori  $\varepsilon_r$ .

Sulla base di  $N$  valori  $\varepsilon_r$  si può costruire un istogramma: si divide l'asse  $\varepsilon$ , attorno a  $\langle \varepsilon \rangle$  in intervalli di ampiezza

$$(A2.3) \quad \Delta\varepsilon = \sigma_N / N^{1/2} \quad (\text{quindi } \Delta\varepsilon / \sigma_N = 1 / N^{1/2}),$$

così che il numero di intervalli che contengono valori è  $v_i \sim \sigma_N / \Delta\varepsilon = N^{1/2}$ , per cui il numero di valori che ci si attende cadano in ciascun intervallo è  $v_t \sim N / N^{1/2} = N^{1/2}$ ; in definitiva sarà  $v_t \sim v_i$ .

Conviene poi definire la funzione  $p_N(\varepsilon)$  che, nel  $t$ -esimo intervallo, assume il valore  $v_t / \sigma_N N^{1/2}$ . In questo modo  $p_N(\varepsilon)$  è normalizzata a 1; infatti risulta

$$\int_{-\infty}^{\infty} d\varepsilon p_N(\varepsilon) = \frac{\Delta\varepsilon}{\sigma_N N^{1/2}} \cdot \sum_t v_t = \frac{1}{N} \cdot N = 1,$$

indipendentemente da  $N$ .

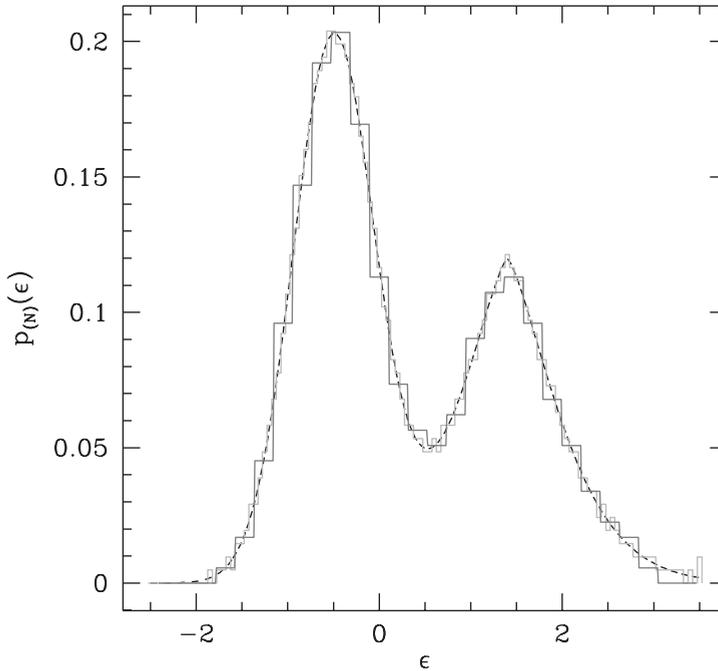
Accrescendo  $N$ , valor medio e varianza possono variare, ma è intuitivo attendersi che oscilleranno in intervalli sempre più ristretti al crescere di  $N$ .

L'assunto, verificato dall'esperienza, è che il **processo**  $p(\varepsilon)$  che distribuisce i valori **coincide col limite cui tende la successione di funzioni**  $p_N(\varepsilon)$ :

$$p(\varepsilon) = \lim_{N \rightarrow \infty} p_N(\varepsilon),$$

come mostrato in figura II.14. Facendo uso delle funzioni  $p_N(\varepsilon)$  si possono dare definizioni alternative di valor medio e varianza:

$$(A2.4) \quad \langle \varepsilon \rangle = \int_{-\infty}^{\infty} d\varepsilon \varepsilon p_N(\varepsilon), \quad \sigma_N^2 = \int_{-\infty}^{\infty} d\varepsilon [\varepsilon - \langle \varepsilon \rangle]^2 p_N(\varepsilon).$$

**FIGURA II.14**

Distribuzioni di 325 e 1600 valori  $\varepsilon$  sulla base del processo  $p(\varepsilon)$  (in tratteggio, parzialmente coperto dagli istogrammi); gli istogrammi dei 325 e 1600 valori sono rispettivamente in grigio 60% e 30%; il secondo istogramma approssima ormai tanto bene il processo da essere quasi indistinguibile alla scala del grafico.

C'è discrepanza tra queste definizioni e le definizioni (A2.1), (A2.2). Discutiamola per il valor medio.

Sulla base dell'eq. (A2.4) e usando la definizione di  $p_N$  si ha:

$$(A2.5) \quad \langle \varepsilon \rangle = \sum_t \frac{v_t}{\sigma_N N^{1/2}} \bar{\varepsilon}_t \Delta \varepsilon = \frac{\Delta \varepsilon}{\sigma_N N^{1/2}} \sum_t v_t \bar{\varepsilon}_t = \frac{1}{N} \sum_t v_t \bar{\varepsilon}_t$$

dove  $\bar{\varepsilon}_t$  è la coordinata del punto centrale del  $t$ -esimo intervallo e, per l'ultimo passaggio, si è utilizzata la relazione (A2.3). Dunque, in ciascun intervallo, invece di prendere le coordinate  $\varepsilon_r$  dei  $v_t$  punti  $\varepsilon_{st}$  che gli appartengono, si è presa  $v_t$  volte la coordinata del punto centrale  $\bar{\varepsilon}_t$ , che sarà, in generale, diverso dalla media  $\langle \varepsilon \rangle_t$  dei valori del  $t$ -esimo intervallo.

La discrepanza tra  $\bar{\varepsilon}_t$  e  $\langle \varepsilon \rangle_t$  sarà dell'ordine di  $\Delta \varepsilon / v_t$ , ma con segni alterni; essa darà luogo, complessivamente, a una discrepanza sul valor medio

$$\Delta \langle \varepsilon \rangle = \frac{1}{N} \sum_t v_t (\langle \varepsilon \rangle_t - \bar{\varepsilon}_t) = \frac{1}{N} \sum_t v_t \frac{\Delta \varepsilon}{v_t^{3/2}} \approx \frac{\Delta \varepsilon}{N} \frac{v_i}{v_t^{1/2}} \approx \frac{\Delta \varepsilon}{N^{1/2}} = \frac{\sigma_N}{N}$$

dove la potenza  $3/2$  su  $v_t$ , al denominatore, tiene conto del fatto che i segni degli addendi, ciascuno dell'ordine di  $\Delta \varepsilon / v_t$ , sono necessariamente alterni. Si è tenuto anche conto del fatto che  $v_t \sim v_i$  e dell'eq. (A2.3).

Al crescere di  $N$ , dunque, la differenza tra le due definizioni di valor medio si riduce e si annulla

quando  $N \rightarrow \infty$ . Un analogo teorema vale per la definizione (A2.4) di varianza. Tali definizioni saranno quelle che useremo d'ora in poi, assieme alla definizione del momento  $n$ -esimo

$$(A2.6) \quad \sigma_N^{(n)n} = \int_{-\infty}^{\infty} d\varepsilon [\varepsilon - \langle \varepsilon \rangle]^n p_N(\varepsilon).$$

La varianza, dunque, coincide con secondo momento.

Ogni  $N$ -upla di valori  $\varepsilon_r$  costituisce una **realizzazione** del processo  $p(\varepsilon)$ . Al crescere di  $N$ , come s'è detto,  $p_N(\varepsilon)$  approssima  $p(\varepsilon)$ . Tuttavia, nessuna realizzazione, in quanto costituita da un insieme finito di valori, può contenere la stessa informazione del processo. Con riferimento al processo si definisce allora il **valor medio**, o valore di aspettazione, di  $\varepsilon$  e i momenti

$$(A2.7) \quad \langle \varepsilon \rangle = \int_{-\infty}^{\infty} d\varepsilon \varepsilon p(\varepsilon), \quad \sigma^{(n)n} = \int_{-\infty}^{\infty} d\varepsilon [\varepsilon - \langle \varepsilon \rangle]^n p(\varepsilon),$$

compreso  $\sigma^{(2)}$ , che è la varianza. Non essendo noto il processo, valor medio e momenti si approssimano sostituendo  $p(\varepsilon)$  con  $p_N(\varepsilon)$ . L'approssimazione sarà tanto migliore, quanto maggiore è  $N$ .

## 2. Distribuzioni Gaussiane

Senza perdita di generalità, con un cambiamento dell'origine sugli assi, si possono sempre rendere nulli i valori medi. Consideriamo allora un processo che distribuisca i valori di una variabile  $\varepsilon$  secondo la legge

$$(A2.8) \quad G(\varepsilon) = \frac{h}{\sqrt{2\pi}} e^{-\frac{1}{2}h^2\varepsilon^2}$$

(con valor medio di  $\varepsilon$  nullo). Questo processo viene detto **Gaussiano** e risulta completamente definito quando è assegnato il «modulo di precisione»  $h$ . Il coefficiente  $h/\sqrt{2\pi}$  garantisce la normalizzazione

$$\int_{-\infty}^{\infty} d\varepsilon G(\varepsilon) = 1.$$

Integrando per parti, otteniamo allora

$$(A2.9) \quad \frac{\sqrt{2\pi}}{h} = \int_{-\infty}^{\infty} d\varepsilon e^{-\frac{1}{2}h^2\varepsilon^2} = \varepsilon e^{-\frac{1}{2}h^2\varepsilon^2} \Big|_{-\infty}^{+\infty} + \int_{-\infty}^{\infty} d\varepsilon \varepsilon h^2 \varepsilon e^{-\frac{1}{2}h^2\varepsilon^2} = h^2 \frac{\sqrt{2\pi}}{h} \sigma^2,$$

sicché  $h\sigma = 1$ : il modulo di precisione è l'inverso della varianza e avremo:

$$G(\varepsilon) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{\varepsilon}{\sigma}\right)^2\right].$$

È inoltre chiaro che i momenti  $\sigma^{(n)}$  con  $n$  dispari sono nulli, per ragioni di simmetria. Integrando ancora per parti si ricava poi:

$$(A2.10) \quad (n+1)[\sigma^{(n)}]^n = [\sigma^{(n+2)}]^{n+2} \quad \text{da cui: } \sigma^{(n)} = [(n-1)!!]^{\frac{1}{n}} \sigma_n^{\frac{2}{n}}.$$

Nota la varianza di una distribuzione Gaussiana, ne conosciamo quindi tutti i momenti.

### 3. Distribuzioni in uno spazio funzionale

Gli argomenti dei paragrafi precedenti si generalizzano al caso di un insieme di  $R$  variabili stocastiche  $\{\varepsilon_i\}$  o  $\{\varepsilon(i)\}$  ( $i = 1, 2, \dots, R$ ). Attenzione: l'indice  $i$  non contrassegna valori di una singola variabile, ma denomina variabili diverse. Così un processo  $P[\varepsilon(i)] \equiv P[\varepsilon_i]$  viene approssimato da un insieme di  $N$   $R$ -uple di valori  $\varepsilon_n(i)$  ( $n = 1, 2, \dots, N$ ). Con tali  $R$ -uple si può costruire un istogramma nello spazio  $R$ -dimensionale di coordinate  $\varepsilon_i$ .

Un processo  $P[\varepsilon_i]$  è Gaussiano se lo è su ciascuna variabile  $\varepsilon_i$ :

$$(A2.11) \quad G[\varepsilon_i] = \frac{\sqrt{\text{Tr } h_{ij}}}{(2\pi)^{n/2}} e^{-\frac{1}{2}\varepsilon_i h_{ij} \varepsilon_j}$$

(si somma sugli indici ripetuti). Qui  $h_{ij}$  è una matrice puramente diagonale e, ponendo  $h_{ii} = h_i^2$ ,  $G[\varepsilon_i]$  ridiviene esplicitamente un prodotto di distribuzioni Gaussiane sulle singole variabili.

La natura del processo non cambia se effettuiamo una «rotazione» di assi nello spazio  $R$ -dimensionale delle  $\varepsilon_i$ . In questo modo otteniamo la più generale distribuzione Gaussiana in tale spazio, che ha forma apparentemente identica a (A2.11), salvo che le nuove  $\varepsilon_j$  non coincidono con le precedenti, ma sono le ascisse dei nuovi assi, mentre  $h_{jk}$  non è più una matrice puramente diagonale. La trasformazione ortogonale non ne ha tuttavia modificato la traccia.

La distribuzione soddisfa chiaramente la condizione di normalizzazione, in quanto la matrice  $h_{ij}$  potrà sempre essere diagonalizzata:

$$(A2.12) \quad \int_{-\infty}^{\infty} d\varepsilon_1 \int_{-\infty}^{\infty} d\varepsilon_2 \dots \int_{-\infty}^{\infty} d\varepsilon_R G[\varepsilon_i] = \prod_{j=1}^R \left\{ \int_{-\infty}^{\infty} d\varepsilon_j \right\} G[\varepsilon_i] = 1.$$

Le precedenti definizioni possono essere ora generalizzate al caso di una successione di variabili  $\{\varepsilon_i\}$  ( $i = 1, 2, \dots, R, \dots$ ). Una distribuzione su tali infinite variabili si dirà Gaussiana se avrà la forma

$$(A2.13) \quad G[\varepsilon_i] = \mathcal{N} e^{-\frac{1}{2}\varepsilon_j h_{jk} \varepsilon_k},$$

in cui tutte le somme sono diventate serie, mentre il coefficiente  $\mathcal{N}$  non viene esplicitato, ma deve essere tale che

$$(A2.14) \quad \prod_{j=1}^{\infty} \left\{ \int_{-\infty}^{\infty} d\varepsilon_j \right\} G[\varepsilon_i] = 1.$$

Oltre alle distribuzioni Gaussiane su 1,  $R$  o  $\infty$  variabili, ci sono anche distribuzioni **non-Gaussiane**, che non hanno la forma (A2.13), ma si assumono comunque normalizzate e quindi soddisfano anch'esse la condizione (A2.14).

La scrittura dell'operazione di integrazione si semplifica definendo

$$(A2.15) \quad \int \mathcal{D}\varepsilon \equiv \prod_{j=1}^{\infty} \left\{ \int_{-\infty}^{\infty} d\varepsilon_j \right\}, \quad \text{per cui } \int \mathcal{D}\varepsilon G[\varepsilon_i] = 1$$

indica la stessa operazione di eq. (A2.14).

Un'ulteriore generalizzazione può essere effettuata se, invece di un insieme numerabile di variabili  $\varepsilon(i)$ , consideriamo un insieme continuo di variabili  $\varepsilon(\mathbf{x})$ , «catalogate» dall'indice continuo  $\mathbf{x}$ . Avremo così definito una distribuzione sul corpo delle funzioni  $\varepsilon(\mathbf{x})$ , una **distribuzione in uno spazio funzionale**.

Una distribuzione Gaussiana, in uno spazio funzionale, ha la forma

$$(A2.16) \quad G[\varepsilon(\mathbf{x})] = \mathcal{N} \exp \left\{ -\frac{1}{2} \int d^3x_1 \int d^3x_2 \varepsilon(x_1) h(x_1, x_2) \varepsilon(x_2) \right\}$$

e si scrive più semplicemente usando la convenzione introdotta nell'ultimo paragrafo di appendice A1, analoga alla convenzione di somma sugli indici ripetuti:

$$(A2.17) \quad G[\varepsilon_{\mathbf{x}}] = \mathcal{N} e^{-\frac{1}{2} \varepsilon_{\mathbf{x}_1} h_{\mathbf{x}_1 \mathbf{x}_2} \varepsilon_{\mathbf{x}_2}}.$$

La normalizzazione di questa o di altra distribuzione, anche non-Gaussiana, nello spazio funzionale, viene ora indicata

$$(A2.18) \quad \int \mathcal{D}\varepsilon G[\varepsilon_{\mathbf{x}}] = 1;$$

l'integrazione si definisce passando al limite dal discreto al continuo.

La possibilità di calcolare integrali in uno spazio funzionale appare remota. Tuttavia, come vedremo tra breve, vi sono casi significativi in cui essa sussiste.

#### 4. Valore di aspettazione, momenti, funzionale di partizione

In eq. (A2.7) si sono definiti valore di aspettazione e momenti di una singola variabile, assegnato il processo con cui essa è distribuita. Se il processo distribuisce una variabile  $\varepsilon_{\bar{k}}$  e, simultaneamente, anche altre variabili  $\varepsilon_j$  ( $j \neq \bar{k}$ ), si deve integrare su tutte le variabili  $\varepsilon_j$ , incluso ovviamente  $j = \bar{k}$ , per avere valore di aspettazione e momenti della variabile  $\varepsilon_{\bar{k}}$ :

$$(A2.19) \quad \langle \varepsilon_{\bar{k}} \rangle = \left\{ \prod_j \int d\varepsilon_j \right\} p(\varepsilon_j) \varepsilon_{\bar{k}}, \quad \sigma_{\bar{k}}^{(n)} = \left\{ \prod_i \int d\varepsilon_i \right\} p(\varepsilon_i) [\varepsilon_{\bar{k}} - \langle \varepsilon_{\bar{k}} \rangle]^n,$$

Tale definizione si estende al continuo:

$$(A2.20) \quad \langle \varepsilon_{\bar{\mathbf{x}}} \rangle = \int \mathcal{D}\varepsilon p(\varepsilon_{\mathbf{x}}) \varepsilon_{\bar{\mathbf{x}}}, \quad \sigma_{\bar{\mathbf{x}}}^{(n)} = \int \mathcal{D}\varepsilon p(\varepsilon_{\mathbf{x}}) [\varepsilon_{\bar{\mathbf{x}}} - \langle \varepsilon_{\bar{\mathbf{x}}} \rangle]^n,$$

Si considera anche una sorta di «trasformata di Fourier» del processo:

$$(A2.21) \quad Z[J_{\mathbf{x}}] = \int \mathcal{D}\varepsilon p(\varepsilon_{\mathbf{y}}) e^{iJ_{\mathbf{x}}\varepsilon_{\mathbf{x}}}$$

definita «funzionale di partizione» (ricordiamo:  $J_{\mathbf{x}}\varepsilon_{\mathbf{x}} = \int d^3x J(\mathbf{x})\varepsilon(\mathbf{x})$ ).

#### 5. Varianza di massa

Assumiamo ora che  $\varepsilon(\mathbf{x})$  rappresenti proprio il campo di fluttuazioni di densità nel cosmo. Detto  $p(\varepsilon_{\mathbf{x}})$  il processo che distribuisce le fluttuazioni nel cosmo, la varianza di massa sarà

$$(A2.22) \quad \sigma_R^2 = \langle \varepsilon_R^2 \rangle = \int \mathcal{D}\varepsilon p(\varepsilon_{\mathbf{x}}) [\varepsilon_{R\bar{\mathbf{x}}}]^2.$$

Concettualmente, invece di mediare su volumi di dimensioni  $R$  e centri  $\mathbf{x}$ , distribuiti in un grande volume  $\mathcal{V}$  in cui si trovano molte realizzazioni del processo  $p(\varepsilon_{\mathbf{x}})$ , qui si fissa un punto ( $\bar{\mathbf{x}}$ ) e si

media su valori di  $\varepsilon$  che il processo genera in un intorno di raggio  $R$  di quel punto. L'equivalenza delle due operazioni è basata sull'assunto che il processo che crea le disomogeneità sia... omogeneo, eguale in ogni punto del cosmo. Perciò, l'integrale (A2.22), che formalmente dipende da  $\bar{\mathbf{x}}$ , deve risultare indipendente dalla scelta del punto.

In eq. (A2.22) portiamo il filtro fuori dall'integrazione funzionale; avremo:

$$(A2.23) \quad \sigma_R^2 = \int \mathcal{D}\varepsilon p(\varepsilon_{\mathbf{x}}) \left\{ W_{\mathbf{x}\mathbf{y}}^{(R)} \varepsilon_{\mathbf{y}} \right\}^2 = \int d^3y W_{\mathbf{x}\mathbf{y}}^{(R)} \int d^3u W_{\mathbf{x}\mathbf{u}}^{(R)} \int \mathcal{D}\varepsilon p(\varepsilon_{\mathbf{x}}) \varepsilon_{\mathbf{y}} \varepsilon_{\mathbf{u}} .$$

Questa espressione è importante per la connessione che permette di ricavare con la funzione a 2 punti.

## 6. Funzione di correlazione a 2 punti

Riportiamo qui l'espressione (II.42) della funzione a due punti nella forma

$$(A2.24) \quad \zeta(r_{12}) = \frac{\langle n(\mathbf{x}_1)n(\mathbf{x}_2) \rangle}{n^2} - 1 \quad \text{in cui} \quad n(\mathbf{x}) = \frac{\delta N}{\delta V}(\mathbf{x})$$

è la densità (in numero) di galassie nel punto  $\mathbf{x}$ . Ricordiamo di aver assunto una massa costante ( $m_g$ ) per le galassie, per cui la densità (in massa) è  $\rho(\mathbf{x}) = m_g n(\mathbf{x})$ . A partire da eq. (A2.24) è allora facile verificare l'espressione

$$(A2.25) \quad \zeta(r_{12}) \equiv \zeta_{\mathbf{x}_1, \mathbf{x}_2} = \langle \varepsilon(\mathbf{x}_1) \varepsilon(\mathbf{x}_2) \rangle,$$

che può essere valutata sulle realizzazioni del processo che genera le disomogeneità, ma che vale

$$(A2.26a) \quad \zeta(r_{12}) = \int \mathcal{D}\varepsilon p(\varepsilon) \varepsilon(\mathbf{x}_1) \varepsilon(\mathbf{x}_2)$$

in termini del processo stesso. Quest'equazione si scrive anche

$$(A2.26b) \quad \zeta_{\mathbf{x}_1, \mathbf{x}_2} = \int \mathcal{D}\varepsilon p(\varepsilon) \varepsilon_{\mathbf{x}_1} \varepsilon_{\mathbf{x}_2} .$$

La preannunziata verifica si effettua a partire da eq. (A.27), che dà

$$\begin{aligned} \zeta(r_{12}) &= \frac{\langle [\rho(\mathbf{x}_1) - \rho][\rho(\mathbf{x}_2) - \rho] \rangle}{\rho^2} = \frac{\langle [n(\mathbf{x}_1) - n][n(\mathbf{x}_2) - n] \rangle}{n^2} = \\ &= \frac{\langle n(\mathbf{x}_1)n(\mathbf{x}_2) \rangle - n\langle n(\mathbf{x}_2) \rangle - \langle n(\mathbf{x}_1) \rangle n + n^2}{n^2} = \frac{\langle n(\mathbf{x}_1)n(\mathbf{x}_2) \rangle - n^2}{n^2} \end{aligned}$$

in quanto, qualsiasi sia il nome della variabile, risulta  $\langle n(\mathbf{x}_1) \rangle = \langle n(\mathbf{x}_2) \rangle = n$ .

## 7. Funzione di correlazione a 2 punti e varianza di massa

Confrontiamo ora le eq. (A2.23) e (A2.26); si vede che deve essere

$$(A2.27) \quad \sigma_R^2 = \int d^3y W_{\mathbf{x}\mathbf{y}}^{(R)} \int d^3u W_{\mathbf{x}\mathbf{u}}^{(R)} \zeta_{\mathbf{y}\mathbf{u}} .$$

Usiamo qui un filtro a top-hat e l'espressione  $\zeta(r) = (r_o/r)^\gamma$ ; otterremo

$$\begin{aligned}
 \sigma_R^2 &= \int_0^\infty d^3y W^{(R)}(|\mathbf{x} - \mathbf{y}|) \int_0^\infty d^3u W^{(R)}(|\mathbf{x} - \mathbf{u}|) \left( \frac{r_o}{|\mathbf{y} - \mathbf{u}|} \right)^\gamma = \\
 \text{(A2.28)} \quad &= V_R^{-2} \int_0^R d^3y \int_0^R d^3u \left( \frac{r_o}{|\mathbf{y} - \mathbf{u}|} \right)^\gamma = \\
 &= \left( \frac{3}{4\pi} \right)^2 \left( \frac{r_o}{R} \right)^\gamma \int_0^1 d^3Y \int_0^1 d^3U \left( \frac{1}{|\mathbf{Y} - \mathbf{U}|} \right)^\gamma = G(\gamma) \left( \frac{r_o}{R} \right)^\gamma
 \end{aligned}$$

dove si sono operati i cambi di variabile  $Y = y/R$ ,  $U = u/R$  e si è posto

$$\text{(A2.29)} \quad G(\gamma) = \left( \frac{3}{4\pi} \right)^2 \int_0^1 d^3Y \int_0^1 d^3U |\mathbf{Y} - \mathbf{U}|^{-\gamma}.$$

Se ne conclude che

$$\text{(A2.30)} \quad \sigma_R^2 = G(\gamma) \zeta(R)$$

almeno finché  $\zeta(r)$  è una legge di potenza. Per valori e andamento dettagliato di  $G(\gamma)$  (comunque  $G(\gamma) \sim 1.6-2$ ) vedi il testo del capitolo.

## 8. Funzione a 2 punti per processi Gaussiani

Si può provare che la funzione a 2 punti si ricava dal funzionale di partizione con un'operazione di derivazione:

$$\text{(A2.31)} \quad \zeta_{\mathbf{x}_1 \mathbf{x}_2} = (-i)^2 \frac{\delta}{\delta J_{\mathbf{x}_1}} \frac{\delta}{\delta J_{\mathbf{x}_2}} Z[J_{\mathbf{x}}] \Big|_{J_{\mathbf{x}}=0}.$$

$J_{\mathbf{x}_1}$  e  $J_{\mathbf{x}_2}$  sono due funzioni dello spazio funzionale delle  $J_{\mathbf{x}}$ ; le due derivate considerate sono quindi l'analogo funzionale delle derivate parziali rispetto a due coordinate in uno spazio di punti.

In uno spazio  $\infty$ -dimensionale, ma con  $\infty$  discreto, ogni punto sarebbe contrassegnato da una successione di valori delle coordinate  $J_r$  ( $r = 1, 2, \dots, n, \dots$ ) e il senso della derivazione rispetto a una particolare coordinata  $J_{\vec{r}}$  sarebbe quello usuale. In eq. (A2.31) si considera la stessa operazione al limite continuo. Tutto ciò risulterà più chiaro seguendo i passaggi qui sotto.

Proviamo ad esempio l'eq. (A2.31). Sarà

$$\begin{aligned}
 \frac{\delta}{\delta J_{\mathbf{x}_1}} \frac{\delta}{\delta J_{\mathbf{x}_2}} Z[J_{\mathbf{x}}] &= \frac{\delta}{\delta J_{\mathbf{x}_1}} \frac{\delta}{\delta J_{\mathbf{x}_2}} \int \mathcal{D}\varepsilon p(\varepsilon_{\mathbf{x}}) e^{iJ_{\mathbf{x}}\varepsilon_{\mathbf{x}}} = \int \mathcal{D}\varepsilon p(\varepsilon_{\mathbf{x}}) \frac{\delta}{\delta J_{\mathbf{x}_1}} \frac{\delta}{\delta J_{\mathbf{x}_2}} e^{iJ_{\mathbf{x}}\varepsilon_{\mathbf{x}}} = \\
 &= \int \mathcal{D}\varepsilon p(\varepsilon_{\mathbf{x}}) \frac{\delta}{\delta J_{\mathbf{x}_1}} i\varepsilon_{\mathbf{x}_2} e^{iJ_{\mathbf{x}}\varepsilon_{\mathbf{x}}} = \int \mathcal{D}\varepsilon p(\varepsilon_{\mathbf{x}}) i\varepsilon_{\mathbf{x}_1} i\varepsilon_{\mathbf{x}_2} e^{iJ_{\mathbf{x}}\varepsilon_{\mathbf{x}}}
 \end{aligned}$$

Se  $J_{\mathbf{x}} \equiv 0$  sarà  $e^{iJ_{\mathbf{x}}\varepsilon_{\mathbf{x}}} = 1$ , e ricordando la definizione (A2.26) l'eq. (A2.31) è provata.

Nel caso di matrici  $N$ -dimensionali, se la matrice  $A_{rs}$  ha determinante non nullo, l'inversa  $I = A^{-1}$  è la matrice  $I_{uv}$  tale che  $\sum_b I_{ab} A_{bc} = \delta_{ac}$ .

Questa notazione si estende subito al caso di  $\infty$ -discrete dimensioni. La notazione è la medesima, le differenze sono che la somma sugli indici si estende da 1 a  $\infty$ , e che le condizioni per l'esistenza dell'inversa sono più complesse.

La funzione a 2 punti  $\zeta_{\mathbf{x}_1, \mathbf{x}_2}$  può essere vista come matrice quadrata a  $\infty$ -continue componenti. È possibile che essa abbia una matrice inversa tale che

$$(A2.32) \quad \int d^3y I(\mathbf{x}, \mathbf{y}) \zeta(\mathbf{y}, \mathbf{z}) = I_{\mathbf{xy}} \zeta_{\mathbf{yz}} = \delta^{(3)}(\mathbf{x} - \mathbf{z}).$$

Non sempre esiste l'inversa di una matrice continua. Una matrice che soddisfi eq. (A2.32) si indica  $I(\mathbf{x}_1, \mathbf{x}_2) = \zeta^{-1}(\mathbf{x}_1, \mathbf{x}_2)$ .

Consideriamo ora un processo Gaussiano

$$G[\varepsilon_{\mathbf{x}}] = \mathcal{N} e^{-\frac{1}{2} \varepsilon_{\mathbf{x}} h_{\mathbf{xy}} \varepsilon_{\mathbf{y}}}.$$

Si dimostra che la funzione a 2 punti della distribuzione generata dal processo è l'inverso della matrice  $h_{\mathbf{xy}}$ :

$$(A2.33) \quad \zeta_{\mathbf{yz}} = h_{\mathbf{xy}}^{-1}.$$

In questo caso, infatti, l'integrale di partizione ha la forma

$$(A2.34) \quad Z(J_{\mathbf{x}}) = \int \mathcal{D}\varepsilon \mathcal{N} e^{-\frac{1}{2} \varepsilon_{\mathbf{x}} h_{\mathbf{xy}} \varepsilon_{\mathbf{y}} + i J_{\mathbf{x}} \varepsilon_{\mathbf{x}}}$$

e si può operare un cambio di variabile d'integrazione che semplifica in modo decisivo la quantità all'esponente. La nuova variabile d'integrazione sarà

$$(A2.35) \quad \tilde{\varepsilon}_x = \varepsilon_x - i h_{xx'}^{-1} J_{x'}$$

dove  $h_{xx'}$  è simmetrica nello scambio degli indici (pur rappresentando vettori, per semplicità gli indici non sono qui indicati in grassetto). Siccome la differenza  $\tilde{\varepsilon} - \varepsilon$  non dipende da  $\varepsilon$ , sarà  $\mathcal{D}\tilde{\varepsilon} = \mathcal{D}\varepsilon$ . Inoltre:

$$\begin{aligned} -\frac{1}{2} \varepsilon_x h_{xy} \varepsilon_y + i J_x \varepsilon_x &= -\frac{1}{2} (\tilde{\varepsilon}_x + i h_{xx'}^{-1} J_{x'}) h_{xy} (\tilde{\varepsilon}_y + i h_{yy'}^{-1} J_{y'}) + i J_x (\tilde{\varepsilon}_x + i h_{xx'}^{-1} J_{x'}) = \\ &= -\frac{1}{2} (\tilde{\varepsilon}_x h_{xy} \tilde{\varepsilon}_y + i J_{x'} h_{xx'}^{-1} h_{xy} \tilde{\varepsilon}_y + i \tilde{\varepsilon}_x h_{xy} h_{yy'}^{-1} J_{y'} - h_{xx'}^{-1} J_{x'} h_{xy} h_{yy'}^{-1} J_{y'}) + \\ &+ i J_x \tilde{\varepsilon}_x - J_x h_{xx'}^{-1} J_{x'}. \end{aligned}$$

Essendo  $h_{xx'}$  una matrice simmetrica, sarà allora

$$\begin{aligned} -\frac{1}{2} \varepsilon_x h_{xy} \varepsilon_y + i J_x \varepsilon_x &= -\frac{1}{2} (\tilde{\varepsilon}_x h_{xy} \tilde{\varepsilon}_y + 2 i J_{x'} \tilde{\varepsilon}_x + -J_{x'} h_{xx'}^{-1} J_{x'}) + i J_x \tilde{\varepsilon}_x - J_x h_{xx'}^{-1} J_{x'} = \\ &= -\frac{1}{2} \tilde{\varepsilon}_x h_{xy} \tilde{\varepsilon}_y - \frac{1}{2} J_{x'} h_{xx'}^{-1} J_{x'}. \end{aligned}$$

e quindi il funzionale di partizione è

$$(A2.36) \quad Z(J_{\mathbf{x}}) = e^{-\frac{1}{2} J_{x'} h_{xx'} J_{x'}} \int \mathcal{D}\tilde{\varepsilon} \mathcal{N} e^{-\frac{1}{2} \tilde{\varepsilon}_x h_{xy} \tilde{\varepsilon}_y} = e^{-\frac{1}{2} J_{x'} h_{xx'}^{-1} J_{x'}};$$

l'integrale vale 1, in quanto la distribuzione Gaussiana è normalizzata.

Si noti che, in questo caso, si è riusciti a «calcolare» l'integrale funzionale: si è partiti da un'espressione di  $Z$  in cui compariva un tale integrale e si è arrivati a un'espressione in cui ci sono solo integrazioni ordinarie (sugli indici ripetuti, all'esponente).

Usiamo ora questa espressione di  $Z$  per ricavare la funzione a 2 punti, conformemente all'eq. (A2.30):

$$\zeta_{x_1 x_2} = - \frac{\delta}{\delta J_{x_1}} \frac{\delta}{\delta J_{x_2}} e^{-\frac{1}{2} J_{x'} h_{x'x}^{-1} J_x} \Big|_{J_x \equiv 0} .$$

Riportiamo alcuni passaggi, per mostrare una procedura che sarà ancora seguita in altri casi:

$$\begin{aligned} \text{(A2.37)} \quad - \frac{\delta}{\delta J_{x_2}} e^{-\frac{1}{2} J_{x'} h_{x'x}^{-1} J_x} &= \frac{1}{2} \left( h_{x_2 x}^{-1} J_x + J_{x'} h_{x'x_2}^{-1} \right) e^{-\frac{1}{2} J_{x'} h_{x'x}^{-1} J_x} = h_{x_2 x}^{-1} J_x e^{-\frac{1}{2} J_{x'} h_{x'x}^{-1} J_x} , \\ - \frac{\delta}{\delta J_{x_1}} \frac{\delta}{\delta J_{x_2}} e^{-\frac{1}{2} J_{x'} h_{x'x}^{-1} J_x} &= \frac{\delta}{\delta J_{x_1}} \left( h_{x_2 x} J_x e^{-\frac{1}{2} J_{x'} h_{x'x}^{-1} J_x} \right) = \\ &= h_{x_2 x_1}^{-1} e^{-\frac{1}{2} J_{x'} h_{x'x}^{-1} J_x} + \left( h_{x_2 x}^{-1} J_x h_{x_1 x'}^{-1} J_{x'} \right) e^{-\frac{1}{2} J_{x'} h_{x'x}^{-1} J_x} . \end{aligned}$$

Quindi, una volta posto  $J_x \equiv 0$ , risulta  $\zeta_{x_1 x_2} = h_{x_1 x_2}^{-1}$ , come indicato in eq. (A2.33). La conclusione è che ogni processo Gaussiano avrà espressione

$$\text{(A2.38)} \quad G[\varepsilon_x] = \mathcal{N} e^{-\frac{1}{2} \varepsilon_x \zeta_{xy}^{-1} \varepsilon_y} ,$$

sicché la matrice «modulo di precisione» è l'inverso della matrice «funzione di correlazione a 2 punti».

# Appendice A3

## Le funzioni a $n$ -punti

### 1. La funzione a 4 punti

Quando si considerano funzioni di correlazione a  $n$  punti con  $n > 3$  compaiono una serie di complicazioni di varia natura, che discuteremo in questa appendice. Mostreremo anche la relazione tra le funzioni a  $n$  punti e il funzionale di partizione  $Z[J_x]$  definito in appendice A2.

Molte delle citate complicazioni già compaiono nel caso della funzione a 4 punti. Allo stesso modo della funzione a 3 punti, quella a 4 punti è legata alla probabilità di avere simultaneamente 4 galassie in 4 volumetti  $\delta V_i$  ( $i = 1, 2, 3, 4$ ), probabilità data dal valore di aspettazione

$$(A3.1) \quad \langle \delta N_1 \delta N_2 \delta N_3 \delta N_4 \rangle = n_g^4 \delta V_1 \delta V_2 \delta V_3 \delta V_4 [ 1 + \\ + \zeta_{12} + \zeta_{13} + \zeta_{14} + \zeta_{23} + \zeta_{24} + \zeta_{34} + \zeta_{123}^{(3)} + \zeta_{124}^{(3)} + \zeta_{134}^{(3)} + \zeta_{234}^{(3)} + \\ + \zeta_{12} \zeta_{34} + \zeta_{13} \zeta_{24} + \zeta_{14} \zeta_{23} + \zeta^{(4)}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) ].$$

Si sono qui introdotte notazioni frequentemente in uso, quando non sussiste il rischio di confusione:

$$(A3.2) \quad \zeta(r_{ij}) \equiv \zeta_{ij}, \quad \zeta^{(3)}(r_{ij}, r_{jk}, r_{ki}) \equiv \zeta_{ijk}^{(3)}, \quad \text{analogamente: } \zeta^{(4)}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k, \mathbf{x}_l) \equiv \zeta_{ijkl}^{(4)}$$

I termini nella seconda riga di eq. (A3.1) sono contributi dati da correlazioni a 2 e 3 punti. Anche nella terza riga di eq. (A3.1) compaiono contributi dati da combinazioni di funzioni a 2 punti, che però connettono tutti i 4 punti considerati. Se, oltre tali contributi, esiste un sovrappiù di probabilità in ciascun punto, dovuta non alla semplice presenza di galassie in alcuni o tutti gli altri 4 punti, ma alla loro *simultanea* presenza, esso viene espresso dalla funzione (di correlazione) *ridotta* a 4 punti  $\zeta_{1234}^{(4)}$ . L'insieme dei termini dell'ultima riga, che connettono tutti i punti, costituiscono la funzione (di correlazione) *connessa*, ma *riducibile*, a 4 punti.

Sulla stessa linea si può procedere per la definizione di funzioni a un numero  $n$ , qualsiasi, di punti.

### 2. Espressioni alternative delle funzioni a 3 e 4 punti

Riscriviamo anzitutto l'espressione (II.54)

$$\langle \delta N_1 \delta N_2 \delta N_3 \rangle = n_g^3 \delta V_1 \delta V_2 \delta V_3 [ 1 + \zeta(r_{12}) + \zeta(r_{23}) + \zeta(r_{31}) + \zeta^{(3)}(r_{12}, r_{23}, r_{31}) ],$$

facendo uso delle densità locali (in numero)  $n(\mathbf{x}_i) \equiv n_i = \delta N_i / \delta V_i$  ( $i = 1, 2, 3$ ). Esprimendo le funzioni a 2 punti tramite l'espressione (A2.25), avremo

$$\left\langle \frac{n_1 n_2 n_3}{n_g n_g n_g} \right\rangle = 1 + \zeta_{123}^{(3)} + \left\langle \frac{n_1 n_2}{n_g n_g} \right\rangle + \left\langle \frac{n_2 n_3}{n_g n_g} \right\rangle + \left\langle \frac{n_3 n_1}{n_g n_g} \right\rangle - 3$$

Consideriamo ora il valore di aspettazione

$$\begin{aligned} \langle \varepsilon_1 \varepsilon_2 \varepsilon_3 \rangle &= \left\langle \left( \frac{n_1}{n_g} - 1 \right) \left( \frac{n_2}{n_g} - 1 \right) \left( \frac{n_3}{n_g} - 1 \right) \right\rangle = \\ &= \left\langle \frac{n_1 n_2 n_3}{n_g n_g n_g} \right\rangle - \left\langle \frac{n_1 n_2}{n_g n_g} \right\rangle - \left\langle \frac{n_2 n_3}{n_g n_g} \right\rangle - \left\langle \frac{n_3 n_1}{n_g n_g} \right\rangle + 2 ; \end{aligned}$$

questa relazione si riordina nella forma

$$\left\langle \frac{n_1 n_2 n_3}{n_g n_g n_g} \right\rangle = 1 + \langle \varepsilon_1 \varepsilon_2 \varepsilon_3 \rangle + \left\langle \frac{n_1 n_2}{n_g n_g} \right\rangle + \left\langle \frac{n_2 n_3}{n_g n_g} \right\rangle + \left\langle \frac{n_3 n_1}{n_g n_g} \right\rangle - 3 ,$$

mostrando chiaramente che il valore di aspettazione del prodotto delle fluttuazioni nei 3 punti è la funzione a 3 punti:

$$(A3.3) \quad \langle \varepsilon_1 \varepsilon_2 \varepsilon_3 \rangle = \zeta_{123}^{(3)} .$$

Nel caso della funzione a 4 punti, troveremo un risultato leggermente diverso:

$$(A3.4) \quad \langle \varepsilon_1 \varepsilon_2 \varepsilon_3 \varepsilon_4 \rangle = \zeta_{1234}^{(4)} + \zeta_{12} \zeta_{34} + \zeta_{13} \zeta_{24} + \zeta_{14} \zeta_{23}$$

e il valore di aspettazione del prodotto delle fluttuazioni nei 4 punti sarà la funzione a 4 punti *riducibile*.

La prova segue la stessa linea di quanto visto per la funzione a 3 punti, con qualche ovvio appesantimento. Le complicazioni diventano maggiori salendo con l'ordine della funzione. Si può comunque provare che, a tutti gli ordini, **il valore di aspettazione del prodotto delle fluttuazioni in n punti è la funzione a n punti riducibile**.

Riportiamo la verifica per il caso a 4 punti. Per semplicità di scrittura useremo qui l'abbreviazione

$$\left\langle \frac{n_i}{n_g} \dots \frac{n_j}{n_g} \right\rangle \equiv [i \dots j]$$

L'espressione (A3.1) assume allora la forma

$$\begin{aligned} 1 + \zeta_{1234}^{(4)} + \zeta_{12} \zeta_{34} + \zeta_{13} \zeta_{24} + \zeta_{14} \zeta_{23} &= \\ &= 10 + [1234] - [12] - [13] - [14] - [23] - [24] - [34] - [123] + [12] + [13] + [23] - \\ &- [124] + [12] + [14] + [24] - [134] + [13] + [14] + [34] - [234] + [23] + [24] + [34] = \\ &= 10 + [1234] - [123] - [124] - [134] - [234] + [12] + [13] + [14] + [23] + [24] + [34] , \end{aligned}$$

da confrontare all'espressione

$$\begin{aligned} \langle \varepsilon_1 \varepsilon_2 \varepsilon_3 \varepsilon_4 \rangle &= \left\langle \left( \frac{n_1}{n_g} - 1 \right) \left( \frac{n_2}{n_g} - 1 \right) \left( \frac{n_3}{n_g} - 1 \right) \left( \frac{n_4}{n_g} - 1 \right) \right\rangle = \\ &= 9 + [1234] - [123] - [124] - [134] - [234] + [12] + [13] + [14] + [23] + [24] + [34] , \end{aligned}$$

verificando così l'eq. (A3.4).

### 3. Espressioni del modello gerarchico

Abbiamo visto nel testo del capitolo l'espressione del modello gerarchico per la funzione a 3 punti,

$$\zeta_{123}^{(3)} = Q [\zeta_{12}\zeta_{23} + \zeta_{23}\zeta_{31} + \zeta_{31}\zeta_{12}]$$

in cui  $Q \sim 1.2 \pm 0.1$ . Di questa espressione, in figura II.15, si dà una rappresentazione grafica; analoghe rappresentazioni sono particolarmente utili per funzioni di ordine più elevato.

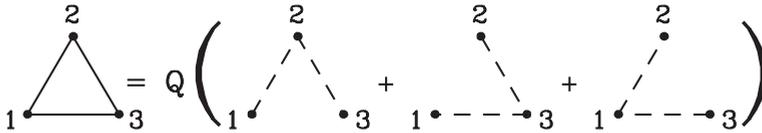


FIGURA II.15

Funzione di correlazione a 3 punti. I grafi rappresentano gli indici di  $\zeta_{ijk}^{(3)}$  (linee compatta) e delle  $\zeta_{ij}$  tra cui si eseguono i prodotti (linee tratteggiate).

Nel caso della funzione a 4 punti, l'espressione del modello gerarchico si complica, per il presentarsi di diverse possibili topologie, ma si costituisce comunque di prodotti di 3 funzioni a 2 punti, completamente simmetrizzate rispetto ai 4 punti considerati:

$$\zeta_{1234}^{(4)} = Q_a [\zeta_{41}\zeta_{12}\zeta_{23} + \zeta_{31}\zeta_{12}\zeta_{24} + \text{term. simm. con 2 indici ripetuti: } 1-3 \ 1-4 \ 2-3 \ 2-4 \ 3-4] + Q_b [\zeta_{12}\zeta_{13}\zeta_{14} + \text{term. simm. con ripetuto 3 volte l'indice: } 2 \ 3 \ 4]$$

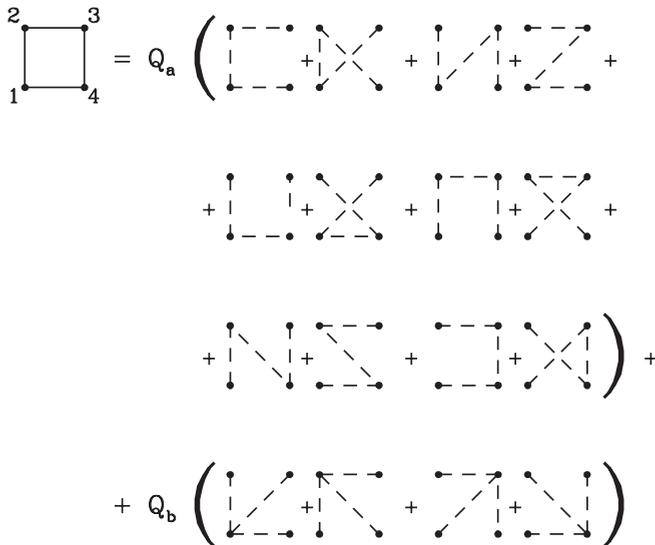


FIGURA II.16

Funzione di correlazione a 4 punti. Nel grafo a linea compatta, che rappresenta  $\zeta_{ijkl}^{(4)}$ , sono riportati i numeri associati ai punti considerati; nei grafi a linea tratteggiata, che rappresentano le  $\zeta_{ij}$  tra cui si esegue un prodotto, i numeri non sono ripetuti.

La novità è che qui intervengono grafi di due possibili topologie. Nel primo tipo di grafi abbiamo 2 punti che sono ripetuti 2 volte. Nel secondo tipo di grafi 1 punto è ripetuto 3 volte. La rappresentazione grafica di questa espressione è data in figura II.16.

I coefficienti  $Q_a$  e  $Q_b$  sono noti con minore precisione di  $Q$ , ma si ritengono ancora compatibili con l'unità,

Per le funzioni a  $n$  punti ( $n > 4$ ) valgono simili espressioni. Al crescere di  $n$  aumenta anche il numero di topologie possibili.

#### 4. Funzioni a $n$ punti per processi Gaussiani e non Gaussiani

Le espressioni (A3.3) e (A3.4) permettono di mostrare la relazione tra funzioni a 3 o 4 punti e funzionale di partizione. In stretta analogia con le eq. (A2.25)-(A2.29), si può provare, con passaggi simili, che risulta

$$(A3.5) \quad \langle \varepsilon_{x_1} \dots \varepsilon_{x_n} \rangle = (-i)^n \frac{\delta}{\delta J_{x_1}} \dots \frac{\delta}{\delta J_{x_n}} Z[J_x] \Big|_{J_x=0}.$$

Mostriamo inoltre che, per un processo Gaussiano, per  $n = 3$  e 4, si ha

$$(A3.6) \quad \langle \varepsilon_{x_1} \varepsilon_{x_2} \varepsilon_{x_3} \rangle = 0, \quad \langle \varepsilon_{x_1} \dots \varepsilon_{x_4} \rangle = \zeta_{x_1, x_2} \zeta_{x_3, x_4} + \zeta_{x_1, x_3} \zeta_{x_2, x_4} + \zeta_{x_1, x_4} \zeta_{x_2, x_3}.$$

Ciò significa che, nelle funzioni a 3 e 4 punti ricavate da distribuzioni Gaussiane, ci sono solo i termini riducibili. Ciò si può generalizzare a ogni ordine: **per processi Gaussiani, le funzioni ridotte a  $n$  punti sono nulle.**

La verifica riguarderà solo le funzioni a 3 e 4 punti e parte dall'espressione (A2.38). Essa si può derivare ulteriormente, prima rispetto  $\varepsilon_{x_3}$ , poi rispetto  $\varepsilon_{x_4}$ . [C'è simmetria tra gli indici, sicché l'ordine di derivazione rispetto  $\varepsilon_{x_k}$  ( $k = 1, \dots, 4$ ) è irrilevante.] Si ottengono così le espressioni

$$(A3.7) \quad (-i)^3 \frac{\delta}{\delta J_{x_3}} \frac{\delta}{\delta J_{x_1}} \frac{\delta}{\delta J_{x_2}} e^{-\frac{1}{2} J_x h_{x'x}^{-1} J_x} =$$

$$= -i \frac{\delta}{\delta J_{x_3}} \left\{ h_{x_2, x_1}^{-1} e^{-\frac{1}{2} J_x h_{x'x}^{-1} J_x} + \left( h_{x_2, x_1}^{-1} J_x h_{x_1, x'}^{-1} J_{x'} \right) e^{-\frac{1}{2} J_x h_{x'x}^{-1} J_x} \right\} =$$

$$= i \left\{ h_{x_2, x_1}^{-1} h_{x_3, x'}^{-1} J_x + h_{x_3, x'}^{-1} J_{x'} h_{x_2, x}^{-1} J_x h_{x_1, x'}^{-1} J_{x'} - h_{x_2, x_3}^{-1} h_{x_1, x}^{-1} J_x - h_{x_2, x_4}^{-1} J_x h_{x_1, x_3}^{-1} \right\} \times e^{-\frac{1}{2} J_x h_{x'x}^{-1} J_x},$$

$$(A3.8) \quad (-i)^4 \frac{\delta}{\delta J_{x_4}} \frac{\delta}{\delta J_{x_3}} \frac{\delta}{\delta J_{x_1}} \frac{\delta}{\delta J_{x_2}} e^{-\frac{1}{2} J_x h_{x'x}^{-1} J_x} = \frac{\delta}{\delta J_{x_4}} \left( \left\{ h_{x_2, x_1}^{-1} h_{x_3, x'}^{-1} J_x + \right. \right.$$

$$\left. \left. + h_{x_3, x''}^{-1} J_{x''} h_{x_2, x}^{-1} J_x h_{x_1, x'}^{-1} J_{x'} - + h_{x_2, x_3}^{-1} h_{x_1, x}^{-1} J_x - h_{x_2, x}^{-1} J_x h_{x_1, x_3}^{-1} \right\} \times e^{-\frac{1}{2} J_x h_{x'x}^{-1} J_x} \right) =$$

$$= \left\{ h_{x_2, x_1}^{-1} h_{x_3, x_4}^{-1} + h_{x_3, x_4}^{-1} h_{x_2, x}^{-1} J_x h_{x_1, x'}^{-1} J_{x'} + h_{x_3, x''}^{-1} J_{x''} h_{x_2, x_4}^{-1} h_{x_1, x'}^{-1} J_{x'} + \right.$$

$$\left. + h_{x_3, x''}^{-1} J_{x''} h_{x_2, x}^{-1} J_x h_{x_1, x_4}^{-1} + h_{x_2, x_3}^{-1} h_{x_1, x_4}^{-1} + h_{x_2, x_4}^{-1} h_{x_1, x_3}^{-1} \right\} e^{-\frac{1}{2} J_x h_{x'x}^{-1} J_x} +$$

$$+ \left\{ h_{x_2, x_1}^{-1} h_{x_3, x''}^{-1} J_x + h_{x_3, x''}^{-1} J_{x''} h_{x_2, x}^{-1} J_x h_{x_1, x'}^{-1} J_{x'} + \right.$$

$$\left. + h_{x_2, x_3}^{-1} h_{x_1, x}^{-1} J_x + h_{x_2, x}^{-1} J_x h_{x_1, x_3}^{-1} \right\} h_{x_4, x}^{-1} J_x e^{-\frac{1}{2} J_x h_{x'x}^{-1} J_x}.$$

In eq. (A3.7), tutti i coefficienti dell'esponenziale contengono un fattore  $J$  e quindi l'espressione si annulla quando valutata per  $J \equiv 0$ . Invece nell'espressione (A3.7), nella prima parentesi grafa, vi sono 3 termini che non contengono  $J$ . Conformemente a quanto indicato in eq. (A3.6), in definitiva, risulta

$$(A3.9) \quad \begin{aligned} \langle \varepsilon_1 \varepsilon_2 \varepsilon_3 \varepsilon_4 \rangle &= h_{x_2 x_1}^{-1} h_{x_3 x_4}^{-1} + h_{x_2 x_3}^{-1} h_{x_1 x_4}^{-1} + h_{x_2 x_4}^{-1} h_{x_1 x_3}^{-1} = \\ &= \zeta_{x_1 x_2} \zeta_{x_3 x_4} + \zeta_{x_1 x_4} \zeta_{x_2 x_3} + \zeta_{x_1 x_3} \zeta_{x_2 x_4} , \end{aligned}$$

Quindi, la presenza di termini irriducibili nelle funzioni a  $n$  punti significa che il processo è non-Gaussiano.

### 5. Funzioni di correlazione come coefficienti dello sviluppo del funzionale di partizione.

Nel caso di una funzione regolare  $Z(J_i)$  con  $i = 1, \dots, N$ , vale lo sviluppo in serie di Mac Laurin:

$$\begin{aligned} Z(J_i) &= Z(0) + \frac{1}{1!} \sum_k \frac{\partial}{\partial J_k} Z(J_i) \Big|_{J_k=0} J_k + \frac{1}{2!} \sum_{k_1, k_2} \frac{\partial}{\partial J_{k_1}} \frac{\partial}{\partial J_{k_2}} Z(J_i) \Big|_{J_k=0} J_{k_1} J_{k_2} + \dots + \\ &+ \frac{1}{n!} \sum_{k_1, k_2, \dots, k_n} \frac{\partial}{\partial J_{k_1}} \frac{\partial}{\partial J_{k_2}} \dots \frac{\partial}{\partial J_{k_n}} Z(J_i) \Big|_{J_k=0} J_{k_1} J_{k_2} \dots J_{k_n} + \dots \end{aligned}$$

dove sono state esplicitate le sommatorie sugli indici ripetuti. Questo sviluppo si generalizza al caso dei funzionali  $Z[J_x]$ , nella forma

$$\begin{aligned} Z(J_x) &= Z(0) + \frac{1}{1!} \frac{\partial}{\partial J_{x'}} Z(J_x) \Big|_{J_x=0} J_{x'} + \frac{1}{2!} \frac{\partial}{\partial J_{x_1}} \frac{\partial}{\partial J_{x_2}} Z(J_x) \Big|_{J_x=0} J_{x_1} J_{x_2} + \dots + \\ &+ \frac{1}{n!} \frac{\partial}{\partial J_{x_1}} \frac{\partial}{\partial J_{x_2}} \dots \frac{\partial}{\partial J_{x_n}} Z(J_x) \Big|_{J_x=0} J_{x_1} J_{x_2} \dots J_{x_n} + \dots \end{aligned}$$

in cui non è stato invece esplicitato l'integrale sugli indici continui ripetuti.

Se scegliamo come funzionale  $Z$  il funzionale di partizione

$$Z[J_x] = \int \mathcal{D}\varepsilon p(\varepsilon_y) e^{iJ_x \varepsilon_x}$$

per cui vale l'espressione (A2.40)

$$\langle \varepsilon_{x_1} \dots \varepsilon_{x_n} \rangle = (-i)^n \frac{\delta}{\delta J_{x_1}} \dots \frac{\delta}{\delta J_{x_n}} Z[J_x] \Big|_{J_x=0} ,$$

varrà l'espansione

$$(A3.10) \quad Z(J_x) = 1 + \frac{i^2}{2} \langle \varepsilon_{x_1} \varepsilon_{x_2} \rangle J_{x_1} J_{x_2} + \dots + \frac{i^n}{n!} \langle \varepsilon_{x_1} \varepsilon_{x_2} \dots \varepsilon_{x_n} \rangle J_{x_1} J_{x_2} \dots J_{x_n} + \dots ,$$

per cui **le funzioni a  $n$  punti connesse riducibili sono i coefficienti dello sviluppo in serie di McLaurin del funzionale di partizione** (si noti che è  $Z[0] = 1$ , mentre la derivata funzionale prima del funzionale di partizione di annulla).

La conoscenza della completa sequenza di tutte le funzioni di correlazione a  $n$  punti è dunque equivalente alla conoscenza dell'integrale di partizione che, a sua volta, è equivalente alla conoscenza del processo.

# Appendice A4

## Lo spettro delle fluttuazioni

### 1. Lo spettro

Il processo che distribuisce le fluttuazioni di densità

$$(A4.1) \quad \varepsilon(\mathbf{x}) = (2\pi)^{-3} \int d^3k \delta(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}$$

ne distribuisce anche le componenti di Fourier  $\delta(\mathbf{k})$ . Il loro spettro si definisce mediando sul processo:

$$(A4.2) \quad P(k) \equiv \langle |\delta(\mathbf{k})|^2 \rangle$$

e dipenderà solo dal modulo  $k$  di  $\mathbf{k}$ , altrimenti il processo violerebbe l'isotropia dello spazio.

Rispetto a altre medie qui troviamo qualche problema in più. Concettualmente sarà

$$\langle |\delta(\mathbf{k})|^2 \rangle = \int \mathcal{D}\varepsilon_{\mathbf{x}} p(\varepsilon_{\mathbf{x}}) |\delta(\mathbf{k})|^2$$

in cui si assume correttamente che  $|\delta(\mathbf{k})|^2$  è esprimibile come funzione di  $\varepsilon_{\mathbf{x}}$ . Il fatto è, tuttavia, che l'operazione (A4.1) che definisce tale funzione ricorre a un integrale su «tutto lo spazio». La media sul processo non può quindi trasformarsi in una media sulle realizzazioni; qualsiasi sia il valore di  $\mathbf{k}$  che ci interessa, per ottenere il  $\delta$  corrispondente, in linea di principio, ho bisogno di tutto lo spazio.

Detto altrimenti, per effettuare la media, oltre che sul «nostro» spazio dovremo valutare  $\delta(k)$  su «altri» spazi.

Si cerca di ovviare a questo problema campionando le componenti di Fourier del campo  $\varepsilon(\mathbf{x})$  in ampi volumi  $\mathcal{V} = L^3$  diversi. Si può assumere che i volumi siano cubi e allora  $\varepsilon(\mathbf{x})$  dovrebbe avere gli stessi valori sulle facce opposte dei cubi. Lo sviluppo di Fourier è allora una serie:

$$(A4.3) \quad \varepsilon(\mathbf{x}) = \sum_{\mathbf{k}} \delta(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} \quad \text{con} \quad \mathbf{k} = \mathbf{n} \, 2\pi/L \quad (\mathbf{n}: \text{vettore con componenti intere}).$$

e, se  $L$  è abbastanza grande, i possibili moduli  $k$  saranno comunque «fitti», così da potersi considerare distribuiti con continuità.

Ovviamente questa procedura concettuale non funziona più all'avvicinarsi dell'orizzonte  $R_H$ . A  $R_H/10$  essa chiaramente già fallisce, se non altro perché la distribuzione «discreta» su  $k$  non ricalca più quella «continua». Ciononostante, seguiremo questo procedimento concettuale, salvo poi passare al limite per sfruttare al meglio i dati effettivamente disponibili.

Vediamo allora come varianza e funzione a 2 punti si possono esprimere tramite lo spettro delle fluttuazioni.

### 2. Espressioni per la varianza di massa e la funzione a 2 punti

Atteniamoci alla definizione (II.18) per la varianza di massa:

$$(A4.4) \quad \sigma_R^2 = \frac{1}{\mathcal{V}} \int_{\mathcal{V}} d^3r \varepsilon_R^2(\mathbf{r}),$$

con  $\mathcal{V} = L^3$ . Usando coerentemente l'espressione (A4.3), avremo allora

$$\begin{aligned}\sigma_R^2 &= \frac{1}{\mathcal{V}} \int_{\mathcal{V}} d^3r \sum_{\mathbf{k}} \delta_R(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} \sum_{\mathbf{k}'} \delta_R^*(\mathbf{k}') e^{-i\mathbf{k}'\cdot\mathbf{r}} = \\ &= \sum_{\mathbf{k}, \mathbf{k}'} \delta_R(\mathbf{k}) \delta_R^*(\mathbf{k}') \frac{1}{\mathcal{V}} \int_{\mathcal{V}} d^3r e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}} = \sum_{\mathbf{k}, \mathbf{k}'} \delta_R(\mathbf{k}) \delta_R^*(\mathbf{k}') \delta_{\mathbf{k}\mathbf{k}'} = \sum_{\mathbf{k}} |\delta_R(\mathbf{k})|^2.\end{aligned}$$

Se passiamo al limite  $\mathcal{V} \rightarrow \infty$ , avremo:

$$\sigma_R^2 = (2\pi)^{-3} \int_0^\infty d^3k |\delta(k)|^2 \tilde{W}^2(kR) = \frac{1}{2\pi^2} \int_0^\infty \frac{dk}{k} k^3 |\delta(k)|^2 \tilde{W}^2(kR)$$

ricordando che è

$$|\delta_R(\mathbf{k})|^2 = |\delta(k) \tilde{W}(kR)|^2.$$

La quantità  $|\delta_R(\mathbf{k})|^2$ , peraltro, contiene la massima informazione che possiamo ottenere sullo spettro. Assumeremo quindi che la procedura di media sia implicita nel passaggio al limite per  $\mathcal{V} \rightarrow \infty$  e scriveremo

$$(A4.5) \quad \sigma_R^2 = \frac{1}{2\pi^2} \int_0^\infty \frac{dk}{k} k^3 P(k) \tilde{W}^2(kR)$$

Il filtro, di fatto, stabilisce un limite superiore «soffice» all'integrazione su  $k$ , tagliando le armoniche con  $kR \gg 1$ , sicché sarà

$$\sigma_R^2 = \frac{1}{2\pi^2} \int_0^{\sim 1/R} \frac{dk}{k} k^3 P(k)$$

In modo analogo, per la funzione a 2 punti, avremo

$$(A4.6) \quad \begin{aligned}\xi(r) &= \langle \varepsilon(\mathbf{x}) \varepsilon(\mathbf{x} + \mathbf{r}) \rangle = \sum_{\mathbf{k}} \sum_{\mathbf{k}'} \delta_R(\mathbf{k}) \delta_R^*(\mathbf{k}') \frac{1}{\mathcal{V}} \int_{\mathcal{V}} d^3r e^{i\mathbf{k}\cdot\mathbf{x}} e^{-i\mathbf{k}'\cdot(\mathbf{x}+\mathbf{r})} = \\ &= \frac{1}{(2\pi)^3} \int_0^\infty dk k^2 P(k) \int_{-1}^1 dz e^{ikrz} = \frac{1}{2\pi^2} \int_0^\infty \frac{dk}{k} k^3 P(k) \frac{\sin kr}{kr}.\end{aligned}$$

Questa relazione può essere invertita:

$$(A4.7) \quad P(k) = \frac{1}{2\pi^2} \int_0^\infty dr r^2 \xi(r) \frac{\sin kr}{kr}.$$

Le eq. (A4.6) e (A4.7) mettono in evidenza il fatto che **la conoscenza della funzione a 2 punti è equivalente alla conoscenza dello spettro.**

Accanto allo spettro si definisce la «funzione spettrale»

$$(A4.8) \quad \Delta^2(k) = \frac{1}{2\pi^2} k^3 P(k)$$

con la quale possiamo scrivere

$$(A4.9) \quad \sigma_R^2 = \int_0^{\infty} \frac{dk}{k} \Delta^2(k) \tilde{W}^2(kR)$$

$$(A4.10) \quad \zeta(r) = \int_0^{\infty} \frac{dk}{k} \Delta^2(k) \frac{\sin kr}{kr}.$$

La funzione spettrale mette in evidenza i contributi alla varianza e alla funzione a 2-punti pesati in modo logaritmico sulle scale.

Se lo spettro è ben approssimabile con una legge di potenza,

$$(A4.11) \quad P(k) = A k^{n_{eff}}, \quad \text{per cui anche} \quad \Delta^2(k) = \frac{A}{2\pi^2} k^{n_{eff}+3},$$

sarà

$$(A4.12) \quad \begin{aligned} \sigma_R^2 &= \frac{A}{2\pi^2} \int_0^{\infty} dk k^{n_{eff}+2} \tilde{W}^2(kR) = \frac{A}{2\pi^2} R^{-(n_{eff}+3)} \int_0^{\infty} dx x^{n_{eff}+2} \tilde{W}^2(x) \\ \zeta(r) &= \frac{A}{2\pi^2} \int_0^{\infty} dk k^{n_{eff}+2} \frac{\sin kr}{kr} = \frac{A}{2\pi^2} r^{-(n_{eff}+3)} \int_0^{\infty} dx x^{n_{eff}+2} \frac{\sin x}{x}. \end{aligned}$$

Nella regione in cui  $\zeta(r) \propto r^{-\gamma}$  avremo pertanto

$$(A4.13) \quad n_{eff} = \gamma - 3 \sim 1.2, \quad \Delta^2(k) \propto k^\gamma.$$

Si può verificare che queste relazioni si applicano anche con  $n_{eff}$  moderatamente variabile in funzione della scala.

## Appendice A5

### Il bias

L'uso di distribuzioni negli spazi funzionali agevola lo studio della possibilità che gli oggetti cosmici siano distribuiti in modo difforme dalla materia.

Si considera in particolare l'ipotesi che gli oggetti cosmici di scala  $R$  nascano nei punti  $\mathbf{r}_i$  ove la fluttuazione supera di  $\nu$  volte  $\sigma_R$ :  $\varepsilon_R(\mathbf{r}_i) > \nu\sigma_R$ . Gli oggetti si collocherebbero dunque in corrispondenza di *picchi eccezionali* nel campo di densità;  $\nu$  è un parametro che va opportunamente fissato.

Nota il campo di fluttuazioni

$$(A5.1) \quad \varepsilon(\mathbf{r}) = \rho(\mathbf{r})/\bar{\rho} - 1$$

( $\bar{\rho}$ : densità media della materia) se ne deduce il **campo di densità *biased***, cui si attribuisce valore 1 dove  $\varepsilon_R(\mathbf{r}) > \nu\sigma_R$  e 0 altrove:

$$(A5.2) \quad \rho_{R,\nu} = \theta\{\nu\sigma_R - \varepsilon_R(\mathbf{r})\};$$

$\theta(u)$  è la «distribuzione a gradino».

Nota il processo  $P[\varepsilon_x]$ , il valore di aspettazione della densità «biased» sarà ottenibile dall'espressione

$$(A5.2) \quad \langle \rho_{R,\nu} \rangle \equiv \Pi_1(\mathbf{r}) = \int \mathcal{D}\varepsilon_x P[\varepsilon_x] \int_{\nu\sigma}^{\infty} d\alpha \delta[\varepsilon_{R,\mathbf{r}} - \alpha],$$

in cui  $\theta(u)$  è stata espressa come integrale della «distribuzione di Dirac»

$$\delta(u) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\phi \exp(iu\phi).$$

Il calcolo è semplificato per un processo Gaussiano, per cui risulta

$$(A5.4) \quad \Pi_1(\mathbf{r}) = \frac{1}{2\pi} \int_{\nu\sigma}^{\infty} d\alpha \int_{-\infty}^{\infty} d\phi e^{-i\alpha\phi} \int \mathcal{D}\varepsilon_x \mathcal{N} e^{-\frac{1}{2}\varepsilon_x \xi_{xy}^{-1} \varepsilon_y + i\varepsilon_x W_{xr}^{(R)} \phi},$$

dove  $\xi_{xy} = \xi(|\mathbf{x} - \mathbf{y}|)$  è la funzione a 2 punti e  $\xi_{xy}^{-1}$  è la sua «matrice inversa». L'esponente si può scrivere (vedi A2.36, A2.37)

$$-\frac{1}{2} \varepsilon_x \xi_{xy}^{-1} \varepsilon_y + i\varepsilon_x J_x^* = -\frac{1}{2} J_x^* \xi_{xy} J_y^* + \varepsilon'_x \xi_{xy}^{-1} \varepsilon'_y$$

con  $J_x^* = W_{xr}^{(R)} \phi$ ,  $\varepsilon'_x = \varepsilon_x - i\xi_{xy} J_y$  e, quindi

$$J_x^* \xi_{xy} J_y^* = \phi W_{rx}^{(R)} \xi_{xy} W_{yr}^{(R)} \phi = \phi \xi^{(R)}(0) \phi,$$

per cui, essendo  $\xi^{(R)}(0) = \sigma_R^2$ , risulta:

$$(A5.5) \quad \Pi_1(\mathbf{r}) = \frac{1}{2\pi} \int_{\nu\sigma}^{\infty} d\alpha \int_{-\infty}^{\infty} d\phi e^{-\frac{1}{2}\alpha^2 \sigma_R^2 \phi - i\alpha\phi}.$$

Risolto l'integrale funzionale, il resto del calcolo non presenta difficoltà. Si richiedono un cambio di variabile d'integrazione  $\phi \rightarrow \phi + i\alpha/\sigma^2$ , per risolvere l'integrale tra  $-\infty$  a  $\infty$  e, successivamente, l'applicazione della relazione approssimata

$$(A5.6) \quad \int_v^\infty dx x^k \exp(-x^2/2) \simeq v^{k-1} \exp(-v^2/2).$$

In questo modo si ottiene

$$(A5.7) \quad \Pi_1(\mathbf{r}) \simeq (2\pi)^{-1/2} v^{-1} e^{-\frac{v^2}{2}}$$

e, di qui, si può ricavare

$$(A5.8) \quad \langle n_{v,R} \rangle = \Pi_1(\mathbf{r})/V_R = 3(2\pi)^{-3/2} R^{-3} v^{-1} e^{-\frac{v^2}{2}},$$

il **numero atteso di oggetti cosmici** di scala  $R$ .

Per ricavare la funzione a 2 punti, si parte analogamente dal calcolo di

$$(A5.9) \quad \begin{aligned} \Pi_2(\mathbf{r}_1, \mathbf{r}_2) &= \int \mathcal{D}\varepsilon_{\mathbf{x}} P[\varepsilon_{\mathbf{x}}] \int_{v\sigma}^\infty d\alpha_1 \delta[\varepsilon_{R,r} - \alpha_1] \int_{v\sigma}^\infty d\alpha_2 \delta[\varepsilon_{R,r} - \alpha_2] = \\ &= (2\pi)^{-2} \int_{v\sigma}^\infty d\alpha_1 \int_{-\infty}^\infty d\phi_1 e^{-i\alpha_1 \phi_1} \int_{v\sigma}^\infty d\alpha_2 \int_{-\infty}^\infty d\phi_2 e^{-i\alpha_2 \phi_2} \times \\ &\quad \times \int \mathcal{D}\varepsilon_{\mathbf{x}} \mathcal{N} e^{-\frac{1}{2}\varepsilon_{\mathbf{x}} \zeta_{\mathbf{x}\mathbf{y}}^{-1} \varepsilon_{\mathbf{y}} + i[\varepsilon_{\mathbf{x}} W_{\mathbf{x}\mathbf{r}_1}^{(R)} \phi_1 + \varepsilon_{\mathbf{x}} W_{\mathbf{x}\mathbf{r}_2}^{(R)} \phi_2]}. \end{aligned}$$

L'argomento dell'esponenziale si può ancora riscrivere

$$(A5.10) \quad -\frac{1}{2} \varepsilon_{\mathbf{x}} \zeta_{\mathbf{x}\mathbf{y}}^{-1} \varepsilon_{\mathbf{y}} + i[\varepsilon_{\mathbf{x}} W_{\mathbf{x}\mathbf{r}_1}^{(R)} \phi_1 + \varepsilon_{\mathbf{x}} W_{\mathbf{x}\mathbf{r}_2}^{(R)} \phi_2] = -\frac{1}{2} \varepsilon_{\mathbf{x}} \zeta_{\mathbf{x}\mathbf{y}}^{-1} \varepsilon_{\mathbf{y}} + iJ_{\mathbf{x}}^* \varepsilon_{\mathbf{x}},$$

con una diversa definizione di

$$(A5.11) \quad J_{\mathbf{x}}^* = W_{\mathbf{x}\mathbf{r}_1}^{(R)} \phi_1 + W_{\mathbf{x}\mathbf{r}_2}^{(R)} \phi_2.$$

Ciò permette di calcolare ancora l'integrale funzionale, ottenendo

$$(A5.12) \quad \int \mathcal{D}\varepsilon_{\mathbf{x}} \mathcal{N} e^{-\frac{1}{2}\varepsilon_{\mathbf{x}} \zeta_{\mathbf{x}\mathbf{y}}^{-1} \varepsilon_{\mathbf{y}} + iJ_{\mathbf{x}}^* \varepsilon_{\mathbf{x}}} = e^{-\frac{1}{2}[W_{\mathbf{r}_1}^{(R)} \phi_1 + W_{\mathbf{r}_2}^{(R)} \phi_2] \zeta_{\mathbf{y}\mathbf{y}} [W_{\mathbf{y}\mathbf{r}_1}^{(R)} \phi_1 + W_{\mathbf{y}\mathbf{r}_2}^{(R)} \phi_2]} = e^{-\frac{1}{2}\phi_i \zeta_{r_i r_j}^{(R)} \phi_j},$$

(è sottointesa la somma da 1 a 2 su  $i$  e  $j$ ). Inserita questa relazione in eq. (A5.9), si ricava

$$(A5.13) \quad \Pi_2(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{(2\pi)^2} \int_{v\sigma}^\infty d\alpha_1 \int_{v\sigma}^\infty d\alpha_2 \int_{-\infty}^\infty d\phi_1 \int_{-\infty}^\infty d\phi_2 e^{-\frac{1}{2}\phi_i \zeta_{r_i r_j}^{(R)} \phi_j - i\alpha_i \phi_i}$$

dove si è usata la matrice  $2 \otimes 2$

$$(A5.14) \quad \zeta_{ij}^{(R)} \equiv \begin{vmatrix} \sigma_R^2 & \zeta_{r_1 r_2}^{(R)} \\ \zeta_{r_2 r_1}^{(R)} & \sigma_R^2 \end{vmatrix} = \sigma_R^2 \begin{vmatrix} 1 & w_{r_1 r_2}^{(R)} \\ w_{r_2 r_1}^{(R)} & 1 \end{vmatrix},$$

intendendosi  $w_{r_i r_j}^{(R)} = \zeta_{r_i r_j}^{(R)} / \sigma_R^2$ . Pertanto, per  $|\mathbf{r}_1 - \mathbf{r}_2| > R$ , risulta  $w_{r_i r_j}^{(R)} \ll 1$  e, trascurando i

termini in  $w^{(R)2}$ , l'inversa della matrice (A5.13) ha l'espressione

$$(A5.15) \quad \zeta_{ij}^{(R)-1} = \sigma_R^{-2} \begin{vmatrix} 1 & -w_{r_1 r_2}^{(R)} \\ -w_{r_2 r_1}^{(R)} & 1 \end{vmatrix},$$

Gli integrali in  $\phi_i$  si trattano «come» l'integrale funzionale, salvo che ora lo spazio è a 2 dimensioni (invece che a  $\infty$ -continue dimensioni). In effetti, ricordando che deve essere

$$\int \prod_{i=1}^2 d\phi_i \frac{\sqrt{Tr_{\zeta_{r_i r_j}^{(R)}}}}{2\pi} e^{-\frac{1}{2}\phi_i \zeta_{r_i r_j}^{(R)} \phi_j} = 1,$$

vista l'eq. (A5.14) che porge

$$\sqrt{Tr_{\zeta_{r_i r_j}^{(R)}}} = \sigma_R^2,$$

avremo, tenuto conto delle componenti della matrice (A5.15),

$$(A5.16) \quad \begin{aligned} \Pi_2(\mathbf{r}_1, \mathbf{r}_2) &= \frac{1}{2\pi\sigma^2} \int_{\nu\sigma}^{\infty} d\alpha_1 \int_{\nu\sigma}^{\infty} d\alpha_2 e^{-\frac{1}{2}\alpha_i \zeta_{ij}^{(R)-1} \alpha_j} = \\ &= \frac{1}{2\pi} \int_{\nu}^{\infty} d\alpha_1 \int_{\nu}^{\infty} d\alpha_2 e^{-\frac{1}{2}\sigma_R^2 \alpha_i \alpha_j \zeta_{ij}^{(R)-1}} = \\ &= \frac{1}{2\pi} \int_{\nu}^{\infty} d\alpha_1 \int_{\nu}^{\infty} d\alpha_2 e^{-\frac{1}{2}(\alpha_1^2 + \alpha_2^2 - 2\alpha_1 \alpha_2 \zeta_{12}^{(R)}/\sigma_R)}. \end{aligned}$$

Applicando ora la generalizzazione dell'approssimazione (A5.6) agli integrali doppi, si ha

$$(A5.17) \quad \Pi_2(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{2\pi\nu^2} e^{-\nu^2} e^{\frac{\nu^2}{\sigma_R^2} \zeta_{r_1 r_2}^{(R)}}$$

e, di qui, vista l'eq. (A5.7), si ricava che

$$(A5.18) \quad \frac{\Pi_2(\mathbf{r}_1, \mathbf{r}_2)}{\Pi_1^2(\mathbf{r})} = \exp\left(\frac{\nu^2}{\sigma_R^2} \zeta_{r_1 r_2}^{(R)}\right).$$

Dall'espressione della funzione a 2 punti, si ottiene facilmente

$$(A5.19) \quad \zeta_{r_1 r_2}^{(v,R)} = \frac{\langle \rho_{R,v}(\mathbf{r}_1) \rho_{R,v}(\mathbf{r}_2) \rangle}{\rho_{R,v}^2} - 1 = \frac{\Pi_2(\mathbf{r}_1, \mathbf{r}_2)}{\Pi_1^2(\mathbf{r})} - 1$$

e quindi, in definitiva

$$(A5.20) \quad \zeta^{(v,R)}(r) = \exp\left(\frac{\nu^2}{\sigma_R^2} \zeta^{(R)}(r)\right) - 1$$

Se  $\nu^2 \zeta^{(R)}(r) / \sigma_R^2 \ll 1$ , se ne deduce che  $\zeta^{(v,R)}(r) \simeq \frac{\nu^2}{\sigma_R^2} \zeta^{(R)}(r)$ , e che

$$(A5.21) \quad b = \nu/\sigma$$

è il rapporto tra  $\varepsilon_R$  e  $\varepsilon_{R,b}$  (vedi testo).

## Appendice A6

### Profondità dei campioni e equazione di Limber

Riportiamo anzitutto la relazione tra  $M$  (magnitudine assoluta) e  $m$  (magnitudine apparente):

$$(A6.1) \quad m = M + 5 \log(r/Mpc) + 25$$

e ricordiamo che la magnitudine assoluta del Sole è  $M_{\odot} \simeq 5.5$ . La funzione di luminosità delle galassie può essere espressa in termini di  $M$ :

$$(A6.2) \quad \langle \delta N \rangle = \Phi(M) \delta M \delta V$$

sarà allora il valore di aspettazione del numero di galassie con magnitudine assoluta tra  $M$  e  $M + \delta M$  in  $\delta V$ .

Se invece cerchiamo il numero di galassie con magnitudine apparente tra  $m$  e  $m + \delta m$  in un angolo solido  $\delta\Omega$ , esso sarà dato da

$$(A6.3) \quad \langle \delta n \rangle = \delta m \delta\Omega \int_0^{\infty} ds s^2 \Phi(m - 5 \log s - 25)$$

( $s = r/Mpc$ ). Si può qui operare un cambio di variabile, definendo

$$\log u = \log s - m/5.$$

Si ottiene così:

$$\langle \delta n \rangle = \delta m \delta\Omega 10^{3m/5} F \times 0.6 \ln 10$$

$$(A6.4) \quad \text{con } F = (0.6 \ln 10)^{-1} \int_0^{\infty} du u^2 \Phi(-5 \log u - 25)$$

Integrando sulle magnitudini apparenti, fino a una determinata magnitudine limite  $m_l$ , si ha allora la densità angolare in numero

$$(A6.5) \quad n_{\Omega}(< m_l) = F 10^{3m_l/5}$$

che, riespressa in termini di flusso, ridà  $N(> S) \propto S^{-3/2}$ .

Nella espressione di Schechter per la funzione di luminosità delle galassie, compare la «luminosità critica»  $L^*$ , attorno alla quale la funzione cessa l'andamento a legge di potenza per avere un brusco decremento esponenziale.

La magnitudine corrispondente a  $L^*$  è

$$(A6.6) \quad M^* = -19.5 + 5 \log h$$

e, se una galassia di magnitudine  $M^*$  ci appare alla magnitudine limite  $m_l$  di un campione, vista eq. (A6.1), la sua distanza sarà

$$(A6.7) \quad D^*/Mpc = 10^{(m_l - M^*)/5 - 5}.$$

Questa è una definizione della «profondità» del campione.

In luogo di esprimere la densità angolare delle galassie in un campione in funzione della magnitudine limite, tramite  $n_{\Omega}(< m_l)$  (eq. A6.5), può essere conveniente esprimerla in funzione della profondità, tramite una funzione  $n_{\Omega}(D^*)$ .

Dalle eq. (A6.4) vediamo allora che

$$(A6.8) \quad n_{\Omega}(D^*) = G D^{*3}$$

con

$$(A6.9) \quad G = F 10^{3M^*/5-15}$$

Ciò pone in evidenza un chiaro effetto di volume. Se lo spazio non fosse euclideo, la relazione (A6.1) non sarebbe valida, in quanto equivalente a  $S = L/4\pi r$ . Non si sarebbe allora potuta ricavare questa legge di proporzionalità al volume. Lo stesso accade se la funzione di luminosità dipende da  $r$ .

La funzione di luminosità espressa tramite le magnitudini entra anche nei coefficienti dell'equazione di Limber. Nel caso in cui sia  $\zeta(r) = (r_o/r)^\gamma$ , essa dà la funzione angolare a 2 punti

$$(A6.10) \quad w_D(\theta) = A(\gamma) \left( \frac{r_o}{D^*} \right)^\gamma \theta^{1-\gamma},$$

noti la lunghezza di correlazione, la profondità del campione e il coefficiente

$$(A6.11) \quad A(\gamma) = \int_0^\infty dq q^4 \mathcal{G}^2(q) \int_{-\infty}^\infty \frac{du}{(u^2 + q^2)^{\gamma/2}}$$

in cui

$$(A6.12) \quad \mathcal{G}(q) = \frac{1}{G} \int_{-\infty}^{5 \log_{10} q} d\mu \Phi(M^* + \mu).$$

Per la dimostrazione di queste relazioni rimandiamo ai testi specializzati.

Per quanto concerne l'espressione di  $A(\gamma)$  è importante notare che la funzione di luminosità  $\Phi$  entra sia nell'espressione di  $\mathcal{G}(q)$ , che è a numeratore, sia in quella di  $G$ , che è al denominatore. Ne segue che variazioni col sito della normalizzazione della funzione di luminosità, tendono a essere riassorbite. Viceversa, variazioni col sito di  $M^*$ , e/o della pendenza  $\alpha$  del tratto a legge di potenza, hanno impatto nel valore del coefficiente  $A(\gamma)$ . Variazioni di  $M^*$  hanno un impatto anche nella definizione della profondità  $D^*$ .

Questo meccanismo attutisce dunque l'impatto dell'evoluzione dell'ampiezza della funzione di luminosità. Variazioni evolutive di  $M^*$  e  $\alpha$ , invece, si riflettono sulle relazioni di scala.

Non riportiamo in questo testo le espressioni ancora più complesse delle generalizzazioni dell'equazione di Limber alle funzioni a  $n$  punti. È peraltro significativo che esse forniscano ulteriori relazioni di scaling, con le quali, in linea di principio, si possono confrontare diversi modelli geometrici e evolutivi.

## CAPITOLO III

# Le equazioni di Friedmann

### Come variano le distanze sulle più grandi scale cosmiche

#### 1. Metrica e geometria

Il primo dato cosmologico è l'espansione del cosmo: a distanze maggiori di  $\sim 10$  Mpc, si comincia a osservare un moto coerente di allontanamento, con velocità  $v \propto d$ , denominato *flusso di Hubble*. A esso sono sovrapposti moti locali, dovuti a *velocità peculiari*  $v_p$  delle singole galassie. La costante di proporzionalità tra velocità e distanza è andata precisandosi nei decenni. Usando tutti i dati disponibili e le relative incertezze statistiche, nel capitolo precedente abbiamo riportato il valore

$$(III.1) \quad H_0 = v/d = 100 h \text{ (km/s)/Mpc} \quad \text{con } h = 0.72 \pm 0.03 .$$

Noto  $H_0$  dal redshift  $z$  si stima la distanza, sempre che essa sia abbastanza elevata da permettere di trascurare le velocità peculiari. Il *redshift* ci fornisce per la velocità un valore numerico  $v$ /(km/s); la relazione (III.1) ci dice allora che la distanza è  $d = 100 h \times [v/(km/s)]$  Mpc. Oppure, conglobando la residua incertezza su  $H_0$  nelle unità di distanza, diremo che è

$$(III.2) \quad d = [v/(100 \text{ km/s})] h^{-1} \text{ Mpc} ;$$

per questo le distanze intergalattiche sono usualmente espresse in  $h^{-1}$  Mpc.

Alcuni esempi: se il *redshift* di una galassia ci dice che essa si allontana da noi a 3000 km/s, la sua distanza si stima in  $30 h^{-1}$  Mpc. Se un ammasso di galassie ha mediamente una velocità di allontanamento di 43 000 km/s, la sua distanza si stima in  $430 h^{-1}$  Mpc.

#### *La geometria spaziale*

Le galassie dunque si allontanano in tutte le direzioni e, indipendentemente dalla direzione, se sono alla stessa distanza hanno (mediamente) la stessa velocità di allontanamento. È noto che ciò non significa che l'osservatore terrestre si situi in posizione privilegiata. È ormai consuetudine dimostrarlo riferendosi a dei punti sulla

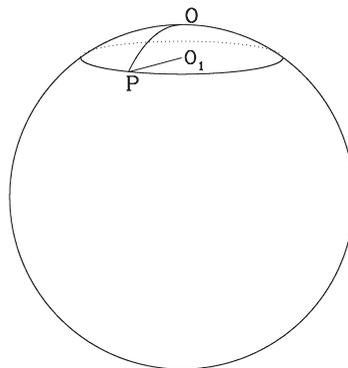
superficie di un palloncino. Se il palloncino si gonfia i punti si allontanano e un osservatore «bidimensionale» che si trovasse su uno qualunque dei punti del palloncino, vedrebbe gli altri punti allontanarsi, isotropicamente, come fosse lui il centro dell'Universo.

Per passare da quest'esempio alla realtà fisica dobbiamo compiere un'operazione intrigante: aumentare di 1 il numero delle dimensioni dello spazio. Il flusso di Hubble potrebbe dunque indicare che ci troviamo sulla superficie di un'ipersfera 3-dimensionale.

Il palloncino dell'esempio è immerso in uno spazio 3-dimensionale; l'ipersfera, di conseguenza, dovrebbe trovarsi in un ambiente 4-dimensionale. Come l'osservatore sul palloncino vede solo il suo mondo a 2 dimensioni, così noi non avremmo cognizione delle 4 dimensioni dell'ambiente (che nulla hanno a che fare con le 4 dimensioni del continuo spazio-temporale relativistico).

Questo esempio è però utile anche per cominciare a delineare quali possano essere le proprietà geometriche dell'Universo. L'osservatore sulla superficie del palloncino, che usualmente coprisse solo piccolissime distanze, confonderebbe i suoi dintorni con un elemento di superficie piana. Tutta la storia precolombiana certifica che si tratta di un fraintendimento quasi inevitabile. La confusione cessa solo quando Colombo – o il nostro osservatore – percorre tratti confrontabili col raggio della Terra-palloncino.

Ma, se l'osservatore avesse a che fare con un raggio di curvatura estremamente maggiore, penserebbe senz'altro di vivere in un mondo piano, in cui la geometria è esattamente *euclidea*. Se poi gli venisse il sospetto che, per qualche bizzarra ragione, la geometria del suo mondo fosse sferica (e la cosa lo interessasse), egli comincerebbe a coprire distanze sempre maggiori, confrontando per esempio il raggio  $R$  di grandi cerchi con la loro circonferenza  $C$ . Sulla superficie di una sfera, il rapporto  $C/R$  è minore di  $2\pi$  (si veda a proposito la figura III.1). Esplorando cerchi sempre maggiori, egli potrebbe



**FIGURA III.1**

Superficie sferica 2-dimensionale su cui si trova una circonferenza con centro in  $O$ . Il rapporto tra la circonferenza e il segmento  $PO_1$  è  $2\pi$ ; il rapporto tra la circonferenza e il segmento sferico  $PO$  è allora evidentemente  $< 2\pi$ .

porre dei limiti sempre crescenti al raggio della sfera. Magari, alla fine, stanco di esplorare, deciderebbe che lo spazio bidimensionale da lui abitato è effettivamente piano, perché *tutto è come se* il centro della sfera del palloncino si trovasse all'infinito.

Torto o ragione che avesse, in tutta l'operazione esplorativa egli avrebbe fatto riferimento a un criterio *metrico* per stabilire la geometria del suo spazio.

Un altro criterio per stabilire la *piattezza* dello spazio, che d'ora in poi considereremo sempre a 3 dimensioni, è di verificare se valga il teorema di Pitagora. Se l'osservatore sta nell'origine e considera due punti  $P$  e  $P'$ , le cui coordinate cartesiane differiscono di  $dr_i$  ( $i = 1, 2, 3$ ), affermando che la loro distanza è

$$(III.3) \quad d\lambda^2 = \sum_{i=1}^3 dr_i^2, \quad \text{spazio piatto}$$

sosterrà la validità del teorema di Pitagora. In coordinate polari  $r, \theta, \phi$ , la stessa distanza si indica

$$(III.4) \quad d\lambda^2 = dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2). \quad \text{spazio piatto}$$

Questa seconda espressione si presta a essere modificata in modo semplice, per indicare quale è la distanza  $PP'$  se lo spazio **non** è piatto. In questo caso sarà

$$(III.5) \quad d\lambda^2 = \frac{dr^2}{1 - \kappa r^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2) \quad \text{spazio curvo}$$

e, in questa espressione, la costante  $\kappa$  prende il nome di *curvatura* dello spazio.

Queste espressioni permettono di sottolineare ancora la relazione tra le nozioni metriche e la geometria e, in effetti, le espressioni (III.4) e (III.5) sono chiamate *metriche* dello spazio.

Elenchiamo qui brevemente alcune quantità che abbiamo appena considerato: (i) le distanze lungo gli assi coordinati, denominate  $dr_i$ ; (ii) la lunghezza del segmento  $PP'$ , denominata  $d\lambda$ ; (iii) gli angoli  $d\theta$  e  $d\phi$ , che esprimono la differenza tra le direzioni in cui l'osservatore «vede»  $P$  e  $P'$ . Dal punto di vista matematico si tratta di *differenziali*, quantità infinitesime. Ma non occorre fare appello ai canoni dell'analisi differenziale per trattarle: qualsiasi distanza su scala subgalattica è tranquillamente assimilabile a un tratto di lunghezza infinitesima. Al lettore meno familiare con la matematica basterà considerare queste e simili quantità come tratti piccoli rispetto alle scale cosmologiche.

Torniamo allora all'osservatore sito nell'origine. Per esso, il luogo dei punti con  $r = \text{costante}$ , sul piano  $\theta = \pi/2$ , è una circonferenza. Immaginiamo però che  $r$  sia molto maggiore della distanza  $PP'$ , di fatto una distanza su scala *cosmologica*. La lunghezza della circonferenza, su questa scala, si ottiene allora «sommando» trattini di lunghezza  $r d\phi$  con  $r$  costante, fino a ottenere l'usuale lunghezza  $r2\pi$ . L'osservatore ottiene invece il raggio della circonferenza «sommando» trattini a  $\theta$  e  $\phi$  costanti, di lunghezza  $dr/\sqrt{1 - \kappa r^2}$ . Se è  $\kappa = 0$ , sommando tanti trattini di lunghezza  $dr$  si ottiene  $r$ ; se è invece  $\kappa \neq 0$ , ciascun trattino è maggiore o minore di  $dr$  a seconda che il termine

$\kappa r^2$ , al denominatore, sottragga o aggiunga qualcosa all'unità. Quindi la somma è maggiore o minore di  $r$  a seconda che sia  $\kappa > 0$  o  $\kappa < 0$ .

Nel caso  $\kappa > 0$  è anche evidente che gli elementi infinitesimi di distanza radiale,  $dr/\sqrt{1-\kappa r^2}$ , sono reali solo per  $r < 1/\sqrt{\kappa}$ . Su una sfera o ipersfera, esiste quindi un limite alle distanze fisiche che si possono considerare, dato dalla semicirconferenza della sfera stessa.  $1/\sqrt{\kappa}$  è dunque il valore massimo accessibile lungo l'asse  $r$ . Per ricavare la distanza di un punto per cui sia  $r = 1/\sqrt{\kappa}$  occorre però «sommare» tanti trattini infinitesimi di lunghezza  $dr/\sqrt{1-\kappa r^2}$ .

*Si tratta in realtà di calcolare l'integrale*

$$\lambda_{max} = \int_0^{\kappa^{-1/2}} \frac{dr}{\sqrt{1-\kappa r^2}} = \frac{1}{\sqrt{\kappa}} \int_0^1 \frac{dx}{\sqrt{1-x^2}} = \frac{1}{\sqrt{\kappa}} \int_0^{\pi/2} \frac{\cos \theta d\theta}{\sqrt{1-\sin^2 \theta}} = \frac{\pi}{2\kappa^{1/2}}$$

Rimarchiamo dunque che un punto che ha coordinata radiale  $1/\sqrt{\kappa}$  ha distanza  $(\pi/2)/\sqrt{\kappa}$ . Si tratta della *semicirconferenza* dell'ipersfera, che ha quindi raggio

$$(III.6) \quad R_c = \frac{1}{2\sqrt{\kappa}}.$$

$R_c$  è il *raggio di curvatura* dello spazio. Nella espressione della «metrica» (III.5) si può sostituire  $\kappa$  col raggio di curvatura e si ottiene

$$(III.7) \quad d\lambda^2 = \frac{dr^2}{1-(r/2R_c)^2} + r^2(d\theta^2 + \sin^2 \theta d\phi^2) \quad \text{spazio curvo}$$

L'eq. (III.5) mette però in evidenza una possibilità ulteriore: quella che sia  $\kappa < 0$ . In questo caso, lo spazio non è piatto, ma non ci troviamo neppure sulla superficie di una ipersfera. Questa geometria si può assimilare a quella che, in 2 dimensioni, vige nell'intorno di un *punto sellare*.

### ***Dalla geometria spaziale alla geometria spazio-temporale***

Sin qui abbiamo *fotografato* il cosmo in un determinato istante temporale, che potrebbe essere quello attuale. Le osservazioni ci dicono anche che la lunghezza  $\kappa^{-1/2}$  è comunque molto superiore delle distanze più grandi esplorate, su cui non scorgiamo deviazioni dalla *piattezza*.

Ma sappiamo che le distanze stanno crescendo nel tempo, e lo possiamo esprimere ricorrendo a un *fattore di scala*  $a(t)$ , di modo che le distanze, a un tempo generico, non siano  $d\lambda$  ma  $a(t) d\lambda$ ; conviene quindi indicarle con

$$(III.8) \quad d\lambda^2 = a^2(t) \left[ \frac{dr^2}{1-\kappa r^2} + r^2(d\theta^2 + \sin^2 \theta d\phi^2) \right],$$

dove  $a(t)$  sarà una funzione crescente di  $t$ .

Bisogna però sottolineare che, al passaggio dall'espressione (III.5) all'espressione

(III.8), le «coordinate»  $r, \theta, \phi$  hanno subito una mutazione genetica. In particolare, in eq. (III.8), si conviene di attribuire al fattore di scala  $a(t)$  le «dimensioni fisiche» di lunghezza, considerando  $r$  come *numero puro*, alla stregua degli angoli.

Un punto  $P$ , la cui coordinata radiale è  $r$ , si trova pertanto a una distanza  $l = a \times r$  dall'osservatore. Nel caso  $\kappa = 0$  questa espressione si semplifica e diventa

$$(III.9) \quad l = a \times r.$$

Per distanze piccole (infinitesime) rispetto al raggio di curvatura, questa relazione è comunque valida.

Del tutto in generale, una distanza fisica andrà scissa nel prodotto di due fattori: (i) una distanza coordinata o *comovente* data dalla somma di intervallini  $d\lambda$ ; (ii) un fattore di scala, che è una lunghezza, e fissa la sua crescita nel tempo. È chiaro che i due fattori non sono separatamente misurabili; ciò che si misura è il loro prodotto. Il fattore  $a(t)$  è quindi sempre fissato a meno di una normalizzazione. D'altro canto, ciò che conta non è il suo valore assoluto, ma il modo in cui esso evolve.

### La geometria spazio-temporale

La crescita delle distanze nel tempo va di pari passo con la crescita delle velocità con la distanza. Se è vero che il flusso di Hubble emerge dal rumore delle velocità peculiari a una distanza  $\sim 10$  Mpc, se ne ricava che il flusso di Hubble è riconoscibile, in una cosmologia con  $h \simeq 0.7$ , quando porta a velocità  $\sim 700$  km/s; si tratta di velocità già assai elevate, oltre 1/500 della velocità della luce. Se la legge di proporzionalità si mantenesse indefinitamente valida, la velocità della luce si raggiungerebbe a  $\sim 4300$  Mpc.

Non occorre molto altro per capire che siamo in ambito decisamente relativistico. Non potremo quindi ovviare alla necessità di considerare l'intervallo spazio-temporale come unica quantità invariante. Esso sarà allora

$$(III.10) \quad ds^2 = c^2 dt^2 - a^2(t) \left[ \frac{dr^2}{1 - \kappa r^2} + r^2 (d\theta^2 + \sin^2 \theta d\phi^2) \right].$$

Sulla distanza temporale «infinitesima»  $dt$  vale quanto detto per le distanze spaziali. In particolare, tutti i tempi su scala umana sono assimilabili a quantità infinitesime  $dt$ .

L'espressione (III.10) costituisce una *metrica spazio-temporale*, denominata *metrica di Friedmann-Robertson-Walker* (FRW). Si può dimostrare che essa è la metrica più generale in cui la sezione spaziale sia isotropa e omogenea.

L'intervallo  $ds^2$  deve essere invariante per *trasformazioni di Lorentz*, cioè al passaggio da un sistema di riferimento  $K$  a un sistema di riferimento  $K'$ , in moto tra loro. Il punto è che la *forma* (III.10) dell'intervallo *non* è invariante. L'introduzione nella metrica di un fattore di scala  $a(t)$  non distrugge l'invarianza relativistica, ma mostra che la metrica ha una forma più «semplice» in un solo sistema di riferimento «privilegiato».

Ciò è ben comprensibile: Il sistema di riferimento privilegiato  $K$  è quello in cui vale omogeneità e isotropia; se passiamo a  $K'$ , l'isotropia è rotta dalla direzione della velocità di  $K'$  rispetto a  $K$  e, di conseguenza, la metrica non appare isotropa.

Resterà comunque vero che le stesse leggi dinamiche valgono in  $K$  e in  $K'$ , ma diventa possibile distinguere tra un sistema di riferimento *fermo* e un sistema di riferimento *in moto*; non per ragioni dinamiche, ma per la distribuzione delle masse e delle velocità degli oggetti che l'osservatore *vede*.

Va poi aggiunto che il sistema di riferimento *fermo* varia di sito in sito. A  $100 h^{-1}$  Mpc dall'osservatore sito nella Via Lattea, un altro osservatore si scoprirebbe «fermo» se si muovesse a 10 000 km/s (1/30 della velocità della luce) rispetto a noi.

## 2. Dalla geometria alla cinematica

La descrizione geometrica del cosmo spazio-temporale non può dunque evitare il riferimento alle velocità. In questo paragrafo approfondiremo la relazione tra geometria e cinematica.

Una volta introdotto il fattore di scala, variabile nel tempo, possiamo provare la relazione

$$(III.11) \quad H_o = \frac{da/dt}{a}$$

che vale per piccoli incrementi temporali  $dt$  e le conseguenti piccole variazioni nel fattore di scala  $da$ .

In effetti, in un tempo  $dt$ , l'espansione del cosmo fa crescere una distanza  $a \times \lambda$  della quantità infinitesima  $da \times \lambda$ . La velocità con cui il punto di allontana (spazio/tempo), rapportata alla sua distanza è allora

$$H_o = \frac{(da \times \lambda)/dt}{(a \times \lambda)} = \frac{da/dt}{a}.$$

È chiaro che questi passaggi si riferiscono a un cosmo ideale, dove il flusso di Hubble non è inquinato dalle velocità peculiari.

C'è un altro elemento che va tenuto ben presente, quando scrutiamo i cieli remoti: l'osservazione a grandi distanze è anche un'osservazione del passato. Quando le distanze si esprimono in anni-luce, la cosa è subito evidente. La luce di *Proxima Centauri*, a 4 anni-luce dal Sole, è stata emessa 4 anni orsono. Ma *Proxima Centauri* è una nostra vicina di casa. 10 kpc sono una tipica distanza galattica; una stella che noi vediamo a 10kpc emetteva la luce che oggi vediamo 30 000 anni orsono, quando sulla Terra gli uomini di Neanderthal si cibavano degli uomini *sapiens-sapiens* (e viceversa). Ma l'ipotetico osservatore, in moto rispetto a noi alla velocità di 10 000 km/s, considerato nel precedente paragrafo, si trova in un ambiente da cui la luce è partita verso di noi circa 500 milioni di anni fa, quando sulla Terra vivevano specie animali e vegetali oggi in gran parte estinte.

Misurare una velocità di allontanamento, un *redshift*, non significa dunque solo misurare una distanza, ma anche un tempo nel passato, quando il fattore di scala  $a(t)$  aveva un valore inferiore al valore  $a_0$  di oggi.

C'è quindi una precisa relazione tra *redshift* e fattore di scala: un fotone giunto con un *redshift*  $z$  è stato emesso quando il fattore di scala era

$$(III.12) \quad a = \frac{a_0}{1+z}.$$

Il *redshift*, una volta depurato dalle velocità peculiari, è dunque una misura dell'età cosmologica, allo stesso modo del tempo  $t$  e del rapporto  $a/a_0$  (tra il fattore di scala a un dato tempo e quello di oggi). Crescendo  $z$ , dunque,  $a$  e  $t$  diminuiscono.

Per provare la relazione (III.12), immaginiamo allora un fotone, emesso con energia  $\varepsilon$ , da una sorgente in allontanamento a velocità  $v$  rispetto all'osservatore. Per determinare l'energia  $\varepsilon_0$  che esso avrà al momento della ricezione si deve usare la relazione relativistica

$$(III.13) \quad \varepsilon_0 = \varepsilon + \Delta\varepsilon = \gamma\varepsilon(1 - \beta) \quad [\text{ove } \beta = v/c, \gamma = (1 - \beta^2)^{-1/2}].$$

$\Delta\varepsilon$  evidentemente è negativo.

Il *redshift* si definisce tramite le lunghezze d'onda  $\lambda$  e  $\lambda_0$  della luce *emessa* e *osservata*:

$$(III.14) \quad z = (\lambda_0 - \lambda)/\lambda$$

I redshift misurati da Hubble e collaboratori erano  $\sim 0.001-0.002$ , valori che sembravano enormi. Un trentennio dopo, quando vennero scoperti i QSO, il loro spettro rimase a lungo misterioso, finché si comprese che per essi era  $z \sim 1$ , sicché la sorgente si allontanava a velocità relativistica, ricavabile tramite l'eq. (III.13). Attualmente si misurano redshift superiori a 5 e valori ancora maggiori potranno essere misurati in futuro.

Per dimostrare l'eq. (III.12), tuttavia, non dobbiamo considerare velocità elevate, ma passare al limite  $v \rightarrow 0$ , come si potrebbe fare in un mondo ideale, del tutto omogeneo, dove le uniche velocità sono dovute al flusso di Hubble.

In un cosmo strettamente omogeneo, dunque, ha senso considerare anche sorgenti in moto a velocità  $dv$  infinitesime rispetto all'osservatore e, in questo caso, le correzioni relativistiche sono irrilevanti e si ha

$$(III.15) \quad dv/c = -d\varepsilon/\varepsilon = -dv/v = d\lambda/\lambda$$

in quanto tra energia di un fotone, e frequenza o lunghezza dell'onda cui esso «appartiene», valgono le relazioni di proporzionalità  $\varepsilon \propto \nu \propto \lambda^{-1}$ .

D'altronde, sul cono luce e usando eq. (III.11), risulta

$$(III.16) \quad dv = H_0 dr = H_0 c dt = c \frac{da}{a} \quad \text{ovvero} \quad \frac{dv}{c} = \frac{da}{a}$$

sicché, tenuto conto di eq. (III.15), si ha la relazione

$$(III.17) \quad d\lambda/\lambda = da/a$$

che è corretto estendere anche a variazioni «grandi» di  $a$  e  $\lambda$ , ottenendo

$$(III.18) \quad z = \frac{\lambda_0 - \lambda}{\lambda} = \frac{a_0 - a}{a} = \frac{a_0}{a} - 1;$$

di qui si ottiene immediatamente l'eq. (III.12) tra  $a$  e  $z$ .

### 3. Il parametro di decelerazione

Osservando il flusso di Hubble viene spontaneo chiedersi se il ritmo di espansione, misurato da  $H_0$ , sia sempre lo stesso di oggi. Una diminuzione nel tempo di  $H_0$  significherebbe che l'espansione è *decelerata*;  $H_0$  crescente, invece, indicherebbe un'espansione *accelerata*.

Senza entrare nei dettagli della dinamica, sembrerebbe spontaneo immaginare che l'attrazione gravitazionale dei materiali cosmici rallenti l'espansione, logicamente su scale di tempo cosmologiche. Ciò fu quanto i primi cosmologi arguirono e per questo essi definirono un *parametro di decelerazione*, dando per scontato che l'espansione dovesse comunque rallentare o, al più, proseguire a ritmo costante.

Questa scelta serve anche a valutare quanto pericoloso sia affidarsi al *senso comune* in un contesto di Relatività Generale. Le osservazioni dell'ultimo decennio sembrano affermare, in modo ormai inequivocabile, che l'attuale espansione del cosmo procede in maniera accelerata. Siccome la storia della scienza ci ha fornito un «parametro di decelerazione», esso risulta negativo.

In un cosmo rigorosamente omogeneo e quindi privo di «velocità peculiari», il parametro di decelerazione si definisce considerando un punto a distanza  $D$  dall'osservatore e supponendo di poter misurare, oltre che la sua velocità  $V$ , anche la sua accelerazione  $\alpha$ . Si definisce allora il parametro di decelerazione

$$(III.19) \quad q_0 = -\frac{\alpha}{D} \frac{1}{H_0^2}.$$

È importante sottolineare che, per il modo in cui è definito,  $q_0$  non dipende dalla scelta di  $D$ . In effetti, sarà comunque  $\alpha \propto D$  e l'eq. (III.19) si può riscrivere riferendosi unicamente al fattore di scala:

$$(III.20) \quad q_0 = -\frac{1}{a} \frac{d(da/dt)}{dt} \times \frac{1}{H_0^2} \left( = -\frac{d^2a/dt^2}{a} \frac{1}{H_0^2} \right)$$

Si tratta di una relazione dall'inevitabile sapore differenziale:  $d(da/dt)$  esprime la variazione del ritmo  $da/dt$  con cui cresce il fattore di scala; si tratta della variazione che avviene nel tempuscolo  $dt$ , cui essa viene poi rapportata.

Prima di concludere questa parte del capitolo, vogliamo sottolineare che tutte le definizioni e le conclusioni tratte in relazione al tempo attuale, valgono anche per un

altro tempo cosmologico. Così, se  $H_0$  diventa  $H$  e  $q_0$  diventa  $q$ , significherà che staremo trattando il ritmo dell'espansione cosmica e la sua accelerazione a un tempo diverso dall'attuale.

#### 4. Dalla cinematica alla dinamica

Dalle osservazioni sull'espansione del cosmo ci proporremo ora di dedurre le proprietà della *sostanza cosmica*. La necessità di muoverci in ambito relativistico è stata già chiarita, sulla base della semplice rilevazione delle velocità relative degli oggetti cosmici, molto più vicine a  $c$  delle velocità con cui operiamo su scala terrestre o planetaria. Se le forze che agiscono in quest'ambito sono (anche) gravitazionali, diventa allora imprescindibile l'uso della Relatività Generale.

A questo proposito ha molto rilievo la definizione stessa di *galassia*, come rilevato nel precedente capitolo. Ricordiamo che la scala galattica separa l'ambito in cui la dinamica risente delle forze *dissipative*, dalle grandi scale, su cui le forze dissipative non hanno tempo a dispiegare una significativa azione dinamica: l'emissione di radiazione  $X$ , da parte di un ammasso di galassie, è il segnale più cospicuo che esso ci invia; ma la dinamica complessiva del gas e delle galassie che in esso si muovono, è regolata dalla gravità. (Anche se la comprensione dettagliata dei dati richiederà di tenere conto di opportune correzioni da effetti non gravitazionali.)

Abbiamo quindi una notizia buona e una notizia cattiva. Quella buona è che, cercando di scrivere la dinamica del cosmo su base puramente gravitazionale, ci aspettiamo di andare molto vicini all'essenziale. Quella cattiva è che, nel contesto in cui operiamo, la gravità è *hard*, l'uso della Relatività Generale sembra imprescindibile. Il punto di partenza sono dunque le equazioni di Einstein. Applicate a un continuo isotropo e omogeneo, come più volte preconizzato, esse diventano le equazioni di Friedmann, di cui tratteremo nel prossimo paragrafo.

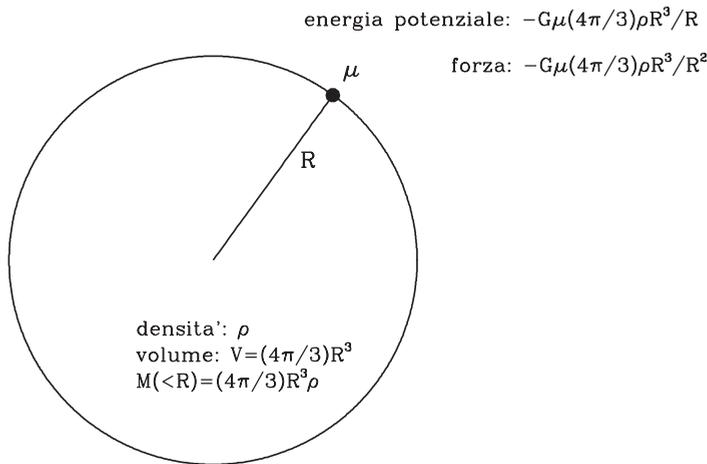
#### 5. Le equazioni di Friedmann

Utilizzando il parametro di Hubble e il parametro di decelerazione, possiamo ora scrivere le *equazioni di Friedmann* e discuterne il significato. Si veda appendice A1 per il metodo da seguire per la loro derivazione. La loro forma è la seguente:

$$(III.21) \quad H^2 = \frac{8\pi}{3}G\rho - \frac{\kappa c^2}{a^2}, \quad -H^2q = \frac{4\pi}{3}G(\rho + 3p/c^2).$$

$G$  è la costante di gravitazione;  $a$  è il fattore di scala;  $\kappa$  è la curvatura, che entra nell'espressione (III.10) della metrica FRW;  $\rho$  e  $p$  sono densità e pressione della sostanza omogenea e isotropa contenuta nell'ambiente cosmico.

Una prova di *ragionevolezza* delle equazioni di Friedmann si ottiene tramite gli argomenti esposti in figura III.2, dove viene considerata una particella di prova di



$$\begin{aligned} \text{energia cinetica} + \text{energia potenziale} &= \text{costante} \\ \mu v^2/2 \quad -\mu G(4\pi/3)\rho R^2 &= K \\ \text{forza} &= \text{massa} \times \text{accelerazione} \\ -\mu G(4\pi/3)\rho R &= \mu \times \alpha \end{aligned}$$

**FIGURA III.2**

*Giustificazione newtoniana delle equazioni di Friedmann.*

All'interno della sfera sono indicate la sua densità  $\rho$ , il suo volume  $V$  e la sua massa.

La stessa massa entra nelle espressioni dell'energia potenziale della particella di massa  $\mu$  e della forza con cui essa è attratta (in alto).

massa  $\mu$  ai bordi di una sfera omogenea di densità  $\rho$  e raggio  $R$ . L'equazione di conservazione dell'energia meccanica, divisa per  $\mu R^2$  e moltiplicata per 2, dà l'equazione

$$(III.22) \quad (v/R)^2 = (8\pi/3)G\rho + (2K/\mu)/R^2$$

che coincide con la prima eq. (III.21), se si prende  $H = v/R$  e si definisce  $\kappa c = 2K/\mu$ . Il senso fisico apparente della prima equazione di Friedmann è dunque la «conservazione dell'energia».

Ciò pone in relazione  $\kappa$  con  $K$ , *energia totale* in prospettiva newtoniana. Viene allora naturale considerare l'esempio di un *grave*, lanciato verso l'alto, che si allontana definitivamente dalla Terra, oppure vi ricade, a seconda che sia  $v \geq v_f$  (velocità di fuga) e energia totale  $\varepsilon_{tot} \geq 0$ , oppure  $v < v_f$  e  $\varepsilon_{tot} < 0$ . In eq. (III.22),  $K$  è l'energia totale e il suo segno è opposto a quello di  $\kappa$ . Questa analogia proposta in molti testi di cosmologia, porterebbe a concludere che, se  $\kappa \leq 0$ ,  $a(t)$  è sempre crescente; avremmo invece espansione seguita da ricollasso se  $\kappa > 0$ .

Queste conclusioni risultano oggi parzialmente superate, a causa del ruolo assunto dall'*energia oscura*.

Dividendo poi per  $\mu R$  la seconda equazione in figura, si ottiene

$$(III.23) \quad -\alpha/R = (4\pi/3)G\rho$$

e quindi la seconda eq. (III.21), sulla base di eq. (III.19), *nel caso di pressione nulla*. Il senso apparente della seconda eq. di Friedmann è dunque quello di « $f = ma$ ».

Accanto alle due eq. (III.21) va riportata un'ulteriore equazione:

$$(III.24) \quad d(\rho a^3) = -pd(a^3)/c^2$$

che si interpreta facilmente come  $dU = -p dV$ : variazione di energia pari al lavoro delle forze di pressione (ricordiamo che, per definizione di densità, l'energia nel volume  $V$  è  $U = \rho c^2 V$ ). Ebbene, in appendice A2, partendo dalla prima eq. (III.21) e dall'eq. (III.24), si ricava la seconda eq. (III.21), con il giusto termine di pressione.

Ciò equivale a dire che le tre equazioni (III.21), (III.24) sono tra loro consistenti solo in un contesto general-relativistico, non lo sarebbero in fisica newtoniana.

Per le sostanze ordinarie, comunque, la densità d'energia, dell'ordine di  $mc^2/\text{volume}$ , è assai maggiore della pressione che è dell'ordine di  $mv^2/\text{volume}$ . Per sostanze composte di particelle non relativistiche, quindi, il termine di pressione ha un rilievo modesto. Ma, nel contesto cosmologico, i casi in cui la pressione è rilevante sono significativi. È quindi giusto concludere che non si potrebbe descrivere l'evoluzione osservata del cosmo senza Relatività Generale.

In appendice A2 vengono ricavate altre forme per le equazioni di Friedmann (III.21). Per esempio, l'eq. (III.24) può facilmente essere messa nella forma

$$(III.25) \quad d\rho = -dt 3H(\rho + p/c^2),$$

che esprime direttamente la variazione della densità d'energia  $\rho$ . Per  $p = -\rho c^2$ , dunque,  $\rho$  rimarrebbe costante: il lavoro della pressione negativa compenserebbe esattamente la perdita di densità conseguente all'espansione.

Questo caso apparentemente bizzarro corrisponde alla realtà osservativa della sostanza oggi dominante nel cosmo. Dovremo quindi discutere a fondo la possibilità di sostanze a pressione negativa.

## 6. Densità critica e parametro di densità

Se è  $\kappa = 0$ , le eq. (III.21) ci dicono che è

$$(III.26) \quad H^2 = \frac{8\pi}{3}G\rho.$$

Invertendo questa equazione di ricava per la densità il valore

$$(III.27) \quad \rho_{cr} = \frac{3H^2}{8\pi G},$$

detto *densità critica*. Dal parametro di Hubble,  $H_o = h \text{ 100 km/s/Mpc}$ , si ricava, per la densità critica attuale, il valore

$$(III.28) \quad \rho_{o,cr} = 1.8791 h^2 \times 10^{-29} \text{ g/cm}^{-3} = 2.778 h^2 \times 10^{11} M_{\odot}/\text{Mpc}^3,$$

che coinciderebbe con la densità reale del cosmo solo nel caso  $\kappa = 0$ .

Definita  $\rho_{cr}$ , la prima eq. (III.21) si può riscrivere nella forma

$$(III.29) \quad \frac{8\pi}{3} G \rho_{cr} - \frac{8\pi}{3} G \rho = -\frac{\kappa c^2}{a^2},$$

$$(III.30) \quad \text{e si definisce} \quad \Omega = \frac{\rho}{\rho_{cr}} \quad (\text{parametro di densità}).$$

Sulla base di queste relazioni si stabiliscono delle corrispondenze precise, raccolte nella tabella qui sotto.

curvatura	densità	par. di dens.
$\kappa < 0$	$\rho < \rho_{cr}$	$\Omega < 1$
$\kappa = 0$	$\rho = \rho_{cr}$	$\Omega = 1$
$\kappa > 0$	$\rho > \rho_{cr}$	$\Omega > 1$

In generale,  $\Omega$  varia al trascorrere di  $t$ . Il suo valore attuale  $\Omega_o$  è comunque uno dei parametri più importanti a definire il modello cosmologico più aderente ai dati.

Nel cosmo possono essere presenti diverse componenti. Per esempio barioni e materia oscura fredda, denominata CDM (*Cold Dark Matter*). Per ciascuna di esse si definisce il relativo parametro di densità. Per esempio  $\Omega_b = \rho_b/\rho_{cr}$  per i barioni,  $\Omega_c = \rho_c/\rho_{cr}$  per CDM. Infine, in presenza di DE (*Dark Energy*: energia oscura), avremo il parametro di densità  $\Omega_{de} = \rho_{de}/\rho_{cr}$ .

C'è una relazione tra  $\Omega$  e parametro di decelerazione. Dividendo membro a membro le equazioni

$$H^2 = (8\pi/3)G\rho_{cr}, \quad H^2q = (4\pi/3)G(1+3w)\rho$$

(definizione di densità critica e seconda eq. (III.21) con  $w = p/\rho c^2$ ), si ottiene

$$(III.31) \quad q = \frac{1+3w}{2} \Omega \quad \text{ovvero} \quad q = \sum_i \frac{1+3w_i}{2} \Omega_i$$

(la seconda equazione vale quando ci sono più componenti di densità confrontabili, con parametri di stato e di densità  $w_i$  e  $\Omega_i$ ).

Queste relazioni sono ricche di conseguenze:

- 1) Se la sostanza dominante oggi nel cosmo è la materia (non relativistica), per cui  $w = 0$ , allora  $q_o = \Omega_o/2$ . Una relazione analoga vale a ogni tempo  $t$  a cui la materia sia dominante.
- 2) Per avere  $q < 0$  (espansione accelerata) occorre che sia  $w < -1/3$ .

3) Supponiamo sia  $\Omega = 1$  e ci siano due sostanze, con parametri di stato e di densità  $w_{1,2}$  e  $\Omega_{1,2}$ . Questo può essere il caso del cosmo odierno, per cui i dati osservativi indicano che  $q \lesssim -0.6$ . Volendo ottenere questo risultato con  $w_1 = 0$  e  $w_2 = -1$ , si ha:

$$(III.32) \quad 1 - \frac{3}{2}\Omega_2 < -0.6 \quad \text{da cui: } \Omega_2 \gtrsim 0.73.$$

Se ne conclude che  $\Omega_{de} > 0.7$ : circa il 70% della “sostanza” del cosmo è oggi costituita di DE.

## 7. Evoluzione del parametro di densità

L'evoluzione del parametro di densità può essere ricavata da eq. (III.29). Al passaggio del fattore di scala e della densità da valori  $a_o$  e  $\rho_o$  a valori  $a$  e  $\rho$ , il parametro di densità evolve come segue:

$$(III.33) \quad \Omega^{-1} - 1 = (\Omega_o^{-1} - 1) \frac{\rho_o a_o^2}{\rho a^2}.$$

Infatti, divisi ambo i membri di eq. (III.29) per il secondo termine, si ottiene

$$(III.34) \quad \frac{\rho_{cr}}{\rho} - 1 = -\frac{\kappa c^2}{\frac{8\pi}{3}G\rho a^2}, \quad \text{ovvero } \Omega^{-1} - 1 \propto 1/(\rho a^2)$$

e quindi, normalizzando la proporzionalità a un tempo  $t_o$ , segue l'eq. (III.33).

Per servirsi pienamente di eq. (III.33), occorre conoscere l'andamento delle densità  $\rho$  in funzione di  $a$ . Se assumiamo una legge di proporzionalità  $\rho \propto a^{-s}$ , lasciando a priori indeterminata la potenza  $s$ , abbiamo

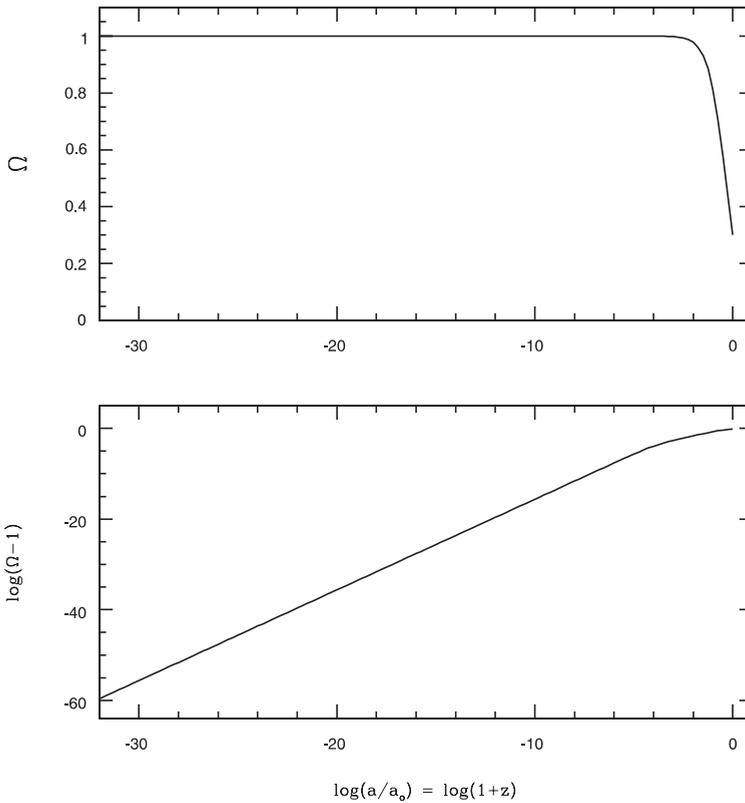
$$(III.35) \quad \Omega^{-1} - 1 = (\Omega_o^{-1} - 1) \left(\frac{a_o}{a}\right)^{2-s},$$

sicché il valore  $s = 2$  costituisce il discriminante tra due comportamenti opposti. Va peraltro notato che l'eq. (III.35) mette in evidenza che, se è  $\Omega = 1$  a un tempo, è  $\Omega = 1$  sempre.

Diamo comunque alcune informazioni su fatti che saranno provati nei due prossimi capitoli: (i) Per materia ordinaria e radiazione si ha  $s = 3$  e 4. (ii) Per  $a/a_o \lesssim 10^{-4}$  (salvo tempi molto prossimi al presente, ma irrilevanti ai fini del ragionamento che andremo a fare), la materia ordinaria è la sostanza prevalente nel cosmo; a tempi anteriori, quando era  $a/a_o \gtrsim 10^{-4}$ , la sostanza dominante è invece la radiazione.

Viste queste prevalenze, nel caso  $\Omega_o \neq 1$ , l'andamento di  $\Omega(a)$ , assumendo che sia  $\Omega_o = 0.3$ , è quello mostrato in figura III.3.

C'è un limite inferiore al fattore di scala, sotto il quale non ha più senso considerare il grafico in figura III.3. A tempi «anteriori» al *tempo di Planck*, le eq. di Einstein perdono la loro validità e si è in regime di *gravità quantistica*. Il tempo di Planck è  $\sim 10^{-43}$ sec. e da ciò si deduce che solo per  $a > 10^{-32}a_o$ , valgono le equazioni di Friedmann, ricavate dalle equazioni di Einstein. Maggiori dettagli sul tempo di Planck sono forniti nell'ultimo paragrafo del capitolo.

**FIGURA III.3**

*Fine tuning* di  $\Omega$ . Andamento del parametro di densità in funzione del fattore di scala, a partire dal tempo di Planck. I grafici mostrano come, per ottenere oggi un valore di  $\Omega \simeq 0.3$ , si debba partire con un valore che differisce dall'unità di circa 1 parte su  $10^{60}$ . Una situazione analoga si ha per qualsiasi valore di  $\Omega_0 \neq 1$ , ma dell'ordine dell'unità. L'ascissa (logaritmica)  $a_0/a$  si può correlare col *redshift*  $z$ , come mostrato in figura.

Usando l'eq. (III.35), si ricava allora che, alla fuoriuscita dal regime di gravità quantistica,  $\Omega$  doveva differire dall'unità di  $\sim 1 : 10^{60}$ : una parte su un milione di miliardi di miliardi di miliardi di miliardi di miliardi di miliardi.

È questo un primo esempio di *fine-tuning* (accordatura perfetta) che si richiede alla «condizioni iniziali» dell'espansione, per produrre l'Universo osservato, *a meno che sia*  $\Omega = 1$ .

Se ne può concludere che *deve essere*  $\Omega \equiv 1$ ? Questa opzione è senz'altro compatibile coi dati osservativi. La credibilità di argomenti volti a evitare «fine tunings» è molto dibattuta. Comunque, su questo punto, l'argomentazione andrà completata nell'ambito dello studio delle teorie inflazionarie.

## 8. Il valore della curvatura

Dall'eq. (III.29), ricordando la definizione (III.27), si ricava la relazione

$$(III.36) \quad \frac{\kappa c^2}{a_o^2} = -H_o^2(1 - \Omega_o),$$

valida al tempo attuale  $t_o$ . Per  $\kappa = 0$  la normalizzazione del fattore di scala è evidentemente arbitraria. La scelta usuale è allora  $a(t_o) = a_o = 1$  (Mpc).

Supponiamo invece di avere effettuato misure e osservazioni che determinino  $H_o$  e  $\Omega_o$ , e che sia  $\Omega_o \neq 1$ . Da queste misure si ricava anche il rapporto  $\kappa/a_o^2$ , ma non  $\kappa$  e  $a_o$  singolarmente.

Tra le infinite possibili scelte, due sono le opzioni più comuni:

1) Si fissa  $a_o = 1$  (Mpc) e si determina di conseguenza il valore di

$$(III.37) \quad \kappa/a_o^2 = (H_o/c)^2(1 - \Omega_o) = [3000 h^{-1}(\text{Mpc})]^{-2}(1 - \Omega_o)$$

2. Si decide che  $\kappa$ , oltre al valore 0 per  $\Omega_o = 1$ , possa avere solo i valori  $+1$  o  $-1$ , rispettivamente nei casi di curvatura positiva o negativa. Ciò permette alcune semplificazioni discusse in appendice A3. Questa scelta, in generale, sarà incompatibile con la normalizzazione  $a_o = 1$ , a meno di non scegliere un'unità di lunghezza *ad-hoc*, diversa dal Mpc.

La relazione tra  $\kappa$  e  $\Omega_o$  permette di esprimere anche il raggio di curvatura dello spazio in termini di  $\Omega_o$ , facendo uso dell'eq. (III.6). Si ottiene così

$$(III.38) \quad R_c = \frac{c}{H_o} \frac{1}{2\sqrt{\Omega_o - 1}},$$

Va qui ribadito che il «raggio dell'Universo» esiste solo in caso di curvatura positiva. Esso non va confuso con l'*orizzonte cosmologico*, di cui parleremo nel prossimo capitolo, che è pure  $\sim \mathcal{O}(c/H_o)$ .

## 9. La scala di Planck

Nel paragrafo 7 e in altre occasioni ci siamo riferiti alla scala di Planck. In questo paragrafo vogliamo approfondirne il significato.

In fisica ci sono molte costanti con precisi valori numerici, ma *tutte possono essere riportate a tre sole costanti dimensionali*:  $G, c, \hbar$  (costante di interazione gravitazionale, velocità della luce e costante di Planck). Il valore di queste costanti dipende quindi dalle unità di misura. Se scegliamo unità *c.g.s.*, legate all'esperienza dell'*homo faber*, l'ordine di grandezza di  $G, c, \hbar$  è, rispettivamente,  $10^{-7}$ ,  $10^{11}$  e  $10^{-27}$ . Si tratta dunque di valori espressi tramite potenze di 10 lontane dall'unità. Ciò esprime la lontananza dell'esperienza ordinaria dal mondo celeste, in cui la gravità ha effetti che vanno oltre l'*accelerazione di gravità g*, e ancor più dai regimi rela-

tivistico e quantistico. La stessa sequenza storica della scoperta e determinazione di  $G$ ,  $c$ ,  $\hbar$  segue grosso modo la rispettiva distanza dall'unità.

Legata a  $G$  c'è la scoperta della *gravitazione universale*; all'invarianza di  $c$  è legata la scoperta della *Relatività* e della struttura non-euclidea dello spazio-tempo; legata alla determinazione di  $\hbar$ , infine, c'è la meccanica *quantistica*.

Ancor prima che  $G$ ,  $c$  e  $\hbar$  fossero scoperte e misurate, era stato possibile fissare a 3 il numero delle costanti dimensionali della fisica. I sistemi di unità di misura possono fissare grandezze di riferimento in modo più o meno arbitrario. Per esempio, pur misurando le lunghezze in metri o centimetri, si possono misurare quote in *pièdi*, le profondità in *braccia*, le superfici in *acri*, senza preoccuparsi troppo, per esempio, della relazione tra piede<sup>2</sup> e acro. È tuttavia più sensato misurare le superfici usando come unità il quadrato dell'unità di lunghezza. Le tribù primitive non lo fanno, i popoli moderni sì.

Alcune scelte arbitrarie permangono, per ragioni di comodità. Per esempio, l'unità di carica elettrica, che si potrebbe fissare usando la legge di Coulomb, come si faceva nel sistema *c.g.s.* elettrostatico, nel S.I. viene invece stabilita in modo arbitrario. Abbiamo allora il sistema *MKSQ* in cui le unità considerate *primitive* sono quattro. Ma, mentre si possono usare più di 3 unità fondamentali, apparentemente *non si può scendere al di sotto di 3*.

In realtà un modo ci sarebbe, quello di sfruttare le leggi in cui intervengono direttamente  $G$ ,  $c$  e  $\hbar$ . Ciò è equivalente a non usare più, come grandezze primitive, lunghezze, tempi e masse, ma a scegliere come primitive le grandezze fisiche che hanno le *dimensioni* di gravità, velocità e azione. Sarebbe cioè «naturale» fissare come unitario il valore delle 3 costanti dimensionali e misurare tutte le altre grandezze fisiche in multipli di esse o di loro combinazioni.

In particolare, si possono costruire, con  $G$ ,  $c$  e  $\hbar$  delle combinazioni che hanno *dimensioni fisiche*  $L$ ,  $T$  e  $M$  (lunghezze, tempi e massa). Tali combinazioni vengono indicate col suffisso  $_P$  e dette «*di Planck*». Esse sono le *unità naturali* di lunghezze, tempi e masse. Allo stesso modo si possono costruire accelerazione, viscosità, energia, ecc. *di Planck*.

Lunghezze, tempi e masse di Planck si ottengono con le combinazioni:

$$l_P^2 = G \hbar c^{-3}, \quad t_P^2 = G \hbar c^{-5}, \quad m_P^2 = G^{-1} \hbar c.$$

Per energia di Planck risulta:

$$E_P^2 = G^{-1} \hbar c^5.$$

Naturalmente tutte queste definizioni possono essere modificate, moltiplicando le quantità indicate per arbitrari fattori numerici.

Discuteremo ora il ruolo cosmologico del *tempo di Planck*. Supponiamo di compiere idealmente una *misura di energia* ( $\varepsilon$ ) a un tempo  $t_P$ , dopo il *big-bang*. Tale misura sarà soggetta a un'indeterminazione quantistica tale che

$$\Delta \varepsilon t_P \simeq \hbar,$$

in quanto dobbiamo effettuarla in un tempo breve. Se l'energia misurata è quella contenuta in un volume di dimensioni  $l_P = ct_P$  (lunghezza di Planck), la massa corri-

spondente  $m = \varepsilon/c^2$  sarà indeterminata entro un intervallo di ampiezza  $\Delta m \simeq \hbar/t_P c^2$ .

Ma il raggio di Schwartzschild  $R_S$  per  $\Delta m$ , a sua volta, è dell'ordine della lunghezza di Planck. Ciò risulta dai seguenti passaggi:

$$R_S \sim G\Delta m/c^2 = G\hbar/t_P c^4 = G\hbar/l_P c^3 = l_p^2/l_P = l_p$$

Ricordiamo che il raggio di Schwartzschild o *raggio gravitazionale* è il raggio della sfera entro cui deve concentrarsi la massa  $M$ , perché la velocità di fuga dalla sua superficie sia la velocità della luce: quando una massa è concentrata entro il suo raggio gravitazionale, quindi, essa non può più emettere radiazione e diviene un *buco nero*. Una «fluttuazione quantistica» della massa contenuta nel volume  $l_p^3$  è dunque sufficiente a trasformare tale volume in un *buco nero*.

L'emersione del cosmo dal mondo pre-quantistico appare quindi legata a una problematica distribuzione di probabilità, che trasforma alcune «zone cosmiche» in aree destinate a chiudersi su se stesse, mentre lancia altre «zone cosmiche» verso un futuro di espansione illimitata.

Per arrivare a questa conclusione siamo ricorsi a argomenti validi ma sommari. La massa in un volume  $l_p^3$  dovrebbe assumere un valore determinato solo in conseguenza di una sua *misurazione*. A priori, la meccanica quantistica dovrebbe essere in grado di fornire dettagli sulle *distribuzioni di probabilità*, dei possibili *autovalori* della massa. Essi si troveranno entro intervalli definiti dalle relazioni di indeterminazione, ma si otterranno tramite la soluzione di una equazione dinamica, in forma opportuna.

L'equazione dinamica è stata prima ricavata per sistemi discreti, a numero finito o numerabile di gradi di libertà (equazione di Schrödinger). Essa è stata poi ottenuta anche per sistemi a infiniti gradi di libertà, dando luogo alla teoria quantistica dei campi, che illustra anche la relazione tra aspetti particellari e ondulatori e la loro incompatibilità osservativa. Nulla di analogo è invece stato ancora ben individuato nel caso della *gravità quantistica*.

Negli ultimi anni, un passo avanti significativo si è basato sulla sostituzione del *punto*, come elemento topologico di base nella struttura spazio-temporale, con la *corda* (*string*). Dopo un iniziale entusiasmo per questa formulazione, che pareva destinata a dipanare la matassa della gravità quantistica, portando a teorie non solo rinormalizzabili, ma addirittura prive di infinità, l'analisi si è un po' impantanata in problematiche matematiche e non sembrano imminenti sbocchi significativi di carattere sperimentale.

In assenza di una fisica consolidata da applicare in regime di gravità quantistica, la cosmologia fisica può solo prendere inizio al tempo  $t_P$ . Molti problemi fisici sono quindi rimandati a un ambito certamente insondabile, ma apparentemente minimo. Torneremo su questo punto nel capitolo dedicato al significato del *tempo*, anche se è opportuno precisare subito che la scarsa significanza dell'ambito *preplanckiano*, il suo ridursi alla durata «infima» di  $10^{-43}$  sec., potrebbero essere legati a un giudizio basato su categorie strettamente antropocentriche.

# Appendice A1

## Dalle equazioni di Einstein alle equazioni di Friedmann

Le equazioni di Einstein mettono in relazione il campo gravitazionale e la sua sorgente, il tensore energia-impulso  $T_{\mu\nu}$ . La metrica spazio-temporale si può sempre scrivere nella forma

$$(A1.1) \quad ds^2 = g_{\mu\nu} dx^\mu dx^\nu,$$

e il campo gravitazionale è dato dalla deviazione di  $g_{\mu\nu}$  dalla forma Minkowskiana che, se si usano coordinate cartesiane ortogonali è diagonale ha componenti  $+1, -1, -1, -1$ . In coordinate polari, da (A1.1) si passa all'espressione

$$(A1.2) \quad ds^2 = c^2 dt^2 - a^2(t) \left[ \frac{dr^2}{1 - \kappa r^2} + r^2 (d\theta^2 + \sin^2\theta d\phi^2) \right],$$

sulla base di chiare considerazioni di simmetria. Sulla sola base di considerazioni di simmetria, si può poi anche dimostrare che, in un cosmo isotropo e omogeneo, il tensore energia-impulso ha la forma (qui, come nella maggior parte delle appendici, si usano unità con  $c = 1$ )

$$(A1.3) \quad T_{\mu\nu} = (\rho + p)u_\mu u_\nu - p g_{\mu\nu},$$

dove il 4-vettore  $u_\mu$ , nel sistema di riferimento cartesiano in cui vale omogeneità e isotropia, ha componenti

$$(A1.4) \quad u_0 = 1, \quad u_\alpha = 0 \quad (\alpha = 1, 2, 3).$$

L'espressione (A1.3) è anche ottenibile in idrodinamica per un fluido perfetto. L'ipotesi che la sostanza cosmica sia un fluido non è però qui necessaria. Peraltro, se il contenuto del cosmo si può approssimare con un fluido perfetto, allora  $\rho$  è la sua densità d'energia e  $p$  è la sua pressione. Useremo comunque questi appellativi per  $\rho$  e  $p$ .

Un esempio di sostanza cosmica che non costituisce un fluido è la radiazione che compone il CMB. Il cammino libero medio dei fotoni di questa radiazione è  $\gg ct_0$  ( $t_0$ : età attuale del cosmo), sicché i fotoni, di norma, non possono cedere impulso a altre particelle o scambiarlo tra loro. Non c'è quindi un'azione pressoria, «ponderomotrice» del CMB.

Le equazioni di Einstein contengono le componenti di  $T_{\mu\nu}$  e si scrivono:

$$(A1.5) \quad R_\mu^\nu = \frac{8\pi}{c^2} G \left( T_\mu^\nu - \frac{1}{2} T g_\mu^\nu \right), \quad \text{con } T = T_\mu^\mu.$$

Ricordiamo anche che:

$$(A1.6a) \quad R_{\mu\nu} = \partial_\mu \Gamma_{\nu\alpha}^\alpha - \partial_\alpha \Gamma_{\mu\nu}^\alpha + \Gamma_{\mu\alpha}^\beta \Gamma_{\nu\beta}^\alpha - \Gamma_{\alpha\beta}^\alpha \Gamma_{\mu\nu}^\beta,$$

$$(A1.6b) \quad \text{con } \Gamma_{\alpha\beta}^\gamma = \frac{1}{2} g^{\gamma\nu} (\partial_\alpha g_{\nu\beta} + \partial_\beta g_{\nu\alpha} - \partial_\nu g_{\alpha\beta});$$

mentre, sulla base di eq. (A1.2), le componenti di  $g_{\mu\nu}$  sono:

$$(A1.7a) \quad g_{00} = 1, \quad g_{11} = -\frac{a^2}{1 - \kappa r^2}, \quad g_{22} = -a^2 r^2, \quad g_{33} = -a^2 r^2 \sin^2\theta,$$

intendendosi  $x_0 = ct$ ,  $x_1 = r$ ,  $x_3 = \theta$ ,  $x_4 = \phi$ . Essendo la metrica diagonale, le componenti con-

trovarianti del tensore metrico sono gli inversi:

$$(A1.7b) \quad g^{00} = 1, \quad g^{11} = -\frac{1 - \kappa r^2}{a^2}, \quad g^{22} = -\frac{1}{a^2 r^2}, \quad g^{33} = -\frac{1}{a^2 r^2 \sin^2 \theta}.$$

Il calcolo di  $R_{\mu\nu}$  può essere un po' semplificato tramite argomenti, analoghi a quelli che permettono di porre il tensore  $T_{\mu\nu}$  nella forma (A1.3), che semplificano il calcolo delle componenti  $R_{ij}$  ( $i, j = 1, 2, 3$ ).

Preliminarmente procediamo al calcolo delle componenti necessarie per i simboli di Christoffel  $\Gamma_{ij}^k$ , che sono quelle in cui almeno uno degli indici è 0:

$$(G1) \quad \Gamma_{00}^0 = (1/2)g^{00} \partial_0 g_{00} = 0 \quad e \quad \Gamma_{z0}^0 = (1/2)g^{00} \partial_z g_{00} = 0 \\ \text{in quanto } g_{00} = 1$$

$$(G2) \quad \Gamma_{z\beta}^0 = -(1/2)g^{00} \partial_0 g_{z\beta} \quad e, \text{ quindi :} \\ \Gamma_{z\beta}^0 = 0, \text{ se } \alpha \neq \beta, \\ \Gamma_{11}^0 = a\dot{a}/c(1 - \kappa r^2), \quad \Gamma_{22}^0 = a\dot{a} r^2/c, \quad \Gamma_{33}^0 = a\dot{a} r^2 \sin^2 \theta/c$$

$$(G3) \quad \Gamma_{0\beta}^\alpha = (1/2)g^{\alpha\gamma} \partial_0 g_{z\beta} \quad e, \text{ quindi :} \\ \Gamma_{0\beta}^\alpha = 0 \text{ se } \alpha \neq \beta, \Gamma_{0z}^\alpha = -(1/c)(\dot{a}/a) \quad (\alpha = 1, 2, 3)$$

$$(G4) \quad \text{infine :} \quad \Gamma_{0z}^0 = \frac{1}{2}g^{00}(\partial_0 g_{0z} + \partial_z g_{00} - \partial_0 g_{0z}) = 0, \\ \Gamma_{00}^\alpha = g^{\alpha\gamma} \partial_0 g_{0z} - \frac{1}{2}g^{\alpha\gamma} \partial_z g_{00} = 0.$$

CALCOLO DELLE COMPONENTI DI  $R_{ij}$  ( $i, j = 1, \dots, 3$ ). Per queste componenti conviene separatamente due addendi: nel primo, che chiamiamo  ${}^1R_{ij}$ , sono coinvolte solo componenti di  $g_{\mu\nu}$  e  $\Gamma_{\nu\rho}^\mu$  con  $\mu, \nu, \rho$  da 1 a 3. Nel secondo, che chiamiamo  ${}^0R_{ij}$ , entreranno i termini coinvolgenti l'indice 0.

Considerazioni di simmetria permettono allora per provare che

$${}^1R_{ij} = -2\kappa g_{ij}/a^2.$$

Si calcola poi

$${}^0R_{ii} = -\partial_0 \Gamma_{ii}^0 - \Gamma_{0j}^j \Gamma_{ii}^0 = -(\ddot{a}/a + 2\dot{a}^2/a^2)g_{ii}/c^2$$

mentre è  ${}^0R_{ij} = 0$  se  $i \neq j$ . Pertanto, in generale, risulta:

$$(A1.8) \quad R_{ij} = {}^0R_{ij} + {}^1R_{ij} = -\frac{1}{c^2} \left( \frac{\ddot{a}}{a} + 2\frac{\dot{a}^2}{a^2} \right) g_{ij} - 2\frac{\kappa}{a^2} g_{ij}$$

CALCOLO DI  $R_{00}$ . Per questa componente del tensore di Riemann, sarà:

$$(A1.9) \quad R_{00} = \partial_0 \Gamma_{0z}^\alpha - \partial_z \Gamma_{00}^\alpha + \Gamma_{0z}^\beta \Gamma_{0\beta}^\alpha - \Gamma_{z\beta}^\alpha \Gamma_{00}^\beta = \frac{3}{c^2} \left( \frac{\dot{a}^2}{a^2} - \frac{\ddot{a}}{a} \right) - \frac{3}{c^2} \frac{\dot{a}^2}{a^2} = -\frac{3}{c^2} \frac{\ddot{a}}{a}$$

LE EQUAZIONI DI FRIEDMANN. Usfruiamo ora dell'espressione (A1.3) per il calcolo delle componenti diagonali di

$$S_{\mu\nu} = T_{\mu\nu} - \frac{1}{2} T g_{\mu\nu}$$

Risulta, anzitutto:

$$T = T_0^0 + T_i^i = \rho - 3p,$$

in quanto  $g_i^i = \delta_i^i = 3$ . Pertanto:

$$(A1.10a) \quad S_{00} = \rho - \frac{1}{2}(\rho - 3p) = \frac{1}{2}(\rho + 3p),$$

$$(A1.10b) \quad S_{ii} = \left[ -p - \frac{1}{2}(\rho - 3p) \right] g_{ii} = -\frac{1}{2}(\rho - p)g_{ii}$$

L'espressione (A1.10a), assieme all'espressione (A1.9), ci dà:

$$(A1.11) \quad -\frac{3}{c^2} \frac{\ddot{a}}{a} = \frac{1}{2} \frac{8\pi}{c^2} G(\rho + 3p) \quad \text{e, quindi:} \quad -\frac{\ddot{a}}{a} = \frac{4\pi}{3} G(\rho + 3p).$$

L'espressione (A1.10b), assieme all'espressione (A1.8), ci dà invece:

$$\frac{\ddot{a}}{a} + 2 \frac{\dot{a}^2}{a^2} + 2 \frac{\kappa}{a^2} = 8\pi G \frac{1}{2} (\rho - p).$$

Sostituendovi poi l'eq. (A1.11), si ricava:

$$2 \frac{\dot{a}^2}{a^2} = \frac{4\pi}{3} G(3\rho - 3p) + \frac{4\pi}{3} G(\rho + 3p) - 2 \frac{\kappa}{a^2}$$

da cui:

$$(A1.12) \quad \left( \frac{\dot{a}}{a} \right)^2 = \frac{8\pi}{3} G\rho - \frac{\kappa}{a^2}$$

Le eq. (A1.11) e (A1.12) si riconducono facilmente alle equazioni di Friedmann nella forma delle eq. (III.21) di questo capitolo, in quanto  $H = \dot{a}/a$ ,  $Hq = \ddot{a}/a$ .

## Appendice A2

### Altre forme delle equazioni di Friedmann

Come conseguenza delle equazioni di Einstein e dell'identità di Bianchi, il tensore energia-impulso deve soddisfare alla relazione

$$(A2.1) \quad T^{\mu}_{\nu;\mu} = 0,$$

in cui il simbolo; indica la derivazione covariante. Usufruento dell'espressione (A1.3) del tensore energia-impulso, questa equazione si può porre nella forma

$$(A2.2) \quad d(\rho a^3) = -p d(a^3),$$

evocativa dell'espressione del primo principio della termodinamica, in assenza di scambi di calore,  $dU = -p dV$  (vedi eq. III.24, nel testo del capitolo).

Questa equazione può anche essere derivata direttamente dalle eq. (A1.11) e (A1.12). Alternativamente, come faremo in questa appendice, si può ricavare eq. (A1.11) dalle eq. (A2.2) e (A1.12), partendo da quest'ultima nella forma

$$\dot{a}^2 = (8\pi/3)G\rho a^2 - \kappa,$$

e derivandola rispetto al tempo; in questo modo si ottiene

$$\dot{a}\ddot{a} = \frac{4\pi}{3}G \left[ -\frac{\dot{a}}{a^2}(\rho a^3) - \frac{1}{a} \frac{d}{dt}(\rho a^3) \right]$$

e, facendo uso nell'ultimo termine di eq. (A2.2), si ha

$$\dot{a}\ddot{a} = \frac{4\pi}{3}G \left[ -\frac{\dot{a}}{a^2}(\rho a^3) - \frac{1}{a}p \frac{d}{dt}(a^3) \right] = \frac{4\pi}{3}G\dot{a}[-\rho a - 3pa].$$

Questa relazione, divisa per  $\dot{a}a$  e cambiata di segno coincide con l'eq. (A1.11).

Nell'eq. (A2.2) non è direttamente coinvolto il tempo. Se l'equazione di stato può essere posta nella forma  $p = w\rho$ , con  $w$  indipendente dal tempo, essa permette di ricavare  $\rho(a)$ .

Di questa equazione si possono dare le forme equivalenti:

$$(A2.2a) \quad d[(\rho + p)a^3] = a^3 d\rho, \quad a^3 d\rho = -(\rho + p)d(a^3)$$

Sostituendo il differenziale con la derivata rispetto a  $t$ , l'ultima relazione dà:

$$(A2.2b) \quad \dot{\rho} = -3H(\rho + p)$$

Questa equazione mostra direttamente che, per  $w = -1$ , cioè  $p = -\rho$ , la densità  $\rho$  rimane costante al variare del fattore di scala e al trascorrere del tempo.

## Appendice A3

### Espressioni della metrica FRW nei casi $\kappa = \pm 1$

Se normalizziamo il fattore di scala così che sia  $\kappa = \pm 1$ , la parte spaziale della metrica può assumere altre forme significative. Per  $\kappa = 1$ , essa diviene

$$(A3.1) \quad d\ell^2 = a^2(t) \left( \frac{dr^2}{1-r^2} + r^2 d\Omega^2 \right)$$

e, dovendo essere  $r < 1$ , si può porre  $r = \sin \chi$ , per cui avremo:

$$dr = d\chi \cos \chi, \quad dr^2/(1-r^2) = d\chi^2$$

e, in definitiva,

$$d\ell^2 = a^2(t) (d\chi^2 + \sin^2 \chi d\Omega^2).$$

Oltre a non richiedere alcuna limitazione «a hoc» sui valori delle coordinate ( $r < 1$  in eq. A3.1), questa espressione rende pienamente conto della natura «sferica» della sezione spaziale. Infatti,  $\chi$  varia da 0 a  $\pi/2$  quando  $r$  va da 0 a 1. Se poi si prende  $\chi > \pi/2$  la variabile  $r$  e la distanza  $\ell$  ridiscendono verso 0 per raggiungerlo quando  $\chi = \pi$ : la variabile ha realizzato un giro completo attraverso la geometria sferica. Crescendo ulteriormente  $\chi$  si realizzano altri «giri» lungo l'ipersuperficie.

In modo analogo, se  $\kappa = -1$ , si può assumere  $r = \sinh \chi$  e scrivere la parte spaziale della metrica tramite

$$d\ell^2 = d\chi^2 + \sinh^2 \chi d\Omega^2.$$

Anche in questo caso sono ammessi tutti i valori di  $\chi$ .

## CAPITOLO IV

# La sostanza cosmica e l'espansione dell'universo

### 1. Le sorgenti dell'espansione

Nel precedente capitolo si sono ricavate le equazioni di Friedmann. In questo ci proponiamo di risolverle. Ciò che distingue i vari casi è l'*equazione di stato* della sostanza cosmica. Preliminare è quindi una discussione sulle possibili relazioni tra la pressione, la densità  $\rho$ , in taluni casi, la temperatura.

Per questa discussione, una prima alternativa che si pone concerne la scala di riferimento. Per esempio, non c'è dubbio che una parte significativa della sostanza cosmica attuale si componga di galassie. A loro volta le galassie si compongono di stelle, polveri, gas, buchi neri, ecc. Si può scendere ulteriormente di scala: gran parte dei sistemi contenuti in una galassia si possono a loro volta risolvere in molecole, atomi, particelle elementari. Cosa ha più senso: considerare il *gas di galassie* come sorgente di gravità nelle equazione di Friedmann o riferirsi alle distribuzioni delle particelle elementari che le compongono?

Sembrerebbe banale fare la prima scelta e, in prima battuta, ci atterremo alla soluzione banale di riferirci al gas di galassie. Avremo allora un gas di mega-particelle in moto a velocità ragguardevoli, ma comunque almeno 2 ordini grandezza sotto  $c$ , e quindi non-relativistiche. Questa risposta dà comunque per scontato che si possano *smussare* tutte le disomogeneità su scala sub-galattica, incluse non-linearità estreme, come quelle dovute ai *buchi neri*, senza implicazioni sulle scale superiori. È stato sollevato il dubbio che taluni conflitti coi dati, che portano alla necessità della cosiddetta *energia oscura* (DE: *dark energy*), siano conseguenza di questa operazione. È una problematica aperta, ma è bene anche sottolineare che non ci sono evidenze conclusive in questa direzione. Considereremo quindi la classe di sorgenti composta di *particelle non relativistiche*, senza anettere rilevanza alla scala.

L'alternativa costituita da un gas di particelle ultrarelativistiche, certamente microscopiche, ha altrettanto rilievo. Nel prossimo paragrafo mostreremo che, in entrambi i casi le equazioni di stato si possono scrivere

$$(IV.1) \quad p = w\rho,$$

con  $w$  costante, che vale 0 nel caso non relativistico e 1/3 in quello relativistico.

Tra le particelle ultrarelativistiche ci sono quelle a massa nulla, in primo luogo i fotoni del CMB. Peraltro, se si fa girare all'indietro nel tempo il film dell'espansione cosmica, si giunge a uno stato primordiale di altissima temperatura, quando era  $k_B T \gg mc^2$ , per la massa  $m$  di un gran numero di particelle. Tali particelle erano allora presenti, assieme ai fotoni, nel *brodo termico* primordiale. Si tratta di un contesto remoto dal quadro attuale del cosmo, ma tutt'altro che irrilevante ai fini della sua comprensione. A tale contesto si applica ancora un'equazione di stato di forma (IV.1) con  $w = 1/3$ .

Nella trattazione cosmologica, una sostanza con  $w = 0$ , che i relativisti indicavano come *polvere*, è spesso chiamata *materia*, tralasciando di precisare *non relativistica*. Una sostanza con  $w = 1/3$  viene denominata *materia radiativa* o semplicemente *radiazione*.

Un'ovvia generalizzazione, rispetto ai casi  $w = 0, 1/3$ , è quella di trattare valori arbitrari di  $w$ , in particolare valori negativi fino a  $w = -1$ . In questo capitolo daremo le soluzioni per il caso di  $w$  arbitrario, discutendo brevemente del loro possibile significato.

Studieremo più diffusamente, invece, l'andamento dell'espansione cosmica durante *transizioni di fase*, dopo aver rivisto la fisica delle transizioni di fase in ambito cosmologico. Si tratta di un argomento divenuto di notevole interesse dopo la scoperta delle transizioni di fase a alta energia, che presiedono alle transizioni elettrodebole e di grande unificazione.

## 2. Distribuzioni di particelle

In questo paragrafo di proponiamo di legare le proprietà microscopiche delle particelle all'equazione di stato della sostanza che esse vengono a costituire. Così mostreremo che, per la materia non relativistica, è  $w = 0$  e, per la materia ultrarelativistica, è  $w = 1/3$ . Sottolineiamo sin d'ora che ciò vale, *indipendentemente da come sono distribuiti gli impulsi delle particelle* purché la distribuzione sia isotropa e omogenea.

Ciò è specificato nella distribuzione  $f(\mathbf{x}, \mathbf{p}, t)$  delle particelle nello *spazio delle fasi*: uno spazio 6-dimensionale le cui coordinate sono i possibili valori delle coordinate spaziali ( $\mathbf{x}$ ) e degli impulsi ( $\mathbf{p}$ ) delle particelle. Un elemento di volume dello spazio delle fasi è dato dal prodotto

$$(IV.2) \quad dx_1 dx_2 dx_3 \times dp_1 dp_2 dp_3 \equiv d^3x d^3p$$

mentre

$$(IV.3) \quad dN = f(\mathbf{x}, \mathbf{p}, t) d^3x d^3p$$

è il numero di particelle che al tempo  $t$  si trovano nel volume  $d^3x$  attorno a  $\mathbf{x}$ , con impulso prossimo a  $\mathbf{p}$ , entro limiti fissati da  $d^3p$ . Per intervalli di coordinate ( $dx_i$ ) o di impulso ( $dp_i$ ) piccoli, questa definizione è ben approssimata. *Essa diviene esatta per intervalli infinitesimi.*

Una distribuzione è omogenea quando  $f$  non dipende dal sito  $\mathbf{x}$ . Essa è isotropa se  $f$  non dipende dalla direzione:  $f$  può quindi dipendere solo dal modulo  $p$  dell'impulso. *Attenzione: il simbolo  $p$  indica sia la pressione che il modulo dell'impulso; sarà necessario specificare a cosa ci si riferisce, ogni volta che il senso di  $p$  non fosse chiaro dal contesto.*

Utilizzando le distribuzioni nello spazio delle fasi si dimostra:

- 1) Per *qualunque* distribuzione omogenea e isotropa di particelle relativistiche sarà  $p = \rho c^2/3$ .
- 2) Per *qualunque* distribuzione omogenea e isotropa di particelle non relativistiche sarà  $p \simeq 0$ .
- 3) Se le distribuzioni hanno la forma

$$(IV.4) \quad f(p) = \frac{\mathcal{N}_s}{(2\pi\hbar)^3} \frac{1}{e^{E(p)/k_B T} \pm 1} \quad \text{con} \quad E^2(p) = (pc)^2 + (mc^2)^2,$$

( $\hbar$ : costante ridotta di Planck;  $k_B$ : costante di Boltzmann) in cui

$\mathcal{N}_s$ : numero di *stati di spin* della particelle considerate;

$m$ : loro massa,

ne risultano espressioni semplici della pressione  $p$ , della densità d'energia  $\rho$  e della densità in numero  $n$  in funzione della *temperatura*  $T$ .

- 4) In questo caso, la dipendenza dal tempo di  $p$ ,  $\rho$ ,  $n$  sarà esclusivamente determinata da come la temperatura  $T$  dipende da  $t$ .

Le distribuzioni (IV.4), dette di *Fermi* e di *Bose*, valgono in equilibrio termodinamico per particelle di spin semi-intero (col segno +) o intero (col segno -), sotto la condizione che tra esse non si esercitino forze, salvo quelle che entrano in gioco in collisioni *istantanee*. Per esempio: in equilibrio, i fotoni (spin 1) hanno distribuzione di Bose, gli elettroni (spin 1/2) hanno distribuzione di Fermi.

A partire dalle distribuzioni (IV.4) e usando unità con  $\hbar$ ,  $c$ ,  $k_B = 1$ , per un gas di particelle relativistiche in equilibrio a temperatura  $T$ , miscela di diverse specie di particelle con  $m \ll T$ , si ricavano le espressioni (vedi appendice A1)

$$(IV.5) \quad n = \frac{\zeta(3)}{\pi^2} \tilde{g} T^3 \quad \text{con} \quad \tilde{g} = \mathcal{N}_{bos} + \frac{3}{4} \mathcal{N}_{fer}$$

$$(IV.6) \quad \rho = 3p = \frac{\pi^2}{30} g T^4 \quad \text{con} \quad g = \mathcal{N}_{bos} + \frac{7}{8} \mathcal{N}_{fer}$$

( $\zeta(3)$  è uno dei *numeri di Riemann* e, con buona approssimazione, vale 1.20);  $\mathcal{N}_{bos,fer}$  sono i numeri di stati di spin complessivi attribuibili a particelle di natura bosonica, fermionica. Esplicitando  $\hbar$ ,  $c$ ,  $k_B$ , si ha invece:

$$(IV.7) \quad n = \frac{\zeta(3)}{\pi^2} \tilde{g} \left( \frac{k_B T}{\hbar c} \right)^3, \quad \rho c^2 = 3p = \frac{\pi^2}{30} g \frac{(k_B T)^4}{(\hbar c)^3}$$

Il rapporto densità d'energia/numero è l'energia media per particella:

$$(IV.8) \quad \begin{aligned} \langle \varepsilon \rangle_{bosoni} &= \frac{\pi^4}{\zeta(3)30} k_B T \simeq 2.7 k_B T, \\ \langle \varepsilon \rangle_{fermioni} &= \frac{7}{6} \frac{\pi^4}{\zeta(3)30} k_B T \simeq 3.15 k_B T. \end{aligned}$$

Le relazioni (IV.7) valgono in particolare per il CMB, residuo del brodo termico primordiale, contenente solo fotoni, che sono bosoni con due possibili stati di spin. Densità d'energia e numero dei fotoni del CMB sono dunque

$$(IV.9) \quad \rho_\gamma = \frac{\pi^2 (k_B T)^4}{15 (\hbar c)^3} \simeq 4.8 \times 10^{-34} \left( \frac{T}{2.73 \text{ K}} \right)^4 \frac{\text{g}}{\text{cm}^3},$$

$$(IV.10) \quad n_\gamma = \frac{2\zeta(3)}{\pi^2} \left( \frac{k_B T}{\hbar c} \right)^3 \simeq 412 \left( \frac{T}{2.73 \text{ K}} \right)^3 \text{ cm}^{-3}.$$

Nel caso di particelle non relativistiche, cioè per temperature  $k_b T \ll mc^2$ , le distribuzioni (IV.4) cessano di avere differenze sostanziali, in quanto è

$$(IV.11) \quad \begin{aligned} E^2(p) &\simeq mc^2 + p^2/2m \gg k_B T && \text{e, quindi, sarà} \\ \frac{1}{e^{E(p)/k_B T} \pm 1} &\simeq e^{-E(p)/k_B T} \simeq e^{-mc^2/k_B T} e^{p^2/2mk_B T}. \end{aligned}$$

Ne seguono le espressioni della densità in numero,

$$(IV.12) \quad \begin{aligned} n = \mathcal{N}_s \left( \frac{mT}{2\pi} \right)^{3/2} e^{-E(p)/k_B T} &= \mathcal{N}_s \left( \frac{mk_B T}{2\pi \hbar^2} \right)^{3/2} e^{-mc^2/k_B T}; \\ \uparrow & \qquad \qquad \qquad \uparrow \\ (\hbar = c = k_B = 1) & \qquad \qquad \qquad (\text{costanti esplicitate}) \end{aligned}$$

la densità d'energia è semplicemente  $\rho = mc^2 n$  e la pressione è trascurabile.

Nelle espressioni di questo paragrafo non si è sinora considerata la presenza di *potenziali chimici* legati a *numeri quantici conservati*. Nel caso relativistico cosmologico, i contributi dei potenziali chimici sono trascurabili.

Il numero delle particelle non relativistiche in equilibrio, invece, è fortemente depresso dal fattore esponenziale. Per esempio, a temperature  $\sim 1$  MeV, molto sotto la massa dei barioni  $m_b \simeq 940$  MeV, nel cosmo dovranno esserci protoni e neutroni, portatori di *numero barionico*  $B$ , un numero quantico conservato. Se ne tiene conto modificando l'espressione della densità nella forma

$$(IV.13) \quad n = \mathcal{N}_s \left( \frac{mT}{2\pi} \right)^{3/2} e^{-(m-\mu)/T},$$

$\mu$  essendo il potenziale chimico legato a  $B$ . In pratica, il valore di  $\mu$  si deduce proprio da eq. (IV.13), imponendo alla densità in numero  $n$  di eguagliare la densità di numero barionico  $n_b$  imposta dalla conservazione di  $B$ .

### 3. Entropia e numero nell'espansione cosmica

Accanto a densità d'energia, pressione e numero, c'è un'ulteriore variabile termodinamica che si dimostrerà critica nel permetterci di seguire l'evoluzione del cosmo, la *densità di entropia*. A ciascuna distribuzione di particelle corrisponde un valore dell'entropia. L'entropia è dunque una funzione di stato definita anche per configurazioni non di equilibrio. Quando la distribuzione è di equilibrio termodinamico, l'entropia delle particelle assume il massimo valore.

In appendice A2 si dimostra che la densità d'entropia, in equilibrio termodinamico, ha la seguente espressione:

$$(IV.14) \quad \sigma = \frac{\rho + p}{T} = \frac{4}{3} \frac{\pi^2}{30} g T^3$$

(vedi eq. IV.6). Questa espressione rapportata alla densità in numero dà

$$(IV.15) \quad \frac{\sigma}{n} = \frac{4\pi^4}{90 \zeta(3)} \frac{g}{\tilde{g}}$$

e il rapporto ha una blanda dipendenza dalla temperatura, legata unicamente alla diversa definizione di  $\tilde{g}$  e  $g$  (eq. IV.5-IV.6). Questa equazione si applica anzitutto al brodo primordiale, quando la temperatura  $T$  è maggiore della massa di molte particelle elementari. Tra le particelle in equilibrio termodinamico, con  $m \ll T$ , alcune saranno di natura fermionica (per esempio: elettroni, neutrini, i quark dei vari *colori* e *sapori*, ecc.), altre di natura bosonica (con spin=1: i *bosoni di gauge*; con spin=0, i *bosoni di Higgs*). Di grande importanza potrebbero poi essere i partner supersimmetrici delle varie particelle.

Tutte queste particelle hanno valori diversi della massa. Via via che  $T$  cala, il suo valore passerà al di sotto della massa di determinate particelle che, quindi, usciranno dal brodo termico. Sia  $g$  che  $\tilde{g}$  caleranno quindi simultaneamente. La sottile differenza tra  $g$  e  $\tilde{g}$  (un fattore 7/8 rispetto a 3/4, davanti a  $\mathcal{N}_f$ ) sarà l'unica possibile ragione di variazione nel rapporto  $\sigma/n$ . Se si tralascia questa variazione, si può utilmente affermare che la densità di entropia misura il numero complessivo di particelle nel brodo termico, per unità di volume.

Si possono applicare considerazioni analoghe alla componente fotonica, a partire da quando la temperatura è scesa sotto la massa dell'elettrone ( $\simeq 0.51$  MeV) e non ci sono più particelle massive in equilibrio termodinamico coi fotoni. Il rapporto  $\sigma_\gamma/n_\gamma$ , tra entropia e numero di fotoni, vale allora  $4\pi^4/90 \zeta(3)$  e questo valore si mantiene costante fino all'epoca attuale, per i fotoni del CMB.

#### 4. Produzione di entropia nelle transizioni di fase

A alte temperature, la fisica delle interazioni fondamentali prevede *transizioni di fase*, cioè mutazioni nello stato di aggregazione nel brodo termico. In prossimità e durante le transizioni di fase la legge di espansione subisce significativi scossoni e alcuni di questi eventi hanno riflessi osservativi. Viceversa, studiando le condizioni attuali del cosmo, si mira a risalire alla dinamica delle interazioni per temperature inaccessibili in laboratorio.

Un teorema riguardante il brodo termico, in queste condizioni, riguarda la *conservazione dell'entropia* comovente

$$(IV.16) \quad S = \sigma a^3 ;$$

con  $\sigma$  dato in eq. (IV.14). Al contrario, nel corso dell'espansione l'*energia* in un volume comovente, in genere, *non* si conserva. Ciò non sorprende: l'espansione implica un lavoro.

Vi sono tuttavia casi specifici, in cui si verifica *produzione di entropia*. Un input anche abbondante di entropia si verifica appunto in associazione a *transizioni di fase del primo ordine, che abbiano luogo a temperatura diversa della temperatura di transizione  $T_c$ , dopo un periodo di sovraraffreddamento (supercooling)*. La dimostrazione è in appendice A2.

Va però sottolineato che, se una transizione di fase del primo ordine avviene a temperatura  $\simeq T_c$ , non ci sarà praticamente produzione di entropia.

#### *Stati di aggregazione e ordine della transizione di fase*

Le transizioni di fase sono uno degli eventi più comuni, non solo in laboratorio. Per esempio, la transizione da acqua liquida a acqua solida (ghiaccio) è una transizione di fase; nelle condizioni ordinarie essa definisce la temperatura di  $0^\circ\text{C}$ . Quando ci sono due possibili fasi o *stati di aggregazione* (1 e 2), le loro pressioni e densità d'energia dipendono da  $T$  in modo diverso, tramite funzioni  $p_1(T)$ ,  $p_2(T)$  e  $\rho_1(T)$ ,  $\rho_2(T)$ .

Nota la funzione  $p(T)$  e in assenza di potenziali chimici si può ricavare  $\rho(T)$  per via analitica (vedi appendice A2). Sottolineiamo che il valore di  $\rho$ , a una temperatura  $T$ , non dipende solo dal valore di  $p$  alla temperatura  $T$ , ma anche dalla *pendenza* della curva  $p(T)$ , a quella temperatura.

La transizione di fase avviene quando le curve  $p_1(T)$  e  $p_2(T)$  si intersecano, a una temperatura (di transizione)  $T_c$ . Lo stato di aggregazione stabile è quello con pressione maggiore. È facile capirlo: supponiamo che una sostanza si trovi accidentalmente nella fase a pressione minore; in essa nucleerebbero bolle dell'altra fase che, grazie alla maggiore pressione, si espandono rapidamente, finendo con l'occupare tutto lo spazio. In figura IV.1 sono rappresentati gli andamenti delle pressioni di due ipotetiche fasi della medesima sostanza, in funzione della temperatura. Con tratti lunghi è marcata la curva che rappresenta la fase stabile, diversa a seconda che  $T$  sia maggiore o minore di  $T_c$ .

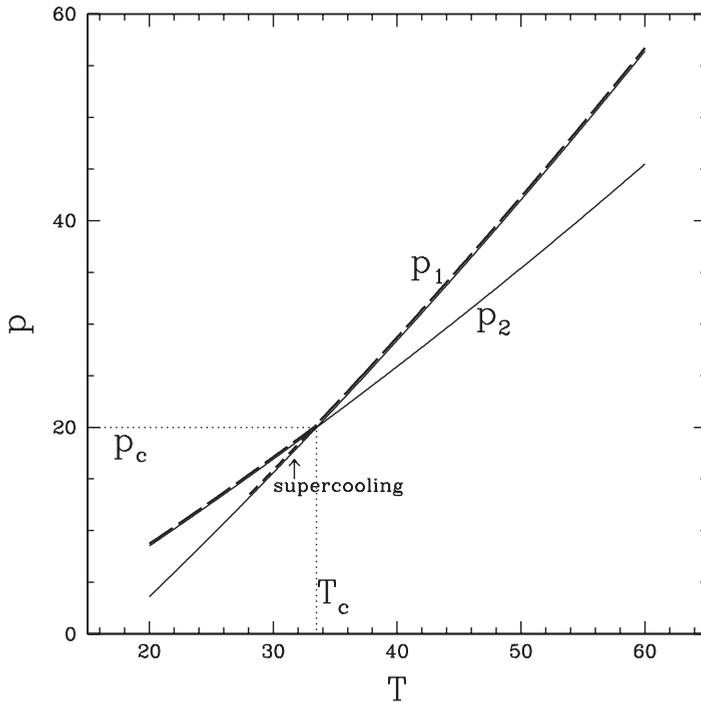


FIGURA IV.1

Andamento della pressione per le due fasi, in una transizione di fase del primo ordine. La transizione di fase, in condizioni di equilibrio, avviene alla temperatura  $T_c$ , quando la pressione è  $p_c$ ; se la transizione non avviene al punto critico, si entra in supercooling; dopo un tratto di *supercooling* la transizione avverrà comunque, ma con generazione di entropia. In ascissa e ordinata le unità sono arbitrarie.

Il grafico mostra, in particolare, che la pendenza delle due curve  $p(T)$  è diversa, alla temperatura  $T_c$ . Allora, le due fasi hanno la stessa pressione  $p_c$ , ma densità d'energia  $\rho_{1,c}$  e  $\rho_{2,c}$  diverse, com'è mostrato in figura IV.2.

Nelle figure è descritta una *transizione di fase del primo ordine*, compreso l'eventuale tratto in sovraraffreddamento (*supercooling*). Se il tratto effettivamente percorso in *supercooling* è significativo, quando la transizione finalmente avviene, in forma più o meno esplosiva, si ha produzione di entropia.

Per esempio, avremo produzione di entropia alla fine del *periodo inflazionario*. Ulteriori contributi entropici possono essere avvenuti alla transizione elettro-debole e alla transizione quark-adroni, se quest'ultima è una transizione di fase.

In una *transizione di fase del secondo ordine* le curve  $p_1(T)$  e  $p_2(T)$  hanno la stessa pendenza nel punto in cui si incrociano. Tuttavia, in tale punto la loro pendenza varia in modo diverso.

Il concetto si esprime con chiarezza in termini differenziali: una transizione di fase è d'ordine  $N$  quando, a  $T_c$ , le derivate  $N$ -sime di  $p_1(T)$  e  $p_2(T)$  sono diverse. Essa è

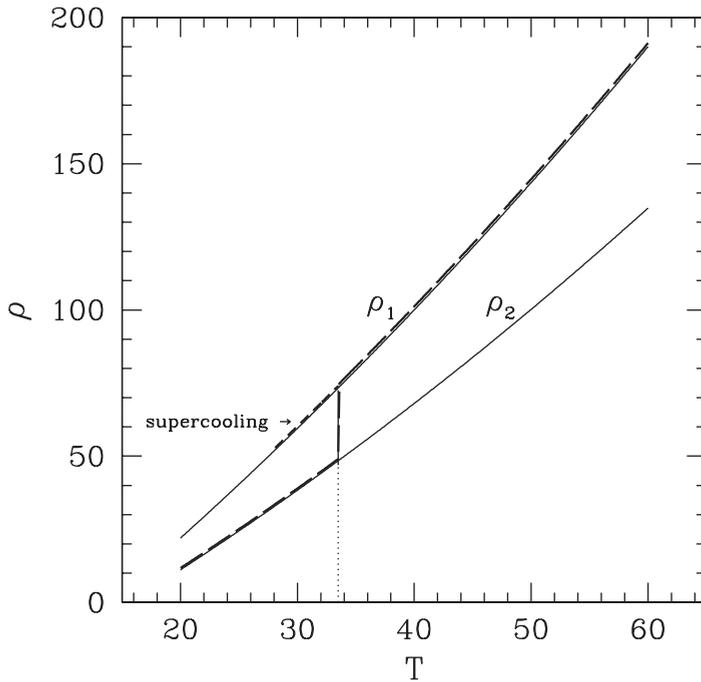


FIGURA IV.2

Andamento della densità per le due fasi in una transizione di fase del primo ordine. La transizione alla temperatura  $T_c$  non può essere istantanea; vedi testo (unità arbitrarie).

quindi di primo ordine se sono diverse le derivate prime, di secondo ordine se sono diverse le derivate seconde, e così via. Paradossalmente, se fossero eguali tutte le derivate, uno sviluppo in serie a partire da  $T_c$  mostrerebbe che le due funzioni  $p_1(T)$  e  $p_2(T)$  coincidono.

In appendice A3 è determinato l'andamento del fattore di scala durante una transizione di fase del primo ordine senza *supercooling*, per cui esiste una precisa forma analitica.

Si tratta comunque di un caso ormai abbastanza peculiare, prima del quale, nel prossimo paragrafo, saranno ricavati i regimi di espansione quando nel cosmo è presente un'unica sostanza con  $w$  costante.

## 5. Integrazione delle equazioni di Friedmann per $w$ costante

Se un'unica sostanza con  $w$  costante è responsabile dell'espansione o, comunque, essa è la componente dominante nel cosmo, l'equazione di Friedmann,

$$(IV.17) \quad d(\rho a^3) = -p d(a^3) \quad \text{diviene:} \quad d(\rho a^3) + w p d(a^3) = 0$$

e implica che sia

$$(IV.18) \quad \rho a^\alpha = \text{costante}, \quad \text{con} \quad \alpha = 3(w + 1).$$

La prova è data in appendice A4. Di conseguenza

$$(IV.19) \quad \begin{array}{ll} (i) & \text{se } w = 0 \quad \text{risulta: } \rho = \rho_0(a_0/a)^3 \\ (ii) & \text{se } w = 1/3 \quad \text{risulta: } \rho = \rho_0(a_0/a)^4. \end{array}$$

A differenza del solito,  $a_0$  qui indica semplicemente il fattore di scala a un generico tempo  $t_0$ , quando la densità ha valore  $\rho_0$ .

Questi casi meritano uno specifico commento.

1) Un andamento  $\rho \propto a^{-3}$  significa che il prodotto

$$\begin{array}{l} \text{densità} \times \text{volume} \quad \text{è costante.} \\ \propto a^{-3} \quad \quad \quad \propto a^3 \end{array}$$

L'energia (di massa) in un volume *comovente* resta dunque costante. Nelle eq. (IV.17) ciò consegue all'annullarsi della pressione  $p$ , sicché il *lavoro di espansione* è nullo; possiamo vedere la cosa anche da un altro punto di vista: il numero di particelle si conserva, mentre la loro energia individuale, data esclusivamente dalla loro massa, non varia con  $a$ .

2) L'andamento  $\rho \propto a^{-4}$  è proprio della *materia radiativa*. La conservazione del numero di particelle, che fissa la proporzionalità  $n \propto a^{-3}$ , unita all'eq. (IV.5), che fissa la proporzionalità  $n \propto T^3$ , significa che è

$$(IV.20) \quad T \times a = \text{const.}$$

In questo caso, conformemente alle eq. (IV.8) e (IV.20), risulta  $\langle \varepsilon \rangle \propto T \propto a^{-1}$ : energia media e frequenza dei fotoni, nel corso dell'espansione, calano; si tratta del *redshift*. Ne segue che  $\rho = n \langle \varepsilon \rangle \propto a^{-4}$ ; la stessa conclusione è raggiunta in eq. (IV.19) considerando il lavoro della *pressione* della sostanza radiativa.

I risultati in eq. (IV.19), immessi nella seconda equazione di Friedmann con  $\kappa = 0$ , nel caso  $w = 0$  (tramite inevitabili passaggi differenziali) danno

$$(IV.21) \quad \dot{a}/a = H_0(a_0/a)^{3/2} \quad \text{dove} \quad H_0^2 = 8\pi G\rho_0/3$$

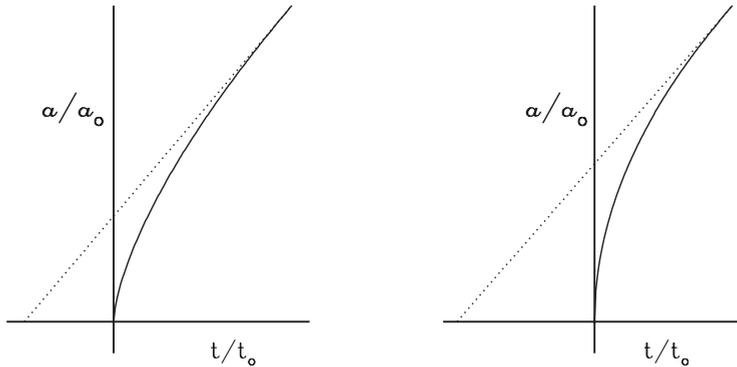
e, quindi

$$(IV.22) \quad (a/a_0)^{1/2} d(a/a_0) = H_0 dt \quad \text{da cui:} \quad (a/a_0)^{3/2} = 3t H_0/2$$

Se ne conclude che, per  $w = 0$ , varranno le relazioni

$$(IV.23) \quad a/a_0 = (t/t_0)^{2/3}, \quad H_0 = 2/3t_0, \quad t_0/t = (1+z)^{3/2},$$

assumendo che l'espansione sia sempre dominata dalla materia (non relativistica) fino al tempo  $t_0$ . L'andamento del fattore di scala in questo caso viene ulteriormente discusso in figura IV.3.

**FIGURA IV.3**

Andamento del fattore di scala nei casi  $w = 0$  (sinistra) e  $w = 1/3$  (destra). La linea punteggiata rappresenta l'andamento che il fattore di scala avrebbe avuto se il parametro di Hubble avesse sempre il valore  $H_0$ . Conformemente alle eq. (IV.23) e (IV.25) il tempo  $t_0$  è una frazione del *tempo di Hubble*  $H_0^{-1}$  e, più precisamente,  $t_0 = (2/3)H_0^{-1}$  e  $t_0 = (1/2)H_0^{-1}$  nei due casi; lo si vede anche nelle rispettive figure, in cui la tratteggiata raggiunge l'ascissa rispettivamente a  $t/t_0 = -1/2$  e  $t/t_0 = -1$ .

Nel caso  $w = 1/3$ , analogamente, avremo

$$(IV.24) \quad \dot{a}/a = H_0(a_0/a)^2$$

e, quindi

$$(IV.25) \quad (a/a_0)d(a/a_0) = H_0 dt \quad \text{da cui:} \quad (a/a_0)^2 = 2t H_0$$

Nel caso  $w = 1/3$ , in cui l'espansione è dominata dalla *radiazione* fino a  $t_0$ , varranno dunque le relazioni

$$(IV.26) \quad a/a_0 = (t/t_0)^2, \quad H_0 = 1/2t_0, \quad t_0/t = (1+z)^2.$$

L'andamento del fattore di scala viene ulteriormente discusso in figura IV.3.

### **Altre leggi di espansione per $w$ costante**

I casi in eq. (IV.19) sono quelli con rilievo fisico più immediato, tra tutti i casi considerati in eq. (IV.18). La stessa procedura analitica realizzata per i primi due casi può essere peraltro generalizzata a  $\rho = \rho_0(a_0/a)^{3(w+1)}$ , con  $w$  generico, con la sola esclusione del caso  $w = -1$ , da trattare a parte. Posto allora

$$(IV.27) \quad x = \frac{3}{2}(1+w)$$

si integra l'equazione differenziale

$$(IV.28) \quad \dot{a}/a = H_0(a_0/a)^x, \quad \text{dove} \quad H_0 = 8\pi G\rho_0/3,$$

ottenendo

$$(IV.29) \quad a/a_o = [xH_o t]^{1/x} = 1/(z+1), \quad t_o = [xH_o]^{-1}$$

Nel caso  $w = -1$ , invece, avremo

$$(IV.30) \quad \rho \equiv \rho_o = \text{costante}$$

e quindi sarà costante lo stesso parametro di Hubble:

$$(IV.31) \quad H^2 = \frac{8\pi}{3} G\rho_o = H_o^2 = \text{costante}$$

L'equazione differenziale

$$(IV.32) \quad \dot{a} = Ha$$

ha allora integrale

$$(IV.33) \quad a(t) = a_o e^{H(t-t_o)} :$$

l'espansione cosmica avviene con legge esponenziale.

### ***Espansione «dominata» dalla curvatura***

Tra le leggi qui considerate c'è anche quella che corrisponde a  $w = -1/3$ . In questo caso, vista eq. (IV.27), risulta  $x = 2$  e la seconda equazione di Friedmann assume l'aspetto

$$(IV.34) \quad H^2 = \frac{8\pi}{3} G\rho_o \left(\frac{a_o}{a}\right)^2 = \frac{K^2}{a^2}$$

Confrontiamo allora questa equazione con l'equazione di Friedmann nel caso  $\kappa < 0$  ( $\Omega < 1$ ) ricordando che

$$(IV.35) \quad \frac{\kappa c^2}{a_o^2} = -H_o^2(1 - \Omega_o),$$

conformemente a eq. (24) del precedente capitolo. Tale equazione assume in generale l'aspetto

$$(IV.36) \quad H^2 = H_o^2 \Omega_o \left(\frac{a_o}{a}\right)^3 + H_o^2(1 - \Omega_o) \left(\frac{a_o}{a}\right)^2$$

se la sostanza cosmica si compone di particelle non relativistiche.

Si vede allora che il secondo addendo del secondo membro dell'equazione prevale sul primo per

$$(IV.37) \quad \frac{a_o}{a} = 1 + z < \Omega_o^{-1} - 1$$

Per esempio, per  $\Omega_0 = 0.2$ , prevale il secondo addendo fino a  $z \simeq 3$ . Se, in questo regime, si trascuria il primo addendo, l'eq. (IV.36) assume la forma (IV.34).

L'eq. (IV.34) descrive dunque il comportamento del fattore di scala sia nel caso di una sostanza con parametro di stato  $w = -1/3$ , che nel caso di prevalenza del termine di curvatura.

Essendo poi  $H^2 = (\dot{a}/a)^2$ , eq. (IV.34) diviene

$$(IV.38) \quad \dot{a} = K :$$

la velocità di espansione è costante,

$$(IV.39) \quad a = Kt ;$$

si tratta del regime in cui il cosmo si espande senza sentire l'effetto della gravità della sostanza in esso contenuta. Tornando all'analogia col moto dei gravi, il caso in questione è l'analogo del moto di un grave con energia positiva che si sia liberato definitivamente dall'attrazione terrestre.

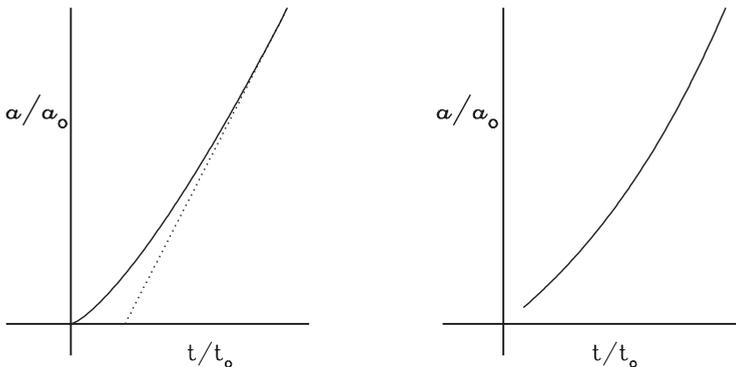
### ***Espansione decelerata e accelerata***

Il caso  $w = -1/3$ ,  $x = 2$  è anche il caso di confine tra regimi di espansione decelerata e accelerata.

Ricordiamo infatti che, conformemente a eq. (20) del capitolo precedente, il parametro di decelerazione

$$(IV.40) \quad q = \frac{1 + 3w}{2} \Omega$$

assume valori positivi o negativi a seconda che sia  $w > -1/3$  o  $w < -1/3$ .



**FIGURA IV.4**

Andamento del fattore di scala nei casi  $w = -1/2$  (sx) e  $w = -1$  (dx). La linea tratteggiata nel grafico a sx, rappresenta l'andamento che il fattore di scala avrebbe avuto se il parametro di Hubble avesse sempre il valore  $H_0$  assunto al tempo  $t_0$ . Il grafico a sx termina al tempo  $t_0$ ; in quello di dx, invece, l'espansione esponenziale  $a(t) = a_0 e^{H(t-t_0)}$  è supposta iniziare al tempo  $t_0$ .

Ciò chiarisce ulteriormente la situazione: l'espansione è decelerata se il parametro di stato  $w$  è  $> -1/3$ . Nella figura IV.3 la concavità è allora rivolta verso il basso e diminuisce al calare di  $w$ . Per  $w = -1/3$ , il grafico dell'espansione sarebbe dato da una retta: esso ricalcherebbe le tratteggiate in figura. Al di sotto di  $w = -1/3$ , la concavità si rivolge verso l'alto fino a avere una crescita addirittura esponenziale nel caso  $w = -1$  (vedi Fig. IV.4).

## 6. Integrazione delle equazioni di Friedmann in presenza di più componenti

Considereremo ora due casi in cui sono presenti simultaneamente due componenti diverse, di densità confrontabile.

L'equazione (IV.18) esprime l'andamento con fattore di scala di ciascuna componente, anche quando vi siano più componenti simultaneamente presenti *a condizione che tra loro non vi sia interazione*. I problemi che ci porremo riguardano quindi solo la determinazione di  $a(t)$ , essendo già note le dipendenze delle diverse densità  $\rho$  dal fattore di scala.

### $w = 0$ e $w = -1$

In questo caso la seconda equazione di Friedmann assume la forma

$$(IV.41) \quad H^2 = H_o^2[\Omega_{om}(a_o/a)^3 + \Omega_{oA}]$$

in cui il suffisso  $_o$  si riferisce a un tempo determinato, che potrebbe essere l'epoca attuale.  $\Omega_{om}$  e  $\Omega_{oA}$  sono, rispettivamente, i parametri di densità delle componenti con  $w = 0$  e  $w = -1$ , come del resto è evidente dall'andamento delle rispettive densità col fattore di scala.

In appendice A6 si dimostra che, da questa equazione, ponendo

$$(IV.42) \quad c = (\Omega_{oA}/\Omega_{om})^{1/2}(a/a_o)^{3/2},$$

si ricava l'andamento del fattore di scala col tempo, nella forma

$$(IV.43) \quad \ln(\sqrt{1+c^2} + c) = (3/2)\Omega_{oA}^{1/2}H_o t.$$

Da tale espressione generale si estrapolano due regimi:

1) Per  $c \gg 1$ , ovvero  $a \gg a_o$ :

$$(IV.44) \quad a/a_o \simeq \exp(\Omega_{oA}^{1/2}H_o t)$$

2) Per  $c \ll 1$ , ovvero  $a \ll a_o$ :

$$(IV.45) \quad (a/a_o)^{3/2} = (3/2)\Omega_{om}^{1/2}H_o t$$

Si noti che ciascuno dei due regimi risente solo del parametro di densità della sostanza che allora è dominante.

Formalmente le soluzioni valgono per qualsiasi valore di

$$(IV.46) \quad \Omega_o = \Omega_{om} + \Omega_{oA},$$

ma hanno senso fisico solo per  $\Omega_o = 1$ . Se interessa un caso con  $\Omega_o \neq 1$ , si deve aggiungere sulla destra dell'equazione (IV.41) il termine che tiene conto della curvatura, ottenendo un'equazione di cui non si conosce una soluzione analitica.

Un'esame numerico, tuttavia, dimostra che l'espansione diviene esponenziale anche per  $\Omega_o > 1$ , senza che a essa segua, come nei modelli *classici*, dominati dalla materia non-relativistica, una fase di contrazione.

Infine, l'eq. (IV.42) mostra che, al tempo attuale  $t_o$ , avremo per  $c$  il valore

$$(IV.47) \quad c_o = (\Omega_{oA}/\Omega_{om})^{1/2}$$

da cui ricaviamo l'espressione

$$(IV.48) \quad t_o = \frac{2}{3H_o} \frac{\Omega_{oA}^{1/2}}{\ln(\sqrt{1 + \Omega_{oA}/\Omega_{om}} + \Omega_{oA}^{1/2}/\Omega_{om}^{1/2})}.$$

Al limite  $\Omega_{oA} \rightarrow 0$ ,  $\Omega_{om} \rightarrow 1$  essa restituisce la relazione  $t_o = 2/3H_o$ , che vale per modelli critici in cui sia presente solo materia.

### $w = 0$ e $w = 1/3$

Se abbiamo due componenti con densità confrontabili  $\rho_m \propto a^{-3}$  e  $\rho_r \propto a^{-4}$ , definiamo  $a_{eq}$ , valore del fattore di scala di *equivalenza*, quando esse sono esattamente eguali.

In appendice A5 si mostra che, attorno all'equivalenza vale l'espressione

$$(IV.49) \quad t = \frac{2}{3\tilde{H}_{eq}} [(1 + R)^{1/2}(R - 2) + 2] \quad \text{con} \quad R = a/a_{eq}.$$

e  $\tilde{H}_{eq}^2 = (8\pi/3)G\rho_{m,eq} = (8\pi/3)G\rho_{r,eq}$  ( $\rho_{m,eq}$  e  $\rho_{r,eq}$  sono le densità delle due componenti considerate all'equivalenza). L'inversione di questa espressione, per ricavare  $a(t)$  implica la risoluzione di un'equazione algebrica di terzo grado.

## 7. Orizzonti

Il concetto di orizzonte può apparire piuttosto semplice: Se il cosmo esiste da un tempo  $t_o$ , nessun segnale luminoso può aver percorso una distanza superiore a  $ct_o$ . Se non vi sono segnali capaci di propagarsi a velocità superluminale, come richiede la Relatività, ne segue che il sito cosmico occupato dall'osservatore non può essere mai stato in contatto causale con le regioni del cosmo situate a distanza  $> ct_o$ . In definitiva, il nostro sito cosmico può aver subito l'influenza di eventi verificatisi entro una sfera di raggio  $ct_o$ , centrata su di noi, mentre ogni altro evento nel cosmo non lo può avere influenzato.

Questa argomentazione, in realtà, potrebbe essere grossolanamente fuorviante. Essa non tiene conto del fatto che, finché la luce viaggiava, il cosmo stava espandendosi. Se la distanza attuale tra due determinati punti è  $D_o$ , a un tempo  $t$  nel passato essa era  $D(t) = D_o a(t)/a_o$  e quindi minore di  $D_o$ . Il tempo impiegato dalla luce a coprire quella distanza è quindi stato  $< D_o/c$ . In altri termini, nel tempo  $D_o/c$  il segnale luminoso ha potuto coprire una distanza significativamente maggiore di quella che c'era tra i due punti che oggi si trovano a distanza  $D_o$ .

Segnali luminosi possono quindi essere giunti all'osservatore anche da distanze  $\gg ct_o$ . L'argomentazione, in linea di principio, mette persino in dubbio la stessa esistenza di un orizzonte, nel senso anzidetto. È chiaro che, in sostanza, tutto il problema è rimandato a una specie di competizione: *al crescere del fattore di scala crescono più velocemente le distanze o i tempi?* Ovvero, se è

$$(IV.50) \quad a \propto t^\alpha$$

è  $\alpha < 1$  o  $\alpha > 1$ ? Conformemente alle eq. (IV.27) e (IV.29), risulta

$$(IV.51) \quad \alpha = \frac{1}{x} = \frac{2/3}{1+w}$$

e avremo

$$(IV.52) \quad \begin{array}{ll} \alpha > 1 & \text{per } 1+w < 2/3 \rightarrow w < -1/3 \\ \alpha < 1 & \text{per } 1+w > 2/3 \rightarrow w > -1/3 \end{array}$$

Si preannuncia quindi l'assenza di orizzonte nel caso di espansione accelerata, con  $w < -1/3$ .

Un'analisi completa del problema richiede di seguire la progressiva crescita delle distanze al trascorrere del tempo. Il calcolo è quindi intrinsecamente differenziale e viene riportato in appendice A6. Anche quando l'espansione è decelerata, comunque, la distanza percorsa dalla luce è  $> ct_o$ . Il risultato del calcolo dà, per l'*orizzonte di particella* il valore

$$(IV.53) \quad l_p = \frac{1}{1-\alpha} ct_o = \frac{3(w+1)}{3w+1} ct_o = \frac{2}{3w+1} cH_o^{-1} \quad (\text{per } w > -1/3).$$

Per esempio:

$$\begin{array}{ll} \text{nel caso } w = 0, & \text{espansione dominata dalla materia, } l_p = 3ct_o = 2c/H_o \\ \text{nel caso } w = 1/3, & \text{espansione radiativa, } l_p = 2ct_o = c/H_o. \end{array}$$

Accanto all'orizzonte di particella va definito un *orizzonte degli eventi*: la massima distanza cui può giungere un segnale luminoso inviato da un determinato sito. Potrebbe sembrare che una tale distanza non esista: in un tempo infinitamente grande la distanza raggiunta dovrebbe essere infinita. Ebbene, ciò non è vero in caso di espansione accelerata, quando le distanze crescono più rapidamente dei tempi. Si

trova allora un orizzonte degli eventi

$$l_e = \frac{1}{\alpha - 1} ct_o = -\frac{3(w+1)}{3w+1} ct_o = -\frac{2}{3w+1} cH_o^{-1} \quad (\text{per } -1 < w < -1/3). \quad (\text{IV.54})$$

Per esempio, per  $w = -2/3$ , si ottiene  $l_e = ct_o = 2c/H_o$ .

Invece, nel caso  $w = -1$ , in cui il fattore di scala cresce con legge esponenziale:

$$a(t) = a_o \exp[H(t - t_o)] \quad \text{con } H \text{ costante}$$

abbiamo:

orizzonte degli eventi: finito, costante nel tempo  $l_e = c/H$ ,

orizzonte di particella: finito, crescente nel tempo  $l_p = (c/H)[\exp(Ht) - 1]$ .

Non abbiamo considerato in questo capitolo valori  $w < -1$ , presi in considerazione da qualche autore nella descrizione dell'energia oscura. Per esempio, nell'analisi cosmologica della massa del neutrino, valori  $w < -1$  permettono opzioni altrimenti inconsistenti coi dati.

Ciò è in relazione a quanto segue: in questo capitolo abbiamo fornito una base fisica per parametri di stato  $w = 0$  e  $w = 1/3$ . Nei prossimi capitoli mostreremo che parametri di stato  $w \geq -1$  sono derivabili nell'ambito di precise teorie fisiche, per esempio se la sostanza cosmica ha una componente dovuta a un campo scalare autointeragente. Invece, a tutt'oggi, è assai problematico fornire una teoria fisica ragionevole per la sostanza cosmica, che dia luogo a un parametro di stato  $w < -1$ .

Consideriamo infine l'andamento dell'orizzonte attorno all'*equivalenza*, come definita nel precedente paragrafo, cioè al trapasso tra un'espansione dominata da materia radiativa a un'espansione dominata da materia non-relativistica.

In appendice A6 si dimostra che, in questo caso, l'andamento dell'orizzonte di particella è il seguente:

$$(IV.55) \quad l_p(t) = \frac{2c}{\tilde{H}_{eq}} R(t) \{ [1 + R(t)]^{1/2} - 1 \} \quad \text{con} \quad R = a/a_{eq}$$

e  $\tilde{H}_{eq}$  definito alla fine del precedente paragrafo.

Per  $R = 1$ , sarà allora

$$(IV.56) \quad l_{p,eq} = 2(\sqrt{2} - 1) \frac{c}{\tilde{H}_{eq}}$$

Queste espressioni saranno ampiamente usate nei prossimi capitoli.

## 8. Cosmografia classica

Prima che le misure effettuate nell'ultimo decennio fornissero elementi fortemente discriminanti, le alternative considerate erano assai disparate. Curiosamente, il modello che oggi appare più aderente ai dati fisici non era tenuto in grande considera-

zione. Molta attenzione veniva invece data a modelli in cui la sostanza era ordinaria materia non-relativistica, ma con curvatura  $\kappa \neq 0$ .

In quest'ambito si erano focalizzate le alternative:

- 1)  $\Omega \leq 1$  ( $\kappa \leq 0$ ), espansione senza limiti.
- 2)  $\Omega > 1$  ( $\kappa > 0$ ), espansione seguita da ricontrazione, verso il *big-crunch*.

Abbiamo invece già visto che ci sono modelli con  $\kappa > 0$  che prevedono un'espansione indefinita. In sostanza, le alternative classiche sono basate sull'assunto *a priori* che la sostanza cosmica è dovuta a particelle.

Anche per il loro rilievo storico riportiamo le soluzioni per questi casi.

### **Soluzione per modelli aperti**

Essa si scrive in forma parametrica:

$$(IV.57) \quad \begin{aligned} \frac{a(\theta)}{a_o} &= \frac{\Omega_o}{2(1 - \Omega_o)} (\cosh \theta - 1) \\ H_o t(\theta) &= \frac{\Omega_o}{2(1 - \Omega_o)^{3/2}} (\sinh \theta - \theta), \end{aligned}$$

come si può accertare per sostituzione diretta. Il fattore di scala ha il valore  $a_o$  quando il parametro  $\theta$  vale

$$(IV.58) \quad \cosh \theta_o = 2\Omega_o^{-1} - 1, \quad \text{per cui:} \quad \sinh \theta_o = \frac{2}{\Omega} (1 - \Omega_o)^{1/2}$$

e, in definitiva,

$$(IV.59) \quad H_o t_o = \frac{1}{1 - \Omega_o} - \frac{\Omega_o}{2(1 - \Omega_o)^{3/2}} \operatorname{arccosh}(2\Omega_o^{-1} - 1)$$

Sviluppando in serie di  $\Omega_o$  si ricava di qui

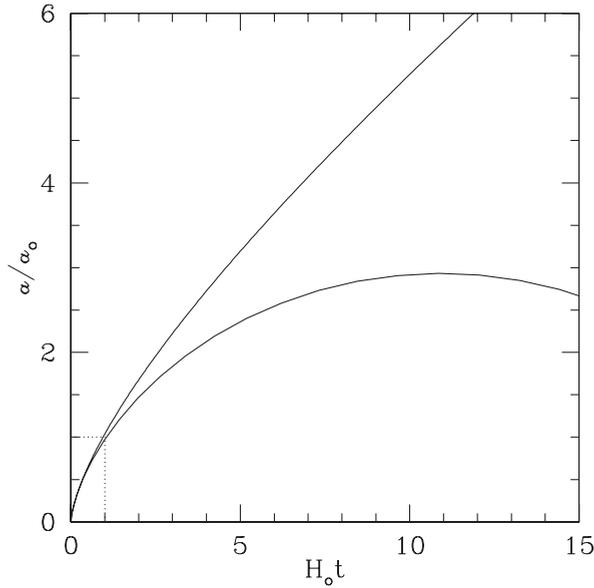
$$(IV.60) \quad H_o t_o = 1 + \Omega_o \ln \Omega_o.$$

L'interesse di questa soluzione è quello di fornire la correzione all'ordine più basso rispetto alla relazione  $H_o t_o = 1$ , trovata per  $w = -1/3$ .

### **Soluzione per modelli chiusi**

Essa si scrive in forma parametrica:

$$(IV.61) \quad \begin{aligned} \frac{a(\theta)}{a_o} &= \frac{\Omega_o}{2(\Omega_o - 1)} (1 - \cos \theta) \\ H_o t(\theta) &= \frac{\Omega_o}{2(\Omega_o - 1)^{3/2}} (\theta - \sin \theta), \end{aligned}$$

**FIGURA IV.5**

Andamento del fattore di scala in modelli chiusi ( $\Omega_0 = 1.09$ ) e aperti ( $\Omega_0 = 0.91$ ) in cui tutta la sostanza è materia non-relativistica e che si espandono alla stessa velocità a  $t \ll t_0$ . Le curve sono normalizzate in modo che un modello con  $\Omega_0 = 1$  e lo stesso ritmo di espansione a  $t \ll t_0$  venga a avere  $a(t_0) = a_0$ .

come si può accertare per sostituzione diretta. Una verifica più dettagliata si può trovare nel capitolo XI, dove si studia l'evoluzione di una sovradensità sferica, tramite un'equazione formalmente identica all'equazione di Friedmann che vale in questo caso e che qui riportiamo:

$$(IV.62) \quad H^2 = H_0^2 [\Omega_0 (a_0/a)^3 + (1 - \Omega_0) (a_0/a)^2].$$

La massima espansione sarà raggiunta per  $\theta = \pi$ ; al tempo dato dalla relazione

$$(IV.63) \quad H_0 t_m = (\pi/2) \Omega_0 (\Omega_0 - 1)^{-3/2},$$

quando il fattore di scala è dato dalla relazione

$$(IV.64) \quad a_m/a_0 = \Omega_0 / (\Omega_0 - 1).$$

Si veda anche figura IV.5 per il confronto tra questi andamenti.

# Appendice A1

## Distribuzioni statistiche: densità di energia e pressione.

Dalla distribuzione nello spazio delle fasi relativistico 8-dimensionale,  $f(x^\alpha, p_\beta)$ , si ottiene il tensore energia-impulso

$$(A1.1) \quad T_{\mu\nu}(x^\sigma) = 2\mathcal{N}_s \int d^4p p_\mu p_\nu f(x^\alpha, p_\beta) \delta(p^\rho p_\rho - m^2);$$

$\mathcal{N}_s$ : è il numero di stati di spin delle particelle, di massa  $m$ . Ricordiamo che nel riferimento di Lorentz in cui si ha isotropia e omogeneità, sarà

$$T_{\mu\nu} = \begin{vmatrix} \rho & & & \\ & -p & & \\ & & -p & \\ & & & -p \end{vmatrix}$$

L'integrale in eq. (A1.1) può essere trasformato in integrale triplo, tramite l'uso della  $\delta$  sul 4-impulso, pena la scelta di un particolare riferimento di Lorentz. A questo scopo si cambia variabile d'integrazione da  $p_o$  a  $p_o^2$ :

$$\begin{aligned} 2 \int d^3p dp_o \delta(p_o^2 - p^2 - m^2) \dots &= 2 \int d^3p \frac{1}{dp_o^2/dp} dp_o^2 \delta(p_o^2 - p^2 - m^2) \dots = \\ &= 2 \int d^3p \frac{1}{2p_o} dp_o^2 \delta(p_o^2 - p^2 - m^2) \dots = \int d^3p \frac{1}{p_o(\mathbf{p})} \dots \end{aligned}$$

ottenendo per densità d'energia  $\rho$  e pressione  $p$  (da non confondere con l'impulso)

$$(A1.2) \quad T_{00} = \rho = \mathcal{N}_s \int d^3p p_o f(\mathbf{x}, \mathbf{p}, t).$$

$$T_z^z = 3p = \mathcal{N}_s \int d^3p \frac{p_x p^x}{p_o} f(\mathbf{x}, \mathbf{p}, t);$$

quest'ultima relazione, per  $m \simeq 0$  e quindi impulso  $p^2 \simeq p_o^2$ , dà

$$(A1.3) \quad 3p = \rho,$$

(la densità d'energia è 3 volte la pressione) *indipendentemente dalla forma della distribuzione*. In generale risulta  $p_x p^x = p^2 = p_o^2 - m^2$  sicché, nel caso non-relativistico, si ottiene

$$3p \simeq \int d^3p \frac{p^2}{p_o} p_o f(\mathbf{x}, \mathbf{p}, t) = \int d^3p \beta^2 p_o(\mathbf{p}) f(\mathbf{x}, \mathbf{p}, t) = \langle \beta^2 \rangle \rho,$$

dove  $\langle \beta^2 \rangle$  è un'opportuna media sulle velocità delle particelle non-relativistiche.

Passiamo ora al caso specifico delle distribuzioni di equilibrio statistico

$$(A1.4) \quad f(p) = \frac{\mathcal{N}_s}{(2\pi)^3} \frac{1}{e^{E(p)/T} \pm 1},$$

di Bose (segno -) o di Fermi (segno +), per particelle di spin intero o semi-intero, rispettivamente. In genere è  $p_o \equiv E(p) = \sqrt{p^2 + m^2}$ . Peraltro, al limite non relativistico, quando

$E(p) \simeq m + p^2/2m \gg T$ , il termine  $\pm 1$  diviene trascurabile e entrambe le distribuzioni divengono la distribuzione di Boltzmann:

$$(A1.5) \quad f_B(p) = \frac{\mathcal{N}_s}{(2\pi)^3} e^{-m/T - p^2/2mT}$$

L'eq. (A1.4) inserita in eq. (A1.2) dà

$$\rho = \mathcal{N}_s \frac{4\pi}{(2\pi)^2} \int_0^\infty dq q^2 \frac{q}{e^{q/T} \pm 1} = \frac{\mathcal{N}_s}{2\pi^2} T^4 \int_0^\infty dx \frac{x^3}{e^x \pm 1}.$$

L'integrale numerico vale  $\pi^4/15$  e  $7\pi^4/120$ , rispettivamente per distribuzioni di Bose e di Fermi.

Per un brodo termico di particelle sarà allora

$$\rho = 3p = \frac{\pi^2}{30} g T^4 \quad \text{con} \quad g = \mathcal{N}_{bos} + \frac{7}{8} \mathcal{N}_{fer}$$

dove  $\mathcal{N}_{bos}$  e  $\mathcal{N}_{fer}$  sono i numeri di stati di spin, rispettivamente di bosoni e fermioni con massa  $m \ll T$ .

Con procedura analoga si valuta la densità in numero. Essendo

$$n = \frac{\mathcal{N}_s}{2\pi^2} \int_0^\infty dp p^2 \frac{1}{e^{p/T} \pm 1} = \frac{\mathcal{N}_s}{2\pi^2} T^3 \int_0^\infty dx \frac{x^2}{e^x \pm 1},$$

si ottiene

$$n = \frac{\zeta(3)}{\pi^2} \tilde{g} T^3 \quad \text{con:} \quad \tilde{g} = \mathcal{N}_b + \frac{3}{4} \mathcal{N}_f$$

in cui  $\zeta(3)$  è un «numero di Riemann» e vale 1.20....

Nel caso di  $T \ll m$ , quando si usa la distribuzione (A1.5), sarà:

$$(A1.6) \quad \begin{aligned} n &= \frac{\mathcal{N}_s}{2\pi^2} e^{-m/T} \int_0^\infty dp p^2 e^{-p^2/2mT} = \\ &= \frac{\mathcal{N}_s}{\sqrt{\pi}} e^{-m/T} \left(\frac{mT}{2\pi}\right)^{\frac{3}{2}} \int_0^\infty dx x^{\frac{1}{2}} e^{-x} = \frac{\mathcal{N}_s}{2} e^{-m/T} \left(\frac{mT}{2\pi}\right)^{\frac{3}{2}}, \end{aligned}$$

in quanto l'integrale numerico vale  $\sqrt{\pi}/2$ .

Nel caso in cui sia presente un potenziale chimico  $\mu$ , le precedenti espressioni si correggono facilmente sostituendo  $m$  con  $m - \mu$ .

# Appendice A2

## Entropia e altre funzioni termodinamiche

L'identità termodinamica,

$$(A2.1) \quad dU = -p dV + T dS,$$

essendo  $dU = \rho dV$ , implica la seguente espressione per la densità d'entropia:

$$(A2.2) \quad \sigma = \frac{dS}{dV} = \frac{\rho + p}{T}$$

Essa permette anche di ricavare anche un'altra espressione di  $\sigma$ . A questo scopo definiamo la funzione di stato

$$(A2.3) \quad F = U - TS$$

per cui sarà  $dF = -p dV + T dS - T dS - S dT = -p dV - S dT$ . Quindi

$$\frac{\partial F}{\partial V} = -p, \quad \frac{\partial F}{\partial T} = -S$$

e, per il lemma di Schwarz,

$$\frac{\partial}{\partial T} \left( \frac{\partial F}{\partial V} \right) = -\frac{\partial p}{\partial T} = \frac{\partial}{\partial V} \left( \frac{\partial F}{\partial T} \right) = -\frac{\partial S}{\partial V} = -\sigma.$$

ovvero

$$(A2.4) \quad \sigma = \frac{\partial p}{\partial T}.$$

Uguagliando allora le espressioni (A2.2) e (A2.4) di  $\sigma$  si ricava:

$$(A2.5) \quad \rho = -p + T \frac{\partial p}{\partial T}$$

e, data  $p(T)$ , eq. (A2.5) permette di ricavare  $\rho(T)$ . Per esempio, se è  $p = (\pi^2/90)gT^4$  (brodo termico), risulta  $p'(T) = 4p/T$  e  $\rho = -p + 4T(p/T) = 3p$ .

Più in generale, se  $p = \Phi(T)T^4$  e assumendo  $\rho = E(T)T^4$ , si ricava che

$$(A2.6) \quad E = 3\Phi + T \frac{\partial \Phi}{\partial T}.$$

## 1. Transizioni di fase. Ordine della transizione

Consideriamo una sostanza con due stati di aggregazione (fasi), le cui pressioni sono date da  $p_1(T)$  e  $p_2(T)$ . Il passaggio tra fasi si ha in genere alla temperatura  $T_c$  per cui  $p_1(T_c) = p_2(T_c)$ , definita *temperatura di transizione*.

È chiaro che non può essere

$$(A2.7) \quad \frac{d^n p_1}{dT^n}(T_c) = \frac{d^n p_2}{dT^n}(T_c)$$

per ogni  $n$ ; altrimenti, in linea di principio, sviluppando in serie di Taylor attorno a  $T_c$ , si otterrebbe che  $p_1(T) \equiv p_2(T)$ , per ogni valore di  $T$ . A partire da un determinato valore di  $n$ , dunque, l'eq. (A2.7) non sarà più soddisfatta. Tale valore di  $n$  si dice *ordine della transizione* di fase.

In una transizione del primo ordine, quindi, sarà  $p'_1(T_c) \neq p'_2(T_c)$ . Se è, per esempio,  $p'_1(T_c) > p'_2(T_c)$ , eq. (A2.5) mostra subito che  $\rho_1(T_c) > \rho_2(T_c)$ . Questo spiega, in particolare i comportamenti illustrati nelle figure VI.1 e VI.2 di questo capitolo.

## 2. Conservazione dell'entropia nell'espansione cosmologica e generazione di entropia nelle transizioni di fase

La conservazione dell'entropia comovente è conseguenza immediata delle equazioni di Friedmann. Dall'eq.  $d(\rho a^3) = -pd(a^3)$ , infatti, si ricava subito che

$$(A2.8) \quad \frac{d}{dt}[(\rho + p)a^3] = a^3 \frac{dp}{dt}$$

e, definita l'entropia comovente  $S = \sigma a^3$ , si vede che

$$(A2.9) \quad \frac{dS}{dt} = \frac{d}{dt} \left[ \frac{\rho + p}{T} a^3 \right] = \frac{a^3}{T} \dot{p} - S \frac{\dot{T}}{T}.$$

$S$  è dunque conservata se l'espansione avviene a  $p$  e  $T$  costanti, come in una transizione di fase del primo ordine alla temperatura  $T_c$ .

In generale, supponiamo che  $p$  dipenda da  $T$  e da un altro parametro  $\theta$ , a riassumere qualsiasi altra possibile dipendenza:  $p(T, \theta)$ . Risulta allora

$$(A2.10) \quad \dot{S} = -S \frac{\dot{T}}{T} + a^3 \frac{\partial p}{\partial T} \frac{\dot{T}}{T} + a^3 \frac{\partial p}{\partial \theta} \frac{\dot{\theta}}{T} = a^3 \frac{\partial p}{\partial \theta} \frac{\dot{\theta}}{T},$$

vista eq. (A2.4)), per cui  $a^3 \partial p / \partial T = S$ .

Se  $p$  dipende solo da  $T$ , dunque, l'entropia comovente  $S$  si conserva. Se abbiamo una transizione di fase cosmologica, e il parametro  $\theta$  è utilizzato per marcare la porzione del brodo termico che si trova in ciascuna fase, l'entropia continua a essere conservata, purché la pressione nelle due fasi sia eguale ( $\partial p / \partial \theta = 0$ ), come è se la transizione avviene alla temperatura  $T_c$ . Quindi, *in una transizione di fase del primo ordine che avviene a  $T_c$ ,  $S$  si conserva.*

Formalmente l'eq. (A2.10) può essere utilizzata per valutare l'aumento di entropia in una transizione di fase che avvenga dopo un periodo di *supercooling*, a condizione che non siano determinanti le deviazioni dalla condizione di omogeneità nella temperatura, che si possa supporre ben approssimata da un'unica funzione  $T(a)$ . In questo caso sarà

$$dS = a^3(t) \{p_1[T(a)] - p_2[T(a)]\} \sin 2\theta(a) dt$$

e le equazioni di Friedmann stabiliranno un ulteriore vincolo differenziale tra le variabili  $a$ ,  $T$  e  $\theta$ .

L'incremento di entropia, anche sotto queste restrizioni, non è dunque totalmente vincolato dalle condizioni cosmologiche, ma dipende ancora da un'informazione fisica, per esempio, sulla velocità di variazione di  $\theta$ .

Il problema risulta invece totalmente determinato dalle condizioni cosmologiche se si impone che la temperatura rimanga costante durante la transizione, sia che la transizione s'inneschi dopo una fase di *supercooling*, sia che essa avvenga alla temperatura  $T_c$ .

## Appendice A3

### Evoluzione del cosmo in una transizione di fase a temperatura $T_c$

Se una transizione di fase del primo ordine avviene a una temperatura prossima a  $T_c$ , la pressione delle due fasi ( $p_c$ ) è eguale. La densità d'energia può invece essere scritta

$$(A3.1) \quad \rho(\theta) = \rho_1 \cos^2 \theta + \rho_2 \sin^2 \theta$$

dove  $\rho_1 > \rho_2$  sono le densità d'energia delle due fasi alla temperatura  $T_c$ . La transizione inizia con  $\theta = 0$  e  $\rho = \rho_1$  per terminare quando  $\theta$  ha raggiunto  $\pi/2$  e  $\rho = \rho_2$ .

Possiamo immaginare che  $\theta$  lasci lo 0 non appena si nucleano bolle di fase 2 all'interno della fase 1. La crescita di dimensioni delle bolle e/o la nucleazione di altre bolle, accompagna la crescita di  $\theta$ . Già prima che sia  $\theta \sim \pi/4$  e metà del cosmo sia occupata dalla fase 2, le bolle cominciano a toccarsi; la tensione superficiale tende allora a trasformare i volumi di fase 1 in sfere. Tuttavia, attorno a  $\theta = \pi/4$ , la fase connessa diviene la fase 2 e in essa vanno gradualmente rimpicciolendo bolle residue della fase 1. Quando  $\theta$  si avvicina a  $\pi/2$  esse sono ridotte a proporzioni minime e spariscono del tutto per  $\theta = \pi/2$ .

Ponendoci su scale molto maggiori delle dimensioni medie delle bolle, su cui valga l'approssimazione di omogeneità, dalla conservazione dell'entropia comovente si ricava:

$$(A3.2) \quad \rho = T_c S / a^3 - p_c = \sigma_1 T_c (a_1 / a)^3 - p_c = \rho_1 \cos^2 \theta + \rho_2 \sin^2 \theta,$$

con  $\sigma_1 = S / a_1^3 = (\rho_1 + p_c) / T_c$ , e questa equazione stabilisce un legame tra  $\theta$  e  $a$ , determinando altresì la dipendenza di  $\rho$  da  $a$ .

La seconda equazione di Friedmann

$$(A3.3) \quad \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi}{3} G \left[ \sigma_1 T_c \left(\frac{a_1}{a}\right)^3 - p_c \right]$$

si può porre allora nella forma

$$(A3.4) \quad d\psi = \frac{du}{(1-u^2)^{1/2}} = (6\pi G p_c)^{1/2} dt, \quad \text{con } u = \left(\frac{p_c}{\rho_1 + p_c}\right)^{1/2} \left(\frac{a}{a_1}\right)^{3/2}, \quad \psi = \sin u$$

ottenendo così l'integrale

$$(A3.5) \quad u = \sin [(6\pi G p_c)^{1/2} t + A]$$

dove la costante  $A$  si ricava in funzione del tempo  $t_1$ , quando il fattore di scala vale  $a_1$ .

Questa espressione va eventualmente dettagliata in relazione alla transizione di fase cui si applica, per esempio quella elettrodebole o quella quark-adroni.

## Appendice A4

### Dipendenza della densità dal fattore di scala

Data l'equazione

$$(A4.1) \quad d(\rho a^3) + w\rho d(a^3) = 0,$$

consideriamo la relazione differenziale

$$d(\rho a^\alpha) = d(\rho a^3) a^{\alpha-3} + \rho \frac{\alpha-3}{3} a^{\alpha-3} d(a^3).$$

Dividendo membro a membro per  $a^{\alpha-3}$ , si ricava

$$(A4.2) \quad a^{-\alpha+3} d(\rho a^\alpha) = d(\rho a^3) + \rho \frac{\alpha-3}{3} d(a^3)$$

e il secondo membro di eq. (A4.2) si annulla grazie a eq. (A4.1) se si pone

$$(A4.3) \quad w = \frac{\alpha-3}{3} \quad \text{ovvero} \quad \alpha = 3(w+1).$$

Pertanto, se  $\alpha$  è dato da eq. (A4.3), sarà  $d(\rho a^\alpha) = 0$ .

## Appendice A5

### Evoluzione del fattore di scala in modelli a 2 componenti

*Miscela di componenti con  $w = 0$  e  $w = -1$*

L'equazione di Friedmann (IV.41) con semplici passaggi può essere messa nella forma

$$(A5.1) \quad a^{1/2} \dot{a} = H_0 a_o^{3/2} \Omega_{om}^{1/2} [1 + (\Omega_{oA}/\Omega_{om})(a/a_o)^3]^{1/2}$$

che, ponendo

$$(A5.2) \quad c = (\Omega_{oA}/\Omega_{om})^{1/2} (a/a_o)^{3/2},$$

diviene

$$(A5.3) \quad \frac{dc}{(1+c^2)^{1/2}} = \frac{3}{2} \Omega_{oA}^{1/2} H_0 dt.$$

Di questa equazione esiste l'integrale analitico

$$(A5.4) \quad \ln(\sqrt{1+c^2}+c) = (3/2) \Omega_{oA}^{1/2} H_0 t$$

Esso si semplifica in due regimi:

1) Nel «futuro», per  $c \gg 1$ , ovvero  $a \gg a_o$  avremo

$$(3/2) \ln(a/a_o) \simeq (3/2) \Omega_{oA}^{1/2} H_0 t$$

ovvero

$$(A5.5) \quad a/a_o \simeq \exp(\Omega_{oA}^{1/2} H_0 t).$$

Si tratta dell'andamento esponenziale, atteso quando la materia sia talmente diluita da non avere più impatto dinamico; allora il valore della costante  $\Omega_{om}$  diviene irrilevante.

2) Nel passato, per

$$(A5.6) \quad 1+z \gg (\Omega_{oA}/\Omega_{om})^{-1/3},$$

quando

$$(A5.7) \quad (a/a_o)^{3/2} = (3/2) \Omega_{om}^{1/2} H_0 t$$

e il valore della costante  $\Omega_{oA}$  diviene irrilevante.

*Miscela di componenti con  $w = 0$  e  $w = 1/3$ ; calcolo dell'orizzonte*

Attorno all'equivalenza, ponendo  $R = a/a_{eq}$ ,  $\bar{H}_{eq}^2 = (8\pi/3)G\rho_{m,eq} = H_{eq}^2/2$ , l'equazione di Friedmann si può convenientemente scrivere nella forma

$$(A5.8) \quad R\dot{R} = \bar{H}_{eq}\sqrt{1+R}.$$

Integrando per separazione di variabili, si ricava

$$(A5.9) \quad t = \frac{2}{3\tilde{H}_{eq}} [(1+R)^{1/2}(R-2) + 2]$$

per cui è facile dare  $t$  in funzione di  $R = a/a_{eq}$ . Per ottenere la funzione inversa si deve risolvere una cubica.

Ciò non è necessario per ricavare l'orizzonte

$$(A5.10) \quad l_h(t) = c a(t) \int_0^t \frac{dt'}{a(t')} = cR(t) \int_0^t \frac{dt'}{R(t')},$$

per cui è invece importante riscrivere l'eq. (A5.8) nella forma

$$\frac{dt'}{R'} = \frac{dR'}{\tilde{H}_{eq}\sqrt{1+R'}},$$

che permette la sostituzione di  $t'$  con  $R'$  nell'eq. (A5.10). Si ha così l'espressione

$$(A5.11) \quad l_h(t) = \frac{c}{\tilde{H}_{eq}} R(t) \int_0^{R(t)} \frac{dR'}{\sqrt{1+R'}} = \frac{2c}{\tilde{H}_{eq}} R(t) \{ [1+R(t)]^{1/2} - 1 \}.$$

che dà l'orizzonte attorno all'equivalenza senza alcuna approssimazione. In particolare, per  $R = 1$  ( $a = a_{eq}$ ), risulta

$$l_{h,eq} = \frac{2c}{\tilde{H}_{eq}} (\sqrt{2} - 1) = 0.8284 \frac{c}{\tilde{H}_{eq}} = 1.172 \frac{c}{H_{eq}}$$

L'andamento (A5.11) può essere considerato anche nei due casi limite:

$$1) R \ll 1: [1+R(t)]^{1/2} - 1 \simeq R(t)/2,$$

$$\text{e, quindi: } l_h \simeq R^2 c / \tilde{H}_{eq} \simeq c/H$$

$$2) R \gg 1: [1+R(t)]^{1/2} - 1 \simeq R^{1/2}(t),$$

$$\text{e, quindi: } l_h \simeq R^{3/2} 2c / \tilde{H}_{eq} = 2c/H$$

come previsto. Tenendo conto dell'eq. (A5.9) si ha altresì

$$(A5.13) \quad \frac{l_h(t)}{ct} = \frac{3R[(1+R)^{1/2} - 1]}{2 + (R-2)(1+R)^{1/2}}$$

che, per  $R \gg 1$  ( $R \ll 1$ ) dà  $l_h \simeq 3ct$  ( $l_h \simeq 2ct$ ) mentre, per  $R = 1$ , dà

$$(A5.14) \quad l_{h,eq} = (3/\sqrt{2}) ct_{eq} = 2.121 ct_{eq}.$$

## Appendice A6

### Orizzonti

Se  $w (\neq -1)$  è costante, il fattore di scala varia nel tempo con la legge di potenza

$$(A6.1) \quad a/a_i = (t/t_i)^\alpha, \quad \text{con } \alpha = 2/[3(w+1)],$$

ove  $a_i$  è il valore del fattore di scala a un tempo di riferimento  $t_i$ . Il valore  $\alpha = 1$  (o  $w = -1/3$ ) delimita due regimi diversi, per quanto concerne l'orizzonte causale: per  $\alpha > 1$  ( $< 1$ ) c'è orizzonte degli eventi (di particella) e non c'è orizzonte di particella (degli eventi). Infatti i due orizzonti sono

$$(A6.2) \quad \begin{aligned} l_p(t) &= c a(t) \int_0^t \frac{dt'}{a(t')} = c t^\alpha \int_0^t \frac{dt'}{t'^\alpha} = c t^\alpha \frac{1}{1-\alpha} \left\{ t'^{1-\alpha} \Big|_0^t \right\} && \text{(particella)} \\ l_{ev}(t) &= c a(t) \int_t^\infty \frac{dt'}{a(t')} = c t^\alpha \int_t^\infty \frac{dt'}{t'^\alpha} = c t^\alpha \frac{1}{1-\alpha} \left\{ t'^{1-\alpha} \Big|_t^\infty \right\} && \text{(eventi)} \end{aligned}$$

La quantità in parentesi graffa diverge se  $\alpha > 1$  ( $\alpha < 1$ ) nel caso di orizzonte di particella (degli eventi): tale orizzonte è illimitato.

In sostanza, per  $\alpha < 1$ , c'è un orizzonte di particella

$$(A6.3) \quad l_p(t) = ct/(1-\alpha),$$

mentre, per  $\alpha > 1$  (espansione accelerata), c'è un orizzonte degli eventi

$$(A6.4) \quad l_{ev}(t) = ct/(\alpha-1).$$

Si è escluso il caso  $w = -1$ . Per esempio, eq. (A6.1) darebbe  $\alpha \rightarrow \infty$  per  $w = -1$ . Per  $w = -1$ , l'equazione di Friedmann

$$d(\rho a^3) = \rho d(a^3)$$

implica sia  $\rho \equiv \text{const}$ . Allora la seconda equazione di Friedmann dà

$$a(t) = a_i \exp[H(t-t_i)],$$

ove  $H = (8\pi/3)G\rho$  è costante. In questo caso risultano finiti sia l'orizzonte di particella

$$l_p = \frac{1}{H} (e^{Ht} - 1)$$

che quello degli eventi

$$l_{ev} = \frac{1}{H},$$

che è costante: la distanza cui è possibile inviare un segnale permane sempre eguale, mentre invece la distanza tra punti cresce esponenzialmente.

## CAPITOLO V

# Un modello per l'universo

### 1. Modello e realtà

Ogni descrizione teorica di fenomeni reali si basa su un modello. Nessun modello pretende di coincidere col reale, anche se il rapporto modello-realtà è molto diverso quando il modello è un *corpo rigido* o *l'equazione di Schrödinger*.

I modelli cosmologici non fanno eccezione. In questo caso l'avvicinamento alla realtà viene dall'alto. Si presume cioè che, sulle grandissime scale, la fisica si semplifichi; abbiamo a lungo discusso come ciò, che all'inizio era un assunto arbitrario, sia gradualmente divenuto un dato fisico. Da questo punto di partenza la trattazione si dipana per gradi.

Nei due capitoli precedenti abbiamo dibattuto varie soluzioni delle equazioni di Einstein, tradotte in equazioni di Friedmann grazie all'assunto di isotropia e omogeneità. In questo capitolo è venuto il momento di operare delle scelte.

Nei capitoli successivi il quadro omogeneo-isotropo resterà come base e contorno. Si cercherà di articolarlo prima introducendo blande disomogeneità, trattabili in approssimazione lineare, per passare poi alle alte non-linearità. È una progressiva immersione in scale decrescenti e in crescenti complicazioni analitiche e numeriche.

Nel corso di questo capitolo, tali complessità sono ancora rimandate o compaiono solo come strumento per costruire i dati atti a operare le scelte.

Possiamo affermare allora che la maggior parte di questo volume fa riferimento a due modelli, denominati  $\Lambda$ CDM (*standard-cold-dark-matter*) e  $\Lambda$ CDM ( *$\Lambda$ -cold-dark-matter*). In entrambi gli appellativi è chiaro il riferimento alla *cold dark matter*, materia oscura fredda. La *materia oscura* (DM: *Dark Matter*) non emette radiazione elettromagnetica né appare soggetta a altre forze non gravitazionali. Più complesso è specificare cosa si intende qualificandola come *fredda*; è opportuno rimarcare che non ci si riferisce a una sua (attuale) temperatura.

Il prossimo paragrafo sarà centrato sulle evidenze fisiche della DM. Una delle maggiori scoperte della cosmologia è che la parte prevalente della materia cosmica non è *barionica*, non si compone cioè delle particelle incluse nel cosiddetto *modello standard* delle interazioni fondamentali. Essa non è quindi fatta di atomi, elettroni, protoni, mesoni, ecc., ma di particelle di natura ancora ampiamente incognita. Questa

componente oscura domina il modello  $\Lambda$ CDM, lasciando spazio a poco più del 4% di materiali barionici.

Ma l'ultimo decennio ha portato a un ulteriore distacco della sostanza cosmica dalla nozione di materia, legata all'esperienza comune e/o dal laboratorio. Accanto alla DM, nel cosmo c'è un'altra componente oscura, distinta dalla DM dall'equazione di stato. Si tratta dell'*energia oscura* (DE: *Dark Energy*). Di essa possiamo affermare subito che *non si compone di particelle*.

Si tratta di una conclusione suffragata da una fenomenologia ormai vastissima, che è quella che ha portato il modello  $\Lambda$ CDM a prevalere sul modello SCDM. SCDM resta importante come palestra semplificatrice di un gran numero di calcoli e valutazioni teoriche.

Il modello  $\Lambda$ CDM, invece, permette attualmente di catalogare tutta la fenomenologia di rilievo cosmologico. Come vedemo, peraltro, esso contiene elementi assai insoddisfacenti, dal punto di vista concettuale. Tutta la questione della natura della DE sarà comunque oggetto di un altro capitolo.

## 2. DM: perché

Ci sono due strade per «pesare» i contenuti materiali di una regione cosmica:

- 1) Dalla luminosità  $L$  risalire alla massa  $M$  di ogni singolo oggetto, attraverso una motivata analisi del rapporto  $M/L$ .
- 2) Dalla velocità dei singoli oggetti risalire alla massa complessiva, assumendo che il sistema sia in equilibrio dinamico, sicché per esso vale il *teorema del viriale*. Energia cinetica  $E_{kin}$  e energia potenziale  $E_{pot}$  totali del sistema saranno allora legati dalla relazione:

$$(V.1) \quad \text{Vir} = 2E_{kin} + E_{pot} = 0.$$

Un primo uso del secondo approccio venne fatto negli anni Trenta da Oort e Zwicky, che si proponevano di verificare se le stelle nei dintorni solari fossero in equilibrio dinamico.

### I dintorni solari

Se applichiamo il teorema (V.1) a un sistema in equilibrio, composto di stelle o altri oggetti cosmici, di massa  $m_i$ , in moto a velocità di modulo  $v_i$ , a distanze reciproche  $r_{ij}$ , deve essere

$$(V.2) \quad \text{Vir} = \sum_i m_i v_i^2 - G \sum_{j>i} \frac{m_i m_j}{r_{ij}} = 0.$$

La prima somma si applica a tutti gli oggetti, etichettati dall'indice  $i$ . La seconda a tutte le coppie di oggetti, identificabili facendo scorrere l'indice  $i$  e prendendo le coppie etichettate da  $i$  e  $j > i$ .

Il viriale delle stelle nei dintorni solari, di cui era possibile determinare la massa dalla luminosità (esse sono in genere piuttosto simili al Sole) e la distanza per triangolazione, misurato da Oort e Zwicky, era invece  $\gg 0$ .

Il loro approccio tuttavia conteneva vari elementi di incertezza:

- 1) Le masse  $m_i$  erano dedotte dalla luminosità tramite un modello politropico, come si usava al tempo.
- 2) Le velocità erano stimate dalla sola componente radiale, misurata tramite il *redshift*.
- 3) Si escludeva un ruolo dei materiali al contorno.

Interessante è peraltro la soluzione da essi data al problema: se tutte le masse  $m_i$  sono moltiplicate per un identico fattore  $f$ , il termine positivo,  $E_{kin}$ , diviene  $fE_{kin}$ ; il termine negativo  $E_{pot}$ , invece, diviene  $f^2E_{pot}$ . Con un  $f$  adeguato, quindi, Vir si riporta allo 0. Al tempo il valore determinato fu  $f = 3$ . C'erano dunque  $2/3$  della massa non visibili, già nei dintorni solari.

È curioso che, dopo aver precisato le masse tramite tecniche di evoluzione stellare, determinato con estrema precisione le distanze, valutata l'incertezza derivante dall'uso di una sola componente della velocità per un sistema di stelle di masse diverse, ecc., si giunga a tutt'oggi a una stima  $f \simeq 3$ .

### **Il disco della Via Lattea**

L'approccio corretto si basa sulla determinazione della *funzione di luminosità* delle stelle  $\Phi(L)$ : nell'unità di volume, ci sono  $\Phi(L)dL$  stelle con luminosità compresa tra  $L$  e  $L + dL$ . Per le stelle il rapporto  $M/L$  è adeguatamente conosciuto, sicché si può dare una funzione  $m(L)$ . La somma del contributo alla massa totale, proveniente da ciascun intervallo di luminosità  $dL$ , si esprime tramite l'integrale

$$(V.3) \quad \rho_{lum} = \int dL m(L) \Phi(L) \simeq 0.1 M_{\odot}/\text{pc}^3$$

che dà la densità in materia luminosa nella Via Lattea.

Per passare alla massa del disco della Via Lattea nel suo assieme occorrono vari altri dati. Anzitutto lo spessore (variabile) del disco e una legge con cui i suoi bordi sfumino verso il vuoto. Poi i contributi dovuti a polveri, gas, ecc. Su questi punti le stime sono piuttosto delicate.

Un'analisi astronomica adeguata di tutta questa problematica porta a una massa totale  $\sim 7-8 \times 10^{10} M_{\odot}$ . Una stima dinamica delle masse, del tipo espresso dalle eq. (V.1) e (V.2), ma adeguata alla variabilità delle condizioni locali, porta invece a una massa  $\simeq 2 \times 10^{11} M_{\odot}$ .

Questo mancato *fit*, per un fattore compreso tra 2 e 3, non va visto come problematico. Se i dati permettessero di determinare i due termini del confronto con assoluta certezza e il *fit* fosse esatto, avremmo una straordinaria conferma della va-

lità della gravitazione universale sulla scala di alcune decine di kpc, diametro della Via Lattea. La gravitazione universale si può invece considerare verificata solo nell'ambito del Sistema Solare, fino a distanze  $10^{-4}$ - $10^{-3}$ pc.

Il mancato *fit* porta invece alla supposizione che, valendo la gravitazione newtoniana, nel disco galattico sia presente altra materia, di quantità almeno pari alla materia luminosa, che non emette radiazione sufficiente da poter essere osservata. Si tratta della prima evidenza di DM.

### **La curva di rotazione delle spirali. Gli aloni massivi**

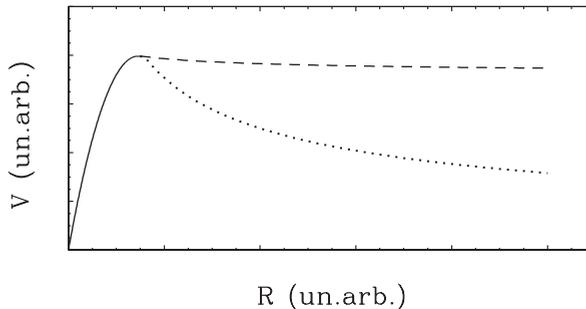
Morfologia e fisica delle galassie non sono semplici e dipendono significativamente dal *redshift*. Tra le galassie, notevoli sono quelle che presentano una morfologia a spirale, ove si distingue un nucleo, un disco, ecc. L'andamento della luminosità del disco declina esponenzialmente a partire dal centro galattico  $C$ , fino al *raggio ottico*  $r_o$ , oltre il quale l'immagine ottica si confonde con la luminescenza del cielo.

Dai dati ottici, a distanze  $R < r_o$  da  $C$ , si costruiscono le *curve di rotazione* e la velocità di rotazione osservata  $V(R)$  appare in crescita al crescere di  $R$ .

Sarebbe comunque legittimo attendersi che, scemata l'intensità luminosa e quindi la densità dei materiali, a partire da un opportuno raggio confrontabile con  $r_o$  o poco maggiore,  $V(R)$  iniziasse a decrescere, come fanno le velocità dei pianeti via via che ci si allontana dal Sole, conformemente alle aspettative kepleriane. Infatti, se la massa  $M_g$  della galassia è essenzialmente contenuta entro  $r_o$ , a distanze  $> r_o$  la velocità di rotazione dovrebbe essere data dalla relazione

$$(V.4) \quad V^2(R) = GM_g/R$$

che esprime l'eguaglianza tra accelerazione centripeta e forza newtoniana su un'orbita circolare, oppure semplicemente l'equilibrio viriale.



**FIGURA V.1**

Andamento della velocità di rotazione nelle galassie a spirale. La linea continua dà la velocità di rotazione nella zona per cui sono disponibili dati ottici. Oltre il raggio ottico sono disponibili dati Radio ottenuti dalla riga a 21 cm; essi mostrano una velocità di rotazione che permane essenzialmente costante (linea tratteggiata) invece di calare in modo  $\propto R^{-1/2}$  come ci si attenderebbe da un ordinario comportamento kepleriano (linea punteggiata). In ascissa e ordinata le unità sono arbitrarie.

Anche quando la luminosità dei materiali galattici va mediamente al di sotto della luminosità del cielo, resta però possibile osservare la galassia su una riga tipica dei materiali galattici e non presente nello spettro del cielo notturno. In effetti, l'analisi delle curve di rotazione delle spirali è stata proseguita oltre  $r_o$  sulla base di dati Radio sulla riga dell'idrogeno a 21 cm.

In figura V.1, il tipico andamento di  $V(R)$  nella zona ottica è rappresentato da una linea continua. La linea punteggiata è un'indicazione delle aspettative kepleriane. La linea tratteggiata, infine, rappresenta l'andamento tipico di  $V(R)$  a  $R > r_o$  quale si deduce dai dati Radio. In sostanza, a  $R > r_o$  l'andamento di  $V(R)$  è costante e tale permane tipicamente fino a  $R \sim 7-8 r_o$ .

I dati sono compatibili con la presenza di due componenti nella velocità:

$$(V.5) \quad \sigma^2 = \langle V_{th}^2 \rangle + \langle V_{kep}^2 \rangle,$$

che diremo “termica” e “kepleriana”, per distinguere moti disordinati locali da flussi coerenti su scala  $\sim R$ . I moti non si osservano solo sul piano galattico, ma in tutte le direzioni attorno al centro della galassia, dove si scopre la presenza di materiali che ne costituiscono l'*alone*. Essendo  $\sigma$  indipendente da  $R$ , si ritiene giustamente che ciò non derivi da una “congiura” tra  $V_{th}$  e  $V_{kep}$ , in tutte le spirali osservate. In particolare, quindi,  $V_{th}$  sarà ovunque costante (alone *isotermo*). La costanza della componente kepleriana richiede poi che vi siano materiali in quantità sostanziali anche a  $R > r_o$ , di modo che l'equazione

$$(V.6) \quad V^2(R) = GM(<R)/R$$

implichi  $V$  costante. In questa relazione  $M(<R)$  è la massa totale entro  $R$ . È chiaro come ciò richieda che sia

$$(V.7) \quad M(<R) \propto R$$

e quindi, se la curva di rotazione con  $\sigma$  costante si estende a  $n$  volte il raggio ottico, la massa della galassia, incluso il suo *alone massivo*, sarà  $n$  volte la massa entro  $r_o$ .

Applicando questo criterio alla Via Lattea, per esempio, la sua massa risulta non inferiore a  $10^{12}M_{\odot}$ .

### **Stime del parametro di densità**

Per stimare la densità complessiva dei materiali galattici nel cosmo, si procede in modo analogo alle stelle. Si parte dalla *funzione di luminosità* delle galassie  $\Phi_g(L)$  e si cerca il rapporto  $M/L$  per le galassie. Mentre l'analogo rapporto per le stelle fornisce una funzione quasi univoca  $m(L)$ , per le galassie si riscontrano sbalzi di ordini di grandezza. Non è quindi banale convergere su un andamento *medio* di  $m_g(L)$ , che si potrà comunque applicare solo a campioni ampi di galassie. Su questa base si cerca di ot-

tenere la *funzione di massa* delle galassie

$$(V.8) \quad \Phi(M) dM = \Phi_g(L) m_g(L) dL .$$

Un ulteriore elemento di incertezza è costituito da previsioni e osservazioni di oggetti galattici di piccola massa e luminosità, per i quali ci sono ulteriori difficoltà nella stima di  $m_g(L)$ .

La massa totale per unità di volume contenuta nelle galassie è comunque espressa da

$$(V.9) \quad \rho_g = \int dL \Phi_g(L) m_g(L)$$

e il suo valore è proporzionale alla normalizzazione della funzione  $m_g(L)$  e quindi, in definitiva, alla massa  $M_{mw}$  attribuita alla Via Lattea. Come termine di confronto possiamo usare la densità critica

$$(V.10) \quad \rho_{cr} = 2.78 \times 10^{11} h^2 M_\odot / \text{Mpc}^3 ,$$

stimata a partire da  $H_o = h 100 \text{ km/s/Mpc}$ , ottenendo i relativi valori del parametro di densità dei materiali galattici. Si veda a questo proposito la seguente tabella che mostra come scalino la densità (in materiali galattici) dell'Universo ( $\rho_g$ ) e il relativo parametro di densità  $\Omega_g$ , al crescere delle stime sulla massa della Via Lattea ( $M_{mw}$ ).

$M_{mw}/M_\odot$	$\rho_g h^2 / (M_\odot / \text{Mpc}^3)$	$\Omega_g h^2$
$7 \cdot 10^{10}$	$\sim 1.5 \times 10^9$	$\sim 0.005$
$2 \cdot 10^{11}$	$\sim 4.5 \times 10^{10}$	$\sim 0.015$
$10^{12}$	$\sim 2.2 \times 10^{10}$	$\sim 0.08$

Con un valore di  $h \sim 0.7$ , il parametro di densità per la massa contenuta nelle galassie è quindi  $\sim 0.15-0.20$ .

### ***Ammassi di galassie***

La sottostima di un fattore 5-10 della massa delle galassie, ove ci si rifaccia unicamente ai dati ottici, è confermata da un'analisi dinamica degli *ammassi di galassie*, ottenibile da una semplice applicazione del teorema del viriale (V.1)-(V.2).

Negli anni Sessanta, quando gli ammassi erano stati da poco scoperti, come picchi di densità numerica di galassie sulla volta celeste, l'osservazione che in essi era  $2E_{kin}/E_{pot} \sim \mathcal{O}(10)$  aveva portato alcuni astrofisici a collocare negli ammassi la sorgente della creazione continua di materia, postulata nei modelli cosmologici dello *stato stazionario*. Oggi gli ammassi sono ben individuati tramite la radiazione  $X$  emessa, per *bremstrahlung*, dal gas caldo in essi racchiuso, e non c'è nessun dubbio sulla loro stabilità, come sul fatto che lo stesso gas ha da solo una massa superiore a quella delle galassie dell'ammasso.

Maggiori dettagli saranno forniti nel capitolo sugli ammassi di galassie. È comunque chiaro che l'analisi degli ammassi di galassie conferma la necessità di ri-normalizzare la massa delle singole galassie per un fattore analogo a quello ricavato dalla presenza degli aloni massivi.

### Verso $\Omega = 1$

Il primo punto da sottolineare è che, a priori, non c'è alcuna ragione di inferire che la densità dei materiali cosmici sia prossima alla densità critica. Mentre  $\rho_{cr}$  è dedotto da dati sull'espansione cosmica su grandissima scala, la stima di  $\rho_g$  si diparte da stime di luminosità, opportunamente rinormalizzate tramite stime dinamiche del rapporto  $M/L$ . Una differenza di 10 ordini di grandezza, per misure di grandezze tanto difformi, per esempio, sarebbe del tutto legittima.

È il fatto che le osservazioni portino invece a  $\Omega_g \sim 0.2$  a far nascere un problema oggettivo: perché le due densità debbono essere così prossime, ma non coincidenti. Si ricordi altresì l'argomentazione tesa a mostrare che ogni valore  $\Omega \neq 1$ , ma prossimo all'unità, richiede un estremo *fine tuning* del parametro di densità al tempo di Planck.

Anche per questa ragione vennero sviluppate tecniche mirate a studiare la distribuzione dei materiali su grandissima scala, volte a ricostruire dal *campo di velocità*, supposto *irrotazionale*, l'andamento della distribuzione delle masse. L'idea era di dimostrare che i materiali esterni agli ammassi concorrevano a coprire il *gap* tra 0.2-0.3 e 1. Tra queste tecniche ricordiamo quella legata al programma POTENT. L'input di un simile programma era costituito da dati sulle velocità peculiari delle galassie, depurati dalla componente di *redshift*. La difficoltà di usufruire di un vasto e omogeneo campione di galassie adeguate ha condotto a output irrisolutivi, che comprendono come *estremo inferiore*  $\Omega \sim 0.3$  e come *estremo superiore*  $\Omega \sim 1$ .

Le osservazioni permettono quindi di affermare che la densità complessiva del cosmo è coerente con un parametro di densità  $\Omega \lesssim 0.15-0.20$ ; non vi sono dati in favore di parametri di densità molto più elevati.

Ciò significa, di per sé, che solo il 3-4% dei materiali cosmici sono luminosi. Il resto è costituito di *materia oscura*.

A fronte di questo dato stanno peraltro le conclusioni di due prossimi capitoli: quello sulla nucleosintesi primordiale (BBN) e quello sulle anisotropie del CMB. Entrambi portano alla stessa valutazione della densità dei *barioni* nel cosmo:  $\Omega_b \sim 0.040-0.045$ . Tra questo valore e la densità oggettivamente esistente c'è una discrepanza di un fattore superiore a 4. Ciò significa che almeno  $\sim 75-80\%$  dei materiali cosmici non sono barioni.

Complessivamente abbiamo quindi che:

- 1) 1% (o un po' meno) della materia del cosmo, di natura barionica, emette luce
- 2) 4-4.5% della densità critica è data da materia barionica
- 3) 20% (o più) della densità critica è data da materiali di natura non barionica.

La conclusione generale è che, nel cosmo, la materia che emette radiazione è  $\sim 1\%$  del totale; la materia oscura restante si suddivide in barionica e non barionica; e il 75% almeno, della materia del cosmo, è *materia oscura non barionica*.

Nessuno dei dati sin qui considerati obbliga a ritenere che il parametro di densità attuale  $\Omega_o$  sia  $\simeq 1$ .

### 3. Le ere cosmiche

Dunque, il cosmo attuale è comunque caratterizzato da un parametro di densità della materia  $\Omega_{om} > 0.2$ . Parte di questa materia sono barioni, per il resto si tratta di DM. Accanto alla materia c'è anche molta radiazione, dovuta al CMB, ma non abbastanza da alterare le stime su  $\Omega_o$ . Le relative densità sono

$$(V.11) \quad \rho_{om} = \Omega_{om} h^2 \times 2.78 \times 10^{11} M_{\odot} / \text{Mpc}^3$$

$$(V.12) \quad \rho_{o\gamma} = 0.74 \times 10^7 M_{\odot} / \text{Mpc}^3$$

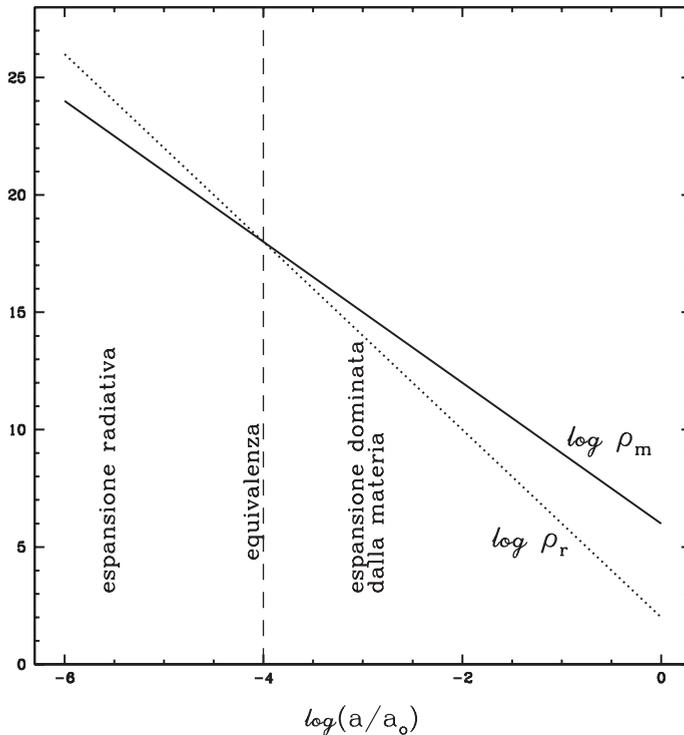


FIGURA V.2

Andamento delle densità di radiazione e materia fino all'epoca attuale. Le densità sono normalizzate arbitrariamente e anche il loro rapporto a  $z = 0$ , sebbene ragionevole, è arbitrario.

Il grafico mette in evidenza come la diversa pendenza di  $\rho(a)$  implichi necessariamente un'espansione radiativa a alti  $z$ .

Andando poi all'indietro nel tempo, mentre è  $\rho_m \propto a^{-3}$ , sarà  $\rho_\gamma \propto a^{-4}$ . Ciò dà luogo al grafico in figura V.2 sicché, al *redshift*

$$(V.13) \quad 1 + z_{e,\gamma} = \Omega_{om} h^2 \times 2.78 \times 10^{11} / 0.74 \times 10^7 \simeq 4.02 \times 10^4 \Omega_{om} h^2,$$

materia e  $\gamma$  hanno eguali densità e, a *redshift*  $z > z_{e,\gamma}$ , la densità della radiazione fotonica supera quella della materia.

Oltre al CMB, nel cosmo odierno di ritiene debba esistere un analogo fondo di neutrini. Nel prossimo capitolo mostreremo che la sua «temperatura» è  $T_\nu = (4/11)^{1/3} T$ . Il numero di neutrini, rapportato al numero di fotoni del CMB, sarà allora

$$(V.14) \quad \frac{n_\nu}{n_\gamma} = \frac{\tilde{g}_\nu}{\tilde{g}_\gamma} \frac{4}{11} = \frac{3}{11}$$

per ciascuna specie neutrinica. Essendo  $n_\gamma \simeq 415$ , sarà  $n_\nu \simeq 112$ . Con  $N_\nu = 3$  specie neutriniche, il numero totale di neutrini risulta  $n_{\nu,3} \simeq 340$ .

La densità d'energia dei neutrini, sempre rapportata a quella del  $\gamma$ , sarà

$$(V.15) \quad \frac{\rho_\nu}{\rho_\gamma} = \frac{7}{8} N_\nu \left( \frac{4}{11} \right)^{4/3} = 0.2271 N_\nu.$$

Ne consegue che la densità complessiva della componente radiativa ( $\rho_r$ ) è accresciuta di un fattore  $1 + 0.2271 N_\nu$  rispetto a  $\rho_\gamma$  e il *redshift* di equivalenza si abbassa, diventando

$$(V.16) \quad 1 + z_e = 2.39 \times 10^4 \Omega_{om} h^2 \frac{1.6813}{1 + 0.2271 N_\nu}$$

A  $z < z_e$ , dunque, la materia è la principale sorgente dell'espansione e quest'epoca viene detta *era dominata dalla materia*. A  $z > z_e$ , invece, la materia relativistica prende il sopravvento e quest'epoca viene detta *era radiativa*.

Sempre andando a ritroso nel tempo si giunge alla temperatura  $T_l = T_o(1 + z_l) = m_e$  (massa dell'elettrone,  $\sim 0.511$  MeV). A  $z \simeq z_l$  siamo nell'*era leptonica*. L'espansione avviene come nell'era radiativa, ma al brodo termico vanno aggiunti gli elettroni. Attorno  $z_l$  abbiamo l'era della nucleosintesi, cui dedichiamo per intero il prossimo capitolo. A una temperatura  $T = m_\mu$  (massa del  $\mu$ ,  $\sim 105$  MeV) al brodo termico si uniscono anche le particelle  $\mu$  che, in sostanza, sono elettroni pesanti.

A una temperatura poco più alta, si giunge alla transizione quark-adroni, i quark entrano a far parte del brodo cosmico, e l'era leptonica può dirsi terminata. Assieme a essa termina il dominio in cui si può costruire la cosmologia partendo da fisica nota dal laboratorio e ci si avvia verso quegli sviluppi in cui il quadro cosmologico è *utilizzato* come laboratorio di verifica di ipotesi fisiche.

#### 4. DE: perché

Al finire degli anni Ottanta due modelli cosmologici si fronteggiavano: SCDM e  $\Lambda$ CDM. In entrambi la componente principale del cosmo odierno era la DM. Nel primo, si supposeva che la quantità complessiva di DM portasse il parametro di densità complessivo a valori  $\Omega_o \simeq 1$ . Di questa densità faceva parte, per  $\sim 4\text{-}5\%$ , la componente barionica. Il modello  $\Lambda$ CDM, invece, assumeva che fosse  $\Omega_o \simeq \Omega_{o,m} \sim 0.2\text{-}0.3$  e quindi che curvatura spaziale  $\kappa$  fosse negativa.

Riassumendo vantaggi e svantaggi: a favore di  $\Lambda$ CDM stavano soprattutto i dati sul numero di sistemi a alto  $z$ . In particolare,  $\Lambda$ CDM prevedeva più *ammassi di galassie* di SCDM e i dati erano in favore di  $\Lambda$ CDM. Ancora più convincenti i numeri delle *Lyman- $\alpha$  clouds*, sistemi pre-galattici, il cui numero eccedeva ampiamente le previsioni di SCDM. A sfavore di  $\Lambda$ CDM stava il problema del *fine-tuning* di  $\Omega$  al tempo di Planck, che però poteva essere alleviato da opportuni modelli inflazionari. Decisamente a favore di SCDM stavano invece i dati sulle anisotropie di grande scala del CMB, provenienti dal satellite COBE.

Le difficoltà di entrambi i modelli erano parzialmente alleviate dai modelli CHDM, in cui una parte sostanziale della DM veniva ricondotta a neutrini massivi. Si tratta di modelli ormai superati dai dati più recenti. La loro importanza oggi è quella di aver fornito le tecniche per stabilire limiti cosmologici alla massa del neutrino.

Gli esperimenti che segnano il punto di svolta riguardano la dipendenza delle luminosità apparenti delle SNIa (Supernovae di tipo Ia) dal *redshift*. Di per sé si tratta di dati dall'apparenza assai discriminante, ma che si basano su una fisica assai delicata. È quindi importante che le conclusioni ricavate dalle SNIa siano oggi corroborate da dati sul CMB e sui campioni profondi di galassie.

Va però subito chiarita una differenza tra la portata dei diversi dati. Tramite le SNIa si misura l'*accelerazione* o decelerazione dell'espansione cosmica. In un modello  $\Lambda$ CDM, dominato dalla curvatura, ci si aspettava un parametro di decelerazione  $q_o \simeq 0$ . Per un modello SCDM, invece, doveva essere  $q_o \simeq 0.5$ . Tramite le SNIa, invece, si accerta che  $q_o$  è negativo.

I campioni profondi di galassie e i dati recenti sulle anisotropie del CMB, invece, permettono una misura separata della densità totale del cosmo ( $\Omega_o$ ) e della densità della materia non-relativistica nel cosmo ( $\Omega_{o,m}$ ). Essi convergono nell'indicare valori  $\Omega_o \sim 1$ ,  $\Omega_{o,m} \sim 0.25\text{-}0.30$ . Il *gap* tra questi parametri di densità, che corrisponde a  $\sim 70\%$  della massa del cosmo, viene attribuito all'energia oscura, una sostanza non composta di particelle.

#### *Le SNIa*

Le Supernovae (SN) sono stelle che esplodono. La loro luminosità al picco è confrontabile con quella di un'intera galassia:  $L \sim 10^9 L_\odot$ .

La prima suddivisione tra SN riguarda le righe di *H* nel loro spettro: Se sono presenti si hanno SNII; nelle SNI, invece, esse sono assenti o molto tenui.

Le SNI si suddividono ulteriormente come segue:

<i>TIPO</i>	caratteristiche osservative	origine
SN Ia:	forte assorb. attorno 6150 Å ( $S_{i^{II}}$ )	distruzione di nana bianca
SN Ib:	prominenti linee da $He^I$	collasso di supergigante
SN Ic:	SNI che non sono a o b	collasso di supergigante b e c <i>distinte da perdita o</i> <i>conservazione strati esterni</i>

Se tutte le stelle di una determinata classe avessero esattamente la stessa luminosità intrinseca, dalla luminosità apparente se ne potrebbe determinare la distanza. La luminosità delle SNIa invece varia nel tempo e ci sono (piccole) differenze tra le diverse SNIa. Entrambe tuttavia sono sotto controllo, in quanto riducibili a altre caratteristiche riconoscibili in modo indipendente.

Più in dettaglio, le SNIa sono state usate come indicatori di distanza perchè:

- 1) Molto luminose.
- 2) Piccola dispersione di luminosità attorno al picco ( $< 0.3$  mag) e pendenza (quasi) sempre eguale della curva di luminosità dopo il picco
- 3) I progenitori e il meccanismo di esplosione sono ragionevolmente noti.
- 4) Ci si attende poca evoluzione nel loro comportamento al variare del *redshift* e tale modesta evoluzione può essere prevista.
- 5) C'è abbondanza di *test* su SNIa locali in diversi ambienti cosmici.

In questo modo Phyllips ha costruito un *template* con 9 curve di luce di raffronto, che collega le piccole variazioni della luminosità al picco alle piccole variazioni nella pendenza dopo il picco.

Queste caratteristiche fanno sì che le SNIa possano essere utilizzate come *candele standard*, determinandone quindi la distanza  $d_L$  (di luminosità). Dallo spettro della galassia ospite si può poi determinare il loro *redshift*, costruendo così un grafico  $d_L$  vs.  $z$ .

### **Misura del parametro di decelerazione tramite candele standard**

I punti base del ragionamento sono semplici.

1. Si definisce la *distanza di luminosità*. In uno spazio euclideo in cui sorgente e osservatore, siti a distanza  $d$ , siano mutuamente a riposo, la luminosità apparente  $l$  e la luminosità intrinseca  $L$  di una sorgente sono collegate dalla relazione

$$(V.17) \quad l = L/4\pi d^2 \quad \text{ovvero} \quad d = \sqrt{L/4\pi l}.$$

Questa relazione applicata alle sorgenti reali nel cosmo ne definisce la distanza di luminosità

$$(V.18) \quad d_L = \sqrt{L/4\pi l}.$$

2. *Connessione con la metrica.* Si dimostra che

$$(V.19) \quad d_L = ra_o^2/a.$$

$r$  è la coordinata radiale del punto ove è sita la sorgente. La metrica è

$$(V.20) \quad ds^2 = c^2 dt^2 - a^2(t) \left[ \frac{dr^2}{1 - \kappa r^2} + r^2 d\Omega \right]$$

ed è  $a_o = a(t_o)$  (fattore di scala e tempo attuali).

Si tratta di un risultato quasi immediato. Anzitutto, a fronte di una *distanza comovente*  $r$ , la distanza fisica attuale della sorgente è  $a_o r$  (eventuali correzioni dovute a uno spazio non euclideo sono di ordine superiore a quello di cui si tiene conto al punto successivo). A quella distanza una sorgente appartenente al flusso di Hubble si allontana a velocità  $v = ca(t)/a_o$  ( $t$  è il tempo di emissione). Ogni fotone è quindi *redshiftato* di un fattore  $a/a_o$  mentre dello stesso fattore cresce l'intervallo temporale  $\Delta_o$  tra gli istanti di ricevimento di due fotoni 1 e 2, rispetto alla durata dell'intervallo temporale  $\Delta$  tra le loro emissioni, semplicemente a causa della maggior distanza che il fotone 2 deve compiere.

Complessivamente, la prima relazione (V.17) diviene

$$(V.21) \quad l = \frac{L}{4\pi(a_o r)^2} \left( \frac{a}{a_o} \right)^2 \equiv \frac{L}{4\pi d^2}$$

e dal confronto tra secondo e terzo membro segue subito eq. (V.19).

3. Si collega la coordinata  $r$  della sorgente emittente al suo *redshift*  $z$ , tenendo conto di come l'andamento di  $a(t)$ , e quindi di  $z(t)$ , varia con la distanza; dal momento che a grandi distanze si è anche nel passato, si possono vedere eventuali effetti di un cambio nella velocità di espansione.

La relazione  $r$ - $z$ , valida per distanze grandi, ma piccole rispetto al raggio di Hubble  $c/H_o$ , è allora la seguente:

$$(V.22) \quad r = \frac{c}{H_o a_o} \left( z - \frac{1 + q_o}{2} z^2 \right);$$

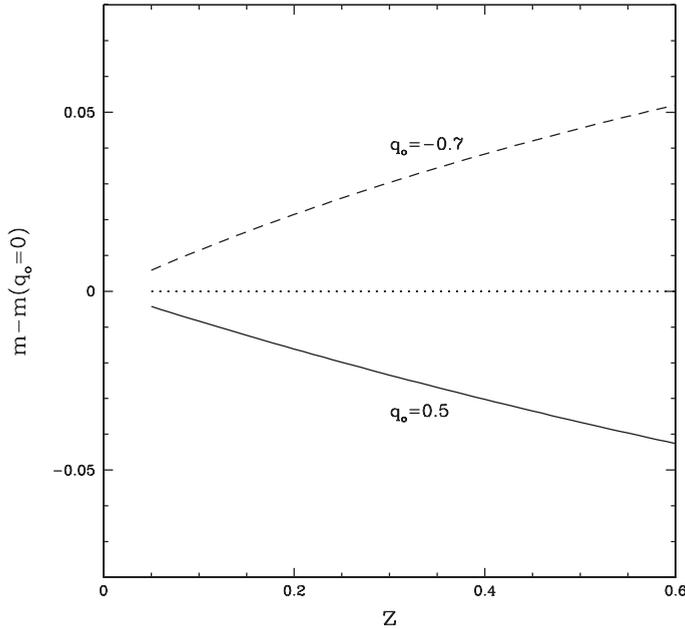
si noti la presenza in questa relazione del parametro di decelerazione  $q_o$ ; l'eq. (V.22), unita a eq. (V.19), permette di ottenere l'espressione per la distanza di luminosità

$$(V.23) \quad d_L = \frac{c}{H_o} \left( z + \frac{1 - q_o}{2} z^2 \right).$$

Solo la prova delle eq. (V.22) e (V.23) richiede degli sviluppi di una certa complessità, che si trovano in appendice A.

Dalla definizione di  $d_L$  segue allora

$$(V.24) \quad \frac{1}{l^{1/2}} \times \frac{H_o \sqrt{L}}{4\pi c} = z + \frac{1 - q_o}{2} z^2$$

**FIGURA V.3**

Andamento previsto della magnitudine apparente di SNIa in funzione del *redshift*. I dati sono normalizzati alle aspettative in un modello con  $q_0 = 0$ .

con  $H_0\sqrt{L}/4\pi c$  che, per candele standard, è costante. Passando ai logaritmi si ricava allora l'andamento della magnitudine apparente in funzione di  $z$ ,

$$(V.25) \quad m = 0.8 \log_{10} \left( z + \frac{1 - q_0}{2} z^2 \right) + K,$$

e il confronto tra eq. (V.25) e i dati osservativi permette la determinazione di  $q_0$ .

In figura V.3 sono mostrati gli andamenti previsti, fino a  $z = 0.6$ , sulla base di queste relazioni. A  $z$  più elevati, il solo parametro  $q_0$  non basta più a formulare una previsione attendibile. In sostanza, l'andamento di  $z$  con  $t$  non è più affidabile se determinato sulla base di parametri che riguardano l'espansione al tempo attuale  $t_0$ .

La figura mette in evidenza che le discrepanze tra modelli sono  $\sim 0.02$  mag, quindi inferiori di un ordine di grandezza alle oscillazioni di luminosità (al picco) tra SNIa. È quindi essenziale l'uso del template che, col riferimento a 9 curve-tipo, permette di dividere in  $\sim 10$  parti l'intervallo di 0.3 mag. A ciò si deve aggiungere il fatto di poter contare su un vasto campione di SNIa a diversi *redshift*.

### **CMB e campioni profondi**

Il risultato su  $q_0$  ottenuto dalle SNIa è chiaramente in favore di  $q_0 < 0$ . Peraltro, questo risultato lascia qualche perplessità, specialmente per la complessità del procedimento osservativo con cui è stato ottenuto. Ne abbiamo descritto qui solo gli elementi principali. Ma, per esempio, il risultato passa attraverso un'accurata disamina dell'assorbimento da polveri, nella Via Lattea e nella galassia ospite.

Questo risultato è però confermato dall'esame delle anisotropie del CMB e della distribuzione delle galassie in campioni profondi.

Una discussione articolata dei dati sul CMB potrà essere fatta solo alla fine del capitolo a esso dedicato. Ancora più complesso è il caso dei campioni profondi. Daremo qui solo le linee guida delle analisi che hanno portato alla conclusione in favore dell'esistenza della DE.

Le anisotropie del CMB danno luogo a fluttuazioni della sua temperatura  $\Delta = \Delta T/T$ , dell'ordine di  $1:10^5$ . L'analisi dei dati fornisce una mappa di  $\Delta(\mathbf{n})$  ( $\mathbf{n}$  contrassegna i siti sulla volta celeste). Uno sviluppo di  $\Delta(\mathbf{n})$  in armoniche sferiche porta allo spettro angolare  $C_l$ . Si tratta di uno sviluppo analogo a quello delle fluttuazioni spaziali  $\varepsilon(\mathbf{x})$ , che fornisce lo spettro  $P(k)$ . La differenza sta nel fatto che la volta celeste ha un'area finita, per cui  $\Delta(\mathbf{n})$  è una funzione periodica e le componenti del suo sviluppo, etichettate da  $l$ , sono un'infinità discreta.

In figura V.4 si mostra l'andamento atteso di  $C$  in funzione di  $l$ . Le curve sono tracciate continue, saldando tra loro i  $C_l$  vicini.

Nel primo riquadro di questa figura viene mostrato il sistema di massimi e minimi che caratterizza lo spettro angolare, per un determinato modello. Negli altri tre riquadri viene magnificata una parte della figura precedente e si procede a confronti con modelli in cui uno dei parametri del modello di partenza è variato. Nella didascalia della figura si descrivono queste variazioni sottolineando la loro diversa natura.

L'ampiezza dello spettro è  $\sim \mathcal{O}(10^{-10})$ , essendo media dei quadrati di fluttuazioni  $\sim \mathcal{O}(10^{-5})$ . Fluttuazioni di quest'ampiezza sono decisamente *lineari* e il confronto osservazioni-previsioni, nell'analisi del CMB, si basa sull'analisi numerica di un sistema di equazioni lineari, il cui risultato è coerente con previsioni puramente analitiche, per cui si veda il capitolo IX.

L'analisi dei campioni di galassie offre una possibilità diretta di confronto col modello, sulle scale in cui le fluttuazioni non sono ancora entrate nella fase fortemente non-lineare. Sappiamo dal capitolo II che per questo si debbono raggiungere scale  $\lesssim 50 h^{-1}\text{Mpc}$ , e averle ben campionate. La profondità di un campione adatto non può quindi essere inferiore a  $\sim 250\text{-}300 h^{-1}\text{Mpc}$ .

Dalla distribuzione delle galassie su queste scale si mira a ricostruire il cosiddetto *spettro trasmesso*, misurato anche tramite la *funzione a 2 punti*. Più dettagli saranno forniti nel capitolo sull'evoluzione lineare delle fluttuazioni.

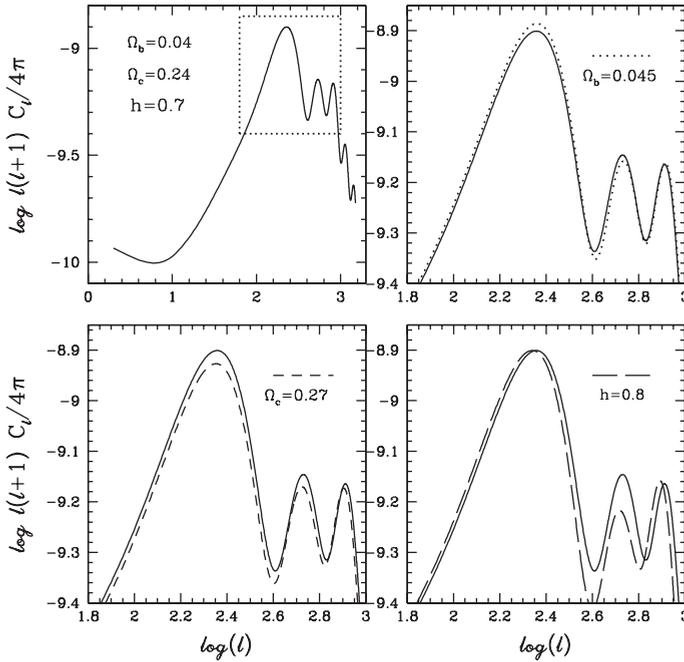


FIGURA V.4

Spettro angolare delle fluttuazioni di temperatura del CMB. Nel primo riquadro in alto lo spettro fino a  $l \simeq 1500$ . Il rettangolo marcato dalla linea punteggiata è magnificato nei riquadri successivi. Nel secondo riquadro in alto gli effetti di una variazione di  $\Omega_b$ : si noti che la curva variata è più alta sul primo massimo e più bassa nei minimi. Nel primo riquadro in basso gli effetti di una variazione di  $\Omega_c$ : si noti che la curva variata si sposta nella stessa direzione su massimi e minimi. Nell'ultimo riquadro gli effetti di una variazione di  $H_0$ : caratteristico (anche se non vistoso) lo spostamento della posizione in  $l$  del primo picco.

La procedura seguita per la valutazione del modello più *verosimile* rispetto ai dati è assai complessa e non verrà trattata in questo volume. Nell'ambito dei modelli  $\Lambda$ CDM con  $\Omega_o = 1$  (vedi il prossimo paragrafo per una completa definizione), la situazione è descritta dalla seguente

### Tabella

	solo CMB	CMB & campione $2dF$
$10 \Omega_b h^2$	$0.2230 \pm 0.0074$	$0.2223 \pm 0.0069$
$\Omega_m h^2$	$0.1265 \pm 0.0080$	$0.1261 \pm 0.0050$
$h$	$0.735 \pm 0.032$	$0.733 \pm 0.020$
$\Omega_m$	$0.237 \pm 0.034$	$0.236 \pm 0.020$
$\sigma_8$	$0.743 \pm 0.051$	$0.737 \pm 0.036$

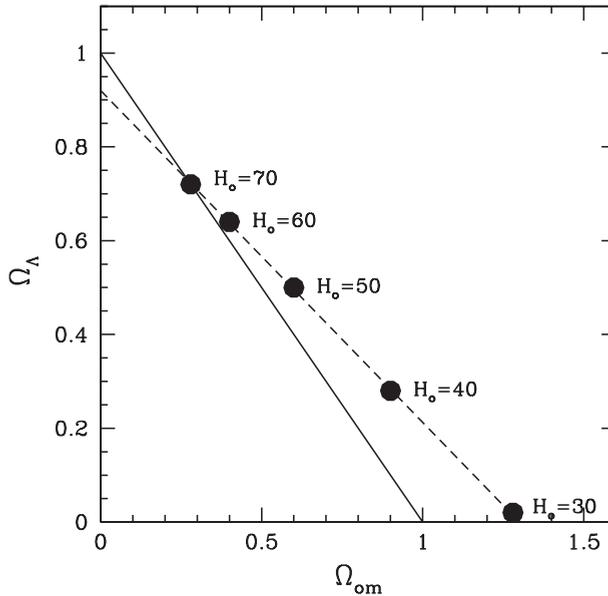


FIGURA V.5

Relazione tra parametro di densità totale e parametri di densità di DE e materia, al variare del parametro di Hubble. La linea continua è il luogo dei modelli spazialmente piatti. La linea tratteggiata è il luogo dei modelli che meglio si adattano ai dati, al variare di  $H_0$ .

I dati del CMB sono quelli del secondo *release* dell'esperimento WMAP. Sulla base degli stessi dati è costruita la figura V.5. Essa mostra come, per avere un modello con  $\Omega_0 \simeq 1$ , occorrono  $\Omega_\Lambda = \Omega_{o,de} \simeq 0.7$  e  $\Omega_{om} \simeq 0.3$ . Ciò prescrive anche  $H_0 \simeq 70$  km/s/Mpc.

Altre possibilità sono però aperte. La linea tratteggiata, in figura, è il luogo di *best-fit* coi dati, al decrescere del valore di  $H_0$ . La figura mostra dunque che un modello chiuso, senza DE, con  $H_0 \sim 30$  km/s/Mpc, non è in disaccordo coi dati sul CMB. Esso è però in assoluto disaccordo coi dati su  $H_0$ , ottenuti dal telescopio spaziale Hubble, che prescrivono valori superiori a  $\sim 0.65$ .

Dai dati della Tabella risalta l'estremo dettaglio con cui i valori dei parametri cosmologici sono oggi ottenibili. Il confronto tra parametri ricavati dai soli dati del CMB, coi parametri ricavati tenendo conto anche del campione profondo di galassie *2dF* mette in evidenza:

- 1) l'assoluta coerenza tra i due sistemi di dati, sicché
- 2) l'effetto di aggiungere i dati *2dF* ai dati del CMB si riduce a un restringimento – peraltro significativo – dell'intervallo d'errore.

Nell'insieme l'incertezza residua sui parametri cosmologici riguarda in genere il terzo decimale e si può veramente concludere che l'analisi dati ha portato ormai a una *cosmologia di precisione*.

## 5. $\Lambda$ CDM

Fino a un decennio fa il modello  $\Lambda$ CDM era un... controesempio, atto a mostrare come *sarebbe evoluto* il cosmo in contesti diversi da quelli considerati più realistici.

Si tratta peraltro di un modello antico, introdotto da Albert Einstein attorno al 1920, quando il cosmo... era fatto di stelle. Alla ricerca di una soluzione cosmologica stazionaria, egli aggiunse una costante nelle sue equazioni – la *costante cosmologica* – per bilanciare l'autogravità dei materiali cosmici. Scoperto il flusso di Hubble, Einstein ricacciò la costante cosmologica nel ricettacolo delle cose inutili, qualificandola anzi come «il più grosso svarione della mia vita».

Invece di introdurre la costante cosmologica come termine *ad-hoc* nelle equazioni, gli stessi effetti dinamici si ricavano immaginando la presenza di una componente con equazione di stato  $p = wp$  e  $w = -1$ . La differenza tra lo schema geometrico e questa visuale più moderna sta nella conseguente necessità di indicare la fisica che sta dietro un'equazione di stato con  $w = -1$ .

La risposta più naturale è che si tratti della cosiddetta *energia del vuoto*. In effetti, nello studio della dinamica degli oscillatori quantistici, si scopre che gli autovalori dell'energia sono  $E_n = 2\pi\hbar\nu(n + 1/2)$ , sicché la transizione tra due livelli implica emissione o assorbimento di un'energia  $2\pi\hbar\nu$ , di norma tramite un fotone. Il livello minimo dell'energia non è però nullo, anche se si dice che la sua energia residua  $E_0 = \pi\hbar$  è *inosservabile*. Questo residuo energetico potrebbe invece giocare un ruolo cosmologico.

Questi dati di prima quantizzazione vanno peraltro riconsiderati nell'ambito della teoria dei campi. Essa mette in evidenza due fatti: (i) Per il vuoto si può effettivamente scrivere un'equazione di stato e il suo parametro di stato è  $w = -1$ . (ii) La densità d'energia del vuoto attesa in teoria dei campi può essere ricondotta a zero ma, se la si assume  $\neq 0$ , risulta maggiore della densità critica  $\rho_{cr}$ , di molti (50, 100) ordini di grandezza.

Una generica sostanza omogenea e isotropa entra nelle equazioni di Einstein tramite il suo tensore energia impulso  $T_{\mu\nu}$ , rappresentato da una matrice  $4 \otimes 4$  sulla cui diagonale principale compaiono  $\rho$ ,  $-p$ ,  $-p$ ,  $-p$ . Una sostanza con

$$(V.26) \quad T_{\mu\nu} = \Lambda g_{\mu\nu} \quad (g_{\mu\nu}: \text{tensore metrico})$$

ha parametro di stato  $w = -1$ . Il “falso vuoto” quantistico sarebbe caratterizzato da  $T_{\mu\nu}$  di questa forma.

Una discussione più approfondita su altre scelte possibili per la DE verrà fatta in un apposito capitolo, dove si dibattono anche i problemi del *fine tuning* (accordatura perfetta) e della *coincidenza*, indotti dalla presenza di DE.

Tra le scelte che si possono operare c'è la semplice possibilità che  $w$  sia negativo ma  $\neq -1$ . Le osservazioni attuali, così stringenti su altri parametri, non lo sono molto sul valore di  $w$ , che può essere scelto nell'intervallo (-1.3,-0.8) senza conflitti coi dati.

Nel caso di  $w > -1$  la DE viene talora denominata *quintessenza*. Nel caso di  $w < -1$ , caso in realtà difficile da inserire nell'ambito di una visuale fisica coerente, la DE viene detta «*phantom DE*».

Tornando al modello  $\Lambda$ CDM, ricordiamo che, nel precedente capitolo, abbiamo considerato gli andamenti del fattore di scala sia in modelli a una sola componente, sia in modelli a 2 componenti con le seguenti scelte di parametri di stato: (i)  $w = 0$ ,  $w = -1$ . (ii)  $w = 0$ ,  $w = 1/3$ .

Questi due casi permettono di trattare il modello  $\Lambda$ CDM. Il caso (ii) sarà adatto al suo studio a cavallo dell'equivalenza tra materia non relativistica e relativistica, che ha luogo a un *redshift*  $z \simeq 10^4$ . Il modello (i) è esattamente il modello  $\Lambda$ CDM, in prosimità dell'epoca attuale, che offre il *best-fit* dei dati disponibili.

Usando i risultati del precedente capitolo, possiamo ricavare il tempo  $t$  in funzione del fattore di scala, una volta noto  $\Omega_{om}$ , tramite la relazione

$$(V.27) \quad \ln \left[ \sqrt{1 + (\Omega_{om}^{-1} - 1) \left(\frac{a}{a_0}\right)^3} + (\Omega_{om}^{-1} - 1)^{1/2} \left(\frac{a}{a_0}\right)^{3/2} \right] = \frac{3}{2} (1 - \Omega_{om})^{1/2} H_0 t.$$

In particolare, per  $a = a_0$ , ricaviamo

$$(V.28) \quad H_0 t_0 = \frac{2}{3} \frac{\ln \left[ \sqrt{\Omega_{om}^{-1} + \sqrt{\Omega_{om}^{-1} - 1}} \right]}{\sqrt{1 - \Omega_{om}}}$$

Per esempio, per  $\Omega_{om} = 0.28$ , si ricava

$$(V.29) \quad t_0 = 1.48 \frac{2}{3H_0}.$$

Per  $H_0 = 71$  (km/s)/Mpc, l'età del cosmo è allora

$$(V.30) \quad t_0 = 13.6 \times 10^9 \text{ anni}.$$

## 6. Conclusioni

I dati di rilevanza cosmologica sono ormai talmente stringenti da determinare molti dei parametri del modello con errori  $\ll 1\%$ . Essi sono chiaramente in favore di un modello  $\Lambda$ CDM in cui il 75% della *sostanza cosmica* è DE, non riconducibile a particelle elementari di alcun tipo. L'80% almeno del restante 25% è poi costituito di DM *non barionica*: una sostanza meno «aliena», nel senso che può essere ricondotta a quanti di campi quantistici, ma pur sempre di natura sconosciuta. Poco più del 4% del cosmo, infine, si compone di *materia barionica*, la sostanza di cui siamo fatti noi stessi e il nostro ambiente.

Si tratta veramente di un ulteriore passo nella direzione del *principio copernicano*: non solo la Terra non è un sito privilegiato nell'Universo, tanto meno il suo centro; non solo l'Uomo non è un Essere con caratteristiche biologiche che lo differenziano dagli altri viventi; anche la sostanza stessa di cui siamo fatti, noi, il nostro pianeta, il nostro ambiente cosmico, è una componente secondaria della sostanza cosmica che tutto pervade.

# Appendice A1

## Dipendenza del *redshift* della distanza e del tempo di emissione di una sorgente

*La distanza di luminosità*

L'andamento del fattore di scala in prossimità di  $t_0$  si può esprimere tramite lo sviluppo in serie di potenze

$$(A1.1) \quad a(t) = a_0 + \dot{a}(t_0)(t - t_0) + \frac{1}{2}\ddot{a}(t_0)(t - t_0)^2 + \dots = a_0 \left( 1 - \varepsilon - \frac{q_0}{2}\varepsilon^2 + \dots \right)$$

dove

$$(A1.2) \quad \varepsilon = H_0(t_0 - t), H_0 = \dot{a}(t_0)/a_0, q_0 = -\ddot{a}(t_0)/a_0 H_0^2.$$

Terremo solo i primi due termini degli sviluppi in serie, in tutta questa appendice, coerentemente alla scelta di misurare solo derivata prima e derivata seconda di  $a(t)$ , al tempo  $t_0$ .

Sarà allora

$$(A1.3) \quad 1 + z = \frac{a_0}{a} \simeq \frac{1}{1 - \varepsilon - \frac{q_0}{2}\varepsilon^2} \simeq 1 + \varepsilon + \left( 1 + \frac{q_0}{2} \right) \varepsilon^2$$

e questa relazione, invertita, darà

$$(A1.4) \quad \varepsilon = z - \left( 1 + \frac{q_0}{2} \right) z^2.$$

Parallelamente, essendo  $cdt \simeq a(t)dr$  sul cono luce (si trascura qui il denominatore di  $dr$ , che sarebbe  $\sqrt{1 - \kappa r^2}$ , in quanto correzioni dovute a curvatura  $\kappa \neq 0$  sono di ordine superiore al secondo), da eq. (A1.3) possiamo anche ricavare

$$r(t) = c \int_t^{t_0} \frac{dt'}{a(t')} = \frac{c}{a_0} \int_t^{t_0} dt' [1 + \varepsilon(t') + \dots] \simeq \frac{c}{H_0 a_0} \left( \varepsilon + \frac{\varepsilon^2}{2} \right)$$

e usare in questo risultato l'espressione (A1.4) di  $\varepsilon$ . Avremo così

$$(A1.5) \quad r \simeq \frac{c}{H_0 a_0} \left( z - \frac{1 + q_0}{2} z^2 \right)$$

cioè l'eq. (V.22) nel testo del capitolo.

Di qui si ricava

$$d_L = r a_0 \times \frac{a_0}{a} = \frac{c}{H_0} \left( z - \frac{1 + q_0}{2} z^2 \right) (1 + z) = \frac{c}{H_0} \left( z + \frac{1 - q_0}{2} z^2 \right),$$

(eq. V.23), sempre trascurando termini di grado superiore al secondo in  $z$ .

## CAPITOLO VI

# L'epoca della nucleosintesi

### 1. Elio nelle stelle

Nelle specole dell'ultimo Ottocento e all'inizio del Novecento, lo spettroscopio divenne lo strumento principe della ricerca astronomica; per un certo numero di anni, la Società Astronomica Italiana cambiò persino nome in *Società degli Spettroscopisti Italiani*. Uno dei frutti di quell'epoca, in cui si faceva spettroscopia dai tetti delle città, fu la suddivisione delle stelle in Popolazione I e II, rispettivamente con o senza importanti righe di assorbimento *metalliche*. Fu in quel periodo che invalse tra gli astronomi l'uso di suddividere la totalità degli elementi in tre parti: idrogeno, elio, e il resto, i metalli appunto.

La suddivisione in Popolazioni è rimasta, anche se l'assorbimento varia con continuità e registra la distribuzione delle età, testimoniando il graduale arricchimento, in precedenti fucine stellari, dei materiali cosmici di cui si compone la stella. Poca cosa, comunque, i metalli nelle stelle; al più l'1% della massa nelle stelle giovani, ma anche meno dello 0.1% nelle stelline di estrema popolazione II. L'abbondanza di elio, invece, è elevata, attorno al 24%.

Sappiamo che le stelle della *Sequenza Principale* sono sede di bruciamento di idrogeno in elio. Può essere allora che il 24% di elio sia nato lì dentro e sia stato poi rimesso poi in circolo da *venti stellari* o dall'esplosione di *supernovae*?

Un fatto è però subito evidente nelle osservazioni: c'è una correlazione tra tasso di elio e tasso di metalli; tuttavia, al tendere a zero del tasso metallico, il tasso di elio cala ma tende verso un valore non nullo (vedi Fig. VI.1). Supernove e venti stellari arricchiscono dunque i materiali galattici non solo di metalli ma anche di elio, forgiati nelle stelle, ma l'elio di origine stellare si accumula su un consistente tasso preesistente.

La stima dell'abbondanza *primordiale* di  ${}^4\text{He}$  utilizza gli spettri delle galassie nane irregolari e delle galassie blu compatte. I tassi di  ${}^4\text{He}$  osservati nelle popolazioni stellari di queste galassie sono confrontati con le (ridotte) abbondanze in metalli e l'andamento viene estrapolato fino a un'abbondanza nulla in metalli, come in figura VI.1. La stima della frazione di massa barionica in elio è

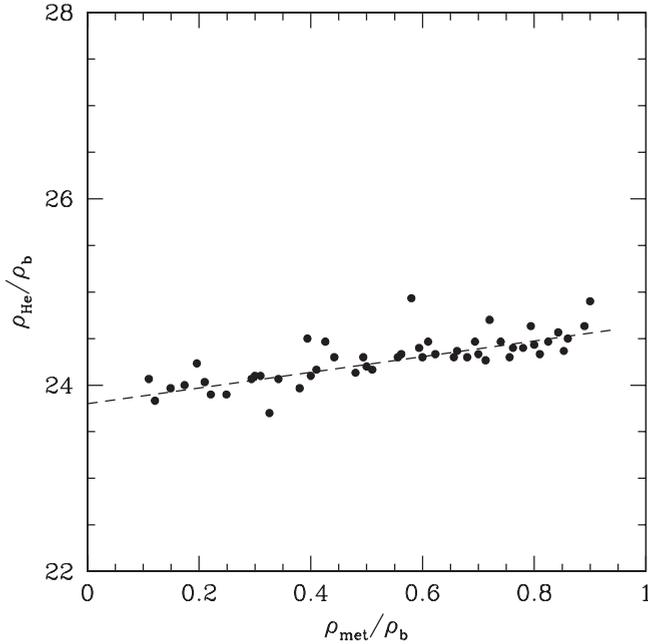


FIGURA VI.1

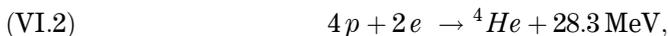
Esempio di relazione tra tasso di elio e tasso di metalli nelle popolazioni stellari. Al tendere a zero del secondo, in popolazioni molto antiche, il tasso di elio cala, verso un valore  $\sim 24\%$ , antecedente al contributo della nucleosintesi stellare. La linea tratteggiata rappresenta un *fit* lineare alle abbondanze osservative.

la seguente

$$(VI.1) \quad Y_p = 0.238 \pm (0.002)_{stat} \pm (0.005)_{sist}.$$

Quindi, la piccola incertezza legata all'attendibilità della procedura (errore sistematico) è nettamente superiore all'incertezza statistica.

Quando i dati osservativi non avevano l'attuale livello di precisione, si era anche argomentato che, se l'origine di tutto l'elio fosse stellare, si dovrebbe vedere la radiazione prodotta nel processo



che avviene nelle stelle e le fa brillare. I 28.3 MeV (*difetto di massa* di  ${}^4He$ ) sono poco più dello 0.75% della massa barionica coinvolta nella reazione, sicché circa un quarto di questa percentuale, pari a  $\sim 0.18\%$  della massa barionica totale, sarebbe divenuta radiazione, quando si formava l'elio.

Dal momento che  $\rho_{o,b}$  è una frazione  $\Omega_{o,b} \simeq 0.04$  della densità critica, la densità della

radiazione in banda ottica, nata assieme a  ${}^4\text{He}$  sarebbe allora

$$(VI.3) \quad \rho_{opt} = \Omega_{o,b} \times 1.8 \times 10^{-3} \times \rho_{o,cr} \simeq 7 \cdot 10^{-5} \rho_{o,cr},$$

laddove il fondo microonde ha densità

$$(VI.4) \quad \rho_{CMB} \sim 10^{-4} \rho_{o,cr}.$$

Un fondo ottico intenso come il fondo microonde (CMB) sarebbe uno dei dati più evidenti all'osservazione astronomica! Questo fondo non esiste.

Prima che i dati divenissero stringenti, si era pensato di spostare le stelle produttrici di elio a alto *redshift* ( $z_{III}$ ), così che la radiazione di questa Popolazione III, attenuata da un fattore  $1/(1+z_{III})^2$ , fosse nell'IR. Ma, anche con valori estremi, come  $z_{III} \sim 50$ , i dati odierni non lasciano scampo.

La Popolazione III è peraltro sopravvissuta al fallimento dell'idea. Oltre a originare i pochi metalli rilevati nelle stelline di estrema Popolazione II, essa è presumibilmente responsabile di una crescente emissione UV, che reionizza l'idrogeno del cosmo, sicché negli spettri dei QSO più remoti (fino a  $z \sim 5$  e anche oltre) ci son pochi segni di assorbimento da idrogeno atomico o molecolare.

Riconoscendo l'elio come un problema, sin dagli anni Quaranta Gamow aveva proposto che esso avesse origine in un *Big-bang caldo*; il suo ragionamento si basa sul confronto tra equilibrio termodinamico e dinamica delle reazioni capaci di produrlo.

Indichiamo le densità in numero di protoni e neutroni con  $n_p$  e  $n_n$ . Nel prossimo paragrafo si dimostra che, in condizioni di equilibrio, la differenza tra masse

$$(VI.5) \quad \Delta m = m_n - m_p \simeq 1.3 \text{ MeV} \simeq 1.5 \cdot 10^{10} \text{ K}$$

(si noti che 1 MeV e  $10^{10}$  K hanno lo stesso ordine di grandezza) fissa il rapporto

$$(VI.6) \quad n_n/n_p \simeq \exp(-\Delta m/T).$$

Finché  $T \gg \Delta m$ , il rapporto non si discosta dall'unità: stesso numero di protoni e neutroni. Quando poi  $T$  approssima  $\Delta m$ , e il rapporto  $n_n/n_p$  cala.

Il problema sta nel *congelare* questo rapporto a 1:7. Se ogni 8 barioni c'è 1 neutrone, quando i neutroni si combineranno con protoni a formare elio, si avrà il 25% di elio. Ebbene, il congelamento è possibile se la simultanea rarefazione del cosmo blocca le reazioni che garantiscono l'equilibrio a una temperatura  $\sim 0.6-0.7 \Delta m$ . Poiché la conversione neutrone-protone è una reazione debole con neutrini, si tratta di tarare il rapporto neutrini/barioni – e quindi fotoni/barioni – in modo che, raggiunto il rapporto 1:7, tutto si congeli.

Non sapendo nulla del CMB, Gamow, dalla temperatura alla nucleosintesi primordiale (BBN o BBNS: *Big-bang nucleosynthesis*) ricavò la temperatura del cosmo di oggi, valutandola attorno ai 4K. Il resto di questo capitolo sarà la versione aggiornata e *corretta* di questo schema. Ciò che a Gamow sfuggì fu che una *radiazione fossile* di qualche K era perfettamente osservabile da terra.

## 2. Processi dovuti a forze deboli nell'era leptonica

L'analisi degli eventi fino alla sintesi di  ${}^4\text{He}$ , in realtà, è un po' più articolata, ma, seguendola in dettaglio, si giunge a precise previsioni quantitative.

I processi rilevanti sono i seguenti:

- 1) conversione neutrone/protone in equilibrio termodinamico;
- 2) conversione neutrone/protone dopo la fuoriuscita dall'equilibrio termodinamico;
- 3) fuoriuscita dall'equilibrio e formazione del fondo neutrinico;
- 4) annichilazione degli elettroni e riscaldamento del fondo fotonico;
- 5) apertura del *deuterium bottleneck* e reazioni di sintesi di  ${}^4\text{He}$ ;
- 6) reazioni collaterali e produzione di altri nuclidi leggeri.

### L'era leptonica

Il punto di partenza è la nascita dei barioni. Essa avviene a una temperatura  $T_{qh} \sim 100 \text{ MeV}$ , quando tutti i *quark* si combinano negli *adroni*, la forma attuale della materia con *interazioni forti*. Se nel cosmo non ci fosse un contenuto netto di *numero barionico*, appena nati, gli adroni sarebbero prossimi all'estinzione: la massa del mesone  $\pi$ , l'adrone più leggero, è infatti  $m_\pi \simeq 140 \text{ MeV} \sim T_{qh}$ .

Con la scomparsa dei mesoni  $\pi$  ha inizio l'*era leptonica*. Il *brodo termico* è allora composto da coppie  $e^+$ ,  $e^-$ ,  $\nu_\tau$ ,  $\nu_\mu$ ,  $\nu_e$  (neutrini dei 3 possibili sapori) e  $\gamma$ . Per un breve tempo, del brodo termico fanno parte anche le particelle  $\mu^+$ ,  $\mu^-$ , la cui massa ( $\simeq 105 \text{ MeV}$ ) è di poco inferiore a  $m_\pi$ . Per  $\mu$  e  $\pi$ , quando  $T < m$  ( $= m_\mu$ ,  $m_\pi$ ), la densità in numero ha l'espressione (di *Boltzmann*)

$$(VI.7) \quad n = \mathcal{N} \left( \frac{mT}{2\pi} \right)^{3/2} e^{-m/T}$$

( $\mathcal{N}$ : numero di stati di carica o di spin).

Oltre al brodo termico rimangono presenti i *protoni* e i *neutroni* che garantiscono una densità  $n_b = n_n + n_p$  di numero barionico. Anche per  $n_n$  e  $n_p$  vale l'espressione di Boltzmann, che va però integrata con un opportuno *potenziale chimico*  $\mu$  associato al numero barionico, sicché risulta

$$(VI.8) \quad n_{p,n} = \mathcal{N} \left( \frac{m_{p,n}T}{2\pi} \right)^{3/2} \exp\left(\frac{-m_{p,n} + \mu_b}{T}\right).$$

$\mu_b$  dipende da  $T$  e si ricava imponendo che sia  $n_n + n_p = n_b$ . Viceversa, nel rapporto

$$(VI.9) \quad \frac{n_n}{n_p} = \left( \frac{m_n}{m_p} \right)^{3/2} \exp\left(\frac{-m_p + m_n}{T}\right),$$

la differenza di masse è significativa solo all'esponente e, finché permangono le con-

dizioni di equilibrio, il rapporto neutroni/protoni è

$$(VI.10) \quad n_n/n_p \simeq \exp(-\Delta m/T).$$

La sezione d'urto debole è peraltro modesta e i processi deboli stanno per uscire dall'equilibrio.

### **Disaccoppiamento dei neutrini**

Vediamo le reazioni deputate a garantire l'equilibrio. Oltre agli urti

$$(n \leftrightarrow p) \quad n + \nu_e \leftrightarrow p + e, \quad n + e^+ \leftrightarrow p + \bar{\nu}_e;$$

va considerato il *decadimento beta*

$$(dec. \beta) \quad n \rightarrow p + e + \bar{\nu},$$

che, come ogni decadimento, risente poco delle condizioni esterne.

Più in generale, si possono confrontare due tipi di reazioni:

- 1) un neutrone o un protone collide con un neutrino o un elettrone, in una delle reazioni ( $n \leftrightarrow p$ );
- 2) neutrini e elettroni collidono tra loro anelasticamente o elasticamente:

$$(annich.) \quad \nu_e + \bar{\nu}_e \leftrightarrow e^+ + e^-,$$

$$(elastic) \quad \nu_e + e^\pm \rightarrow \nu_e + e^\pm,$$

Si noti anche che, finché è  $T > m_e \simeq 0.511$  MeV, le densità in numero di elettroni e neutrini sono circa eguali. Nelle reazioni 2) non intervengono adroni, ma in entrambi i casi la collisione è dovuta a forze deboli e la natura adronica di protoni e neutroni è irrilevante. Nelle espressioni dei tempi di collisione,

$$(VI.11) \quad \tau = 1/\sigma_w n v,$$

compaiono quindi comunque le sezioni d'urto deboli

$$(VI.12) \quad \sigma_w = 2.6 \times 10^{-42} \text{cm}^2 (T/\text{MeV})^2 \equiv \sigma_1 (T/\text{MeV})^2,$$

la densità in numero dei bersagli  $n$  e la velocità della particella incidente  $v$ . Se la particella incidente è relativistica ( $\nu$  o  $e$ ), si ha  $v \simeq c$ ; se la particella incidente è un protone o un neutrone non relativistico, per cui risulta  $mv^2/2 = (3/2)k_B T$ , si ha (con costante di Boltzmann  $k_B = 1$ , cioè usando unità energetiche per la temperatura),

$$(VI.13) \quad v \simeq c \sqrt{3T/m_b} \quad (m_b: \text{massa barione})$$

Per le reazioni (*annich.*) e (*elastic*), risulta  $\tau = t$  alla temperatura

$$(VI.14) \quad T_{v,dg} \simeq 900 \text{ keV} \simeq 1.05 \times 10^{10} \text{ K},$$

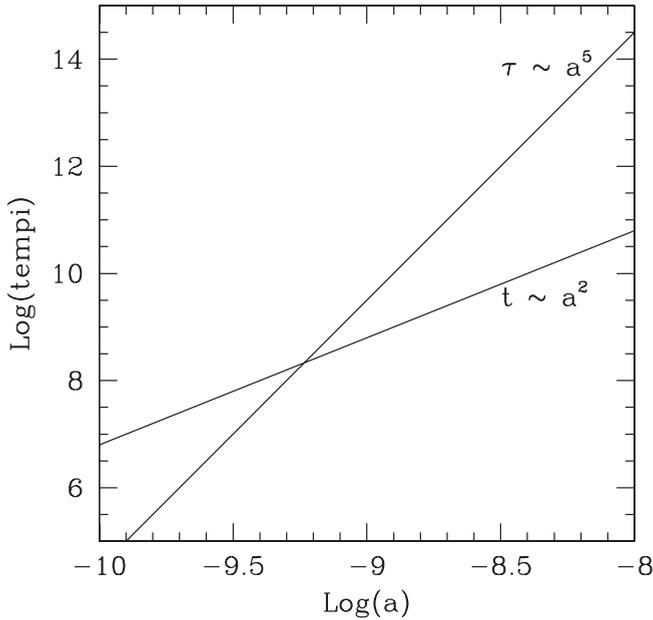


FIGURA VI.2

Crescita del tempo cosmico e del tempo di collisione per processi deboli, nell'era leptonica.

di *disaccoppiamento dei neutrini*. In figura VI.3 è rappresentato l'andamento di  $\tau$  e  $t$  al crescere del fattore di scala. Le densità in numero sono  $\propto a^{-3} \propto T^3$ ; le reazioni deboli, poi, hanno sezione d'urto  $\propto T^2$  (eq. VI.12). Sarà quindi  $\tau \propto T^{-5} \propto a^5$ . In era radiativa sarà anche  $t \propto a^2$  e quindi avremo  $\tau/t \propto a^3 \propto T^{-3}$ . Per esempio, a  $T = 2$  MeV, il tempo medio tra due successive collisioni è ancora  $\sim 0.1 t$ . Viceversa, per  $T = 0.51$  MeV ( $= m_e$ ), il rapporto  $\tau/t$  è salito a  $\sim 3.5$ . Ciò significa, per esempio, che solo un elettrone su 1000, in un tempo confrontabile con  $t$ , subisce un urto capace di trasformarlo in neutrino.

La crescita indefinita del rapporto  $\tau/t$  significa che, una volta disaccoppiati, i neutrini restano «isolati» per sempre. Il fondo neutrino, simile al fondo fotonico, è quindi assai difficile da sondare sperimentalmente; la sua presenza è peraltro verificabile sulla base del contributo dei neutrini all'espansione cosmica.

Un fatto importante, dimostrato in appendice A2, è che lo spettro del fondo neutrino, pur in assenza di interazioni, mantiene la forma

$$(VI.15) \quad I_\nu(p) = \frac{\mathcal{N}_\nu}{(2\pi)^3} \frac{p^3}{\exp(p/T) + 1},$$

assai simile a una planckiana (a parte il segno di somma al denominatore);  $p$  è l'impulso dei neutrini e  $\mathcal{N}_\nu$  è il loro numero di stati di spin; il parametro  $T$ , che in condizioni di equilibrio è la temperatura, evolve indefinitamente col fattore di scala, secondo la relazione  $Ta = \text{const.}$

Se la massa dei neutrini è nulla o trascurabile rispetto a  $T$ , l'espressione (VI.15) coincide con quella di equilibrio termodinamico. Un fatto analogo avviene il CMB, che non è più in equilibrio termodinamico da  $z \sim 10^3$ , ma cui si attribuisce una «temperatura» odierna di 2.72 K, sebbene la temperatura perda il suo significato termodinamico, una volta usciti dall'equilibrio.

I dati sui neutrini atmosferici e solari richiedono che almeno un sapore neutrino abbia massa  $\gtrsim 0.2$  eV. Tale massa è trascurabile, al disaccoppiamento, e lo spettro allora ha comunque la forma (VI.15). In assenza di forze che redistribuiscano poi i neutrini sulle energie, la forma dello spettro rimane invariata, anche quando  $T$  divenga confrontabile o inferiore alla massa.

Tempo e temperatura, in quest'era radiativa, sono legati dalla relazione  $1/2t = (8\pi^3 g/90)^{1/2} T^2/m_p$ , sicché il prodotto  $tT^2 g^{1/2}$  è costante e risulta

$$(VI.16) \quad \frac{t}{\text{sec.}} \times \left( \frac{T}{\text{MeV}} \right)^2 = \frac{2.42}{g^{1/2}}, \quad \text{dove } g = \mathcal{N}_{bos} + \frac{7}{8} \mathcal{N}_{fer}$$

e  $\mathcal{N}_{bos}$  ( $\mathcal{N}_{fer}$ ) è il numero di stati di spin bosonici (fermionici) nel brodo termico al tempo  $t$ . Per  $T \simeq 900$  keV, gli stati di spin bosonici sono 2 (fotoni), quelli fermionici sono 10 (4 per  $e^\pm$ , 6 per 3 sapori neutrini). Sarà quindi  $g = 10.75$ . L'età dell'Universo, per  $T = 900$  keV è quindi  $t_{T=900 \text{ keV}} = 0.904$  sec.

Circa 0.9 secondi dopo il *Big-bang*, dunque, il fondo neutrino si disaccoppia e diviene una realtà sconnessa dalle altre componenti cosmiche. Viceversa, se un esperimento riuscisse a sondare il fondo neutrino, ne trarrebbe informazione sulle condizioni del cosmo 0.9 secondi dopo il *Big-bang*.

### ***Fuoriuscita dall'equilibrio della conversione protone-neutrone***

Passiamo ora alle reazioni che garantiscono l'equilibrio termodinamico tra  $p$  e  $n$ , quelle di cui Gamow studiava il congelamento (*freezing*). Applicando la condizione  $\tau = t$  alle reazioni ( $n \leftrightarrow p$ ), si ricava una temperatura molto prossima a 900 keV, bilanciandosi tra loro la presenza dei 3 quark «in» ciascun nucleone, il fatto che gli elettroni non sono più completamente ultra-relativistici e il valore non relativistico delle velocità di protoni e neutroni.

Per  $T = T_{fr} \simeq 900$  keV, l'eq. (VI.10) dà  $n_n/n_p \simeq 0.22-0.23$  e quindi

$$(VI.17) \quad \bar{X}_{n,fr} \equiv n_n/n_b = 1/(1 + n_p/n_n) \simeq 0.18$$

A  $T < T_{fr}$ , la trasformazione di neutroni in protoni tuttavia continua, prima per un residuo di reazioni a due corpi, poi per decadimento beta, sicché sarà

$$(VI.18) \quad X_n = X_{n,fr} \exp[-(t - t_{fr})/\tau_\beta], \quad \text{con } \tau_\beta = 887(\pm 2) \text{ sec}, \quad X_{n,fr} = 0.165;$$

la riduzione da  $\bar{X}_{n,fr}$  a  $X_{n,fr}$  tiene conto del residuo di reazioni a due corpi.

Il tempo  $\tau_\beta$  è dunque lungo, poco meno di un quarto d'ora, quindi  $\sim 1000$  volte  $t_{fr}$ . Peraltro, se i neutroni non vengono opportunamente «fissati» in nuclidi, dopo un paio

d'ore sarebbero praticamente scomparsi e tutti i nuclidi diversi da  ${}^1H$ , nel cosmo odierno, avrebbero dovuto formarsi nelle stelle.

Prima che avvengano reazioni capaci di «salvare» i neutroni, tuttavia, il cosmo si raffredda al di sotto della massa dell'elettrone, con le conseguenze discusse nel prossimo paragrafo.

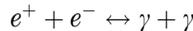
### 3. Annichilazione degli elettroni e temperatura del fondo neutrinico

Quando la temperatura  $T$  scende molto al di sotto di  $m_e \simeq 0.51$  MeV, la densità di equilibrio degli elettroni, conformemente a eq. (VI.8) avrà l'espressione

$$(VI.19) \quad n_e = \mathcal{N}_e \left( \frac{m_e T}{2\pi} \right)^{3/2} e^{(-m_e + \mu_l)/T}$$

e continuerà a declinare esponenzialmente, finché restano solo gli  $e^-$  necessari a mantenere la *neutralità di carica* elettrica, bilanciando la carica positiva dei protoni, e la carica di *numero leptonic*, cui è collegato il potenziale chimico  $\mu_l$ , che bilancia il numero barionico presente nel cosmo.

Al mantenimento del tasso di equilibrio (VI.19) sono deputate le reazioni



e, quando  $T$  scende sotto 0.51 MeV, mentre tutti gli urti tra elettroni «possono» dare fotoni, solo parte degli urti  $\gamma$ - $\gamma$  possono *fotoprodurre* coppie  $e^+$ ,  $e^-$ , quelli con energia superiore alla soglia della reazione:  $\varepsilon_1 \varepsilon_2 > (m_e c^2)^2$  ( $\varepsilon_{1,2}$  sono le energie dei due fotoni che collidono; ricordiamo che il valore medio dell'energia dei fotoni è  $\langle \varepsilon \rangle \simeq 2.8 k_B T$ ). Di per sè, tuttavia, questo fatto non è determinante: finché il tempo di collisione per la reazione tra  $\gamma$  «*sopra-soglia*» è minore dell'età del cosmo, l'abbondanza di elettroni è dettata dalla condizione di equilibrio.

Di fatto, a causa dell'alto valore della sezione d'urto elettromagnetica, le proporzioni di equilibrio sono mantenute fino all'esaurimento di (quasi) tutti gli  $e^+$ . Lasciamo al lettore il calcolo dell'abbondanza residua di  $e^+$ , in violazione di eq. (VI.19), quando il tempo di collisione con gli elettroni, necessario a produrre la reazione di annichilazione in  $2 \gamma$ , supera l'età del cosmo.

### Riscaldamento del fondo $\gamma$

Gradualmente, quindi, il cosmo si svuota delle coppie  $e^+$ ,  $e^-$ , e il fondo fotonico si riscalda, a causa del significativo input di energia e entropia. La conseguenza principale di tale processo è che oggi il fondo neutrinico avrà una «temperatura»

$$(VI.20) \quad T_\nu = (4/11)^{1/3} T_\gamma.$$

Se oggi è  $T_{\nu,\gamma} \simeq 2.72$  K, sarà anche  $T_{\nu,\nu} \simeq 1.94$  K.

Ciò si dimostra sulla base della conservazione dell'entropia comovente. (Ricordiamo che nell'espansione cosmica l'entropia è conservata; input di entropia si possono avere solo quando il cosmo recupera condizioni di equilibrio termodinamico dopo lunghe fasi di disequilibrio, per esempio quando una transizione di fase del primo ordine avviene dopo un lungo *supercooling*. L'energia, invece, non si conserva, nel corso dell'espansione cosmica, come si vede banalmente dal *redshift* dei fotoni.) Infatti: l'entropia attuale dei  $\gamma$  del CMB coincide con l'entropia di  $\gamma$  e elettroni alla temperatura  $T_{v,dg}$  (disaccoppiamento dei  $\nu$ ):

$$(VI.21) \quad \text{entropia di } \gamma \text{ e } e^\pm: \quad S_{\gamma,e} = \frac{4\pi^2}{3 \cdot 30} g_{\gamma,e} (T_{v,dg} a_{v,dg})^3 = \frac{4\pi^2}{3 \cdot 30} g_\gamma (T_{\gamma,o} a_o)^3$$

*prima dell'annich. }  $e^+, e^-$                       *dopo l'annich. }  $e^+, e^-$**

Invece, l'entropia dei neutrini del fondo neutrinico resta invariata, tra la temperatura di disaccoppiamento e quella attuale:

$$(VI.22) \quad \text{entropia dei } \nu: \quad S_\nu = \frac{4\pi^2}{3 \cdot 30} g_\nu (T_{v,dg} a_{v,dg})^3 = \frac{4\pi^2}{3 \cdot 30} g_\nu (T_{v,o} a_o)^3$$

In queste relazioni:

$a_{v,dg}$ : fattore di scala alla temperatura  $T_{v,dg}$ ;  $a_o$ : fattore di scala attuale;

$$g_\nu = (7/8) \times 6, \quad g_\gamma = 2, \quad g_{\gamma,e} = 2 + (7/8) \times 4 = 11/2.$$

Inoltre,  $T_{\gamma,o}$  è la temperatura attuale del CMB, altrove indicata come  $T_o$ .

Dividendo le due equazioni membro a membro si ricava facilmente che  $(g_{\gamma,e}/g_\nu) = (g_\gamma/g_\nu)(T_{\gamma,o}/T_{v,o})^3$ , ovvero:  $T_{\gamma,o} = T_{v,o}(g_{\gamma,e}/g_\gamma)^{1/3} = T_{v,o}(11/4)^{1/3}$ , relazione che coincide con l'espressione (VI.20).

Lo stesso rapporto tra temperature dei due fondi permane a tutti i tempi successivi all'annichilazione dei neutrini. Ciò si ricava dalle stesse eq. (VI.21), (VI.22), sostituendo, al fattore di scala  $a_o$  e alle corrispondenti temperature  $T_{\gamma,o}$  e  $T_{v,o}$ , un generico fattore di scala  $a$  e le corrispondenti temperature  $T_\gamma$  e  $T_\nu$ .

La densità di energia delle componenti radiative, dopo l'annichilazione degli elettroni, sarà data dall'espressione

$$(VI.23) \quad \rho_r(T) = \frac{\pi^2}{30} \left[ 2T^4 + \frac{7}{8} N_\nu 2T_\nu^4 \right] = \frac{\pi^2}{30} g_{eff} T^4$$

con  $g_{eff} = 2 + 2N_\nu(7/8)(4/11)^{4/3} = 2 + 0.4542 N_\nu,$

valida sin tanto che l'eventuale massa dei neutrini possa essere trascurata ( $N_\nu$ : numero di sapori neutrinici supposti presenti nel fondo neutrinico; se ne fanno parte i neutrini elettronico, muonico, tauonico, sarà  $N_\nu = 3$ ).

Si noti come, dopo il disaccoppiamento dei neutrini e, comunque, finché la materia non relativistica ha una densità trascurabile rispetto alle componenti relativistiche, la

relazione tempo-temperatura (VI.16) si possa scrivere

$$(VI.24) \quad \frac{t}{\text{sec.}} \times \left( \frac{T}{\text{MeV}} \right)^2 = 1.31 u(N_\nu) \quad \text{con} \quad u(N_\nu) = \sqrt{\frac{3.363}{2 + 0.4542 N_\nu}}.$$

Si noti che  $u(3) = 1$ .

### ***Entropia per barione***

Il valore del rapporto  $T_\gamma/T_\nu$  si calcola, dunque, assumendo la conservazione dell'entropia comovente, dall'annichilazione degli elettroni a oggi, e quindi anche la conservazione del numero comovente di fotoni. Ciò può apparire abbastanza paradossale: a ogni istante, fotoni vengono prodotti e/o assorbiti. Il fatto è che i fotoni del CMB sono così tanti, che tutto ciò non conta.

All'inizio di questo capitolo, per esempio, abbiamo confrontato la radiazione delle stelle col CMB, trovando che l'intensità di un fondo ottico d'origine stellare sarebbe confrontabile a quella del CMB, se *tutto l'elio* venisse prodotto nelle stelle. Ne viene invece prodotto, al più, qualche per cento. Per la conta dei fotoni, bisogna poi tenere presente che ciascun  $\gamma$  ottico ha energia  $\sim 10^4$  volte maggiore dell'energia dei  $\gamma$  del CMB. Per ogni fotone prodotto dalle stelle, ci sono quindi quasi un milione di fotoni nel CMB.

A parte la smaccata inferiorità numerica, i fotoni stellari avrebbero poi l'imprecindibile connotato distintivo della frequenza. Quando contiamo i fotoni del CMB, non c'è ragione di preoccuparsi delle accrezioni di origine stellare, diretta o indiretta, eventualmente degradati fino al *range* delle microonde.

Un problema più serio potrebbe essere invece costituito da  $\gamma$  iniettati nel cosmo a *redshift*  $z_{in} \geq 10^7$ . Fino a  $z = 10^7$ , infatti, sono attivi processi capaci di redistribuire termicamente gli input energetici. Un input fotonico riscalderebbe quindi ulteriormente il CMB, riducendo in proporzione la temperatura attuale del fondo neutrinico.

Un simile input non è un'ipotesi del tutto peregrina. Immaginiamo, per esempio, particelle neutre massive con tempi di vita  $\sim 10^5$ - $10^6$ sec., magari prodotte quando si ha la rottura *soft* delle SUSY e prive di interazioni apprezzabili fino al momento del decadimento. Sulla base di eq. (VI.24) vediamo che, al loro decadimento, la temperatura del cosmo sarebbe  $\sim 10^{-2}$ - $10^{-3}$ eV  $\sim 10^7$ - $10^8$ K, corrispondenti appunto a un *redshift*  $z_{in} \sim 10^7$ .

Se, al momento del decadimento, queste particelle avessero densità d'energia  $\rho_o \simeq (\pi^2/15)T_\gamma^4$  e il decadimento ha luogo in 2 fotoni, dopo il decadimento la densità dei fotoni sarebbe  $\rho_1 \simeq 2\rho_o$  e  $T_\gamma$  sarebbe salita del  $\sim 20\%$ . Si noti che la densità di un fondo non-relativistico è  $\propto a^{-3}$ , mentre quella di  $\gamma$  e  $\nu$  è  $\propto a^{-4}$ . Per  $T \sim 10^8$ - $10^9$  K, fino alla fine della BBN, quindi, il contributo di questo fondo alla densità d'energia complessiva del cosmo, sarebbe stato  $< 10\%$ . Ciò che cambia radicalmente, invece, è la densità d'energia dei  $\gamma$ , durante la BBN, che viene ridotta di un fattore 1/2, sicché tutto lo scenario della BBN va rivisto.

Un input di questa natura è un esempio delle modifiche proposte alla BBN, e che portano alla modifica delle sue previsioni. Il fatto che queste ultime collimino straordinariamente coi dati osservativi sulle abbondanze di nuclidi leggeri è il vincolo più potente a questo genere di speculazioni.

Se dunque, dall'annichilazione degli elettroni a oggi, non ci sono altri *input* di entropia, sulla base delle osservazioni di oggi possiamo calcolare il rapporto tra  $n_\gamma$  e  $n_b$  («fotoni per barione»), che in tutto il periodo considerato sarà invariante, come l'entropia e il numero barionico comoventi.

Detta  $T_o$  la temperatura attuale del CMB, la densità in numero dei suoi fotoni CMB sarà

$$(VI.25) \quad n_\gamma(T_o) = \frac{\zeta(3)}{\pi^2} 2 T_o^3 = 0.243 \left( \frac{k_B T_o}{\hbar c} \right)^3 = 0.243 \times (13.7 \text{ cm}^{-1})^3 \simeq 415 \text{ cm}^{-3}.$$

Il numero medio di barioni per  $\text{cm}^3$  è invece dato da

$$(VI.26) \quad n_b \simeq \frac{\rho_{cr}}{m_b} \Omega_b = \frac{1.88 \times 10^{-29} \text{ g/cm}^3 \Omega_b}{1.67 \times 10^{-24} \text{ g}} = 1.13 \times 10^{-5} \Omega_b h^2 \text{ cm}^{-3}.$$

Ne segue che

$$(VI.27) \quad \frac{b}{n_\gamma} = 2.72 \times 10^{-8} \Omega_b h^2 = 10^{-9.26} \frac{\Omega_b h^2}{0.02}$$

e, se è  $\Omega_b h^2 \simeq 0.02$ , ci sono poco meno di due miliardi di fotoni per ogni barione.

Nel periodo considerato, il brodo termico appare costituito di soli  $\gamma$ , più gli inosservabili  $\nu$ . Ma, se andiamo a prima dell'annichilazione degli elettroni, avremo un numero di coppie  $e^+$ ,  $e^-$  confrontabile col numero di  $\gamma$ . Cui si deve aggiungere quel piccolo residuo di  $e^-$  destinati a sopravvivere a  $T \ll m_e$ , per garantire oggi la neutralità di carica. Si tratta di un residuo che annovera dunque 1 elettrone extra per ogni miliardo di coppie  $e^+$ ,  $e^-$ .

Una situazione analoga si ha per il numero barionico. Risalendo a prima della transizione quark-adroni, quando è comunque  $T \gg m_q$  (massa dei quark più leggeri), dovevano esserci un miliardo di coppie  $\bar{q}$ ,  $q$  per ogni  $q$  extra, necessario a rendere conto del numero barionico presente oggi nel cosmo.

La rottura della simmetria particella-antiparticella, di  $\sim 1 : 10^9$ , implicita nei dati odierni su CMB e numero di barioni, tipica dei modelli di *Big-bang* caldo, è uno dei dati di *fine tuning* che la cosmologia deve spiegare.

#### 4. Il «collo di bottiglia» del Deuterio

I neutrini si disaccoppiano quando l'energetica cosmica è tipica della fisica nucleare. Ma l'istante del disaccoppiamento, più che dalle leggi della microfisica, dipende proprio dal rapporto  $10^9 : 1$ , tra particelle del brodo termico e barioni. Questo stesso

rapporto sarà nuovamente critico nel determinare il momento in cui le reazioni tra barioni possono dar luogo alla sintesi di nuclidi.

Per ricostruire i passi verso la formazione di nuclei di  ${}^4\text{He}$ , si devono tenere presenti due fatti essenziali:

- 1) Le reazioni avvengono in un bagno termico, cioè in presenza di un preciso numero di  $\gamma$  con spettro di energia planckiano.
- 2) Si devono considerare solo reazioni a 2 corpi:  ${}^4\text{He}$  si compone di 2 protoni e 2 neutroni; ma una reazione di «assemblaggio», con 4 nucleoni che convergono in un nuovo nucleo, ha probabilità zero, nonostante le elevate densità.

Bisogna quindi procedere per passi e il primo passo è forzato; neutroni e protoni debbono formare nuclei di deuterio o *deutoni* ( ${}^2\text{H}$  o  $d$ ):



la reazione può avvenire anche a rovescio, se ci sono  $\gamma$  con energia  $> B_d = 2.2 \text{ MeV} \simeq 2.5 \cdot 10^{10} \text{ K}$  (energia di legame o difetto di massa del deuterio).

### Valutazione semi-quantitativa

Potrebbe sembrare allora che il deuterio si possa formare appena la temperatura del cosmo scende sotto  $2.5 \cdot 10^{10} \text{ K}$ , addirittura prima della fuoriuscita dagli equilibri tra  $p$  e  $n$ ; non è così.

Com'è ragionevole, le sezioni d'urto delle reazioni di *deutosintesi* diretta e inversa sono simili; il loro ritmo sarà quindi simile quando saranno eguali i prodotti tra i numeri delle particelle reagenti:

$$(VI.28) \quad n_p \times n_n \simeq n_d \times n_\gamma. \quad (n_d: \text{densità di deutoni})$$

È qui che entra in gioco ancora il numero dei  $\gamma$  del brodo termico. Finché i  $\gamma$  sono (quasi) tutti attivi nelle reazioni (*deutosintesi*), sarà  $n_\gamma \sim 10^9 n_p, n_n$ , sicché l'eq. (VI.28) impone che ci sia 1  $d$  ogni  $\sim 10^9 p, n$ , praticamente niente deuterio. Attorno a  $2.5 \times 10^{10} \text{ K}$ , questa è ancora la situazione. Perché la sintesi di  $d$  possa iniziare, la temperatura deve essere scesa al punto di «disattivare» la quasi totalità dei  $\gamma$ ; può restare «attivo», al più, 1 fotone su  $10^9$ .

La frazione dei fotoni del CMB con energia superiore a  $B_d$ , per  $T \ll B_d$ , è

$$(VI.29) \quad f(> B_d) \simeq \frac{1}{2\zeta(3)} \left[ \left( \frac{B_d}{T} \right)^2 + 2 \frac{B_d}{T} - 1 \right] \exp\left(-\frac{B_d}{T}\right)$$

(vedi appendice A2) e, quindi, l'eq. (VI.28) va riscritta nella forma

$$(VI.30) \quad \frac{n_p \times n_n}{n_d \times n_\gamma} = \frac{1}{2\zeta(3)} \left[ \left( \frac{B_d}{T} \right)^2 + 2 \frac{B_d}{T} - 1 \right] \exp\left(-\frac{B_d}{T}\right)$$

Quando si ha  $n_\gamma f(> B_d) \sim n_b$ , ci possiamo aspettare densità  $n_p$ ,  $n_n$ ,  $n_d$  tra loro confrontabili, tutte  $\sim n_b$ . Se si richiede che  $f(> B_d) \sim n_b/n_\gamma$ , utilizzando l'espressione (VI.27) del rapporto barioni/fotoni, si calcola che la sintesi del deuterio inizia quando  $T \simeq 84$  keV. (Richiedendo semplicemente  $B_d/T = 9.25 \ln(10)$ , si avrebbe  $T \sim B_d/21 \simeq 103$  keV; il risultato è un po' minore a causa delle potenze di  $B_d/T$  a fattore.)

Si tratta di una valutazione intuitiva e efficace, ma poco più che qualitativa.

### Equazione di Saha

Una valutazione più corretta si ottiene richiedendo l'equilibrio termodinamico, da cui si ricava l'equazione di Saha

$$(VI.31) \quad \frac{n_p \times n_n}{n_d \times n_\gamma} = R(T) \equiv \frac{\pi^{1/2}}{2\zeta(3)} \left(\frac{m_b}{4T}\right)^{3/2} \exp\left(-\frac{B_d}{T}\right)$$

(vedi appendice A3). Il fattore chiave in eq. (VI.31), come in eq. (VI.30), è l'esponenziale, derivante dalla distribuzione planckiana dei fotoni. Esso è moltiplicato per un coefficiente, che in eq. (VI.31) è minore di quello in eq. (VI.30); ne consegue che la quantità prevista di deuterio, a ciascuna temperatura, è maggiore.

La richiesta che  $n_\gamma R(T)$  sia  $\sim n_b$ , ovvero che il secondo membro di eq. (VI.31) sia  $\sim \mathcal{O}(n_\gamma/n_b)$ , porta a una temperatura di apertura del collo di bottiglia del deuterio, così definita, che varia da 101 a 104 keV al variare di  $\Omega_b h^2$  da  $1.4 \times 10^{-2}$  a  $2.8 \times 10^{-2}$ ,

Da eq. (VI.31), tenendo conto che  $n_n = n_p = n_b - n_d$ , si può derivare che

$$(VI.32) \quad (n_d/n_b - 1)^2 = (n_d/n_b) R(T) (n_\gamma/n_b)$$

e la prescrizione  $R = n_b/n_\gamma$  significa  $n_d/n_b \simeq 0.38$ . Per avere  $n_d/n_b = 0.5$ , cioè tutto il numero barionico sotto forma di deuterio,  $R(T)$  deve diminuire ulteriormente fino al valore dato dalla temperatura  $T_d \simeq 70$  keV.

Sulla base di eq. (VI.24), alla temperatura di 70 keV, l'età del cosmo è  $\sim 270$  sec., se  $\sim N_\nu = 3$ . Il decadimento beta dei neutroni, conformemente a eq. (VI.18), dà allora

$$(VI.33) \quad X_n \simeq 0.165 \exp(-267/887) = 0.122$$

Questa deduzione di  $X_n$  deve essere presa con qualche riserva:

- 1) La sintesi del deuterio inizia a temperature  $> T_d$ , come abbiamo visto, quando la frazione di neutroni è maggiore.
- 2) Il deuterio così sintetizzato intraprende subito la catena di reazioni che portano a nuclidi più pesanti; ma, come vedremo nel prossimo paragrafo, esse hanno anche l'effetto di rigenerare  $p$  e  $n$  liberi; inoltre, il deuterio può anche essere ritrasformato in  $p + n$  dall'inverso della reazione di sintesi.
- 3) Quando il numero di protoni e neutroni diventa piccolo, si pone il problema di «farli incontrare» e, a 70 keV, una frazione significativa di neutroni è ancora disaccoppiata e destinata al decadimento.

L'assieme di questi effetti porta a indicare a  $\sim 50$  keV la temperatura a cui tutto il deuterio che si poteva formare si è formato, ma non altera di molto il valore (VI.33).

Il calcolo rende inoltre evidente un fatto essenziale: la quantità di deuterio formatasi dipende da  $N_\nu$ . Per esempio, se  $N_\nu = 2$ , la temperatura  $T_d$  si raggiunge al tempo  $t_d = 287$  sec. e ne deriva che, in questo maggior tempo, altri neutroni sarebbero decaduti, portando  $X_n$  a 1.19. Le ragioni 1), 2), 3), riportate qui sopra, che alterano di poco il valore di  $X_n$  per  $N_\nu = 3$ , hanno un impatto leggermente superiore per valori diversi di  $N_\nu$ , per cui conviene attenersi all'espressione

$$(VI.34) \quad X_n = 0.138 (\Omega_b h^2 / 0.02)^{0.04} (N_\nu / 3)^{0.2}$$

che dà valori con precisione superiore al 2%, per  $N_\nu$  da 1 a 4 e  $\Omega_b h^2$  da 0.14 a 0.30. Per  $\Omega_b h^2 = 0.02$ , si ricavano i valori qui riportati:

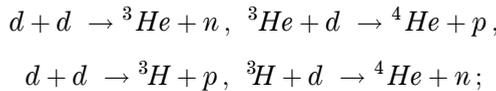
$N_\nu :$	1	2	3	4
$X_d :$	0.211	0.228	0.244	0.259

La quasi totalità del deuterio è destinata a trasformarsi in  ${}^4\text{He}$ . A fronte di un valore osservativo  $Y_{\text{He}} = 0.238 \pm 0.006$  si vede che  $N_\nu = 2$  è consistente coi dati entro  $\sim 1.6 \sigma$ ;  $N_\nu = 3$  entro  $\sim 1 \sigma$ ;  $N_\nu = 1$  si trova oltre  $\sim 4 \sigma$ . Il massimo valore entro  $3 \sigma$  è  $X_d \sim 0.255$ ; oltre ai 3 neutrini noti c'è quindi «spazio» per poco più di un altro *mezzo stato di spin*, ipoteticamente dovuto a qualche fondo di altra natura. Questi risultati si modificano leggermente al cambiare di  $\Omega_b h^2$ , ma anche dimezzando o raddoppiando questo parametro, gli spostamenti non raggiungono il 10%.

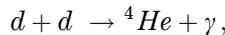
**I risultati della nucleosintesi primordiale permettono quindi di prevedere che non ci possano essere più di 3 neutrini e non meno di 2.** Questo limite fu ottenuto tramite le misure di abbondanza di elio molto prima degli esperimenti al LEP di Ginevra.

## 5. L'elio e gli altri nuclidi

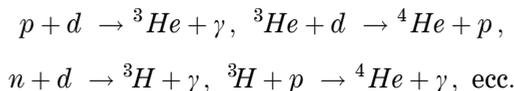
Le principali catene di reazioni che mutano  ${}^2\text{H}$  in  ${}^4\text{He}$  sono le seguenti:



molto meno frequente è la reazione



che va tenuta presente, a fini quantitativi, assieme agli altri *canali radiativi*



che diventano importanti quando il numero di deutoni è già fortemente ridotto, sicché un deutone incontra più facilmente un protone (o anche un neutrone) piuttosto che un altro deutone.

Mentre è relativamente semplice valutare l'abbondanza finale di  ${}^4\text{He}$ , che dipende essenzialmente dalla quantità di neutroni «salvati», meno semplice è comprendere le previsioni descritte in figura VI.3 per gli altri elementi leggeri.

Che, alla fine della sintesi dell'elio, resti comunque un residuo di deuterio, è cosa ragionevole; la sua densità sarà quella per cui il tempo di collisione tra deutoni diviene più lungo dell'età del cosmo ( $\tau_{dd} > t$ ). Una parte consistente di questo residuo finirà comunque nelle stelle dove viene distrutto. È questo un esempio dei processi di

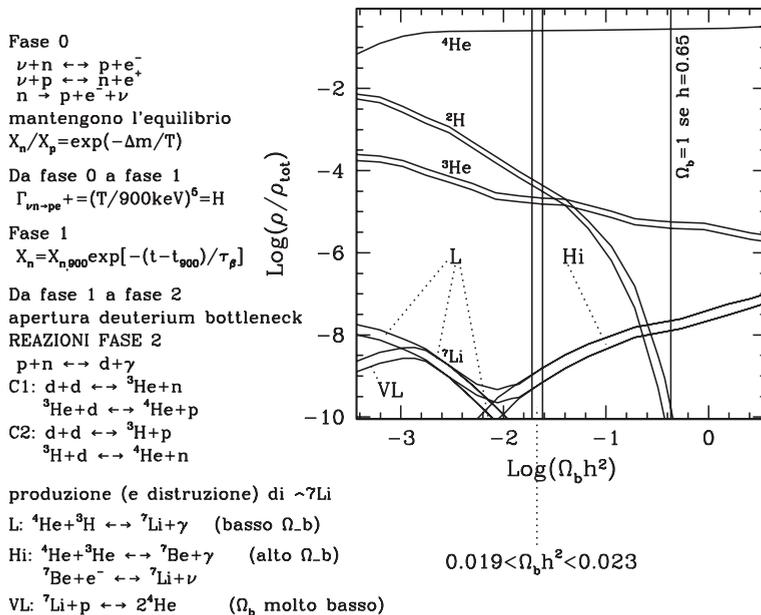


FIGURA VI.3

Reazioni e previsioni della BBNS. Nella Fase 0, le reazioni indicate in figura mantengono l'equilibrio termodinamico tra abbondanze di neutroni e protoni ( $X_{n,p}$ ). Il passaggio alla Fase 1 ha luogo a una temperatura di circa 900 keV, quando si congelano le reazioni provocate dai  $\nu$ ; prosegue il decadimento  $\beta$  che fa calare la frazione residua di neutroni. Il passaggio alla Fase 2, con l'apertura del «deuterium bottleneck», si ha a una temperatura  $\sim 70$  keV, quando la reazione di sintesi del deuterio diviene più rapida della sua fotodisintegrazione. Nella Fase 2 avvengono varie catene di reazioni. Le catene C1 e C2 sono le principali reazioni di sintesi dell'elio. Le reazioni L e Hi portano invece alla produzione di piccolissime quantità di  ${}^7\text{Li}$ . Per  $\Omega_b$  molto bassi, il  ${}^7\text{Li}$  può essere tuttavia distrutto dalla reazione VL. L'intervallo di  $\Omega_b h^2$  mostrato in figura permette l'accordo con le abbondanze osservate di tutti i nuclidi indicati. Per avere  $\Omega_b \simeq 1$  (con  $h = 0.65$ ) l'abbondanza di  ${}^2\text{H}$  dovrebbe essere  $\sim 10^{-10}$ , invece del valore osservato  $\simeq 10^{-4}$ . Questa discrepanza di 6 ordini di grandezza garantisce definitivamente che la componente barionica del cosmo ha una densità  $\ll \rho_{cr}$ , sicché diviene inescapabile la presenza di componenti oscure di natura non-barionica.

*astrazione*, che in molti casi comporta una significativa riduzione delle abbondanze primeve. Per confrontare le abbondanze primeve di nuclidi con quelle osservate oggi, si debbono valutare gli effetti dei possibili processi di astrazione.

Qualitativamente, tuttavia, si può capire perché un valore maggiore di  $\Omega_b h^2$  implica un minor residuo di  $d$ . Alzare  $\Omega_b$  equivale a abbassare la scala delle temperature. Se il deuterio è mediamente sintetizzato a  $T$  inferiore e quindi a tempi  $t$  maggiori, la condizione  $\tau_{dd} \simeq t$  si verificherà a tempi più elevati. Un tempo di collisione  $d-d$  maggiore corrisponde allora a una densità  $n_d$  minore. In sostanza, c'è più tempo per «consumare» il deuterio.

Formatosi  ${}^4\text{He}$  e in presenza di  ${}^3\text{H}$  e  ${}^3\text{He}$  in quantità non del tutto trascurabili, si rendono possibili varie catene di reazioni che portano a nuclidi più massicci (vedi Fig. VI.3). Tuttavia, le quantità di elementi massicci prodotte sono modeste.

Importante è la sintesi di  ${}^7\text{Li}$ , derivante da catene diverse di reazioni, a seconda del valore di  $\Omega_b h^2$ . Per  $\Omega_b h^2 \gtrsim 10^{-2}$ , la catena di reazioni

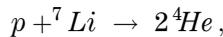


dà il contributo dominante. Per  $\Omega_b h^2$  minore, diviene dominante invece la reazione diretta



La prevalenza di  $(Hi)$  o  $(L)$  è in parte dovuta alle abbondanze di  ${}^3\text{He}$  e  ${}^3\text{H}$ , che ci sono durante la BBN e che sono diverse da quelle oggi osservate e riportate in figura VI.3.

Infine, quando  $\Omega_b h^2$  fosse molto basso, l'azione disgregante dei protoni, secondo la reazione



riduce l'ammontare finale di  ${}^7\text{Li}$ .

Si noti che, tra il *range* di  $\Omega_b h^2$  dove  $(Hi)$  è importante e quello dove  $(L)$  prevale, c'è un significativo *gap* e un minimo di  ${}^7\text{Li}$ , per  $\Omega_b h^2 \simeq 9 \cdot 10^{-3}$ .

L'abbondanza osservativa di  ${}^7\text{Li}$  non è facile da determinare. Gli specialisti dibattono sull'utilizzo di stelle di diverse popolazioni e sulla rilevanza osservativa di stelle in veloce rotazione. Incontestabilmente, tuttavia, il valore osservativo è prossimo al valore di  ${}^7\text{Li}$  in corrispondenza del *gap* ed è per questo che l'abbondanza osservativa di  ${}^7\text{Li}$  è un parametro difficile da riprodurre, negli scenari alternativi della BBN.

Notiamo anche che le condizioni in cui si verifica la BBN sono assai diverse da quelle che si realizzano nelle stelle dove sono sintetizzati nuclidi anche molto pesanti; mentre la BBN si arresta a  $A \sim 7$ .

Nel *core* delle stelle della sequenza principale l'idrogeno è gradualmente trasformato in elio. Per esempio, nel *core* del Sole, la frazione di  ${}^1\text{H}$  si è ormai ridotta da un valore  $\sim 75\%$ , immediatamente successivo alla formazione della stella, poco meno

di 5 miliardi di anni orsono, a un valore attuale  $\sim 40\%$ . Quando una stella finisce di sintetizzare  ${}^4\text{He}$ , cessa anche l'*input* termico relativo. Non più sostenuto dal calore di reazione, il plasma del core si contrae, *la densità cresce*, finché si verificano le condizioni per il bruciamento di  ${}^4\text{He}$  in  ${}^{12}\text{C}$  e, successivamente, altre catene di reazioni che sintetizzeranno nuclidi ancora più pesanti. L'evoluzione stellare, dopo la sintesi dell'elio, viaggia quindi verso densità sempre più elevate, fino a raggiungere talora lo stadio di pre-supernova e supernova, durante la cui esplosione si verificano ulteriori reazioni essenziali alla compressione delle abbondanze di metalli negli spazi galattici.

Nel cosmo primordiale, invece, indipendentemente dal ritmo delle reazioni nucleari, *la densità cala*. Formatosi l'elio, le condizioni cosmiche si allontanano da quelle che permetterebbero il suo bruciamento; i nuclidi che si formano dopo l'elio sono quindi scarsi e comunque leggeri.

La parte essenziale della rete di reazioni che porta a nuclidi più complessi è anche riportata in figura VI.3. Tuttavia, il peso relativo delle varie reazioni può essere seguito in dettaglio solo con metodi numerici. Si tratta veramente di *cosmologia di precisione* che porta a un'eccellente concordanza con i dati osservativi per  $\Omega_b h^2 = 0.021 \pm 0.002$  (2 deviazioni standard). **Questo accordo riguarda nuclidi presenti in frazioni che vanno da valori confrontabili con 1 a valori attorno a  $1 : 10^{10}$**  ed è uno dei risultati più significativi della cosmologia.

Da sottolineare, per la loro rilevanza, le conclusioni su  $\Omega_b$ : i barioni sono pochi e non adeguati a rendere conto di tutta la materia presente nel cosmo: quindi è  $\Omega_m \gg \Omega_b$ . La BBN è uno degli elementi inescapabili, che portano a diagnosticare l'esistenza della **materia oscura non-barionica**.

## 6. Modifiche alla nucleosintesi primordiale

Fisici e astrofisici hanno sempre nutrito perplessità su questa conclusione e hanno cercato in vari modi di rendere i dati compatibili con  $\Omega_b \sim 1$  o almeno con  $\Omega_b \simeq \Omega_m$ . Recentemente, tuttavia, molte evidenze osservative legate principalmente all'analisi delle fluttuazioni del CMB, indipendenti dall'abbondanza di nuclidi leggeri, confermano il valore previsto per  $\Omega_b h^2$  dalla BBN.

Una prima possibilità di modifica della BBN, cui abbiamo già accennato, è che il cosmo venga riscaldato successivamente alla nucleosintesi, dal decadimento di particelle massicce. I risultati di questo scenario vanno nella direzione opposta a quella auspicata. Incidentalmente, la concordanza tra valori di  $\Omega_b$  ricavati da diverse osservazioni pone limiti severi a tale riscaldamento.

Una seconda possibilità è che un simile decadimento sia avvenuto *durante* la nucleosintesi stessa, contribuendo a spaccare il deuterio e gli altri nuclidi appena formati, così da rallentare tutto il meccanismo della sintesi degli elementi leggeri. Si forma allora meno deuterio, nonostante più alti valori di  $\Omega_b$ , ma si perde la concordanza generale, da  ${}^2\text{H}$  a  ${}^7\text{Li}$ , con le abbondanze osservate.

Una terza ipotesi merita invece una considerazione particolare e sarà dibattuta un

po' più a lungo in un prossimo capitolo. Questo approccio prende il nome di *nucleosintesi disomogenea*. In questo scenario, all'inizio della sintesi del deuterio, *la distribuzione dei protoni* presentava delle *disomogeneità* pronunciate, mentre la distribuzione dei neutroni era omogenea. Ampiezza, distribuzione e forma delle disomogeneità protoniche sono collegabili alla dinamica della transizione quark-adroni.

Questo quadro consente di rilassare i limiti su  $\Omega_b$  di un fattore  $\sim 2$ , ma a costo di ipotesi sulla transizione quark-adroni che sembrano non confermate dai risultati numerici più recenti.

# Appendice A1

## Evoluzione cosmologica delle distribuzioni nello spazio delle fasi

Data una distribuzione  $v(x^i, p^i, t)$  ( $i = 1, 2, 3$ ) nello spazio delle fasi, sarà

$$(A1.1) \quad \frac{dv}{dt} = \frac{\partial v}{\partial x^i} \dot{x}_i + \frac{\partial v}{\partial p^i} \dot{p}_i + \frac{\partial v}{\partial t}.$$

Consideriamo ora il caso in cui  $v$  sia spazialmente omogenea e dipenda esplicitamente dal tempo solo tramite un parametro  $T$ . L'eq. (A1.1) diviene allora

$$(A1.2) \quad \frac{dv}{dt} = \frac{\partial v}{\partial p^i} \dot{p}_i + \frac{\partial v}{\partial T} \dot{T}.$$

Passiamo infine a un caso ancora più particolare, in cui  $v$  sia isotropa, dipenda quindi solo dal modulo  $p$  del vettore di componenti  $p^i$  e, inoltre, la dipendenza da  $p$  e  $T$  sia esprimibile come dipendenza dal rapporto  $u = p/T$ . Sarà allora

$$(A1.3) \quad \frac{dv}{dt} = \frac{\partial v}{\partial p} \dot{p} + \frac{\partial v}{\partial T} \dot{T} = \frac{\partial v}{\partial u} \left( \frac{\partial u}{\partial p} \dot{p} + \frac{\partial u}{\partial T} \dot{T} \right) = \frac{\partial v}{\partial u} p \left( \frac{\dot{p}}{p^2} - \frac{\dot{T}}{T^2} \right)$$

per cui, se è  $p \propto T$ , risulta  $dv/dt = 0$  e la distribuzione è invariante. Ciò avviene se  $T$  e  $p$  dipendono solo dall'andamento del fattore di scala  $a$ , essendo

$$(A1.4) \quad p a = \text{const.}, \quad T a = \text{const.},$$

come nel caso di particelle libere con o senza massa, il cui impulso (non l'energia, nel caso  $m \neq 0$ ) varia comunque, a causa del *redshift*, secondo le eq. (A1.4).

In conclusione, se una distribuzione nello spazio delle fasi è funzione del solo rapporto  $p/T$ , essa è invariante nell'espansione cosmologica. In questa categoria rientrano le distribuzioni di Fermi e Bose, per particelle con  $m = 0$ :

$$v = [\mathcal{N}/(2\pi)^3] (e^{p/T} \pm 1)^{-1}$$

Un caso significativo è quello di un neutrino, con massa  $m_\nu \ll T_{v,dg}$  che non subisce quindi più azioni dinamiche (a parte la gravità) a temperature  $< T_{v,dg}$ .

Alla temperatura di disaccoppiamento,  $T_{v,dg}$ , la sua distribuzione è

$$v = \frac{2}{(2\pi)^3} \frac{1}{e^{p/T} + 1} \quad \text{e, quindi,} \quad v \simeq \frac{2}{(2\pi)^3} \frac{1}{e^{p/T} + 1},$$

tranne una correzione del tutto trascurabile. La forma data dalla seconda espressione si mantiene quindi anche quando  $T$  divenisse  $< m_\nu$ .

## Appendice A2

### Integrali della distribuzione di Bose

Per determinare la frazione di  $\gamma$  con energia maggiore di  $B$ , si deve calcolare il rapporto

$$f(> B) = \frac{\frac{4\pi}{(2\pi)^3} \int_B^\infty dq q^2 \frac{1}{e^{q/T} - 1}}{\frac{4\pi}{(2\pi)^3} \int_0^\infty dq q^2 \frac{1}{e^{q/T} - 1}} = \frac{\frac{1}{2\pi^2} T^3 \int_{B/T}^\infty dx x^2 \frac{1}{e^x - 1}}{\frac{\zeta(3)}{\pi^2} T^3} \simeq \frac{1}{2\zeta(3)} \int_{B/T}^\infty dx x^2 e^{-x};$$

l'ultimo passaggio è lecito solo per  $B \gg T$ . L'integrale si esegue per parti e dà:

$$f(> B) \simeq \frac{1}{2\zeta(3)} \left[ \left(\frac{B}{T}\right)^2 + 2\frac{B}{T} - 1 \right] \exp\left(-\frac{B}{T}\right).$$

## Appendice A3

### Equazione di Saha

Alla temperatura  $T$ , la densità in numero (prendendo  $c = \hbar = k_B = 1$ ) di determinate particelle *non-relativistiche* è data dall'espressione di Boltzmann (A1.5)

$$(A3.1) \quad n = \mathcal{N}(mT/2\pi)^{3/2} \exp[-(m - \mu)/T].$$

$m$  e  $\mathcal{N}$  sono massa e numero di stati di spin delle particelle;  $\mu$  è il *potenziale chimico* associato ai loro numeri quantici. La conservazione del potenziale chimico garantisce la conservazione dei numeri quantici a esso collegati.

Dall'eq. (A3.1) si ricava l'espressione della densità di nuclidi di massa  $m$ , con  $\mathcal{N}$  stati di spin, peso atomico  $A$ , numero atomico  $Z$ , in equilibrio termodinamico con protoni ( $p$ ) e neutroni ( $n$ ) di densità  $n_{n,p}$ :

$$(A3.2) \quad n(Z, A) = \frac{\mathcal{N}}{2^A} \left( \frac{m}{m_p^Z m_n^{A-Z}} \right)^{\frac{3}{2}} \left( \frac{2\pi}{T} \right)^{\frac{3}{2}(A-1)} n_p^Z n_n^{A-Z} \exp\left(\frac{B}{T}\right);$$

$$(A3.3) \quad B = -m + Zm_p + (A - Z)m_n$$

è il difetto di massa del nuclide considerato.

Infatti, anzitutto, per la conservazione dei numeri quantici associati a  $p$  e  $n$  (numero barionico e carica elettrica), il potenziale chimico del nuclide sarà

$$(A3.4) \quad \mu = Z\mu_p + (A - Z)\mu_n,$$

$\mu_p$  e  $\mu_n$  essendo i potenziali chimici di  $p$  e  $n$ . Nel rapporto

$$(A3.5) \quad \frac{n(A, Z)}{n_p^Z n_n^{A-Z}} = \frac{\mathcal{N}}{2^A} \frac{(2\pi/mT)^{3/2}}{(2\pi/m_p T)^{(3/2)Z} (2\pi/m_n T)^{(3/2)(A-Z)}} \times \\ \times \exp\left(-\frac{m - Zm_p + (A - Z)m_n}{T} + \frac{\mu - Z\mu_p - (A - Z)\mu_n}{T}\right),$$

riconosciamo poi, al primo rigo, le diverse potenze di  $m$ ,  $T$ ,  $m_p$  e  $m_n$  del fattore  $(m/m_p^Z m_n^{A-Z})^{\frac{3}{2}} (2\pi/T)^{(3/2)(A-1)}$  di eq. (A3.2). Nel secondo rigo, grazie a eq. (A3.3), il primo termine all'esponente diviene  $B/T$ , mentre, grazie a eq. (A3.4), il secondo termine si annulla. Così eq. (A3.5) diviene eq. (A3.2).

Dall'eq. (A3.2) si ricava, come caso particolare, l'abbondanza di deuterio ( $d$ ) in equilibrio termodinamico alla temperatura  $T$ . Un deutone ha  $A = 2$ ,  $Z = 1$  e spin  $3/2$ , per cui è  $\mathcal{N}_d = 4$ . La sua densità sarà allora data da:

$$(A3.6) \quad \frac{n_d}{n_p n_n} = \left( \frac{m_d}{m_p m_n} \right)^{\frac{3}{2}} \left( \frac{2\pi}{T} \right)^{\frac{3}{2}} \exp\left(\frac{B_d}{T}\right);$$

$B_d = 2.2$  MeV è il difetto di massa di  $d$ . Nei rapporti tra masse si può prendere  $m_d \simeq 2m_p \simeq 2m_n$  o usare la «massa barionica media»

$$(A3.7) \quad m_b \simeq 2m_p m_n / m_d$$

ottenendo

$$(VI.31) \quad \frac{n_p n_n}{n_d n_\gamma} = \frac{\pi^2}{2\zeta(3)} \left( \frac{m_b}{4\pi T} \right)^{\frac{3}{2}} \exp\left(-\frac{B_d}{T}\right)$$

dove si è tenuto conto che  $n_\gamma = (2\zeta(3)/\pi^2)T^3$ .

## CAPITOLO VII

# Evoluzione lineare delle disomogeneità

## Approccio analitico nel caso a una sola componente

### 1. Lo scenario dell'instabilità gravitazionale

La densità della materia, nel cosmo, conosce valori estremi. Ci sono siti in cui è straordinariamente elevata: all'interno di una stella a neutroni si raggiunge la densità nucleare,  $\sim 10^{15}\text{g/cm}^3$ ; in ogni  $\text{cm}^3$  c'è un milione di miliardi di grammi. Una goccia di stella a neutroni, portata sulla Terra, peserebbe svariati milioni di tonnellate. Ma, nel cosmo, troviamo anche gli spazi intergalattici; vuoti, freddi, oscuri, dove la densità residua si può esprimere solo con un esponente:  $10^{-30}\text{g/cm}^3$ ; se fosse dovuta esclusivamente a atomi di idrogeno, troveremmo 1 atomo per metro cubo. Non meno di 45 ordini di grandezza, dunque, tra densità massima e minima.

La sfida lanciata dai cosmologi è mostrare che questi 45 ordini di grandezza non contrastano con l'idea di un cosmo primordiale omogeneo e isotropo, un'ipotesi da cui si debbono poter desumere le proprietà statistiche dei sistemi cosmici presenti sulle varie scale. Questo programma è tuttora in elaborazione, ma ci sono tutti gli elementi per affermare che, a grandi linee e con l'aggiunta di uno specifico ingrediente, lo schema funziona e le proprietà statistiche previste sono confermate dai dati.

L'idea implicita nelle prime speculazioni era che la fuoriuscita dall'omogeneità primeva fosse figlia dei lunghi tempi e di mere fluttuazioni statistiche. Oggi sappiamo, anzitutto, che il tempo disponibile è veramente enorme, tra 13 e 14 miliardi di anni. Ma anche questo abisso temporale non basta: le fluttuazioni statistiche son troppo poco ed è qui che la semplicità dell'assunto un po' s'incrina. L'ingrediente da aggiungere è dunque una violazione primeva dell'omogeneità. Ma, come vedremo, anche questa «violazione» viene inserita in modo isotropo e omogeneo, sicché l'informazione contenuta nelle condizioni iniziali è veramente quasi nulla.

Tutta l'articolazione del mondo osservato è dunque conseguenza della legge fisica, non di *input* specifici. Cercheremo ora di spiegare come si giunge a questa conclusione.

### Le fluttuazioni statistiche

Vediamo anzitutto cosa possono essere le fluttuazioni statistiche. In un ambiente omogeneamente riempito di elementi atomici – molecole, atomi, nuclei, componenti di plasma – ritagliamo un volume  $R^3$ . Esso conterrà

$$(VII.1) \quad N \simeq n R^3$$

elementi;  $n$  è la *densità in numero* degli elementi in questione. Se prendiamo volumi eguali in siti diversi, troveremo valori di  $N$  lievemente diversi. La discrepanza  $\Delta N$  sarà tanto minore quanto maggiore è il volume (e quindi  $N$ ). Una stima delle discrepanze è

$$(VII.2) \quad \frac{\Delta N}{N} \simeq \frac{1}{N^{1/3}}.$$

Per verificarlo, partiamo dal valore della distanza media tra gli elementi,

$$(VII.3) \quad a \simeq R/N^{1/3},$$

ottenuta come lato di un cubetto di volume  $a^3 \simeq R^3/N$ , come dire il volume che ciascun elemento ha «a disposizione». Il risultato è sostanzialmente vero anche se gli elementi sono distribuiti casualmente, a condizione che la superficie che delimita  $R^3$  non sia studiata *ad hoc*, con buchi e convessità. Se si bara, si possono persino trovare volumi  $R^3$  che non contengono alcun elemento.

Se invece tracciamo a caso la superficie di  $R^3$ , gli elementi più prossimi alla superficie ne distano, mediamente,  $a/2$ . Immaginiamo allora uno strato superficiale di profondità  $a$ , di volume

$$(VII.4) \quad R^2 \times a = R^3/N^{1/3}$$

come si ottiene usando eq. (VII.3). Mediamente, esso contiene  $n \times R^3/N^{1/3} = N/N^{1/3} = N^{2/3}$  elementi.

Questo è il numero medio, che si trova tra due casi estremi:

- 1) Se tutti gli elementi distano dalla superficie del volume  $R^3$  un infinitesimo più di  $a$ , lo strato superficiale è vuoto.
- 2) Se tutti gli elementi sono a distanza infinitesima dalla superficie, la densità, nello strato superficiale, potrebbe arrivare al doppio della media: a ogni elemento aderente la superficie esterna dello strato ne corrisponde un secondo, appena entro la sua superficie interna.

Il numero di elementi del volume  $R^3$  è quindi compreso tra  $N - N^{2/3}$  (strato superficiale vuoto) e  $N + N^{2/3}$  (strato superficiale con densità doppia), sicché è  $\Delta N \simeq N^{2/3}$ . Sarà allora chiaramente  $\Delta N/N \simeq N^{-1/3}$  (eq. VII.2).

Applichiamo questo risultato a una grossa galassia, con massa  $\sim 10^{12} M_{\odot}$ , mille miliardi di masse solari. Ricordiamo che  $1 M_{\odot} \simeq 2 \times 10^{33} \text{g}$  e che ogni grammo contiene  $\sim 10^{24}$  particelle atomiche. Una grossa galassia si compone quindi di  $\sim 10^{69}$  elementi.

Se pensiamo che essa si sia distaccata da un plasma inizialmente omogeneo, possiamo contare su una fluttuazione statistica di  $1:10^{23}$ . In 14 miliardi di anni, un particolare volume dove, casualmente, c'era un simile eccesso di densità dovrebbe aver visto la crescita della fluttuazione di densità  $\delta = \Delta N/N$  da  $10^{-23}$  a  $10^7$ ; 30 ordini di grandezza.

### ***Bastano le fluttuazioni statistiche?***

A fronte di questa stima, sta lo studio dell'evoluzione reale delle fluttuazioni che costituirà una gran parte del contenuto di questo capitolo e del prossimo. Abbiamo però già anticipato la conclusione: non basta.

Già in questo capitolo vedremo che la crescita di una fluttuazione, fin tanto che essa evolve in *regime lineare*, avviene dapprima con legge  $\delta \propto a^2$ , poi  $\delta \propto a$ , dove  $a$  è il fattore di scala. Il passaggio da un ritmo di crescita all'altro ha luogo quando  $a/a_0 \sim 10^{-4}$  ( $a_0$  è il parametro di scala oggi). Tralasciamo pure il fatto che questi sono i massimi ritmi di crescita possibili, che ci sono altri *modi* evolutivi, che una fluttuazione può trascorrere un lungo periodo oscillando come un'onda sonica, o in una fase di stagnazione. Non si tratta di dettagli, queste fasi sono effettivamente presenti nella storia evolutiva di una fluttuazione che dà origine a una galassia. Aggiungiamo poi un ulteriore ingrediente ottimistico: che, raggiunto un valore  $\delta \simeq 1$ , e uscita la fluttuazione dal regime lineare, il passaggio a  $\delta \simeq 10^7$  sia rapidissimo. Resta da chiedersi se il 23 ordini di grandezza precedenti possono essere comunque coperti.

La crescita può iniziare non appena l'orizzonte causale raggiunge la scala galattica, quando è  $a/a_0 \sim 10^{-6}$ . Pochi giorni sono trascorsi dal *Big-bang* e restano disponibili praticamente tutti i 14 miliardi di anni. Ma, in questo tempo, la fluttuazione cresce al più di 8 ordini di grandezza: 4 fino a  $a/a_0 \simeq 10^4$  e altri 4 da allora a quando  $a = a_0$ . Partendo da  $10^{-23}$  arriva  $10^{-15}$ ; 1 è lontano.

La conclusione che trarremo alla fine di questi due capitoli è che occorre un meccanismo *fuori dall'orizzonte*, che pre-confezioni un'ampiezza iniziale  $\delta_i \sim 10^{-5}$  ( $10^{-8}$ , visti i conti dettagliati, non basterebbe). In attesa di dibattere questa richiesta, che sembra violare il principio di causalità e rende cogente il meccanismo inflazionario, precisiamo ulteriormente la conclusione cui arriveremo: l'omogeneità primeva deve essere accompagnata da perturbazioni, *esse stesse omogeneamente distribuite*, la cui ampiezza all'orizzonte è  $\sim 1 : 10^5$ .

## **2. Il problema di Jeans**

Il problema di Jeans consiste semplicemente nel determinare quando le fluttuazioni crescono e con che legge la crescita avviene. Il suo studio inizia sul finire dell'Ottocento, quando nulla si sapeva del cosmo in espansione. Jeans parte dall'idea di un *background* stazionario, un assunto internamente non consistente: le sue stesse equazioni permettono di verificare che un *background* stazionario è impossibile. Trascurando questa di-

scrasia di base, Jeans trovò che le piccole disomogeneità di un *background* stazionario avrebbero crescita esponenziale.

Fosse realmente così, le disomogeneità statistiche potrebbero farcela. L'ostacolo sta dunque nell'espansione del cosmo. La trattazione di Jeans, oltre che semplice, è comunque interessante: essa illustra l'esistenza dei due regimi (sonico e di instabilità gravitazionale) e fissa correttamente la scala che li separa.

Successivamente il problema venne riconsiderato da Lifshitz, nell'ambito di un cosmo in espansione; la trattazione diviene consistente e... più complicata.

Una trattazione relativistica si rende poi indispensabile quando si verifichi almeno una delle seguenti circostanze: (i) Si trattino scale prossime all'orizzonte. (ii) Si tratti con *sostanze* composte di particelle relativistiche. Tale trattazione richiede un uso tecnicamente evoluto della Relatività Generale.

Dedicheremo poi un cenno al caso importante in cui una sostanza cosmica non possa essere vista come «fluido».

### 3. Instabilità di Jeans e crescita delle fluttuazioni

L'esperienza ordinaria e di laboratorio ci dice che le fluttuazioni di densità in un fluido si propagano come onde *soniche*. Ciò avviene perché la loro autogravità è trascurabile. Su scale astrofisiche, invece, si può arrivare a un *regime di instabilità gravitazionale*, in cui le fluttuazioni di densità tendono a crescere sotto l'azione della propria gravità.

L'instabilità gravitazionale riguarda scale molto grandi. Su una fluttuazione di densità agiscono infatti: (i) robuste forze di pressione, che contrastano lo squilibrio dovuto alla variazione di densità; (ii) tenui forze gravitazionali, che attraggono ulteriore massa laddove la densità è aumentata.

Le forze di pressione sono però *forze di superficie*, quelle gravitazionali sono *forze di volume*; le prime crescono con  $\lambda^2$ , le seconde con  $\lambda^3$  ( $\lambda$  indica la scala di della fluttuazione). Ci sarà quindi una scala  $\lambda_J$ , sopra la quale la gravità vince la pressione.  $\lambda_J$  è comunque molto grande e il regime di instabilità gravitazionale ha senso solo su scala astronomica.

L'espressione più generale della scala di Jeans, assumendo che le fluttuazioni riguardino una sola sostanza, è la seguente (vedi appendice A1):

$$(VII.5) \quad \lambda_J = c_s \sqrt{\frac{\pi}{G\rho}} \times \sqrt{\frac{1}{(1+w)(1+3c_s^2/c^2)}}.$$

Per materiali non relativistici il secondo fattore è  $\simeq 1$ . Nel primo fattore entrano  $G$ : costante gravitazionale;  $\rho$ : densità di massa;  $c_s$ : velocità del suono.

Nel secondo entra anche il rapporto  $w = p/\rho c^2$ : *parametro di stato* del fluido.

L'espressione inizialmente trovata da Jeans coincide col primo fattore. Essa venne confermata da Lifshitz, per una sostanza non relativistica. La soluzione di

Lifschitz si applica quindi direttamente a una cosmologia SCDM e anche a modelli con energia oscura (DE), ma solo a alti  $z$  (finché la densità della DE può essere trascurata).

### ***Decomposizione di un segnale in componenti armoniche***

Jeans e Lifshitz utilizzano nella definizione di  $\lambda_J$  lo sviluppo di una fluttuazione in componenti armoniche. Tale sviluppo è alla base di tutta la trattazione lineare e ricorderemo quindi, brevemente, in cosa esso consista.

Una fluttuazione di densità

$$(VII.6) \quad \varepsilon(\mathbf{x}) = \frac{\rho(\mathbf{x}) - \bar{\rho}}{\bar{\rho}}$$

( $\bar{\rho}$  è la densità media) può dipendere dalla coordinata spaziale  $\mathbf{x}$  come segue:

$$(VII.7) \quad \varepsilon(\mathbf{x}) = \delta e^{\frac{2\pi i}{\lambda} \mathbf{n} \cdot \mathbf{x}}.$$

In questo caso essa costituisce un'onda *piana*. Su tutti i punti di ciascun piano normale al versore  $\mathbf{n}$ , infatti,  $\varepsilon$  ha lo stesso valore. Si dice poi che essa è *monocromatica*, per significare che contiene la sola lunghezza d'onda  $\lambda$ : crescendo  $\mathbf{n} \cdot \mathbf{x}$  di  $\lambda$ ,  $\varepsilon$  riassume gli stessi valori.

Si possono ovviamente considerare fluttuazioni di densità che sono sovrapposizioni di più segnali della forma (VII.7), di ampiezze diverse ( $\delta_j$ ), lunghezze d'onda diverse ( $\lambda_j$ ) e diretti in direzioni diverse ( $\mathbf{n}_j$ ):

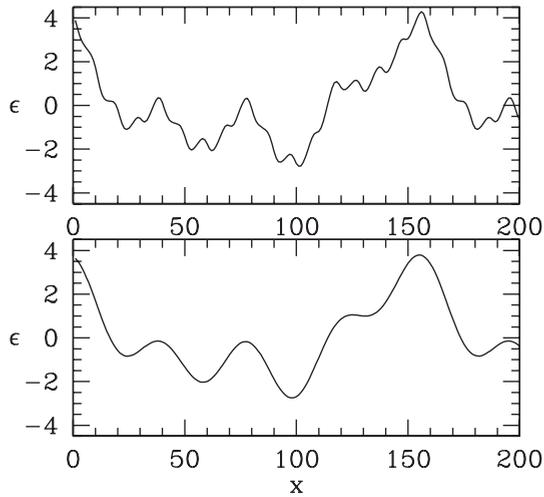
$$(VII.8) \quad \varepsilon(\mathbf{x}) = \sum_j \delta_j e^{\frac{2\pi i}{\lambda_j} \mathbf{n}_j \cdot \mathbf{x}}.$$

La somma può estendersi a un'infinità discreta di segnali, diventando una serie, o anche a un'infinità continua, nel qual caso sarà espressa da un opportuno integrale.

Se l'ampiezza  $\delta$  in eq. (VII.7) o le ampiezze  $\delta_i$  in eq. (VII.8) variano nel tempo, ovviamente  $\varepsilon$  dipenderà anch'essa dal tempo.

Del tutto in generale si dimostra ora che:

- 1) *Qualunque segnale periodico  $\varepsilon(\mathbf{x})$  può essere posto nella forma di una somma (VII.8); ma qualunque segnale, anche non periodico, può essere posto della stessa forma, se si estende la somma dal discreto al continuo.*
- 2) *Fin tanto che il modulo delle ampiezze  $\delta_i$  resta  $\ll 1$  (regime lineare), l'evoluzione complessiva di un segnale si ricava sovrapponendo le evoluzioni delle ampiezze  $\delta_i$  e questa è la più generale forma di evoluzione temporale possibile per il segnale (VII.8).*
- 3) *Ciascuna delle ampiezze  $\delta_i$ , dette componenti armoniche, evolve separatamente dalle altre.*

**FIGURA VII.1**

Esempio unidimensionale dell'articolazione di un segnale (periodico) al crescere del numero delle componenti armoniche. Nella figura in basso ci sono 3 componenti armoniche. Esse salgono a 6 nella figura in alto. La strutturazione si estende a segnali anche non periodici, in cui si ha però un insieme continuo di componenti armoniche.

I punti 2) e 3) qui sopra costituiscono il *principio di sovrapposizione*, che discende dalla linearità delle equazioni dinamiche, e che permette di studiare un qualunque segnale preoccupandosi di vedere come evolve un segnale di tipo (VII.7), al variare di  $\lambda$ .

Dato un segnale generico  $\varepsilon(\mathbf{x})$  ci sono delle precise prescrizioni matematiche per trovarne le componenti armoniche. Al di là del loro dettaglio, in figura VII.1 si dà un esempio di come possa articolarsi un segnale al crescere del numero delle sue componenti armoniche. In un segnale periodico di lunghezza d'onda  $\lambda$ , la presenza di struttura su scale  $\sim D$  ( $\ll \lambda$ ) richiede inevitabilmente la presenza di componenti armoniche di lunghezza d'onda  $\lambda/n$ , con  $n \lesssim \lambda/D$ .

### ***Evoluzione temporale dei segnali***

L'evoluzione delle componenti armoniche dipende criticamente dal loro raffronto con la scala di Jeans  $\lambda_J$ . In appendice A1 si dimostra che componenti con  $\lambda \ll \lambda_J$  hanno l'andamento temporale

$$(VII.9) \quad \delta(\mathbf{k}_i, t) = \delta_0(\mathbf{k}_i) e^{3i\omega t}$$

con  $\omega = c_s k$ . Si faccia molta attenzione al fattore 3, all'esponente, caratteristico di un'espansione del *background* con legge  $a \propto t^{2/3}$ . In assenza di espansione il fattore 3

non ci sarebbe. Un fattore siffatto è essenziale a comprendere lo spettro delle anisotropie del CMB (vedi capitolo successivo).

Le componenti con  $\lambda \gg \lambda_J$  hanno invece tutte la stessa evoluzione che, nel modello SCDM, o in una cosmologia approssimabile da tale modello, avviene con la legge:

$$(VII.10) \quad \delta(t) = \delta_{+,i} (t/t_i)^{2/3} + \delta_{-,i} (t/t_i)^{-1}.$$

Nell'espressione (VII.10)  $t_i$  è un «tempo iniziale» a cui la fluttuazione ha ampiezza  $\delta_i = \delta_{+,i} + \delta_{-,i}$ . A partire da  $t_i$ , la fluttuazione si ripartisce tra *modo crescente*  $\delta_+ \propto t^{2/3}$  e *modo decrescente*  $\delta_- \propto t^{-1}$ . Non basta quindi conoscere l'ampiezza iniziale, per sapere come la fluttuazione evolverà nel tempo. Comunque, la crescita non può avvenire più rapidamente della potenza  $2/3$  del tempo. Nel modello SCDM risulta  $a \propto t^{2/3}$  e il modo crescente evolve quindi con legge:

$$\delta_+ \propto a \quad (\text{crescita durante espansione dominata dalla materia}).$$

Se la sostanza dominante è relativistica, l'evoluzione lineare è la seguente:

$$(VII.11) \quad \delta(t) = \delta_{+,i} (t/t_i) + \delta_{-,i} (t/t_i)^{-1}.$$

Questa legge si applica prima dell'equivalenza (redshift  $z_{eq} \sim 10^4$ ), quando è  $a^2 \propto t$ . Nel modo crescente, abbiamo allora:

$$\delta_+ \propto a^2 \quad (\text{crescita durante espansione dominata da radiazione}).$$

Il fatto che questo risultato si applichi a scale maggiori di  $\lambda_J$  ha implicazioni cosmologiche essenziali. Infatti, nel caso relativistico, risulta

$$(VII.12) \quad \lambda_J = \frac{c}{\sqrt{3}} \sqrt{\frac{\pi}{G\rho}} \times \sqrt{\frac{3}{8}} = \frac{\pi}{\sqrt{3}} 2ct,$$

come si vede usando l'eq. (VII.5) con  $w = c_s^2/c^2 = 1/3$ , ciò che è corretto nel caso in cui i fotoni siano la sostanza cosmica dominante, e tenendo conto che in quest'epoca, per l'equazione di Friedmann, sarà  $\sqrt{3/8} \sqrt{\pi/G\rho} = 2\pi t$ .

La scala di Jeans è dunque lievemente superiore all'orizzonte  $2ct$ ; la crescita con legge  $\propto a^2$  si applicherebbe quindi *prima dell'entrata nell'orizzonte* (in questo caso ci sono però notevoli ambiguità connesse al problema della definizione della *gauge gravitazionale*). Dopo l'entrata nell'orizzonte, invece, le fluttuazioni sarebbero comunque in regime sonico.

In realtà la situazione è ancora più complessa e, per trattarla correttamente, si deve tenere conto opportunamente della presenza di più componenti. Torneremo sulla questione, con opportuni dettagli, nel prossimo capitolo.

#### 4. Oltre il modello di fluido

Anche a questo scopo, comunque, va considerata la possibilità che la sostanza cosmica non sia un *fluido*. Una sostanza si comporta come un fluido se composta di particelle

che hanno cammini liberi medi molto piccoli, al di sotto delle scale di lunghezza che caratterizzano l'osservazione o l'esperimento. Non possiamo considerare fluidi, quindi, le sostanze in cui i cammini liberi medi delle particelle risultino formalmente superiori alla scala di Jeans.

In questo caso, comunque, ci sono differenze sostanziali solo *sotto* la scala di Jeans, dove esse sono inescappabili: la propagazione di onde soniche avviene solo nei fluidi; un'onda di compressione si regge sul trasporto di energia e impulso, da un punto all'altro dello spazio, da parte degli elementi microscopici del fluido. Se il trasporto avvenisse su distanze infinitesime, il fluido sarebbe «perfetto», a viscosità e conduzione termica nulle. Di fatto, il trasporto avviene sulle scale dei cammini liberi medi  $\lambda_c$  e le scale su cui la sostanza è un fluido sono quelle  $\gg \lambda_c$ . Su lunghezze d'onda  $\lambda > \lambda_c$ , ma prossime a  $\lambda_c$ , le fluttuazioni risentono fortemente di viscosità e conduzione termica, si smorzano rapidamente e l'energia sonica si trasforma in disordinata energia termica.

Su scale  $> \lambda_J$ , invece, non ci aspettiamo propagazione sonora, non c'è problema di trasferimento di impulso e energia. Tuttavia, nell'espressione di  $\lambda_J$  è presente la *velocità del suono*  $c_s$  e si tratta di discutere come «sostituirla».

Mostreremo ora che ciò si può fare abbastanza facilmente. La conclusione cui si perviene, per i mezzi non collisionali, sarà comunque la seguente: al di sopra di una scala  $\lambda_J$ , opportunamente definita ma assai prossima alla scala di Jeans dei paragrafi precedenti, siamo ancora in regime di instabilità gravitazionale; le leggi di crescita delle fluttuazioni sono quelle già ricavate. *Fluttuazioni con lunghezza inferiore a  $\lambda_J$ , invece, sono destinate a una progressiva dissipazione.* Questo secondo punto è caratteristico dei materiali non collisionali. Nei fluidi, in questo regime si propagano onde soniche.

Il problema dell'instabilità gravitazionale, nei mezzi non collisionali, si imposta tramite il *teorema del viriale* che prescrive che, in equilibrio, deve essere

$$(VII.13) \quad Vir = 2E_k + V = 0,$$

con  $E_k$ : energia cinetica,  $V$ : energia potenziale.

Consideriamo un materiale omogeneo e, al suo interno, un volume  $R^3$  contenente una massa  $M = \rho R^3$ ; sia  $\bar{v}$  la velocità quadratica media dei suoi componenti elementari, che supporremo per semplicità avere tutti eguale massa  $m$ . L'energia cinetica entro  $M$  è data da

$$2E_k = \sum_p mv_p^2 = M\bar{v}^2;$$

la somma si estende a tutti i componenti elementari e  $\bar{v}$  è la *velocità quadratica media*. L'energia potenziale gravitazionale invece risulta

$$V = -Gf_o \frac{M^2}{R} = -Gf_o M \rho R^2;$$

$f_o$  è un fattore che dipende dalla forma del volume e, salvo forme assai peculiari, è prossimo all'unità. Richiedendo allora che sia  $Vir/M = 0$  si ottiene:

$$(VII.14) \quad \bar{v}^2 - Gf_o\rho R^2 = 0.$$

Perciò, un materiale di densità  $\rho$  è in equilibrio se occupa un volume di dimensioni

$$(VII.15) \quad R_J = \frac{\bar{v}}{\sqrt{f_o\pi}} \sqrt{\frac{\pi}{G\rho}}.$$

Se le dimensioni sono invece  $R > R_J$ , il viriale diventa negativo: c'è troppo poca energia cinetica per contenere gli effetti dell'elevata energia potenziale. Il sistema è quindi gravitazionalmente instabile, le sue dimensioni tendono a ridursi e la sua densità aumenta. Tutto ciò è strettamente analogo con quanto accade a un fluido su scale superiori alla lunghezza di Jeans.

Invece, se la scala del volume occupato è  $R < R_J$  e il viriale è positivo. L'eccesso di energia cinetica significa velocità quadratiche medie superiori alla «velocità di fuga».

In un fluido, l'eccesso di energia cinetica si riconduce a moti «ordinati» e dove esse sono presenti abbiamo un *nodo*, in un'onda di compressione. Le particelle, urtando contro altre particelle, provocheranno presto un addensamento e quindi una situazione di *ventre*, dove  $\rho$  cresce bilanciando il termine cinetico. Poi la successione nodo-ventre continuerà, nel fenomeno di propagazione dell'onda.

In un materiale non collisionale, non si hanno urti e le velocità, superiori alla *velocità di fuga*, fanno sì che la fluttuazione si dissipi per fuoriuscita (*free-streaming*) delle particelle.

L'espressione (VII.15) di  $R_J$  coincide con l'espressione di  $\lambda_J$  se  $c_s^2 = \bar{v}^2/f_o\pi$ . Ebbene, nei fluidi, la velocità quadratica media delle particelle e la velocità del suono sono rispettivamente

$$\bar{v}^2 = k_B T / 3m, \quad c_s^2 = (2/l + 1)k_B T / m$$

( $l$  è il numero di gradi di libertà delle particelle). Per  $l = 3$  (particelle puntiformi) sarà  $R_J = \lambda_J$  se

$$f_o = 1/5\pi.$$

È un esercizio di geometria differenziale dimostrare che questo è il fattore di forma nel caso di una sfera omogenea.

# Appendice A1

## Trattazione analitica

### 1. Le equazioni di Jeans

Consideriamo un fluido, di cui  $\rho$  e  $p$  sono densità e pressione, mentre  $\mathbf{v}$  e  $\mathbf{g}$  sono campo di velocità e campo gravitazionale. Useremo unità naturali ( $c = 1$ ).

Valgono le seguenti equazioni:

$$(a) - \text{continuità: } \frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot (\rho \mathbf{v}) = 0$$

$$(b) - \text{Eulero: } \frac{\partial \mathbf{v}}{\partial t} = -(\mathbf{v} \cdot \vec{\nabla})\mathbf{v} - \vec{\nabla} \left( \frac{p}{\rho} \right) + \mathbf{g}$$

$$(c) - \text{Poisson: } \vec{\nabla} \cdot \mathbf{g} = -4\pi G \rho$$

L'EQUAZIONE DI EULERO è l'estensione dell'equazione di Bernoulli,

$$\frac{1}{2} \mathbf{v} \cdot \mathbf{v} + p/\rho - \int \mathbf{g} \cdot d\mathbf{r} = \text{const.},$$

valida in regime stazionario ( $\mathbf{v}$  non dipende da  $t$ ), al caso dinamico. Considerandone il gradiente, si ricava

$$(\mathbf{v} \cdot \vec{\nabla})\mathbf{v} + \vec{\nabla} p/\rho - \mathbf{g} = 0.$$

Se  $\mathbf{v}$  dipende da  $t$ , questa quantità non si annulla, ma vale  $-\partial \mathbf{v} / \partial t$ .

SOLUZIONE STATICA. Sostituiamo, nelle eq. (a), (b), (c), valori costanti  $\bar{\rho}$ ,  $\bar{p}$  e  $\mathbf{v} = 0$ . (a) è soddisfatta. (b) ci dà  $\mathbf{g} = 0$ . (c) allora implica  $\bar{\rho} = 0$ . Una soluzione statica è quindi *vuota*. Ciononostante Jeans assume

$$\rho = \bar{\rho} + \delta\rho, \quad p = \bar{p} + \delta p, \quad \mathbf{v} = \delta\mathbf{v}, \quad \mathbf{g} = \delta\mathbf{g}$$

con  $\bar{\rho}$  e  $\bar{p}$  costanti e finiti,  $\delta\rho$ ,  $\delta p$ ,  $\mathbf{v}$ ,  $\mathbf{g}$  infinitesimi. Assunto che i termini finiti risolvano il sistema, si passa a considerare le equazioni al primo ordine:

$$\frac{\partial}{\partial t} \delta\rho + \bar{\rho} \vec{\nabla} \cdot \delta\mathbf{v} = 0, \quad \frac{\partial}{\partial t} \delta\mathbf{v} = -\frac{1}{\bar{\rho}} \vec{\nabla} \delta p + \delta\mathbf{g}, \quad \vec{\nabla} \cdot \delta\mathbf{g} = -4\pi G \delta\rho$$

che, assunto costante il rapporto  $c_s^2 = \delta p / \delta\rho$ , sono equivalenti all'equazione

$$(A1.1) \quad \partial^2 \delta\rho / \partial t^2 - c_s^2 \nabla^2 \delta\rho - 4\pi G \bar{\rho} \delta\rho = 0.$$

**Verifica:** Prendendo il gradiente della seconda equazione si ha:

$$\vec{\nabla} \cdot (\partial / \partial t) \delta\mathbf{v} = -\nabla^2 \delta p / \bar{\rho} + \vec{\nabla} \cdot \delta\mathbf{g},$$

sicché, usando la terza equazione, si ottiene:

$$(\partial / \partial t) \vec{\nabla} \cdot \delta\mathbf{v} = -\nabla^2 \delta p / \bar{\rho} - 4\pi G \delta\rho.$$

Questa equazione va associata alla derivata temporale della prima equazione:

$$(\partial/\partial t)\vec{\nabla} \cdot \delta\mathbf{v} = -(\partial^2/\partial t^2)\delta\rho/\bar{\rho}$$

Uguagliando i secondi membri e ricordando che  $\delta p = c_s^2\delta\rho$ , si ricava l'eq. (A1.1).

Se ci riferiamo ora alla *fluttuazione*

$$\varepsilon(\mathbf{x}) = \delta\rho(\mathbf{x})/\bar{\rho} = [\rho(\mathbf{x}) - \bar{\rho}]/\bar{\rho},$$

essa assume la forma

$$(A1.2) \quad \partial^2\varepsilon/\partial t^2 - c_s^2\nabla^2\varepsilon - 4\pi G\bar{\rho}\varepsilon = 0.$$

Se ne cercano soluzioni separatamente, per le diverse componenti di Fourier

$$(A1.3) \quad \varepsilon(\mathbf{r}, t) = \delta(\mathbf{k}, t) \exp(i\mathbf{k} \cdot \mathbf{r}),$$

che, sostituite nell'eq. (A1.2), danno l'equazione

$$\ddot{\delta} + (c_s^2k^2 - 4\pi G\bar{\rho})\delta = 0,$$

la cui soluzione dipende dal segno del coefficiente  $\omega^2 = c_s^2k^2 - 4\pi G\bar{\rho}$ , che si annulla quando la lunghezza d'onda  $\lambda = 2\pi/k$  vale

$$\lambda_J \equiv c_s\sqrt{\pi/G\bar{\rho}},$$

(lunghezza d'onda di Jeans). Per  $\lambda < \lambda_J$ , sarà

$$\varepsilon(\mathbf{r}, t) = A_+e^{i(\omega t + \mathbf{k} \cdot \mathbf{r})} + A_-e^{-i(\omega t - \mathbf{k} \cdot \mathbf{r})}$$

e, se  $\lambda \ll \lambda_J$  e  $\omega \simeq c_s k$ , posto  $\mathbf{k} = n\mathbf{k}$  e detta  $x = \mathbf{n} \cdot \mathbf{r}$  (ascissa nella direzione di  $\mathbf{n}$ ), l'andamento della fluttuazione di densità diviene

$$\varepsilon(\mathbf{r}, t) = A_+e^{ik(c_s t + x)} + A_-e^{-ik(c_s t - x)};$$

ci sono due onde di compressione, propagantisi nelle direzioni  $\pm\mathbf{n}$  a velocità  $c_s$ .

Nel caso opposto,  $\lambda > \lambda_J$ , risulta  $-\omega^2 \equiv \mu^2 > 0$ . In questo caso

$$\delta(\mathbf{k}, t) = A_+e^{\mu t} + A_-e^{-\mu t},$$

per cui una frazione della fluttuazione (da stabilirsi tramite la condizione iniziale su  $\delta$ ) apparentemente intraprende una crescita esponenziale.

**Conclusione.** Come preannunciato, sussistono due regimi:

- 1) Fluttuazioni di piccola scala si propagano come onde di compressione.
- 2) Per scale  $\lambda > \lambda_J$ , invece, domina l'autogravitazione. In questo regime Jeans dà una crescita esponenziale. Questo risultato erroneo deriva dall'assunzione di un background statico.

## 2. Evoluzione di perturbazioni rispetto un fondo in espansione

È possibile trovare soluzioni esatte delle equazioni (a), (b), (c),

$$(A1.4) \quad \frac{\partial\rho}{\partial t} + \vec{\nabla} \cdot (\rho\mathbf{v}) = 0, \quad \frac{\partial\mathbf{v}}{\partial t} = -(\mathbf{v} \cdot \vec{\nabla})\mathbf{v} - \frac{1}{\rho}\vec{\nabla}p + \mathbf{g}, \quad \vec{\nabla} \cdot \mathbf{g} = -4\pi G\rho,$$

che siano omogenee e isotrope, ma dipendenti dal tempo. Esse si usano come *background*, rispetto cui si studia il comportamento delle disomogeneità.

Fissata un'origine e detta  $R$  la distanza dall'origine, assumiamo che sia

$$(A1.5) \quad \rho = \rho_o (R_o/R)^3, \quad \mathbf{v} = \mathbf{r}\dot{R}/R, \quad \mathbf{g} = -\mathbf{r}(4\pi/3)G\rho, \quad p = 0$$

e  $R$  dipenda da  $t$ , sicché  $\rho_o$  è il valore della densità al tempo  $t_o$ , quando  $R$  vale  $R_o$ . Si mostra che le eq. (A1.4) sono simultaneamente risolte, se

$$(A1.6) \quad R/R_o = (t/t_o)^{2/3}.$$

**Verifica.** Notiamo anzitutto i valori delle seguenti espressioni:

$$\begin{aligned} \vec{\nabla} \cdot \mathbf{r} &= \partial_i x_i = 3 \quad (\text{si omette il simbolo di somma sugli indici ripetuti}), \\ (\mathbf{r} \cdot \vec{\nabla}) \mathbf{r} &= (x_i \partial_i)(x_j \mathbf{n}_j) = x_i \delta_{ij} \mathbf{n}_j = \mathbf{r} \quad (\text{gli } \mathbf{n}_i \text{ sono i versori degli assi}). \end{aligned}$$

I due termini della prima equazione (A1.4), indipendentemente dalla legge  $R(t)$ , danno:

$$\partial \rho / \partial t = -3\rho \dot{R}/R, \quad \vec{\nabla} \cdot (\rho \mathbf{v}) = \rho (\dot{R}/R) \vec{\nabla} \cdot \mathbf{r} = \rho 3(\dot{R}/R)$$

e si cancellano comunque. Allo stesso modo, per qualsiasi  $R(t)$ , la terza equazione è soddisfatta. Veniamo infine alla seconda equazione, da cui si ricava:

$$\mathbf{r} \partial (\dot{R}/R) / \partial t = -(\dot{R}/R)^2 (\mathbf{r} \cdot \vec{\nabla}) \mathbf{r} - \mathbf{r}(4\pi/3)G\rho;$$

essendo poi  $(\mathbf{r} \cdot \vec{\nabla}) \mathbf{r} = \mathbf{r}$ ,

$$\ddot{R}/R - (\dot{R}/R)^2 = -(\dot{R}/R)^2 - 4\pi G\rho/3 \quad \text{e, quindi:} \quad -\ddot{R}/R = 4\pi G\rho/3.$$

Questa equazione è analoga a una delle equazioni di Friedmann per il fattore di scala  $a(t)$ ; la soluzione è la stessa dell'equazione di Friedmann per  $\rho \propto R^{-3}$  (materia non relativistica), data da eq. (A1.6).

**Conclusione.** Se la sostanza cosmica ha pressione trascurabile, possiamo assumere che densità media  $\rho$ , campo di velocità  $\mathbf{v}$  e gravità  $\mathbf{g}$  abbiano le espressioni (A1.5), purchè  $R(t)$  soddisfi l'eq. (A1.6). Tutte le sue caratteristiche fisiche varieranno nel tempo (ma non nello spazio) al variare nel tempo di  $R$ .

L'andamento trovato è quello valido in un modello di Friedmann se la sorgente dell'espansione ha equazione di stato  $p = 0$ . Le soluzioni cosmologiche di *background*, omogenee e isotrope, sono valide soluzioni delle eq. (a), (b), (c).

**EQUAZIONI PER LE PERTURBAZIONI AL PRIMO ORDINE.** Nel seguito denomineremo con l'indice  $B$  (*background*) le soluzioni omogenee isotrope, che descrivono un cosmo in espansione dominata da materia non relativistica,

$$\rho = \rho_B + \delta\rho, \quad \mathbf{v} = \mathbf{v}_B + \delta\mathbf{v}, \quad \mathbf{g} = \mathbf{g}_B + \delta\mathbf{g}, \quad p = \delta p = c_s^2 \delta\rho$$

e studieremo le fluttuazioni  $\varepsilon(\mathbf{x}, t) = \delta\rho(\mathbf{x}, t)/\rho_B(t)$  dipendenti dal tempo e dalle coordinate spaziali. Notiamo che  $\mathbf{x}$  indica le coordinate dei punti a un dato tempo, nel corso dell'espansione cosmica. Se la metrica dello spazio-tempo è

$$ds^2 = c^2 dt^2 - R^2(t)(dr^2 + r^2 d\Omega^2);$$

sarà  $\mathbf{x} = R(t)\mathbf{r}$  e il vettore  $\mathbf{r}$  dà le *coordinate comoventi*.

Il sistema di equazioni differenziali lineari del primo ordine può essere ricondotto a una sola equazione lineare del secondo ordine in  $\varepsilon$ :

$$(A1.7) \quad \partial^2 \varepsilon / \partial t^2 - c_s^2 \nabla^2 \varepsilon + 2(\dot{R}/R) \partial \varepsilon / \partial t - 4\pi G \rho_B \varepsilon = 0.$$

La verifica sarà attuata sull'equazione che si ottiene espandendo in modi di Fourier:

$$\varepsilon(\mathbf{x}, t) = \delta(t) \exp[i\mathbf{k} \cdot \mathbf{r}], \quad \delta\mathbf{v}(\mathbf{x}, t) = \mathbf{v}_1(t) \exp[i\mathbf{k} \cdot \mathbf{r}], \quad \delta\mathbf{g}(\mathbf{x}, t) = \mathbf{g}_1(t) \exp[i\mathbf{k} \cdot \mathbf{r}],$$

cioè:

$$\varepsilon(\mathbf{x}, t) = \delta(t) \exp\left[i\frac{\mathbf{k} \cdot \mathbf{x}}{R(t)}\right], \quad \delta\mathbf{v}(\mathbf{x}, t) = \mathbf{v}_1(t) \exp\left[i\frac{\mathbf{k} \cdot \mathbf{x}}{R(t)}\right], \quad \delta\mathbf{g}(\mathbf{x}, t) = \mathbf{g}_1(t) \exp\left[i\frac{\mathbf{k} \cdot \mathbf{x}}{R(t)}\right].$$

Notiamo anzitutto che:

$$\vec{\nabla}\varepsilon = i\mathbf{k}\varepsilon/R, \quad \text{sicchè:} \quad \nabla^2\varepsilon = -(k^2/R^2)\varepsilon;$$

ne segue che l'equazione (A1.7) diviene:

$$(A1.8) \quad \ddot{\delta} + 2(\dot{R}/R)\dot{\delta} + [(c_s k/R)^2 - 4\pi G \rho_B] \delta = 0.$$

Questa è l'equazione da verificare e anche quella che mette più direttamente in evidenza i possibili regimi di evoluzione delle fluttuazioni.

**Verifica.** Tenendo conto che  $(\delta\mathbf{v} \cdot \vec{\nabla})\mathbf{v}_B = \delta\mathbf{v}\dot{R}/R$  e  $\vec{\nabla} \cdot \mathbf{v}_B = 3\dot{R}/R$ , e essendo

$$\dot{\varepsilon} = \frac{1}{\rho_B} \frac{\partial}{\partial t} \delta\rho - \frac{\dot{\rho}_B}{\rho_B \rho_B} \delta\rho = \frac{1}{\rho_B} \frac{\partial}{\partial t} \delta\rho + 3\frac{\dot{R}}{R} \frac{\delta\rho}{\rho_B},$$

le eq. (A1.4) danno, al primo ordine perturbativo:

$$\dot{\varepsilon} + \vec{\nabla} \cdot \delta\mathbf{v} + (\mathbf{v}_B \cdot \vec{\nabla})\varepsilon = 0,$$

$$\delta\dot{\mathbf{v}} + c_s^2 \vec{\nabla}\varepsilon - \delta\mathbf{g} + (\mathbf{v}_B \cdot \vec{\nabla})\delta\mathbf{v} + \delta\mathbf{v} \frac{\dot{R}}{R} = 0,$$

$$\vec{\nabla} \cdot \delta\mathbf{g} = -4\pi G \delta\rho.$$

Conviene quindi espandere sui modi di Fourier tenendo conto che:

$$\dot{\varepsilon} = \left[ \dot{\delta} - i\delta \frac{\dot{R}\mathbf{k} \cdot \mathbf{x}}{R R(t)} \right] \exp\left[i\frac{\mathbf{k} \cdot \mathbf{x}}{R(t)}\right] = \left[ \dot{\delta} - i\frac{\mathbf{k} \cdot \mathbf{v}_B}{R(t)} \delta \right] \exp\left[i\frac{\mathbf{k} \cdot \mathbf{x}}{R(t)}\right],$$

$$(\mathbf{v}_B \cdot \vec{\nabla})\varepsilon = i\frac{\mathbf{k} \cdot \mathbf{v}_B}{R} \delta \exp\left[i\frac{\mathbf{k} \cdot \mathbf{x}}{R(t)}\right],$$

di modo che risulta

$$\dot{\varepsilon} + (\mathbf{v}_B \cdot \vec{\nabla})\varepsilon = \dot{\delta} \exp\left[i\frac{\mathbf{k} \cdot \mathbf{x}}{R(t)}\right]$$

e, analogamente:

$$\delta\dot{\mathbf{v}} + (\mathbf{v}_B \cdot \vec{\nabla})\delta\mathbf{v} = \dot{\mathbf{v}}_1 \exp\left[i\frac{\mathbf{k} \cdot \mathbf{x}}{R(t)}\right].$$

Le equazioni dinamiche al primo ordine perturbativo diventano allora:

$$\dot{\delta} + \frac{i}{R}\mathbf{k} \cdot \mathbf{v}_1 = 0, \quad \dot{\mathbf{v}}_1 + \frac{i}{R}c_s^2\mathbf{k} \delta - \tilde{\mathbf{g}}_1 + \mathbf{v}_1 \frac{\dot{R}}{R} = 0, \quad \frac{i}{R}\mathbf{k} \cdot \mathbf{g}_1 = -4\pi G \rho_B \delta.$$

Definendo  $\theta = \mathbf{k} \cdot \mathbf{v}_I$  esse si semplificano ancora e assumono la forma

$$(A1.9) \quad \dot{\delta} + \frac{i}{R}\theta = 0, \quad \frac{i}{R}\dot{\theta} = \left(\frac{c_s k}{R}\right)^2 \delta - 4\pi G \rho_B \delta - \frac{i}{R} \frac{\dot{R}}{R} \theta$$

in cui è sparita la variabile  $\mathbf{g}_1$ . Nella derivata della prima equazione

$$\ddot{\delta} + i\dot{\theta}/R - (\dot{R}/R)i\theta/R = 0.$$

si può poi sostituire la seconda, ottenendo

$$\ddot{\delta} - 2\frac{\dot{R}}{R}\frac{i}{R}\theta + \left[\left(\frac{c_s k}{R}\right)^2 - 4\pi G \rho_B\right]\delta = 0.$$

Qui si sostituisce ancora la prima delle eq. (A1.9) ottenendo l'eq. (A1.8).

**Conclusione.** L'eq. (A1.8) è simile a quella di Jeans. La differenza principale è la presenza del termine in  $\dot{\delta}$ . La differenza nel coefficiente del termine in  $\delta$  si riassume quando si passa da numero a lunghezza d'onda, come vedremo subito all'inizio del prossimo paragrafo.

### 3. Scala di Jeans e soluzioni di Lifshitz

Infatti, la relazione tra numero a lunghezza d'onda si scrive

$$\lambda = 2\pi R/k,$$

e mostra essere  $\lambda \propto R$  (le lunghezze crescono col fattore di scala) se  $k$  è indipendente da  $R$ . Il coefficiente di  $\delta$  si annulla per lunghezza d'onda

$$(A1.10) \quad \lambda_J = c_s \sqrt{\pi/G\rho}$$

e questa espressione è identica a quella di Jeans; vedremo subito che essa separa ancora un regime di instabilità gravitazionale da quello oscillatorio.

Usando  $a$ , invece di  $R$ , come fattore di scala, l'eq. (A1.8) diviene

$$(A1.11) \quad \ddot{\delta} + 2(\dot{a}/a)\dot{\delta} + [(c_s k/a)^2 - 4\pi G \rho_B]\delta = 0.$$

Teniamo poi conto del fatto che, se  $p = 0$ , per le equazioni di Friedmann risulta  $\dot{a}/a = 2/3t$  e  $4\pi G \rho = 2/3t^2$ . L'eq. (A1.11) assume perciò la forma

$$(A1.12) \quad \ddot{\delta} + (4/3t)\dot{\delta} + [(c_s k/a)^2 - (2/3t^2)]\delta = 0,$$

di cui esiste un integrale generale esprimibile tramite funzioni di Bessel.

È tuttavia fisicamente più significativo considerare separatamente i casi:

- 1)  $\lambda \gg \lambda_J$ , che mostreremo dar luogo a instabilità gravitazionale, sebbene la crescita delle fluttuazioni avvenga con legge di potenza (e non esponenziale).
- 2)  $\lambda \ll \lambda_J$ , che si tratta sulla linea di quanto esposto all'inizio del par. A1.6 e verrà discusso ampiamente nell'ambito del caso fisico riguardante il fluido fotoni-barioni, nel prossimo capitolo. In questo caso si ha una soluzione oscillatoria.

Consideriamo anzitutto il caso

$$(A1.13) \quad (c_s k/a)^2 \ll 2/3t^2,$$

quando, l'equazione (A1.11) si può approssimare:

$$(A1.14a) \quad \ddot{\delta} + (4/3t)\dot{\delta} - (2/3t^2)\delta = 0.$$

Di tale equazione si cercano integrali della forma  $A t^\alpha$  con  $A$  e  $\alpha$  costanti. Dovrà allora essere:

$$\alpha(\alpha - 1)\frac{\delta}{t^2} + \frac{4}{3t}\alpha\frac{\delta}{t} - \frac{2}{3t^2}\delta = 0, \quad \text{e, quindi:} \quad \alpha(\alpha - 1) + \frac{4}{3}\alpha - \frac{2}{3} = 0.$$

Questa equazione algebrica si riscrive:

$$\alpha^2 + \alpha/3 - 2/3 = 0 \quad \text{e ha radici:} \quad \alpha_{\pm} = -\frac{1}{6} \pm \sqrt{\frac{1}{36} + \frac{24}{36}} = -\frac{1}{6} \pm \frac{5}{6}.$$

L'integrale generale dell'equazione differenziale risulta allora:

$$\delta = \delta_+ (t/t_i)^{\frac{5}{3}} + \delta_- (t/t_i)^{-1}$$

dove  $t_i$  è un tempo di riferimento e  $\delta_{\pm}$  sono costanti da fissare dalle condizioni iniziali.

La crescita di una fluttuazione gravitazionalmente instabile nel corso di un'espansione dominata dalla radiazione può essere estrapolata in modo simile. In questo caso, essendo  $\dot{a}/a = 1/2t$  e  $4\pi G\rho = 1/4t^2$ , l'eq. (A1.14a) diviene

$$(A1.14b) \quad \ddot{\delta} + (1/t)\dot{\delta} - (1/t^2)\delta = 0.$$

e, tramite passaggi analoghi a quelli visti per eq. (A1.14a), si ha la soluzione

$$\delta = \delta_+ (t/t_i) + \delta_- (t/t_i)^{-1}$$

Si noti che in questo caso è  $t \propto a^2$ , per cui il modo crescente cresce col quadrato del fattore di scala.

**Conclusione.** Anche rispetto a un *background* in espansione, le fluttuazioni possono trovarsi in regime sonico (sotto la scala  $\lambda_J$ ) o in regime di instabilità gravitazionale (sopra la scala  $\lambda_J$ ). In questo secondo caso ci sono un modo crescente e uno decrescente. Il primo, che interessa per lo sviluppo di strutture, prevede una crescita a una legge di potenza  $\delta\rho/\rho_B \propto t^{2/3} \propto a$ , molto diversa dalla dipendenza esponenziale di Jeans. Sul primo caso, torneremo in un ambito più generale.

#### 4. Caso relativistico. La gauge gravitazionale

Nel caso relativistico il punto di partenza saranno le equazioni di Einstein linearizzate. In un background omogeneo e isotropo si studierà l'autogravitazione di una sostanza fluida, con parametro di stato  $w = p/\rho$  e velocità del suono  $c_s = \sqrt{\delta p/\delta\rho}$ , in cui ci siano fluttuazioni di densità  $\varepsilon(\mathbf{r}, \tau) = \delta\rho(\mathbf{r}, \tau)/\rho$  ( $\ll 1$ ); le  $\delta(\mathbf{k}, \tau)$ , con numero d'onda

$$(A1.15) \quad \mathbf{k} = k\mathbf{n}$$

saranno le loro componenti di Fourier.

Nella metrica di background conviene comunque usare il "tempo conforme"  $\tau$ :

$$(A1.16) \quad ds^2 = g_{\mu\nu}dr^\mu dr^\nu = a^2(\tau)[d\tau^2 - \delta_{ij}dr^i dr^j];$$

risulta così

$$c dt = a(\tau) d\tau \quad \text{ovvero} \quad d\tau = c dt/a(t).$$

da cui si vede facilmente che è

$$(A1.17) \quad \begin{array}{ll} \text{per } a \propto t^{1/2} & \tau \propto t^{1/2} \propto a \\ \text{per } a \propto t^{2/3} & \tau \propto t^{1/3} \propto a^{1/2} \end{array}$$

In eq. (A1.16) le  $r^i$  sono coordinate cartesiane ortogonali. Se si scelgono altre coordinate spaziali, la parte spaziale del tensore metrico dovrà essere opportunamente cambiata.

### ***Il tempo universale. La gauge sincrona.***

In presenza di fluttuazioni, la sorgente della gravità non è più omogenea, sicché le componenti del tensore metrico (A1.15) dovranno dipendere dalle coordinate  $x^\mu$ . Se le disomogeneità sono lineari, anche le perturbazioni del tensore metrico saranno lineari.

Le stesse perturbazioni fisiche possono peraltro essere descritte in forme diverse, operando opportune scelte delle coordinate  $x^\mu$ . Questa scelta di sistema di riferimento è denominata, per precise ragioni fisiche su cui qui non entreremo, *scelta della gauge gravitazionale*.

Si dimostra anzitutto che è sempre possibile, *nel caso di fluttuazioni lineari*, scegliere la gauge di modo che il tempo  $\tau$  sia «universale» e si possano sincronizzare tutti gli eventi. (Anzi, esistono  $\infty$  possibili scelte di questo tipo.)

In questo caso, l'espressione (A1.15) diviene

$$(A1.18) \quad ds^2 = g_{\mu\nu} dr^\mu dr^\nu = a^2(\tau) [d\tau^2 - (\delta_{ij} + h_{ij}) dr^i dr^j]$$

e il campo gravitazionale è descritto dal tensore  $h_{ij}(\tau, \mathbf{r})$ , la cui trasformata di Fourier si dimostra avere necessariamente la forma

$$(A1.19) \quad \begin{aligned} h_{ij}(\tau, \mathbf{r}) &= \int d^3k e^{i\mathbf{k}\cdot\mathbf{r}} \left[ n_i n_j h(\tau, \mathbf{k}) + \left( n_i n_j - \frac{1}{2} \delta_{ij} \right) 6\eta(\tau, \mathbf{k}) \right] \equiv \\ &\equiv \int d^3k e^{i\mathbf{k}\cdot\mathbf{r}} [n_i n_j \omega(\tau, \mathbf{k}) - 3\delta_{ij} \eta(\tau, \mathbf{k})], \end{aligned}$$

per cui i potenziali gravitazionali relativistici si possono ricondurre a due delle 3 funzioni  $h(\tau, \mathbf{k})$ ,  $\eta(\tau, \mathbf{k})$ ,  $\omega(\tau, \mathbf{k})$ .

### ***La gauge newtoniana***

Definiamo allora la funzione

$$(A1.20) \quad \beta(\tau, \mathbf{x}) = \int d^3k e^{i\mathbf{k}\cdot\mathbf{r}} \frac{1}{2k^2} \omega(\tau, \mathbf{k})$$

che utilizzeremo per effettuare il seguente cambio di sistema di riferimento o trasformazione di gauge:

$$(A1.21) \quad \hat{x}^0 = x^0 + \dot{\beta}, \quad \hat{\mathbf{x}} = \mathbf{x} + \vec{\nabla}\beta.$$

La trasformata di Fourier della funzione  $\beta$ , dipendente da  $\mathbf{k}$ , e che continueremo a denominare  $\beta$ , coincide allora con  $\omega/2k^2$ . Essa si usa per definire i “potenziali”

$$(A1.22) \quad \Psi(\tau, \mathbf{k}) = \beta' + (\dot{a}/a)\beta', \quad \Phi(\tau, \mathbf{k}) = \eta - (\dot{a}/a)\beta'$$

(la derivata rispetto a  $\tau$  viene indicata con  $'$ , qui e nel seguito) e si dimostra allora che la metrica perturbata può essere posta nella forma

$$(A1.23) \quad ds^2 = \hat{g}_{\mu\nu} d\hat{r}^\mu d\hat{r}^\nu = a^2(\hat{\tau}) \{ [1 + 2\Psi(\hat{\tau}, \hat{\mathbf{r}})] d\hat{\tau}^2 - [1 - 2\Phi(\hat{\tau}, \hat{\mathbf{r}})] \delta_{ij} d\hat{r}^i d\hat{r}^j \}.$$

in cui, di nuovo, si usano le trasformate delle  $\Psi$  e  $\Phi$ , dipendenti ora da  $\hat{\mathbf{r}}$ , senza variare la denominazione.

La gravità è allora descritta dai potenziali  $\Psi(\tau, \mathbf{r})$  a  $\Phi(\tau, \mathbf{r})$  che, in assenza di stress anisotropi, si possono dimostrare essere coincidenti.

Questa scelta del tempo implica tempi propri diversi punti per punto, ma, al limite non relativistico, il potenziale  $\Psi(\hat{\tau}, \hat{\mathbf{r}})$  coincide col potenziale newtoniano. Questa gauge viene perciò denominata «gauge newtoniana conforme», per sottolineare altresì l'uso di una coordinata temporale diversa dal tempo ordinario.

### Caso relativistico. Equazioni del moto

In questo libro, ove non altrimenti specificato, ci porremo sempre nella gauge sincrona. Trascurando termini  $\mathcal{O}(c^2)$ , dalle equazione di Einstein, in tale gauge, si ottengono le equazioni

$$(A1.24) \quad \begin{aligned} \delta' + (1+w)(kv + h'/2) + 3(a'/a)(c_s^2 - w)\delta &= 0, \\ v' + \left(1 - 3w + \frac{w'}{1+w}\right) \frac{a'}{a} v - \frac{c_s^2}{1+w} k\delta &= 0, \end{aligned}$$

per la cui prova rimandiamo a testi specializzati.

Sotto la condizione  $w = c_s^2 = \text{const.}$ , dalla prima delle eq. (A1.24) si ottiene

$$(A1.25) \quad kv = -h'/2 - \delta'/(1+w)$$

e si ricava, derivando, l'equazione

$$\delta'' + (1+w)kv' + (1+w)h''/2 = 0.$$

Si usa qui la seconda delle (A1.24) per sostituire  $v'$  e, nell'equazione così ricavata, si sostituisce  $kv$  con l'espressione (A1.25). Si ottiene così l'equazione

$$(A1.26) \quad \delta'' + (1-3w)(a'/a)\delta' + c_s^2 k^2 \delta + (1+w)(1/2)[h'' + (1-3w)h'a'/a] = 0,$$

in cui non compare  $v$ , nè dipendenza da  $\mathbf{n}$ , sicchè  $\delta$  e  $h$  sono funzioni di  $\tau$  e del solo modulo  $k$ . Accanto all'eq. (A1.26) serve l'equazione

$$(A1.27) \quad h'' + (a'/a)h' = -8\pi G a^2 \rho [1 + 3c_s^2] \delta$$

derivata anch'essa dalle equazione di Einstein; non occorre invece esplicitare l'equazione soddisfatta da  $\omega$  o da  $\eta$ . Le eq. (A1.26)-(A1.27) determinano l'andamento temporale delle fluttuazioni.

Per una sostanza con  $w \simeq 0$  (e quindi  $c_s \simeq 0$ ), il secondo membro di eq. (A1.22) può essere sostituito in eq. (A1.21), ricavando l'equazione

$$\delta'' + (a'/a)\delta' - 4\pi G a^2 \rho \delta = 0,$$

in cui non compare il campo gravitazionale. Questa equazione coincide con l'eq. (A1.14b).

Se invece è  $c_s^2 = w = 1/3$ , avremo

$$\delta'' + k^2\delta/3 + (2/3)h'' = 0,$$

sicché l'eq. (A1.27) non è utilizzabile per eliminare la variabile  $h$ .

## 5. Il problema di Jeans relativistico

Tuttavia, quando  $c_s^2$  e  $w$  sono costanti o variano lentamente, si può trattare il problema di Jeans da una configurazione iniziale stazionaria. Se, a un certo istante, una fluttuazione è completamente «potenziale», non c'è in essa alcun campo di velocità ( $v = 0$ ) ed è  $\delta' = 0$ . La prima delle eq. (A1.24) dà allora è anche  $h' = 0$  e, all'istante considerato, dalle eq. (A1.26)-(A1.27), si ricava:

$$\delta'' + c_s^2 k^2 \delta + (1+w)(1/2)h'' = 0, \quad h'' = -8\pi G a^2 \rho [1 + 3c_s^2] \delta.$$

Si può così eliminare comunque la variabile  $h$  ottenendo la relazione

$$(A1.28) \quad \delta'' + \{c_s^2 k^2 - (1+w)4\pi G a^2 \rho [1 + 3c_s^2]\} \delta = 0$$

che non è un'equazione differenziale, ma esprime la «forza totale» che agisce sulla fluttuazione nel momento in cui essa è puramente potenziale. La lunghezza di Jeans si ottiene annullando la quantità in parentesi graffa, ottenendo

$$\lambda_J = \frac{2\pi a}{k_J} \simeq c_s \sqrt{\frac{\pi}{G\rho(1+w)(1+3c_s^2)}}.$$

## 6. Modi di fluttuazione relativistici

Per la risoluzione del problema dinamico vanno ancora distinte le scale al di sotto e al di sopra della scala di Jeans.

Al di sotto della scala di Jeans si possono trascurare i termini gravitazionali sicché, a partire dall'eq. (A1.21), abbiamo l'equazione

$$(A1.29) \quad \delta'' + (1-3w)(a'/a)\delta' + c_s^2 k^2 \delta = 0$$

che, per  $w = c_s^2 = 1/3$ , diviene semplicemente

$$(A1.30) \quad \delta'' + c_s^2 k^2 \delta = 0 \quad \text{e ha soluzioni:} \quad \delta(\tau) = A_+ e^{ikc_s\tau} + A_- e^{-ikc_s\tau}$$

( $A_{\pm}$ : costanti arbitrarie) in funzione del tempo conforme  $\tau$ . La dipendenza di  $\delta$  da  $t$  verrà discussa nel prossimo capitolo, per specifici casi fisici. Atteniamoci qui al caso  $w = 1/3$ , in cui  $a = (t/\bar{t})^{1/2}$ , dove  $\bar{t}$  è un opportuno tempo di normalizzazione. Essendo  $\tau = \int^t dt'/a(t')$ , è facile vedere che l'argomento dell'esponenziale in eq. (A1.30) si può scrivere  $i2ck\bar{t}/a = 2\pi i(2c_s\bar{t}/\lambda)$ . Si noti il fattore 2 nell'ultima parentesi, conseguenza del regime espansivo ed assente nel caso statico.

Il caso  $w = 0$  è un po' più complesso e porterebbe a sostituire tale fattore 2 con un fattore 3.

Formalmente, a partire da eq. (A1.26)-(A1.27), si possono definire tutti i modi di fluttuazione possibili, al di sopra della scala di Jeans, al di là del puro modo adiabatico, assumendo un background in espansione radiativa. In tale contesto le eq. (A1.26)-(A1.27) divengono

$$(A1.31) \quad \delta'' + (2/3)h'' = 0, \quad h'' + h'/\tau + 6\delta/\tau^2 = 0.$$

Derivando due volte la seconda equazione moltiplicata per  $\tau^2$ , e sostituendo in essa la prima, si ottiene

$$\tau \frac{d^4 h}{d\tau^4} + 5 \frac{d^3 h}{d\tau^3} = 0.$$

Questa equazione ha soluzione

$$(A1.32) \quad h = A\tau^{-2} + B + C\tau + D\tau^2$$

ove  $A, B, C, D$  sono costanti arbitrarie. Tra questi 4 «modi» sceglieremo quello in crescita più rapida, destinato quindi a prevalere nel tempo. In altri contesti, tuttavia, anche gli altri 3 modi risultano significativi.

Assumendo anche  $\delta = \tilde{D}\tau^\alpha$  e sostituendo nelle eq. (A1.31) si ricava

$$(A1.33) \quad \alpha(\alpha - 1)\tilde{D} = -(4/3)D, \quad 4D + 6\tilde{D} = 0$$

da cui si ha

$$(A1.34) \quad \tilde{D} = -(2/3)D, \quad \alpha = 2, -1$$

**Conclusioni.** Anche nel caso relativistico valgono i due regimi, di instabilità gravitazionale e di oscillazioni acustiche. Nel modo instabile crescente col ritmo maggiore, essendo ora  $t \propto \tau^2 \propto a^2$ , avremo  $\delta \propto t \propto a^2$ . Esistono però altri modi, anche crescenti. Questo risultato conferma quanto estrapolato a partire da eq. (A1.14b).

## CAPITOLO VIII

# Evoluzione lineare delle disomogeneità

## Le scale cosmiche e lo sviluppo delle fluttuazioni nelle ere intermedie

### 1. Scale e disomogeneità

Cinque secondi dopo il *Big-bang*, la temperatura del cosmo scende sotto la massa dell'elettrone, gli elettroni che si annichilano non vengono più rimpiazzati e «ri-scaldano» il CMB. Da allora il numero degli stati di spin nel brodo termico non varia più e la relazione tempo-temperatura diviene stabilmente

$$(VIII.1) \quad \frac{t}{\text{sec.}} \times \left( \frac{T}{\text{MeV}} \right)^2 = 1.31 u_\nu \quad \text{con} \quad u_\nu = \sqrt{\frac{3.36}{2 + 0.454 N_\nu}}$$

dove  $N_\nu$  è il numero di neutrini (per 3 famiglie,  $u(3) = 1$ ; cfr. l'eq. (24) del capitolo VI). Questa relazione si riscrive anche nella forma

$$(VIII.2) \quad \frac{t}{10^3 y} \left( \frac{1+z}{10^4} \right)^2 = 7.50 u_\nu u_T^2 \quad \text{con} \quad u_T = \frac{2.73 \text{ K}}{T_0}, \quad (y: \text{anni}),$$

più maneggevole quando i tempi si allungano e le temperature si abbassano. Entrambe le relazioni valgono finché la materia non-relativistica resta trascurabile, quindi fino a *redshift* significativamente superiori a

$$(VIII.3) \quad 1 + z_{eq} = 2.39 \times 10^4 \Omega_{om} h^2 u_\nu^2 u_T^4 = 3.35 \times 10^3 (\Omega_{om} h^2 / 0.14) u_\nu^2 u_T^4,$$

quando *materia* e *radiazione* avranno eguali densità (equivalenza).

La relazione (VIII.1) ci dice anche che la nucleosintesi primordiale (BBN) si conclude una decina di minuti dopo il *Big-bang* ( $T \simeq 50 \text{ keV}$ ). Entriamo nelle *ere intermedie* e la chimica del cosmo non muterà sin quando non s'accenderanno le prime stelle; tranne che per la transizione da plasma a gas nella componente barionica, a  $z_{ric} \sim 1100$  ( $T \sim 2900 \text{ K} \sim 0.23 \text{ eV}$ ), comunque dopo  $z_{eq}$ .

Usando l'eq. (VIII.2) fino a  $z_{eq}$ , si stima in 70 000 anni il tempo dal *Big-bang* all'equivalenza. A  $z < z_{eq}$ , la relazione tempo-*redshift* diviene  $t^2(1+z)^3 = \text{const.}$ : da  $z_{eq}$  a

$z \simeq 10^3$  il tempo cresce allora di un fattore  $3.35^{1.5} \simeq 6.13$ ; quindi a  $z \sim 10^3$  l'età del cosmo sarà  $\sim 430\,000$  anni. Possiamo considerare concluse le ere intermedie a  $z \sim 100$ , alla vigilia dell'accensione delle prime stelle. L'età del cosmo è allora di 13-14 milioni di anni, un millesimo dei 13-14 miliardi di anni attuali.

I 70 000 anni iniziali sono essenziali nel plasmare le strutture nel cosmo che oggi osserviamo. Come mostreremo in questo capitolo:

- 1) Tutte le scale su cui oggi osserviamo strutture astronomiche sono entrate nell'orizzonte tra BBN e  $z_{eq}$ .
- 2) In questo periodo scala di Jeans e orizzonte sono praticamente coincidenti. Le fluttuazioni nei barioni e nei fotoni, sulle scale su cui si formeranno oggetti cosmici, appena raggiunte dall'orizzonte entrano in regime sonico.
- 3) Le fluttuazioni nella DM, invece, restano *congelate* dall'entrata nell'orizzonte all'equivalenza. Ciò rompe definitivamente l'invarianza di scala nel cosmo.

## 2. La designazione delle scale

Prima di proseguire diamo un quadro riassuntivo di come si classificano le *scale* cosmiche: il punto di riferimento è la massa  $M$  di materia non-relativistica che una scala contiene (*scala di massa*). Accanto a essa si usano:

- 1) La *scala di lunghezza*

$$(VIII.4) \quad \lambda = \left( \frac{3M}{4\pi\rho_m} \right)^{\frac{1}{3}}, \quad \text{dalla relazione} \quad M = \frac{4\pi}{3} \lambda^3 \rho_m \quad (\text{geometria sferica}).$$

$\rho_m$  è la densità media della materia, *non la densità interna delle strutture*.

- 2) La *scala di lunghezza comovente*

$$(VIII.5) \quad \lambda_c = \lambda \times a_o/a = \lambda \times (1+z),$$

lunghezza odierna che corrisponde a  $\lambda$  allora (il fattore di scala  $a$  vale oggi  $a_o$ ).

Ad esempio, per una galassia di massa  $M \simeq 3 \times 10^{11} M_\odot$ , la scala comovente  $R_g$  si ottiene imponendo:  $3 \times 10^{11} = (4\pi/3) \times 2.78 \times 10^{11} \Omega_{om} h^2 \times R_g^3$ . Risulta quindi  $R_g \simeq 1/(4\Omega_{om} h^2)^{1/3} \simeq 1-2$  Mpc.

- 3) Il *numero d'onda*

$$(VIII.6) \quad k = 2\pi a/\lambda = 2\pi a_o/\lambda_c.$$

Trattando anisotropie del CMB si considerano poi

- 1) La *scala angolare*

$$(VIII.7) \quad \theta = \lambda_c/R_H,$$

dove  $R_H$  è la distanza comovente della banda dell'ultimo scattering che, con buona approssimazione, coincide con l'orizzonte di particella.

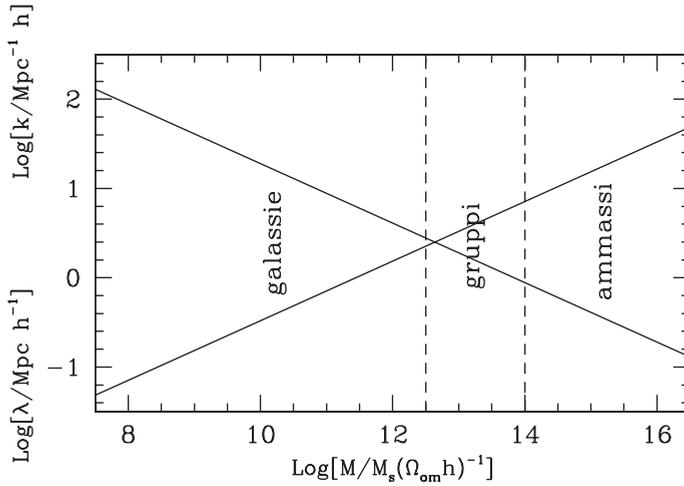


FIGURA VIII.1

Relazione tra scale di massa, lunghezze comoventi e numeri d'onda. I confini tra galassie, gruppi e ammassi di galassie sono approssimativi. Oltre  $\sim 10^{16}M_{\odot}$  passiamo nel campo dei superammassi.

2) Il *numero d'onda angolare* (discreto; il cielo ha un'area angolare finita)

$$(VIII.8) \quad l = 2\pi/\theta = 2\pi R_H/\lambda_c,$$

equivalente angolare di  $k$ . Massa e numero barionico si assumono conservati. Detta  $\rho_{om}$  la densità attuale della materia sarà allora  $\rho_m = \rho_{om}(a_0/a)^3$  e quindi

$$(VIII.9) \quad M = \frac{4\pi}{3} \lambda_c^3 \rho_{om} = \frac{4\pi}{3} \left( \frac{2\pi a_0}{k} \right)^3 \rho_{om}$$

sicché  $M$ ,  $\lambda_c$  e  $k$  designano una scala in modo indipendente da  $a$  o  $z$ . In genere le scale sono quindi designate da  $M$  o  $k$  (vedi anche Fig. VIII.1).

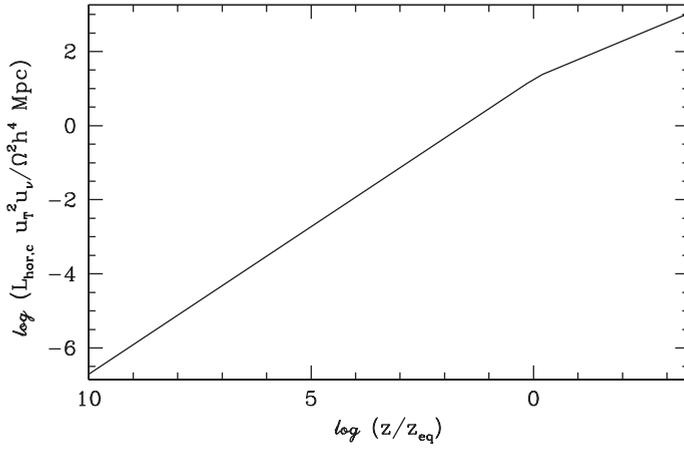
### 3. La scala dell'orizzonte

In figura VIII.2 diamo la scala di lunghezza (comovente) dell'orizzonte, prima, durante e dopo l'equivalenza. In figura VIII.3 è rappresentato l'andamento delle scale di massa corrispondenti, magnificando la zona attorno all'equivalenza.

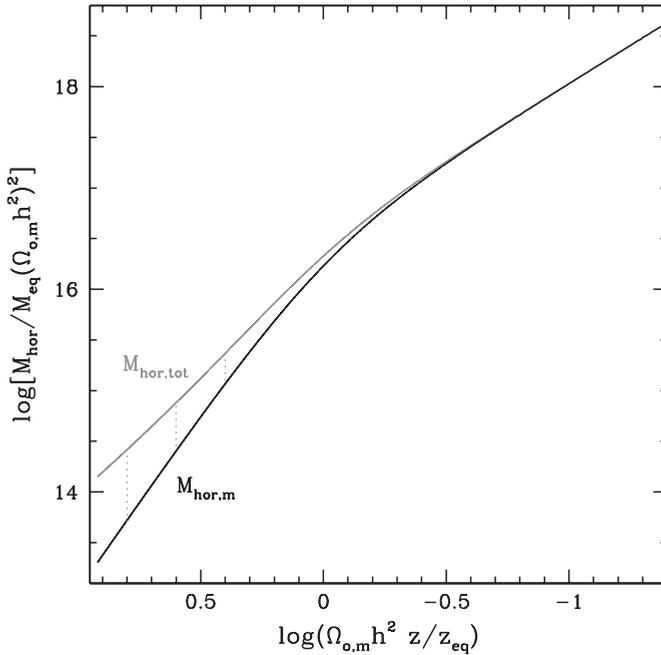
Questi grafici logaritmici mostrano come l'orizzonte vada gradualmente crescendo, con legge diversa prima e dopo l'equivalenza. In effetti, le espressioni

$$(VIII.10) \quad L_{hor} \simeq \frac{2c}{H} = 3ct \quad (\text{per } z \ll z_{eq}), \quad L_{hor} \simeq \frac{c}{H} = 2ct \quad (\text{per } z \gg z_{eq})$$

differiscono solo per un piccolo fattore numerico, ma l'elemento discriminante è la



**FIGURA VIII.2**  
Lunghezza (comovete) della scala dell'orizzonte attorno all'equivalenza e alla ricombinazione.



**FIGURA VIII.3**  
Orizzonte attorno all'equivalenza e alla ricombinazione. La curva scura dà il *redshift* a cui ciascuna scala di massa entra nell'orizzonte. Spostandosi lungo la verticale, si incontra la curva più chiara, che indica la corrispondente energia totale (inclusa quella della componente radiativa). (si veda l'appendice TAVOLE A COLORI).

diversa dipendenza di  $t$  da *redshift*  $z$ , temperatura  $T = T_o(1+z)$ , fattore di scala  $a/a_o = 1/(1+z)$ , prima e dopo l'equivalenza.

Si considerino le definizioni:

$$(VIII.11) \quad H_{eq}^2 = \frac{8\pi}{3} G (\rho_{m,eq} + \rho_{r,eq}) = 2 \frac{8\pi}{3} G \rho_{m,eq} = 2 \frac{8\pi}{3} G \rho_{r,eq} = 2 \tilde{H}_{eq}^2.$$

Inoltre, essendo  $c/H_o = [300\,000 \text{ (km/s)}]/[100 h \text{ (km/s)/Mpc}]$ , risulta

$$(VIII.12) \quad \frac{c}{\tilde{H}_{eq}} = \frac{3000 \text{ Mpc}}{(\Omega_{om} h^2)^{1/2}} [2.39 \times 10^4 \Omega_{om} h^2 u_T^4 u_V^2]^{-3/2} = \\ = 8.12 \times 10^{-4} \text{ Mpc } (\Omega_{om} h^2)^{-2} u_T^{-6} u_V^{-3}$$

Questo valore sarà utilizzato per normalizzare tutte le scale, nel resto di questo paragrafo.

### Scale di lunghezza dell'orizzonte

Nel capitolo IV si sono ricavati integrali analitici delle equazioni di Friedmann attorno all'equivalenza. In particolare, l'orizzonte all'equivalenza è

$$(VIII.13) \quad L_{hor}(z_{eq}) = 2(\sqrt{2} - 1)c/\tilde{H}_{eq} = 0.8284 c/\tilde{H}_{eq}.$$

Finché la DE può essere trascurata, sarà  $H^2 \propto \rho_m$  o  $\rho_r$ , rispettivamente dopo o prima dell'equivalenza; avremo quindi

$$(VIII.14) \quad z \ll z_{eq}: \frac{c}{H} = \frac{c}{\tilde{H}_{eq}} \left( \frac{1+z_{eq}}{1+z} \right)^{3/2}, \quad z \gg z_{eq}: \frac{c}{H} = \frac{c}{\tilde{H}_{eq}} \left( \frac{1+z_{eq}}{1+z} \right)^2,$$

Ne segue che, oltre all'espressione (VIII.13), valida all'equivalenza, avremo:

$$(VIII.15) \quad \begin{aligned} z \ll z_{eq}: \quad L_{hor} &= \frac{2c}{\tilde{H}_{eq}} \left( \frac{1+z_{eq}}{1+z} \right)^{3/2} = \frac{2c}{\tilde{H}_{eq}} \left( \frac{a}{a_{eq}} \right)^{3/2} \\ z \gg z_{eq}: \quad L_{hor} &= \frac{c}{\tilde{H}_{eq}} \left( \frac{1+z_{eq}}{1+z} \right)^2 = \frac{c}{\tilde{H}_{eq}} \left( \frac{a}{a_{eq}} \right)^2. \end{aligned}$$

L'espressione delle scale *comoventi* dell'orizzonte sarà quindi:

$$(VIII.16) \quad \begin{aligned} z \ll z_{eq}: \quad L_{hor,c} &= \frac{2c}{\tilde{H}_{eq}} \frac{(1+z_{eq})^{3/2}}{(1+z)^{1/2}} = \frac{6000 h^{-1} \text{Mpc}}{\Omega_{om}^{1/2} (1+z)^{1/2}} \\ z = z_{eq}: \quad L_{hor,c} &= 0.8284 \times \frac{c}{\tilde{H}_{eq}} (1+z_{eq}) = \frac{16.1 h^{-1} \text{Mpc}}{\Omega_{om} h u_V u_T^2} \\ z \gg z_{eq}: \quad L_{hor,c} &= \frac{c}{\tilde{H}_{eq}} \frac{(1+z_{eq})^2}{1+z} = 46.4 (h u_V u_T^2) \frac{10^4}{1+z} h^{-1} \text{Mpc} \end{aligned}$$

Per ricavare l'espressione numerica si è tenuto conto dell'eq. (VIII.3) e in particolare che, sulla base delle eq. (VIII.3) e (VIII.12), risulta

$$(VIII.17) \quad \frac{c}{\bar{H}_{eq}}(1+z_{eq}) = \frac{19.4 h^{-1} \text{Mpc}}{\Omega_{om} h u_v u_T^2}.$$

### Numeri d'onda dell'orizzonte

Ricordiamo che  $k = 2\pi/L_c$ . Useremo inoltre l'espressione (VIII.17) assumendo che sia  $T_o = 2.73$  e  $u_v = 1$ ; eventuali valori «anomali» della temperatura del CMB o del numero di neutrini richiederanno modifiche. Avremo:

$$(VIII.18) \quad \begin{aligned} z \ll z_{eq} : & \quad k_{hor}(z) \simeq 0.162 \text{ Mpc}^{-1} \Omega_{om} h^2 [(1+z)/(1+z_{eq})]^{1/2} \\ z = z_{eq} : & \quad k_{hor,eq} \simeq 0.391 \text{ Mpc}^{-1} \Omega_{om} h^2 \\ z \gg z_{eq} : & \quad k_{hor}(z) \simeq 0.324 \text{ Mpc}^{-1} \Omega_{om} h^2 [(1+z)/(1+z_{eq})] \end{aligned}$$

### Scala di massa dell'orizzonte

La scala di massa dell'orizzonte

$$(VIII.19) \quad M_{hor}(z) = \frac{4\pi}{3} L_{hor}^3(z) \rho(z)$$

ha un significato non ambiguo a  $z \ll z_{eq}$ , dove si assumerà  $\rho(z) \equiv \rho_m(z)$ . A  $z \sim z_{eq}$  o  $z \gg z_{eq}$ , la densità della componente radiativa è confrontabile o supera quella della componente materiale. Tuttavia, mentre l'energia totale della componente materiale, entro una determinata scala comovente, è costante, l'energia della componente radiativa è soggetta a *redshift*. Ci riferiremo quindi anzitutto alla massa data dalla componente materiale, a tutti i *redshift*, per poi considerare anche la massa totale a  $z \gg z_{eq}$ .

Sarà allora

$$(VIII.20) \quad M_{hor,m}(z) = \frac{4\pi}{3} L_{hor}^3(z) \rho_m(z) = \frac{4\pi}{3} \left[ \frac{2\pi}{k(z)} \right]^3 \rho_{o,cr} \Omega_{o,m}$$

e, quindi,

$$(VIII.21) \quad \begin{aligned} z \ll z_{eq} : & \quad M_{hor,m}(z) \propto (1+z)^{-3/2} \propto T^{-3/2} \propto a^{3/2} \\ z \gg z_{eq} : & \quad M_{hor,m}(z) \propto (1+z)^{-3} \propto T^{-3} \propto a^3 \\ & \quad M_{hor,tot}(z) \propto (1+z)^{-2} \propto T^{-2} \propto a^2 \end{aligned}$$

La massa (in materia) nell'orizzonte all'equivalenza sarà infine:

$$(VIII.22) \quad M_{hor,m,eq} = \frac{4\pi}{3} \left[ \frac{2\pi}{k_{eq}} \right]^3 \rho_{o,cr} \Omega_{o,m} = 4.84 \times 10^{15} (\Omega_{o,m} h^2)^{-2} M_{\odot}$$

Ad esempio, per  $\Omega_{o,m} \simeq 0.3$  e  $h \simeq 0.7$ , sarà  $M_{hor,m,eq} \simeq 1.57 \times 10^{17} h^{-1} M_{\odot}$ , sicuramente al di sopra della massa del maggiore degli ammassi di galassie.

L'andamento della massa nell'orizzonte è rappresentato in figura VIII.3.

#### 4. Interazione barioni-fotoni nelle ere intermedie

Nelle ere intermedie, le forze rilevanti sono quelle gravitazionali e elettromagnetiche. Queste ultime causano la diffusione Thomson fotone-elettrone (carica  $e$ , massa  $m_e$ ), che ha sezione d'urto

$$(VIII.23) \quad \sigma_T = \frac{8\pi}{3} \left( \frac{e^2}{m_e c^2} \right)^2 = 0.67 \times 10^{-24} \text{cm}^2;$$

La sezione d'urto Thomson del protone sarebbe quindi minore di un fattore  $(m_e/m_p)^2 \sim 3 \times 10^{-7}$  e viene legittimamente trascurata. I barioni però risentono degli urti elettrone-fotone. Difatti, sebbene la materia sia ionizzata, elettroni e nuclei carichi debbono essere distribuiti allo stesso modo, a evitare la formazione di fluttuazioni di carica elettrica. Le modeste discrepanze possibili sono quelle che danno un'energia potenziale dell'ordine dell'energia cinetica media delle particelle  $\sim k_B T$ .

Detta  $d$  la distanza media tra particelle, la discrepanza tra la posizione cui corrisponde assoluta neutralità di carica e quella effettivamente occupata potrà al più essere  $\delta \simeq 10^{-5} d$ .

**Verifica.** La distanza tra barioni (o elettroni), oggi  $d_o \sim 10^2 \text{cm}$ , a redshift  $z$  è  $d \sim d_o/(1+z)$ . Tra cariche  $e$ , a distanza  $\delta$ , si ha invece energia potenziale

$$E_p = e^2/\delta \sim 0.4 \sigma_T^{1/2} m_e c^2 / \delta \sim k_B T \quad \text{sicché risulta:}$$

$$\delta \sim 0.4 \sigma_T^{1/2} \times (m_e c^2 / k_B T) \simeq 3 \times 10^{-13} \text{cm} \times 2 \times 10^9 / (1+z) \simeq 6 \times 10^{-4} \text{cm} / (1+z).$$

Indipendentemente da  $z$ , finché il cosmo è ionizzato, sarà quindi  $\delta/d \sim 10^{-5}$ .

(Anche campi di velocità diversi, nelle componenti elettronica e barionica, richiederebbero l'energia necessaria a creare il conseguente campo magnetico.)

Pertanto, tutto avviene come se le particelle che subiscono urti Thomson avessero la sezione d'urto dell'elettrone e l'inerzia del protone (o di una particella  $\alpha$ ).

Le particelle cariche, però, non risultano apprezzabilmente deviate dall'urto con un singolo fotone: l'impulso ceduto da un  $\gamma$ , a  $T \sim 5000-6000 \text{K}$ , è  $\sim 0.5 \text{eV} \sim 10^{-6} m_e c^2 \sim 10^{-9} m_b c^2$ . Tuttavia, ci sono  $n_\gamma/n_b \simeq 10^{9.26} (0.02/\Omega_{ob} h^2)$  (eq. (VII.15), capitolo VII) fotoni per protone e l'impulso ceduto coerentemente dai  $\gamma$  è confrontabile con quello delle particelle cariche, sicché inevitabilmente la loro distribuzione tende a ricalcare quella dei fotoni.

Discrepanze tra le due distribuzioni sono possibili su scale pari al cammino libero medio di un fotone ( $\lambda_\gamma$ ), in quanto  $\lambda_\gamma$  è la distanza sulla quale  $\sim n_\gamma/n_b$  fotoni urtano ciascuna particella carica.

Questa espressione mostra che, mentre le discrepanze tra distribuzioni di elettroni e protoni sono piccole, le discrepanze tra distribuzioni di fotoni e barioni (ed elettroni), a un *redshift*  $\sim 10^4\text{-}10^3$ , avvicinano ormai scale *astronomiche*.

Infatti, per il cammino-libero-medio di un  $\gamma$  vale anche l'espressione

$$(VIII.24) \quad \lambda_\gamma = \frac{m_b}{\sigma_T \rho_{ob}(1+z)^3 x} = \frac{4.38 \times 10^4}{\Omega_{ob} h^2 (1+z)^3 x} \text{ Mpc},$$

che corrisponde alla distanza comovente

$$(VIII.25) \quad \lambda_{\gamma,c} = 2.19 \left( \frac{0.02}{\Omega_{ob} h^2} \right) \left( \frac{10^3}{1+z} \right)^2 \frac{1}{x} \text{ Mpc}$$

Detto  $\tilde{\rho}_{o,cr} = 2.78 \times 10^{11} M_\odot / \text{Mpc}^3$ , in una sfera di raggio  $\lambda_{\gamma,c}$  si trova la massa

$$(VIII.26) \quad M_\gamma = \frac{4\pi}{3} \lambda_{\gamma,c}^3 \tilde{\rho}_{o,cr} \Omega_{om} h^2 = 5.3 \times 10^{10} \frac{\Omega_{om} h^2}{0.1} \left( \frac{0.02}{\Omega_{ob} h^2} \right)^3 \left( \frac{10^3}{1+z} \right)^6 \frac{1}{x^3} M_\odot$$

sicché, a un *redshift*  $\sim 2 \times 10^3$  (assumendo  $x = 1$ , ma siamo alla vigilia della ricombinazione), in una sfera di raggio  $\lambda_\gamma$  si trovano  $\sim 8 \times 10^8 M_\odot$ , la massa di una piccola galassia.

Tale massa è destinata a salire, all'avvicinarsi della ricombinazione, quando la ionizzazione  $x$  diviene  $< 1$ , fino a superare la massa delle maggiori galassie.

## 5. La ricombinazione

A temperature  $T > B_H$  ( $= 13.6 \text{ eV}$ : energia di ionizzazione di  $H$ ), i fotoni possono ionizzare l'idrogeno. Per l'elio occorrono energie un po' superiori. A temperature  $T > B_H$ , la componente barionica sarà quindi ionizzata, grazie anche al gran numero di  $\gamma$  ( $n_b/n_\gamma = 10^{9.26} 0.02 / \Omega_{ob} h^2$ ).

Si potrebbe poi pensare che i materiali barionici passino allo stato atomico (ricombinazione) non appena sia  $T < B_H$ ; ciò porterebbe a stimare il passaggio da plasma a gas a un *redshift*  $z_{rec}^{(0)} \simeq B_H/T_o \sim 5\text{-}6 \times 10^4$ . La ricombinazione, comunque, proprio per il gran numero di fotoni, che qui interviene in modo analogo a quanto visto per la BBN, inizia alla temperatura  $T_{rec} \sim 4000 \text{ K} \sim 0.3 \text{ eV}$  e *redshift*  $z_{rec} \simeq 1000\text{-}1500$ , un fattore  $\sim 40$  sotto  $z_{rec}^{(0)}$ .

In effetti, la reazione



può avvenire da sinistra a destra o viceversa e sarà sostanzialmente in equilibrio quando  $n_\gamma \times n_H = n_p \times n_e$ . Essendo  $n_p \simeq n_e$ , sinché è  $T \sim B_H$  ci saranno oltre un miliardo di protoni liberi per ogni atomo di  $H$ .

Perché si formi  $H$  atomico, devono quindi restare pochi fotoni su  $\sim 10^9$  con energia  $> B_H$ . Conformemente a quanto ricavato in appendice A2 del capitolo sulla BBN, a

$T \ll B_H$ , la frazione di  $\gamma$  con energia  $> B_d$  è

$$(VIII.27) \quad f(> B_H) \simeq \frac{1}{2\zeta(3)} \left[ \left( \frac{B_H}{T} \right) + 2 \left( \frac{B_H}{T} \right) - 1 \right] e^{-B_H/T} \equiv \frac{P(B_H/T)}{2\zeta(3)} e^{-B_H/T}$$

ed ha un rapido decremento esponenziale, secondo la coda della planckiana. La temperatura  $T_{rec}^{(1)}$  a cui  $f(> B_H) = 1/2$  si ottiene risolvendo l'equazione

$$(VIII.28) \quad 1 \sim \frac{1}{\zeta(3)} \frac{n_\gamma}{n_b} P(B_H/T) e^{-B_H/T}.$$

Approssimativamente, perciò, la ricombinazione avverrebbe quando fosse

$$e^{-B_H/T_{rec}^{(1)}} \sim 10^{-10} \quad \text{e, quindi, alla temperatura } T_{rec}^{(1)} \sim B_H/23$$

e a un *redshift*  $z_{rec}^{(1)} \sim 2400$ .

Ci si avvicina ulteriormente a una valutazione precisa tramite l'equazione di Saha, che darà le abbondanze in equilibrio statistico di protoni, elettroni,  $\gamma$  e atomi. Il procedimento, analogo a quello descritto nel capitolo sulla BBN, porta all'equazione

$$(VIII.29) \quad \frac{n_p n_e}{n_\gamma n_H} \simeq \frac{\pi^{1/2}}{2\zeta(3)} \left( \frac{m_e}{2T} \right)^{\frac{3}{2}} \exp\left(-\frac{B_H}{T}\right).$$

Essendo  $n_p = n_e = x n_b$  e  $n_H = (1-x)n_b$ , l'espressione

$$(VIII.30) \quad \frac{x^2}{1-x} \simeq \frac{n_\gamma}{n_b} \frac{\pi^{1/2}}{2\zeta(3)} \left( \frac{m_e}{2T} \right)^{\frac{3}{2}} \exp\left(-\frac{B_d}{T}\right)$$

permette il confronto con l'eq. (VIII.28); infatti, se si pone  $x = 1/2$ , si ha

$$(VIII.31) \quad 1 \simeq \frac{\pi^{1/2} n_\gamma}{\zeta(3) n_b} \left( \frac{m_e}{2T} \right)^{\frac{3}{2}} \exp\left(-\frac{B_d}{T}\right).$$

e si vede che  $P(B_H/T) \sim (B_H/T)^2$  è qui sostituito da  $(m_e/T)^{3/2}$ . Si valuta quindi una temperatura di ricombinazione minore, di un fattore  $\sim 0.5$ .

L'espressione (VIII.30) fornisce poi l'andamento di  $x$  con  $T$  e, noto l'andamento di  $T$  con  $t$ , la legge  $x(t)$ . È anche opportuno riscriverla nella forma

$$(VIII.32) \quad \frac{x^2}{1-x} \simeq \frac{1}{n_b} \left( \frac{m_e T}{2\pi} \right)^{\frac{3}{2}} \exp\left(-\frac{B_d}{T}\right).$$

Tuttavia essa è ancora insufficiente. Oltre a trascurare i nuclidi diversi da  $H$ , compreso il 24% di  ${}^4He$ , essa trascura il fatto che le reazioni di ricombinazione possono divenire troppo lente per mantenere l'equilibrio statistico.

In sé ciò è inevitabile: la ricombinazione consiste nel calare di  $x$  e, quindi, del numero di elettroni liberi; ciò rende l'incontro protone-elettrone sempre meno probabile. A un certo punto, inevitabilmente, il cosmo diviene addirittura *trasparente* agli elet-

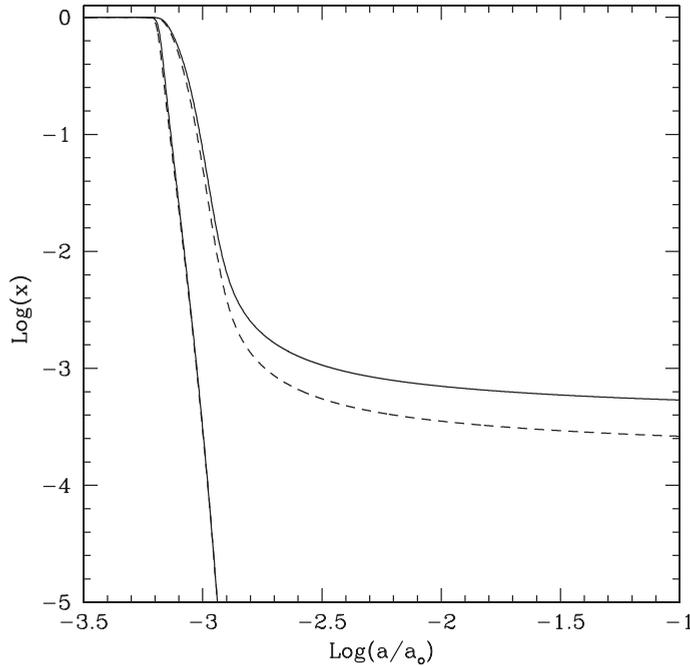


FIGURA VIII.4

Andamento della ionizzazione calcolata secondo la legge di Saha (equilibrio statistico) e tenendo conto del ritmo delle reazioni di ricombinazione. Le linee continue (tratteggiate) corrispondono a  $\Omega_b h^2 = 0.02$  (0.03). Le curve in rapida discesa sono calcolate secondo l'eq. di Saha.

troni; in altri termini, il loro c.l.m.  $\lambda_e$  raggiunge e supera l'orizzonte  $3ct$  e, quando ciò accade, il tasso di espansione,  $\simeq 2/3t$ , supera il tasso di ricombinazione  $\simeq c/\lambda_e$ .

Tuttavia, le reazioni di ricombinazione divengono troppo lente molto prima che si verifichi questa situazione «di stallo»: ben prima che la piena *trasparenza* sia raggiunta,  $x$  si mantiene su valori sempre più al di sopra di quelli dati da eq. (VIII.32).

Il tasso effettivo di ionizzazione si ricava allora risolvendo un'equazione differenziale, che tiene conto dei ritmi dei processi elementari e dell'espansione globale. Tale equazione, discussa in appendice A1, porta alla curva in Fig. VIII.4, dove si dà anche il confronto con la curva dedotta dall'equazione di Saha.

### La banda dell'ultimo scattering

Utilizzando la curva di ionizzazione si ricava la distribuzione sul *redshift* detta *banda dell'ultimo scattering*  $\nu(z)$ . Il sito in cui i fotoni del CMB hanno subito l'ultimo processo di diffusione, durante la ricombinazione, è distribuito su  $z$  secondo  $\nu(z)$ .

È chiaro che, via via che il cosmo si ricombina, sempre più fotoni riescono a arrivare all'osservatore, sito a  $z = 0$ , senza subire ulteriori processi di diffusione. Il loro numero

crebbe al calare di  $z$ , per raggiungere un massimo quando la metà dei fotoni può considerarsi liberata da ulteriori urti. Successivamente la distribuzione  $\nu(z)$  necessariamente cala, con l'esaurirsi del numero stesso dei fotoni che subiscono un urto a  $z$  via via inferiori.

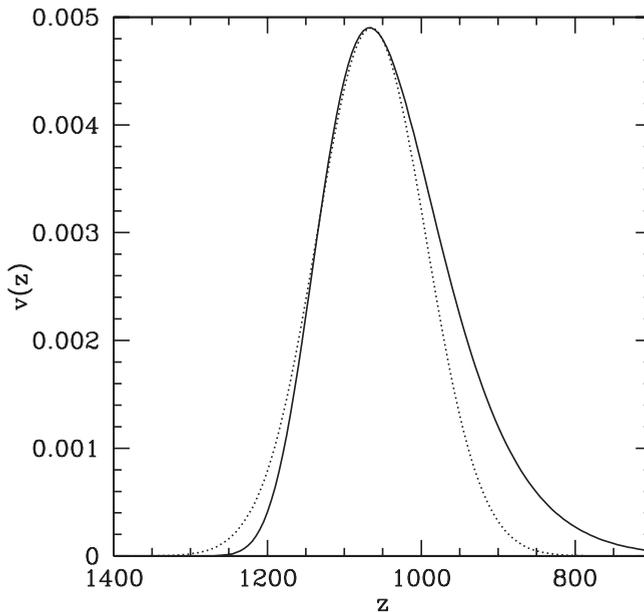
La forma della banda dell'ultimo scattering è studiata in appendice A2. Nella stessa appendice sono motivate alcune approssimazioni, sulla cui base si prova che l'andamento di  $\nu(z)$  è sostanzialmente indipendente dal modello cosmologico. Le motivazioni, espresse in appendice A2 per modelli spazialmente piatti, possono essere estese anche a modelli con densità totale non critica.

In figura VIII.5 viene mostrata la forma della banda dell'ultimo scattering. Essa è anche confrontata con la distribuzione Gaussiana

$$(VIII.33) \quad \mathcal{G}(z) = \mathcal{N} \exp \left[ - \left( \frac{z - 1060}{100} \right)^2 \right]$$

in cui  $\mathcal{N}$  è una costante di normalizzazione.

Si vede dunque che la ricombinazione avviene essenzialmente in una banda di profondità  $\Delta z \sim 200$ , centrata attorno a  $z = 1060$ ; e questo risultato è sostanzialmente indipendente dal modello cosmologico.



**FIGURA VIII.5**

Banda dell'ultimo scattering. La linea punteggiata rappresenta una distribuzione Gaussiana vicina all'andamento della banda stessa (vedi testo).

Nelle analisi più recenti delle anisotropie e della polarizzazione del CMB si è andati al di là di questi risultati. Il grado di precisione dei dati permette infatti ormai di apprezzare elementi più raffinati.

Un esempio è l'espressione del *redshift*  $z^*$  a cui l'opacità  $\tau_{op}$  raggiunge l'unità;  $z^*$  può essere definito *redshift della ricombinazione*. Con un'approssimazione  $\sim 1\%$  vale l'espressione

$$(VIII.34) \quad z^* = 1048 (1 + 1.24 \times 10^{-3} / \omega_b^{0.738}) (1 + b \omega_m^b)$$

dove 
$$a = \frac{0.56}{1 + 21.1 \omega_b^{1.81}}, \quad b = \frac{7.83 \times 10^{-2}}{\omega_b^{0.238} (1 + 39.5 \omega_b^{0.736})}$$

con  $\omega_{b,m} = \Omega_{b,m} h^2$ . Un esame di queste espressioni mostra tuttavia quanto modesta sia la residua dipendenza dai parametri cosmologici.

## 6. Natura delle fluttuazioni di densità

Una sostanza può essere trattata come *fluido* su scale molto maggiori dei cammini liberi medi (c.l.m.) delle particelle che la compongono. In cosmologia può accadere che i c.l.m. superino l'orizzonte e, allora, la sostanza *non* è un fluido.

All'avvicinarsi della transizione da plasma a gas,  $\lambda_\gamma$  cresce e la sostanza barioni-fotoni cessa di essere un fluido su scale via via crescenti. Sinché, a un *redshift*  $z^*$ , il c.l.m. dei fotoni supera l'orizzonte; la radiazione elettromagnetica allora si *disaccoppia*, com'è avvenuto ai neutrini a  $T_\nu \simeq 900$  keV.

Se ci poniamo a *redshift*  $z \gg z^*$  e su scale  $\gg \lambda_\gamma$ , la sostanza barioni-fotoni è un fluido. In essa possono allora sussistere due tipi di disomogeneità, caratterizzate da diverso rapporto tra le fluttuazioni  $\varepsilon_b = \delta\rho_b/\rho_b$  e  $\varepsilon_\gamma = \delta\rho_\gamma/\rho_\gamma$ , nei barioni e nella radiazione:

- 1) Fluttuazioni  $\varepsilon_b = \delta\rho_b/\rho_b$ ,  $\varepsilon_\gamma = \delta\rho_\gamma/\rho_\gamma$ , in cui il rapporto  $n_\gamma/n_b$  invariato rispetto al valore medio (fluttuazioni *adiabatiche*).
- 2) Fluttuazioni  $\varepsilon_b = \delta\rho_b/\rho_b$ ,  $\varepsilon_\gamma = \delta\rho_\gamma/\rho_\gamma$ , in cui il rapporto  $n_\gamma/n_b$  assume un valore costante, ma diverso dal valore medio (fluttuazioni *isoterme* o di *isocurvatura*).

In entrambi i tipi di fluttuazione, dunque, il rapporto  $n_\gamma/n_b$  non varia nel tempo, conformemente a quanto discusso nel paragrafo. 4.

### *Fluttuazioni adiabatiche*

L'adiabaticità è una caratteristica generica delle onde di compressione e le fluttuazioni di tipo 1) si comportano, in genere, come onde di compressione. Esse sono adiabatiche perché  $n_\gamma = [30\zeta(3)/\pi^4]\sigma$  ( $\sigma$ : densità d'entropia, vedi capitolo IV) e la costanza del rapporto  $n_\gamma/n_b$  assicura quindi la costanza di  $\sigma/n_b$ : nelle fluttuazioni adiabatiche l'*entropia-per-barione* permane costante. È facile vedere allora che l'invarianza nel

rapporto  $n_\gamma/n_b$  implica che sia

$$(VIII.35) \quad \varepsilon_b = (3/4) \varepsilon_\gamma.$$

**Verifica.** Densità in numero e densità d'energia dei fotoni sono (capitolo IV)

$$n_\gamma = [2\zeta(3)/\pi^2]T^3 \quad e \rho_\gamma = (\pi^2/15)T^4.$$

Quindi:

$$\frac{\delta n_\gamma}{n_\gamma} = 3 \frac{\delta T}{T}, \quad \frac{\delta \rho_\gamma}{\rho_\gamma} = 4 \frac{\delta T}{T} \quad \text{e, allora:} \quad \frac{3}{4} \varepsilon_\gamma = \frac{3}{4} \frac{\delta \rho_\gamma}{\rho_\gamma} = \frac{\delta n_\gamma}{n_\gamma}.$$

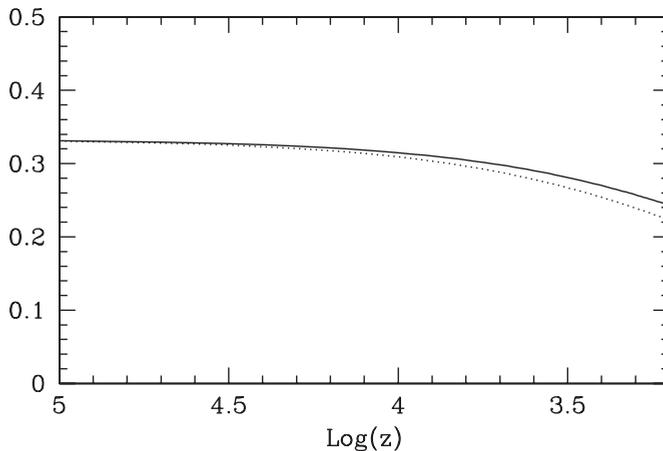
Per i barioni, invece, è  $\rho_b = m_b n_b$  e quindi  $\varepsilon_b = \delta n_b/n_b$ . Sarà allora

$$\varepsilon_b = \delta n_b/n_b = \delta n_\gamma/n_\gamma = (3/4) \varepsilon_\gamma.$$

Le fluttuazioni adiabatiche si propagano come onde di compressione (o *soniche*). Pur coinvolgendo sia la componente fotonica che quella barionica, esse hanno intrinsecamente natura radiativa. I barioni sono però importanti, in quanto forniscono un tramite per l'interazione  $\gamma$ - $\gamma$ . Infatti il *principio di sovrapposizione* esclude ogni interazione diretta  $\gamma$ - $\gamma$  (in elettrodinamica quantistica, la sezione d'urto  $\gamma$ - $\gamma$  non è nulla, ma si tratta comunque di un processo di quart'ordine, dinamicamente trascurabile) e, senza il tramite dei barioni, la pura sostanza fotonica sarebbe non-collisionale.

Per il fluido barioni-fotoni, la densità d'energia è  $\rho_b + \rho_\gamma$ , mentre la pressione, dovuta in pratica alla sola radiazione, è  $\rho_\gamma/3$ . La velocità del suono sarà allora

$$(VIII.36) \quad c_s^2 = c^2 \frac{\delta p}{\delta \rho} = \frac{c^2}{3} \frac{\varepsilon_\gamma \rho_\gamma}{\varepsilon_\gamma \rho_\gamma + \varepsilon_b \rho_b} = \frac{1}{3} \frac{1}{1 + 3\rho_b/4\rho_\gamma}$$



**FIGURA VIII.6**

Andamento di  $c_s^2$  (linea continua) e  $w$  (linea tratteggiata) vs. *redshift*, fino alla ricombinazione, in un modello con  $\Omega_{ob} = 0.04$ ,  $\Omega_{oc} = 0.026$  e  $h = 0.7$ .

(vedi eq. VIII.35) con

$$(VIII.37) \quad \frac{\rho_b}{\rho_\gamma} = \frac{\rho_b \rho_m}{\rho_m \rho_\gamma} = \frac{\Omega_{ob} a}{\Omega_{om} a_{eq}} \simeq \frac{800}{1+z} \frac{\Omega_{ob} h^2}{0.02}.$$

A  $z \simeq 10^3$  sarà dunque  $(3/4)\rho_b/\rho_\gamma \simeq 0.6$ . Ma allora la transizione plasma-gas è ormai in fase avanzata e l'insieme barioni-fotoni non è più un fluido.

In figura VIII.6 è riportato l'andamento di  $(c_s/c)^2$  e quello di  $w = p_\gamma/(\rho_\gamma + \rho_b)$ , parametro di stato per la miscela fotoni-barioni.

### Fluttuazioni isoterme o di isocurvatura

Accanto alle fluttuazioni adiabatiche, va considerata la possibilità di fluttuazioni in cui sia alterato il rapporto  $n_b/n_\gamma$ ; conformemente a quanto discusso nel Par.4, esse non possono essere generate da azioni fluidodinamiche, ma debbono essere primordiali. Ad esempio, potrebbero risalire a una diversa rottura della simmetria barione-antibarione nei diversi siti.

Considereremo anche un'altra possibilità, che non attiene il rapporto barioni/fotoni: fluttuazioni primordiali nel valore del rapporto  $n_c/n_\gamma$  ( $n_c$ : numero di particelle di DM).

In tutte le fluttuazioni di questa natura, le fluttuazioni nella componente fotonica,  $\varepsilon_\gamma$ , sono nulle o trascurabili. Procedendo verso alti *redshift*, quando  $\rho_\gamma \gg \rho_m$ , si avranno quindi fluttuazioni nella densità d'energia tendenti a zero. Non ci sono quindi campi gravitazionali ovvero, in termini relativistici, non vi sono fluttuazioni nella curvatura spazio-temporale.

Un valore invariante di  $n_\gamma = [2\zeta(3)/\pi^2]T^3$  implica una temperatura costante e ciò spiega la denominazione di *fluttuazioni isoterme*. L'assenza di curvatura, al limite  $z \rightarrow \infty$ , spiega la denominazione di *fluttuazioni di isocurvatura*.

Un rapporto  $n_c/n_\gamma$  o  $n_b/n_\gamma$  alterato rispetto alla media, fuori dall'orizzonte, non implica ovviamente alcuna azione dinamica. All'entrata nell'orizzonte, le particelle di DM, non relativistiche e quindi in moto a velocità  $v_c \ll c$ , non possono coprire che distanze  $\ll ct$  e permangono quindi entro la fluttuazione, le cui caratteristiche permangono invariate fino all'equivalenza.

All'avvicinarsi dell'equivalenza, tuttavia, una fluttuazione  $\varepsilon_c$  comincia a dar luogo a un significativo campo gravitazionale e, dopo l'equivalenza, queste fluttuazioni nella DM sarebbero destinate a crescere per autogravità.

Anche fluttuazioni primordiali  $\varepsilon_b \neq 0$  con  $\varepsilon_\gamma = 0$ , vale a dire fluttuazioni nel rapporto  $n_b/n_\gamma$ , permangono sostanzialmente invariate, cioè senza che ne nasca una significativa fluttuazione  $\varepsilon_\gamma$ , anche dopo l'entrata nell'orizzonte.

Ciò è conseguenza dell'elevata pressione della radiazione, a confronto della piccolissima pressione del mezzo barionico, per cui  $w_b = p_b/\rho_b \ll w_\gamma = p_\gamma/\rho_\gamma$ . Basta quindi una piccolissima fluttuazione «negativa» nei fotoni a generare una pressione che bilancia la pressione «positiva» nei barioni: Se è  $\varepsilon_\gamma \rho_\gamma = -(w_b/w_\gamma)\varepsilon_b \rho_b$ , sarà

$$(VIII.38) \quad \delta p_\gamma = w_\gamma \delta \rho_\gamma = -w_b \delta \rho_b = -\delta p_b$$

e, essendo complessivamente  $\delta p = 0$ , non ci sarà alcuna azione fluidodinamica tendente a modificare la distribuzione delle particelle.

Fluttuazioni adiabatiche e isoterme costituiscono i possibili *modi di fluttuazione*. Ogni altro tipo di fluttuazione può essere sempre decomposto in una combinazione dei due modi, cui corrisponde una precisa evoluzione dinamica.

Nei modelli di formazione delle strutture cosmiche, le fluttuazioni di isocurvatura assumono in genere un ruolo secondario. Limiti severi al loro contributo possono essere ricavati da vari osservabili, in particolare dalle anisotropie del CMB. Nel prosieguo di questo capitolo esse non verranno trattate.

## 7. Spettri delle fluttuazioni

Le fluttuazioni di densità  $\varepsilon(\mathbf{x})$  in una determinata componente, ad esempio nella componente barionica, si possono utilmente considerare come somma di segnali armonici sovrapposti

$$(VIII.39) \quad \varepsilon_{\mathbf{k}}(\mathbf{x}) = \delta(k)e^{i\mathbf{k}\cdot\mathbf{r}},$$

come abbiamo già visto più volte. Le componenti armoniche sono esattamente come le note musicali semplici, che si sovrappongono per dare un suono. Tramite le ampiezze  $\delta(k)$  si costruisce lo spettro delle fluttuazioni

$$(VIII.40) \quad P(k) = \langle |\delta(k)|^2 \rangle;$$

in linea di principio, la media dovrebbe essere fatta tra molte realizzazioni indipendenti. Se, prima di entrare nell'orizzonte, nessuna scala  $\mathbf{k}$  è *privilegiata*, mentre la distribuzione è isotropa, è ragionevole assumere che l'ampiezza dello spettro dipenda dal modulo di  $\mathbf{k}$  con una legge di potenza:

$$(VIII.41) \quad P(k) = A k^{n_s} \quad (\text{con } A, n_s \text{ costanti}).$$

Ne conseguono due fatti che avranno varie conseguenze:

1) Al momento dell'entrata nell'orizzonte sarà

$$(VIII.42) \quad P_{hor}(k) = \langle |\delta_{hor}(k)|^2 \rangle \propto k^{n_s-4}$$

2) La varianza di massa sulla scala dell'orizzonte sarà

$$(VIII.43) \quad \sigma_{R_{hor}}^2 \propto R_{hor}^{1-n_s}$$

Quindi, per  $n_s = 1$ , la varianza di massa sull'orizzonte,  $\sigma_{R_{hor}}$ , è *indipendente dalla scala*.

Lo spettro  $P(k)$  con  $n_s = 1$  viene detto *spettro di Harrison-Zel'dovich*.

**Verifica.** Nel precedente capitolo abbiamo mostrato che la crescita delle fluttuazioni (fuori dall'orizzonte e in una gauge sincrona) avviene con leggi diverse prima e dopo

l'equivalenza:

$$(VIII.44) \quad \delta = \delta_i(a/a_i)^2 \quad \text{per } z \gg z_{eq}, \quad \delta = \delta_i(a/a_i) \quad \text{per } z \ll z_{eq}$$

(l'indice  $i$  denota un determinato istante «iniziale»). Quindi, sull'orizzonte sarà

$$(VIII.45) \quad \begin{aligned} \delta_{hor} &\simeq \delta_i(a_{hor}/a_i)^2 && \propto \delta_i a_{hor}^2 && \text{per } z \gg z_{eq}, \\ \delta_{hor} &\simeq \delta_i(a_{eq}/a_i)^2 (a_{hor}/a_{eq}) && \propto \delta_i a_{hor} && \text{per } z \ll z_{eq}. \end{aligned}$$

Al Par. 3 di questo capitolo si è invece mostrato che

$$(VIII.46) \quad \begin{aligned} k_{hor}(a) &\propto a^{-1} && \text{per } z \gg z_{eq}, && k_{hor}(a) \propto a^{-1/2} && \text{per } z \ll z_{eq}, \\ \text{e quindi: } a_{hor}(k) &\propto k^{-1} && \text{per } z \gg z_{eq}, && a_{hor}(k) \propto k^{-2} && \text{per } z \ll z_{eq}. \end{aligned}$$

Ponendo assieme le eq. (VIII.45) e (VIII.46), si ottiene

$$(VIII.47) \quad \begin{aligned} \delta_{hor}(k) &\propto \delta_i(k) a_{hor}^2(k) && \propto \delta_i(k) k^{-2} && \text{per } z \gg z_{eq}, \\ \delta_{hor}(k) &\propto \delta_i(k) a_{hor}(k) && \propto \delta_i(k) k^{-2} && \text{per } z \ll z_{eq}; \end{aligned}$$

risultato eguale, quindi, sia che la scala sia maggiore o minore dell'orizzonte all'equivalenza. Tramite le eq. (VIII.40) e (VIII.41) si ricava allora l'espressione (VIII.42):

$$(VIII.48) \quad |\delta_{hor}(k)|^2 = |\delta_i(k)|^2 k^{-4} = A k^{n_s-4}$$

Il passaggio successivo richiede inevitabilmente il calcolo integrale, vista l'espressione della varianza di massa (capitolo II):

$$(VIII.49) \quad \sigma_{R_{hor}}^2 = \frac{1}{2\pi^2} \int_0^\infty dk k^2 P_{hor}(k) W(kR_{hor}).$$

Sostituendo lo spettro (VIII.48) e con un cambio di variabile si ha allora:

$$(VIII.50) \quad \sigma_{R_{hor}}^2 = \frac{A}{2\pi^2} \int_0^\infty dk k^{n_s-2} W(kR_{hor}) = R_{hor}^{1-n_s} \frac{A}{2\pi^2} \int_0^\infty dx x^{n_s-2} W(x).$$

Non occorre dare il valore dell'integrale numerico, che dipende comunque dalla forma del «filtro»  $W(x)$ , per dedurne che  $\sigma_R^2 \propto R^{1-n_s}$ , conformemente a eq. (VIII.43).

## 8. Instabilità di Jeans fino alla ricombinazione

La scala di instabilità gravitazionale, in modelli a più componenti, viene ricavata in appendice A3, assumendo la presenza di barioni e radiazione accoppiati a formare un singolo fluido (velocità del suono  $c_s = f_s c$ ) e di DM. L'espressione

(A3.9), qui riportata,

$$(VIII.51) \quad \lambda_J = c_s \sqrt{\frac{\pi}{G\rho}} \sqrt{\frac{1}{(1+w)[\Omega_c + (1+3f_s^2)\Omega_{by}]}}$$

trascura invece i neutrini. Sulla base dei dati recenti sulle loro masse, sappiamo che in quest'epoca, essi sono ultrarelativistici e non collisionali; fuoriescono quindi liberamente da qualsiasi fluttuazione al di sotto dell'orizzonte e l'espressione (VIII.51) non risente significativamente della loro presenza.

A  $z \gg z_{eq}$ , l'espressione (VIII.51) si semplifica, essendo  $\Omega_c \simeq 0$ ,  $\Omega_{by} \simeq 1$ ,  $f_s^2 \simeq 1/3$ ,  $1+w \simeq 4/3$ ; avremo perciò

$$(VIII.52) \quad \lambda_J = c_s \pi \sqrt{3/8\pi G\rho} = \pi \times 2c_s t = \pi f_s L_{hor}$$

e, numericamente,  $\lambda_J \simeq 1.81 L_h$ . Notiamo che, a  $z < z_{eq}$  sarà

$$(VIII.53) \quad \lambda_J = \pi f_s L_{hor} \times \sqrt{\frac{2}{1+3f_s^2\Omega_{by}}} > 1.81 L_{hor},$$

in quanto  $3f_s^2 < 1$  e  $\Omega_{by} < 1$ ; quindi *tutte le scale al di sotto dell'orizzonte sono comunque in regime sonico*.

La quantità

$$(VIII.54) \quad L_{shor} = f_s L_{hor}$$

si ottiene come l'orizzonte di particella, per un segnale che viaggia alla velocità del suono, invece della velocità della luce. Essa viene perciò detta *orizzonte acustico* o sonico.

## 9. Dinamica delle fluttuazioni tra orizzonte e ricombinazione. Componente oscura e plasma

L'evoluzione lineare delle fluttuazioni nelle ere intermedie determina una gran parte delle proprietà di anisotropia e polarizzazione del CMB, espresse tramite gli *spettri angolari*. Essa determina inoltre il *coefficiente di trasferimento* o di trasmissione per le fluttuazioni nella materia, da cui si ottiene lo spettro delle fluttuazioni dopo la ricombinazione, detto *spettro trasmesso*.

La coerenza tra i risultati riguardando il CMB e quelli sulla formazione delle strutture, dedotti dallo spettro trasmesso, è uno dei maggiori successi della cosmologia. È bene quindi rimarcare alcune differenze di base tra i due risultati.

1) *Essi attengono, rispettivamente, dati lineari o non-lineari.*

Le fluttuazioni osservate nel CMB hanno ampiezza  $\sim 1 : 10^5$ . I sistemi cosmici, su cui si verifica l'andamento del coefficiente di trasferimento, hanno contrasti di densità  $\gtrsim 100$ , rispetto alla densità media. In effetti, mentre l'analisi lineare (in verità estesa

al di là delle ere intermedie) porta direttamente alla previsione degli spettri angolari, la conoscenza della funzione di trasmissione è solo un ingrediente, anche se molto importante, per ottenere previsioni sulle strutture non lineari. Molte delle loro caratteristiche osservative si riconnettono alla genesi «cosmologica», sicché la loro osservazione permette di porre vincoli al modello, ma molte altre sono generiche e non portano il marchio di una particolare cosmologia.

2) *Essi attengono per lo più scale diverse.*

La scala comovente che entra nell'orizzonte all'equivalenza, in una cosmologia  $\Lambda$ CDM è  $\sim 75 h^{-1}$  Mpc e contiene una massa  $\sim 1.6 \times 10^{17} h^{-1} M_{\odot}$ . I grandi ammassi di galassie hanno una massa inferiore di 2 ordini di grandezza, che corrisponde a una scala comovente  $\sim 15 h^{-1}$  Mpc. Le scale delle strutture osservate entrano infatti nell'orizzonte prima dell'equivalenza.

Nel paragrafo 2 abbiamo definito il numero d'onda angolare  $l$ , in funzione del quale vengono forniti gli spettri angolari del CMB. Sulla base delle relazioni in paragrafo 2 e paragrafo 3 è facile dedurre che, in un modello SCDM, il valore di  $l$  corrispondente alla scala che entra nell'orizzonte a  $z \sim 10^3$  (ricombinazione) è

$$(VIII.55) \quad l = \frac{R_H 2\pi}{L_{hor,c}(z = 10^3)} = \frac{6000 h^{-1} \text{Mpc} \times 2\pi}{6000 10^{-3/2} h^{-1} \text{Mpc}} = 2\pi 10^{3/2} \simeq 200.$$

Ciò vale anche in una cosmologia  $\Lambda$ CDM. Essendo poi, dopo l'equivalenza,  $L_{hor,c} \propto (1+z)^{1/2}$ , è facile dedurre che a  $z \sim 10^4$ , cioè all'incirca all'equivalenza, corrisponde  $l \simeq 200 \times \sqrt{10} \simeq 600$ . Siccome poi le maggiori strutture corrispondono a un fattore  $\sim 100^{1/3}$  in lunghezza (100 in massa) in meno, il valore di  $l$  che corrisponde alle strutture maggiori sarà  $\sim 2700$ .

I dati sulle anisotropie e polarizzazione del CMB, per  $l > 2000$ , sono tuttora incerti e corrispondono comunque a ampiezze di fluttuazione  $< 1 : 10^5$ . La connessione tra i dati CMB e quelli sulla struttura su grande scala è quindi in gran parte legata all'univocità dello spettro.

3) *Essi derivano da componenti cosmiche diverse.*

Le strutture nascono dall'evoluzione delle fluttuazioni nel DM,  $\delta_c$ , con qualche sovrapposta influenza della piccola componente barionica. Le anisotropie del CMB, invece, nascono dalle fluttuazioni  $\delta$  nel plasma barioni-fotoni e specialmente dalla loro componente fotonica.

Fino all'entrata nell'orizzonte,  $\delta$  e  $\delta_c$  evolvono parallelamente. Entrate nell'orizzonte, le fluttuazioni nel fluido e nella DM interagiscono ancora gravitazionalmente, ma il fluido risente pure di forze pressorie, mentre la DM ha un comportamento legato ancora alla sola gravità.

In appendice A3 viene analizzato, in forma semplificata (essenzialmente assumendo che la velocità del suono sia costante), l'andamento delle fluttuazioni adiabatiche, nella DM e nel plasma barioni/fotoni. Una trattazione più completa del problema si trova nella

letteratura specialistica e richiede l'analisi di un sistema di molte equazioni differenziali accoppiate. Le approssimazioni di appendice A3 sono peraltro sufficienti a prevedere molte delle caratteristiche osservative e i loro risultati verranno qui riportati.

### **Funzione di trasferimento: definizione**

L'andamento del modo crescente delle fluttuazioni si può applicare al limite  $k \rightarrow 0$ ; per la DM, ad esempio, vale l'espressione ben approssimata

$$(VIII.56) \quad \delta_c \simeq \delta_{c,i} \left( \frac{t/a}{t_i/a_i} \right)^2$$

$$\begin{aligned} \text{che diviene:} \quad \delta_c &= \delta_{c,i} (a/a_i)^2 = \delta_{c,i} (t/t_i) \quad , \quad \text{per } z \gg z_{eq}; \\ \delta_c &= \delta_{c,i} (a/a_i) = \delta_{c,i} (t/t_i)^{2/3} \quad , \quad \text{per } z \ll z_{eq}; \end{aligned}$$

$a_i$  è il fattore di scala al tempo  $t_i$ , quando l'ampiezza della fluttuazione è  $\delta_{c,i}$ . Complessivamente, quindi, possiamo definire un *fattore di crescita lineare* delle fluttuazioni al limite  $k \rightarrow 0$  (in una gauge sincrona), che si può approssimare

$$(VIII.57) \quad \mathcal{L}(t_{in}, t_{out}) = \frac{\delta_{out}}{\delta_{in}} \simeq \left( \frac{t_{out}}{t_{eq}} \right)^{\frac{2}{3}} \frac{t_{eq}}{t_{in}} = \frac{a_{out}}{a_{eq}} \left( \frac{a_{eq}}{a_{in}} \right)^2,$$

se non si richiede una particolare precisione per il comportamento attorno all'equivalenza.

La crescita delle fluttuazioni, sulle scale rilevanti, sarà minore (o uguale) a  $\mathcal{L}$ . Essa dipenderà innanzitutto dal modo di fluttuazione. In quanto segue ci riferiremo di norma al modo adiabatico. Si tiene conto della crescita effettiva nel modo considerato definendo la *funzione di trasferimento*

$$(VIII.58) \quad \mathcal{T}(k) = \frac{\delta_{out}(k)/\delta_{in}(k)}{\mathcal{L}(t_{in}, t_{out})} \simeq \frac{\delta_{out}(k)/\delta_{in}(k)}{(a_{out}/a_{eq})(a_{eq}/a_{in})^2},$$

tra i tempi  $t_{in}$  e  $t_{out}$ ;  $t_{in}$  sarà comunque molto prima dell'entrata nell'orizzonte delle scale rilevanti;  $t_{out}$  sarà scelto opportunamente, in genere dopo la ricombinazione, anche coincidente col tempo attuale  $t_0$ . Le fluttuazioni  $\delta_{in}$  e  $\delta_{out}$  riguarderanno in genere la DM, i barioni, o la materia nel suo complesso, ai tempi  $t_{in}$  e  $t_{out}$  rispettivamente. Nel definire la funzione di trasferimento, quindi, si deve precisare a quale/i componente/i ci si riferisce.

### **L'effetto di stagnazione.**

Vediamo ora l'elemento caratteristico di tutte le funzioni di trasferimento, la *stagnazione* o *effetto Meszaros*.

Esso riguarda la legge di crescita delle fluttuazioni della CDM su tutte le scale raggiunte dall'orizzonte prima dell'equivalenza ( $k > k_{eq}$ ). Su queste scale, tra l'entrata nell'orizzonte e l'equivalenza, l'ampiezza delle fluttuazioni nel CDM,  $\delta_c$ , cresce po-

chissimo. Ciò è dovuto al fatto che, nel contempo, le fluttuazioni  $\delta$ , nella componente radiativa, oscillando come onde soniche, hanno ampiezza mediamente nulla. Non c'è quindi una sorgente efficace di *autogravitazione*: da un lato,  $\langle \delta \rangle$  è nullo; dall'altro, la CDM, prima dell'equivalenza, ha densità  $\ll \rho_{\gamma b}$ .

Più specificamente, in appendice A2 si ricava la legge di crescita

$$(VIII.59) \quad \delta_c = \delta_{hor} [\ln(t/t_{hor}) + 1].$$

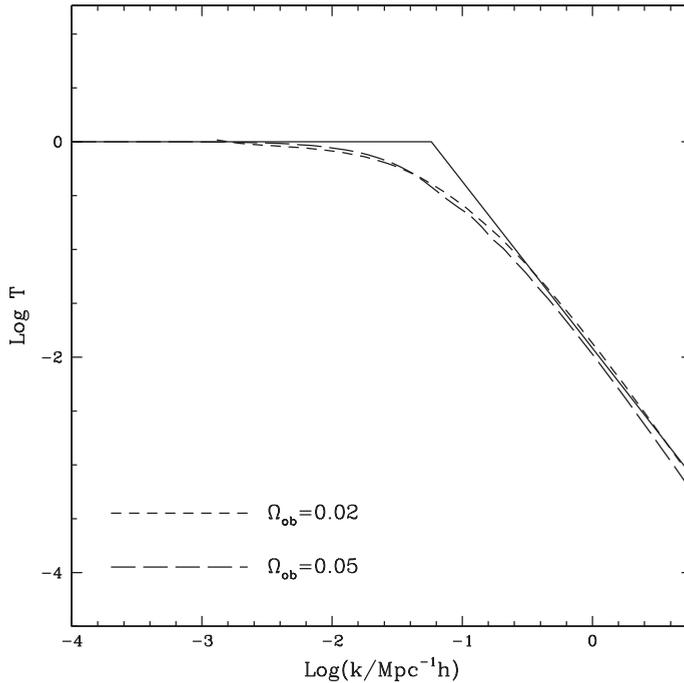
Qui  $\delta_{hor}$  è l'ampiezza della fluttuazione all'orizzonte, raggiunto al tempo  $t_{hor}$ . Questa legge, in sé approssimata, è riconducibile, per le scale rilevanti, alla legge di potenza

$$(VIII.60) \quad \delta_c = \delta_{hor} [t/t_{hor}(k)]^u \simeq \delta_{hor} [a/a_{hor}(k)]^{2u} \quad \text{con } u = 0.25 - 0.30$$

Tenendo conto del solo effetto Meszaros, vediamo che deve essere

$$(VIII.61) \quad \mathcal{T}(k) \simeq \frac{[a_{out}/a_{eq}][a_{eq}/a_{hor}(k)]^{2u}[a_{hor}(k)/a_{in}]^2}{[a_{out}/a_{eq}][a_{eq}/a_{in}]^2} = \left[ \frac{a_{hor}(k)}{a_{eq}} \right]^{2(1-u)} \quad (\text{per } k > k_{eq}).$$

Evidentemente, quando  $a_{hor}(k) = a_{eq}$ , cioè per scale  $k < k_{eq}$ , sarà  $\mathcal{T} = 1$ .



**FIGURA VIII.7**

Funzione di trasmissione in modelli con CDM; l'andamento previsto analiticamente, trascurando i barioni e modellando l'effetto Meszaros, è confrontato coi risultati ottenuti tramite un codice numerico.

Trascurando quindi completamente quanto accade alla componente barionica, la dipendenza del coefficiente di trasferimento da  $k$  è tutta in  $a_{hor}(k)$ . Essendo allora  $c t_{hor}(k) \propto a_{hor}(k)/k$  (e  $t \propto a^2$ ) sarà  $a_{hor}(k) \propto k^{-1}$ , sicché risulta

$$(VIII.62) \quad \mathcal{T}(k) \simeq (k_{eq}/k)^{2(1-u)} \quad \text{per } k > k_{eq}$$

In figura VIII.7 questo andamento del coefficiente di trasferimento è confrontato con quello ottenuto numericamente per modelli con  $\Omega_{om} = 0.3$  e  $h = 0.7$ , prendendo  $u = 0.27$ . Al di là della presenza di «punti angolosi», inevitabili visto il carattere schematico delle approssimazioni, la figura mostra come l'elemento base della funzione di trasferimento sia proprio l'effetto di stagnazione.

La valutazione precisa di  $\mathcal{T}(k)$  dovrà necessariamente tener conto delle altre componenti e, in particolare, dell'evoluzione in prossimità di  $a_{hor}$  e  $a_{eq}$ , non facile da approssimare analiticamente.

### ***Evoluzione delle fluttuazioni adiabatiche nel plasma.***

Lo studio delle fluttuazioni nel plasma barioni/radiazione è utile a sia a questo scopo, che a capire la formazione delle anisotropie del CMB.

Le fluttuazioni adiabatiche nel plasma evolvono assieme a quelle nella materia oscura, fin tanto che la scala  $\lambda = 2\pi a/k$  supera l'orizzonte acustico. Quando l'orizzonte supera  $\lambda$ , nel plasma si hanno onde di compressione, o *soniche*. Finché poi  $\lambda \gg \lambda_\gamma$ , le onde soniche sono ben approssimate tramite l'espressione

$$(VIII.63) \quad \delta(t, k) = \delta_{hor}(k) \cos \left[ 2\pi \left( \frac{vc_s t}{\lambda} - f_s \right) \right] + \bar{\delta}(k)$$

(vedi appendice A3) valida a partire da quando l'argomento del coseno diviene positivo; questa espressione e le seguenti sono ricavate sotto l'ipotesi di poter trascurare le variazioni nel tempo di  $c_s$ . In questa espressione sarà:

$$\begin{aligned} \delta_{hor}(k) &: \text{ampiezza flutt. all'orizzonte} \\ \nu &: \sim 2 \text{ o } 3 \text{ nell'espansione dominata da rad. o mat.,} \\ &\text{da interpolare al passaggio tra regimi espansivi} \\ f_s = c_s/c & \\ \lambda = 2\pi a/k &: \text{lunghezza d'onda (variabile)} \end{aligned}$$

$$(VIII.64) \quad \begin{aligned} \text{mentre, per } z \gtrsim z_{eq} &: \bar{\delta}(k) = -[8\nu^2/27(2\pi)^2](\lambda/vc_s t)^2 \delta_c \\ \text{per } z \lesssim z_{eq} &: \bar{\delta}(k) \equiv 0 \\ &\text{andamenti da interpolare attorno } z_{eq} \end{aligned}$$

L'argomento del coseno è quello che ci si attende per una normale onda sonica, a parte il fattore  $\nu$ . Esso deriva dal fatto che la lunghezza d'onda *non* è costante. In effetti l'onda *non* è periodica rispetto al tempo ordinario, ma lo è, con buona approssimazione, rispetto al tempo conforme.

Un'espressione alternativa a eq. (VIII.63) è la seguente:

$$(VIII.65) \quad \delta(t, k) = \delta_{hor}(k) \cos(v c_s k t / a - 2\pi f_s) + \bar{\delta}(k).$$

Oltre che più semplice, essa mette in evidenza la reale dipendenza dal tempo, ma non ha l'aspetto immediato di un'onda sonora. In effetti, il rapporto  $t/a$ , durante l'espansione dominata dalla radiazione o dalla materia, è rispettivamente  $\propto t^{1/2}$  o  $t^{1/3}$ .

Nell'espressione (VIII.64) il rapporto  $\lambda/v c_s t$  può essere valutato a qualunque tempo, dopo l'equivalenza e l'entrata dell'onda nell'orizzonte sonico. Per le onde che entrano nell'orizzonte sonico dopo l'equivalenza, tale rapporto vale quindi 1. Per le onde che entrano nell'orizzonte prima dell'equivalenza, invece, sarà  $\lambda/v c_s t = k_{shor,eq}/k$ , dove  $k_{shor}$  è il numero d'onda corrispondente alla lunghezza dell'orizzonte acustico all'equivalenza.

Le espressioni (VIII.56) e (VIII.63) sono state utilizzate, ricavando  $\delta_{shor}$  in eq. (VIII.63) dalla stessa espressione (VIII.56), per tracciare l'evoluzione delle fluttuazioni nel plasma, su 3 diversi numeri d'onda, in figura VIII.8. Le ampiezze sono state assunte eguali all'orizzonte.

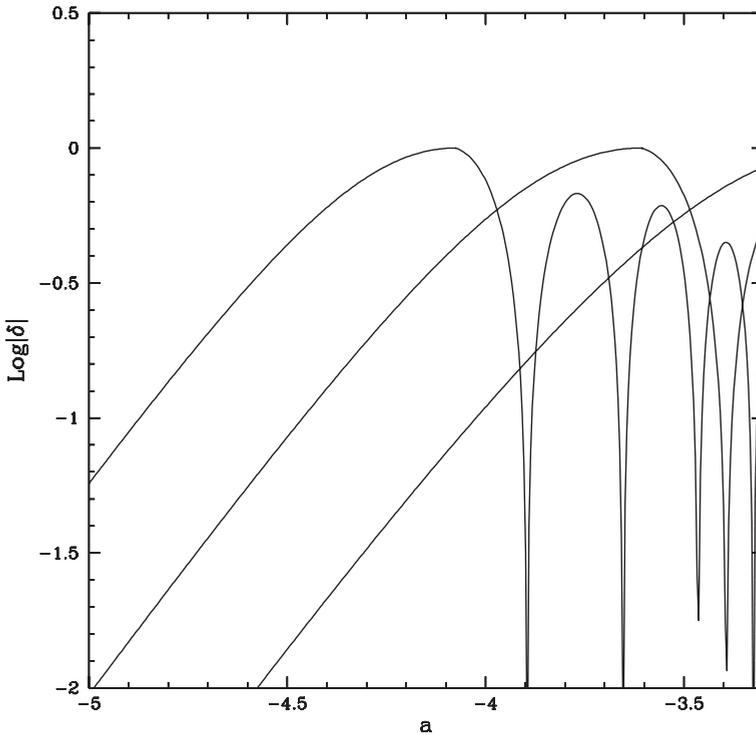


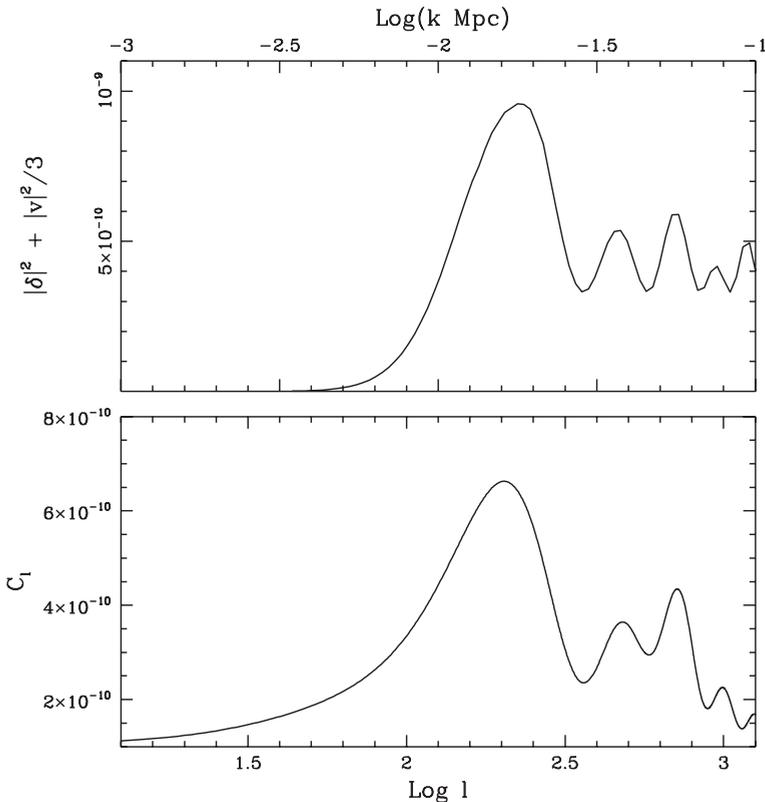
FIGURA VIII.8

Evoluzione delle fluttuazioni nel plasma su 3 diversi numeri d'onda. L'ampiezza delle fluttuazioni è normalizzata a 1 sull'orizzonte.

Tramite l'espressione (VIII.63), usata fino alla ricombinazione, è anche possibile fare una previsione qualitativa dello spettro delle anisotropie del CMB.

A questo scopo possiamo tracciare il grafico di  $|\delta|^2 + (1/3)|v|^2$ , alla ricombinazione, in funzione di  $k$ ; la *velocità*  $v$  è stata definita nel capitolo precedente ed è usata in appendice A3.

Questa espressione mira a valutare, sulla *banda dell'ultimo scattering* dopo la quale i fotoni proseguono (praticamente) indisturbati fino all'osservatore, le variazioni di densità e l'effetto Doppler dovuti alle onde soniche. È una procedura grossolana che dà però luogo ai risultati in figura VIII.9, in cui è riportato anche il confronto con lo spettro angolare delle anisotropie (vedi capitolo sul CMB) in funzione dei valori di  $l$  corrispondenti ai  $k$  considerati (paragrafo 2).



**FIGURA VIII.9**

Contributo delle onde soniche all'anisotropia del CMB, calcolato per via analitica assumendo arbitrariamente un'ampiezza  $\delta_m = 10^{-5}$  per le fluttuazioni sull'orizzonte (in alto), confrontato con lo spettro delle anisotropie del CMB, calcolato numericamente per lo stesso modello (in basso). Nel secondo grafico si tiene conto di svariati altri effetti, tra cui è particolarmente evidente lo smorzamento delle ampiezze a alti  $k$  (piccole scale), dovuto all'allungarsi del c.l.m. dei  $\gamma$  in prossimità della ricombinazione. Il confronto pone peraltro in evidenza il ruolo dominante delle onde soniche nel plasmare la forma delle anisotropie del CMB.

Si noti in figura VIII.9 la sequenza di picchi di altezza diversa. I picchi più alti (bassi) si hanno quando i due addendi a destra dell'espressione (VIII.63) hanno segno eguale (opposto). Fisicamente, le forze di compressione dovute al plasma e le forze gravitazionali dovute alla DM sommano i loro effetti se il numero delle fluttuazioni soniche, dall'entrata nell'orizzonte alla ricombinazione, è pari. Nel caso di numero dispari di fluttuazioni, invece, la fase alla ricombinazione risulta opposta a quella iniziale, all'orizzonte, e quindi la forza gravitazionale della DM, che ha mantenuto la fase iniziale, contrasta le forze di pressione; la conseguenza è un'ampiezza minore del massimo complessivo.

Richiamiamo l'attenzione del lettore sul fatto che il campo di gravità della DM, qui considerato, agisce sulle onde di densità, modificando la geografia dei picchi, e non direttamente sui fotoni che vengono emessi nei picchi o nei nodi. Anche questo secondo effetto esiste ed è importante su varie scale angolari, ma verrà considerato a parte nel capitolo XIII sul CMB.

Dal punto di vista analitico, l'argomento del coseno

$$(VIII.66) \quad \vartheta \simeq 2\pi f_s \left( \frac{ct}{\lambda/3} - 1 \right)$$

varrà  $n\pi$  per

$$(VIII.67) \quad 3ct_{rec}/\lambda = 1 - n/2f_s$$

( $t_{rec}$ : tempo della ricombinazione) sicché il coseno avrà modulo massimo (valori = 1) su numeri d'onda

$$(VIII.68) \quad k_n \simeq k_{hor,rec}(1 + n/2f_s).$$

( $k_{hor,rec}$ : numero d'onda corrispondente alla scala dell'orizzonte alla ricombinazione).

Abbiamo già visto in eq. (VIII.55) che il numero d'onda  $k_{hor,rec}$  corrisponde a  $l \simeq 200$ . Avremo allora altri massimi per

$$(VIII.69) \quad l_n \simeq 200 \left( 1 + \frac{n}{2f_s} \right)$$

Ad esempio, i grafici in figura VIII.9, coi massimi distanziati di  $\Delta l \sim 255$  implicano che sia  $f_s \sim 0.39$ , invece di  $\sim 0.58$ , come sarebbe se i barioni fossero trascurabili.

Lasciamo al lettore il calcolo del valore di  $\Omega_{ob}$  che, per  $\Omega_{om} \simeq 0.3$ , dà luogo a questo valore di  $f_s$ , tenendo conto che è  $f_s^{-2} = 3(1 + \delta\rho_b/\delta\rho_\gamma)$  e che l'onda è adiabatica.

## 10. Oscillazioni acustiche nei barioni: BAO

Passando dalle anisotropie del CMB alla funzione di trasferimento, notiamo che, in questo caso, le perturbazioni nella densità barionica lasciano segno solo se raggiungono la ricombinazione in fase di ventre. Sui numeri d'onda  $k_n$  dati da eq. (VIII.68) si manifestano quindi delle tenui oscillazioni nella funzione di trasferimento, la cui evo-

luzione post-ricombinazione verrà più dettagliatamente discussa nel prossimo capitolo.

Con la ricombinazione, il tempo di collisione  $\gamma\text{-}\gamma$  diventa praticamente  $\infty$  e i fotoni non costituiscono più un fluido. La sezione d'urto tra atomi di  $H$ , invece, assicura un tempo di collisione  $\ll t$  e il gas barionico mantiene le caratteristiche di fluido.

Ciò che cambia radicalmente con la ricombinazione, nel fluido barionico, è la velocità del suono. Le onde soniche, nel plasma, sono essenzialmente radiative e la loro velocità di propagazione  $f_s c$  riflette la velocità  $c$  dei fotoni. I neoformati atomi di  $H$ , invece, hanno velocità non-relativistiche  $v_H \sim \sqrt{k_B T / m_b}$ ; con temperature  $\sim 0.1$  eV, si ha già  $v_H \sim 3 \cdot 10^{-5} c$ . (È importante segnalare che, dopo la ricombinazione, la temperatura dei barioni non è più eguale al valore di  $T$  desumibile dallo spettro del CMB.) A un crollo di oltre 4 ordini di grandezza nella velocità del suono corrisponde un crollo nella massa di Jeans che, dalla scala dell'orizzonte  $M_{hor} \simeq 1.7 \times 10^{19} \Omega_m^{-1/2} h^{-1} M_\odot$ , cala a  $\sim 3\text{-}5 \times 10^5 \Omega_m^{-1/2} h^{-1} M_\odot$ . In un modello senza DM, tutte le fluttuazioni sopravvissute alla ricombinazione sarebbero quindi gravitazionalmente instabili.

In un modello realistico, la dinamica dei barioni dopo la ricombinazione è in dominata dalla presenza simultanea di fluttuazioni nella DM, che creano «buche di potenziale» più significative delle modeste fluttuazioni (inferiori di un fattore  $\sim 10$  rispetto alle fluttuazioni nella DM, che sono cresciute  $\propto a$ , tra equivalenza e ricombinazione) nella modesta componente barionica (che conta al più per il 18-20%).

L'ampiezza massima delle fluttuazioni barioniche (BAO: *baryonic acoustic oscillation*) nella funzione di trasferimento sarà quindi  $\sim 2\text{-}3\%$  dell'ampiezza della funzione stessa.

Tramite l'uso di campioni profondi, le BAO sono state individuate osservativamente e costituiscono uno degli elementi discriminatori sui valori dei parametri cosmologici.

## 11. Soppressione delle onde soniche di piccola scala

Il confronto tra la stima delle anisotropie e il loro spettro angolare, in figura VIII.9, mette in evidenza la soppressione nell'altezza dei picchi oltre il terzo-quarto. Essa non può essere capita se si trattano le fluttuazioni in un modello a due componenti, fluido barioni/fotoni e CDM. I limiti del modello sono però già stati segnalati: l'insieme barioni/fotoni costituisce un fluido solo su determinate scale e a seconda del *redshift*. All'avvicinarsi della ricombinazione, il c.l.m. dei fotoni,  $\lambda_\gamma$ , cresce; su scale  $\lambda$  sempre maggiori vengono allora meno le condizioni per la propagazione di onde soniche.

Dal punto di vista quantitativo si consideri il rapporto tra le scale:

$$(VIII.70) \quad \lambda_{\gamma,c} \simeq 2.19 \frac{0.02}{\Omega_{ob} h^2} \left( \frac{1000}{1+z} \right)^2 \frac{1}{x} \text{ Mpc} \quad (\text{c.l.m. comovente}),$$

$$(VIII.71) \quad L_{hor,c} = 190 (\Omega_{om} h^2)^{-1/2} \left( \frac{1000}{1+z} \right)^{1/2} \text{ Mpc} \quad (\text{orizzonte}).$$

È facile vedere che il rapporto

$$(VIII.72) \quad \frac{\lambda_{\gamma,c}}{L_{hor,c}} x \simeq 4 \times 10^{-3} \left( \frac{\Omega_{om}}{0.12} \right)^{1/2} \frac{0.02}{\Omega_{ob} h^2} \left( \frac{1000}{1+z} \right)^{3/2},$$

raggiunge l'unità a  $z \sim 950-1000$ , quando la ionizzazione è  $x \sim 10^{-2}-10^{-3}$ .

Il fatto che si osservino anisotropie provenienti da scale  $\lambda_4 \sim L_{hor}/2.4$ , sta a indicare che il cosmo è sostanzialmente trasparente a partire da un *redshift* un po' superiore a 1000. Infatti, il picco di emissione di fotoni del CMB si ha per  $z \simeq 1060$ .

Il confronto  $\lambda-\lambda_\gamma$  ci dice quando una fluttuazione esce dal regime sonico. Di fatto, già quando è  $\lambda_\gamma \sim \lambda/10$ , il trasferimento di impulso tra  $\gamma$  – che avviene grazie agli urti con la componente barionica – si svolge tra punti sfasati di  $\sim 2\pi/10$  (invece che localmente). Ne deriva quindi uno smorzamento già significativo. Quando poi  $\lambda_\gamma \sim \lambda/2\pi$ , i barioni trasferiscono impulso tra punti in fasi sostanzialmente scorrelate.

La perdita della coerenza sonica ha conseguenze che dipendono dalla rapidità con cui questo processo si verifica. Se la perdita di coerenza fosse «istantanea», le particelle resterebbero «cristallizzate» nella condizione della fase in cui l'onda si trova.

La permanenza del segnale nel CMB si ha allora per quelle lunghezze d'onda che risultano confrontabili o maggiori della profondità della *banda dell'ultimo scattering*, da cui provengono i fotoni osservati.

Definiamo allora i tempi:  $t_1$ , quando  $\lambda_\gamma \simeq \lambda/10$ ;  $t_2$  quando  $\lambda_\gamma = \lambda/2\pi$ . La condizione si può esprimere richiedendo che sia

$$(VIII.73) \quad c_s(t_2 - t_1) \gtrsim \lambda.$$

La lunghezza d'onda  $\lambda_D$ , per cui questa relazione vale come eguaglianza, fissa una scala-limite per la produzione di anisotropie significative. Lo smorzamento è progressivo e un'analisi di questo punto richiede il superamento graduale del modello di fluido. La trattazione sofisticata di questi processi ha prodotto raffinati algoritmi numerici, sulla cui struttura discuteremo ancora nel capitolo XIII.

# Appendice A1

## La ricombinazione dell'idrogeno

L'equilibrio termodinamico porta all'eq. (VIII.32), che qui riportiamo

$$(A1.1) \quad \frac{x^2}{1-x} \simeq \frac{1}{n_b} \left( \frac{m_e T}{2\pi} \right)^{3/2} \exp\left(-\frac{B_d}{T}\right).$$

Viceversa, il bilancio ricombinazione-reionizzazione avviene secondo l'equazione

$$(A1.2) \quad -\dot{x} = A n_b^2 \left[ x^2 - \frac{1}{n_b} (1-x) \left( \frac{m_e T}{2\pi} \right)^{3/2} e^{-\frac{B_d}{T}} \right]$$

dove  $A$  è un'opportuna costante di proporzionalità. Notiamo che, se l'eq. (A1.1) fosse soddisfatta, si avrebbe  $\dot{x} = 0$ . Il senso dei due addendi a secondo membro nell'equazione (A1.2) è chiaro: il primo addendo,  $\propto n_e \times n_p$ , rende conto della probabilità che elettroni e protoni si incontrino e si combinino. Il secondo addendo, invece, è proporzionale a  $n_H$  e alla frazione di  $\gamma$  con energia sufficiente a ionizzare; esso rende conto del processo inverso al precedente. L'eq. (A1.2) entra comunque in gioco quando la variazione di  $x$  prescritta dalla legge di equilibrio (A1.1) causi un valore di  $\dot{x}$  maggiore di quello calcolabile dall'eq. (A1.2). Il ruolo essenziale nel confronto è dunque dettato proprio dalla costante  $A$ .

Per calcolare  $A$  si deve tener conto che, se un elettrone con energia cinetica  $E_k$  va direttamente allo stato fondamentale, c'è un'elevata probabilità che il fotone prodotto, che ha energia  $\varepsilon = B_d + E_k$ , urti un altro atomo e lo reionizzi, prima che  $\varepsilon$  sia *redshiftata* al di sotto di  $B_d$ . Questo processo è quindi trascurabile. Se invece un elettrone è catturato in un livello energetico eccitato, a energia  $-B^*$ , la situazione è diversa. Il fotone prodotto, con energia  $E_k + B^*$ , può reionizzare solo atomi di  $H$  eccitati. Ma la permanenza di un qualsiasi atomo nello stato eccitato è breve; una successiva transizione lo porta allo stato fondamentale con produzione di un ulteriore fotone della serie di Lyman, di energia  $B_b - B^*$ . Basta allora un blando *redshift* di questo fotone (monocromatico) per renderlo inabile a riportare un altro atomo a livelli eccitati.

È anche possibile che l'iniziale ricombinazione porti l'atomo nello stato metastabile 2s, sicché nella transizione allo stato fondamentale si abbia l'emissione di 2 fotoni. Il tempo caratteristico per questo secondo processo è  $\tilde{\tau} \simeq 0.12$  s e, finché permane nello stato 2s, l'atomo è suscettibile di reionizzazione. Quindi questo processo ha un peso lievemente inferiore al precedente.

Tenendo conto di questi effetti, risulta

$$A = \alpha_e C \quad \text{con} \quad \alpha_e = 1.65 \cdot 10^{-12} (10^3 K/T)^{0.8} \text{ cm}^3/\text{s}, \quad C = \left( 1 + \frac{K\beta_e n_{1s}}{1 + Kn_{1s}/\tilde{\tau}} \right)^{-1}$$

$$\text{dove} \quad K = (\lambda_\alpha/2)^3 / (\pi \dot{a}/a), \quad n_{1s} \simeq (1-x)n_b, \quad \beta_e = \alpha_e (m_e T/2\pi)^{3/2} \exp(-B_2/T)$$

mentre  $\lambda_\alpha = 1.22 \cdot 10^{-8}$  cm e  $B_2 = 3.4$  eV sono la lunghezza d'onda della riga Lyman- $\alpha$  e l'energia di legame del livello  $n = 2$  dell'atomo di  $H$ .

## Appendice A2

### La banda dell'ultimo scattering

La ricombinazione fa crescere il c.l.m. dei  $\gamma$ , permettendo a un numero crescente di essi di giungere fino all'osservatore senza ulteriori urti Thomson.

Per stabilire la distribuzione su tempi o *redshift* dell'ultimo scattering dei fotoni, conviene considerare il processo opposto. Immaginiamo cioè che un fascio di  $\gamma$  proceda a ritroso nel tempo, partendo da  $z = 0$  e andando incontro al cosmo che si sta «ionizzando» con la legge  $x(t)$  ricavata nella precedente appendice. Via via che  $x$  cresce, un numero crescente di  $\gamma$  urterà un elettrone.

Il processo è completamente invertibile: la distribuzione degli assorbimenti riproduce esattamente la distribuzione degli «ultimi scattering» nel mondo reale.

Detto allora  $n(l)$  il numero di  $\gamma$  che hanno raggiunto una distanza fisica  $l$  dall'osservatore, procedendo a ritroso nel tempo, e  $n_o$  il loro numero a  $z = 0$ , ovvero il loro numero totale, avremo

$$(A2.1) \quad \frac{dn}{dl} = -n n_e \sigma_T$$

( $\sigma_T$  : sezione d'urto Thomson) e, poiché siamo sul «cono-luce», sarà

$$(A2.2) \quad dl = c dt = a(\tau) d\tau$$

dove  $\tau$  è il «tempo conforme». Otteniamo così un'equazione il cui integrale è

$$(A2.3) \quad \frac{n(\tau)}{n_o} = e^{-\int_0^\tau d\bar{\tau} n_e(\bar{\tau}) a(\bar{\tau}) \sigma_T}.$$

La distribuzione su  $\tau$  dell'ultimo scattering dei  $\gamma$  sarà allora

$$(A2.4) \quad V(\tau) = \frac{d[n(\tau)/n_o]}{d\tau} = n_e(\tau) a(\tau) \sigma_T e^{-\int_0^\tau d\bar{\tau} n_e(\bar{\tau}) a(\bar{\tau}) \sigma_T}.$$

Questa distribuzione dipende dal modello cosmologico. Non dipende invece, in pratica, dal modello cosmologico la distribuzione sul *redshift*  $z$ .

Ciò è frutto di un certo numero di fattori concomitanti che permettono, in definitiva, di approssimare l'andamento di  $x$  con  $z$ , attorno a  $z = 10^3$ , come ricavato nella precedente appendice, tramite l'espressione

$$(A2.5) \quad x(z) = \frac{2.4 \times 10^{-3}}{\Omega_b h} \left( \frac{z}{10^3} \right)^{12.75}$$

di modo che  $n_e \propto \rho_b x$  è indipendente da  $\Omega_b$  e

$$(A2.6) \quad d\tau_{op} = \sigma_T n_e (a/a') da$$

con  $\mathcal{H} = a'/a = (da/d\tau)/a \propto h$ , è anche indipendente da  $h$ . Numericamente risulta allora

$$(A2.7) \quad \tau_{op}(z) = 0.37(z/10^3)^{14.25}.$$

La distribuzione su  $z$

$$(A2.8) \quad \nu(z) = V[\tau(z)] d\tau/dz = (d\tau_{op}/dz) e^{-\tau_{op}(z)}$$

ha allora l'andamento mostrato in figura VIII.5.

# Appendice A3

## CDM e fluido barioni-fotoni accoppiati gravitazionalmente

### 1. Equazioni di Einstein linearizzate

Nell'appendice A1 del capitolo precedente si sono già usate le equazioni

$$(A3.1) \quad \begin{aligned} \delta' + (1+w)(kv + h'/2) + 3(a'/a)(c_s^2 - w)\delta &= 0, \\ v' + \left(1 - 3w + \frac{w'}{1+w}\right) \frac{a'}{a} v - \frac{c_s^2}{1+w} k\delta &= 0, \end{aligned}$$

ricavate dalle equazioni di Einstein linearizzate, per un fluido con parametro di stato  $w = p/\rho$  e velocità del suono  $c_s = \sqrt{\delta p/\delta \rho}$ , in cui sono presenti fluttuazioni  $\varepsilon(\mathbf{x}, \tau) = \delta\rho(\mathbf{x}, \tau)/\rho$  ( $\ll 1$ ), sviluppate in componenti di Fourier  $\delta(\mathbf{k}, \tau)$  ( $\mathbf{k} = k\mathbf{n}$ : numero d'onda). La derivata ' è effettuata rispetto a  $\tau$  (tempo conforme). Nel caso di simultanea presenza di più componenti, accoppiate dalla sola gravità, le eq. (A3.1) si applicano alle fluttuazioni  $\delta_i$  di ciascuna componente, usando i relativi valori  $c_{s_i}$  e  $w_i$  di velocità del suono e parametro di stato. L'auto-gravità delle fluttuazioni, in una gauge sincrona, è descritta da un 3-tensore  $h_{ij}(\tau, k)$ , essendo

$$(A3.2) \quad ds^2 = a^2(\tau)[d\tau^2 - (\eta_{ij} + h_{ij})dx^i dx^j];$$

con  $\eta_{ij} = \delta_{ij}$  nel caso di coordinate spaziali cartesiane ortogonali ( $\eta_{ij}$  non ha nulla a che fare col potenziale gravitazionale  $\eta = (h - \omega)/6$  definito in eq. (A1.19) nel capitolo precedente). Tutta l'azione della gravità si riporta alle funzioni  $h(\tau, \mathbf{k})$  e  $\omega(\tau, \mathbf{k})$  (richiamiamo l'attenzione del lettore sul diverso significato dei simboli  $w$  e  $\omega$ ) nello sviluppo di Fourier:

$$(A3.3) \quad h_{ij}(\tau, \mathbf{x}) = \int d^3k e^{-ik \cdot \mathbf{x}} [h \delta_{ij}/2 + \omega(n_i n_j - \delta_{ij}/2)].$$

Sotto la condizione  $w_i, c_{s_i} = \text{const.}$ , delle eq. (A3.1) si ottiene (per i dettagli vedi appendice A1 del capitolo precedente)

$$(A3.4) \quad \delta_i'' + (1 - 3w_i)(a'/a)\delta_i' + c_{s_i}^2 k^2 \delta_i + (1 + w_i)(1/2)[h'' + (1 - 3w_i)h'a'/a] = 0,$$

in cui non compare  $v$ , né dipendenza da  $\mathbf{n}$ , sicché  $\delta_i$  e  $h$  sono funzioni di  $\tau$  e del solo modulo  $k$ . Ricordiamo poi che vale anche l'equazione

$$(A3.5) \quad h'' + (a'/a)h' = -8\pi G a^2 \sum_i \rho_i [1 + 3c_{s_i}^2] \delta_i,$$

in cui la somma è ora estesa a tutte le componenti, di densità  $\rho_i$ . Le eq. (A3.4)-(A3.5) determinano l'andamento temporale delle fluttuazioni.

Per una sostanza con  $w_r = 0$ , il secondo membro di eq. (A3.5) può essere sostituito in eq. (A3.4), ricavando l'equazione ( $i$  sarà ora indice di somma)

$$(A3.6) \quad \delta_r'' + (a'/a)\delta_r' - 4\pi G a^2 \sum_i \rho_i [1 + 3c_{s_i}^2] \delta_i = 0,$$

in cui non compare il campo gravitazionale. Se invece è  $c_{s_r}^2 = w_r = 1/3$ , avremo

$$(A3.7) \quad \delta_r'' + k^2 \delta_r/3 + (2/3)h'' = 0,$$

sicché l'eq. (A3.5) non è direttamente utilizzabile per eliminare la variabile  $h$ .

## 2. Il problema di Jeans nei modelli con CDM

Con passaggi analoghi a quelli svolti per una sola sostanza, in presenza di fluido barioni-fotoni e CDM, la lunghezza d'onda di Jeans per il fluido sarà:

$$(A3.8) \quad \lambda_J = \frac{2\pi a}{k_J} \simeq c_s \sqrt{\frac{\pi}{G\rho}} \sqrt{\frac{1}{(1+w)[\Omega_c + (1+3c_s^2)\Omega_{b\gamma}]}}$$

$\Omega_{b\gamma}$  e  $\Omega_c$  sono i parametri di densità di fluido e CDM. Ricordiamo che risulta

$$(A3.9) \quad c_s^2 \simeq \frac{\delta p_\gamma}{\delta \rho_\gamma + \delta \rho_b} \simeq \frac{c^2}{3} \frac{1}{1+3a/4a_{eq,b}}, \quad w = \frac{p_\gamma}{\rho_\gamma + \rho_b} = \frac{1}{3} \frac{1}{1+a/a_{eq,b}};$$

$a_{eq,b}$  è il fattore di scala quando risulta  $\rho_\gamma = \rho_b$ ; sarà quindi

$$a_{eq,b} = a_{eq} \frac{\Omega_{ob}}{\Omega_{ob} + \Omega_{oc}}$$

e, ad esempio, in un ragionevole modello con  $\Omega_{oc} = 0.26$ ,  $\Omega_{ob} = 0.04$ ,  $h = 0.7$ , avremo  $a_o/a_{eq,b} \simeq 80$ . Per questo modello valgono i grafici in figura VIII.6.

## 3. Il problema dinamico

Per un mix di fluido e CDM, sappiamo che le eq. (A3.6)-(A3.7) non permettono di eliminare il campo  $h$ , come nel caso  $w = 0$ . Tuttavia, se trascuriamo i termini gravitazionali, l'eq. (A3.7) diviene

$$(A3.10) \quad \delta'' + c_s^2 k^2 \delta = 0$$

e il suo integrale generale

$$(A3.11) \quad \delta = A_+ e^{i c_s k \tau} + A_- e^{-i c_s k \tau}$$

vale nella stessa forma prima e dopo l'equivalenza. Vedremo tra breve che i termini gravitazionali non si possono sempre trascurare completamente, ma che in molti casi questa è una buona approssimazione. Un'altra approssimazione contenuta nell'integrale (A3.11) è l'assunzione che  $c_s$  sia costante. Nei casi fisici  $c_s$  varia molto blandamente. Un'espressione della forma

$$(A3.11a) \quad \delta = A_+ e^{i k \int c_s dt'} + A_- e^{-i k \int c_s dt'}$$

può tenere conto di tali variazioni, ma è comunque ancora approssimata.

Dalla relazione  $d\tau = dt/a$ , scegliendo tempi di normalizzazione  $\bar{t}$  di modo che sia  $a = (t/\bar{t})^\alpha$  con  $\alpha = 1/2$  e  $2/3$ , rispettivamente quando l'espansione è dominata dalla radiazione e dalla materia, si vede che è

$$(A3.12) \quad \int d\tau' = \frac{1}{1-\alpha} \bar{t}^\alpha t^{1-\alpha} = \frac{1}{1-\alpha} \left(\frac{\bar{t}}{t}\right)^\alpha t = \frac{1}{1-\alpha} \frac{t}{a}$$

L'integrale (A3.11) può quindi essere posto nella forma

$$(A3.13) \quad \delta = A_+ e^{i\nu c_s k t/a} + A_- e^{-i\nu c_s k t/a} \quad \text{con} \quad \nu = 2 \text{ o } 3$$

per espansione dominata dalla radiazione o dalla materia. Questa espressione interpola anche efficacemente tra i due regimi, a patto di ricavare un opportuno andamento di  $v(t)$  al momento della transizione da 2 a 3 .

Essendo poi  $\lambda = 2\pi a/k$ , l'integrale (A3.13) si potrà scrivere anche

$$(A3.14) \quad \delta = A_+ e^{2v\pi i \frac{c_s t}{\lambda}} + A_- e^{-2v\pi i \frac{c_s t}{\lambda}}$$

Questo integrale generale dovrà soddisfare le condizioni di avere un'opportuna ampiezza  $\delta(t_{hor}, k) \equiv \delta_{hor}(k)$  al tempo  $t_{hor} \simeq \lambda/vc$  in cui l'orizzonte raggiunge la scala  $\lambda$ , nonché la condizione  $\dot{\delta}(t_{hor}, k) = 0$ , che contrassegna il passaggio dalla fase di crescita alla fase oscillatoria. Fin tanto che barioni e fotoni sono ben accoppiati, vale allora l'espressione

$$(A3.15) \quad \delta(t, k) = \delta_{hor}(k) \cos \left[ 2\pi \left( \frac{vc_s t}{\lambda} - f_s \right) \right] \quad \text{con } f_s = c_s/c,$$

fin tanto che i termini gravitazionali sono trascurabili.

#### 4. Fluttuazioni nella CDM a $z > z_{eq}$ . Effetto di stagnazione

L'andamento delle fluttuazioni nella DM, tra entrata nell'orizzonte e equivalenza, è l'elemento più importante a delineare la «funzione di trasferimento».

In questo periodo, assumendo che le fluttuazioni nel fluido barioni-fotoni abbiano l'andamento (A3.15), per cui si ha  $\langle \delta \rangle \simeq 0$ , si trascura l'azione gravitazionale del fluido. L'eq. (A3.6) diviene allora

$$(A3.16) \quad \delta_c'' + (a'/a)\delta_c' - 4\pi G a^2 \rho \Omega_c \delta_c = 0$$

e, espressa in funzione del tempo ordinario, darà

$$(A3.16a) \quad \ddot{\delta}_c + 2(\dot{a}/a)\dot{\delta}_c - 4\pi G \rho \Omega_c \delta_c = 0$$

Infatti

$$\frac{d}{d\tau} \frac{d\delta_c}{d\tau} = a \frac{d}{dt} \left( a \frac{d\delta_c}{dt} \right) = a^2 \left( \ddot{\delta}_c + \frac{\dot{a}}{a} \dot{\delta}_c \right)$$

D'altronde, in era radiativa, risulta  $2\dot{a}/a = 1/t$  mentre è  $4\pi G \rho = 3/2t^2$ ; se  $\dot{\delta}_c \sim \delta_c/t$  il terzo termine si può allora trascurare e l'eq. (A3.16) diviene

$$d\dot{\delta}_c/\dot{\delta}_c = -dt/t, \quad \text{da cui } \dot{\delta}_c = A/t, \quad \delta_c = A \ln(t/\bar{t})$$

con  $A$  e  $\bar{t}$  costanti arbitrarie. La crescita della fluttuazione nella CDM, dall'entrata nell'orizzonte all'equivalenza, segue dunque una legge logaritmica.

Dell'eq. (A3.16) si può anche dare un'integrale apparentemente meno approssimato, trascurando comunque i barioni e definendo la variabile

$$y = \rho_c/\rho = a/a_{eq}$$

(si noti che è  $0 < y \leq 1$ ) usando la quale, l'eq. (A3.16) assume la forma

$$(A3.17) \quad \ddot{\delta}_c + \frac{1 + (3/2)y}{y(1+y)} \dot{\delta}_c - \frac{3/2}{y(1+y)} \delta_c = 0$$

che ha soluzioni

crescente:

$$\delta_{(c,+)} = B_+[1 + (3/2)y]$$

decrescente:

$$\delta_{(c,-)} = B_-[1 + (3/2)y] \ln \left\{ \frac{(1+y)^{1/2} + 1}{\ln(1+y)^{1/2} - 1} \right\} - 3[1+y]^{1/2}$$

Essendo  $0 < y \leq 1$  tra valore massimo e minimo di  $\delta_c$  risulta un fattore

$$(A3.18) \quad \frac{\delta_{(c,+)}(y=1)}{\delta_{(c,+)}(y=0)} = \frac{5}{2}.$$

Entrambe queste trattazioni stabiliscono un effetto di «stagnazione» tra entrata nell'orizzonte e equivalenza. A fronte di una crescita quasi nulla all'interno dell'intervallo  $a_{hor} - a_{eq}$ , diventano peraltro importanti i transienti all'inizio e alla fine dell'intervallo stesso. In un esame numerico del problema, stagnazione e transienti si possono conglobare in una crescita con legge

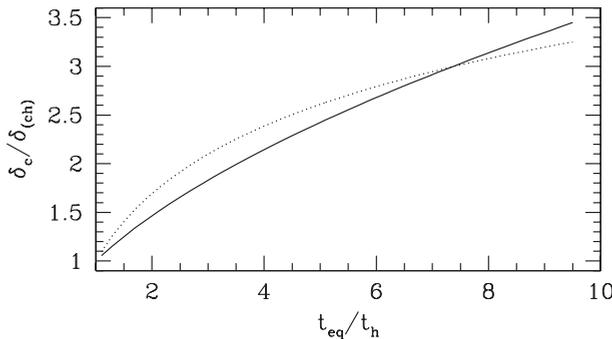
$$(A3.19a) \quad \delta_c \sim \delta_{c,h}(a/a_{hor})^u, \quad \text{con } u \simeq 0.25 - 0.30,$$

che conviene utilizzare in espressioni approssimate e che, per il range di scale significativo, non dà risultati molto diversi dall'espressione

$$(A3.19b) \quad \delta_c = \delta_{c,h} [\ln(t/t_{hor}) + 1]$$

ricavabile da eq. (A3.17) (vedi Fig. VIII.10); entrambe le espressioni, come i risultati numerici, violano lievemente il vincolo (A3.18)

Notiamo infine che la modesta crescita delle fluttuazioni nella DM rende autoconsistente la trattazione del fluido. In effetti, con  $\langle \delta \rangle = 0$  e fluttuazioni gravitazionalmente trascurabili nella DM, l'eq. (A3.10) è l'equazione fisica da risolvere per determinare l'andamento delle fluttuazioni.



**FIGURA VIII.10**

Confronto tra le espressioni (A3.19); la curva punteggiata si ottiene dalla legge logaritmica. La curva continua è ottenuta con  $u = 0.275$ .

### 5. Fluttuazioni a $z < z_{eq}$

Dopo l'equivalenza, lo studio delle fluttuazioni richiede la soluzione del sistema di equazioni

$$(A3.19a) \quad \ddot{\delta}_c + 2(\dot{a}/a)\dot{\delta}_c = 4\pi G\rho_c \delta_c,$$

$$(A3.19b) \quad \ddot{h} + 2(\dot{a}/a)\dot{h} = -8\pi G\rho_c \delta_c,$$

$$(A3.19c) \quad \ddot{\delta} + (\dot{a}/a)\dot{\delta} + (c_s k/a)^2 \delta + (2/3)\ddot{h} = 0,$$

in cui le derivate sono fatte rispetto al tempo ordinario; in effetti, come già sottolineato nel Par. 1 di questa appendice, fra la seconda e la terza di queste equazioni non si può direttamente eliminare la variabile  $h$ . Nello scrivere il sistema si è comunque già fatta l'assunzione che sia  $\langle \delta \rangle = 0$  e le fluttuazioni nel fluido non siano sorgente di gravità. A  $z > z_{eq}$  si tratta di un'assunto autoconsistente; in questo caso vedremo c'è qualche complicazione in più.

L'eq. (A3.19a) si può riscrivere

$$(A3.20) \quad \ddot{\delta}_c + \frac{4}{3t}\dot{\delta}_c = \frac{4}{3t^2}\Omega_c \delta_c$$

e, se si assume  $\delta_c = At^\alpha$ , se ne ricava l'equazione

$$(A3.21) \quad \alpha^2 + \frac{1}{3}\alpha - \frac{2}{3}\Omega_c = 0$$

che sarebbe algebrica se  $\Omega_c$  fosse costante.  $\Omega_c$ , invece, all'equivalenza è  $\sim 1/2$  e assume valori crescenti al calare di  $z$ , avvicinando progressivamente l'unità. Escluso il tratto a ridosso dell'equivalenza, tuttavia, si tratta di una variabilità molto più blanda di una legge di potenza ed è ragionevole trattarlo come costante. Le soluzioni di eq. (A3.21) sono combinazioni di due modi, crescente o decrescente. Per il modo crescente, considerando  $1 - \Omega_c \ll 1$ , abbiamo

$$(A3.22) \quad \alpha = \frac{2}{3} - \frac{2}{5}(1 - \Omega_c)$$

Al variare di  $\Omega_c$ ,  $\alpha$  varia quindi da  $\sim 0.467$  a  $\sim 0.667$ . Sostituiamo allora  $\delta_c = At^\alpha$  e porremo  $\dot{h} = Bt^\beta$  in eq. (A3.19b); continuando coerentemente a trattare  $\Omega_c$  e  $\alpha$  come costanti, ricaviamo allora

$$(A3.23) \quad \dot{h} = -\frac{4\Omega_c}{3\alpha+1} \frac{\delta_c}{t} \quad e \quad \ddot{h} = \frac{4(1-\alpha)\Omega_c}{3\alpha+1} \frac{\delta_c}{t^2} \simeq \frac{4}{9} At^{\alpha-2}$$

Si noti infatti che il coefficiente  $4(1-\alpha)\Omega_c/(3\alpha+1)$  è pressochè indipendente da  $\Omega_c$  e vale comunque  $\simeq 4/9$ .

Nell'eq. (A3.19c), allora, la dipendenza dal tempo del secondo e terzo termine sarà assai simile. In effetti, quando  $\Omega_c \simeq 1$ , è senz'altro  $a^{-2} \propto t^{-4/3}$ , allo stesso modo in cui, per  $\alpha \simeq 2/3$ , è  $t^{\alpha-2} = t^{-4/3}$ . Peraltro, in prossimità di  $z_{eq}$ , quando è  $\alpha - 2 \sim -4/3 - 0.2$ , anche il fattore di scala dipende da  $t$  con una potenza intermedia tra  $2/3$  e  $1/2$ , sicché sarà anche  $a^{-2} \propto t^{-4/3-0.2}$ .

Moltiplicando quindi l'eq. (A3.19c) per  $a^2$  si ricava

$$(A3.24) \quad a^2 \ddot{\delta} + a^2 (\dot{a}/a) \dot{\delta} + c_s^2 k^2 \delta = -(8/27) A a^2 t^{\alpha-2}.$$

L'integrale generale di questa equazione si ottiene come somma dell'integrale generale dell'omogenea associata che, se scritta usando il tempo conforme, altro non è che l'eq. (A3.10), più un integrale particolare dell'equazione qui sopra. Essendo  $a^2/t^{\alpha-2} \simeq \text{const.}$ , si può dunque

scegliere come integrale particolare

$$(A3.25) \quad \bar{\delta}(k) = -\frac{8}{27} \delta_c \left( \frac{a}{c_s k t} \right)^2 = -\frac{8}{27(2\pi)^2} \delta_c \left( \frac{\lambda}{c_s t} \right)^2$$

L'integrale generale coincide quindi con l'espressione (A3.15), salvo uno spostamento del centro dell'oscillazione dato da  $\bar{\delta}(k)$ . Il rapporto  $\lambda/c_s t$ , indipendente da  $z$ , può essere opportunamente valutato (vedi testo del capitolo).

Resta da segnalare che questo spostamento fa sì che  $\langle \delta \rangle$  sia negativo e non nullo. Il coefficiente che moltiplica  $\delta_c$  in eq. (A3.25) è peraltro  $\ll 1$ , sicché questa correzione non è molto significativa rispetto alle approssimazioni comunque coinvolte.

## CAPITOLO IX

# Evoluzione non lineare

## Lo sviluppo delle fluttuazioni dopo la ricombinazione

### 1. Spettro primordiale e spettro trasferito

Nei precedenti capitoli abbiamo visto che, su tutte le scale, all'entrata nell'orizzonte, debbono esser presenti delle fluttuazioni riconducibili a un opportuno *spettro primordiale*. Anche con questa necessaria aggiunta, il quadro delle *condizioni iniziali* è comunque altamente unificante. Le stesse fluttuazioni  $\delta(k)$ , che all'orizzonte assumeranno ampiezze  $\sim \mathcal{O}(10^{-5})$  distribuite Gaussianamente attorno a un valor medio nullo, dipendono solo dal modulo  $k$  di  $\mathbf{k}$ , di modo che la distribuzione delle anisotropie è... isotropa.

In definitiva, se ci poniamo a un tempo  $t_i$  antecedente l'entrata nell'orizzonte delle scale cosmologicamente rilevanti, possiamo dare l'espressione

$$(IX.1) \quad P(t_i, k) = \langle |\delta(k)|^2 \rangle = A(t_i) k^{n_s}$$

dello spettro primordiale, specificata tramite 2 quantità:  $A(t_i)$  e  $n_s$ , ampiezza al tempo  $t_i$  e indice spettrale. Ove si assuma poi  $n_s = 1$ , la fluttuazione quadratica media (o varianza di massa) sarà eguale su tutte le scale, al momento della loro entrata nell'orizzonte, sicché si può dire che tutta l'informazione contenuta dallo spettro si riduce a un'unica quantità.

Dopo l'entrata nell'orizzonte, lo spettro delle varie componenti (radiazione, DM, barioni, ecc.) viene modulato, in sostanza, da 3 effetti:

- 1) la stagnazione, determinante a forgiare la *funzione di trasferimento* per la componente di DM;
- 2) le onde soniche nel plasma, determinanti nel forgiare le anisotropie del CMB e chiaramente visibili nella funzione di trasferimento dei barioni;
- 3) la dissipazione delle onde soniche su scale via via crescenti.

Dopo la ricombinazione, dunque, avremo uno *spettro trasferito*

$$(IX.2) \quad P(t, k) = P(t_i, k) \mathcal{L}^2(t_i, t) T^2(t, k)$$

in cui  $\mathcal{L}(t_i, t)$  esprime la crescita delle fluttuazioni nella DM e nei barioni, tra tempo  $t_i$  e

tempo  $t$ , al limite  $k \rightarrow 0$ , vale a dire su scale molto al di sopra dell'orizzonte;  $\mathcal{T}(t, k)$ , funzione di trasferimento al tempo  $t$ , riassume invece l'azione dei succitati effetti 1), 2), 3) e di tutta la dinamica a essi collegata, sulle fluttuazioni a  $k$  abbastanza elevati da aver avuto una storia, in pratica *sotto l'orizzonte* al tempo  $t$ .

L'azione dei fotoni, dalla fine della ricombinazione, è quasi del tutto trascurabile. Una certa enfasi venne data, fino a un decennio fa, a una possibile componente *hot*, di particelle prima appartenenti alla componente radiativa e che divenissero non-relativistiche tra equivalenza e ricombinazione, o poco dopo. In particolare, neutrini con massa  $\sim eV$  avrebbero giocato un simile ruolo. Dedicheremo un paragrafo, nel capitolo XIII, al ruolo dei neutrini, sia pure alla luce delle restrizioni ormai dettate da una evoluta fenomenologia. Essa permette di affermare che il loro ruolo, nel plasmare il cosmo attuale, è al più marginale.

Tra le ragioni che portavano a invocare una componente *hot* c'era il dato che le componenti *fredde* del cosmo sono difficilmente in grado di portarlo dalla densità critica. L'opzione  $\Omega_{tot} \ll 1$ , invece, oltre ai problemi di fine tuning, fin dall'esperimento COBE è in contrasto coi dati sul CMB. La copertura del *gap* tra le componenti materiali, fredde, e la densità critica è oggi invece attribuita all'energia oscura (DE: *dark energy*), sulla base dei dati che abbiamo riassunto nel capitolo V. Dalla ricombinazione alla vigilia dell'era attuale, peraltro, la DE è irrilevante nella dinamica delle strutture.

La discussione in questo capitolo s'incentra quindi su due classi di modelli cosmologici: il modello  $\Lambda$ CDM (*standard cold-dark-matter*) e il modello delle concordanze ( $\Lambda$ CDM:  *$\Lambda$ -cold-dark-matter*). Per entrambi è  $\Omega_{tot} = 1$ . Per  $\Lambda$ CDM, inoltre, sarà

$$(\Lambda\text{CDM}) \quad \Omega_{o,m} = \Omega_{o,c} + \Omega_{o,b} \simeq 1,$$

mentre, per  $\Lambda$ CDM,

$$(\Lambda\text{CDM}) \quad \Omega_{o,m} = \Omega_{o,c} + \Omega_{o,b} \simeq 0.3;$$

con l'aggiunta di  $\Omega_{o,de} \simeq 0.7$ . (In genere, per  $\Lambda$ CDM si usa un valore  $h \sim 0.5$ , per il *parametro di Hubble*; per  $\Lambda$ CDM, invece, si ha  $h \sim 0.7$ .) Nonostante  $\Lambda$ CDM sia un modello falsificato dalla fenomenologia, esso resta importante per due fatti:

- 1) Fino a una *redshift*  $z \sim 3$ , le cosmologie  $\Lambda$ CDM e  $\Lambda$ CDM sostanzialmente coincidono.
- 2) Molti problemi hanno una soluzione analitica in  $\Lambda$ CDM, mentre in  $\Lambda$ CDM si danno solo risultati numerici, che però si comprendono bene alla luce dei risultati  $\Lambda$ CDM.

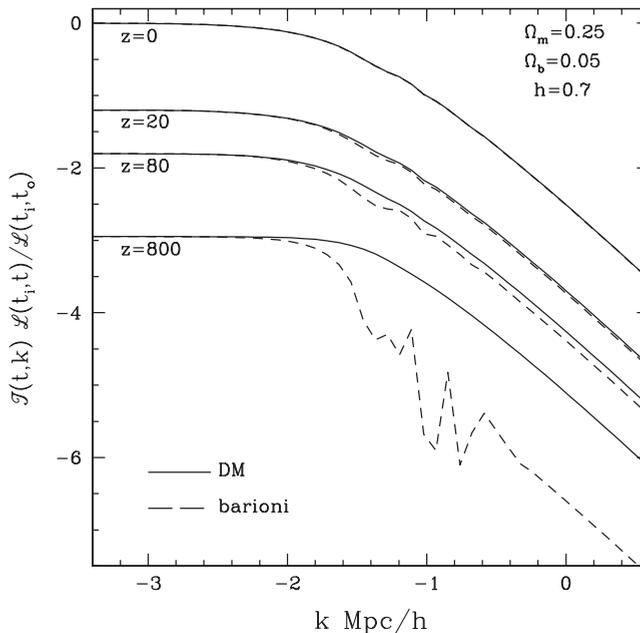
In tutto il quadro delineato nei precedenti capitoli si è sfruttata a fondo una caratteristica della dinamica lineare, quella di agire separatamente sui diversi numeri d'onda  $k$ : spettri e funzioni di trasferimento sono espressi in funzione di  $k$ ; la dinamica su ciascuna scala non dipende dalle altre scale. Quando l'ampiezza  $\delta$  delle fluttuazioni raggiunge valori  $\sim 0.2-0.3$ , per uscire dal regime lineare, questa determinante semplificazione dell'analisi teorica non potrà più essere invocata.

## 2. Dalla ricombinazione alla non-linearità

Subito dopo la ricombinazione, gli spettri nella DM e nei barioni, sulle scale  $k > k_{h,eq}$ , sono molto diversi (vedi Fig. IX.1).

Infatti, a partire dall'equivalenza, a  $z \sim 10^4$ , le fluttuazioni nella DM sono cresciute con legge  $\delta \propto a \propto t^{2/3}$ . Nel frattempo le fluttuazioni nei barioni, strettamente legati alla radiazione, facevano parte di onde soniche. Mediamente, quindi, la loro ampiezza non è cresciuta e, anzi, sulle scale  $k \gtrsim 2.5 k_{hor,rec}$ , il c.l.m. dei  $\gamma$  raggiunge la lunghezza d'onda  $\lambda = 2\pi a/k$  prima della ricombinazione, e una parte crescente (al crescere di  $k$ ) della potenza contenuta nelle onde soniche si dissipa in calore. Nei barioni, qualche traccia della fluttuazione primeva si riscontra fino a  $\sim 15-20 k_{hor,rec}$ . Più oltre, le piccole fluttuazioni barioniche visibili in figura IX.1 a  $z \simeq 800$ , sono indotte dalle fluttuazioni nella DM.

In genere, dunque, dopo la ricombinazione, avremo una fluttuazione *piccola* o nulla nei barioni e una fluttuazione maggiore nella DM. In figura IX.1, a causa della scala logaritmica, non vengono esibiti i segni relativi delle fluttuazioni; è tuttavia chiaro che,



**FIGURA IX.1**

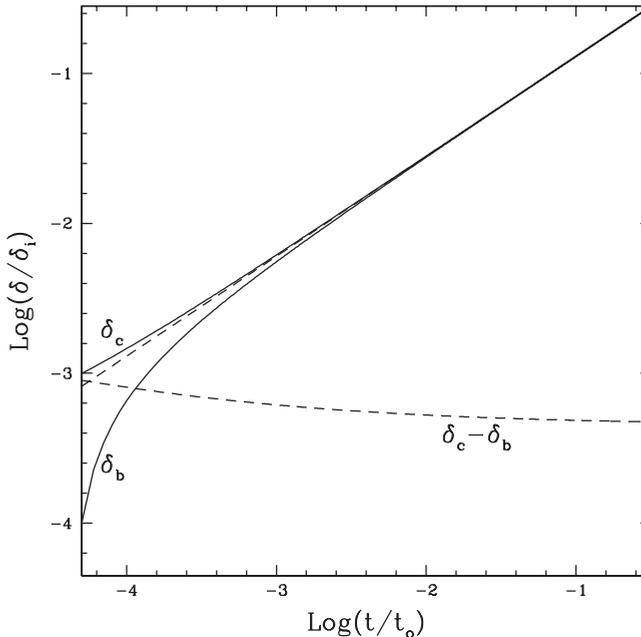
Funzioni di trasferimento e evoluzione delle fluttuazioni nella DM e nei barioni, per il modello indicato nel disegno. A  $z \simeq 800$  e per  $k > k_{h,eq}$ , le fluttuazioni barioniche, per cui si mostrano i moduli, risentono visibilmente del precedente andamento sonico. A  $z = 80$ , i residui delle onde soniche restano più evidenti nei barioni che nella DM. A  $z \simeq 20$ , i due spettri sono praticamente uguali e le oscillazioni barioniche si sono trasferite anche sulla DM, ove permangono anche a  $z = 0$ . Sia ascissa che ordinata sono logaritmiche.

per il gioco delle fasi nelle onde soniche, le fluttuazioni nei barioni su determinate scale sono opposte in segno a quelle nella DM.

In questo quadro le ampie fluttuazioni della DM giocano il ruolo determinante di *seme* per le future disomogeneità. Fissiamoci su un particolare numero d'onda e consideriamo un «tempo iniziale»  $t_i$ , successivo alla ricombinazione, in cui  $\delta_{i,c}$  e  $\delta_{i,b}$  sono le fluttuazioni in DM e barioni; trascorso un tempo ragionevole da  $t_i$ , l'evoluzione avviene secondo la legge

$$(IX.3) \quad \begin{aligned} \delta_c &= \alpha_+(\Omega_c \delta_{c,i} + \Omega_b \delta_{b,i})(t/t_i)^{2/3} + \zeta_+(\delta_{c,i} - \delta_{b,i}), \\ \delta_b &= \alpha_+(\Omega_c \delta_{c,i} + \Omega_b \delta_{b,i})(t/t_i)^{2/3} - \zeta_+(\delta_{c,i} - \delta_{b,i}), \end{aligned}$$

in cui  $\alpha_+$  e  $\zeta_+$  sono numeri compresi tra 0 e 1. Essi conglobano la nostra ignoranza sul modo di «saldare» la soluzione delle equazioni post-ricombinazione con le soluzioni precedenti. Ulteriori dettagli sono forniti in appendice A1. Nel caso illustrato in figura IX.2 si assume  $\delta_{i,b} \simeq 0.1 \delta_{i,c}$ ; inoltre si suppone che sia  $\alpha_+ = 1$ ,  $\zeta_+ = 1/2$ . Le equazioni (IX.3) mostrano chiaramente che il rilievo della differenza  $\delta_c - \delta_b$  va pro-



**FIGURA IX.2**

Evoluzione delle fluttuazioni nella DM e nei barioni, dopo la ricombinazione e fino a un'epoca in cui la DE può divenire importante dinamicamente. In questo grafico ogni interazione residua tra barioni e radiazione è trascurata. Nell'esempio qui esposto si è assunto:  $\delta_{i,c}/\delta_{i,b} = 10$ ,  $\Omega_{o,m}/\Omega_{o,b} = 5$ . Le linee tratteggiate mostrano il calare della differenza tra  $\delta_c$  e  $\delta_b$  e l'andamento regolare, a legge di potenza, della somma  $\Omega_c \delta_c + \Omega_b \delta_b$ .

gressivamente calando, via via che  $\delta_c$  e  $\delta_b$  crescono, fino a divenire  $\sim 10^{-3}$ . In realtà la tenue interazione residua fotoni-barioni rallenta ulteriormente l'avvicinamento tra i due spettri rispetto alla legge (IX.3); lo si intravede anche in figura IX.1.

Nella sostanza, dopo la ricombinazione, i barioni accrescono nelle *bucche di potenziale* create dalle fluttuazioni nella DM. Si noti che le eq. (IX.3) mostrano chiaramente come la fluttuazione complessiva nella materia non-relativistica,

$$(IX.4) \quad \delta = \frac{\delta\rho_c + \delta\rho_b}{\rho} = \frac{\delta_c\Omega_c\rho + \delta_b\Omega_b\rho}{\rho} = \Omega_c\delta_c + \Omega_b\delta_b,$$

cresca in modo  $\propto (t/t_i)^{2/3}$ . Per provarlo basta moltiplicare la prima della eq. (IX.3) per  $\Omega_c$ , la seconda per  $\Omega_b$  e sommarle membro a membro.

In figura IX.2 le ampiezze delle fluttuazioni sono normalizzate arbitrariamente. figura IX.1, invece, mostra l'andamento lineare delle fluttuazioni in funzione di  $z$ , su varie scale  $k$ , ottenuto numericamente; le ampiezze sono qui normalizzate in modo che, per  $k \rightarrow 0$ , sia oggi  $\delta = 1$ . Si ricordi comunque che le equazioni lineari danno soluzioni indipendenti dall'ampiezza; è un fatto di notevole importanza analitica, che entrerà nell'analisi post-lineare di questo capitolo.

Una stima della normalizzazione fisica, invece, si può fare a partire da un'ampiezza «media»  $\sim 10^{-5}$  all'entrata nell'orizzonte, che cresce di un mezzo ordine di grandezza fino all'equivalenza, a  $z \sim 10^4$ , e da allora con legge  $\propto (1+z)^{-1}$ . A  $z \sim 3$ , la fluttuazione «media» sarà quindi  $\sim 10^{-1}$ . Sostanziali effetti non lineari si innescano solo quando le fluttuazioni sono  $\sim 0.2-0.3$ . Una fluttuazione doppia della media, dunque, inizia il suo percorso non-lineare a  $z \sim 3$ . Fluttuazioni eccezionali possono divenire non-lineari a  $z \sim 10$  e, in casi estremi, a  $z$  ancora maggiori; ma un limite attorno a  $z \sim 20$ , che comunque appare ragionevole, è indotto anche da altri meccanismi.

A questi *redshift*, molto lontani dalla ricombinazione, la differenza tra fluttuazioni in DM e barioni, in figura IX.2, è  $\sim \mathcal{O}(10^{-4})$ . Ma, anche tenendo conto dell'accoppiamento residuo barioni-fotoni, figura IX.1 mostra la completa coerenza tra fluttuazioni nei barioni e nella DM a  $z \sim 20$ .

La conclusione è che, quando iniziano le fasi non-lineari, fluttuazioni in DM e barioni sono praticamente eguali; gli effetti residui delle onde soniche, anche se non trascurabili, sono assai modesti. Come vedremo, la dinamica lineare potrà poi nuovamente differenziare DM e barioni. Ma il punto di partenza per l'evoluzione non lineare è lo stesso.

### 3. Evoluzione post-lineare

La non linearità fa cessare la separatezza dinamica delle componenti con diverso numero d'onda  $k$ . Le equazioni lineari determinano l'andamento temporale di funzioni  $\delta_k(t)$ ; componendole poi opportunamente si risale alla dipendenza di  $\varepsilon(\mathbf{x}, t)$  dalla coordinate spaziali. A livello non lineare, la funzione incognita è direttamente  $\varepsilon(\mathbf{x}, t)$ , e le equazioni coinvolgono derivate parziali, sicché una soluzione generica, in forma analitica, non esiste.

C'è tuttavia un problema che può essere affrontato. Se si considera una sovradensità sferica, *inizialmente* di raggio  $R_i$ , lo studio della sua evoluzione si riduce alla ricerca della funzione  $R(t)$ .

Se al tempo *iniziale*  $t_i$  la densità media nel cosmo è  $\bar{\rho}_i$ , possiamo specificare la densità uniforme all'interno della sfera tramite un parametro di densità  $\Omega_i = \rho_i (< R_i) / \bar{\rho}_i$ . *Supponiamo inoltre l'assenza di perturbazioni nelle velocità*: al tempo  $t_i$  il raggio della sfera  $R(t)$  cresce coerentemente al *flusso di Hubble*.

Prendiamo in esame il caso in cui sia  $\Omega_i > 1$ . In figura IX.3 si mostra l'andamento calcolato per il raggio della sfera, in un modello SCDM, per due diversi valori positivi di  $\Omega_i$ . L'ascissa della figura copre  $\sim 13$  miliardi di anni e uno dei valori di  $\Omega_i$  è scelto in modo da causare il ricollasso completo della sfera proprio nella nostra epoca. L'altro valore è scelto leggermente più grande, sicché l'altra sfera ricollasserebbe quando l'età del cosmo fosse 8-9 miliardi di anni, alla vigilia dell'epoca attuale.

### ***Espressioni parametriche e raggio massimo***

Tracciare le curve in figura IX.3 è relativamente semplice, perché il problema ha una soluzione analitica, in un modello SCDM. Essa è ricavata in appendice A2 in forma parametrica:

$$(IX.5) \quad \frac{R(\theta)}{R_i} = \frac{\Omega_i}{2(\Omega_i - 1)} (1 - \cos \theta), \quad H_i t(\theta) = \frac{\Omega_i}{2(\Omega_i - 1)^{3/2}} (\theta - \sin \theta).$$

L'espressione (IX.5) mostra che  $t$  e  $R$  crescono quando  $\theta$  varia da 0 a  $\pi$ . Il raggio massimo è allora  $R_m = R(\pi)$  e viene raggiunto al tempo  $t_m = t(\pi)$ . Il tempo  $t$  continua poi a crescere con  $\theta$  e per  $\theta = 2\pi$  il raggio  $R$  torna a 0, come si vede in figura. Per il massimo avremo, dunque:

$$(IX.6) \quad R_m = R_i \frac{\Omega_i}{\Omega_i - 1}, \quad H_i^2 \Omega_i t_m^2 = \left(\frac{\pi}{2}\right)^2 \left(\frac{\Omega_i}{\Omega_i - 1}\right)^3 = \left(\frac{\pi}{2}\right)^2 \left(\frac{R_m}{R_i}\right)^3$$

e questa seconda relazione, detta  $\rho_m$  la densità interna alla sfera al tempo  $t_m$ , si può riscrivere nella forma

$$(IX.7) \quad \frac{8\pi G}{3} \Omega_i \bar{\rho}_i \left(\frac{R_i}{R_m}\right)^3 t_m^2 = \frac{8\pi G}{3} \rho_m t_m^2 = \left(\frac{\pi}{2}\right)^2,$$

che va confrontata con la relazione

$$(IX.8) \quad \frac{8\pi G}{3} \bar{\rho}(t_m) t_m^2 = \left(\frac{2}{3}\right)^2$$

derivata dell'equazione di Friedmann, utilizzando il fatto che è  $H = 2/3t$ . Ne segue che

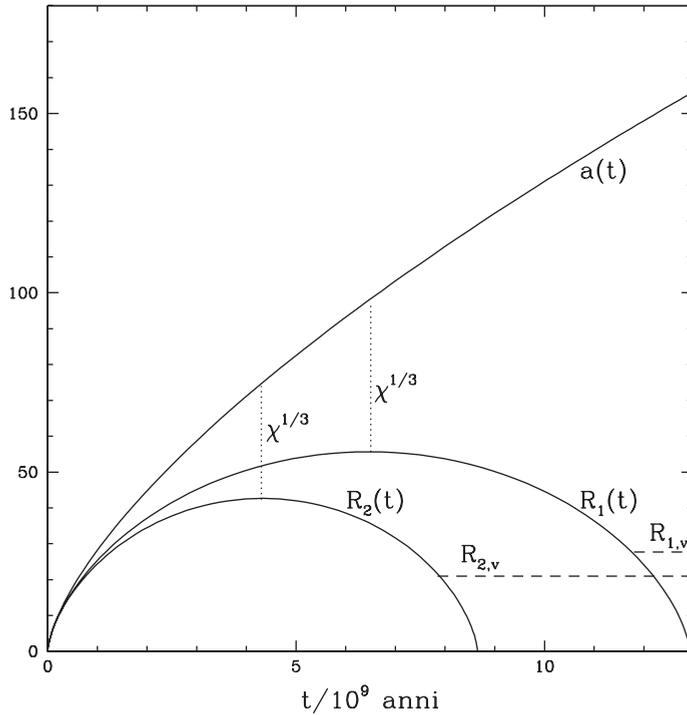


FIGURA IX.3

Evoluzione di sovradensità sferiche in un modello SCDM.  $a(t)$  rappresenta l'andamento del fattore di scala;  $R_1(t)$  e  $R_2(t)$  sono invece i raggi di sfere sovradense, che si sarebbero espanse conformemente a  $a(t)$  se la loro densità fosse stata pari alla densità media del cosmo. Al momento in cui le sfere sovradense raggiungono la massima espansione, il rapporto tra il loro raggio e quello di una sfera eguale, che seguisse il *flusso di Hubble*, è  $\chi^{1/3}$  con  $\chi \simeq 5.6$ , indipendentemente dal livello della sovradensità iniziale. Ciò che differenzia i diversi livelli di sovradensità è solo il tempo a cui la massima espansione viene raggiunta. Il ricollasso successivo tende di norma a arrestarsi quando il raggio si è dimezzato: vedi testo.

il *contrasto di densità*, tra interno della perturbazione e esterno, quando viene raggiunto il massimo raggio, sarà comunque

$$(IX.9) \quad \chi = \frac{\rho_m}{\rho(t_m)} = \left(\frac{3\pi}{4}\right)^2 \simeq 5.6,$$

indipendentemente dal raggio della sfera considerata, dal tempo in cui il massimo viene raggiunto e, quindi, dal valore iniziale  $\Omega_i$ . In figura IX.3 il contrasto di densità è rappresentato dal rapporto tra  $R$  e  $a$ , che avrà valore  $\chi^{1/3}$  per entrambe le sfere, quando esse sono al massimo.

Alternativamente si potrebbe affermare che una fluttuazione sferica, che abbia raggiunto il contrasto di densità 5.6, si trova alla sua massima espansione. Chiaramente ci troviamo ampiamente al di fuori del regime lineare.

### Virializzazione

Crescita e ricollasso, nelle soluzioni in figura IX.3, sono simmetriche, come spesso accade nel moto dei «gravi». Figura e soluzioni (IX.5) attengono comunque il modello SCDM; alla fine del paragrafo discuteremo cosa cambia per  $\Lambda$ CDM.

Non sono invece simmetrici gli effetti di piccole perturbazioni interne alla sfera, quando  $R$  cresce o cala. Durante l'espansione i punti interni alla sfera si allontanano tra loro. Dette  $r_i$  e  $r(t)$  le distanze iniziale e al tempo  $t$ , tra il centro  $O$  e un generico punto  $P_i$  della sfera, il rapporto  $r(t)/r_i$  ha lo stesso andamento di  $R(t)/R_i$ . Ci possono essere piccole perturbazioni locali, rispetto a questo andamento, ma l'espansione provvede a smussarle.

La fase di contrazione, invece, esalta le disomogeneità. Occorrerebbe una sincronizzazione perfetta di tutte le  $r(t)$  perché i moti procedessero coerentemente fino all'annullarsi di  $R$  nel completo ricollasso. Anche piccole disomogeneità, invece, bastano a perturbare l'ordine, dando luogo a moti trasversali. Anche in un sistema prevalentemente costituito da particelle non collisionali di DM, i moti individuali divengono disordinati. Le velocità delle particelle si redistribuiscono e, quando il viriale della sfera

$$(IX.10) \quad Vir = 2E_k + E_p$$

( $E_{k,p}$ : energia cinetica, potenziale) si annulla, la sfera ha raggiunto una configurazione di equilibrio e non si contrae più.

In un modello SCDM, la virializzazione si raggiunge quando il raggio è

$$(IX.11) \quad R_v = R_m/2.$$

Infatti, detta  $M$  la massa della sfera e ponendo  $Vir = 0$  in eq. (IX.10), si ha

$$(IX.12) \quad E_k = -\frac{E_p}{2} = \frac{1}{2}Gf\frac{M^2}{R_v}$$

( $f$  è un fattore numerico che tiene conto della geometria nell'interazione gravitazionale tra punti della sfera) sicché l'energia totale, data da

$$(IX.13) \quad E_k + E_p = -\frac{1}{2}Gf\frac{M^2}{R_v} = -Gf\frac{M^2}{2R_v},$$

deve coincidere con l'energia totale al raggio massimo, che è solo potenziale:

$$(IX.14) \quad E_{p,m} = -Gf\frac{M^2}{R_m}.$$

Il confronto mostra immediatamente che è  $R_m = 2R_v$ .

In figura IX.3 i raggi alla virializzazione sono indicati con linee tratteggiate.

La virializzazione non è un processo istantaneo; le simulazioni numeriche mostrano che essa viene raggiunta mediamente nello stesso tempo che sarebbe occorso per il ricollasso completo della sfera, quindi a un tempo  $t_v = 2t_m$ .

A fronte a una riduzione di un fattore 2 del raggio  $R$  e di una crescita di un fattore 2 nel tempo  $t$ , il fattore di scala cresce di  $2^{2/3}$ . Il rapporto  $a/R$ , quindi, da un valore  $\chi^{1/3}$  passa a un valore  $\chi^{1/3} \times 2^{5/3}$ . Di conseguenza, il contrasto tra la densità interna alla sfera  $\rho_v$  e la densità cosmica media  $\bar{\rho}(t_v)$  sarà

$$(IX.15) \quad \Delta_v = \frac{\rho_v}{\bar{\rho}(t_v)} = \chi \times 2^5 = 2(3\pi)^2 \simeq 178.$$

Questo risultato vale per una sfera di qualsiasi raggio e contrasto di densità iniziale  $\Omega_i$ . Cambiare  $\Omega_i$  significa cambiare il tempo  $t_m$  e il tempo  $t_v = 2t_m$ . Dunque il risultato (IX.15) vale per una sfera che ricollassi a qualsiasi tempo, al momento in cui la virializzazione è raggiunta.

Se ciò accade prima del tempo attuale ( $t_v < t_o$ ) e, dopo la virializzazione, non ci sono ulteriori accrezioni di materia sulla sfera, la data della sua virializzazione ( $t_v$ ) è fissata dal contrasto di densità  $\Delta_o$  osservato a  $t_o$ . Come si vede in figura, l'ulteriore espansione dell'Universo dopo  $t_v$  rende  $\Delta_o > 178$ .

Tutto ciò è vero nell'ambito di un modello SCDM, cioè nel contesto in cui questi risultati sono derivati. Ciononostante il valore 178, spesso arrotondato a 180 o a 200, ha avuto una grande fortuna nello studio delle simulazioni e del materiale osservativo, in cui spesso ci individuano configurazioni con tale contrasto o *eccesso* di densità.

È anche opportuno ricordare che il valore (IX.15) è stato ottenuto considerando la sola gravitazione e trascurando ogni possibile residua azione della radiazione. Una verifica numerica degli effetti residui della radiazione alza il contrasto di densità viriale fino a  $\sim 180$ -181.

### Generalizzazione a modelli $\Lambda$ CDM

La simmetria tra crescita e ricollasso è perduta in modelli in cui diventi importante, da un certo  $z$  in poi, il contributo della DE all'espansione. In un modello  $\Lambda$ CDM, il parametro di densità varia col *redshift*, essendo

$$(IX.16) \quad \Omega_m(z) = \frac{\Omega_{o,m}}{\Omega_{o,m} + (1 - \Omega_{o,m})(1+z)^{-3}}.$$

Il ritmo di espansione risente quindi del fatto che è  $\Omega_{o,m} \neq 1$  e ne risente in modo diverso in funzione di  $z$ . Un sistema raggiunge la massima espansione con contrasto di densità un po' diverso da  $(3\pi/4)^2$  e, soprattutto, raggiunge la virializzazione con contrasto di densità  $\neq 178$ . In più, tali differenze dipendono dal *redshift*: una sfera che virializza a  $z = 1$  raggiunge un contrasto di densità diverso da una sfera che virializza a  $z = 0$ .

Detto  $\Omega_{de} = 1 - \Omega_m$  il parametro di densità della DE al momento della virializzazione, si può dare un'espressione empirica per  $\Delta_v$  in sua funzione,

$$(IX.17) \quad \Delta_v = 18\pi^2 - 82\Omega_{de} - 39\Omega_{de}^2,$$

che vale a qualsiasi *redshift*; vedi anche Fig. IX.4.

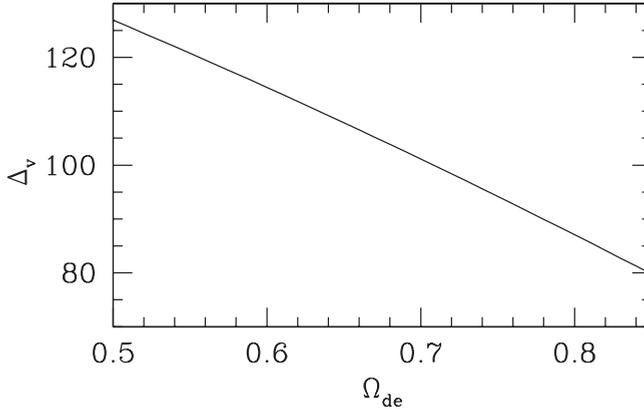


FIGURA IX.4

Contrasto di densità viriale in modelli  $\Lambda$ CDM.

$\Delta_v$  è la densità interna alla sfera rapportata alla densità critica  $\rho_{cr}$ , supposta essere la densità media del cosmo. La densità media della materia nel cosmo è invece  $\Omega_m \rho_{cr}$ . Il rapporto tra densità interna (di materia) e densità esterna di materia è dunque  $\Delta_v / \Omega_m$ .

Ad esempio, per  $\Omega_m = 0.3$ , conformemente a eq. (IX.17) e a figura IX.4, si ha  $\Delta_v = 101$ . Il rapporto tra densità di materia interna e esterna è invece 338.

In figura IX.5, infine, vengono riportati i contrasti di densità viriali, a  $z = 0$ , calcolati numericamente per un alcuni modelli con DE, diversi da  $\Lambda$ CDM; per chiarimenti sulla natura della DE in questi casi, vedi il capitolo XII.

Indipendentemente dal valore di  $\Delta_v$ , comunque, l'esistenza di un *contrasto di densità viriale* implica l'esistenza di un *raggio viriale*  $R_v$  tale che

$$(IX.18) \quad M = \frac{4\pi}{3} R_v^3 \rho_{cr} \Delta_v, \quad \text{ovvero} \quad R_v(M) = (3M/4\pi\rho_{cr}\Delta_v)^{1/3}.$$

Nelle condizioni odierne del cosmo ciò corrisponde a

$$(IX.19) \quad R_v(M) = 2.05 \times 10^{-5} \left( \frac{M/h^{-1}M_\odot}{\Delta_v/100} \right)^{1/3}$$

Ad esempio, in un modello  $\Lambda$ CDM con  $\Delta_v = 100$ , il raggio viriale di una massa di  $4 \times 10^4 h^{-1} M_\odot$  risulta  $\simeq 1.5 h^{-1} \text{Mpc}$ , pari al *raggio di Abell* degli ammassi di galassie (vedi paragrafo 8).

### **Contrasti di densità post-lineari e evoluzione lineare**

La crescita del contrasto di densità, da un valore iniziale  $\Delta_i \equiv \Omega_i = 1 + \delta_i$ , fino a  $\Delta_v$ , può essere confrontata con l'andamento che avrebbe avuto  $\delta_i$  sulla base delle equazioni

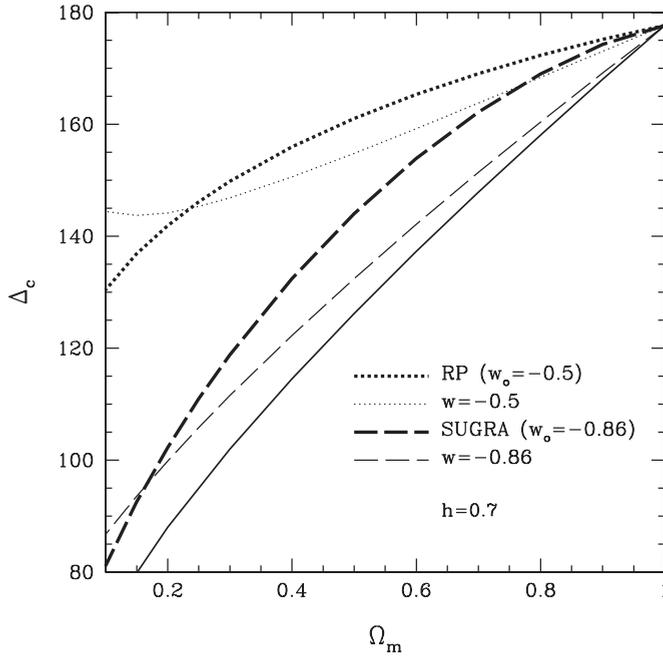


FIGURA IX.5

Contrasto di densità viriale (denominato  $\Delta_c$ ) in modelli con DE in cui  $\Omega_{dm} = 0.3$ ,  $h = 0.7$ . Sono considerati modelli in cui la DE ha parametro di stato  $w = p/\rho = \text{const.}$  con i valori elencati in figura. Inoltre si considerano i modelli con DE dinamica dovuta a potenziali RP e SUGRA (vedi capitolo XII), parametrizzata dal rapporto  $w_0 = p/\rho$  al tempo attuale. Ai fini di raffronto sono riportati anche i valori per  $\Lambda$ CDM (linea compatta).

lineari che, nel caso SCDM, si riconducono alla legge

$$(IX.20) \quad \delta_+(t) = \delta_{i,+}(a/a_i) = \delta_{i,+}(t/t_i)^{2/3},$$

per il modo crescente. Per ipotesi, al tempo  $t_i$  il ritmo di espansione della sovradensità sferica coincide col flusso di Hubble. Per questa ragione, solo i  $3/5$  della fluttuazione iniziale  $\delta_i = 1 - \Omega_i$  sono nel modo crescente (vedi appendice A2).

D'altro canto, sulla base della seconda delle eq. (IX.5) risulta

$$(IX.21) \quad \frac{2t_m}{3t_i} = \frac{\pi\Omega_i}{2\delta_i^{3/2}}$$

e, quindi,

$$(IX.22) \quad \delta_i \left( \frac{t_m}{t_i} \right)^{2/3} = \frac{5}{3} \delta_{i,+} \left( \frac{t_m}{t_i} \right)^{2/3} = \frac{5}{3} \delta_m = \left( \frac{3\pi\Omega_i}{4} \right)^{2/3}.$$

In definitiva avremo che

$$(IX.23) \quad \delta_m \simeq \frac{3}{5} \left( \frac{3\pi}{4} \right)^{2/3} \simeq 1.06$$

trascurando la piccola differenza tra  $\Omega_i$  e 1. Non deve sorprendere di trovare  $\delta_m > 1$ : al tempo  $t_m$  la fluttuazione non è più in regime lineare.

Il contrasto di densità, al tempo  $t_m$ , non è  $\chi_l = 1 + \delta_m \simeq 2.06$  ma  $\chi \simeq 5.6$ . Tuttavia, le equazioni lineari, formalmente, possono sempre essere usate; esse forniscono un valore *fisico* del rapporto  $\delta(t)/\delta_i$  solo se  $\delta(t) \ll 1$ ; altrimenti esse ci danno semplicemente una *mappa* dei valori di  $t$  sui valori di  $\delta$ .

Allo stesso modo, al tempo  $t_v = 2t_m$ , la fluttuazione lineare è cresciuta di un ulteriore fattore  $2^{2/3}$  diventando

$$(IX.24) \quad \delta_v = \frac{3}{5} \left( \frac{3\pi}{2} \right)^{2/3} \simeq 1.68.$$

$\delta_m$  e  $\delta_v$  sono stati calcolati nell'ambito del modello SCDM. A differenza di  $\chi$  e  $\Delta_v$ , essi dipendono solo marginalmente dal modello e si possono utilizzare anche per  $\Lambda$ CDM.

#### 4. Espressioni di Press & Schechter e generalizzazioni

La distribuzione della materia, dopo la ricombinazione, non è certo costituita da sovradensità sferiche. Le strutture sferiche, in astronomia, sono tipiche di sistemi in equilibrio pressorio, la cui superficie coincide con isocore. Le stelle sono sferiche, purché non ruotino troppo velocemente. Nulla del genere è presente alla ricombinazione, quando i materiali cosmici hanno pressioni trascurabili.

A tempo debito, da questa struttura si formano sistemi di dimensioni via via crescenti, tra cui ammassi e gruppi di galassie. Nel tessuto filamentoso della distribuzione cosmica dei materiali, essi si collocano in prossimità delle intersezioni, a costituire specie di nodi, e di fatto, in primissima approssimazione, presentano una simmetria *quasi*-sferica, che si ritiene acquisita in tempi relativamente recenti, nel cosiddetto *rilassamento violento* verso l'equilibrio viriale, seguito all'arresto del collasso gravitazionale.

L'analisi del precedente paragrafo va quindi interpretata a ritroso, a partire dalla configurazione quasi-sferica odierna, risalendo verso tempi antecedenti, in cui le configurazioni potevano essere molto diverse da sfere. Non è tuttavia fuorviante ritenere che il «trend» generale, stabilito per una sovradensità sferica, sia ricalcato in buona misura dalla storia dei materiali destinati a formare una struttura virializzata. In altri termini, dal contrasto di densità misurato in un sistema cosmico (ammasso di galassie, gruppo di galassie, galassia) possiamo pensare di risalire alla densità media del cosmo al momento della sua virializzazione e quindi al *redshift* in cui esso si è formato. Di qui possiamo ricostruire la sua storia precedente, fino alla fluttuazione  $\delta_i$  al tempo  $t_i$  (poco dopo la ricombinazione, ancora in fase lineare) destinata a evolvere in quel sistema.

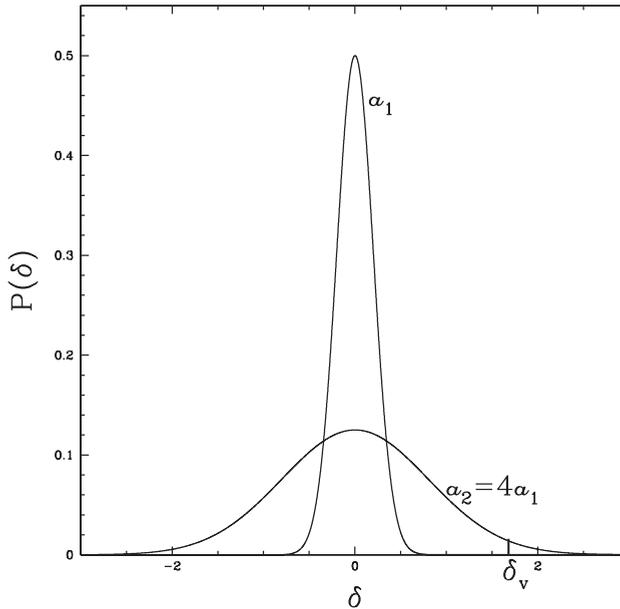


FIGURA IX.6

Evoluzione cosmologica della distribuzione delle ampiezze, se regolata dalle equazioni lineari. Quando il fattore di scala raggiunge il valore  $a_2$  (tempo  $t_2$ ), una frazione piccola ma significativa di fluttuazioni supera (linearmente) il valore critico  $\delta_v = 1.68$ . Essa è data dall'area del piccolo triangoloide tra curva e ascissa, alla destra di  $\delta_v$ , rapportata a tutta l'area tra curva e ascissa. Tra le fluttuazioni sulla scala considerata, questa è la frazione che ha formato oggetti virializzati al tempo  $t_2$ .

Un passo ulteriore ci libererà dalla scelta di un particolare tempo  $t_i$ : invece di dare  $\delta_i$  a un  $t_i$  convenzionale, diamo il valore che avrebbe assunto oggi l'ampiezza della fluttuazione *se avesse seguito il trend lineare*.

In questo modo, se cerchiamo di sapere quanti oggetti di una determinata massa  $M$  si possono formare, ci basterà conoscere la probabilità che ci siano fluttuazioni di ampiezza  $\delta(t_0) > \delta_v$  sulla scala di massa  $M$ .

Questa prassi reca con sé un ulteriore vantaggio: L'evoluzione lineare, agendo con un uguale fattore su tutte le ampiezze, ne preserva la distribuzione Gaussiana iniziale. Ricordiamo che il valore medio delle ampiezze è inizialmente nullo, mentre il loro valore quadratico medio è dato dallo spettro trasmesso. Col trascorrere del tempo, la Gaussiana semplicemente si dilata, come è mostrato in figura IX.6, col fattore di scala che passa da un valore  $a_1$  al valore  $a_2 = 4a_1$ , una crescita limitata per ragioni grafiche.

In figura IX.6 è anche indicato il valore  $\delta_v = 1.68$ . L'area del piccolo triangoloide, tra Gaussiana e ascissa a destra di  $\delta_v$ , rapportata all'area totale tra Gaussiana e ascissa, rappresenta la frazione  $f(t_2, M)$  delle fluttuazioni su scala  $M$  che ha superato  $\delta_v$  nell'evoluzione lineare, al tempo  $t_2$  quando il fattore di scala è  $a_2$ .

Su questa base si può costruire la *funzione di massa* differenziale di Press &

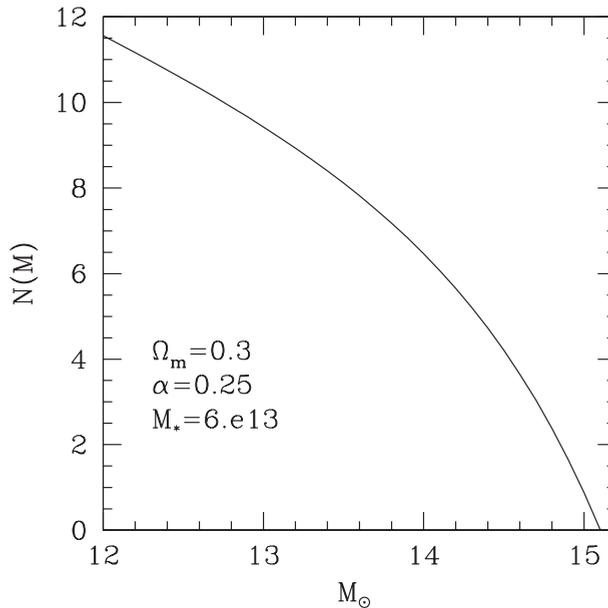


FIGURA IX.7

Funzione di massa PS. Ascissa e ordinata sono logaritmiche.

Schechter (PS), usando lo spettro delle fluttuazioni trasmesse, come è discusso in appendice A3. Nel caso in cui la varianza di massa abbia la forma

$$(IX.25) \quad \sigma_M = (M/M_o)^\alpha,$$

detta  $M_* = M_o(\sqrt{2}/\delta_v)^{1/\alpha}$ , la funzione di massa PS è

$$(IX.26) \quad n(M) = \frac{2\alpha}{\sqrt{\pi}} \frac{\bar{\rho}_M}{M_*^2} e^{-(M/M_*)^{2\alpha}}$$

e il suo andamento è mostrato in figura IX.7, per la scelta di parametri nel riquadro. Ricordiamo che, per definizione di funzione di massa, il numero di sistemi virializzati nell'unità di volume, con massa compresa tra  $M$  e  $M + dM$  sarà  $n(M) dM$ ; il numero di sistemi virializzati con massa  $> \bar{M}$  è quindi espresso dall'area del triangoloide tra la curva e l'ascissa, a destra del valore  $\bar{M}$ .

La funzione di massa PS è stata successivamente perfezionata, tenendo conto in maniera parametrica della non-sfericità del collasso. Per questi sviluppi, che hanno valenza più tecnica e sono di natura differenziale, rimandiamo alla letteratura specializzata.

## 5. Espressioni di Press & Schechter e dati cosmologici

Le funzioni di massa teoriche vanno ora confrontate con risultati di simulazioni e, soprattutto, coi dati.

Sottolineiamo quindi anzitutto che l'efficacia delle espressioni PS è confermata dal confronto coi risultati delle simulazioni, di cui trattiamo nei prossimi paragrafi (vedi in particolare figura IX.10).

Ricavare funzioni di massa dai dati, invece, è cosa abbastanza complessa. Un primo raffronto si può fare tra funzione di massa e *funzione di luminosità*. Esso mostra similarità a grandi linee: le funzioni di massa e di luminosità sono entrambe caratterizzate da una discesa a legge di potenza tagliata poi da un'esponenziale. I dettagli sono però difformi e si può anzi dire che l'uso delle espressioni PS è una comprova che il rapporto  $M/L$  (massa/luminosità) ha un andamento regolare, ma *non* è costante.

Opportune tecniche permettono comunque di ricostruire anche funzioni di massa osservative. Discuteremo quindi l'informazione sul modello cosmologico che può essere ottenuta dall'uso di espressioni tipo PS.

### **PS e dati a $z = 0$**

Gli andamenti (IX.25)-(IX.26) vengono confrontati coi dati sugli ammassi di galassie, deducendo dall'espressione (IX.26) la *funzione di massa integrale*  $N(>M)$ , che dà il numero di ammassi con massa  $> M$ .

Per l'espressione di  $\sigma_M$  si può spesso usare l'andamento previsto dalla teoria lineare. Difatti, nel modello delle concordanze, la lunghezza comovente  $R_8 = 8 h^{-1} \text{Mpc}$  include una scala di massa  $M_8 \simeq 1.8 \times 10^{14} h^{-1} M_\odot$ . Come abbiamo visto nel capitolo II, la varianza  $\sigma_8 \equiv \sigma_{M_8}$ , calcolata sulla distribuzione delle galassie, è  $\sim 1$ , mentre la varianza della distribuzione di materia sulla stessa scala è l'inverso del *bias*  $b$ , il cui valore si stima  $\sim 1.6-1.8$ . Su scale di massa  $\gtrsim 1.8 \times 10^{14} h^{-1} M_\odot$ , perciò, è accettabile riferirsi ai risultati della teoria lineare, come ora comunque faremo.

Un utile punto di riferimento osservativo è costituito allora dal numero di ammassi in  $(100 h^{-1} \text{Mpc})^3$  con  $M > \bar{M} \equiv 4 \times 10^{14} h^{-1} M_\odot$ , a  $z \simeq 0$ , che sono  $\sim 3 \pm 1$ .

L'esponente  $\alpha$  da usare nell'espressione (IX.25) può dunque essere dedotto dalla funzione di trasferimento. Si vede così che esso varia a sua volta in funzione della massa, ma abbastanza lentamente da rendere valido l'uso delle espressioni (IX.25)-(IX.26); esse, comunque, possono essere ulteriormente perfezionate per tener conto fino in fondo di tale variabilità.

Vi sono tre modi in cui la funzione di trasferimento agisce sui risultati:

- 1) tramite il valore di  $\delta_v$ ;
- 2) tramite l'andamento di  $\alpha$ ;
- 3) tramite il valore della scala di massa  $M_o$  tale che  $\sigma_{M_o} = 1$  (per la distribuzione di *background*).

Abbiamo però già rilevato che  $\delta_v$  ha una scarsa dipendenza dal modello e dal *redshift*. Il valore di  $\alpha$ , poi, risente del valore dell'indice spettrale primordiale  $n_s$ , ma non risente significativamente degli altri parametri dei modelli. Ciò che è invece

fortemente significativo è la normalizzazione di  $\sigma_M$ . Variando  $M_o$  si ottengono valori anche molto diversi nel numero di ammassi di galassie.

La normalizzazione di  $\sigma_M$ , d'altronde, riporta essenzialmente all'ampiezza dello spettro primordiale, determinata dalla costante  $A$  nell'espressione  $P(k) = A k^{n_s}$ , e alla sua elaborazione tramite la funzione di trasferimento.

Qualunque modello ragionevole deve essere in accordo col numero di ammassi osservato a massa  $> \bar{M}$ ; questo requisito può comunque essere soddisfatto variando la costante  $A$ . Ciò però non è senza conseguenze:  $A$  fissa anche la normalizzazione delle fluttuazioni del CMB. L'accordo tra ampiezza di tali fluttuazioni e numero di ammassi con massa  $> \bar{M}$  è uno dei primi test cui deve soddisfare qualunque modello.

Viceversa, per le ragioni sopra esposte, l'andamento con  $M$  della funzione  $N(> M)$ , in genere, non distingue tra di loro i modelli.

### **PS e dati a alto $z$**

Un altro test significativo è invece costituito dall'andamento di  $N(> M)$  in funzione del *redshift*  $z$ . Tale andamento rende conto dell'evoluzione con  $z$  della varianza di massa  $\sigma_M$  e, quindi, dell'ampiezza delle fluttuazioni. (Si noti, in particolare, che l'uso delle espressioni lineari è sempre meglio giustificato, al crescere di  $z$ ).

In un modello SCDM, ad esempio, le fluttuazioni crescono con legge  $\delta \propto a$  (fattore di scala), su tutte le scale di massa  $M$  e a tutti i *redshift*  $z$ . In modelli  $\Lambda$ CDM, invece, non appena la DE diventa una componente significativa della densità, la crescita delle fluttuazioni rallenta, per bloccarsi quasi del tutto quando la densità della DE diviene dominante.

In modelli SCDM, quindi, il numero di ammassi cresce rapidamente al crescere di  $a$  sicché, osservativamente, il loro numero dovrebbe calare rapidamente al crescere del *redshift*. Si prospetta inoltre una situazione di *auto-similarità*: le funzioni di massa a  $z_1$  e  $z_2$  sono previste essere assai simili, tranne che per uno spostamento sull'asse delle masse, che tiene conto di come evolve con  $z$  la massa  $M_o$ ; tale spostamento, in questo caso, dipende esclusivamente dal rapporto  $(1 + z_1)/(1 + z_2) = a_2/a_1$ .

Un modelli con DE, invece, la crescita nel numero di ammassi è molto rallentata in tempi recenti. Gli ammassi osservati sono dunque più *vecchi*, si sono necessariamente formati prima.

Ciò ha due chiari risvolti osservativi: (i) L'osservazione del numero di ammassi al variare del *redshift* è un test diretto del modello. (ii) L'ampiezza  $A$  delle fluttuazioni primordiali deve necessariamente essere maggiore, in presenza di DE, e ciò deve accordarsi coi dati sul CMB.

Tramite l'osservazione del numero di ammassi a alto *redshift* si può mirare a stimare l'andamento del rapporto  $\rho_{DE}/\rho_m$ , tra densità della DE e della materia, in funzione di  $z$ . L'andamento del rapporto, previsto dal modello  $\Lambda$ CDM è semplice, dal momento che  $\rho_{DE}$  è costante e  $\rho_m \propto (1 + z)^3$  (vedi anche eq. (IX.16)).

I dati attualmente disponibili sull'andamento del numero di ammassi con  $z$  sono coerenti col modello delle concordanze. Si mira a raffinarli in modo da verificare con maggior esattezza l'andamento di  $\rho_{DE}$  in funzione di  $z$ .

## 6. L'approssimazione di Zel'dovich

Nello studio dell'evoluzione di una sovradensità sferica, il problema non-lineare viene semplificato e reso integrabile (in una cosmologia SCDM), grazie alla riduzione dei gradi di libertà spaziali a un'unica variabile, il raggio  $R$  della sovradensità. L'analisi delle fasi non lineari dello sviluppo delle disomogeneità è però andata ben oltre il caso della sovradensità sferica. Di fatto, tuttavia, le tecniche analitiche cedono gradualmente il passo alle tecniche numeriche.

Le condizioni iniziali di queste ultime si costruiscono tramite l'*approssimazione di Zel'dovich*. Tale approssimazione ha peraltro un notevole significato concettuale di per sé. Essa si basa sul passaggio da una visuale euleriana a una visuale *lagrangiana*: invece di usare le  $\varepsilon(\mathbf{x})$  come insieme continuo di coordinate (in questo contesto,  $\mathbf{x}$  è un *indice continuo* che etichetta le coordinate), si usano come coordinate le *posizioni iniziali* degli elementi di volume, eventualmente «campionandoli» tramite un insieme di punti.

Se si parte da una condizione iniziale di (quasi) omogeneità e si campiona la distribuzione di materia tramite un ampio insieme di punti (distribuiti a caso, ma mediamente equispaziati), le loro traiettorie indicano plasticamente l'evoluzione della densità, la nascita di concentrazioni, il graduale formarsi delle strutture.

Questo quadro è comune all'approssimazione di Zel'dovich e alle tecniche numeriche. Il primo caso, tuttavia, è relativamente semplice e tende a spingere fino agli estremi limiti di applicabilità la teoria perturbativa. In linea di principio, un approccio perturbativo vale sinché le fluttuazioni di densità hanno un'ampiezza quadratica media  $\sigma_R^2 = \langle \varepsilon_R^2(\mathbf{x}) \rangle \ll 1$ , sulle scale  $R$  rilevanti. L'approssimazione di Zel'dovich, invece, è così ben congeniata da rendere conto del quadro fisico, spesso anche oltre questo limite di applicabilità.

L'approssimazione di Zel'dovich si applica, comunque, ai materiali non-relativistici e sotto l'ipotesi che essi siano l'unica sostanza coinvolta nella crescita delle fluttuazioni. Essa è quindi particolarmente adatta allo studio dell'evoluzione delle disomogeneità nei barioni e nella CDM, dopo la ricombinazione.

Il passaggio critico che riesce a estendere le potenzialità dell'approccio perturbativo è l'osservazione che la crescita della densità, punto per punto, non dà luogo a strutture sferoidali, ma a strutture fortemente asimmetriche. Se vogliamo, ciò è in contrasto con il rilievo dato sin qui alla crescita di una sovradensità sferica, che però viene sostanzialmente «usata» solo per stabilire una tempistica di riferimento.

Consideriamo un determinato punto  $\mathbf{x} = a\mathbf{r}$ , nel cui intorno sussiste inizialmente una tenue sovradensità, destinata poi a svilupparsi fino a dar luogo a strutture cosmiche. L'evoluzione perturbativa della densità, attorno a  $\mathbf{x}$ , è consistente con l'e-

spressione

$$(IX.27) \quad \rho(\mathbf{r})a^3[1 + \alpha_1 G(a)][1 + \alpha_2 G(a)][1 + \alpha_3 G(a)] = \\ = \rho_{in}(\mathbf{r})a_{in}^3[1 + \alpha_1 G(a_{in})][1 + \alpha_2 G(a_{in})][1 + \alpha_3 G(a_{in})]$$

dove i coefficienti  $\alpha_i$ , nel caso di *sovradensità* iniziale, saranno negativi (vedi appendice A4). Gli indici  $_{in}$  si riferiscono a un determinato tempo iniziale;  $a(t)$  è il fattore di scala,  $G(a)$  definisce il ritmo di crescita della fluttuazione, in funzione del fattore di scala, usato come coordinata temporale.  $G(a)$  viene ricavato dalla teoria perturbativa, sulla base della legge di proporzionalità:  $\varepsilon(\mathbf{r}, t)/\varepsilon(\mathbf{r}, t_{in}) = G[a(t)]/G[a(t_{in})]$ . In un modello SCDM (o quando la cosmologia si può approssimare con SCDM), si può scegliere semplicemente  $G(a) = a$ .

Il significato dell'espressione (IX.27) è trasparente. Se i coefficienti  $\alpha_i$  si annullano, essa si riduce alla relazione che esprime la conservazione della massa. Per  $\alpha_i \neq 0$ , essa esprime una crescita di  $\rho$ , diversa lungo tre assi ortogonali e ciascun fattore  $1 + \alpha_i G$  rende conto della crescita lungo l'asse  $i$ -esimo. La direzione degli assi non è esplicitata e sarà diversa punto per punto, a seconda di come è inizialmente distribuita la materia (in appendice A4 si discute della loro determinazione).

Se  $\rho_B$  è la densità media (di *background*), in generale sarà

$$(IX.28) \quad \rho(\mathbf{r}) = \rho_B + \delta\rho(\mathbf{r}) \quad \text{con} \quad \delta\rho(\mathbf{r})/\rho_B = \varepsilon(\mathbf{r})$$

e, fin tanto che è  $|\varepsilon(\mathbf{r})| \ll 1$ , da eq. (IX.27) si deriva facilmente che è  $\varepsilon(\mathbf{r}) = -G \sum_{i=1}^3 \alpha_i$ .

Una volta derivati i coefficienti  $\alpha_i$  e la direzione degli assi, dalla distribuzione iniziale della materia, si può conoscere il modo in cui la distribuzione stessa si deforma, al trascorrere del tempo.

In un modello SCDM, in cui è  $G = a$ , risulta

$$(IX.29) \quad \rho(\mathbf{r})[a + \alpha_1 a^2][a + \alpha_2 a^2][a + \alpha_3 a^2] = \text{const.}$$

In ciascuna direzione, cioè per ciascun fattore, si possono definire allora un istante di *turnaround* e un istante di *shell crossing*. Il turnaround significa che il coefficiente  $[a + \alpha_i a^2]$  ( $i = 1, 2, 3$ ) passa da crescente a decrescente; ciò avviene per  $a = -1/2\alpha_i$  (ricordiamo:  $\alpha_i < 0$ ). Fintanto che il coefficiente è crescente, esso induce un calo nella densità  $\rho(\mathbf{r})$ , che sta comunque decrescendo anche se aumenta il contrasto con la densità media  $\rho_B$ .

Se si assume simmetria sferica, il turnaround è fuori dal regime perturbativo, quando il contrasto di densità è  $\sim 5.6$ . Consideriamo invece il caso limite in cui sia  $\alpha_1 \neq 0$  e  $\alpha_2 = \alpha_3 = 0$ . Al turnaround lungo l'asse 1 si avrebbe  $a + \alpha_1 a^2 = a/2$  e, quindi, il contrasto di densità complessivo sarebbe  $\simeq 2$ .

Pur essendo minore, tale contrasto di densità è comunque già al di fuori del regime perturbativo e, tuttavia, il confronto con le simulazioni n-body, di cui discuteremo al prossimo paragrafo, mostra che l'andamento così trovato è assai prossimo a quello corretto.

La validità dell'approssimazione si estende addirittura a contrasti di densità maggiori, fino all'approssimarsi del *shell crossing*, che si avrebbe, in questo caso, per  $a = -1/\alpha_1$  e che corrisponde al divergere della densità.

Non è difficile rendersi conto della ragione per cui l'approssimazione di Zel'dovich estende la validità dei risultati perturbativi. La chiave del successo sta nell'indagine geometrica sulla distribuzione iniziale della materia e nell'individuazione degli assi di simmetria con le rispettive deformazioni. Se si usa l'approssimazione perturbativa assumendo simmetria sferica ci si allontana rapidamente dall'evoluzione reale; non perché l'andamento del contrasto di densità sia brutalmente errato, ma perché il suo effetto spaziale non viene ben calibrato.

Al di là dell'esempio in cui 2 dei 3 coefficienti  $\alpha_i$  si annullano, va detto che, in generale, turnaround e *shell crossing* avverranno lungo un particolare asse prima che lungo gli altri.

Le prime strutture che vanno a formarsi avranno dunque tipicamente una forma piatta, detta *pancake* (letteralmente «frittella»).

Ma l'approssimazione di Zel'dovich rende conto della fisica anche al di là del primo *shell crossing*. Quando esso è avvenuto, spesso si viene a stabilizzare una struttura a pancake, non molto difforme dall'immagine della superficie su cui ha avuto luogo lo *shell crossing*. Prima dello sviluppo estensivo delle tecniche numeriche, erano state formulate approssimazioni atte a schematizzare queste formazioni (in particolare, il *modello di adesione*). Nelle altre due direzioni spaziali, l'evoluzione sostanzialmente procede, guidata dagli altri coefficienti [ $a + \alpha_i a^2$ ]. Quindi, dopo la formazione di una struttura superficiale bidimensionale, si giungerà alla formazione di una struttura filamentosa unidimensionale, e infine di nodi (*knots*) ai punti di giuntura dei filamenti. Il reticolo delle strutture spaziali osservate, o ottenute tramite simulazioni, non differisce di molto da questo schema, con gli ammassi di galassie ai punti di congiunzione di filamenti, tra i quali si hanno tenui strutture bidimensionali. Si veda, a titolo d'esempio, la proiezione bidimensionale di una fetta, profonda  $40 h^{-1} \text{Mpc}$  ricavata da una simulazione in un *box* con  $100 h^{-1} \text{Mpc}$  di lato (Fig. IX.8).

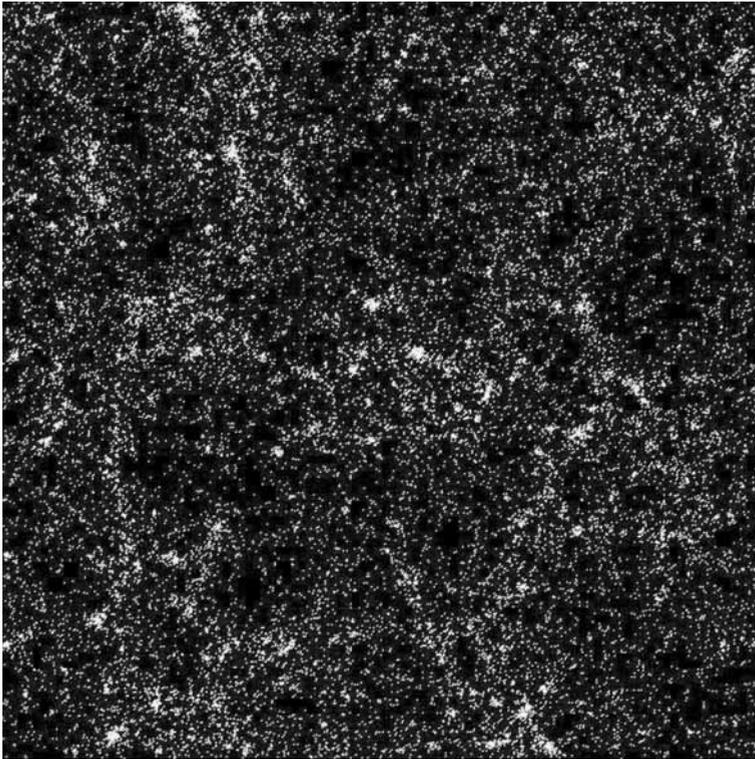
In appendice A4 viene sviluppata l'analisi lagrangiana conseguente all'approssimazione di Zel'dovich. Le coordinate comoventi di ciascun punto, al tempo  $t_{in}$ , vengono denominate  $\mathbf{q}$  e sono le coordinate lagrangiane del sistema.

La posizione di ciascun punto evolve nel tempo secondo la legge

$$(IX.30) \quad \mathbf{x}(t) = a(t)[\mathbf{q} + G(t)\mathbf{s}(\mathbf{q})]$$

dove il vettore  $\mathbf{s}$  si può ricavare come gradiente di un potenziale scalare  $\psi(\mathbf{q})$  (vedi appendice A4) e descrive la distribuzione della materia all'istante iniziale, come si ricava dal campo di velocità, tramite un'equazione di continuità. Definiamo infatti il campo di velocità comoventi  $\mathbf{u} = d\mathbf{r}/dt$  tali che

$$(IX.31) \quad a\mathbf{u} = \mathbf{v} - H\mathbf{x}$$

**FIGURA IX.8**

Filamenti e nodi in una proiezione bidimensionale di una fetta di una simulazione N-body. (si veda l'appendice TAVOLE A COLORI).

(alle velocità si toglie il contributo del flusso di Hubble e si divide il risultato per il fattore di scala).

In una cosmologia SCDM o, comunque, quando la deviazione del modello da SCDM è trascurabile, il campo di velocità comoventi iniziale è dato da

$$(IX.32) \quad \mathbf{u} = aH\mathbf{s},$$

sicché la distribuzione degli assi, rispetto cui vale l'espressione (IX.27), si ricava dal campo di velocità iniziali. In appendice A4, si descrive anche come si possano ricavare gli assi di simmetria e i coefficienti  $\alpha_i$ .

## 7. Simulazioni N-body

Il passaggio dal regime lineare al regime non-lineare richiede anzitutto l'uso di una visuale lagrangiana. Inoltre, non si studia più il generico andamento di un campo di fluttuazioni con determinate caratteristiche (magari distinte per le diverse componenti, come barioni e DM), ma l'evoluzione di una sua *realizzazione* tramite un numero

$N$  di punti o corpi (*body*), possibilmente grande, ma comunque finito. Sfruttando l'enorme potenzialità di calcolo dei moderni *computers*, si simula numericamente l'azione delle forze gravitazionali tra gli  $N$  corpi materiali, realizzando in questo modo una *simulazione N-body*.

La difficoltà principale, nell'attuazione di simulazioni, sta nel raggio d'azione delle forze gravitazionali: una massa sita in un punto  $P$ , a distanza  $r$  dà luogo a una forza  $\propto 1/r^2$ ; ma il flusso di tale forza attraverso una superficie sferica centrata in  $P$ , la cui area cresce in modo  $\propto r^2$ , è costante. Il moto di uno dei corpi della simulazione è quindi determinato dall'azione di tutti gli altri punti.

### Codici diretti

La prima tecnica usata per studiare l'evoluzione di un sistema di  $N$  corpi sfruttava il computer per calcolare le forze

$$(IX.33) \quad \mathbf{f}_i = -Gm_i \sum_{j \neq i} \hat{\mathbf{r}}_{ij} \frac{m_j}{r_{ij}^2}$$

con cui  $N - 1$  corpi agivano sull' $i$ -esimo corpo, via via che la posizione di ciascuno degli  $N$  corpi mutava ( $m_i$ : masse dei corpi;  $\mathbf{r}_{ij}$ : vettore con estremi nei corpi  $i$ -esimo e  $j$ -esimo;  $r_{ij}$ : suo modulo;  $\hat{\mathbf{r}}_{ij}$ : suo versore). L'equazione

$$(IX.34) \quad \mathbf{f}_i dt = m_i d\mathbf{v}_i$$

veniva poi usata per determinare le *variazioni* della velocità di ciascun punto e quindi il suo spostamento, con la velocità *variata*. Considerando intervalli temporali  $dt$  piccoli, il metodo è di sicura efficacia.

Il tempo-macchina  $T_m$ , necessario a calcolare  $\mathbf{f}_i$  tramite l'eq. (IX.33), è  $\propto N^2$ , mentre il numero di eq. (IX.34) da usare è  $\propto N$  e quindi trascurabile. Complessivamente, allora, è  $T_m \propto N^2$ .

Negli anni Settanta, quando il metodo fu tradotto in equazioni numeriche, principalmente da Sverre Aarseth e collaboratori, si riusciva a utilizzarlo per l'evoluzione di sistemi di centinaia di corpi. Per  $\sim 1000$  corpi,  $T_m$  raggiungeva svariate settimane. Le macchine oggi in uso hanno una velocità di calcolo da 1000 a 10000 volte maggiore, ma anche tutto ciò non permette di andare molto oltre  $10^5$  corpi.

### Codici PM e loro perfezionamenti

Il passo avanti determinante venne compiuto grazie all'implementazione di tecniche FFT (*fast-Fourier-transform*: trasformate di Fourier veloci).

Si immaginano gli  $N$  corpi distribuiti in un cubo di lato  $L$  e si riguarda il cubo come elemento di una distribuzione periodica sui 3 assi coordinati. Ciò implica che le fluttuazioni di densità nel cubo avranno componenti discrete: le componenti di  $\mathbf{k}$  da considerare hanno valore  $(2\pi/L)/n$  ( $n = 1, 2, \dots$ ).

Si fissa poi una *risoluzione spaziale*, considerando un valore massimo di  $n$ , per cui si trascurano le componenti di Fourier con  $n > n_{max}$ .

La difficoltà di questo approccio sta nel fatto che le equazioni che ricavano le forze  $f_i$  dalle  $\epsilon(\mathbf{x})$  sono equazioni differenziali alle derivate parziali. È qui che intervengono le FFT, trasformando queste equazioni in equazioni algebriche e permettendo di calcolare algebricamente le forze  $f_i$  dalle  $n_{max}^3$  componenti di Fourier  $\delta_{n_1, n_2, n_3}$  delle fluttuazioni di densità.

Il tempo-macchina di questa tecnica, all'incirca, è  $\propto N \ln N$ ; si possono quindi seguire le evoluzioni di sistemi con numeri di punti assai grandi. Le FFT sono state inizialmente implementate per  $n_{max} = 2^u$ , con  $u$  intero, ma oggi sono disponibili per  $n_{max}$  ottenibili dal prodotto di qualunque combinazione di interi. Taluni importanti algoritmi non hanno comunque ritenuto utile adeguarsi a quest'ultima possibilità e usano tuttora solo i multipli di 2. In questo caso, se il lato del cubo è  $L$ , la *risoluzione nominale* della simulazione sarà  $L/2^u$ . Attualmente si realizzano facilmente simulazioni con  $u = 8$ , ma anche simulazioni con  $n_{max} = 512$  o 1024 non comportano difficoltà insormontabili.

I codici PM sin qui descritti hanno dei limiti operativi, che sono stati superati con tecniche P3M o PM-adattivo, per le quali rinviamo a trattati specializzati.

In generale, una simulazione  $N$ -body è definita quando si danno:

- 1) Il lato  $L$  del cubo entro cui si effettua la simulazione .
- 2) La risoluzione nominale; la scala fisicamente risolta sarà 2-3 volte superiore a quella indicata dalla risoluzione nominale. Il rapporto tra  $L$  e scala risolta si definisce *range dinamico* della simulazione.
- 3) Il numero delle particelle o «corpi» considerati.
- 4) I parametri del modello cosmologico considerato.
- 5) Il *redshift*  $z_{in}$  al quale si inizia la simulazione.

In generale, a  $z_{in}$  il campo di fluttuazioni sarà lineare. Le condizioni iniziali si ricavano dando coordinate e velocità degli  $N$  punti, come desunte in approssimazione di Zel'dovich.

La distribuzione degli  $N$  corpi sarà ricavata a un certo numero di *redshift* intermedi, tra  $z_{in}$  e  $z = 0$ .

### Codici a albero

Una tecnica alternativa è costituita dai codici a albero. Nella sua forma più semplice essa si basa sulla divisione ricorsiva dello spazio in celle, a partire da celle di lato  $D$  (livello 1), che viene comunque arrestata quando una cella ha lato  $D/n$  (livello  $n$ ) per le celle vuote o contenenti una sola particella.

La forza agente sulla  $i$ -esima particella viene allora determinata come in eq. (IX.33), ma associando tra loro insiemi di particelle  $j$ -esime, quando la loro distanza angolare è piccola. Quantitativamente, si fissa un *angolo limite*  $\bar{\theta}$  e si esplorano tutte le celle *popolate*, partendo dal livello 1 (celle di lato  $D$ ). Se la particella  $i$ -esima le «vede»

sotto un angolo  $\theta < \bar{\theta}$  (ciò avverrà per celle «distanti») la forza agente sulla particella *i*-esima si ottiene attribuendo la massa contenuta nella cella al centro di massa delle particelle ivi collocate.

Per le celle «viste» sotto angoli  $\theta > \bar{\theta}$  si passa al livello successivo e si continua così, fin tanto che si giunge al grado di suddivisione adeguato a individuare qualunque particella in una cella «vista» sotto un angolo  $< \bar{\theta}$ .

In sostanza, si opera come in un codice diretto, sulla base di un'equazione di tipo (IX.33), ma si stabilisce la forza originantesi dalle altre particelle raggruppandole in sottoinsiemi sempre più numerosi, via via che si va più lontano. (Si noti che, per  $\bar{\theta} = 0$  questo tipo di codice ridiverrebbe un codice diretto.)

La realizzazione pratica di questi codici porta a tempi macchina e risoluzioni del tutto confrontabili a quelle dei più evoluti codici PM adattivi, con tempi macchina ancora  $\propto N \ln N$ .

## 8. Simulazioni idrodinamiche

Nelle simulazioni *N*-body non si fa differenza tra componente barionica e DM. Il comportamento dei barioni, tuttavia, può essere assai diverso dalla DM, per una serie di ragioni fisiche:

- 1) I materiali barionici costituiscono un fluido.
- 2) I materiali barionici possono riscaldarsi e/o emettere radiazione.
- 3) Di conseguenza essi possono formare sistemi collassati come stelle, buchi neri, ecc.
- 4) Tali stelle possono evolvere e dar luogo a Supernovae, lasciando *remnants* come stelle a neutroni.
- 5) La dinamica dei buchi neri può essere rilevante per la distribuzione dei materiali: da essi si originano getti e *input* di entropia nei sistemi.

Più specificamente ricordiamo che ogni galassia ha almeno un buco nero centrale; che i nuclei galattici attivi (AGN: *active galactic nuclei*) hanno come sorgente energetica l'energia potenziale gravitazionale rispetto a un grosso buco nero centrale; che gli AGN più potenti hanno esaurito i materiali circostanti il buco nero prima della nostra epoca cosmica e sono attualmente quiescenti. Essi hanno dato luogo ai QSO (*quasi-stellar-objects*: oggetti quasi stellari), protagonisti di epoche cosmiche passate. Le proprietà dei sistemi cosmici attuali, tuttavia, possono essere state determinate da fenomeni oggi esauriti.

Sotto il nome di «simulazioni idrodinamiche» vanno tutte le simulazioni che cercano di tenere conto in parte o complessivamente dei fatti fisici suelencati, specialmente al fine di stabilire quali siano i più rilevanti, riuscendo così anche a modelarne la dinamica. L'obiettivo di queste simulazioni è quindi prevalentemente *astrofisico*. Esse mirano a comprendere i meccanismi di formazione degli oggetti celesti, piuttosto che a limitare i parametri della cosmologia e/o, attraverso a essi, la fisica delle alte energie.

L'elemento di base è comunque tenere conto che i barioni sono un fluido e quindi debbono soddisfare le equazioni dell'idrodinamica, le stesse che abbiamo visto nel capitolo VII per ricavare la massa di Jeans.

Si tratta di esprimere tale dinamica in termini numerici e vi sono due classi di metodi, *euleriani* o *lagrangiani*, in cui tale dinamica viene tradotta.

I metodi euleriani si basano su criteri iniziali simili a quelli dei codici PM, lo spazio viene diviso in celle a cui sono associate una densità  $\rho$ , un campo di velocità  $v$  e un'energia  $U$ . Vengono valutati i flussi di tali quantità tra una cella e quelle circostanti e i valori delle quantità vengono aggiornati lungo passi temporali finiti. Come nei codici PM adattativi, il programma si autoregola ai fini di raggiungere una determinata risoluzione.

Questi metodi sono particolarmente adatti a catturare la dinamica del fluido in presenza di *shock*. Dal punto di vista tecnico questo tipo di metodi meglio si adatta a essere usato in concomitanza con codici PM.

I metodi lagrangiani si basano su una descrizione del gas tramite particelle. A ogni particella è associata una dimensione fisica finita e un "profilo". La densità e altri parametri termodinamici pertinenti ciascuna particella vengono a variare nel tempo e sono utilizzati per risolvere le equazioni idrodinamiche, discretizzate tenendo conto del profilo assegnato.

Dal punto di vista tecnico, questo tipo di codici idrodinamici ben si adatta a completare l'analisi basata su codici a albero, di cui condivide il vantaggio-base, la capacità intrinseca di raggiungere maggior risoluzione in corrispondenza dei siti ove la densità è maggiore.

Tra gli effetti legati alla natura dei barioni, quelli derivanti dalle equazioni della idrodinamica sono gli unici per cui esiste una potenza di calcolo adeguata a una risoluzione diretta di simulazioni cosmologiche.

Gli altri effetti suelencati necessitano di una qualche forma di parametrizzazione *sotto griglia*, ovvero di una trattazione semplificata.

Ad esempio, nel caso della formazione di stelle, invece di seguire la formazione e l'evoluzione delle nubi molecolari fredde, si usano le caratteristiche della *particella di gas* per determinare se il sito sia adatto alla genesi di una popolazione stellare. In caso positivo, viene usato un modello stocastico per generare una *particella di stelle*.

Vi sono oggi molti codici, *N*-body e idrodinamici, disponibili in libreria o utilizzabili «su richiesta». In genere sono anche disponibili in rete informazioni sufficienti al loro uso. La potenza di calcolo dei migliori PC è ormai adeguata a realizzare discrete simulazioni; i problemi più significativi, tuttavia, richiedono potenze di calcolo sempre più avanzate.

## 9. Scopo delle simulazioni

Un primo scopo delle simulazioni è quello di ricostruire la distribuzione osservata di galassie, gruppi e ammassi di galassie; quindi: funzione di massa, funzioni di correlazione, probabilità dei vuoti, ecc. Esse mirano altresì a ricavare talune delle caratteristiche morfologiche di questi sistemi *condensati*, nonché quelle dei vuoti che si vengono a formare tra essi.

Le simulazioni  $N$ -body mirano a questo fine, sia pur mettendo sullo stesso piano la componente barionica e quella di DM e considerando solo le forze gravitazionali. Come abbiamo visto, la componente barionica è soggetta a altre forze, anzitutto di natura idrodinamica e dissipativa, e diviene spesso necessario implementare il codice  $N$ -body con algoritmi che ne tengano conto. È comunque opportuno sottolineare che, in ogni caso, la struttura portante delle simulazioni cosmologiche è data dalla fisica della gravitazione, e che la gravità è trattata a livello newtoniano.

La descrizione della morfologia dei sistemi che si vogliono simulare costituisce uno dei rami fondamentali dell'astrofisica e esula in larga parte dagli scopi di questo testo. D'altra parte, sebbene molte di queste caratteristiche non siano strettamente legate al modello cosmologico, ma abbiano una dinamica compatibile con un vasto *range* di modelli, ciò non è necessariamente vero, *a priori*, per qualunque caratteristica. Alcune misure morfologiche derivano certamente dalla sintesi di processi astrofisici e dinamiche cosmologiche, per cui ogni progresso nella comprensione di ciascun elemento dialettico giova a porre limiti sull'altro.

La fenomenologia di galassie, gruppi e ammassi costituisce un articolato capitolo dell'astrofisica. Daremo qui alcuni ragguagli su elementi che spesso sono confrontati con le simulazioni.

### ***Le galassie***

Nel capitolo II abbiamo già considerato alcune proprietà di questi oggetti. In particolare abbiamo visto che:

- 1) le loro masse si estendono da  $\sim 10^7 h^{-1} M_\odot$  a oltre  $\sim 10^{12} h^{-1} M_\odot$ ;
- 2) il contrasto di densità tra l'interno di una galassia e il cosmo è  $\sim 10^7$ - $10^8$ ;
- 3) la distanza media tra galassie è di qualche Mpc ( $d_g \sim 3$ - $4 h^{-1}$  Mpc).

Inoltre abbiamo visto la forma della *funzione di luminosità* delle galassie.

Daremo qui alcuni dettagli ulteriori sui tipi morfologici e sulle relative caratteristiche sottolineando che, in linea di massima, in tutte le galassie sono presenti tre componenti, il *bulge*, il *disco* e l'*alone*, mentre i vari tipi morfologici si possono ricondurre al peso relativo e alle caratteristiche di queste componenti.

Il BULGE costituisce comunque la parte centrale della galassia e ha una forma approssimativamente sferica. In genere esso è in rotazione, con velocità angolare approssimativamente costante, in funzione del raggio; le stelle del bulge si muovono su orbite che hanno una grossa componente radiale, sono prevalentemente vecchie, di *popolazione II*, e hanno quindi bassa metallicità.

Il DISCO è una struttura piatta al cui centro si trova il rigonfiamento più o meno prominente del bulge. Le stelle del disco, in genere, sono più recenti, di *Popolazione I*, e si muovono su orbite che hanno una grossa componente circolare. Nel disco si osservano anche sistemi di natura non stellare, come nebulose di gas e polveri.

L'ALONE è stato già discusso, nelle sue caratteristiche, nel capitolo V, in particolare per la sua velocità di rotazione indipendente dalla distanza  $R$  dal centro galattico.

Nelle *galassie ellittiche* la componente dominante è il bulge. In alcune di esse si può riscontrare una traccia di disco, mentre può risultare difficile distinguere il bulge stesso dall'alone. Il bulge delle galassie ellittiche, in genere, non è sferico ma ellissoidale, con 3 assi più o meno diversi. Proiettandosi sulla volta celeste, esso dà luogo a un'immagine ellittica. Gli astronomi classificano quest'immagine sulla base del rapporto tra asse maggiore e minore dell'ellisse (di lunghezze  $a_M$  e  $a_m$ ), determinata tramite opportune *isofote*, sicché le ellittiche vanno dalle E1 ( $a_M \simeq a_m$ ) alle E6 ( $a_M \simeq 6 a_m$ )

Nelle *galassie a spirale* la struttura a disco è prominente e ha la forma di una spirale, di norma a due braccia. Al crescere della rilevanza delle spirali, queste galassie sono denominate *Sa*, *Sb*, *Sc*, e anche *Sd*, mentre gli astronomi ricorrono talora a denominazioni intermedie come *Sbc*.

Tra le caratteristiche del bulge delle galassie ellittiche, significativa è la possibile presenza di una *barra*, cioè di un piano in cui si accumulano le orbite oscillatorie delle stelle; la barra stessa ruota su tempi scala  $\sim 10^7$ - $10^8$  anni.

La *galassie lenticolari* o SO hanno una morfologia intermedia tra le due precedenti. Vi sono poi le *galassie irregolari*, in cui le tre componenti risultano confuse; in genere si tratta di galassie di piccola massa. Nelle irregolari si riscontra spesso formazione stellare (*galassie blu compatte*).

Tra le caratteristiche che le simulazioni tentano di riprodurre vi sono quelle concernenti il *profilo* del bulge e del disco, nonché le regolarità dette *proprietà dinamiche*.

Osservativamente l'intensità della luce emessa dal disco ha un andamento  $I(r) \propto r^{-\alpha}$  (nessuna relazione tra questo esponente e quelli nella funzione di luminosità o nella varianza di massa) detto *legge di deVaucouleurs* o *legge*  $r^{1/4}$ , in quanto il valore di  $\alpha$  è in genere prossimo a 0.25. Se  $r_{1/2}$  è il raggio dell'area da cui proviene metà della luce e  $I_{1/2}$  è l'intensità luminosa a tale raggio, questa legge si può normalizzare nella forma

$$(IX.35) \quad I(r) = I_{1/2} \exp\{u[1 - (r/r_{1/2})^{0.25}]\}$$

in cui  $u$  è una costante. La luminosità totale del disco risulta allora  $I_t \simeq 3uI_{1/2}r_{1/2}^2$  e il valore di  $u$  è comunque ben approssimato da 7.6;  $r_{1/2}$  invece è variabile ed è il parametro che dà le dimensioni del disco.

Le proprietà dinamiche pongono invece in relazione la luminosità con la velocità di rotazione del disco o la *dispersione di velocità* nel bulge.

Per le galassie ellittiche e nel bulge delle spirali vale con buona approssimazione una regolarità detta *piano fondamentale*. Essa connette la luminosità totale  $L$ , l'intensità luminosa al centro  $I_o$ , la dispersione di velocità  $\sigma$  e ha la forma

$$(IX.36) \quad L \propto I_o^a \sigma^b \quad \text{con} \quad a \simeq -0.7, \quad b \simeq 2.9,$$

che riflette la proprietà del bulge di essere un sistema autogravitante con rapporto massa/luminosità sostanzialmente costante. Il piano fondamentale non appare quindi una proprietà suscettibile di dare informazioni cosmologiche.

Per i dischi delle spirali vale invece la relazione di *Tully-Fisher*

$$(IX.37) \quad L \propto \sigma^c \quad \text{con} \quad c \sim 3 - 4$$

che potrebbe avere una rilevanza cosmologica.

Molte proprietà osservative delle galassie discendono dall'evoluzione della componente barionica e debbono essere studiate approfondendo l'analisi dei fenomeni idrodinamici e dissipativi. Gli elementi di base della loro struttura debbono però essere compresi dal punto di vista gravitazionale.

### Formazione delle galassie

Le galassie si sono mediamente formate quando gruppi e ammassi in pratica non esistevano; molte sono state successivamente inglobate in gruppi o ammassi, quando le fluttuazioni di scala galattica che le avevano generate erano contenute in fluttuazioni più ampie, sulla scala di gruppo o ammasso.

Nel momento in cui i materiali galattici virializzavano, la densità media del cosmo era assai più elevata di oggi. Inoltre, la componente barionica delle galassie si è ulteriormente concentrata, perdendo energia grazie a emissione di radiazione. Ciò spiega perché il contrasto di densità delle galassie sia  $\Delta_g \sim 10^7-10^8$  a fronte di un contrasto di densità  $\sim 200$  per gruppi e ammassi.

Mostreremo ora, per grandi linee, come si arriva al valore  $\Delta_g$ . Il punto di partenza è il grafico di  $\Delta^2(k)$  nel pannello inferiore di figura IX.10.

Ricordiamo anzitutto che la varianza di massa ha l'espressione integrale

$$(IX.38) \quad \sigma^2(R) = \int^{\sim 2\pi/R} \frac{dk}{k} \Delta^2(k),$$

dalla quale ci interessa unicamente dedurre che  $\sigma^2$  cresce proporzionalmente alla funzione spettrale  $\Delta^2(k)$ . Ciò è particolarmente significativo dal momento che  $\Delta^2(k)$  cresce «in blocco», al crescere di  $a$  (fattore di scala), come  $|\delta(k)|^2$ ; in teoria lineare e in un modello SCDM, dunque, la crescita di  $\sigma^2$ , su tutte le scale di interesse cosmologico, sarebbe  $\propto a^2$ .

Come abbiamo visto nel paragrafo 5, la scala di massa tipica di un ammasso,  $\sim 2 \times 10^{14} h^{-1} M_\odot$ , corrisponde a  $\simeq 8 h^{-1} \text{Mpc}$  ed è oggi ancora, sia pure marginalmente, in fase di evoluzione lineare.

Una scala di massa  $\sim 2 \times 10^{11} h^{-1} M_\odot$ , tipica per le galassie, corrisponde dunque a una scala di lunghezza 10 volte minore e a un valore di  $k$  maggiore dello stesso fattore.

Sulla base dell'andamento di figura IX.9, che mostra una pendenza  $\sim 1$  sulla destra del grafico, e tenendo conto dell'espressione (IX.38) per la varianza, possiamo allora

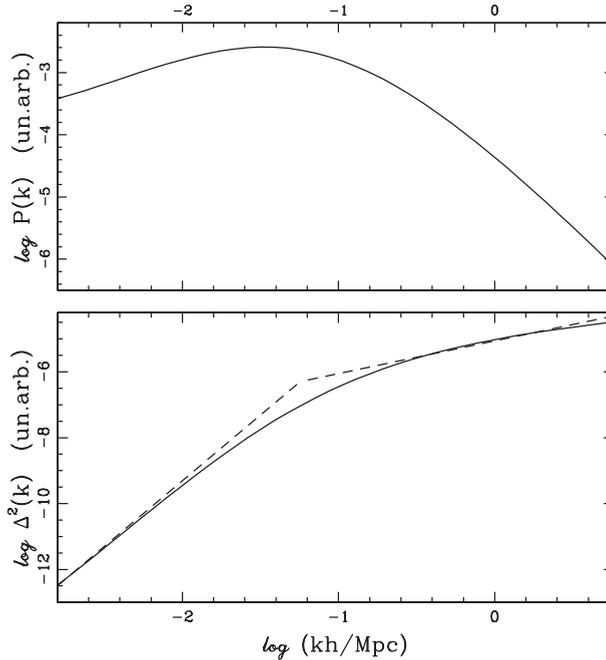


FIGURA IX.9

Spettro  $P(k)$  (pannello superiore) e funzione spettrale  $\Delta^2(k)$  (pannello inferiore) in un modello con  $\Omega_{om} = 0.3$ ,  $\Omega_{ob} = 0.04$ ,  $h = 0.7$ ,  $n_s = 1$ . Le unità in ordinata sono arbitrarie, anche al fine di evidenziare che, a parte una blanda evoluzione nei barioni, i grafici valgono indipendentemente dal *redshift*, in teoria lineare: la crescita del fattore di scala implica semplicemente un rialzo in blocco dei grafici che, in un modello SCDM, sarebbe semplicemente  $\propto a^2$ . Nel pannello inferiore, sopra al grafico ottenuto numericamente, sono tracciate due rette intersecantisi; esse esprimono l'andamento ricavato analiticamente, tenendo solo conto della stagnazione, conformemente a quanto descritto al paragrafo 9 del capitolo VIII. Il valore usato per il parametro  $u$  ivi definito è 0.75. La pendenza risultante per la parte dx del grafico è 1.

facilmente affermare che le galassie si trovavano in situazione simile a quella degli ammassi di oggi quando tutta la funzione  $\Delta^2(k)$ , in blocco, era  $\sim 10$  volte «più bassa». In un modello SCDM, ciò significherebbe  $1+z \sim \sqrt{10}$ ; nel modello  $\Lambda$ CDM considerato, possiamo stimare che ciò significa  $z \simeq 4-5$ .

Al momento della formazione delle galassie, la densità media del cosmo era quindi  $\sim 100$  volte superiore a oggi. Al momento della virializzazione, dunque, un alone destinato a generare una galassia, avrebbe avuto una densità  $\sim 200 \times 100 \times \rho_{o,cr}$  (densità critica oggi). Il passaggio dal contrasto di densità  $\sim 10^4$  così stimato, al contrasto di densità  $\sim 10^7$  osservato, va poi attribuito a una diminuzione di un ulteriore fattore  $\sim 10$  nelle dimensioni della struttura barionica. Ciò è fenomenologicamente confermato dal fatto che gli aloni delle spirali si estendono almeno 5 volte oltre il raggio ottico.

Riteniamo opportuno richiamare l'attenzione del lettore sul fatto che questa argomentazione per ordini di grandezza non significa che tutte le galassie di formino a

$z \simeq 5$ . L'intervallo di *redshift* su cui avviene la formazione di galassie è invece piuttosto ampio e può essere stimato tenendo conto dell'intervallo di masse galattiche e della possibilità di fluttuazioni più o meno eccezionali.

In effetti, accanto all'evoluzione «statica» qui descritta, va considerata tutta una parte «dinamica», che tenga conto dei moti delle galassie in corso di formazione, neoformate o già stabilizzate, delle conseguenti interazioni mutue, del formarsi di satelliti, ecc. Le simulazioni, in linea di principio, sono anche utili a seguire anche questa evoluzione, a patto di disporre di un adeguato *range dinamico*.

In una simulazione *N*-body, puramente gravitazionale, comunque, le galassie compariranno come aloni di raggio maggiore di quello delle galassie osservate. Implementando la simulazione con la dinamica dissipativa della parte barionica, si mira poi a seguire la formazione della struttura barionica. È chiaro che questa seconda parte della dinamica dipende assai meno dai dettagli del modello cosmologico.

### Gruppi e ammassi

L'effetto dinamico dei fenomeni dissipativi è assai meno marcato nella formazione di gruppi e ammassi di galassie.

Ai fini della comprensione dei dati osservativi è utile ricordare che gli ammassi furono inizialmente definiti da *George Abell* sulla base di dati ottici. Abell li selezionò con riferimento a un raggio apparentemente arbitrario  $\simeq 1.5 h^{-1} \text{Mpc}$ , collocandoli nelle regioni della volta celeste in cui, entro un cerchio di tale raggio, si trovavano elevati numeri di galassie. (La lunghezza  $1.5 h^{-1} \text{Mpc} \times 180^{1/3} \sim 8 h^{-1} \text{Mpc}$  corrisponde a un ammasso «medio» con massa  $\sim 4 \times 10^{14} h^{-1} M_{\odot}$ ; vedi anche paragrafo 3.) Il numero di galassie venne definito *ricchezza* e, secondo Abell, le classi di ricchezza *R* corrispondono ai numeri di galassie riportati nella sottostante Tabella

				Classi di ricchezza		
<i>R</i> =	0	1	2			
# gal.	30 – 50	50 – 80	> 80			

Assieme alla controparte dell'emisfero sud, di compilazione successiva, il campione di Abell annovera oltre 4000 ammassi, ma è ragionevolmente completo solo per ricchezze 1 e 2.

Ma gli insiemi principali di dati concernenti ammassi (e gruppi) sono basati su dati *X* e non ottici. Un ruolo importante, in questo campo è tuttora assunto dai dati ottenuti dal satellite *ROSAT*. Tutti i dati *X* sono comunque desunti da rilevazioni satellitari.

L'emissione *X* è prevalentemente causata dal gas caldo intrappolato nella buca di potenziale dell'ammasso. Il meccanismo di emissione è la *bremstrahlung* o effetto *free-free*. Si osservano in sostanza i fotoni emessi dagli elettroni che si incontrano e deviano sotto l'azione del campo repulsivo coulombiano. Il gas è in equilibrio termodinamico e ha quindi una temperatura  $T_a$ , fissata dalla profondità della buca di potenziale e quindi dalla massa dell'ammasso. Come ordine di grandezza risulta

$T_a \sim 10^8$  K. L'intensità dell'emissione sarà proporzionale al volume totale emittente e alla frequenza degli incontri tra elettroni:

$$L \propto R_a^3 n_e^2 \quad (R_a: \text{raggio regione emittente; } n_e: \text{numero elettroni per cc.})$$

Nello studio degli ammassi si assume in genere una simmetria (quasi)-sferica e l'equilibrio idrostatico. Le galassie sono come le particelle di un fluido. Dalla loro dinamica, con opportune assunzioni e tecniche si cerca allora di risalire alle caratteristiche della buca di potenziale e alla massa dell'ammasso.

Per il gas e le galassie si osservano distribuzioni diverse:

$$(IX.39) \quad \rho_{gas} = \frac{\rho_{o, gas}}{\left[1 + (r/R_a)^2\right]^{1.5}}$$

$$\rho_{gal} = \frac{\rho_{o, gal}}{\left[1 + (r/R_a)^2\right]^{1.5\beta}}.$$

La distribuzione del gas approssima la dipendenza da  $r$  della densità di una sfera isoterma. La distribuzione delle galassie differisce sostanzialmente da essa per la presenza del fattore

$$(IX.40) \quad \beta = \mu m_p v^2 / k_B T_a$$

all'esponente. In quest'espressione  $v$  è la velocità quadratica media del «fluido» costituito dalle galassie. Essa viene attribuita a particelle di massa protonica ( $m_p$ ) nonostante la componente dominante la massa delle galassie sia la DM; il fattore  $\mu$  serve a correggere per tener conto del peso molecolare medio delle componenti barioniche delle galassie.

Procedure diverse portano a valutare  $\beta$  nell'intervallo 0.7-1, testimoniando essenzialmente il livello dell'approssimazione fatta nell'assumere  $\beta$  costante.

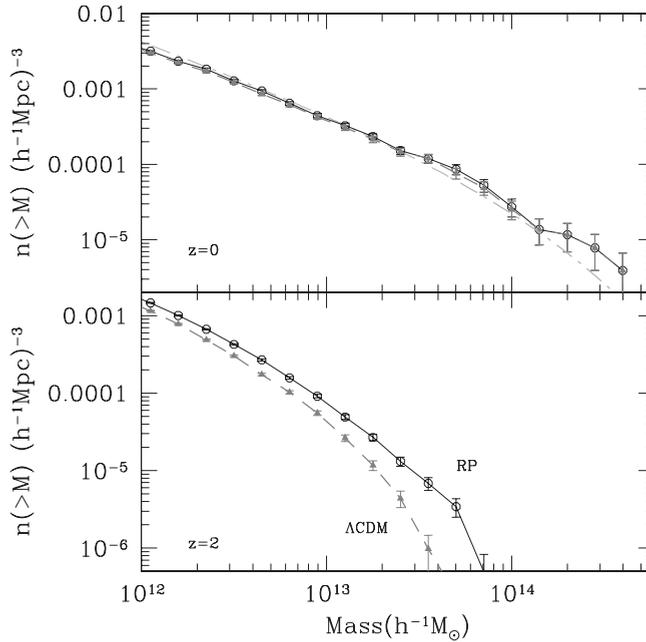
Tra gli scopi delle simulazioni, uno dei principali è la ricostruzione del profilo di densità degli ammassi.

## 10. Alcuni risultati delle simulazioni: il profilo NFW

In questo paragrafo discuteremo essenzialmente risultati di simulazioni  $N$ -body, in cui intervengono quindi essenzialmente forze gravitazionali. Esse sono d'altronde quelle con più diretta rilevanza cosmologica. Assai più difficile risulta una discriminazione tra modelli basata su osservabili in cui intervengono variabili concernenti la fisica del gas barionico o effetti di natura astrofisica ancor più specifica, come l'esplosione di supernovae o l'azione di QSO e/o nuclei galattici attivi.

In figura IX.10 viene confrontata una funzione di massa dedotta da un algoritmo tipo PS con la funzione di massa dedotta, per lo stesso modello, da una simulazione a due diversi *redshift*.

Questi confronti possono essere alternativamente considerati come una verifica degli algoritmi tipo PS o dell'efficacia delle simulazioni  $N$ -body.

**FIGURA IX.10**

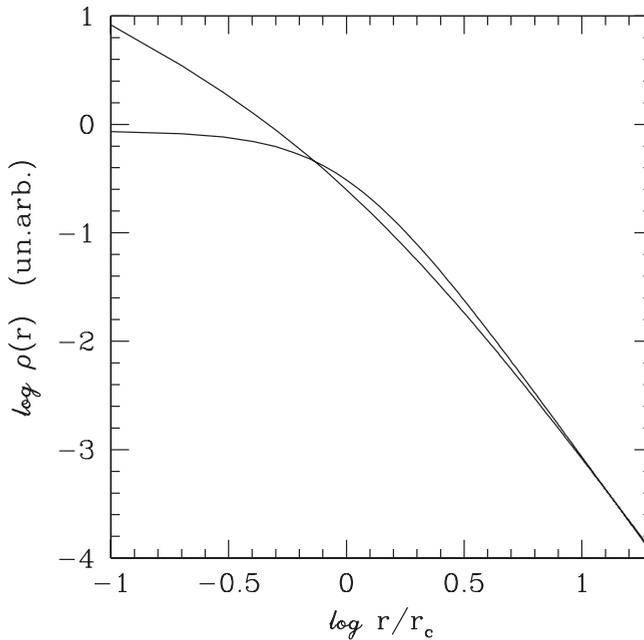
Raffronto tra previsioni sulla funzione di massa e risultati delle simulazioni. Nel riquadro in alto il raffronto è effettuato a  $z = 0$ . In questo caso la linea punto-trattino rappresenta le previsioni ottenute con tecniche PS, mentre i punti con barre di errore poissoniane, collegati da riga continua, rappresentano i risultati delle simulazioni, sostanzialmente indipendenti dal modello cosmologico. Nel riquadro in basso è invece rappresentata la funzione di massa a  $z = 2$ . La linea tratteggiata rappresenta le previsioni teoriche per un modello  $\Lambda$ CDM confrontate con le risultanze dedotte dalla relativa simulazione. La linea compatta rappresenta le previsioni teoriche per un altro modello di DE (RP: vedi capitolo XII), con le rispettive risultanze dedotte dalla simulazione. Le simulazioni usate sono state effettuate col codice ART in un cubo di lato  $80 h^{-1} \text{Mpc}$ . Sia ascissa che ordinata sono logaritmiche. (si veda l'appendice TAVOLE A COLORI).

Dove le simulazioni danno risultati che vanno al di là delle previsioni post-lineari, è sulla forma dei profili degli ammassi.

La densità di un alone virializzato prodotto da una simulazione ha l'andamento

$$(IX.41) \quad \rho(r) = \frac{\rho_c}{(r/r_c)(1 + r/r_c)^2},$$

detto NFW dalle iniziali di Navarro, Frenk e White, gli autori che hanno scoperto questa regolarità. Non è stata ancora trovata una giustificazione teorica veramente soddisfacente di questa regolarità peraltro confermata di un numero enorme di esperimenti numerici.  $\rho_c$  e  $r_c$  sono opportuni parametri che variano da alone a alone; come vedremo alla fine di questo Paragrafo, la virializzazione impone una relazione tra tali parametri, per cui gli aloni di una simulazione sono caratterizzati da un'unico parametro.



**FIGURA IX.11**

Raffronto tra profilo NFW (IX.41) e profilo di un gas isoterma (IX.39). L'andamento a  $r \gtrsim r_c$  è consistente; il profilo NFW prevede invece una cuspidè nell'origine che non si riscontra nei dati.

In figura IX.11 si confronta il profilo NFW col profilo del gas isoterma (IX.39).

Mentre i due profili sono consistenti a grandi  $r$ , la differenza a piccoli raggi è molto significativa. In un gas isoterma, la velocità quadratica media delle particelle è indipendente da  $r$ ; la formazione di un profilo NFW significa quindi una velocità quadratica media delle particelle di DM approssimativamente costante a  $r > r_c$  mentre, per  $r < r_c$ , essa inizia a crescere.

Va però sottolineato che è senz'altro possibile che, nella realtà fisica, gas e DM abbiano i profili suaccennati, consistenti nella parte esterna e difforni verso il centro. In presenza di un profilo NFW per la DM, e delle forze gravitazionali che ne conseguono, il profilo del gas può facilmente rientrare consistentemente nella classe dei profili isoterma.

La differenza tra i due profili può anche essere qualitativamente compresa, seguendo l'evoluzione della distribuzione delle masse nelle simulazioni. La crescita di massa di un alone, infatti, avviene sostanzialmente per accrezione di grossi blocchi e ciascun blocco ha inizialmente una componente barionica e una componente di DM. La prima viene però frenata, nella sua discesa verso il centro della buca di potenziale dell'alone, dall'interazione col gas. Dopo un periodo in cui l'equilibrio del gas è perturbato, la sua energia verrà termalizzata e il profilo del gas tornerà a essere isoterma. La componente di DM, invece, non viene tanto facilmente arrestata nella sua discesa e precipita verso il centro dell'alone, dove contribuisce a creare il picco caratteristico del profilo NFW.

Future misure sistematiche della distribuzione di massa negli ammassi di galassie, basate su fenomeni di lensing gravitazionale, potranno contribuire a accertare quale sia la reale discrepanza tra i profili della DM e dei barioni.

Dall'espressione (IX.41) si deduce che la massa complessiva entro un raggio  $R$  dal centro dell'alone è

$$(IX.42) \quad M(< R) = 4\pi r_c^3 \rho_c \times \left[ \ln(1 + R/r_c) + \frac{1}{1 + R/r_c} \right]$$

**Verifica.** Sarà

$$\begin{aligned} M(< R) &= 4\pi r_c^3 \rho_c \int_0^{R/r_c} d(r/r_c) (r/r_c)^2 \frac{1}{(r/r_c)(1 + r/r_c)^2} = \\ &= 4\pi r_c^3 \rho_c \left( \int_1^{1+R/r_c} \frac{dy}{y} - \int_1^{1+R/r_c} \frac{dy}{y^2} \right) \end{aligned}$$

da cui segue facilmente l'espressione (IX.42).

Questa espressione si può utilmente considerare al raggio viriale  $R_v$  definito in eq. (IX.18): entro il raggio viriale troveremo tutta la massa da attribuire all'alone. Il rapporto

$$(IX.43) \quad c = R_v/r_c$$

si definisce altresì *concentrazione* dell'alone. Per la massa dell'alone valgono allora le espressioni

$$(IX.44) \quad M_h = 4\pi R_v^3 \rho_c \times c^3 \left[ \ln(1 + c) + \frac{1}{1 + c} \right] = (4\pi/3) R_v^3 \rho_{cr} \Delta_v,$$

dal cui confronto si ricava il parametro

$$(IX.45) \quad \rho_c = \rho_{cr} \frac{\Delta_v}{3c^3} [\ln(1 + c) + 1/(1 + c)]$$

in funzione del contrasto di densità viriale e della concentrazione.

Come è del tutto ragionevole e come si riscontra nelle simulazioni, la concentrazione degli aloni cresce al trascorrere del tempo. In simulazioni  $\Lambda$ CDM del cosmo attuale, si riscontrano concentrazioni, sulla scala di massa degli ammassi,  $\sim 5-8$ .

Il valore della concentrazione è maggiore sulla scala delle galassie, dove può giungere a 12-14.

La distribuzione e il valor medio delle concentrazioni sono un parametro dipendente dal modello cosmologico. Ad esempio, se un modello prevede la formazione degli ammassi a un *redshift* maggiore, esso prevede anche una loro maggiore concentrazione media.

# Appendice A1

## Fluttuazioni nella DM e nei barioni dopo la ricombinazione

Finita la ricombinazione, a un tempo  $t_i$ , avremo fluttuazioni nella DM e nei barioni con ampiezze  $\delta_{c,i}$ ,  $\delta_{b,i}$  e derivate delle ampiezze  $\dot{\delta}_{c,i}$ ,  $\dot{\delta}_{b,i}$ , derivanti dall'evoluzione precedente. Una semplice generalizzazione delle equazioni del precedente capitolo ci dice che le fluttuazioni dovranno ora soddisfare le equazioni:

$$(A1.1) \quad \begin{aligned} \ddot{\delta}_c + \frac{4}{3t} \dot{\delta}_c - \frac{2}{3t^2} (\Omega_c \delta_c + \Omega_b \delta_b) &= 0 \\ \ddot{\delta}_b + \frac{4}{3t} \dot{\delta}_b - \frac{2}{3t^2} (\Omega_c \delta_c + \Omega_b \delta_b) &= 0. \end{aligned}$$

Considerando la differenza delle due equazioni nonché la somma della prima equazione moltiplicata per  $\Omega_c$  più la seconda moltiplicata per  $\Omega_b$ , si ottiene

$$(A1.2) \quad \ddot{D} + \frac{4}{3t} \dot{D} = 0, \quad \ddot{S} + \frac{4}{3t} \dot{S} - \frac{2}{3t^2} S = 0$$

(si ricordi che  $\Omega_c + \Omega_b = 1$ ) dove

$$(A1.3) \quad D = \delta_c - \delta_b, \quad S = \Omega_c \delta_c + \Omega_b \delta_b$$

In particolare  $S = \delta\rho/\rho$  è la fluttuazione della densità totale della materia non relativistica. Le due equazioni si integrano facilmente imponendo alle soluzioni di essere  $\propto t^\alpha$  e determinando  $\alpha$ . Si ricava così:

$$(A1.4) \quad S = A(t/t_i)^{1/3} + B(t/t_i)^{-1}, \quad D = V(t/t_i)^{-1/3} + Z,$$

in cui  $A, B, V, Z$  sono da determinare in funzione di  $\delta_{c,i}, \delta_{b,i}, \dot{\delta}_{c,i}, \dot{\delta}_{b,i}$ .

A meno di condizioni iniziali eccezionali, dei due termini di  $S$ , trascorso un tempo ragionevole, rimarrà significativo solo  $A(t/t_i)^{1/3}$ . Dei due termini di  $D$ , invece, ciò che si può dire è che il primo addendo, dopo un tempo ragionevole, non dà più contributo. Il secondo termine invece rimane dell'ordine di  $\delta_{c,i} - \delta_{b,i}$ . Quando  $\delta_c$  e  $\delta_b$  saranno cresciuti, ad esempio, di un fattore 100, esso continua a segnare una differenza tra essi  $\sim \mathcal{O}(1\%)$  o un po' meno. Si tratta di una differenza sempre più trascurabile ma che non può essere fissata tramite un algoritmo che prescindia dalla precedente evoluzione, quando la radiazione non è completamente disaccoppiata.

Del tutto in generale, possiamo comunque definire

$$(A1.5) \quad \alpha_+ = A/(\Omega_c \delta_{c,i} + \Omega_b \delta_{b,i}), \quad \zeta_+ = Z/(\delta_{c,i} - \delta_{b,i})$$

( $0 < \alpha_+, \zeta_+ < 1$ ) sicché l'andamento delle fluttuazioni è

$$(A1.6) \quad \begin{aligned} \delta_c &= \alpha_+ (\Omega_c \delta_{c,i} + \Omega_b \delta_{b,i}) (t/t_i)^{2/3} + \zeta_+ (\delta_{c,i} - \delta_{b,i}), \\ \delta_b &= \alpha_+ (\Omega_c \delta_{c,i} + \Omega_b \delta_{b,i}) (t/t_i)^{2/3} - \zeta_+ (\delta_{c,i} - \delta_{b,i}) \end{aligned}$$

Il problema, dunque, non ha una soluzione che prescindia dai dettagli della precedente evoluzione, ma la nostra ignoranza è tutta conglobata nei valori, compresi tra 0 e 1, dei parametri  $\alpha_+$  e  $\zeta_+$ .

## Appendice A2

### Evoluzione di una sovradensità sferica in una cosmologia SCDM

A un tempo  $t_i$  ( $\gg t_{rec}$ ), ci sia una sovradensità sferica di raggio  $R_i$  ove la densità è  $\rho_i > \bar{\rho}_i$  (densità media al tempo  $t_i$ ); la fluttuazione  $\delta = \rho_i/\bar{\rho}_i - 1$  sia inizialmente  $\ll 1$ . Definiamo anche il *contrasto di densità*  $\Delta_i$  e il *parametro di densità* interno alla sfera

$$(A2.1) \quad \Delta_i = \Omega_i = \rho_i/\bar{\rho}(t_i).$$

A partire dal tempo  $t_i$  fattore di scala cresce secondo l'equazione

$$(A2.2) \quad \left(\frac{\dot{a}}{a}\right)^2 = \frac{4\pi G}{3}\bar{\rho}_i\left(\frac{a_i}{a}\right)^3 = H_i^2\left(\frac{a_i}{a}\right)^3$$

che esprime anche la conservazione dell'energia meccanica (complessivamente nulla) entro una sfera autogravitante di raggio  $a$ , al cui interno la densità iniziale è quella media del cosmo (e tale rimane a tempi successivi).

La stessa equazione può essere scritta per il raggio  $R$  della disomogeneità sferica, tenendo però conto che l'energia meccanica non è complessivamente nulla e  $\bar{\rho}_i$  va sostituito con  $\rho_i$ . Avremo quindi

$$(A2.3) \quad \left(\frac{\dot{R}}{R}\right)^2 = \frac{4\pi G}{3}\rho_i\left(\frac{R_i}{R}\right)^3 - \frac{\kappa}{R^2}$$

e qui aggiungiamo un ingrediente essenziale, che semplifica i risultati: l'espansione della sfera, inizialmente, avviene secondo il flusso di Hubble,  $\dot{R}/R = H_i$ . Avremo allora, all'istante  $t_i$ ,

$$H_i^2 = \frac{4\pi G}{3}\rho_i\Omega_i - \frac{\kappa}{R_i^2} = H_i^2\Omega_i - \frac{\kappa}{R_i^2}$$

e l'eq. (A2.3) diviene

$$(A2.4) \quad \left(\frac{\dot{R}}{R}\right)^2 = H_i^2\Omega_i\left(\frac{R_i}{R}\right)^3 + H_i^2(1 - \Omega_i)\frac{R_i^2}{R^2}$$

ovvero

$$(A2.5) \quad \left(\frac{\dot{R}}{R_i}\right)^2 = H_i^2\left(\Omega_i\frac{R_i}{R} + 1 - \Omega_i\right)$$

Questa equazione è formalmente identica a quella obbedita dal fattore di scala in un Universo chiuso contenente una sostanza con parametro di stato  $w = 0$ . Di questa equazione si può dare un integrale esatto in forma parametrica:

$$(A2.6) \quad R(\theta) = R_i\frac{\Omega_i}{2(\Omega_i - 1)}(1 - \cos\theta), \quad H_i t(\theta) = \frac{\Omega_i}{2(\Omega_i - 1)^{3/2}}(\theta - \sin\theta).$$

come riportato in eq. (IX.5).

## DIMOSTRAZIONE

$$\text{Infatti: } \frac{dR}{d\theta} = R \frac{\sin \theta}{1 - \cos \theta}, \quad \frac{d(H_i t)}{d\theta} = \frac{R}{R_i(\Omega_i - 1)^{1/2}}$$

$$\text{da cui: } \frac{\dot{R}^2}{H_i^2} = \frac{\sin^2 \theta}{(1 - \cos \theta)^2} R_i^2 (\Omega_i - 1)$$

$$\text{ed, essendo: } \frac{1 - \cos^2 \theta}{(1 - \cos \theta)^2} = \frac{2}{1 - \cos \theta} - 1 = \frac{R_i}{R} \frac{\Omega_i}{\Omega_i - 1} - 1$$

$$\text{segue che: } \frac{\dot{R}^2}{H_i^2} = R_i^2 \left( \Omega_i \frac{R_i}{R} - \Omega_i + 1 \right)$$

verificando così che l'espressione (A2.6) è integrale dell'equazione (A2.5).

## CONFRONTO CON LA SOLUZIONE LINEARE

La generica soluzione lineare ha forma

$$(A2.7) \quad \delta(t) = \delta_{i,+}(t/t_i)^{2/3} + \delta_{i,-}(t/t_i)^{-1}$$

con

$$(A2.8) \quad t \dot{\delta}(t) = (2/3)\delta_{i,+}(t/t_i)^{2/3} - \delta_{i,-}(t/t_i)^{-1}$$

Nel nostro caso è  $\dot{\delta}(t_i) = 0$ , in quanto l'espansione iniziale avviene secondo il flusso di Hubble. Sarà quindi  $\delta_{i,-} = (2/3)\delta_{i,+}$  e

$$\delta_{i,+} = (3/5)\delta_i$$

A  $t \gg t_i$ , la soluzione lineare sarà allora

$$(A2.9) \quad \delta(t) \simeq (3/5)\delta_i(t/t_i)^{2/3}.$$

A questa evoluzione della densità si contrappone l'evoluzione della densità entro la fluttuazione sferica:

$$(A2.10) \quad \delta(t) = \Delta(t) - 1 = [R(t)/a(t)]^3 - 1$$

con  $R(t)$  e  $a(t)$  dati dalle eq. (A2.6) e (A2.2). Il confronto non è quindi immediato, in generale, in quanto  $R(t)$  si deve ricavare dalla dipendenza di  $R$  e  $t$  da  $\theta$ ; tuttavia, per alcune configurazioni significative, come quella di massima espansione, il confronto è più agevole ed è discusso nel testo del capitolo.

## Appendice A3

### Funzione di massa di Press & Schechter

Alla base della discussione stanno due fatti discussi nel testo del capitolo:

- 1) Un sistema virializzato a un tempo  $t$  nasce da una fluttuazione che, se fosse evoluta secondo le equazioni lineari, avrebbe ampiezza  $\delta(\bar{t}) > \delta_v$ . Per  $\delta_v$  abbiamo un valore che dipende solo marginalmente dal modello cosmologico:  $\delta_v = 1.68$ .
- 2) Se la distribuzione delle ampiezze di fluttuazione è inizialmente Gaussiana, le equazioni lineari fanno crescere le componenti spettrali secondo la legge

$$(A3.1) \quad P(k) = \langle |\delta(k)|^2 \rangle \propto a^2$$

ma non alterano la natura Gaussiana della distribuzione.

La distribuzione delle ampiezze, sulla scala di massa  $M$ , al tempo considerato, sarà quindi

$$(A3.2) \quad p(\delta, M) = \frac{1}{(2\pi)^{1/2} \sigma_M} e^{-(\delta^2/2\sigma_M^2)}$$

dove

$$(A3.3) \quad \sigma_M^2 = \frac{1}{2\pi^2} \int_0^\infty dk k^2 P(k) W(kR) \quad \text{con} \quad R = (3M/4\pi\rho_m)^{1/3}$$

è la «varianza di massa» sulla scala  $M$ . Ricordiamo che  $W(kR)$  è un «filtro» che stabilisce, in sostanza, un limite superiore soffice all'integrazione (A3.3), attorno a  $k \sim 1/R$ . Espressioni per  $W(kR)$  sono date nel capitolo II. Viceversa, l'espressione (A3.3) permette di ricavare, in funzione dello spettro, lo scarto quadratico medio della massa contenuta in un volume di raggio  $R$ , rispetto al suo valor medio  $M$ .

Al crescere di  $M$  e di  $R$ , il limite superiore nell'integrale (A3.3) si abbassa.  $\sigma_M \equiv \sigma(M)$  è dunque una funzione decrescente di  $M$ , per qualunque tipo di spettro. Di questa funzione utilizzeremo anche l'inversa:

$$(A3.4) \quad \text{se è} \quad \bar{\sigma} = \sigma(M), \quad \text{sia anche} \quad M = \mu(\bar{\sigma})$$

In genere, la funzione decrescente  $\mu(\sigma)$  sarà ricavata numericamente. Se, in un intervallo di masse, vale l'espressione

$$(A3.5) \quad \sigma(M) = (M/M_o)^x$$

sarà anche

$$(A3.6) \quad \mu(\sigma) = M_o \sigma^{1/x}.$$

Data la distribuzione (A3.2), la distribuzione di probabilità di avere una fluttuazione  $> \delta_v$  sulla scala  $M$  si ottiene derivando

$$(A3.7) \quad \Pi(\delta_v, M) = \int_{\delta_v}^\infty d\delta \frac{1}{(2\pi)^{1/2} \sigma(M)} e^{-[\delta^2/2\sigma^2(M)]} = \frac{1}{\sqrt{2\pi}} \int_{\delta_v/\sigma(M)}^\infty dx e^{-x^2/2}$$

rispetto a  $M$ . Ne segue che la distribuzione su  $M$  del numero di oggetti di scala

$M$  si ottiene moltiplicando la derivata di (A3.7) per il rapporto  $\rho_m/M$ , sicché risulta

$$(A3.8) \quad n(M) = 2 \frac{\rho_m}{M} \frac{1}{\sqrt{2\pi}} \frac{\delta_v}{\sigma^2(M)} \frac{d\sigma(M)}{dM} e^{-\delta_v^2/2\sigma^2(M)}$$

con l'aggiunta di un fattore 2 per tener conto dell'accrescimento, verso le zone ove si formano i sistemi, dei materiali delle zone sottodense.

Se si usa ora l'espressione (A3.5) della varianza di massa, si ricava

$$(A3.9) \quad n(M) = \frac{2\alpha}{\sqrt{\pi}} \frac{\rho_m}{M_o^2} \left(\frac{M}{M_o}\right)^{\alpha-2} \frac{\delta_v}{\sqrt{2}} e^{-(\delta_v^2/2)(M/M_o)^{2\alpha}} = \\ = \frac{2\alpha}{\sqrt{\pi}} \frac{\rho_m}{M_*^2} \left(\frac{M}{M_*}\right)^{\alpha-2} e^{-(M/M_*)^{2\alpha}} \quad \text{con} \quad M_* = M_o(2/\delta_v^2)^{1/2\alpha}$$

Espressioni differenziali di questo tipo sono peraltro di uso modesto. Le previsioni teoriche mirano a confrontarsi con dati osservativi che sono di natura integrale e danno il numero di oggetti al di sopra di una determinata massa.

Integrando l'espressione (A3.8) si ricava allora

$$(A3.10) \quad N(>M) = \sqrt{\frac{2}{\pi}} \rho_m \int_M^\infty d\mu \frac{1}{\mu} \frac{\delta_v}{\sigma^2(\mu)} \frac{d\sigma(\mu)}{d\mu} e^{-\delta_v^2/2\sigma^2(\mu)} = \\ = \frac{\rho_m}{M} \sqrt{\frac{2}{\pi}} \int_{\delta_v/\sigma(M)}^\infty dx e^{-x^2/2} \frac{M}{\mu(\delta_v/x)}$$

dove  $\mu(\delta_v/x)$  è la funzione inversa alla varianza, definita in eq. (A3.4).

Il senso di eq. (A3.10) è abbastanza trasparente. Se fosse  $N(>M) = \rho_m/M$ , significherebbe che tutta la materia del cosmo ha formato sistemi di massa  $> M$ . In generale, ovviamente, così non è; la frazione della materia cosmica in sistemi  $> M$  è data dal coefficiente integrale in eq. (A3.10).

## Appendice A4

### L'approssimazione di Zel'dovich

Nell'espressione

$$(A4.1) \quad \mathbf{x}(t) = a(t)[\mathbf{q} + G(t)\mathbf{s}(\mathbf{q})]$$

compare la funzione vettoriale  $\mathbf{s}$ , da cui si può ricavare il «tensore degli sforzi»

$$(A4.2) \quad A_{ij} = \partial s_i / \partial q_j.$$

Risolvendo l'equazione agli autovalori

$$A_{ij}\hat{a}_j = \alpha^{(k)}\hat{a}_i$$

si determinano gli autovettori normalizzati  $\hat{\mathbf{a}}^{(k)}$  e i corrispondenti autovalori  $\alpha^{(k)}$ . I versori  $\hat{\mathbf{a}}^{(k)}$  definiscono gli assi rispetto cui vale l'espressione (IX.27) e gli autovalori  $\alpha^{(k)} \equiv \alpha_k$  ( $k = 1, 2, 3$ ) sono i coefficienti che compaiono nella medesima espressione. Rispetto a tali assi  $A_{ij}$  è diagonale, sicché risulta

$$(A4.3) \quad A_{ij} \equiv \left\| \begin{array}{ccc} \alpha_1 & & \\ & \alpha_2 & \\ & & \alpha_3 \end{array} \right\| \quad \text{ovvero} \quad \frac{\partial s_i}{\partial q_j} = \delta_{ij}\alpha_i$$

o, anche

$$(A4.4) \quad \vec{\nabla} \cdot \mathbf{s} = \sum_{i=1}^3 \alpha_i$$

È facile verificare, come già sottolineato nel paragrafo 5, che è  $\varepsilon(\mathbf{r}) = -G \sum_i \alpha_i$ ; se ne deduce che sarà

$$(A4.5) \quad \varepsilon(\mathbf{r}) = -G\vec{\nabla} \cdot \mathbf{s}.$$

Conformemente all'eq. (IX.31) la «velocità comovente»  $\mathbf{u} = \dot{\mathbf{r}}$  soddisfa all'equazione

$$(A4.6) \quad a\mathbf{u} = \dot{\mathbf{x}} - H\mathbf{x}.$$

(ricordiamo che è  $\mathbf{x} = a\mathbf{r}$ ). Dall'eq. (A4.1) si deduce facilmente essere

$$\frac{1}{a} \left( \dot{\mathbf{x}} - \frac{\dot{a}}{a}\mathbf{x} \right) = \dot{G}(t)\mathbf{s}(\mathbf{q})$$

per cui sarà

$$(A4.7) \quad \mathbf{u} = \dot{G}(t)\mathbf{s}(\mathbf{q})$$

Il confronto tra le eq. (A4.5) e (A4.7) ci dice essere allora

$$(A4.8) \quad \dot{\varepsilon} = -\vec{\nabla} \cdot \mathbf{u}$$

e questa altro non è che l'equazione di continuità in termini di velocità comovente.

Dal campo  $\mathbf{u}(\mathbf{r})$  e dall'andamento del fattore di crescita  $G$  si può quindi ricavare la funzione  $\mathbf{s}$ , la direzione degli assi  $\hat{\mathbf{a}}^{(k)}$  e i coefficienti  $\alpha_i$ .

Nel caso di un modello SCDM, in cui sia  $G \propto a$ , risulta allora

$$(A4.9) \quad \mathbf{u} = HG(t)\mathbf{s}.$$

Se il campo  $\mathbf{u}$  è irrotazionale, allora, esiste un potenziale scalare  $\psi(q_i)$  talché

$$(A4.10) \quad \mathbf{s} = \vec{\nabla}_q \psi(\mathbf{q}).$$

Tale potenziale si può legare al potenziale gravitazionale peculiare delle fluttuazioni lineari iniziali,  $\phi(\mathbf{q})$ , tramite la relazione

$$(A4.11) \quad \phi(\mathbf{q}) = (3/2)H^2(t)a^3(t)G(t)\psi(\mathbf{q})$$

che comunque non dimostreremo.

L'algoritmo POTENT, da noi citato nel capitolo V, si basa sullo studio di un potenziale di tipo  $\psi$ , messo in relazione con la distribuzione delle masse.

## CAPITOLO X

# Cosmologia e forze elementari

### 1. Il dogma di gauge

Alla base dell'attuale comprensione delle forze fondamentali della natura sta la fisica dei *campi di gauge*. Il campo *elettromagnetico* è un campo di gauge. Ci sono poi altre 3 forze, la *gravità* e le forze nucleari *forte* e *debole*. Tutte sono veicolate da campi di gauge.

L'invarianza di gauge dell'elettromagnetismo era nota fin dai tempi di Maxwell. Ma solo negli anni Cinquanta, Yang & Mills rilesero l'elettromagnetismo come *teoria di gauge abeliana* per proporre poi le *teorie di gauge non-abeliane*. Pochi anni dopo, Utiyama provò che anche gravità è una forza di gauge, sebbene di natura molto particolare. Infatti essa è l'unica a essere veicolata da un campo tensoriale; tutti gli altri campi di gauge, incluso il campo elettromagnetico, sono vettoriali. Ma i passi fondamentali furono l'interpretazione delle forze nucleari come forze di gauge. Per ciò si dovette attendere l'approfondimento della *meccanica quantistica a temperatura finita* e del meccanismo di *rottura spontanea delle simmetrie*.

All'inizio degli anni Ottanta, infine, lo schema complessivo era divenuto chiaro e tutte le forze esistenti in natura poterono essere interpretate come *forze di gauge*. Si realizzò così uno schema unificante, non su base geometrica, come Albert Einstein aveva a lungo cercato di fare, ma puntando sull'interpolazione tra gradi di libertà spazio-temporali e gradi di libertà interni dei campi.

Partendo dallo schema unificato, la fisica oggi intravede come e perché, nella storia del cosmo, le varie forze si siano tra loro differenziate e in che senso tutte le interazioni siano riconducibili a un'unica forza.

Il primo passo verso schema unificante consiste nella comprensione del paradigma di gauge. La preistoria delle teorie di gauge venne scritta da una donna, *Emmy Noether* (1882-1935). Val la pena di ricordare brevemente che Emmy Noether era nata in Germania e fu iniziata alla fisica matematica dal padre, docente di fisica prima all'Università di Erlangen poi a Heidelberg. Nel suo paese, Emmy Noether, pur collaborando con personaggi come Hermann Weyl, Felix Klein o David Hilbert, non ricevette mai uno stipendio. Ciò che le università tedesche le concessero fu il «permesso» di accedere al sancta-sanctorum maschile degli istituti scientifici. All'avvento

del nazismo, Emmy Noether si trasferì negli USA. Morì a Princeton, poco dopo che l'*Institute of Advanced Studies* aveva deliberato di offrirle una remunerazione.

Il punto di partenza delle teorie di gauge è il *teorema di Noether*, che dimostra come ciascuna simmetria delle forze portate da campi implichi la conservazione di una quantità fisica. La relazione simmetria-conservazione era ben nota anche prima di Emmy Noether. È il caso, ad esempio, della seconda legge di Keplero. Essa afferma che il momento angolare dei pianeti – e, quindi, la loro *velocità areale* – è costante. Keplero non poteva saperlo, ma ciò risale alla simmetria del campo di gravità solare, eguale in tutte le direzioni; la sua invarianza, rispetto a rotazioni attorno a qualunque asse passante per il Sole, si esprime dicendo che le forze gravitazionali sono forze *centrali*.

Emmy Noether riformulò il rapporto simmetria-conservazione in teoria dei campi. Non è un passo banale, esso esige l'uso di una matematica avanzata e ciò avrà ricadute anche sull'esposizione di questo capitolo. Pur cercando di limitare le espressioni integro-differenziali e escludendo ogni calcolo basato su di esse, non potremo ometterle, neppure nel testo generale.

In teoria dei campi, ciascuna «conservazione» si esprime tramite una *equazione di continuità*, che non si può scrivere senza simboli differenziali; il suo significato può essere però compreso tramite l'esempio della carica elettrica: Se in un volume  $V$  al tempo  $t$  c'è la carica  $Q_V(t)$ , una sua variazione implica il passaggio di corrente attraverso la superficie  $\Sigma_V$  che delimita il volume  $V$ . La conservazione della carica si esprime dunque tramite una relazione carica-correnti, che deve valere per quanto piccolo sia il volume  $V$  considerato. Al limite di volumi  $dV$  differenziali, si ha una relazione valida punto per punto: l'equazione di continuità. Essa mette proprio in relazione, punto per punto, densità di carica ( $\rho$ ) e densità di corrente ( $\mathbf{j}$ ).

Tutto ciò avviene in un contesto relativistico e conviene definire un «4-vettore» di componenti  $j^\mu$  ( $\mu = 0, 1, 2, 3$ ): le  $j^i$  ( $i = 1, 2, 3$ ) sono le componenti di  $\mathbf{j}$ , mentre si definisce  $j^0 = c\rho$  ( $c$ : velocità della luce). Il flusso di  $\mathbf{j}$  attraverso  $\Sigma_V$  è la corrente elettrica. La conservazione della carica elettrica dà luogo all'*equazione di continuità*

$$(X.1) \quad \sum_{\mu=0}^3 \frac{\partial j^\mu}{\partial x^\mu} = 0 \quad \left( \text{che significa: } \frac{\partial \rho}{\partial t} + \frac{\partial j_1}{\partial x_1} + \frac{\partial j_2}{\partial x_2} + \frac{\partial j_3}{\partial x_3} = 0 \right)$$

e ogni equazione della forma (X.1) significa la conservazione di una quantità fisica, la cui densità spaziale è  $\rho = j^0/c$ . (In questo capitolo useremo di norma gli apici per le componenti spazio-temporali; quindi:  $x^0 = ct$ , le componenti di  $\mathbf{x}$  sono le  $x^i$ ;  $i = 1, 2, 3$ ; inoltre, ponendo  $\partial^\mu \equiv \partial/\partial x^\mu$ , si semplificano molte relazioni).

L'eq. (X.1) è un tipico risultato del teorema di Noether. Nel caso specifico il teorema dimostra che la conservazione della carica elettrica trae origine dall'invarianza di campi rispetto a trasformazioni di fase: se  $\psi(t, \mathbf{x})$  è un campo complesso – i suoi quanti potrebbero essere gli elettroni, positivi e negativi – le equazioni del moto risultano invarianti se si operano le trasformazioni

$$(X.2) \quad \psi \rightarrow \psi' = e^{iez} \psi, \quad \psi^* \rightarrow \psi'^* = \psi^* e^{-iez}$$

( $e$ : carica elettrica dell'elettrone;  $\alpha$ : fase arbitraria). In sostanza, le eq. (X.1) e l'invarianza (X.2) sono l'una la conseguenza dell'altra. In appendice A1 diamo l'enunciato generale del teorema di Noether e alcune sue applicazioni.

Su questo risultato di Emmy Noether si basa l'analisi di Yang & Mills. Essi si chiedono se la trasformazione (X.2) abbia senso fisico, con  $\alpha$  eguale in ogni sito e tempo. Essi poi mostrano che, se  $\alpha$  è invece funzione di  $x^\mu$ , all'invarianza (X.2) segue ancora la conservazione (X.1) solo se si assume la presenza di un opportuno campo vettoriale  $a^\mu$ , interagente con  $\psi$  in un modo ben preciso: *se le trasformazioni (X.2) da globali ( $\alpha$  costante) diventano locali ( $\alpha$  dipende da  $t$  e  $\mathbf{x}$ ), si può «salvare» la conservazione solo a patto di introdurre una forza.*

Ogni volta che, sul piano fenomenologico, si individua una quantità conservata, esiste una simmetria che la «causa». Se essa si esprime con un'invarianza *locale*, ci sarà anche una forza. Se non si osservano forze, ciò non comporta una contraddizione: l'invarianza non è locale ma globale.

La circostanza sembra però problematica: che senso fisico si può attribuire a fasi comunque costanti nel tempo e uniformi nello spazio? Viene naturale ritenere che, come questa costanza e uniformità deve inevitabilmente essere violata, così la conservazione osservata non varrà in ogni luogo e condizione.

In questo contesto viene enunciata la congettura detta *dogma di gauge*: **A ogni quantità conservata**, in ogni luogo e condizione, **corrisponde un'invarianza locale e, dunque, una forza**. E viceversa: **Se, a una quantità apparentemente conservata, non corrisponde una forza**, sicché essa si lega a un'invarianza globale, **tale conservazione varrà solo** in determinate circostanze, ad esempio, solo **per un determinato range di energie**. Sul dogma di gauge si basa l'attuale comprensione della dinamica microscopica.

All'invarianza rispetto a trasformazioni di fase e alla conservazione della carica elettrica, corrisponde la forza elettromagnetica. All'invarianza rispetto a trasformazioni di Lorentz e alla conservazione del momento angolare in qualsiasi sistema di riferimento, corrisponde la forza di gravità.

La forza nucleare forte è collegata alla conservazione di una carica denominata *colore* e all'invarianza rispetto a un gruppo denominato SU(3). Ancora più complesso è il caso della forza nucleare debole, cui corrisponde l'invarianza rispetto a un gruppo denominato SU(2)<sub>L</sub> e un teorema di conservazione in qualche modo sovrapposto alla conservazione della carica elettrica.

## 2. Simmetrie discrete

Occorre qui fare una breve parentesi, prima di approfondire queste simmetrie *continue*, rispetto a gruppi di trasformazioni individuate dai valori di opportuni parametri (ad esempio la fase) e la cui struttura è definita da *algebre* di matrici, come vedremo nei prossimi paragrafi. Tratteremo qui le simmetrie *discrete* che si riferiscono sulle operazioni  $\mathcal{P}$ ,  $\mathcal{T}$ ,  $\mathcal{C}$ ; rispettivamente: parità, inversione temporale, coniugazione di carica.

L'operazione di parità consiste nell'inversione dei 3 assi spaziali. Vi sono sistemi invarianti rispetto all'operazione di parità, ad esempio una distribuzione sferica omogenea di particelle centrata nell'origine degli assi, e ovviamente distribuzioni che invarianti non sono.

L'operazione  $\mathcal{P}$  si può anche applicare a una funzione  $f(x_1, x_2, x_3)$ . Se risulta

$$(X.3) \quad \mathcal{P}f(x_1, x_2, x_3) = f(-x_1, -x_2, -x_3) = f(x_1, x_2, x_3)$$

la funzione  $f$  è *invariante* per parità. Può anche essere, per un'altra funzione  $g(x_1, x_2, x_3)$ ,

$$(X.4) \quad \mathcal{P}g(x_1, x_2, x_3) = g(-x_1, -x_2, -x_3) = -g(x_1, x_2, x_3)$$

e, allora,  $f$  è una *autofunzione* di  $\mathcal{P}$  di *autovalore*  $+1$ ,  $g$  è *autofunzione* di  $\mathcal{P}$ , con *autovalore*  $-1$ . Un altro modo di esprimersi è dire che  $g$  è una funzione *dispari*,  $f$  una funzione *pari*: di qui il nome dell'operazione di parità.

Ad esempio: la funzione  $f(x_1, x_2, x_3) = (x_1 - x_2)^4$  è pari; la funzione  $g(x_1, x_2, x_3) = x_3^3 - x_1$  è dispari. Infine, la funzione  $F(x_1, x_2, x_3) = (x_1 - x_2)^4 - x_3$  non è autofunzione di parità.

L'operazione di parità si applica poi anche ai *funzionali*, cioè a quantità che invece di dipendere da variabili dipendono da funzioni. Nelle equazioni del moto (o di propagazione) di un campo  $\phi$  è contenuto un funzionale. Esse sono infatti della forma

$$(X.5) \quad \mathcal{F}[\phi(t, x_1, x_2, x_3)] = 0$$

dove  $\mathcal{F}$  è il *funzionale*, cioè una legge con cui agire sulla funzione di campo  $\phi$ . La soluzione delle equazioni del moto è la ricerca delle «leggi»  $\phi$  che rendano nullo il funzionale.

Un'equazione del moto che coinvolga i campi  $\phi_r$  ( $r = 1, 2, \dots, N$ ) è invariante per parità se risulta

$$(X.6) \quad \mathcal{P}\mathcal{F}[\phi_r(t, x_1, x_2, x_3)] = \mathcal{F}[\phi_r(t, -x_1, -x_2, -x_3)] = \mathcal{F}[\phi_r(t, x_1, x_2, x_3)].$$

Analoghe argomentazioni valgono per l'operazione  $\mathcal{T}$ :

$$(X.7) \quad \mathcal{T}f(t, x_1, x_2, x_3) = f(-t, x_1, x_2, x_3).$$

L'operazione di coniugazione di carica, infine, scambia particelle con anti-particelle. Ad esempio, uno stato fisico composto da 2 protoni, sotto l'azione di  $\mathcal{C}$ , diviene uno stato fisico composto da 2 anti-protoni, posti esattamente nello stesso sito. La «coniugazione di carica» si applica anche alle particelle... senza carica, purché abbiano una anti-particella. Ad esempio, un neutrone è trasformato in anti-neutrone. Ci sono particelle *invarianti* rispetto a  $\mathcal{C}$ , ad esempio i fotoni, che coincidono quindi con la propria antiparticella.

Si possono poi considerare combinazioni delle operazioni  $\mathcal{P}$ ,  $\mathcal{T}$ ,  $\mathcal{C}$ . È spesso assai significativa la combinazione di parità e inversione di carica:  $\mathcal{CP}$ .

Vale infine il seguente teorema: *Tutte le forze «locali» sono invarianti rispetto al prodotto delle tre operazioni CPT. Invertendo i 4 assi e scambiando particelle con anti-particelle, ogni interazione mantiene la stessa forma.*

### 3. Numero barionico e teorema di Sacharov

Queste operazioni entrano nell'enunciato di un teorema di grande importanza cosmologica, dovuto a *Andrei Sacharov* (fisico sovietico, Nobel per la pace).

Il dogma di gauge, infatti, stabilisce che ci sono quantità solo «apparentemente» conservate. Tra di esse c'è il *numero barionico*  $B$ . Ai *barioni*, come protoni e neutroni, si attribuisce numero barionico  $B = +1$ . Agli anti-barioni *coniugati di carica*, numero barionico  $B = -1$ . La conservazione di  $B$  è un evidente dato sperimentale. Essa vieta, ad esempio, a un protone di decadere in elettrone positivo (+ fotone). Sarebbe una reazione fortemente esotermica, che magari risolverebbe per sempre il problema delle risorse energetiche. Ma che renderebbe tragicamente instabile la materia di cui siamo fatti.

Ebbene, non ci sono *forze* legate alla «carica»  $B$ . La differenza tra *carica barionica* e *carica elettrica*, da questo punto di vista, è emblematica. Il dogma di gauge implica allora che, a alte energie, ci siano reazioni  $B$ -violanti.

In ambito cosmologico, quindi, si apre un problema ben definito: il cosmo odierno contiene una significativa quantità di barioni e, quindi, di numero barionico. A alte energie, tuttavia, condizioni iniziali con  $B \neq 0$  verrebbero rimesse in gioco da forze  $B$ -violanti.

Come vedremo, a alte energie, le particelle con numero barionico sono i *quark* ( $q$ ), ciascuno dei quali ha  $B = 1/3$ , mentre gli *antiquark* ( $\bar{q}$ ) hanno  $B = -1/3$ . A alte energie, l'urto  $\gamma\text{-}\gamma$  può produrre una coppia  $q\text{-}\bar{q}$  e, in condizioni di equilibrio la densità in numero di ciascun tipo di  $q$  e  $\bar{q}$ , dettata dalla statistica di Fermi, differirà solo di un fattore  $3/4$  dalla densità in numero di  $\gamma$ , dettata dalla statistica di Bose. Ricordiamo altresì che, oggi, abbiamo  $\sim 10^9$   $\gamma$  per ogni carica  $B$ . La conservazione di  $B$  prescrive quindi che, a alte energie, ci sia 1  $q$  in più ogni  $\sim 10^9$  coppie  $q\text{-}\bar{q}$ . In presenza di una dinamica  $B$ -violante possiamo immaginare che questo eccesso di  $1:10^9$  sarebbe rapidamente riassorbito.

Accanto al problema, tuttavia, si apre una prospettiva: partendo da condizioni iniziali più naturali, simmetriche rispetto all'operazione di *coniugazione di carica* ( $C$ ), in cui il numero di  $q$  e  $\bar{q}$  è eguale, è possibile sia la stessa dinamica dell'espansione, assieme a forze  $B$ -violanti, che origina un surplus di numero barionico, come osservato.

Il *teorema di Sacharov* fissa 3 condizioni per la produzione cosmologica di un eccesso di  $B$ .

- 1) Devono esistere interazioni  $B$ -violanti.
- 2) La forma delle interazioni non deve essere invariante rispetto a  $C$  e  $CP$ .
- 3) Occorre che, nel corso dell'espansione, la sostanza del cosmo sia uscita, per un tempo adeguato, dalle condizioni di equilibrio termodinamico.

Il dogma di gauge ci assicura che la condizione 1) è soddisfatta.

Ma è una condizione necessaria, non sufficiente. Se le reazioni  $B$ -violanti fossero comunque invarianti rispetto a  $C$  o a  $CP$ , a fronte di un processo elementare che

provoca un determinato  $\Delta B > 0$ , ce ne sarebbe un altro, egualmente probabile, che darebbe esattamente  $-\Delta B$ .

Ma, anche se le reazioni sono  $B$ -violanti e  $C$  o  $CP$  sono violati, il semplice permanere dell'equilibrio statistico garantirebbe la simmetria  $B$ - $B$ .

Una situazione primordiale, in cui le condizioni di Sacharov si verificano, ha luogo se il cosmo attraversa una fase di espansione inflazionaria (vedi capitolo successivo). L'equilibrio statistico è allora violato finché il *campo inflatonico* non raggiunge il suo valore di equilibrio. Una «buona inflazione», infatti, si deve concludere con una fase di *reheating*, dove assisteremo a un enorme input di entropia. Per la creazione di un eccesso barionico, alla fine dell'inflazione, si richiede che il *reheating* faccia recuperare al cosmo una temperatura alla quale le reazioni  $B$ -violanti siano efficaci.

#### 4. Particelle e campi

È noto che esistono due modi per trattare la radiazione. In talune condizioni essa va descritta dal campo elettromagnetico. In altre condizioni essa va descritta in termini di quanti, i *fotoni*.

Nella formulazione più semplice della meccanica quantistica, la cosiddetta *prima quantizzazione*, particelle come protoni o elettroni sono trattati in modo diverso dai quanti del campo elettromagnetico. L'equazione di Schrödinger, soddisfatta dalla *funzione d'onda*  $\Psi(t, \mathbf{x})$  dell'elettrone, ad esempio nell'atomo di idrogeno, non è un'equazione di campo;  $|\Psi(t, \mathbf{x})|^2$ , invece, ha il significato di (densità di) probabilità di trovare un elettrone al tempo  $t$  nel sito  $\mathbf{x}$ .

La disimmetria tra trattazione di radiazione e «materia» è riconducibile al numero di gradi di libertà considerati. La funzione d'onda dell'elettrone riguarda un sistema a 3 gradi di libertà. Il 4-potenziale elettromagnetico  $A^\mu$ , invece, che è la «funzione di campo» in elettromagnetismo, descrive un sistema a gran numero di gradi di libertà, almeno pari al numero di fotoni in gioco  $\times 3$ . L'asimmetria è funzionale a applicazioni di bassa energia, finché gli elettroni restano non-relativistici e, comunque, non sono emessi e/o riassorbiti. Altrettanto non vale per i fotoni, che si emettono e riassorbono a tutte le energie.

L'asimmetria si supera usando, accanto al campo  $A^\mu$ , i cui quanti sono i fotoni, un campo  $\psi$ , i cui quanti saranno gli elettroni positivi e negativi. La conservazione del *numero leptonico* assicura poi che, a bassa energia, non ci sia emissione o assorbimento di elettroni; a energie sufficienti, invece, potranno essere emesse o assorbite coppie elettrone-positrone (elettrone positivo).

#### **Rappresentazioni del gruppo di Lorentz**

Tra gli elementi che distinguono i campi tra loro c'è, in primo luogo, il comportamento rispetto a trasformazioni di Lorentz. In una trasformazione di Lorentz, le 4 coordinate  $x^\mu$  di un *evento* si trasformano nei modi noti dalla relatività speciale, che si possono

esprimere tramite un prodotto di matrici:

$$(X.8) \quad x'^{\mu} = \sum_{\nu=0}^3 A^{\mu\nu} x^{\nu}.$$

Nel caso del passaggio tra due sistemi di riferimento in moto relativo lungo l'asse  $x^1$ , con velocità relativa  $V$ , ad esempio, sarà

$$(X.9) \quad A^{\mu\nu} = \begin{vmatrix} \gamma & \beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

con  $\beta = V/c$ ,  $\gamma = (1 - \beta^2)^{-1/2}$ . Per queste (e qualsiasi altre) trasformazioni di Lorentz risulta

$$(X.10) \quad \sum_{\mu=0}^3 A^{\mu\nu} A^{\mu\rho} = \delta^{\nu\rho} \quad (\text{ortonormalità}).$$

Il campo  $A^{\mu}$ , che è un «campo vettoriale», si trasforma allo stesso modo di  $x^{\mu}$ .

Il valore di *campo scalare*  $\phi$  in un determinato punto del 4-spazio, invece, è indipendente dal sistema di riferimento:  $\phi$ , appunto, è uno scalare, come l'intervallo spaziotemporale  $ds^2$ .

Quella scalare e quella vettoriale sono due *rappresentazioni* del gruppo di Lorentz. Ma ce ne sono svariate altre, ciascuna caratterizzata da una determinata *dimensionalità*.

Di particolare importanza è la rappresentazione *spinoriale*. A essa appartiene il campo  $\psi$  che descrive elettroni e positroni. Un campo spinoriale ha 2 o 4 componenti, a seconda che i suoi quanti non abbiano o abbiano massa. Se il neutrino ha massa nulla, il campo «neutrinico» è sicuramente uno spinore a 2 componenti. Il campo spinoriale che descrive elettroni e positroni, invece, ha 4 componenti. Le componenti  $\psi_a$  ( $a = 1, \dots, 4$ ) dello spinore si trasformano, al cambiare del sistema di riferimento, secondo leggi di tipo (X.8):

$$(X.11) \quad \psi'^a = \tilde{A}^{ab} \psi^b$$

con matrici  $\tilde{A}^{ab}$  opportune, diverse dalle  $A^{\mu\nu}$ . La teoria delle rappresentazioni del gruppo di Lorentz è trattata in vari testi specializzati, cui rimandiamo il lettore. Temi critici sono la determinazione delle *rappresentazioni irriducibili* e le proprietà degli *spinori*, su cui comunque forniamo ulteriori informazioni in appendice A2. Nel prosieguo del testo faremo uso, senza dimostrazione, di molti risultati di tale complessa teoria.

Le trasformazioni di Lorentz sono un *gruppo*. Infatti:

- 1) Esiste una trasformazione unità.
- 2) Esiste l'inversa di ciascuna trasformazione.

3) Si sa cosa sia il prodotto di due trasformazioni e esso è una trasformazione di Lorentz. In particolare, il prodotto di 2 trasformazioni  $A_{(1)}{}^{\mu\nu}$  e  $A_{(2)}{}^{\mu\nu}$  è la trasformazione

$$(X.12) \quad A_{(3)}{}^{\mu\nu} = \sum_p A_{(1)}{}^{\mu\rho} A_{(2)}{}^{\rho\nu}.$$

Un esempio di rappresentazione non-scalare, diversa da quella vettoriale, è la rappresentazione tensoriale. Un tensore  $t^{\alpha\beta}$  si trasforma secondo la legge lineare

$$(X.13) \quad t^{\gamma\delta} = \sum_{\alpha,\beta} A^{\gamma\alpha} A^{\delta\beta} t^{\alpha\beta}.$$

Allineiamo le  $4 \times 4 = 16$  componenti dei tensori  $t^{\gamma\delta}$  in un'unica colonna  $t^A$  ( $A = 1, 2, \dots, 16$ ); sarà un semplice problema algebrico determinare gli elementi della matrice  $\bar{A}$  che interviene nella relazione

$$(X.14) \quad t^A = \sum_B \bar{A}^{AB} t^B,$$

capace di sortire gli stessi effetti di eq. (X.13). C'è una corrispondenza biunivoca tra le matrici  $A$  e  $\bar{A}$  e il prodotto di due matrici  $\bar{A}$  si fa righe per colonne. L'insieme delle matrici  $\bar{A}$  costituisce dunque un'altra rappresentazione del gruppo di Lorentz. La matrice a 1 colonna  $t^A$  è la base (16-dimensionale) di tale rappresentazione.

Le rappresentazioni possono essere *irriducibili* o *riducibili*. Ad esempio, una ipotetica rappresentazione  $12 \otimes 12$  è riducibile nella somma diretta di due rappresentazioni  $8 \otimes 8$  e  $4 \otimes 4$ , se tutte le matrici  $\bar{A}$  sono della forma

$$\begin{pmatrix} * & * & * & * & * & * & * & * & * & 0 & 0 & 0 & 0 \\ * & * & * & * & * & * & * & * & * & 0 & 0 & 0 & 0 \\ * & * & * & * & * & * & * & * & * & 0 & 0 & 0 & 0 \\ * & * & * & * & * & * & * & * & * & 0 & 0 & 0 & 0 \\ * & * & * & * & * & * & * & * & * & 0 & 0 & 0 & 0 \\ * & * & * & * & * & * & * & * & * & 0 & 0 & 0 & 0 \\ * & * & * & * & * & * & * & * & * & 0 & 0 & 0 & 0 \\ * & * & * & * & * & * & * & * & * & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & * & * & * \end{pmatrix}$$

in cui gli \* indicano valori  $\neq 0$ : un cambio di sistema di riferimento «mescola» tra loro esclusivamente le prime 8 e le ultime 4 componenti.

Tornando alle 16 componenti della base  $t^A$ , possiamo pensare a sostituirle con 16 loro combinazioni lineari  $f^A = U^{AB} t^B$ , in cui  $U$  è una matrice unitaria ( $U^{-1} = U^+$ ). Anche senza entrare in dettagliate verifiche algebriche, è evidente che, quando si

opera un cambiamento di sistema di riferimento, le  $f^A$  si combineranno anch'esse linearmente, come avviene in eq. (X.14): le matrici della trasformazione di Lorentz  $16 \otimes 16$  possono essere riespresse rispetto alla nuova *base* e avranno l'espressione

$$(X.15) \quad \bar{A}_{(f)}^{AB} = \sum_{A'B'} U^{-1AA'} \bar{A}_{(t)}^{A'B'} U^{B'B},$$

(in indice, tra parentesi, il nome della base) la cui dimostrazione lasciamo al lettore. Immaginiamo allora che le  $\bar{A}_{(t)}$  avessero la forma della matrice  $12 \otimes 12$  qui sopra, mettendo in evidenza una sua riducibilità, ad esempio, in due rappresentazioni 10- e 6-dimensionali. In genere, nella nuova base, la proprietà non sarà più visibile.

Viceversa, se una rappresentazione  $N \otimes N$  non mette in evidenza la sua riducibilità, è possibile che, cambiando base, essa diventi manifesta. La questione della riducibilità delle rappresentazioni è quindi complicata e non si risolve con la semplice ispezione delle matrici di trasformazione.

Allo stesso modo della rappresentazione 16-dimensionale, si possono definire rappresentazioni tensoriali di ordine più elevato.

Infine, è importante sottolineare che rappresentazioni spinoriali non appartengono alla classe delle rappresentazioni tensoriali.

### **Campi e particelle**

Si possono avere campi con diverso comportamento rispetto a trasformazioni di Lorentz, ma le equazioni cui essi obbediscono, se tra loro non intervengono forze, avranno sempre soluzioni di «onda piana». La dipendenza dalle variabili spaziali e temporale sarà quindi riconducibile a sovrapposizione tra esponenziali, seni o coseni, con coefficienti moltiplicativi ( $A$  o  $\psi_o$ ) che fissano le intensità. Usando gli esponenziali, ad esempio, ciascun addendo ha la forma

$$(X.16) \quad \psi = \psi_o \exp [i(\omega t - \mathbf{k} \cdot \mathbf{x})]$$

e una soluzione generica sarà la somma di simili addendi, secondo le regole dello sviluppo di Fourier.

Ciò mette in luce la relazione tra visuale ondulatoria e visuale particellare:

- 1) L'intensità dell'onda,  $|\psi_o|^2$ , dovrà essere riconducibile a  $n$ , densità in numero di particelle, o quanti del campo.
- 2) La frequenza angolare  $\omega$  e il numero d'onda  $\mathbf{k}$  fissano energia e impulso dei quanti, secondo le relazioni

$$(X.17) \quad E = \hbar\omega, \quad \mathbf{p} = \hbar\mathbf{k}.$$

In notazione relativistica si definisce anche  $k^0 = \omega/c$  (mentre è  $x^0 = ct$ ) e l'argomento dell'esponenziale (X.16) diviene  $\sum_{\mu=0}^3 k^\mu x^\mu g_{\mu\mu}$  con

$$g_{\mu\mu} \equiv 1, -1, -1, -1.$$

In quanto segue ometteremo il simbolo di somma, sottintendendola su indici o apici ripetuti. Per le componenti spazio-temporali, indichiate tramite apici, si terrà implicitamente conto della segnatura, data dalle componenti diagonali (omesse) del tensore metrico  $g_{\mu\nu}$ , come suindicato. Si intenderà quindi, in generale,

$$(X.18) \quad w^\mu v^\mu (\equiv w^\mu v^\mu g_{\mu\mu}) = w^0 v^0 - w^1 v^1 - w^2 v^2 - w^3 v^3.$$

In unità naturali, con  $\hbar = 1$ ,  $k^\mu$  coincide con  $p^\mu$ , 4-impulso dei quanti. Quindi, l'argomento dell'esponenziale (X.16) può essere anche scritto  $ip^\mu x^\mu$ .

La convenzione sulla somma degli indici e la collocazione in apice degli indici spazio-temporali non saranno mantenute in contesti di Relatività generale.

### **Lo spin. Parte di spin 0 del campo vettoriale e relazione tra simmetrie spazio-temporali e simmetrie interne.**

I quanti di campi appartenenti a rappresentazioni diverse hanno, in genere, *spin* diverso. Lo spin è il momento angolare intrinseco dei quanti (in unità di  $\hbar$ ).

Tra le conseguenze dello spin, c'è la *statistica* cui i quanti sono soggetti e quindi la loro distribuzione di equilibrio termodinamico. Per spin interi (0,1,2, ...) i quanti sono *bosoni*, per spin semi-interi (1/2, 3/2, ...) essi sono *fermioni*.

TABELLA 1

<i>campo</i>	<i>simbolo usuale</i>	<i># componenti</i>	<i>spin dei quanti</i>
scalare reale	$\phi$	1	0
scalare complesso	$\phi, \phi^*$	2	0
spinore con $m = 0$	$\psi$	2	1/2
spinore con $m \neq 0$	$\psi$	4	1/2
vettoriale	$A^\mu$	4	1, 0
tensoriale	$g^{\mu\nu}$	16	2, 1, 0

In tabella 1 sono riassunte le proprietà dei campi più rilevanti. Per campi vettoriale e tensoriale ci sono quanti con spin diversi. In particolare, nel caso di campi vettoriali, la teoria dei campi dimostra che l'energia dei quanti di spin 0 è *negativa* e che, in presenza di quanti di spin 0 «osservabili», sarebbero paradossalmente possibili degli stati con probabilità superiore all'unità (violazione dell'unitarietà della matrice  $S$ ).

In ogni teoria fisicamente sensata, questi effetti debbono essere neutralizzati. La parte di spin 0 deve essere resa ininteragente con qualsiasi altro campo e quindi intrinsecamente inosservabile.

È proprio su questo punto che si aggancia il paradigma di gauge. In quanto inosservabile, la componente di spin 0 è arbitraria e l'invarianza di gauge significa questo: che **la parte di spin 0 del campo vettoriale è arbitraria.**

Questa condizione costituisce una limitazione precisa alle interazioni possibili e determina in modo univoco tutta la dinamica dell'elettromagnetismo, in quanto descritto da un singolo campo vettoriale: l'interazione tra fotoni e altre particelle non potrebbe avere altra forma che quella osservata.

La scoperta di Yang & Mills consiste nella possibilità di considerare *multipletti* di campi vettoriali, le cui interazioni con gli altri campi comprendono termini ulteriori, rispetto alle interazioni possibili a ciascun campo, pur senza rimettere in gioco la componente di spin 0. Di questo inizieremo a trattare dal prossimo paragrafo.

## 5. Le algebre di Lie

Un multipletto di campi vettoriali è descritto da una funzione di campo a 2 indici, ad esempio  $a_j^\mu$ . L'apice  $\mu$  cataloga le componenti della rappresentazione vettoriale del gruppo di Lorentz. L'indice  $j$ , invece, cataloga le componenti del multipletto. Le componenti  $a_j^\mu$  del campo vettoriale, in un nuovo sistema di riferimento, si ricavano dalle componenti  $a_j^\mu$ , conformemente a quanto mostrato nelle eq. (X.8) e (X.9), usando le matrici  $A^{\mu\nu}$ . Esse «mescolano» le componenti con  $\mu$  diverso, tenendo fisso  $j$ .

L'indice  $j$ , invece, si riferisce a una specie di «spazio interno», in cui non esistono traslazioni ma sono possibili solo operazioni analoghe alle rotazioni di Lorentz.

La validità delle stesse leggi del moto in ogni sistema di riferimento è garantita dal fatto che i vettori  $a_j^\mu$  sono utilizzati per costruire un *integrale variazionale* scalare, invariante. Allo stesso modo, l'integrale variazionale dovrà essere invariante per le trasformazioni nello «spazio interno».

Nel caso del gruppo di Lorentz non abbiamo sin qui avuto necessità di riferirci alla sua *algebra*. Ogni trasformazione di Lorentz può peraltro essere vista come prodotto di rotazioni su uno dei 6 piani ortogonali di uno spazio 4-dimensionale (specificamente: rotazioni spaziali sui piani 1-2, 1-3, 2-3; *boosts* sui piani 0-1, 0-2, 0-3). In ciascuna rappresentazione del gruppo di Lorentz si possono introdurre 6 matrici, corrispondenti a ciascuna di tali rotazioni. Combinate con 6 opportuni coefficienti, esse danno luogo a tutte le  $\infty^6$  possibili trasformazioni di Lorentz. Le proprietà di commutazione tra tali matrici ne definiscono l'*algebra*. A ogni rappresentazione  $N$ -dimensionale del gruppo di Lorentz corrisponderà una rappresentazione  $N$ -dimensionale della sua algebra e le proprietà di commutazione delle matrici saranno ovviamente le stesse in ogni rappresentazione.

Nel caso delle simmetrie interne dei campi, ci conviene invece partire proprio dall'algebra dei gruppi di trasformazione.

Per costruire quantità invarianti rispetto ai gruppi di simmetria interna si dovrà allora tenere conto della particolare rappresentazione dell'algebra cui si riferiscono gli indici come  $j$ . Così,  $a_j^\mu$  appartiene alla rappresentazione vettoriale del gruppo di Lorentz (per l'apice  $\mu$ ) e apparterrà alla *rappresentazione auto-aggiunta* di un gruppo interno di simmetria basato su un'opportuna *algebra di Lie* (per l'indice  $j$ ).

Sin qui, tuttavia, lo spazio di Lorentz e lo «spazio interno» appaiono *ortogonali*, fra

loro indipendenti. L'interdipendenza nasce dalla parte di spin 0 di  $a_j^\mu$ . L'invarianza rispetto alle trasformazioni definite utilizzando algebre di Lie permette di escludere ogni ruolo dinamico per le componenti di spin 0 dei campi vettoriali. Tali componenti, pur presenti, risulteranno prive di qualsiasi ruolo dinamico.

Le forze presenti in natura risultano essere tutte del tipo reso possibile dalle simmetrie definite dalle algebre di Lie.

Le algebre di Lie si costruiscono tramite insiemi di matrici quadrate. Tra esse è definito il prodotto (non commutativo) «righe per colonne», di modo che il prodotto delle matrici  $A^1$  e  $A^2$  è la matrice

$$(X.19) \quad P_{AB} = A_{AZ}^1 A_{ZB}^2.$$

### Cosa sono le algebre di Lie

Prendiamo un insieme  $\mathcal{M}$  di matrici  $A_{ab}$ , esprimibili come combinazione lineare di  $N$  matrici-base  $B^J$  ( $J = 1, \dots, N$ ) tramite opportuni coefficienti  $\alpha^J$ :

$$(X.20) \quad A_{ab} = \alpha^J B_{ab}^J,$$

(si somma da 1 a  $N$  sull'apice ripetuto  $J$ ). È possibile che anche la matrice

$$(X.21) \quad C_{ab} \equiv B_{az}^1 B_{zb}^2 - B_{az}^2 B_{zb}^1 \quad \text{sia della forma} \quad C_{ab} = c^J B_{ab}^J$$

con opportuni coefficienti  $c^J$ , cioè appartenga all'insieme  $\mathcal{M}$ . La prima parte di eq. (X.21) viene anche scritta in modo compatto tramite il commutatore

$$(X.22) \quad C = [B^1, B^2]$$

che sarà l'operazione di composizione nell'ambito dell'algebra. L'eq. (X.21) esprime la principale proprietà che le matrici  $B^J$  debbono avere, per costituire un'algebra di Lie. Le altre proprietà sono che esista una matrice unità e l'inverso di ogni matrice.

### *SU(2): matrici di Pauli, rappresentazioni fondamentale e autoaggiunta*

Un esempio di algebra è costituito da tutte le matrici  $2 \otimes 2$  semplici (a traccia nulla) e unitarie (la matrice hermitiana coniugata coincide con la matrice inversa), a componenti complesse. [L'hermitiana coniugata di una matrice di componenti  $A_{ab}$  ha componenti  $(A^+)_{ab} = (A_{ba})^*$ ]. In questo caso si può scegliere come base l'insieme delle matrici di Pauli (divise per 2)

$$(X.23) \quad \sigma^1 = \frac{1}{2} \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}, \quad \sigma^2 = \frac{1}{2} \begin{vmatrix} 0 & -i \\ i & 0 \end{vmatrix}, \quad \sigma^3 = \frac{1}{2} \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix},$$

per le quali risulta

$$(X.24) \quad \sigma_{ab}^i \sigma_{bc}^j - \sigma_{ab}^j \sigma_{bc}^i = i \epsilon^{ijk} \sigma_{ac}^k.$$

Le matrici  $\varepsilon^{ijk}$  si denominano *costanti di struttura* dell'algebra. Che esse siano antisimmetriche rispetto lo scambio  $i \leftrightarrow j$  segue subito alla loro definizione. In realtà si prova che esse sono antisimmetriche per ogni scambio di apici; tutte le  $\varepsilon^{ijk}$  con 2 apici uguali valgono quindi 0 e le  $\varepsilon^{ijk}$  non nulle valgono +1 o -1, a seconda siano ottenute con un numero pari o dispari di scambi da  $\varepsilon^{123}$ .

La relazione (X.24) implica che le matrici  $\sigma^i$  costituiscono un'algebra. Esse stesse ne costituiscono una rappresentazione *fondamentale*: non esiste rappresentazione dell'algebra a dimensionalità minore di 2. Se definiamo poi

$$(X.25) \quad E_{bc}^i = -i\varepsilon^{ibc},$$

si verifica facilmente che vale la relazione

$$(X.26) \quad [E^i, E^j] = i\varepsilon^{ijk}E^k,$$

sicché le matrici  $E^i$ , avendo le stesse proprietà di commutazione delle  $\sigma^i$ , costituiscono un'altra rappresentazione della stessa algebra. È questa la rappresentazione che si dice *autoaggiunta*.

Il numero di matrici della base coincide con la dimensionalità della rappresentazione autoaggiunta. Per SU(2), la rappresentazione fondamentale è 2-dimensionale, quella autoaggiunta 3-dimensionale. Nella base di SU(2) solo  $\sigma^3$  è diagonale. Ciò si esprime dicendo che SU(2) è di *rango* 1.

**Verifica di eq. (X.26)** *Se consideriamo l'espressione*

$$(X.27) \quad [[\sigma^i, \sigma^j], \sigma^k] + [[\sigma^k, \sigma^i], \sigma^j] + [[\sigma^j, \sigma^k], \sigma^i] = \\ = i\varepsilon^{ijr}[\sigma^r, \sigma^k] + i\varepsilon^{kir}[\sigma^r, \sigma^j] + i\varepsilon^{jkr}[\sigma^r, \sigma^i] = -(\varepsilon^{ijr}\varepsilon^{rks} + \varepsilon^{kir}\varepsilon^{rjs} + \varepsilon^{jkr}\varepsilon^{ris})\sigma^s$$

è facile vedere, esplicitando la prima riga, che essa è nulla. La verifica si fa direttamente sugli indici:

$$(X.28) \quad [[i, j], k] + [[j, k], i] + [[k, i], j] = \\ (ij - ji)k - k(ij - ji) + (jk - kj)i - i(jk - kj) + (ki - ik)j - j(ki - ik) = \\ = ijk^{-a} - jik^{-b} - kij^{-c} + kji^{-d} + jki^{-e} - kji^{-d} - \\ - ijk^{-a} + ikj^{-f} + kij^{-c} - ikj^{-f} - jki^{-e} + jik^{-b} = 0$$

(l'apice con la freccetta indica i termini omologhi che si semplificano; in questi passaggi si è fatta attenzione a non invertire l'ordine degli addendi). Ne segue che le costanti di struttura obbediscono la condizione

$$(X.29) \quad \varepsilon^{ijr}\varepsilon^{rks} + \varepsilon^{kir}\varepsilon^{rjs} + \varepsilon^{jkr}\varepsilon^{ris} = 0$$

che si può riscrivere, invertendo indici ma mantenendo il posto a ciascun addendo e

ponendo  $\varepsilon^{abc} = iE_{bc}^a$ ,

$$(X.30) \quad E_{jr}^i E_{rs}^k - i\varepsilon^{ikr} E_{js}^r - E_{jr}^k E_{rs}^i = 0$$

che si risistema nella forma

$$(X.31) \quad [E^i, E^k] = i\varepsilon^{ikr} E^r,$$

provando che le matrici  $E^i$  hanno le stesse proprietà di commutazione delle  $\sigma^i$ .

### SU(3)

Di notevole importanza sono le algebre SU(N), che hanno rappresentazione fondamentale di dimensionalità  $N$ , rappresentazione auto-aggiunta di dimensionalità  $N^2 - 1$  e rango  $N - 1$ . In particolare, le matrici della rappresentazione fondamentale di SU(3) sono le  $\lambda^i = A^i/2$  con

$$(X.32) \quad \begin{aligned} A^1 &= \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix}, & A^2 &= \begin{vmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix}, & A^3 &= \begin{vmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{vmatrix}, \\ A^4 &= \begin{vmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{vmatrix}, & A^5 &= \begin{vmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{vmatrix}, \\ A^6 &= \begin{vmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix}, & A^7 &= \begin{vmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{vmatrix}, & A^8 &= \frac{1}{\sqrt{3}} \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{vmatrix}. \end{aligned}$$

Esse soddisfano a proprietà di commutazione  $[\lambda^i, \lambda^j] = i\varepsilon^{ijk} \lambda^k$ , con  $i, j, k$  variabili da 1 a 8. Le costanti di struttura sono ancora completamente antisimmetriche e con componenti di valore 0 o  $\pm 1$ . La novità è che tutte le costanti di struttura in cui compaiono assieme gli apici 3 e 8, anche se hanno tutti gli apici diversi, sono nulle:  $\lambda^3$  e  $\lambda^8$ , matrici diagonali, debbono commutare tra loro.

Tramite le costanti di struttura si possono ancora costruire le 8 matrici  $8 \otimes 8$  della rappresentazione auto-aggiunta.

## 6. Algebre di Lie e interazione dei campi vettoriali

In teoria dei campi la presenza di forze è espressa da termini di interazione nelle densità lagrangiane, spesso chiamati semplicemente (densità) potenziali.

Ad esempio, la forza elettromagnetica nasce dall'interazione tra il 4-vettore densità di corrente  $j^\mu$  (vedi eq. X.1) e la «funzione di campo»  $A^\mu$ . Il potenziale corrispondente

ha la forma

$$(X.33) \quad V_{EM} = j^\mu A^\mu .$$

Se le particelle con carica elettrica sono descritte da un campo scalare complesso  $\phi$  o da un campo spinore  $\psi$ , il vettore  $j^\mu$  assume le forme

$$(X.34) \quad j^\mu = e(\partial^\mu \phi^* \phi - \phi^* \partial^\mu \phi), \quad j^\mu = ie\bar{\psi}\gamma^\mu\psi$$

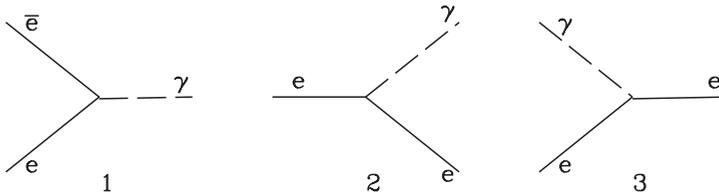
dove  $e$  è la carica dei quanti dei campi, le  $\partial^\mu$  sono simboli di «derivata parziale», lo spinore  $\bar{\psi}$  è costruito con le stesse componenti di  $\psi$  e le  $\gamma^\mu$  sono opportune matrici numeriche  $4 \otimes 4$ , che qui non occorre esplicitare. Va aggiunto che, nel caso di un campo scalare, esiste un ulteriore termine del potenziale, della forma

$$(X.35) \quad \tilde{V}_{EM} = e^2 \phi^* \phi A^\mu A^\mu .$$

In teoria quantistica del campo si può provare una precisa relazione tra le forme dei potenziali e i possibili processi tra quanti dei campi che vi intervengono. Ad esempio, dalle eq. (X.33) e (X.34) ricaviamo il termine potenziale

$$(X.36) \quad V(\psi, A^\mu) = ie\bar{\psi}\gamma^\mu\psi A^\mu g_{\mu\mu}$$

cui seguono i possibili processi in figura X.1.



**FIGURA X.1**

I grafici 1, 2, 3 simboleggiano rispettivamente: l'annichilazione di elettrone e positrone in un fotone, l'emissione di un fotone da parte di un elettrone (o positrone), l'assorbimento di un fotone da parte di un elettrone (o positrone).

Il potenziale è prodotto di 3 funzioni di campo, due campi spinori per elettrone e/o positrone, un campo vettoriale per il fotone. Coerentemente, da ciascun vertice si dipartono due «gambe» spinoriali e una «gamba» vettoriale. I grafi vanno letti da sinistra a destra.

I grafi in figura X.1, considerati isolatamente, non rappresentano processi realmente possibili. Tuttavia, processi come la diffusione Compton o l'interazione elettrone-elettrone si costruiscono con due grafi del tipo in figura X.1, come mostrato nella sottostante figura X.2. A titolo d'esempio discutiamo l'interpretazione da dare al grafo a sinistra, che contribuisce al calcolo dell'effetto Compton.

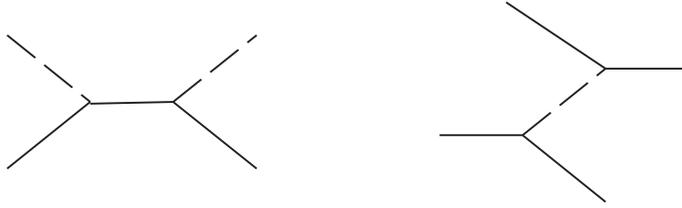


FIGURA X.2

Questi grafi, da leggere da sinistra a destra, contribuiscono al calcolo dell'effetto Compton (destra) e della diffusione elettrone-elettrone (sinistra). Entrambi si compongono dei grafi 2 e 3 di figura X.1, presi in ordine opposto.

CINEMATICA DELL'EFFETTO COMPTON. Siano  $p^\mu, k^\mu$  i 4-impulsi di elettrone e  $\gamma$  incidenti. La relazione impulso-energia-massa implica che sia  $p^\mu p_\mu = m_e^2$ ,  $k^\mu k_\mu = 0$ . La diffusione Compton inizia con l'elettrone che assorbe il fotone; il suo 4-impulso diviene  $p'_\mu = p_\mu + k_\mu$ ; sarà allora

$$(X.37) \quad p'^\mu p'_\mu = (p^\mu + k^\mu)(p_\mu + k_\mu) = p^\mu p_\mu + k^\mu k_\mu + 2k^\mu p_\mu = m_e^2 + 2E_e k(1 - v_e \cos \theta)$$

dove  $E_e$  e  $v_e$  sono l'energia e la velocità dell'elettrone incidente e  $\theta$  è l'angolo tra le direzioni d'incidenza. La quantità in parentesi è sicuramente  $> 1$ ; ne segue che dovrebbe essere  $p'^\mu p'_\mu > m_e^2$ . Di fatto l'energia dell'elettrone, dopo aver assorbito il fotone, sarà minore di quanto prescritto dalla conservazione dell'energia, di una quantità  $\Delta E$ , come è possibile grazie alle relazioni di indeterminazione, per un tempuscolo  $\tau \simeq \hbar/\Delta E$ .

Nel grafo, la «gamba» interna descrive un elettrone che può esistere solo per un breve tempo  $\tau$  in violazione della conservazione dell'energia. L'elettrone «virtuale» così prodotto, prima che  $\tau$  sia trascorso, ri-decade in  $\gamma$  + elettrone, sicché alla fine conservazione di impulso e energia saranno rispettate.

Vediamo ora la forma delle interazioni legate a SU(2). Vi sono anzitutto dei potenziali analoghi a (X.33), che esplicitiamo tenendo presente anche la forma delle correnti costituite da spinori, come in eq. (X.36):

$$(X.38) \quad V_{SU2}(\text{spinori vettore}) = ig \bar{\psi}_a \gamma^\mu \sigma_{Jab} \psi_b \alpha^\mu_J.$$

A parte la diversa denominazione della costante di interazione o carica, indicata con  $g$ , invece di  $e$ , sono qui presenti:

- 1) Il multipletto di spinori  $\psi_a$ ; l'indice  $a$  si somma con analogo indice nella matrice  $2 \otimes 2 \sigma_J$  ed è dunque fatto di 2 componenti:  $\psi_a$  appartiene alla rappresentazione fondamentale dell'algebra.
- 2) Il multipletto di vettori  $\alpha^\mu_J$ , in cui compare l'indice spazio-temporale  $\mu$  e l'indice di SU(2)  $J$ . Il multipletto di campi vettoriali si compone di 3 campi vettori, appartenenti alla rappresentazione auto-aggiunta di SU(3).

Questo potenziale può dunque essere interpretato come somma di 3 termini, ciascuno dei quali esprime l'interazione della componente di  $a_J^\mu$  ( $J = 1, 2, 3$ ). I quanti di ciascun campo  $a_J^\mu$  sarebbero in qualche modo analoghi al fotone e ciascuno di essi avrebbe interazioni con spinori della forma mostrata in figura X.1. La novità consisterebbe nel fatto che le *costanti di interazione*, vale a dire le *cariche* [ $\sim \mathcal{O}(g)$ ] delle forze portate da ciascuno di questi *pseudo-fotoni* sarebbero tra loro «correlate», conformemente al valore delle componenti delle matrici  $\sigma_J$ . Ci sarebbe dunque una precisa «simmetria» tra le interazioni, esprimibile anche tramite proprietà di invarianza, cui corrisponde un'equazione di continuità e una o più quantità conservate, conformemente al teorema di Noether.

A ciò aggiunge un'altra novità, forse più radicale. Oltre al potenziale (X.39), l'uso di SU(2) richiede un'interazione diretta *tra campi vettoriali*. Le trasformazioni rispetto cui le forze sono invarianti sono dettagliate in appendice A3.

Le forze si complicano. Un fotone rilasciato sulla banda dell'ultimo scattering, a un redshift  $z \sim 1000$ , ha il 90% di probabilità di giungere fino all'osservatore terrestre. Uno pseudo-fotone di SU(2), invece, potrebbe avere vita ben più complicata. La sua interazione con altri pseudo-fotoni prescrive che esso stesso abbia una carica, sia sorgente della stessa forza che esso porta. Quindi, nel cammino dal *Big-bang* a oggi, un tale oggetto non sarebbe rimasto tranquillo, ma avrebbe emesso ulteriori oggetti a lui simili, e certo non arriverebbe imperturbato fino a noi.

Riportiamo qui la complessa forma del potenziale vettore-vettore

$$(X.39) \quad V_{SU(N)}(\text{vettore vettore}) = g \varepsilon_{JRS} \partial^\mu a_J^\mu a_R^\mu a_S^\nu + g^2 \varepsilon_{JRS} \varepsilon_{JUV} a_R^\nu a_S^\mu a_U^\nu a_V^\mu,$$

valida per qualsiasi SU(N), salvo usare le relative costanti di struttura. Anche il potenziale spinore-vettore può essere facilmente esteso a SU(3) (e a altre algebre) sostituendo le matrici  $\sigma_J$  con le  $\lambda_J$  o altre opportune matrici.

Tornando nell'ambito di SU(2), le espressioni delle 4-correnti, che soddisfano all'equazione di continuità saranno

$$(X.40) \quad j_J^\mu = ig \bar{\psi}_a \gamma^\mu \sigma_{Jab} \psi_b + g \varepsilon_{JRS} \partial^\mu a_R^\nu a_S^\nu \quad (J = 1, 2, 3),$$

sicché  $j_1^0/c, j_2^0/c, j_3^0/c$  sembrano le densità di 3 cariche conservate. Si noti la presenza in ciascuna  $j_J^\mu$  di un addendo dovuto alle cariche dei quanti vettoriali.

In realtà, solo 1 carica è osservabile, perché solo 1 delle  $\sigma_J$  (ed  $E_J$ , vedi eq. X.25) è diagonale. Di norma la rappresentazione delle matrici rende osservabile  $j_3^0/c$ .

In generale, per SU(N), *il numero effettivo di cariche conservate è pari al rango dell'algebra*.

## 7. La natura delle forze

Queste premesse algebriche serviranno ora a comprendere le forze che osserviamo in natura.

Lasciando a parte la gravità, i cui effetti sono immediatamente discernibili, la quasi

totalità delle altre azioni che sperimentiamo fuori dal laboratorio sono di natura elettromagnetica. Ad esempio, sono effetti elettromagnetici le forze di coesione intermolecolari che tengono aggregato un solido o i covolumi molecolari che danno il volume ai liquidi; allo stesso modo sono di natura elettromagnetica tutte le forze *chimiche* e quindi combustione, ossidazione, evaporazione, condensazione, ecc., sono fatti di natura elettromagnetica.

### **Fenomenologia delle forze nucleari**

Le altre forze sono essenzialmente collocate in ambito nucleare. La forza nucleare *forte* è anzitutto responsabile della coesione dei nuclei atomici, che si manifesta anche tramite il *difetto di massa* dei nuclidi stabili. I protoni e neutroni dei nuclei si scambiano *mesoni*, principalmente mesoni  $\pi$ , la cui massa è  $\sim 138$  MeV. La forza che ne nasce è molto più intensa di quella elettromagnetica; si tratta però di una forza a corto raggio, a causa della massa dei mesoni  $\pi$ . Esistono altre particelle, analoghe a protoni e neutroni, ma con massa maggiore, denominate *iperoni*. Raramente, esse rimpiazzano protoni e neutroni nei nuclei che, in questo caso, sono denominati *ipernuclei*. Esistono anche altri mesoni, il più comune tra essi è il mesone  $\kappa$  ( $m_\kappa \sim 495$  MeV).

La relazione tra massa delle particelle scambiate e *range* della forza fu sottolineata da Yukawa, che notò come il potenziale elettrostatico

$$(X.41) \quad \Phi(r) = e^2 \frac{1}{r}$$

sia a lungo raggio d'azione, grazie alla massa nulla del fotone. Se esso viene sostituito con una particella di massa  $\mu$ , si ha un potenziale

$$(X.42) \quad \Phi(r) = g^2 \frac{e^{-\mu r}}{r},$$

in unità naturali, per cui  $\mu$  è l'inverso di una lunghezza. In unità c.g.s.:

$$(X.43) \quad \mu = \hbar / cr_0$$

ed è facile verificare che, per  $\mu \sim 100$ -200 MeV, si ha  $r_0 \sim 2$ -3  $\times 10^{-13}$  cm. L'andamento esponenziale del *potenziale di Yukawa* (X.42) garantisce che le forze forti si esauriscano a distanze  $\sim r_0$ . Dal *range* sperimentale delle forze forti, che determina le dimensioni dei nuclei, Yukawa fu così in grado di prevedere esistenza e massa del mesone  $\pi$ .

Le simmetrie fenomenologiche delle forze forti sono state all'origine dell'introduzione delle algebre di SU(2) e SU(3). Tali simmetrie, sebbene approssimate, caratterizzano le forze nucleari forti rispetto a quelle deboli. La moderna comprensione delle simmetrie della forza forte ci spiega anche il perché della loro «approssimazione». Ciò che si rileva fenomenologicamente è una specie di *riflesso medio-scopico* di simmetrie esatte a un livello ancora più microscopico.

Sperimentalmente si osserva anche che le forze forti sono invarianti rispetto alle operazioni  $\mathcal{C}$  e  $\mathcal{P}$ .

Le forze nucleari *deboli*, come dice il nome, hanno intensità nettamente inferiore alle forze forti. Esse si osservano perché permettono processi vietati alle forze forti, che violerebbero le simmetrie correlate a SU(2) e/o SU(3) o l'invarianza rispetto a  $\mathcal{C}$  e  $\mathcal{P}$ .

Le particelle osservate si distinguono in *adroni* e *leptoni*, sulla base delle forze che risentono. I leptoni sentono solo le forze deboli. Gli adroni sentono sia le forze forti che quelle deboli. Non va dimenticato il fotone, che non è né adrone né leptone. Ogni processo che coinvolga leptoni, quindi, è reso possibile dalle forze deboli.

Tra gli adroni, riportiamo qui i componenti dell'ottetto barionico di spin 1/2, composto di particelle con *stranezza* 0, 1 e 2. Gli iperoni in questa lista sono stabili rispetto alle forze forti, e decadono in particelle più leggere solo a causa delle forze deboli.

TABELLA 2

<i>nome</i>	<i>m/MeV</i>	<i>carica</i>	<i>stranezza</i>
<i>p</i>	938,27	+1	0
<i>n</i>	939,57	0	0
<i>Λ</i>	1115,68	0	1
$\Sigma^+$	1189.37(±0.07)	+1	1
$\Sigma^0$	1192.55(±0.08)	0	1
$\Sigma^-$	1197.55(±0.03)	-1	1
$\Xi^0$	1314.9(±0.6)	0	2
$\Xi^-$	1321.3(±0.13)	-1	2

I leptoni più comuni sono l'elettrone e il neutrino elettronico. Gli altri leptoni noti sono le particelle  $\mu$  e  $\tau$ , che sono essenzialmente elettroni pesanti, e i rispettivi neutrini muonico e taonico. Esistono quindi 3 *famiglie* di leptoni, che differiscono solo per la massa.

La conta del numero delle famiglie è uno dei problemi ancora insoluti della fisica fondamentale. Che ci siano *almeno* 3 famiglie è un fatto. Non sono state rivelate particelle estranee alle 3 famiglie note, ma nessun dato o teoria vieterebbero ulteriori famiglie.

Commentiamo brevemente sulla possibile esistenza di ulteriori famiglie. Una ispezione della tabella 3 mostra come, lasciando in disparte i neutrini, la particella più leggera di ciascuna famiglia è il leptone carico; un leptone carico con massa 100 volte superiore al  $\tau$  sarebbe sinora sfuggito all'ispezione sperimentale, così come i corrispondenti quark. Un vincolo al numero di famiglie nasce anche dalla cosmologia: la BBN non può permettere l'esistenza di più di  $\sim 3.5$  neutrini con massa inferiore alle decine di keV. Tuttavia, ancora una volta, resta aperta la possibilità che il neutrino della quarta famiglia abbia effettivamente massa e che il suo valore sia al di sopra di questi limiti.

I processi di nucleosintesi stellare e primordiale sono resi possibili dalle forze deboli. Ad esempio, nel *core* del Sole avvengono reazioni che portano 4 protoni e 2 elettroni a trasformarsi in un nucleo di  ${}^4\text{He}$ , con emissione di energia pari a  $\sim 0.7\%$  della massa delle particelle interagenti. L'insieme delle reazioni che rendono possibile la sintesi dell'elio è dovuto a forze deboli.

### ***Origine della forza forte. La quanto-cromo-dinamica***

C'è un iato significativo tra questa fenomenologia delle forze forti e la loro comprensione teorica. Ciò corrisponde alla notevole «fatica» occorsa a trovare il bandolo di una matassa assai ingarbugliata.

La descrizione delle forze forti si basa sul fatto che gli adroni sono essi stessi particelle composte. Nel sistema di scatole cinesi che va dalla molecola all'atomo e dall'atomo al nucleo, si apre qui un altro gradino di scala e i nucleoni stessi vengono interpretati come realtà composite.

Le particelle *veramente* elementari, protagoniste delle forze forti, sono dunque i *quark* e i *gluoni*. C'è una stretta similitudine tra forze elettromagnetiche e forti. Quark e gluoni sono, rispettivamente, gli elettroni e i fotoni delle forze forti. Ciò che rompe brutalmente la similitudine, tuttavia, è lo *status* di queste particelle, nella nostra epoca cosmica. Mentre elettroni e fotoni sono particelle comuni, quark e gluoni non possono essere isolati e ciò mette in causa, in qualche modo, lo stesso significato del termine «particella» applicato a queste entità.

TABELLA 3

	<i>Quark: denominazione (masse in MeV)</i>		<i>leptoni carichi</i>	<i>neutrini</i>
cariche:	$2/3$	$-1/3$	$-1$	$0$
<i>u</i> : up (2-8)	<i>d</i> : down (17-25)		elettrone (0.511)	$\nu_e$
<i>c</i> : charmed ( $1-1.6 \times 10^3$ )	<i>s</i> : strange (100-300)		$\mu$ (106)	$\nu_\mu$
<i>t</i> : top ( $130-150 \times 10^3$ )	<i>b</i> : bottom ( $4.1-4.5 \times 10^3$ )		$\tau$ (1777)	$\nu_\tau$

Un'ulteriore differenza tra elettroni-fotoni e quark-gluoni sta nelle simmetrie delle interazioni di cui sono partecipi. Nel primo caso la simmetria di fase (vedi eq. X.2) può essere anche ricondotta a un'algebra, che si chiama U(1); si tratta dell'algebra delle rotazioni sul piano complesso. La simmetria delle forze forti, invece, è SU(3). Quark e antiquark saranno nelle due rappresentazioni fondamentali, 3-dimensionali, di SU(3). I gluoni saranno nella rappresentazione autoaggiunta, 8-dimensionale, di SU(3). I tre quark si distinguono per diversi valori di *numeri quantici* associati a  $\lambda_3$  e  $\lambda_8$ . Quark e antiquark si collocano sul piano  $\lambda_3$ ,  $\lambda_8$  come descritto in figura X.3.

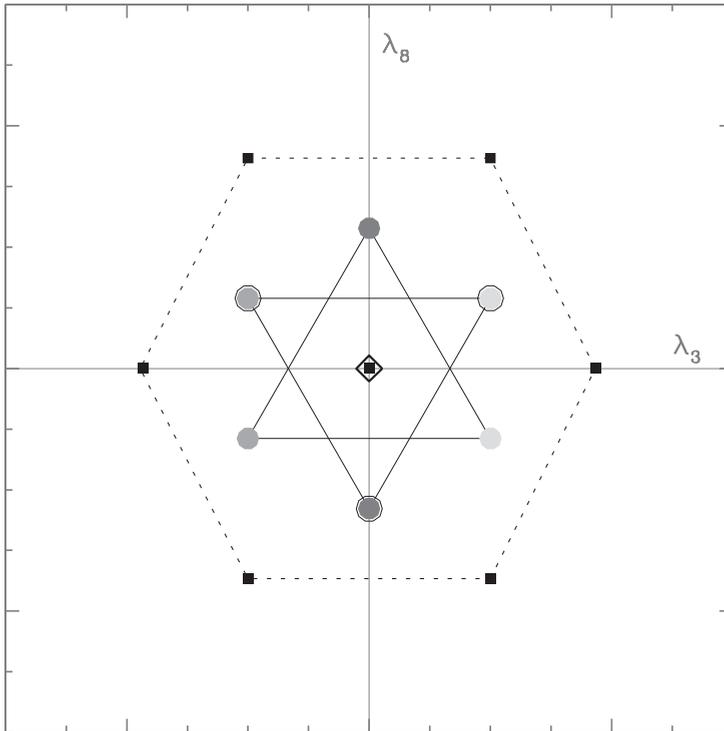


FIGURA X.3

Quark e antiquark distinti per colore. Sui lati dell'esagono si trovano invece le posizioni sullo stesso piano dei gluoni colorati. Altri 2 gluoni si collocano nel punto  $\lambda_3 = \lambda_8 = 0$ . (si veda l'appendice TAVOLE A COLORI).

I siti del piano in cui si collocano i quark possono essere direttamente ricavati dai valori delle componenti diagonali di  $A_3$  e  $A_8$  (eq. X.32). La collocazione degli anti-quark e dei gluoni si può dedurre allo stesso modo. Non daremo dettagli in merito, salvo osservare il parallelismo tra i lati dell'esagono esterno e quelli del triangolo interno, che mostrano come il passaggio da un vertice all'altro del triangolo si possa ricavare tramite lo scambio di un gluone.

Le 3 collocazioni dei quark vengono individuate da un *colore* e, per questo, la carica forte – una carica bi-dimensionale, a differenza della carica elettrica – viene denominata colore.

### Confinamento

L'elemento chiave della quanto-cromo-dinamica (QCD) è che i gluoni sono colorati. Un quark può emettere un gluone, come un elettrone emette un fotone nel grafo 2 di figura X.1 Il gluone può poi essere assorbito da un altro quark, come un elettrone assorbe un fotone nel grafo 3 di figura X.1. In questo modo due elettroni interagiscono

e, a seconda della carica, si attraggono o respingono. Allo stesso modo due quark interagiscono, con forza attrattiva o repulsiva a seconda del colore. Tutto ciò è vero se la distanza tra quark è limitata. Se la distanza che un gluone deve coprire è troppo elevata, invece, interviene la carica del gluone stesso, che diviene sorgente di ulteriori gluoni. Denominiamo  $\bar{d}$  l'ordine di grandezza di tale distanza critica. Il processo si ripete per i gluoni eventualmente figliati, in una *cascata gluonica* che può esistere solo con grande spesa energetica.

Queste considerazioni vanno prese assieme a un altro elemento chiave della QCD, la cosiddetta *asymptotic freedom* (libertà asintotica): in sostanza, se  $T$  è molto elevata, quark e gluoni si distribuiscono come particelle libere.

Abbiamo quindi senz'altro due possibili stati per la materia con interazioni forti. A alta temperatura c'è un *plasma quark-gluonico*, in cui quark e gluoni sono praticamente liberi. A densità  $< \bar{n} \simeq \bar{d}^{-3}$ , i quark si raccolgono invece in sistemi di 2 o 3 particelle (*adroni*) la cui carica-colore è nulla. Il legame interno agli adroni è dovuto allo scambio di gluoni. All'esterno permangono solo blande azioni, tipo «forze di *Van der Waals*» tra particelle neutre, ma non c'è scambio di gluoni. Abbiamo allora un *gas adronico*.

Nel cosmo primordiale deve essere avvenuta una transizione da plasma a gas, a una determinata temperatura. Ne tratteremo al paragrafo 11.

Per formare un sistema con autovalori nulli di  $\lambda_3$  e  $\lambda_8$ , si possono:

- 1) Raccogliere assieme 3 quark  $q$  dei 3 colori, che si sovrappongono dando un sistema «bianco», un barione.
- 2) Mettere assieme un quark  $q$  e un anti-quark  $\bar{q}$  su due vertici opposti dei due triangoli; i due colori si cancellano dando un sistema «nero», un mesone.

La transizione da plasma a gas porta così dal colorato mondo primordiale al mondo adronico, in bianco e nero.

Sebbene non ci sia un valore preciso per la distanza  $\bar{d}$ , oltre cui si attiva il colore dei gluoni, una stima può derivare dalle dimensioni fisiche degli adroni. In effetti, i barioni e, in misura minore, i mesoni hanno una zona centrale praticamente impenetrabile, denominata *hard core* (nucleo duro), di dimensioni  $r_{hc} \sim 10^{-13}$  cm, distanza probabilmente prossima al valore di  $\bar{d} \simeq \bar{n}^{-1/3}$ .

Entro l'*hard core* si trovano 3 quark o una coppia  $q\bar{q}$ , a seconda che l'adrone sia un barione o un mesone. Dal momento che un sistema  $qqq$  è un barione, ogni quark deve avere numero barionico  $B = 1/3$ .

In figura X.3 vediamo come si interpretano, in questo schema, l'emissione e l'assorbimento di un mesone da parte di un barione.

È opportuno ribadire, quindi, che il tipo di legame tra quark, che dà origine agli adroni, è assai diverso da quello che permette ai nucleoni di formare i nuclei atomici. Un nucleo è stabile grazie al difetto di massa  $\Delta m$ , rispetto alla somma delle masse dei nucleoni che lo compongono, e può essere «slegato» fornendo al nuclide l'*energia di legame*  $\Delta m c^2$ . Le masse dei quark più leggeri, invece, sono dell'ordine di qualche

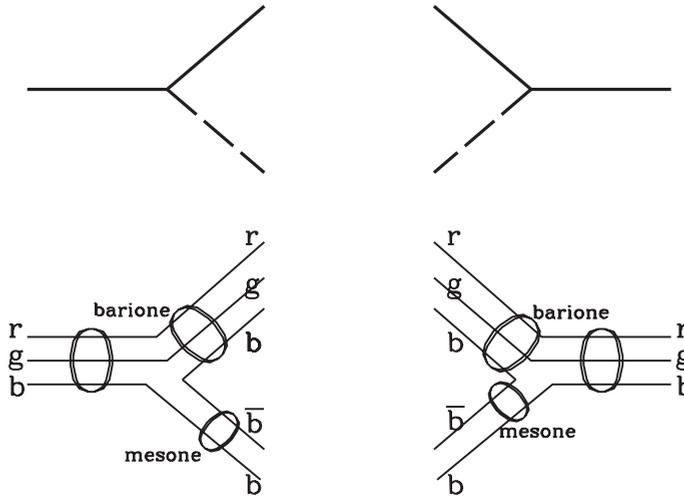


FIGURA X.4

Emissione (a sinistra) e assorbimento (a destra) di un mesone da parte di un barione. I grafi in alto rappresentano il processo *medio*-scopico. I grafi in basso lo interpretano in termini di quark [red ( $r$ ), green ( $g$ ) e blue ( $b$ )] e antiquark  $\bar{b}$ . L'emissione avviene tramite una simultanea creazione di  $b, \bar{b}$ . L'assorbimento si realizza tramite un'annichilazione  $b, \bar{b}$ .

MeV (vedi tabella 3) e la loro somma è assai minore della massa dell'adrone che si forma tramite il legame gluonico. Ciò che si osserva come massa è dunque essenzialmente energia potenziale e cinetica interna al sistema. Il legame non è dovuto a un difetto di massa, ma all'impossibilità di creare sistemi isolati colorati.

### Il sapore

Nella nostra epoca cosmica, dunque, non esistono quark (o gluoni) liberi. Tuttavia, negli urti profondi, la struttura *partonica* degli adroni può essere sperimentalmente osservata.

La simmetria di colore è una simmetria esatta. Tuttavia, i quark hanno un ulteriore numero quantico, il *sapore*. Tale numero riguarda le forze deboli e la distribuzione su famiglie. A ciascun sapore corrisponde una massa diversa. I quark di un adrone hanno colori opportunamente distribuiti; i loro sapori, invece, possono essere eguali o diversi. A seconda delle distribuzioni di sapori, avremo sistemi leggermente diversi che però manifestano ancora, in forma distorta, la fondamentale simmetria di colore. SU(3) fu storicamente introdotta per rendere conto delle simmetrie nelle interazioni di adroni contenenti i quark  $u, d$  e  $s$ , che sono i più leggeri, in tutte le possibili distribuzioni (vedi tabella 4).

I sapori possono essere raccolti in coppie nello schema di *famiglie* di particelle elementari di tabella 3. Ci sono dunque 6 sapori di quark.

A questo proposito è importante segnalare che, in unità di carica elettronica, i quark hanno carica frazionaria: nella prima colonna la carica è  $2/3$ ; nella seconda

colonna 1/3. Qui sotto riportiamo la composizione delle particelle dell'ottetto barionico (Tab. 2).

Esse comprendono tutte le possibili combinazioni dei 3 quark più leggeri. Barioni contenenti i quark più pesanti – e quindi essi stessi assai più massicci – sono stati successivamente scoperti.

TABELLA 4

	<i>composizione</i>	<i>cariche</i>
p	<i>u u d</i>	$2/3 + 2/3 - 1/3 = 1$
n	<i>u d d</i>	$2/3 - 1/3 - 1/3 = 0$
$\Lambda$	<i>u s d</i>	$2/3 - 1/3 - 1/3 = 0$
$\Sigma^+$	<i>u u s</i>	$2/3 + 2/3 - 1/3 = 1$
$\Sigma^0$	<i>u d s</i>	$2/3 - 1/3 - 1/3 = 0$
$\Sigma^-$	<i>d d s</i>	$-1/3 - 1/3 - 1/3 = -1$
$\Xi^0$	<i>u s s</i>	$2/3 - 1/3 - 1/3 = 0$
$\Xi^-$	<i>d s s</i>	$-1/3 - 1/3 - 1/3 = -1$

## 8. Il modello elettrodebole

Consideriamo un processo dovuto alle forze deboli, come il decadimento  $\beta$ :

$$n \rightarrow p + e + \bar{\nu}$$

Il processo è dovuto allo scambio di un bosone di gauge carico, denominato  $W$ , come nel grafico in figura X.5. Il processo «inizia» quando un neutrone, con massa  $m_n \sim 0.9 \text{ GeV}$ , «decade» in un protone e in una particella  $W$ , di massa  $M_W \simeq 80 \text{ GeV}$ , circa 90 volte la particella «madre». Ciò viola clamorosamente la conservazione dell'energia ed è possibile per un tempuscolo  $\tau \sim \hbar/M_W c^2$ , solo grazie alle relazioni di indeterminazione. Il  $W$  «successivamente» decade in elettrone e anti-neutrino. La brevissima vita media del  $W$  è la prima ragione che rende «deboli» le interazioni deboli.

Se indichiamo con  $g_2$  la carica debole del  $W$ , la costante fenomenologica delle forze

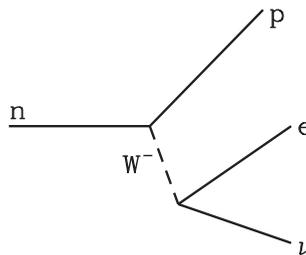


FIGURA X.5

Decadimento  $\beta$  con l'intervento del bosone di gauge  $W^-$ .

deboli, detta *costante di Fermi* (vedi appendice A4), risulta

$$(X.44) \quad \frac{G_F}{\sqrt{2}} = \frac{g_2^2}{8M_W}.$$

$G_F$  e  $M_W$  hanno valori sperimentalmente noti; si può quindi valutare  $g_2$  e si scopre che questa costante è molto prossima alla carica  $e$  dell'elettrone. Ciò si esprime nelle relazioni

$$(X.45) \quad e = g_2 \sin \theta_W \quad \text{con} \quad \sin^2 \theta_W \simeq 0.23$$

che mettono in evidenza la contiguità tra forze deboli e elettromagnetiche, collegando la fenomenologia di bassa energia, in cui le forze deboli sembrano non aver nulla a che fare con l'elettromagnetismo, e dinamica di gauge, in cui le forze elettromagnetiche e deboli si dimostrano connesse.

L'unificazione di forze deboli e elettromagnetiche è però già dietro la porta. Assumiamo dunque che le forze deboli siano simmetriche rispetto a SU(2) e siano portate dai 3 bosoni di gauge  $W_J^\mu$  ( $J = 1, 2, 3$ ) appartenenti alla sua rappresentazione auto-aggiunta. Accanto a SU(2) assumiamo anche una simmetria per trasformazioni di fase, che si denomina U(1), e si potrebbe pensare all'unico bosone di questa gauge abeliana come al fotone. Il meccanismo, in natura, è più complesso. Il campo di gauge di U(1) va denominato  $Y^\mu$ . Cosa sia il fotone, in questo schema, lo vedremo tra poco e la frammistione elettro-debole nasce proprio da questo.

Le forze deboli e elettromagnetiche hanno infatti una dinamica diversa a alta o bassa temperatura. I due regimi sono separati da una temperatura  $T_{EW} \sim 100$  GeV, a cui si pensa avvenga una transizione di fase del primo ordine (la natura della transizione elettrodebole è legata al valore della massa del bosone di Higgs, non ancora nota). Come  $H_2O$  cambia stato di aggregazione alla temperatura di 273.15 K, così esistono campi e quanti diversi al di sotto o al di sopra di  $T_{EW}$ .

I campi  $W_J^\mu$  e  $Y^\mu$ , sin qui indicati, sono i campi di alta temperatura. I loro quanti, a alta temperatura, non hanno massa, come oggi il fotone. Le forze di alta temperatura sono simmetriche rispetto a un ampio gruppo di trasformazioni, che dipende da 3 parametri per SU(2) e da un quarto parametro per U(1).

Quando il cosmo, raffreddandosi, passa attraverso la temperatura  $T_{EW}$ , la dinamica delle forze cambia radicalmente.

1) I 4 campi bosonici  $W_J^\mu$  e  $Y^\mu$  si «mescolano» tra loro, dando origine ai campi

$$(X.46) \quad W^{\mu\pm} = \frac{1}{\sqrt{2}}(W_1^\mu \mp W_2^\mu), \quad Z^\mu = \cos \theta_W W_3^\mu - \sin \theta_W Y^\mu,$$

$$(X.47) \quad A^\mu = \sin \theta_W W_3^\mu + \cos \theta_W Y^\mu$$

Al di sotto di  $T_{EW}$ , le particelle sperimentalmente osservabili saranno i quanti dei campi  $W_\pm^\mu$ ,  $Z_0^\mu$  e il fotone.

2) Nella transizione, i quanti dei campi  $W$  e  $Z$  acquisiscono masse il cui rapporto è  $M_W/M_Z = \cos \theta_W$ , mentre il campo  $A^\mu$  – il campo elettromagnetico – è l'unico a restare a massa nulla.

Questa è dunque la natura del campo elettromagnetico: la sua straordinaria rilevanza fenomenologica nasce dall'essere l'unica miscela dei campi primordiali  $W$  e  $Y$  a conservare la capacità di determinare forze a lungo raggio d'azione.

Miscelamento e masse distruggono l'ampia simmetria  $SU(2) \otimes U(1)$  di alta temperatura. L'unica simmetria residua è definita dall'algebra di  $U(1)$  per la gauge elettromagnetica.

Il ruolo delle diverse particelle e lo schema complessivo del modello elettrodebole sono ulteriormente dettagliati in appendice A4.

L'angolo  $\theta_W$  (angolo di Weinberg) ricorre in tutte queste relazioni e, comunque lo si misuri, si ottiene lo stesso valore. Dei dettagli del processo che porta all'acquisizione di massa è responsabile il *meccanismo di Higgs* (vedi appendice A5), che implica un'interazione con un campo scalare *ad-hoc* i cui quanti, come già segnalato, non sono stati ancora scoperti.

L'ispezione di questa fisica ha richiesto l'uso di acceleratori capaci di provocare collisioni in cui si raggiungano i 100 GeV nel sistema di riferimento del centro massa. Attualmente viene costruito un acceleratore capace di giungere fino alle decine di TeV. Oltre a questo livello energetico, tuttavia, la tecnologia attuale degli acceleratori difficilmente potrà spingersi.

Ciò è causa di non poco disappunto, dal momento che lo stesso meccanismo che ha permesso di prevedere la transizione di fase alla temperatura  $T_{EW}$ , suggerisce un'ulteriore transizione a una temperatura  $T_{GUT} \sim 10^{15}$  GeV, come vedremo nel prossimo paragrafo.

## 9. Il gruppo di rinormalizzazione

Nella relazione tra dinamica delle interazioni e temperatura sta la chiave dell'importanza assunta dalla cosmologia, come campo di verifica della fisica delle particelle elementari. Alla base di questa relazione sta poi il *gruppo di rinormalizzazione*, che prescrive alle forze di cambiare in funzione di  $T$  secondo leggi ben precise. Andando a ritroso nel tempo, dunque, la dinamica delle forze cambia. Allo stesso modo, quando in laboratorio si esplora la dipendenza delle forze dalle energie, si può legittimamente affermare che si tenta di ricostruire le condizioni dell'Universo primordiale.

Per trattare del gruppo di rinormalizzazione occorre anzitutto discutere in che senso la dinamica delle interazioni fondamentali richieda di essere «rinormalizzata». La dinamica di un campo  $\psi$ , di norma, si ottiene risolvendo un'equazione differenziale, la cui forma dipende dalla rappresentazione del gruppo di Lorentz cui  $\psi$  appartiene. In tutte le possibili equazioni, comunque, possono comparire una massa  $m$ , interpretabile come massa dei quanti del campo, e le costanti di interazione  $g$  tra  $\psi$  e altri campi (o anche di  $\psi$  con se stesso). Sono massa  $m$  e carica  $g$ , *in primis*, che debbono essere rinormalizzate. Un esempio permetterà di capire la natura del problema.

Se il campo  $\psi$  ha come quanti gli elettroni, ci si può chiedere quali siano le «di-

mensioni fisiche» dei quanti stessi. In effetti, se immaginiamo la carica  $e$  dell'elettrone omogeneamente distribuita sulla superficie di una sferetta di raggio  $r_e$ , il potenziale a cui l'elettrone viene a trovarsi è  $e/r_e$ , per cui esso è dotato di un'energia potenziale  $\simeq e^2/r_e$ . Naturalmente,  $r_e$  sarà piccolo e l'energia potenziale grande. Uguagliandola all'energia di massa  $m_e c^2$ , si ottiene

$$(X.48) \quad r_e = e^2/m_e c^2,$$

il cosiddetto *raggio classico dell'elettrone*, che entra, tra l'altro, nell'espressione della sua sezione d'urto. Esperimenti di alta energia, tuttavia, mostrano che l'elettrone ha raggio  $\ll r_e$ . La sua auto-energia rischia quindi di essere infinita.

Il processo responsabile per questa divergenza energetica è descritto dal grafo (a) di figura X.6: l'elettrone emette un fotone «virtuale» e poi lo riassorbe, in tempo per stare entro le trasgressioni ammesse dalle relazioni di indeterminazione.

Il contributo del processo è tanto più elevato, quanto più vicini sono emissione e riassorbimento del fotone: piccola distanza significa carica concentrata in un piccolo raggio. Se si devono considerare contributi fino a distanza nulla, l'auto-energia dell'elettrone diverge: la sua massa riceverebbe un contributo infinito. Se si trova invece un modo fisico di porre un *cut-off* alle distanze, il contributo sarà grande ma non  $\infty$ . In realtà, invece delle distanze si limitano i *numeri d'onda* delle onde elettromagnetiche emesse e riassorbite, ovvero gli impulsi dei fotoni. Il *cut-off* non deve introdurre scale spurie nel risultato e il procedimento accurato che si segue prende il nome di *regolarizzazione*.

Divergenze analoghe sarebbero presenti anche nel calcolo di «vertici» e della «polarizzazione del vuoto» (grafi (b) e (c) di figura X.6). La regolarizzazione viene applicata a questi processi e a ogni altro processo elementare. Come conseguenza avremo una massa e una carica normalizzate:  $m(\mu)$ ,  $g(\mu)$ ;  $\mu$  è la scala di regolarizzazione, che appare quindi nella valutazione dei singoli grafi, *ma da cui non deve dipendere alcuna quantità fisica*: tra i diversi grafi deve avvenire una opportuna compensazione, possibile se la teoria non ha una scala intrinseca. Il cambio di  $\mu$  equivale allora a un cambio nella scala dei numeri d'onda  $k$ .

Nell'ambito della meccanica quantistica a temperatura finita, invece, entra in gioco la scala fissata dalla temperatura  $T$ . A ciascuna temperatura potrà esserci cancellazione, ma i risultati dipenderanno dal rapporto  $\mu/T$ .

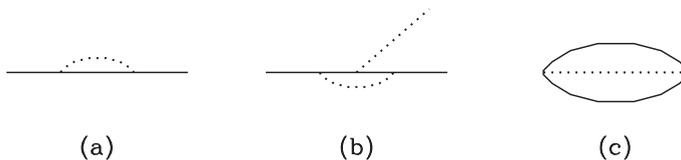


FIGURA X.6

Processi elementari che portano a divergenze: (a) divergenza sulla massa; (b) divergenza sulla costante di interazione; (c) divergenza dovuta a polarizzazione del vuoto.

Nelle forze elettrodeboli entrano due costanti di interazione  $g_1$  e  $g_2$ , legate ai gruppi U(1) e SU(2). In QCD entra la costante  $g_3$  legata a SU(3). L'andamento con  $T$  delle  $g_i$  ( $i = 1, 2, 3$ ) rinormalizzate si esprime tramite le quantità

$$(X.49) \quad \alpha_i = g_i^2/4\pi;$$

e dipende dalle costanti

$$(X.50) \quad \beta_{2,3} = \frac{1}{4\pi} \left( N_f \frac{4}{3} - N \frac{11}{3} \right), \quad \beta_1 = \frac{1}{4\pi} N_f \frac{20}{9},$$

nel caso di teorie di gauge basate su SU(N) ( $N=2,3$ ) e U(1);  $N_f$  è il numero delle famiglie o generazioni di fermioni. Più specificamente risulta

$$(X.51) \quad \begin{aligned} 4\pi\beta_3 &= -7 \\ 4\pi\beta_2 &= -10/3 \\ 4\pi\beta_1 &= +20/3 \end{aligned}$$

Queste costanti entrano nelle leggi

$$(X.52) \quad \frac{1}{\alpha(T)} = \frac{1}{\alpha(\mu)} - \beta \ln \left( \frac{T}{\mu} \right).$$

che determinano la dipendenza delle costanti di interazione dalla temperatura. Chiaramente solo  $1/\alpha_1$ , cui corrisponde  $\beta_1 > 0$ , cala al crescere di  $T$ , per cui  $g_1$  cresce. Viceversa, sia  $g_3$  che  $g_2$  dovranno calare al crescere di  $T$  e la rapidità del calo di  $g_3$  sarà maggiore di quella di  $g_2$ .

A questo punto interviene un dato sperimentale: a  $T \sim 0$ , oggi, è  $g_3 > g_2 > g_1$ . Al crescere di  $T$ , dunque, le 3 costanti si avvicinano; in particolare,  $g_3$  incrocerà sia  $g_2$  che  $g_1$ . Ebbene, i valori osservati delle 3 costanti fan sì che gli incontri  $g_3$ - $g_1$  e  $g_3$ - $g_2$  avvengano alla stessa temperatura  $T_{GUT}$ .

Possiamo dire altrimenti: esiste una temperatura  $T_{GUT}$  a cui le 3 costanti hanno un comune valore  $\alpha_{GUT} \equiv \alpha_i(T_{GUT})$  ( $i = 1, 2, 3$ ). Da  $\alpha_{GUT}$ , usando eq. (X.53), si possono dedurre i valori

$$(X.53) \quad \frac{1}{\alpha_i(T)} = \frac{1}{\alpha_{GUT}} - \beta_i \ln \left( \frac{T}{T_{GUT}} \right)$$

delle costanti a ogni temperatura  $T$ .

**Commenti.** Anzitutto, le costanti  $g_i$  sono espresse tramite le  $\alpha_i$  (eq. X.49) allo stesso modo in cui, in elettromagnetismo,  $\alpha = e^2/4\pi$  rappresenta la cosiddetta *costante di struttura fine* (valore spettroscopico:  $\alpha = 1/137.036$ ).

Va poi sottolineato che le costanti  $\beta_i$  hanno l'espressione (X.50) se calcolate in modo perturbativo, all'ordine più basso. Il loro valore è oggi noto con maggior precisione, sebbene le correzioni siano piccole. Tuttavia, anche blande modifiche nei coefficienti  $\beta$  possono compromettere il «triplo incrocio» delle curve  $\alpha_i(T)$  e/o modificare le previ-

sioni su  $T_{GUT}$ . Va quindi detto che la «precisione» del triplo incrocio era consistente con la precisione con cui le  $\beta$  erano note dai primi calcoli, fatti senza tener conto delle correzioni agli ordini successivi. Al livello di precisione determinato da previsioni più rigorose, che tengono conto di processi di ordine più elevato, il «triplo incrocio» è mancato, seppure di poco. Non si tratta però di un risultato negativo; lo si supera tenendo conto delle *supersimmetrie* (vedi sotto), che anzi trovano in questo dato una forte motivazione sperimentale.

Queste conclusioni sono uno stringente invito a una ispezione sperimentale dell'effettiva convergenza tra costanti di interazione e della nuova fisica a  $T \sim T_{GUT}$ . Il problema è che  $T_{GUT}$  può essere «facilmente» calcolata e ha un valore  $\sim 10^{15}$  GeV, molto al di là dei  $\sim 10^4$  GeV che appaiono come l'ultima frontiera delle attuali capacità sperimentali. Finite le potenzialità della fisica di laboratorio, tuttavia, rimane aperta la strada dell'ispezione cosmologica.

## 10. Supersimmetrie in cosmologia

Alle energie concesse alle esperienze dei grandi laboratori, invece, potrebbe già essere osservabile la *nuova fisica* prevista dalla *supersimmetrie*.

I campi che ci sono sperimentalmente noti hanno quanti fisici con un unico spin. Elettrone e fotone, ad esempio, hanno rispettivamente spin  $s_e = 1/2$  e  $s_\gamma = 1$ . La componente di  $s = 0$  del campo elettromagnetico, che abbiamo discusso, deve essere accuratamente eliminata dalla realtà fisica.

In regime supersimmetrico, invece, elettrone e fotone – come tutte le altre particelle del modello standard – dovrebbero acquisire ulteriori stati di spin, oggi portati da ipotetiche particelle, mai rivelate sperimentalmente ma per ragioni consistenti con la teoria, denominate *fotino* e *seletrone*. A questo punto, la differenza tra elettrone e fotone consisterebbe solo nelle forze che essi attivano e sentono, non nella loro statistica, che è poi il principale elemento che li rende oggi così diversi fenomenologicamente.

Varie ipotesi si confrontano sulla temperatura ( $T_{SUSY}$ ) a cui ci si attende la transizione graduale (*soft*) tra mondo supersimmetrico e mondo attuale, ma le prospettive, in questo caso, non sono del tutto inaccettabili. Il progetto LHC, che implementa il LEP ginevrino, è basato sull'aspettativa che  $T_{SUSY}$  si possa collocare tra il TeV e la decina di TeV: una previsione con discreto fondamento teorico; ma, se l'aspettativa andasse delusa, nessuno sarebbe poi così sorpreso.

## *Neutralini come DM*

Alcune conseguenze delle supersimmetrie, invece, potrebbero essere già state osservate in astrofisica. Tra le ipotesi sulla natura della materia oscura (DM), una delle più accreditate fa proprio riferimento alle SUSY. La DM è una sostanza che oggi appare priva di ogni interazione, a parte le forze gravitazionali; la sua necessità è stata am-

piamente dibattuta nei precedenti capitoli. Ebbene, la DM potrebbe essere costituita da residui cosmologici di partner supersimmetrici di particelle note. A temperature superiori a  $T_{SUSY}$  i partner supersimmetrici fanno parte del *brodo termico* primordiale, con tutti i loro stati di spin bosonici e fermionici. Con la rottura *soft* di SUSY, la maggior parte dei partner supersimmetrici acquisisce elevata massa e decade rapidamente in particelle più leggere, di norma in particelle ordinarie. È però possibile che si formi un residuo di partner SUSY, composto di particelle che denomineremo  $\chi$ , incapaci di decadere in particelle ordinarie, la cui capacità di interagire con le altre componenti cosmiche diviene estremamente ridotta, specialmente a causa dell'elevata massa acquisita da altre particelle SUSY che, allo stato virtuale, dovrebbero intervenire nei processi elementari che assicurano l'interazione tra  $\chi$  e particelle ordinarie.

Se i  $\chi$  esistono, debbono avere carica e colore nulli, essere quindi partner di particelle neutre e non colorate. Inoltre i  $\chi$  saranno fermioni, di modo che sia vietato il decadimento diretto di  $\chi$  in coppie particella-antiparticella ordinarie.

Come fermioni essi saranno partner SUSY di bosoni ordinari. L'elenco dei bosoni neutri del modello standard supersimmetrizzato comprende:

- 1) due bosoni di gauge: il fotone e la particella  $Z$ ;
- 2) i bosoni di Higgs (in una teoria supersimmetrica si richiede, per la rottura della simmetria elettrodebole, l'esistenza di 2 bosoni di Higgs  $H_1$  e  $H_2$ ).

I fermioni corrispondenti sono quindi due *gaugini* (*fotino* e *zino*) e due *higgsini*. La teoria non permette di definire quale sia la più leggera di queste particelle e tanto meno di determinarne le masse. Un'ulteriore complicazione nasce dal fatto che gli *autostati* di massa potrebbero essere miscele quantiche delle particelle qui elencate, che denomineremo  $\chi_i$  ( $i = 0, \dots, 3$ , in ordine di massa crescente). La particella più leggera tra esse,  $\chi_0 \equiv \chi$ , viene denominata *neutralino* o tramite l'acronimo *LSP* (lightest supersymmetric particle).

Lo schema dettato da questa microfisica è abbastanza univoco. Tutti i  $\chi_i$  sono fermioni simmetrici per coniugazione di carica (spinori *di Majorana*, vedi appendice A2):  $\chi_i$  e  $C\chi_i$  coincidono.

Le reazioni che portano al formarsi del fondo di  $\chi_0$  sono anzitutto del tipo

$$\chi_i + \chi_i \rightarrow \chi_{i-1} + \chi_{i-1},$$

possibili evidentemente finché  $i \geq 1$ . Una volta che si arriva a  $\chi_0$ , un ulteriore degrado non è possibile. Accanto ai canali SUSY sono possibili canali del tipo

$$\chi_i + \chi_i \rightarrow \text{particella} + \text{antiparticella ordinarie},$$

cui corrisponde una sezione d'urto  $\sigma_{an}$  assai più piccola e quindi un ritmo nettamente inferiore, essendo essi dovuti allo scambio di altre particelle SUSY di altissima massa, allo stato virtuale. Una volta arrivati alla formazione di  $\chi_0$ , tuttavia, solo questi canali restano aperti.

È stata anche considerata la possibilità di reazioni del tipo

$$\chi + \chi \rightarrow \chi + \text{particella o antiparticella ordinaria.}$$

che sarebbero assai più veloci. Nella maggior parte dei modelli, tuttavia, esiste un numero quantico, detto  $\mathcal{R}$ -parità, che distingue particelle SUSY da particelle ordinarie. Una reazione di questo tipo, violando l' $\mathcal{R}$ -parità, sarebbe esclusa. Se essa è possibile, anche a ritmo estremamente basso, il fondo cosmico di  $\chi$  praticamente scompare.

Sulla base della loro sezione d'urto  $\sigma_{an}$  si valuta allora il tempo caratteristico

$$(X.54) \quad \tau_\chi = 1 / \langle n_\chi \sigma_{an} v_T \rangle$$

che si confronta col tempo cosmico  $t$ . La temperatura  $T_{\chi,dg}$ , di «disaccoppiamento» dei  $\chi$ , è quella a cui  $\tau_\chi \sim t$ . Per  $T < T_{\chi,dg}$  vanno considerati trascurabili:

- 1) il numero di particelle che si trasformano in  $\chi$ , ma, soprattutto,
- 2) il numero di  $\chi$  che si trasformano in altre particelle (per ulteriori chiarimenti quantitativi vedi appendice A6).

Su questa base si studiano i limiti cosmologici alla massa del neutralino ( $m_\chi$ ). Essi sono particolarmente stringenti se  $T_{\chi,dg} \gg m_\chi$ , cioè se il disaccoppiamento avviene quando i  $\chi$  sono ancora relativistici (*disaccoppiamento caldo*). In questo caso  $m_\chi$  dovrebbe essere al più di qualche centinaio di keV. Altrimenti si avrebbe un parametro di densità attuale, per i  $\chi$ ,  $\Omega_{o,\chi} \gg 1$ : un problema molto più serio della natura della DM.

La situazione è molto più promettente se  $T_{\chi,dg} < m_\chi$  (ma non troppo minore, come vedremo), cioè se i  $\chi$  disaccoppiano dopo essere divenuti non relativistici (*disaccoppiamento freddo*). In particolare, se è  $T_{\chi,dg}/m_\chi \sim 0.1-0.2$ , la DM può essere costituita di neutralini. Vediamo più in dettaglio i due casi.

### **Disaccoppiamento caldo**

La densità in numero di particelle  $\chi$ , di massa  $m_\chi$ , fino alla temperatura  $T_{\chi,dg}$  ( $\gg m_\chi$ ), è

$$(X.55) \quad n_\chi(T_{\chi,dg}) = \frac{\zeta(3)}{\pi^2} \tilde{g}_\chi T_{\chi,dg}^3;$$

ricordiamo che  $\tilde{g}_\chi$  è il numero di stati di spin della particella, moltiplicato per 3/4 se essa è un fermione (ricordiamo anche che  $\zeta(3) \simeq 1.20$  è un valore della *funzione di Riemann*).

A  $T < T_{\chi,dg}$ , la densità in numero dei  $\chi$ , cala semplicemente col cubo del fattore di scala. Il loro numero oggi è dato allora da

$$(X.56) \quad n_\chi(T_o) = n_\chi(T_{\chi,dg}) (a_{\chi,dg}/a_o)^3$$

e il rapporto  $a_{\chi,dg}/a_o$  è circa pari a  $T_o/T_{\chi,dg}$ , per la conservazione dell'entropia. Più

precisamente, deve essere:

$$(X.57) \quad g_{\chi,dg} T_{\chi,dg}^3 a_{\chi,dg}^3 = g_{o,eff} T_o^3 a_o^3$$

dove  $g_{\chi,dg} = \mathcal{N}_{bos} + (7/8)\mathcal{N}_{fer}$  è ricavabile dal numero di stati di spin bosonici ( $\mathcal{N}_{bos}$ ) e fermionici ( $\mathcal{N}_{fer}$ ) nel brodo termico, escludendo le particelle  $\chi$ , a  $T_{\chi,dg}$ . L'entropia comovente (che è data da queste espressioni  $\times 4\pi^2/90$ ) a  $T_{\chi,dg}$  si ritrova a  $T_o$ : in parte nel CMB, in parte nell'analogo fondo neutrinico, possibilmente in altri fondi, costituiti da particelle che si fossero disaccoppiate dopo del  $\chi$  e prima del neutrino. Se usiamo l'espressione

$$(X.58) \quad g_{o,eff} = 2 \left( 1 + \frac{7}{22} N_\nu \right) \quad (\text{con } N_\nu: \text{ numero di famiglie neutriniche}),$$

assumiamo l'assenza di fondi oltre a quelli di  $\gamma$  e  $\nu$ , un'ipotesi probabilmente corretta.

Mettendo assieme le prescrizioni delle ultime 4 relazioni, otteniamo

$$(X.59) \quad n_\chi(T_o) = n_\gamma(T_o) \frac{\tilde{g}_\chi}{g_{\chi,dg}} \left( 1 + \frac{7}{22} N_\nu \right) \quad \text{dove} \quad n_\gamma(T_o) = \frac{\zeta(3)}{\pi^2} 2 T_o^3.$$

I neutralini, oggi, sarebbero comunque non relativistici e la loro densità in massa sarebbe

$$(X.60) \quad \rho_\chi = \frac{\tilde{g}_\chi}{\tilde{g}_{\chi,dg}} \frac{g_{o,eff}}{2} n_{o,\gamma} m_\chi \quad \text{ovvero} \quad \frac{\rho_\chi}{\text{GeV}/\text{cm}^3} \simeq \frac{\tilde{g}_\chi}{\tilde{g}_{\chi,dg}} g_{o,eff} \frac{200}{\text{cm}^{-3}} \frac{m_\chi}{\text{GeV}}$$

(si è preso  $n_{o,\gamma} = 400/\text{cm}^{-3}$ ) e il loro parametro di densità sarebbe dato da

$$(X.61) \quad \Omega_\chi h^2 \simeq 2 \times 10^7 \frac{\tilde{g}_\chi}{\tilde{g}_{\chi,dg}} g_{o,eff} \frac{m_\chi}{\text{GeV}}$$

sicché, con  $N_\nu = 3$ , anche prendendo  $\tilde{g}_{\chi,dg} \sim 100$ , per avere  $\Omega_\chi h^2 \sim 0.15$  occorre che sia  $m_\chi \tilde{g}_\chi \simeq 600 \text{ keV}$ . Con  $\tilde{g}_\chi = 1.5$  sarebbe  $m_\chi \sim m_e$ .

Siamo completamente al di fuori del range di masse per i partner supersimmetrici, che dovrebbero essere almeno confrontabili con la temperatura  $T_{SUSY}$  a cui avviene la rottura di SUSY.

Le espressioni trovate sono comunque rilevanti a definire la densità di fondi costituiti di particelle di altro tipo, oggi non relativistiche, disaccoppiatesi quando erano relativistiche, indipendentemente dalla loro rilevanza come componenti della DM.

### Disaccoppiamento freddo

La conclusione è che i neutralini debbono essersi disaccoppiati quando erano già non relativistici e la loro abbondanza in numero si era ridotta per il loro progressivo decadimento in particelle ordinarie più leggere. La permanenza di un fondo di  $\chi$  che dia la DM è allora legata a un opportuno *tuning* tra ritmo delle reazioni di decadimento e ritmo di espansione, che blocchi l'ulteriore decadimento prima che il residuo di  $\chi$  divenga insignificante.

Di fatto, i parametri della microfisica ricadono in un range che rende questo *tuning* non improbabile. Il numero di  $\chi$  non relativistici è dato dall'espressione di Boltzmann

$$(X.62) \quad n_\chi(T) = \mathcal{N}_\chi \left( \frac{m_\chi T}{2\pi} \right)^{3/2} e^{-m_\chi/T},$$

che si può riscrivere utilizzando l'espressione relativistica della densità in numero, che chiameremo ora  $n_{\chi,\text{rel}}$ , data in eq. (X.56), e un fattore di soppressione  $S$ , che dipende dal rapporto  $m_\chi/T_{\chi,dg}$  tra massa del  $\chi$  e temperatura di disaccoppiamento. Ripercorrendo i passaggi che portano a  $N_{\chi,\text{rel}}$ , avremo allora

$$(X.63) \quad n_\chi(T_0) = n_{\chi,\text{rel}} S\left(\frac{m_\chi}{T}\right) \quad \text{con} \quad S(u) = \sqrt{\frac{\pi}{8}} \frac{1}{\zeta(3)} u^{3/2} e^{-u}$$

sicché il loro parametro di densità sarebbe dato da

$$(X.64) \quad \Omega_\chi h^2 \simeq 4 \times 10^7 \frac{\tilde{g}_\chi}{g_{\chi,dg}} \frac{m_\chi}{\text{GeV}} S\left(\frac{m_\chi}{T_{\chi,dg}}\right)$$

e il limite  $m_\chi \tilde{g}_\chi = 600 \text{ keV}$  diviene

$$(X.65) \quad m_\chi \tilde{g}_\chi S\left(\frac{m_\chi}{T_{\chi,dg}}\right) \simeq 600 \text{ keV}.$$

Perché possa esistere un  $\chi$  con massa  $m_\chi \simeq 4 \text{ GeV}$  si richiede allora che sia  $S < 10^{-4}$  e, quindi,  $T_{\chi,dg} < 0.1-0.2 m_\chi$ . Se  $T_{\chi,dg} \ll 0.1-0.2 m_\chi$ , pertanto, il fondo di  $\chi$  esiste, ma è una componente trascurabile della sostanza cosmica. Se, invece,  $T_{\chi,dg}/m_\chi \simeq 0.1-0.2$ , i  $\chi$  costituiscono la DM delle osservazioni astrofisiche.

Si noti che la dipendenza dalla massa di  $T_{\chi,dg}/m_\chi$  è logaritmica e, quindi, l'ordine di grandezza del rapporto  $T_{\chi,dg}/m_\chi$ , perché i neutralini siano la DM, deve essere comunque 0.1-0.2.

Un calcolo più preciso della relazione tra  $m_\chi$  e  $\Omega_\chi$  è riportato in appendice A6. Esso permette di ricavare il valore atteso di  $\Omega_\chi h^2$  per ogni possibile alternativa sulle SUSY.

## 11. Transizione quark-adroni

Facciamo ora un passo indietro sulla scala termica. Sulla base di quanto discusso in paragrafo 5 è chiaro che la materia con interazioni forti si presenta in due forme diverse, a seconda della densità. Trattando gli adroni come particelle puntiformi, la loro densità in numero, a una determinata temperatura  $T$ , è

$$(X.66) \quad n_h = \frac{\zeta(3)}{\pi^2} g_h T^3,$$

in analogia con eq. (X.59) del paragrafo precedente. La distanza media tra adroni è quindi  $d_h \sim 2\pi\hbar c/k_B T$  ( $k_B$ : costante di Boltzmann). A temperature

$T \sim 2\pi\hbar c/k_B(10^{-13}\text{cm}) \simeq 150\text{ MeV}$ , quindi, la distanza tra quark appartenenti allo stesso adrone o a due adroni diversi diventa confrontabile (ricordiamo che il raggio dell'hard-core adronico è  $\sim 10^{-13}\text{ cm}$ ). Non c'è quindi più ragione perché i quark siano confinati negli adroni.

Questo calcolo è viziato dall'aver trascurato il volume proprio e le mutue interazioni degli adroni, proprio nelle circostanze in cui essi divengono significativi, ma dà comunque una stima delle condizioni in cui lo stato di aggregazione della materia con interazioni forti deve cambiare.

Esiste un altro contesto astrofisico in cui la materia con interazioni forti può avere alte densità, quello delle stelle di massa inferiore al limite di Chandrasekar (quindi non destinate a raggiungere lo stadio di *buchi neri*). Dopo aver raggiunto lo stadio di *stelle a neutroni*, l'evoluzione in ipotetiche *stelle di quark* non è impossibile. La differenza tra questo contesto e quello cosmologico sta nel fatto che qui l'elevata densità ( $\sim 10^{15}\text{g/cm}^3$ ) è conseguenza della conservazione del numero barionico e non della temperatura. Le considerazioni cosmologiche non possono quindi essere estese, *tout-court*, alle stelle di quark.

### Gas adronico e plasma quark-gluonico

Nel contesto cosmologico ci aspettiamo dunque sia avvenuta, a una temperatura  $T_{qh}$ , una transizione da plasma quark-gluonico a gas adronico.

In approssimazione zero, per i due stati di aggregazione, densità d'energia e pressione hanno espressione

$$(X.67) \quad \rho_{qg} = 3 p_{qg} = \frac{\pi^2}{30} g_{qg} T^4, \quad \rho_h = 3 p_h = \frac{\pi^2}{30} g_h T^4.$$

Per il plasma si deve tener conto del gran numero di particelle relativistiche, distinte tra loro da numeri quantici: ci saranno 2 stati di spin per ciascun quark dei sapori  $u, d, s$  (la cui massa è  $< T_{qh}$ ) con 3 possibili colori; i gluoni, che hanno massa nulla, hanno anch'essi 2 stati di spin. Sarà quindi:

$$(X.68) \quad g_{qg} = \underset{\substack{\text{spin} \\ \text{gluoni}}}{2} \times \underset{\substack{\text{colore} \\ \text{gluoni}}}{8} + \frac{7}{8} \underset{\substack{\text{spin} \\ \text{quark}}}{2} \times \underset{\substack{\text{colore} \\ \text{quark}}}{3} \times \underset{\substack{\text{sapore} \\ \text{quark}}}{3} = 32 - 1/4$$

I barioni, invece, hanno tutti massa  $\gg T_{qh}$ ; solo la massa del mesone  $\pi$  è  $\sim \mathcal{O}(T_{qh})$ ; il gas adronico apparirebbe quindi caratterizzato da  $g_h < 3$ .

La realtà del gas è tuttavia più complessa, in quanto il volume proprio della particelle non è trascurabile, mentre le mutue interazioni causano la presenza di *risonanze* che, in pratica, sono altre particelle adroniche dalla vita brevissima, continuamente rigenerate e riassorbite. Nell'assieme, possiamo immaginare che la stima  $g_h \sim 3$  possa anche essere in difetto.

Ancora più complessa è la situazione per il plasma, secondo quanto ci si aspetta sulla base della fisica del confinamento. Se i quark si allontanano eccessivamente, i

gluoni «ricordano» di essere colorati e possono generare essi stessi altri gluoni. La richiesta energetica necessaria a rifornire la cascata gluonica, teoricamente, potrebbe divenire persino superiore all'energia necessaria a distribuire nello spazio nuove coppie quark-antiquark, di modo che la distanza media tra quark non cresca. Una densità costante nel corso dell'espansione significa  $p = -\rho$ . Ma, anche se ciò non avviene, la pressione sarà sostanzialmente inferiore a eq. (X.67).

L'espansione del plasma, di fatto, non porterà alla creazione di coppie, ma alla creazione di multipletti di quark confinati. All'avvicinarsi delle condizioni per il confinamento, comunque, la pressione subirà una significativa riduzione, rispetto a un gas perfetto, dovuta a forze di coesione generate dal colore.

### Modello a bag e sue generalizzazioni

Queste correzioni possono essere espresse tramite il modello a bag e le sue generalizzazioni.

Le espressioni (X.67) vanno perciò modificate. Introducendo una correzione negativa sulla pressione, nella forma di legge di potenza, avremo:

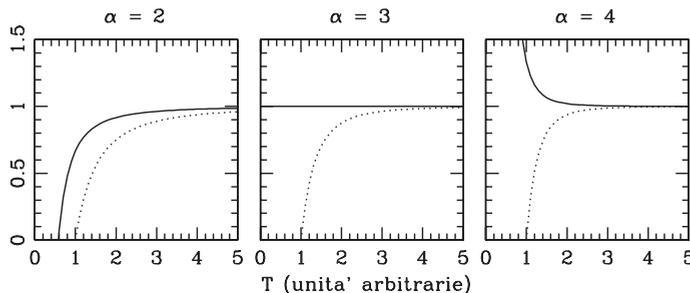
$$(X.69) \quad p_{qg}(T) = \frac{\pi^3}{90} g_{qg} \left[ 1 - \left( \frac{\tilde{T}}{T} \right)^\alpha \right] T^4.$$

Dall'andamento della pressione con la temperatura si può ricavare anche quello della densità d'energia:

$$(X.70) \quad \rho_{qg}(T) = \frac{\pi^3}{30} g_{qg} \left[ 1 + \frac{\alpha - 3}{3} \left( \frac{\tilde{T}}{T} \right)^\alpha \right] T^4$$

(vedi eq. (A2.5) nell'appendice A2 del capitolo IV).

In figura X.7 sono riportati gli andamenti di  $p_{qg}(T)$  e  $\rho_{qg}(T)$  per  $\alpha = 2, 3, 4$ . Per  $\alpha = 3$  c'è



**FIGURA X.7**

Andamento della pressione e della densità d'energia del plasma quark-gluonico in prossimità del passaggio a gas adronico. La curva continua rappresenta  $\rho_{qg}/(\pi/30)g_{qg}T^4$ ; la curva tratteggiata rappresenta  $\rho_{qg}/(\pi/90)g_{qg}T^4$ . Gli andamenti sono ottenuti dalle espressioni (X.69) e (X.70) per i valori di  $\alpha$  indicati sulle figure. Il caso  $\alpha = 4$  corrisponde all'MIT bag-model.

correzione sulla pressione, mentre l'andamento della densità d'energia è lo stesso che per un gas perfetto. Questo caso segna il discrimine tra i casi  $\alpha > 3$  o  $\alpha < 3$ , in cui la densità d'energia del plasma è, rispettivamente, maggiore o minore di quella di un gas libero.

L'andamento per  $\alpha = 4$  corrisponde al classico MIT bag-model. In questo modello, la densità d'entropia  $\sigma_{gg} = (\rho_{gg} + p_{gg})/T$  ha lo stesso andamento che in un gas perfetto: le due correzioni  $\propto (\bar{T}/T)^\alpha$  si cancellano.

### QCD su reticolo

I risultati delle recenti analisi su reticolo sono invece prossimi al caso  $\alpha = 2$ . I calcoli su reticolo sono basati su una tecnica *non perturbativa*, quindi valida indipendentemente dall'elevata intensità delle forze «forti», basata su espressioni di meccanica quantistica ricavate con tecniche di *path-integrals*.

Il reticolo discretizza le lunghezze d'onda e quindi gli impulsi. Nascono inoltre significative complicazioni nella trattazione dei fermioni dinamici. I calcoli di QCD su reticolo sono estremamente laboriosi e vengono realizzati tramite macchine calcolatrici parallele, con adeguata architettura. Tra di esse vanno ricordate APE e le macchine da essa derivate.

Tra i risultati dei calcoli su reticolo, vi sono espressioni di  $\rho$  e  $p$  in funzione della temperatura. La presenza di una discontinuità in  $\rho$  marcherebbe la presenza di una *transizione di fase* di primo ordine.

Va anche notato che vi sono almeno due fenomeni fisici associati alla transizione quark-adroni. Da un lato c'è il *confinamento* dei quark negli adroni; dall'altro vi è la *rottura della simmetria chirale* quando gli adroni acquisiscono una massa «forte». La transizione quark-adroni potrebbe quindi suddividersi in due stadi, con caratteristiche termodinamiche tra loro diverse.

Nell'impossibilità di diretta ispezione sperimentale, la comunità di fisici che persegue tecniche di QCD numerica spesso finisce per considerare i risultati del calcolatore alla stregua di dati di laboratorio. Un esempio che riguarda direttamente la cosmologia è la natura della transizione quark-adroni. La QCD su reticolo, in presenza di 3 sapori di quark, trattati come fermioni dinamici, giunge univocamente alla conclusione che non si tratta di una transizione di fase, ma di una *transizione continua*, sul tipo del passaggio da plasma a idrogeno atomico, quando il cosmo ha una temperatura di circa 3000 K.

Nonostante la convergenza dei risultati è però corretto che la cosmologia mantenga ampie riserve.

## 12. Transizione quark-adroni in cosmologia

Il modello a bag o una delle sue generalizzazioni, oppure la QCD su reticolo, forniscono espressioni fisiche per  $E(T) = \rho_{gg}/T^4$  e  $\Phi(T) = p_{gg}/T^4$ , con cui risolvere le equazioni di Friedmann e determinare i tempi e l'andamento del fattore di scala alla vigilia della transizione.

Una prima stima dell'età del cosmo al momento della transizione può essere comunque ottenuta estrapolando all'indietro la relazione

$$(X.71) \quad \frac{t}{\text{sec}} \left( \frac{T}{\text{MeV}} \right)^2 = 1.31 \sqrt{\frac{3.363}{g(T)}} \quad \text{con} \quad g(T) = \left( 31.75 + 2 + \frac{7}{8} 10 \right) = 42.5$$

(oltre che del plasma (eq. X.68), si deve tenere conto dei  $\gamma$  e dei 10 stati di spin per elettroni e neutrini) che, per  $T = 200 \text{ MeV}$  dà  $t \simeq 9 \times 10^{-6} \text{ sec}$ . La relazione è approssimata perché assume che sia  $\dot{a}/a = 1/2t$  (espansione radiativa).

Resta da chiarire la natura della transizione, se non si vuole passare in giudicato il risultato della QCD su reticolo. Del tutto in generale, le alternative sono:

- 1) transizione continua, come prevederebbe la QCD su reticolo;
- 2) transizione di fase del secondo ordine;
- 3) transizione di fase del primo ordine senza significativo *supercooling*;
- 4) transizione di fase del primo ordine con *supercooling*.

Nei casi 1)-3), in cui c'è conservazione dell'entropia comovente, vanno confrontate le sue espressioni a temperature  $T_1 \gg T_{qh}$  e  $T_2 \ll T_{qh}$ :

$$(X.72) \quad S_1 = \frac{4\pi^2}{90} 42.5 T_1^3 a_1^3, \quad S_2 = \frac{4\pi^2}{90} 9.75 T_2^3 a_2^3.$$

Tenendo conto del valore di  $g(T_1)$  (eq. X.71) e  $g(T_2)$ , valutato per  $\gamma$ , elettroni e 3 neutrini, risulta che  $T_2 a_2 \simeq 1.63 T_1 a_1$ .

In questo computo di  $g(T_2)$  non entra nessun adrone. Per temperature molto prossime a  $T_{qh}$  ciò potrebbe essere scorretto: anche se è  $m_\pi \sim T_{qh}$  e i  $\pi$  non sono relativistici, il contributo di risonanze adroniche potrebbe rendere non trascurabile in numero complessivo di stati di spin effettivi della componente adronica. D'altronde l'effetto probabilmente si è già esaurito se si considera un valore  $T_2 \sim 100 \text{ MeV}$ , quando i  $\pi$  stessi sono comunque già fuori dal *brodo termico*. In pratica, dunque, la transizione quark-adroni implica il passaggio dell'entropia delle particelle con interazioni forti su fotoni e leptoni.

Nelle diverse ipotesi 1)-4) si può impostare la soluzione delle equazioni di Friedmann e, in molti casi, si ottengono anche integrali analitici.

Sulla base di queste conclusioni fisiche ci dobbiamo allora porre due quesiti:

- 1) La transizione quark-adroni risolve qualche problema della cosmologia?
- 2) Ci sono dati cosmologici che permettono di aprire una finestra sulla dinamica fisica della transizione?

Di fatto, le risposte non sono molto confortanti. Il contesto più significativo, nel quale la transizione potrebbe avere impatto sulla cosmologia, riguarda il caso di una transizione di fase del primo ordine. Lo scenario proposto è denominato *nucleosintesi disomogenea*. Ne daremo qui una descrizione qualitativa.

### Nucleosintesi disomogenea

Vedremo che, come conseguenza della transizione quark-adroni, all'apertura del «collo di bottiglia» del deuterio, la distribuzione dei protoni potrebbe essere disomogenea. I neutroni, invece, sarebbero distribuiti omogeneamente. La ragione fisica è che il numero barionico passerebbe con difficoltà dal plasma al gas, a causa della massa dei barioni  $\gg T_{qh}$ , mentre la massa dei quark è  $\ll T_{qh}$ .

Seguiamo allora le fasi della transizione. Il cosmo, raffreddandosi, raggiunge la temperatura di transizione  $T_c$ . A quel punto nucleano bolle di gas all'interno della fase di plasma che, all'inizio, è la fase fisicamente connessa. Essendo  $\rho_h \ll \rho_{qq}$ , la discesa della temperatura cosmica si blocca e l'espansione avviene grazie all'aumento della frazione di volume occupata dalla fase di gas. La distanza iniziale delle bolle è fissata dalla dinamica della nucleazione. Crescendo di volume, tuttavia, le bolle entrano in mutuo contatto. I volumi occupati dal gas tendono comunque a riarrangiarsi in forma sferica, per l'azione della tensione superficiale, almeno sinché le forze di superficie dovute alla tensione superficiale sono sufficienti a smuovere l'inerzia dei materiali, che cresce col volume. Se la tensione superficiale è  $\sigma = \sigma_0 T^3$ , il valore del coefficiente  $\sigma_0$  fissa la scala caratteristica del tessuto di bolle che, alla fine, venute ampiamente in contatto, costituiranno la fase connessa.

L'ulteriore espansione cosmica è poi accompagnata dal calo del raggio delle residue bolle di plasma. Quando esso va a zero, la temperatura riprende la sua discesa e la transizione è finita. Un limite superiore alla distanza tra bolle di plasma, nelle fasi finali, è fissato dall'orizzonte  $\sim 2ct \simeq 10^5\text{-}10^6\text{ cm}$ .

Nella fase di plasma, prima della transizione, c'è un elevato numero di coppie  $q, \bar{q}$ . Il numero barionico è allora responsabile di 1 quark in più ogni  $\sim 10^9$  coppie. La maggior parte delle coppie si annichilano secondo le reazioni

$$q + \bar{q} \rightarrow \gamma + \gamma \qquad q + \bar{q} \rightarrow \text{leptone} + \text{antileptone},$$

che avvengono alla superficie interfase. I quark in eccesso, invece, dovrebbero dar luogo a barioni. In condizioni di equilibrio, la densità di numero barionico, nelle due fasi, sarebbe comunque diversa, a causa dell'elevata massa del barione più leggero, superiore di un fattore  $\sim 5$  alla temperatura  $T_c$ . Ma il problema non si risolve andando a considerare le condizioni di equilibrio. Ciò che rende veramente difficile il passaggio di  $B$  è che esso richiede l'incontro tra 3 quark. Si deve quindi supporre che  $B$  permanga nel plasma finché le bolle sono tanto piccole che, in esse, la densità di  $B$  è così elevata da causare reazioni a 3 corpi.

Alla conclusione della transizione, quindi, ci saranno picchi di  $B$  nei siti in cui «muoiono» le ultime bolle. Essendo  $T_c \sim 100\text{ MeV}$ , tale eccesso sarà inizialmente dovuto, in egual misura, a protoni ( $p$ ) e neutroni ( $n$ ). Diffondendo nel mezzo cosmico, i  $n$  porteranno con sé  $B$ . Alla diffusione si oppongono le forze elettromagnetiche che danno luogo a urti con  $\gamma$  e leptoni carichi. La sezione d'urto di  $n$  e  $b$ , in realtà, è confrontabile: un insieme di cariche  $2/3, 2/3, -1/3$  non dà una sezione d'urto molto diversa da un insieme di cariche  $2/3, -1/3, -1/3$ . Ciò che blocca il moto dei  $p$  è la necessità di portare con sé un

elettrone, al fine di preservare la neutralità di carica. Il volume occupato dai  $p$  aumenterà quindi sulla base della capacità degli elettroni a diffondere.

Un calcolo dettagliato implica che, dalla transizione quark-adroni all'inizio della nucleosintesi, i protoni si sono omogeneizzati solo se la distanza tra bolle, durante la transizione, è inferiore a  $\sim 10$  cm, una distanza  $\sim \mathcal{O}(10^{-5}ct)$ .

La BBN è stata studiata a partire da queste condizioni iniziali. I risultati sono interessanti perché si ottiene consistenza tra abbondanze previste e osservate, per i diversi nuclidi, per un unico valore di  $\Omega_b$  e questo, di per sé, è un notevole successo. Ci sono stati molti altri tentativi di modificare la BBN, ma essi hanno vanificato la consistenza tra abbondanze calcolate e osservate, tramite un unico parametro.

Inoltre, l'intervallo di  $\Omega_b$  consistente coi dati è maggiore di quello della BBN omogenea, per in fattore 1,5-1,8. Un tale incremento non permette certo di attribuire natura barionica a tutta la materia non relativistica. Esso è però abbastanza significativo da essere confrontabile coi dati osservativi, ottenendo così qualche limite sulle caratteristiche della transizione quark-adroni.

### 13. La fisica alla scala $T_{GUT}$

Al di sopra di  $T_{GUT}$  si ritiene che le forze elettromagnetiche, debole e forte siano unificate in un'unica forza di gauge. A  $T < T_{GUT}$  la simmetria delle interazioni è data dal prodotto delle algebre  $U(1) \otimes SU(2) \otimes SU(3)$ . La simmetria delle forze di gauge, a  $T > T_{GUT}$ , dovrà essere dettata da un'algebra che includa, come sottoinsieme, le matrici di  $U(1) \otimes SU(2) \otimes SU(3)$ . Essendo  $U(1)$  e  $SU(2)$  di rango 1 (1 matrice diagonale) e  $SU(3)$  di rango 2 (2 matrici diagonali), la simmetria unificata dovrà riferirsi a un'algebra almeno di rango 4.

Le possibili algebre sono state tutte classificate e l'unica algebra di rango 4 che soddisfa i requisiti fenomenologici è  $SU(5)$ . Più precisamente  $SU(5)$  permette di prevedere correttamente i numeri quantici, rispetto a  $U(1) \otimes SU(2) \otimes SU(3)$  dei 15 quark e leptoni (destrorsi e sinistrorsi). Inoltre, essa permette di prevedere con buona approssimazione, il valore di  $\theta_W$ .

Nell'ambito di  $SU(5)$ , quark e leptoni sono collocati in due multipletti, nelle rappresentazioni 5-dimensionale e 10-dimensionale (che si ottiene dalla parte anti-simmetrica di un prodotto diretto  $5 \otimes 5$ ):

$$\psi_5 = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ e^c \\ \nu^c \end{pmatrix}, \quad \psi_{10} = \begin{pmatrix} 0 & u_3^c & -u_2^c & -u_1^c & -d_1 \\ -u_3^c & 0 & u_1^c & -u_2^c & -d_2 \\ u_2^c & -u_1^c & 0 & -u_3 & -d_3 \\ u_1 & u_2 & u_3 & 0 & -e^+ \\ d_1 & d_2 & d_3 & e^+ & 0 \end{pmatrix}$$

$u, d$  sono i nomi dei sapori; gli indici 1, 2, 3 si riferiscono al colore. Ci sono poi 24 bosoni di gauge, tra cui 8 gluoni, 3  $W$ , 1  $Y$ . Gli altri 12 bosoni sono estranei a

$U(1) \otimes SU(2) \otimes SU(3)$  e forniscono accoppiamenti tra barioni e leptoni che violano la conservazione di numero barionico e leptonico. Tale violazione è causata da 6 bosoni vettori  $X$  e 6 bosoni  $Y$ .

A temperature  $T < T_{GUT}$  tali mesoni danno luogo a forze in qualche modo simili alle forze deboli, con una specie di «costante di Fermi»  $G \simeq g^2/M_{X,Y}^2$ , molto più piccola di  $G_F \simeq g^2/M_W^2$ , in quanto  $M_{X,Y} \sim T_{GUT}$ , mentre  $M_W \sim T_{EW}$ . Ciononostante, questa forza super-debole permette processi che violano la conservazione di  $B$ , altrimenti vietati. Il calcolo del tempo di vita dei protoni, rispetto a un decadimento in  $\pi^0$ ,  $e^+$ , è stato calcolato in  $4.5 \times 10^{29 \pm 2}$  anni. Pur trattandosi di un tempo  $\gg t_0$  (età attuale del cosmo), esso prevede un numero di decadimenti del protone che supera i limiti sperimentali.

Questa formulazione di  $SU(5)$ , detta  $SU(5)$  minimale, è perciò sicuramente errata. Si prospetta un superamento di questo problema sia nell'ambito di  $SU(5)$ , in formulazioni non-minimali, sia passando a altre algebre, come quella di  $SU(10)$ , per le quali rinviamo a trattati specializzati.

C'è un ulteriore problema, detto *problema di gerarchia*, legato al meccanismo di Higgs, che provocherebbe le rotture spontanee delle simmetrie EW e GUT. Il punto è che, partendo da  $SU(5)$ , ci sono due successivi meccanismi di rottura, che agiscono alle scale  $T_{GUT}$  e  $T_{EW}$ , il cui rapporto è  $\sim \mathcal{O}(10^{12})$  e le scale stesse sono immesse come parametri dall'esterno.

Ciò non è solo «antiestetico». Di fatto, quando si intraprende il percorso di regolarizzazione e rinormalizzazione (vedi paragrafo 7), il rapporto tra scale è immediatamente distrutto: la rinormalizzazione riconduce le due scale a valori analoghi.

Il rapporto tra le scale può essere preservato solo se tutto il modello è preventivamente *supersimmetrizzato*. In questo caso, essendo associati dei fermioni ai bosoni di Higgs, si riesce a sfruttare opportune proprietà dei fermioni per bloccare gli effetti della rinormalizzazione.

Va poi notata la vicinanza tra le scale di Planck e GUT. In un certo senso, si sarebbe potuto ritenere «ragionevole» che il gruppo di rinormalizzazione portasse la scala GUT a coincidere con quella di Planck, di modo da realizzare così l'unificazione delle 4 forze, compresa la gravità. La diversa sistemazione fenomenologica può essere segno di una fisica assai articolata, ma ahimè difficilmente esplorabile, tra la scala di Planck e la scala GUT.

Va anche aggiunto che il meccanismo di Higgs viene usato alle scale  $T_{GUT}$  e  $T_{EW}$  in due contesti diversi. A  $T_{EW}$  non c'è una vera unificazione delle forze, ma solo una sistemazione delle forze legate a  $U(1)$  e  $SU(2)$ : le forze deboli e elettromagnetiche, a  $T < T_{EW}$ , sono sprovviste dei connotati di alta simmetria delle forze portate dai campi  $Y$  e  $W$ , quando non avevano massa. L'aumento della complessità e la perdita di simmetria, al calare della temperatura, è un fatto che si riscontra in molte transizioni di fase. Tuttavia, a  $T > T_{EW}$  ci sono 2 costanti di interazione ( $g_1$  e  $g_2$ ). A  $T < T_{EW}$  le stesse costanti continuano a evolvere con  $T$ , seppure nelle combinazioni istituite dall'angolo  $\theta_W$ .

Ciò che accade a  $T_{GUT}$  è molto diverso. Anche qui la discesa verso le basse temperature importa perdita di simmetria. Ma a  $T_{GUT}$  si ha veramente il passaggio da un'unica costante a 3 costanti; viceversa, procedendo verso valori alti di  $T$ , l'unificazione delle forze è un dato reale.

L'uso del meccanismo di Higgs, alla scala  $T_{EW}$  potrà presto essere giustificato quando si scoprirà il quanto del campo di Higgs. Non è però da scartare l'idea che la fisica alla scala  $T_{GUT}$  sia assai diversa dalla fisica legata a un bosone di Higgs e che, anzi, sia proprio questa fisica a innescare il meccanismo di Higgs a una scala  $10^{12}$  volte minore.

Si tratta di speculazioni legittime in attesa della problematica esplorazione dell'intervallo di temperature  $10^{15}$ - $10^{19}$  GeV. Esplorare di queste scale potrebbe diventare possibile, sia pure in modo indiretto, sulla base del paradigma inflazionario, di cui parleremo al prossimo capitolo.

#### 14. La $\mathcal{CP}$ forte e l'assione

Tra i molti problemi aperti dal programma di rinormalizzazione va citato quello che nasce dall'osservazione che le forze forti rispettano la simmetria  $\mathcal{CP}$ . Ne tratteremo qui in modo succinto perché la principale soluzione proposta a questo problema ha condotto all'introduzione dell'*assione*, che è una particella candidata a rendere conto della DM.

La formulazione del problema è molto tecnica e il suo approfondimento va oltre l'ambito di questo libro. In effetti, gli assioni sarebbero di per sé particelle molto leggere, con masse  $\sim 10^{-5}$  eV, e apparirebbero come candidati improbabili per una DM non-relativistica. La questione è che la loro produzione avviene in modo non-termico e che essi non vengono mai termalizzati. La loro distribuzione è coerente e viola quindi i limiti della statistica di Fermi.

Il problema che porta all'introduzione dell'assione sta nell'esistenza di stati multipli di vuoto,  $|0_n\rangle$ , in QCD. Infatti l'insieme delle trasformazioni di gauge del gruppo SU(3) di colore può essere suddiviso in sottinsiemi  $\Omega_n(x^\mu)$ , il cui comportamento asintotico dipende dall'indice  $n$ . Nell'ambito di  $\Omega_n$  si può passare dall'una all'altra trasformazione in modo continuo; non si passa invece in modo continuo da una trasformazione di  $\Omega_n$  a una di  $\Omega_m$  se  $n \neq m$ . Ciascun vuoto  $|0_n\rangle$  pertiene l' $n$ -esimo sottinsieme delle possibili gauge.

Non c'è dunque nessuna comunicazione tra i diversi  $n$ -settori di gauge, a livello classico. La comunicazione diviene invece possibile a livello quantistico grazie a effetti *istantonici*. Gli stati fisici di vuoto, dal punto di vista quantistico, saranno una sovrapposizione degli  $|0_n\rangle$ ,

$$(X.73) \quad |0_\theta\rangle = \sum_n |0_n\rangle e^{in\theta},$$

con coefficienti univocamente determinati da una fase  $\theta$ . Ogni variazione di  $\theta$  può però essere riletta come variazione di un termine aggiuntivo alla Lagrangiana di Yang e

Mills per SU(3) di colore, «creato» dalla rinormalizzazione, e della forma

$$(X.74) \quad \mathcal{L}_\theta = \frac{g_3^2}{16\pi^2} \theta \varepsilon^{\mu\nu\rho\sigma} f_J^{\mu\nu} f_J^{\rho\sigma}$$

( $\varepsilon^{\mu\nu\rho\sigma}$  è il tensore spazio-temporale completamente antisimmetrico; ricordiamo, ad esempio, che nel caso elettromagnetico lo si usa per definire l'invariante  $\varepsilon^{\mu\nu\rho\sigma} F^{\mu\nu} F^{\rho\sigma} = 2\mathbf{E} \cdot \mathbf{B}$ ). Come l'invariante  $\mathbf{E} \cdot \mathbf{B}$ , anche questo termine è pseudo-scalare; esso cambia segno per parità ed è quindi un'autofunzione di  $\mathcal{P}$  di autovalore  $-1$ . Siccome  $\mathcal{C}$  è una simmetria delle interazioni forti e gli altri termini della lagrangiana della QCD sono invarianti per  $\mathcal{P}$ , ne segue che  $\mathcal{CP}$  appare violata.

Viceversa, i limiti sul momento magnetico del neutrone,

$$(X.75) \quad d_n \simeq 10^{-20} \text{ e cm},$$

impongono che la violazione di  $\mathcal{CP}$  sia estremamente piccola, in quanto  $d_n \simeq 5 \times 10^{-16} \theta$  e ciò che corrisponde al limite

$$(X.76) \quad \theta \simeq 10^{-10}.$$

Di per sè, comunque, questa limitazione sarebbe un accettabile vincolo alla rinormalizzazione, se non fosse che un ulteriore contributo allo stesso termine viene anche alla scala elettrodebole, a seguito della *rottura della simmetria chirale* generata dai termini di massa. Quando la matrice di massa del quark,  $\mathcal{M}$ , viene allora diagonalizzata, il parametro  $\theta$  riceve un altro contributo, che risulta di natura elettrodebole, e diviene

$$\theta_{\text{eff}} = \theta + \text{Arg Det } \mathcal{M}.$$

Diventa così problematico accettare che ci sia una «congiura» tra forze forti e deboli, capace di mantenere comunque  $\theta$  entro i limiti (X.76).

Peccei & Quinn hanno proposto una via d'uscita a questo problema, introducendo una simmetria globale rispetto a un ulteriore gruppo U(1) che si rompe spontaneamente a una scala  $V_{PQ}$ . L'*assione* è il bosone di Goldstone che nasce dalla rottura della simmetria e che sarà opportunamente accoppiato all'insieme di quark, con dettagli dipendenti dalla formulazione specifica prescelta. La simmetria U(1) soffre di anomalia chirale e quindi gli assioni acquisiscono una (piccolissima) massa, per effetti non perturbativi. Essa cresce al momento della transizioni quark-adroni, a causa dell'incremento nella rottura della simmetria chirale.

L'anomalia si manifesta quando si effettua una trasformazione chirale sul campo assionico e dà luogo a una lagrangiana esattamente della forma (X.74), che è quindi un termine potenziale per il campo assionico  $\theta$ .

Tra i limiti alla scala di Peccei-Quinn c'è quello dettato dalla densità cosmica di assioni. Viceversa, per

$$(X.77) \quad 10^{10} \text{ GeV} < V_{PQ} < 10^{12} \text{ GeV}$$

è possibile ottenere una componente di assioni che si comporti efficacemente da DM, con opportuni valori di  $\Omega_c$ .

# Appendice A1

## Il teorema di Noether

**Invarianze e conservazioni per il punto materiale.** La dinamica di un punto materiale di massa  $m$  soggetto all'azione di forze derivanti da un potenziale  $V(x^i)$  può essere ricavata dall'integrale variazionale

$$(A1.1) \quad S = \int_{t_1}^{t_2} dt L[x^i, \dot{x}^i], \quad \text{con} \quad L = \frac{m}{2} \sum_{i=1}^3 (\dot{x}^i)^2 - V(x^i).$$

Consideriamo una trasformazione a tempo costante

$$(A1.2) \quad x^i \rightarrow x^i + \delta x^i$$

e la conseguente variazione dell'integrale variazionale:

$$(A1.3) \quad \delta S = \int_{t_1}^{t_2} dt \left( \frac{\partial L}{\partial x^i} \delta x^i + \frac{\partial L}{\partial \dot{x}^i} \frac{d}{dt} \delta x^i \right) = \int_{t_1}^{t_2} dt \left( \frac{\partial L}{\partial x^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}^i} \right) \delta x^i + \int_{t_1}^{t_2} dt \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}^i} \delta x^i \right) = \\ = \int_{t_1}^{t_2} dt \left( \frac{\partial L}{\partial x^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}^i} \right) \delta x^i + \left. \frac{\partial L}{\partial \dot{x}^i} \delta x^i \right|_{t_1}^{t_2}.$$

Essa verrà ora considerata in due casi:

- 1)  $\delta x^i$  arbitrarie, ma nulle agli estremi  $t_1$  e  $t_2$ .
- 2)  $\delta x^i$  opportunamente definite, ma non necessariamente nulle in  $t_1$  e  $t_2$ .

Per il principio di d'Alambert, se si parte dalle  $x_i(t)$  fisiche, il funzionale  $S$  sarà invariante rispetto a una trasformazione infinitesima (A1.2) di tipo 1). La richiesta  $\delta S = 0$ , applicata all'ultima riga dell'eq. (A1.3), è automaticamente soddisfatta per il secondo addendo ( $\delta x_i(t_1, t_2) = 0$ ). Perché si annulli il primo addendo, vista l'arbitrarietà delle  $\delta x^i$  e dell'intervallo di integrazione, deve essere

$$(A1.4) \quad \frac{\partial L}{\partial x^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}^i} = 0.$$

Applicata al lagrangiano (A1.1), questa equazione dà l'eq. del moto del punto:

$$(A1.5) \quad \dot{\mathbf{p}} = -\nabla V, \quad \text{con} \quad p^i = \frac{\partial L}{\partial \dot{x}^i} = m\dot{x}^i.$$

Supponiamo invece che  $V(x^i)$  sia invariante per traslazioni lungo l'asse  $x^3$ :

$$(A1.6) \quad \delta x^i = \varepsilon \delta^{i3}.$$

Anche  $L$  e  $S$  sono allora invarianti. D'altro canto, valendo l'eq. del moto, sarà

$$(A1.7) \quad \delta S = p^i \delta^{i3} \varepsilon \Big|_{t_1}^{t_2} = 0 \quad \text{e, quindi,} \quad p^3(t_2) = p^3(t_1);$$

vista l'arbitrarietà degli estremi, ciò significa la conservazione di  $p^3$ . L'eq. (A1.3), che si usa per ricavare le equazioni di Lagrange dal principio di d'Alambert, prova quindi anche che, se c'è invarianza per traslazioni lungo una coordinata, l'impulso in quella direzione è costante.

Se invece c'è invarianza per rotazioni attorno a un asse, il momento angolare rispetto all'asse è costante, come ora verificheremo.

Una generica rotazione infinitesima dà luogo alla trasformazione

$$(A1.8) \quad x'^i = x^i + \delta x^i, \quad \text{con} \quad \delta x^i = \varepsilon^{ij} x^j, \quad \text{essendo} \quad \varepsilon^{ij} = -\varepsilon^{ji}$$

e quindi, come in (A1.7), sarà

$$\delta S = p^i \varepsilon^{ij} x^j \Big|_{t_1}^{t_2} = 0, \quad \text{per cui} \quad \ell^{ij} \varepsilon^{ij} = \text{const.}, \quad \text{con} \quad \ell^{ij} = x^j p^i - x^i p^j$$

Vista l'arbitrarietà delle  $\varepsilon^{ij}$ , ciò richiede che il vettore momento angolare  $\mathbf{L}$ , di componenti  $L^i = \ell^{ik}$  ( $i, j, k$  circolarmente da 1 a 3), sia costante.

Più in generale, in un sistema discreto, l'invarianza per traslazioni lungo una coordinata lagrangiana qualunque (quindi non necessariamente coordinate cartesiane) significa che il momento coniugato a quella coordinata è costante.

## 1. Invarianze e conservazioni nei sistemi continui

Noether generalizza questa relazione al continuo, quando risulta

$$(A1.9) \quad S = \int_{\Omega_4} d^4x \mathcal{L}(\phi_A, \partial^\mu \phi_A)$$

e le forze sono descritte da una (densità di) potenziale che fa parte di  $\mathcal{L}$ .

Le  $\phi_A(x^\mu)$  sono le funzioni di campo e l'analogo della trasformazione (A1.2) deve tener conto della loro dipendenza dalle coordinate spazio-temporali che esprime sia la dinamica del sistema che i suoi  $\infty$  gradi di libertà.

La generica trasformazione infinitesima è allora della forma

$$(A1.10) \quad x'^\mu \rightarrow x^\mu = x^\mu + \delta x^\mu, \quad \phi_A(x'^\mu) \rightarrow \phi'_A(x'^\mu) = \phi_A(x^\mu) + \delta \phi_A$$

dove  $\delta \phi_A$  è la differenza tra i valori di  $\phi_A$ , dopo e prima della trasformazione, nello stesso punto fisico che, nel frattempo, ha cambiato «nome».

È utile poi definire la variazione infinitesima

$$(A1.11) \quad \delta_o \phi_A = \phi'_A(x'^\mu) - \phi_A(x'^\mu) = \phi'_A(x'^\mu) - \phi'_A(x'^\mu) + \phi'_A(x'^\mu) - \phi_A(x'^\mu) = -\partial^\nu \phi_A \delta x^\nu + \delta \phi_A$$

tra valori di  $\phi_A$  nei punti che hanno lo stesso «nome».

Si dimostra che, in generale, le trasformazioni (A1.8) causano la variazione

$$(A1.12) \quad \delta S = \int_{\Omega_4} d^4x \left( \frac{\partial \mathcal{L}}{\partial \phi_A} - \partial^\mu \frac{\partial \mathcal{L}}{\partial \partial^\mu \phi_A} \right) \delta_o \phi_A + \int_{\Omega_4} d^4x \partial^\mu \left( \frac{\partial \mathcal{L}}{\partial \partial^\mu \phi_A} \delta_o \phi_A + \mathcal{L} \delta x^\mu \right)$$

dell'integrale variazionale. La dimostrazione dettagliata è riportata nei testi di teoria dei campi.

Come per il caso discreto, questa variazione va considerata in due casi:

- 1)  $\delta x^\mu$  e  $\delta \phi_A$  arbitrarie ma nulle al contorno  $\Sigma_3$  di  $\Omega_4$ .
- 2)  $\delta x^\mu$  e  $\delta \phi_A$  opportunamente definite, ma non necessariamente nulle sull'ipersuperficie (3-volume)  $\Sigma_3$ .

Nel caso 1) abbiamo una variazione sincrona che, se si parte dalle  $\phi_A$  fisiche, non deve modificare  $S$ . Nell'eq. (A1.12), il secondo integrale a secondo membro, per il teorema della divergenza, diviene un integrale sull'ipersuperficie  $\Sigma_3$  di  $\Omega_4$ , su cui è  $\delta_o \phi_A = \delta x^\mu = 0$ . Perché  $S$  sia invariante, per  $\Omega_4$  e  $\delta \phi_A$  generici, occorre e basta allora che sia

$$(A1.11) \quad \partial^\mu \frac{\partial \mathcal{L}}{\partial \partial^\mu \phi_A} - \frac{\partial \mathcal{L}}{\partial \phi_A} = 0,$$

cioè che valga l'equazione di Euler-Lagrange.

Nel caso 2), invece, annullandosi il primo integrale a secondo membro di eq. (A1.12), grazie a Euler-Lagrange, deve valere l'equazione

$$(A1.13) \quad \partial^\mu J^\mu = 0 \quad \text{con} \quad J^\mu = - \frac{\partial \mathcal{L}}{\partial \partial^\mu \phi_A} \delta_o \phi_A - \mathcal{L} \delta x^\mu,$$

in cui il segno è stato scelto in funzione di futuri passaggi. Resta da stabilire, caso per caso, per quali  $\delta_o \phi_A$  e  $\delta x^\mu$  vi sia invarianza; se, ad esempio, essi sono definiti dai valori di una  $n$ -upla di parametri  $\varepsilon_q$ ,

$$(A1.14) \quad \delta_o \phi_A = \Phi_{Aq} \varepsilon_q, \quad \delta x^\mu = \Psi_q^\mu \varepsilon_q,$$

all'eq. (A1.13) seguono  $n$  equazioni di continuità

$$(A1.15) \quad \partial^\mu j_q^\mu = 0 \quad \text{con} \quad j_q^\mu = - \frac{\partial \mathcal{L}}{\partial \partial^\mu \phi_A} \Phi_{Aq} - \mathcal{L} \Psi_q^\mu \quad (q = 1, \dots, n)$$

e, in apparenza,  $n$  quantità conservate

$$(A1.16) \quad Q_q = \int_V d^3x j_q^0$$

che però, talora, non sono simultaneamente «osservabili».

Vedremo ora le espressioni delle quantità conservate in alcuni casi specifici: 1) Invarianza per traslazioni spazio-temporali. 2) Invarianza per trasformazioni di Lorentz, ovvero rotazioni spaziali e spazio-temporali. 3) Invarianza per trasformazioni di fase nel caso di campi complessi. 4) Invarianza per trasformazioni appartenenti a gruppi definiti da algebre di Lie. Le espressioni, di norma, si semplificano quando si passa dal caso generico a un caso specifico.

1) Invarianza per traslazioni e tensore impulso-energia.

Assumiamo ci sia invarianza per trasformazioni con

$$(A1.17) \quad \delta x^\mu = x'^\mu - x^\mu = \varepsilon^\mu, \quad \delta \phi_A = \phi'_A(x'^\mu) - \phi_A(x^\mu) = 0$$

e, quindi, vista eq. (A1.11),

$$(A1.18) \quad \delta_o \phi_A = -\partial^\nu \phi_A \varepsilon^\nu.$$

Si tratta delle trasformazioni causate da un semplice mutamento dell'origine degli assi spazio-temporali. Conformemente a eq. (A1.13), allora, sarà

$$(A1.19) \quad J^\mu = T^{\mu\nu} \varepsilon^\nu \quad \text{con} \quad T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial \partial^\mu \phi_A} \partial^\nu \phi_A - \mathcal{L} g^{\mu\nu}$$

e varranno le equazioni di continuità

$$(A1.20) \quad \partial^\mu T^{\mu\nu} = 0.$$

Le quantità conservate legate alle traslazioni spazio-temporali sono impulso e energia. Nel volume  $dV$ , centrato in  $x^i$ , al tempo  $x^0$ , troviamo dunque energia e (componenti di) impulso (infinitesimi)

$$(A1.21) \quad dE = T^{00}(x^\mu) dV, \quad dP^i = T^{0i}(x^\mu) dV,$$

dovuti ai campi  $\phi_A$ . Gli elementi della diagonale principale di  $T^{\mu\nu}$  danno altresì

$$(A1.22) \quad T^{00} = \rho, \quad T^{ii} = p$$

(densità d'energia e pressione). Per un campo scalare, con lagrangiana

$$(A1.23) \quad \mathcal{L} = \frac{1}{2} \partial^\mu \phi \partial^\mu \phi - V(\phi),$$

avremo allora

$$(A1.24) \quad T^{00} = \rho = \frac{\partial \mathcal{L}}{\partial \partial^0 \phi} \partial^0 \phi - \mathcal{L} = \partial^0 \phi \partial^0 \phi - \frac{1}{2} \partial^\nu \phi \partial^\nu \phi + \frac{1}{2} \partial^i \phi \partial^i \phi + V(\phi) = \\ = \frac{1}{2} \dot{\phi}^2 + \frac{1}{2} \nabla \phi \cdot \nabla \phi + V(\phi)$$

e, al limite di omogeneità,

$$(A1.25) \quad \rho = \frac{1}{2} \dot{\phi}^2 + V(\phi).$$

Sempre al limite di omogeneità abbiamo allora anche

$$(A1.26) \quad p = \mathcal{L}[\phi(t)] = \frac{1}{2} \dot{\phi}^2 - V(\phi).$$

Queste sono le espressioni di densità d'energia e pressione dei campi scalari che intervengono nell'inflazione e nello studio dell'energia oscura.

2) Invarianza per rotazioni, momento angolare e spin.

Assumiamo ci sia invarianza per rotazioni degli assi. Le rotazioni danno luogo alle variazioni

$$(A1.27) \quad \delta x^\mu = x'^\mu - x^\mu = \varepsilon^{\mu\nu} x^\nu, \quad (\text{con } \varepsilon^{\mu\nu} = -\varepsilon^{\nu\mu}) \\ \delta_o \phi_A = \phi'_A(x^\mu) - \phi_A(x^\mu) = M_{AB}^{\mu\nu} \varepsilon^{\mu\nu} \phi_B - \partial_\mu \phi_A \delta x^\mu$$

in cui il primo addendo, contenente la matrice  $M_{AB}^{\mu\nu}$ , è  $\delta \phi_A$ . In quanto segue terremo presente l'esistenza di questo termine, ma daremo risultati espliciti per un campo scalare  $\phi(x^\mu)$ , per il quale si ha  $\delta \phi = 0$  e

$$(A1.28) \quad \delta_o \phi = -\partial^\mu \phi x^\nu \varepsilon^{\mu\nu}$$

Conformemente a eq. (A1.13) e sfruttando l'antisimmetria dei parametri  $\varepsilon^{\mu\nu}$ , si ha allora

$$(A1.29) \quad \partial^\rho M^{\rho\mu\nu} = 0 \quad \text{con} \quad M^{\rho\mu\nu} = x^\mu T^{\rho\nu} - x^\nu T^{\rho\mu}.$$

Vista l'antisimmetria di  $M^{\rho\mu\nu}$  per lo scambio  $\mu \leftrightarrow \nu$ , si tratta di 6 teoremi di conservazione.

Consideriamo i 3 teoremi che si hanno quando  $\mu, \nu = 1, 2, 3$ . Le quantità fisicamente conservate hanno densità

$$\ell^{ij} = x^i T^{0j} - x^j T^{0i}$$

sicché, tenendo conto di eq. (A1.21), in  $dV$  troviamo quantità «conservate»

$$(A1.30) \quad dL^{ij} = x^i dP^j - x^j dP^i,$$

che sono manifestamente legate al momento angolare  $\mathbf{L}$ , la cui componente  $i$ -esima sarà data da  $L^{jk}$ , con  $i, j, k$  presi circolarmente da 1 a 3.

Nel caso di campi descritti da più componenti spazio-temporali  $\phi_A$ , al momento angolare (A1.30) va aggiunto un termine ulteriore, dovuto al modo in cui i campi trasformano rispetto a trasformazioni di Lorentz. Una volta quantizzati i campi, questo termine esprime la somma degli spin dei quanti del campo.

### 3) Trasformazioni di fase abeliane.

Si considerino qui le trasformazioni

$$(A1.33) \quad \begin{aligned} \psi &\rightarrow \psi' = e^{ie\alpha}\psi \simeq \psi + \delta\psi & \text{con} & \quad \delta\psi = ie\alpha\psi \\ \bar{\psi} &\rightarrow \bar{\psi}' = \bar{\psi} + \delta\bar{\psi} & \text{con} & \quad \delta\bar{\psi} = -ie\alpha\bar{\psi} \end{aligned}$$

in cui il campo  $\psi$  è un campo spinore con lagrangiana

$$(A1.34) \quad \mathcal{L} = \bar{\psi}\gamma^\mu\partial^\mu\psi - m\bar{\psi}\psi$$

(vedi eq. A2.10). È allora facile vedere, essendo nulle le  $\delta x^\mu$ , che risulta

$$(A1.35) \quad J^\mu = ie\bar{\psi}\gamma^\mu\psi\alpha$$

per cui vale l'equazione di continuità per la corrente elettrica

$$(A1.36) \quad \partial^\mu j^\mu = 0 \quad \text{essendo} \quad J^\mu = j^\mu\alpha$$

### 4) Trasformazioni di gauge non abeliana.

Si considerino ora le trasformazioni descritte in appendice A3, per un campo spinoriale e il campo vettoriale di gauge. L'argomento che ha portato alle teorie di gauge è la necessità che le  $\alpha_J$  dipendano dalle coordinate; tuttavia, tra le possibili scelte, quella di  $\alpha_J = \text{const.}$  non è vietata ed è questa che considereremo, per la sua semplicità. Le correnti così trovate sono le stesse che si hanno nel caso generale.

Se le matrici dell'algebra di Lie che dà la teoria di gauge sono le  $\lambda_J$  e il campo spinoriale appartiene alla rappresentazione fondamentale dell'algebra, sarà

$$(A1.37) \quad \delta\psi = ig\alpha_J\lambda_J\psi, \quad \delta\bar{\psi} = -ig\bar{\psi}\alpha_J\lambda_J, \quad \delta a_J^\mu = \epsilon_{RSJ}a_S^\mu\alpha_R.$$

Si avrà allora

$$J^\mu = \bar{\psi}\gamma^\mu ig\alpha_J\lambda_J\psi + \partial^\mu a_J^\nu g\epsilon_{RSJ}a_S^\nu\alpha_R$$

e, riordinando,

$$(A1.38) \quad J^\mu = j_J^\mu\alpha_J \quad \text{con} \quad j_J^\mu = ig\bar{\psi}\gamma^\mu\lambda_J\psi - g\partial^\mu a_R^\nu\epsilon_{RSJ}a_S^\nu,$$

confermandosi così la presenza di una carica portata dai campi di gauge.

## Appendice A2

### La rappresentazione spinoriale

La rappresentazione spinoriale del gruppo di Lorentz è stata scoperta da P.A.M. Dirac, al fine di generalizzare l'equazione di Schrödinger al caso relativistico. Il suo principale utilizzo si ha in teoria quantistica dei campi: i fermioni di spin=1/2 (elettroni, neutrini, barioni, quark) sono quanti di campi spinori.

Lo spinore di Dirac è un oggetto a 4 componenti  $\psi_a$  che, in assenza di interazioni, soddisfa all'equazione

$$(A2.1) \quad (i\gamma^\mu \partial^\mu - m)\psi = 0.$$

Alle matrici  $\gamma^\mu$  può essere data la forma

$$(A2.2) \quad \gamma^0 = \begin{vmatrix} 0 & \mathbf{1} \\ \mathbf{1} & 0 \end{vmatrix}, \quad \gamma^1 = \begin{vmatrix} 0 & \sigma^1 \\ -\sigma^1 & 0 \end{vmatrix}, \quad \gamma^2 = \begin{vmatrix} 0 & \sigma^2 \\ -\sigma^2 & 0 \end{vmatrix}, \quad \gamma^3 = \begin{vmatrix} 0 & \sigma^3 \\ -\sigma^3 & 0 \end{vmatrix};$$

$\mathbf{1}$  e  $\sigma^i$  (matrici di Pauli, vedi eq. X.23) sono a loro volta matrici  $2 \otimes 2$ . Si può poi facilmente verificare, per ispezione diretta, la regola di anticommutazione

$$(A2.3) \quad \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu \equiv \{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$$

( $g^{\mu\nu}$  diagonale con  $g^{\mu\mu} = +1, -1, -1, -1$ , per  $\mu = 0, \dots, 3$ )

tenendo conto delle proprietà delle matrici di Pauli:

$$(A2.4) \quad \sigma^i \sigma^j = i\epsilon^{ijk} \sigma^k \quad (i, j, k : \text{da } 1 \text{ a } 3).$$

Si definisce poi la matrice  $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$  che risulta diagonale, con componenti

$$(A2.5) \quad \gamma^5 = \begin{vmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{vmatrix},$$

per cui la rappresentazione delle matrici  $\gamma^\mu$ , che stiamo considerando, è detta «rappresentazione  $\gamma^5$ ».

Notiamo che è  $(i\gamma^\mu \partial^\mu + m)\psi = 2m\psi$ , per cui sarà  $\{(i\gamma^i \partial^i - m), (i\gamma^i \partial^i + m)\}\psi = 0$ , mentre, usando eq. (A2.3), si vede che

$$\{(i\gamma^\mu \partial^\mu - m), (i\gamma^\nu \partial^\nu + m)\} = 2(\partial^\mu \partial^\mu - m^2).$$

In definitiva, ciascuna componente di  $\psi$  soddisfa l'equazione di Klein Gordon:

$$(A2.6) \quad [(\partial^2 / \partial t^2) - \nabla^2 - m^2]\psi_a = 0$$

Si noti anche che:

$$(A2.7) \quad \gamma^{0+} = \gamma^0, \quad \gamma^{i+} = -\gamma^i \quad (i = 1, 2, 3)$$

( $\gamma^0$  hermitiana,  $\gamma^i$  antihermitiane). Definito allora  $\bar{\psi} \equiv \psi^\dagger \gamma^0$ , sarà

$$(A2.8) \quad i\partial^\mu \bar{\psi} \gamma^\mu + m\bar{\psi} = 0,$$

in quanto è nulla l'hermitiana coniugata di questa equazione:

$$(A2.9) \quad \begin{aligned} (i\partial^\mu \bar{\psi} \gamma^\mu + m\bar{\psi})^+ &= [(i\partial^0 \psi^+ \gamma^0 + i\partial^i \psi^+ \gamma^i + m\psi^+) \gamma^0]^+ = \\ &= \gamma^0 (-i\gamma^0 \partial^0 \psi + i\gamma^i \partial^i \psi + m\psi^+) = -\gamma^0 (\gamma^\mu \partial^\mu - m)\psi = 0. \end{aligned}$$

Entrambe le eq. (A2.1) e (A2.8) si possono ricavare dalla densità lagrangiana

$$(A2.10) \quad \mathcal{L} = \bar{\psi}(i\gamma^\mu \partial^\mu - m)\psi$$

tramite le equazioni di Euler-Lagrange.

Se suddividiamo lo spinore a 4 componenti in 2 blocchi a 2 componenti,

$$(A2.11) \quad \psi = \begin{vmatrix} \psi_L \\ \psi_R \end{vmatrix},$$

le eq. (A1.1) si possono scrivere nella forma

$$(A2.12) \quad \partial^0 \psi_R + \sigma \cdot \nabla \psi_R = -im\psi_L, \quad \partial^0 \psi_L - \sigma \cdot \nabla \psi_L = -im\psi_R$$

siché, per  $m = 0$ ,  $\psi_L$  e  $\psi_R$  soddisfano equazioni separate. Una generica trasformazione di Lorentz, agendo su  $\psi_L$  o  $\psi_R$ , li trasforma in modi diversi:

$$(A2.13) \quad \begin{aligned} \psi_{La} \rightarrow \psi'_{La} &= A_{Lab} \psi_{Lb}, \quad \psi_{Ra} \rightarrow \psi'_{Ra} = A_{Rab} \psi_{Rb} \\ \text{con} \quad A_L &= e^{\frac{i}{2}\sigma_i(\lambda_i - i\nu_i)}, \quad A_R = e^{\frac{i}{2}\sigma_i(\lambda_i + i\nu_i)} \end{aligned}$$

dove  $\lambda_i$  e  $\nu_i$  sono, rispettivamente, gli «angoli» delle rotazioni spaziali attorno ai 3 assi, e delle rotazioni spazio-temporali tra ciascun asse e l'asse temporale (boost). Ciò mette in evidenza il significato degli indici  $L,R$  (Left: sinistrorso; Right: destrorso) che si riferiscono al verso dei boost. Espressioni più immediatamente operative per i boost saranno date alla fine dell'appendice.

Va considerata poi l'operazione di parità  $\mathcal{P}$ , che scambia le componenti  $R$  e  $L$  degli spinori e quindi il relativo comportamento destrorso o sinistrorso:

$$(A2.14) \quad \mathcal{P} \begin{vmatrix} \psi_L \\ \psi_R \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} \begin{vmatrix} \psi_L \\ \psi_R \end{vmatrix} = \gamma^0 \begin{vmatrix} \psi_L \\ \psi_R \end{vmatrix} = \begin{vmatrix} \psi_R \\ \psi_L \end{vmatrix}.$$

Essendo  $\sigma^2$  è l'unica matrice di Pauli a componenti immaginarie, eq. (A2.13) mette in evidenza che, in trasformazioni di Lorentz, gli spinori  $\sigma^2 \psi_R^*$  e  $-\sigma^2 \psi_L^*$  si comportano rispettivamente come spinore  $L$  o  $R$ . In generale, lo spinore

$$(A2.15) \quad \psi_{(c)} = \mathcal{C} \begin{vmatrix} \psi_L \\ \psi_R \end{vmatrix} = \begin{vmatrix} \sigma^2 \psi_R^* \\ -\sigma^2 \psi_L^* \end{vmatrix},$$

si dice «coniugato di carica» dello spinore a componenti  $\psi_{L,R}$ .

Si può quindi anche costruire uno spinore che ha apparentemente 4 componenti, tutte derivabili da due componenti di partenza, della forma

$$(A2.16) \quad \psi_M = \begin{vmatrix} \psi_L \\ -\sigma^2 \psi_L^* \end{vmatrix},$$

detto «spinore di Majorana». Ci sono dunque due tipi di spinori a 2 componenti: gli «spinori di Weyl»  $\psi_L$  e  $\psi_R$ , necessariamente di massa nulla, non sono autostati di  $\mathcal{C}$ ; gli spinori di Majorana, non necessariamente a massa nulla, sono invece autostati di  $\mathcal{C}$ . Ad esempio, si ritiene che i «neutralini», possibili componenti della DM, siano spinori di Majorana con massa ( $\gtrsim$  TeV).

Consideriamo ora gli operatori di proiezione (di chiralità)

$$(A2.17) \quad \Pi^\pm = \frac{1}{2}(\mathbf{1} \pm \gamma^5)$$

che selezionano la parte  $L$  o  $R$  di uno spinore. Infatti risulta

$$(A2.18) \quad \Pi^+ \begin{vmatrix} \psi_L \\ \psi_R \end{vmatrix} = \begin{vmatrix} \psi_L \\ \mathbf{0} \end{vmatrix}, \quad \Pi^- \begin{vmatrix} \psi_L \\ \psi_R \end{vmatrix} = \begin{vmatrix} \mathbf{0} \\ \psi_R \end{vmatrix},$$

e gli spinori  $\begin{vmatrix} \psi_L \\ \mathbf{0} \end{vmatrix}$  e  $\begin{vmatrix} \mathbf{0} \\ \psi_R \end{vmatrix}$  sono autovettori, rispettivamente, di  $\Pi^+$  e  $\Pi^-$ .

Sostituendo a  $\psi_{L,R}$  un'onda piana di argomento  $\pm i(p^\mu x^\mu)$  nelle (A2.6) e (A2.12), si ricava

$$(p^0)^2 - p^2 - m^2 = 0 \quad \text{con} \quad (p = |\mathbf{p}|),$$

$$(A2.19) \quad (p^0 + \mathbf{p} \cdot \boldsymbol{\sigma})\psi_R = -im\psi_L, \quad (p^0 - \mathbf{p} \cdot \boldsymbol{\sigma})\psi_L = -im\psi_R.$$

Sia  $\mathbf{n} = \mathbf{p}/|\mathbf{p}|$  la direzione del moto di una particella di impulso  $\mathbf{p}$ . A livello quantistico,  $\vec{\sigma}/2$  è l'operatore di spin dei quanti del campo spinoriale. Le eq. (A2.19) significano allora che, se è  $m = 0$ , la componente dello spin nella direzione del moto (elicità) è  $\pm 1/2$ , rispettivamente per  $\psi_L$  e  $\psi_R$ .

Per particelle spinoriali di massa nulla, dunque, lo spin è allineato alla direzione del moto, parallelo o anti-parallelo a seconda che lo spinore sia  $L$  (left: sinistrorso) o  $R$  (right: destrorso). Lo spin di particelle spinoriali ultrarelativistiche si comporta essenzialmente allo stesso modo. Componenti dello spin non allineate a  $\mathbf{n}$  si manifestano con sempre maggior probabilità al calare di  $p/m$ .

Sin qui si è usata per le matrici  $\gamma^\mu$  la «rappresentazione di Weyl». Si può passare a altre rappresentazioni tramite una trasformazione unitaria, che quindi «miscela» anche le 4 componenti dello spinore, sicché le parti  $L$  e  $R$ , in genere, non appaiono più separate. È importante la rappresentazione «di Dirac», non solo per ragioni storiche (è stata la prima a essere studiata), in cui le  $\gamma^i$  sono invariate (vedi eq. A2.2), mentre

$$(A2.20) \quad \gamma^0 = \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix}, \quad \gamma^5 = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}.$$

Diamo ora l'espressione di un boost sul piano 0-1 in questa rappresentazione. Quando un 4-vettore trasforma secondo la trasformazione di Lorentz (4), uno spinore nella rappresentazione di Dirac trasforma come segue:

$$(A2.21) \quad \psi \rightarrow S\psi \quad \text{con} \quad S = \sqrt{\frac{\gamma+1}{2}} \begin{vmatrix} \mathbf{1} & \frac{\gamma}{\gamma+1} \vec{\beta} \cdot \vec{\sigma} \\ \frac{\gamma}{\gamma+1} \vec{\beta} \cdot \vec{\sigma} & \mathbf{1} \end{vmatrix}$$

( $\gamma = (1 - \beta^2)^{-1/2}$  e  $\vec{\beta} = \mathbf{v}/c$  con  $\mathbf{v}$ : velocità relativa).

Elettroni e barioni sono quanti di campi spinori con massa. I neutrini (se senza massa) possono essere spinori  $L$ , o spinori di Majorana. Dopo la scoperta del «mixing» tra sapori neutrinici, che rende necessaria una (piccola) massa neutrinica, l'ipotesi più accreditata è che i neutrini siano spinori di Majorana.

Il comportamento dei quark è più complesso, anche perché essi possono essere considerate particelle libere solo a temperatura molto elevata, quando sono comunque ultrarelativistiche.

## Appendice A3

### Trasformazioni di gauge abeliane e non abeliane

La lagrangiana di  $N$  campi spinoriali  $\psi_a$  ( $a = 1, 2, \dots, N$ ) contiene i termini

$$(A3.1) \quad \mathcal{L}_0 = \bar{\psi}_a \gamma^\mu \partial^\mu \psi_a$$

ove si somma su apici e indici ripetuti. L'apice  $\mu$  ( $= 0, \dots, 3$ ) riguarda componenti spazio-temporali, su cui si somma con segnatura  $+- -$ , come consueto. L'indice  $a$ , invece, non etichetta le componenti spazio-temporali dello spinore, ma cataloga gli  $N$  spinori componenti un *multipletto*, nel cui ambito si assume invarianza rispetto alle trasformazioni

$$(A3.2) \quad \psi_a \rightarrow \psi'_a = (\delta_{ab} + ig\alpha_J \lambda_{Jab}) \psi_b.$$

In eq. (A3.2) abbiamo:

- 1) la costante  $g$ , che diventerà la costante di interazione;
- 2) i parametri reali infinitesimi  $\alpha_J$ , che definiscono l'elemento del gruppo di trasformazioni, combinando le matrici hermitiane  $\lambda_J$  che compongono la sua algebra.

Ove non sia indispensabile esplicitare gli indici, l'eq. (A3.2) si può scrivere semplicemente

$$(A3.3) \quad \psi \rightarrow \psi' = (1 + ig\alpha_J \lambda_J) \psi.$$

Per  $\alpha_J$  finiti, la trasformazione si sarebbe scritta

$$(A3.3) \quad \psi \rightarrow \psi' = e^{ig\alpha_J \lambda_J} \psi,$$

espressione coincidente con l'eq. (A3.3) per  $\alpha_J$  infinitesimi.

Essendo gli  $\alpha_J$  reali e le  $\lambda_J$  hermitiane, sarà anche

$$(A3.4) \quad \bar{\psi} \rightarrow \bar{\psi}' = \bar{\psi} e^{-ig\alpha_J \lambda_J} = \bar{\psi} (1 - ig\alpha_J \lambda_J)$$

sicché risulta

$$(A3.5) \quad \mathcal{L}_0 \rightarrow \bar{\psi} e^{-ig\alpha_J \lambda_J} \gamma^\mu e^{ig\alpha_J \lambda_J} \partial^\mu \psi = \bar{\psi} \gamma^\mu \partial^\mu \psi \equiv \mathcal{L}_0$$

purché sia lecito scambiare tra loro gli operatori  $\partial^\mu$  e  $e^{ig\alpha_J \lambda_J}$ . Se gli  $\alpha_J$  sono funzione di  $x^\mu$ , ciò non è lecito. Si noti che, invece, le matrici numeriche  $\gamma^\mu$  commutano con l'operatore  $e^{\pm ig\alpha_J \lambda_J}$ , che attiene uno «spazio» diverso.

Se gli  $\alpha_J$  sono funzione di  $x^\mu$ , avremo che

$$(A3.6) \quad \partial^\mu \psi \rightarrow \partial^\mu \psi' = \partial^\mu (1 + ig\alpha_J \lambda_J) \psi = (1 + ig\alpha_J \lambda_J) \partial^\mu \psi + ig \partial^\mu \alpha_J \lambda_J \psi =$$

sicché risulta

$$(A3.7) \quad \mathcal{L}_0 \rightarrow \mathcal{L}_0 + ig \bar{\psi} \gamma^\mu \partial^\mu \alpha_J \lambda_J \psi.$$

Il secondo termine, che rompe l'invarianza, può allora essere cancellato da un contro-termino se si assume l'esistenza di un potenziale di interazione

$$(A3.8) \quad V = ig \bar{\psi} \gamma^\mu \alpha_J^\mu \lambda_J \psi$$

contenente il multipletto di campi «ausiliari»  $\alpha_J^\mu$  che, quanto  $\psi$  trasforma secondo l'eq. (A3.3),

trasformerà in

$$(A3.9) \quad a_J^\mu = a_J^\mu + \delta a_J^\mu.$$

Ci proponiamo ora di ricavare l'espressione di  $\delta a_J^\mu$  che permette di preservare l'invarianza. A questo scopo vediamo come trasforma  $V$ , quando  $\psi$  e  $a_J^\mu$  trasformano, all'ordine più basso in  $\alpha_J$  (e assumendo  $\delta a_J^\mu$  dello stesso ordine  $\alpha_J$ ).

Questo calcolo è semplificato schematizzando  $V$  nella forma  $\bar{\psi} a_J \lambda_J \psi$ , senza costante di accoppiamento  $g$  e indici spazio-temporali, irrilevanti vista l'assenza di derivate. Sarà allora

$$\begin{aligned} \bar{\psi} a_J \lambda_J \psi &\rightarrow \bar{\psi} (1 - ig \alpha_R \lambda_R) (a_J + \delta a_J) (1 + ig \alpha_R \lambda_R) \lambda_J \psi = \\ &= \bar{\psi} a_J \lambda_J \psi + \bar{\psi} \delta a_J \lambda_J \psi + ig \alpha_R \bar{\psi} [\lambda_R, \lambda_J] a_J \psi \end{aligned}$$

vale a dire

$$(A3.10) \quad V \rightarrow V + ig \bar{\psi} \gamma^\mu \{ \delta a_J^\mu \lambda_J + \alpha_R ig [\lambda_R, \lambda_J] a_J^\mu \} \psi$$

e, mettendo assieme eq. (A3.7) e (A3.10), si ha che

$$\mathcal{L}_o + V \rightarrow \mathcal{L}_o + V + ig \bar{\psi} \gamma^\mu \{ \partial^\mu \alpha_J \lambda_J + \delta a_J^\mu \lambda_J + \alpha_R ig [\lambda_R, \lambda_J] a_J^\mu \} \psi$$

sicché l'invarianza è recuperata a condizione che sia

$$(A3.11) \quad \lambda_J \delta a_J^\mu = -\lambda_J \partial^\mu \alpha_J - ig [\lambda_R, \lambda_S] a_S^\mu \alpha_R$$

dove, nell'ultimo addendo, si è cambiato il nome dell'indice di somma. È qui che si vede perché una teoria di gauge richieda un'algebra di Lie. Infatti, solo se  $[\lambda_R, \lambda_S] = i \epsilon_{RSJ} \lambda_J$ , si può «raccoigliere»  $\lambda_J$  e definire la trasformazione dei campi vettoriali,

$$(A3.11) \quad \delta a_J^\mu = -\partial^\mu \alpha_J + g \epsilon_{RSJ} a_S^\mu \alpha_R,$$

che, in concomitanza alla trasformazione (A3.3) per le componenti spinoriali, rende invariante  $\mathcal{L}_o + V$  anche per  $\alpha_J$  «locali».

Si noti che questa trasformazione, valida per  $\lambda_J$  che costituiscono un'algebra, ha come sottocaso quello in cui ci sia un'unica matrice  $\lambda$  diagonale. In questo caso, nella trasformazione (A3.11), manca il secondo addendo a secondo membro: la trasformazione riassume la forma di una trasformazione di gauge in elettromagnetismo. Non si tratta neppure di un caso del tutto «banale», in quanto gli elementi della diagonale possono essere diversi tra loro e ciò corrisponde a differenti «cariche» per i termini del multipletto  $\psi_a$ .

Questa algebra prende il nome di  $U(1)$  e dà luogo a trasformazioni che coincidono con rotazioni sul piano complesso. Si parla in questo caso di una teoria di gauge commutativa o «abeliana», mentre la teoria di gauge generica, per il fatto che le matrici  $\lambda_J$  non commutano, viene denominata non-commutativa o «non-abeliana».

Per completare la trattazione si deve dare anche l'equivalente della «lagrangiana libera» del campo di gauge, quella che per il caso elettromagnetico assume la forma  $(1/4)F^{\mu\nu}F^{\mu\nu}$ , con  $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ .

Per campi di gauge non-abeliani, avremo

$$(A3.12) \quad \mathcal{L}_g = \frac{1}{4} f_J^{\mu\nu} f_J^{\mu\nu} \quad \text{con} \quad f_J^{\mu\nu} = \partial^\mu a_J^\nu - \partial^\nu a_J^\mu - g \epsilon_{JRS} a_R^\mu a_S^\nu$$

Usando le trasformazioni (A3.11) si verifica che

$$(A3.13) \quad \delta f_J^{\mu\nu} = g \epsilon_{RSJ} f_S^{\mu\nu} \alpha_R$$

e che  $\mathcal{L}_g$  è invariante.

# Appendice A4

## Il modello standard delle interazioni elettrodeboli

L'attuale comprensione delle forze elettrodeboli è nata dalla ricerca di una teoria rinormalizzabile, in cui le divergenze legate ai grafici di figura X.6 fossero regolarizzabili. La teoria di Fermi delle interazioni deboli permette valide previsioni, al più basso ordine perturbativo, ma non è rinormalizzabile. Se le forze, invece, sono portate da campi di gauge, la rinormalizzabilità è assicurata. Anche quando questo punto fu compreso, tuttavia, restava un grosso ostacolo all'idea che le forze deboli fossero dovute a campi di gauge non-abeliani: pareva escluso che essi potessero avere massa e questo sembrava implicare un raggio di azione infinito per le forze, come nell'elettromagnetismo.

Solo la scoperta del meccanismo di Higgs, che causa una rottura spontanea della simmetria, dando massa ai campi di gauge, senza violare la rinormalizzabilità, ha permesso di arrivare a una teoria delle forze elettrodeboli.

Tale meccanismo salva completamente tutti gli aspetti che si possono esprimere a livello di teoria di gauge. Quindi conviene trattare prima il modello elettrodebole «pulito», fissando così tutte le caratteristiche delle particelle implicate, per poi passare alla parte «sporca», l'azione dei bosoni di Higgs.

Questa appendice è dedicata essenzialmente alla parte «pulita» e in essa discuteremo le interazioni elettrodeboli sulla base di una teoria di gauge riferita all'algebra di  $SU(2)_L \otimes U(1)_Y$ . Spiegheremo tra breve il significato dei suffissi.

Le particelle di spin 1/2, della «famiglia» di massa minore, si dispongono nelle rappresentazioni di dimensioni 2 o 1 (primo numero tra parentesi) dell'algebra  $SU(2)$  nel modo sotto elencato:

$$\begin{array}{l} \left| \begin{array}{c} u \\ d \end{array} \right|_L \equiv (2, 1/6) \quad \left| \begin{array}{c} \nu_e \\ e \end{array} \right|_L \equiv (2, -1/2) \\ u_R \equiv (1, 2/3) \quad d_R \equiv (1, -1/3) \quad e_R \equiv (1, -1) \end{array}$$

Le altre famiglie, di massa maggiore, ricalcheranno questa distribuzione. Il secondo numero tra parentesi fornisce le cariche  $Y$  nell'interazione col campo di gauge di  $U(1)$ . Esso è scelto in modo che la loro carica elettrica, coerentemente con lo schema di rottura della simmetria, sia

$$(A4.1) \quad Q = T_3 + Y .$$

$T_3$  è l'autovalore della matrice diagonale in  $SU(2)$  e vale quindi  $\pm 1/2$  per la rappresentazione di dimensione 2, 0 per la rappresentazione di dimensione 1. Risulta perciò

$$\begin{array}{l} Q(\nu_e) = +1/2 - 1/2 = 0 \\ Q(e_L) = -1/2 - 1/2 = -1 \\ Q(u_L) = +1/2 + 1/6 = 2/3 \\ Q(d_L) = -1/2 + 1/6 = -1/3 \\ Q(e_R) = -1 \\ Q(u_L) = 2/3 \\ Q(d_L) = -1/3 \end{array}$$

Prima di Higgs, tutte le particelle hanno massa nulla. Quindi, gli spinori L (left) e R (right) hanno equazioni separate. Ad esempio, gli elettroni L e R, entrambi a massa nulla, sono par-

ticelle distinte. Le parti R sono sempre singoletti di  $SU(2)_L$ . Le parti L, invece, sono organizzate in doppietti. Ciò significa che i campi R, invarianti per trasformazioni di  $SU(2)$ , sono isolati dalle forze prodotte dai bosoni di gauge di  $SU(2)$ . Le parti L, invece, trasformano e interagiscono sicché, prima della rottura di simmetria, il potenziale delle forze è

$$V = ig_1 J^\mu Y^\mu + ig_2 J_J^\mu W_J^\mu$$

con

$$J^\mu = \frac{1}{6} |\bar{u} \bar{d}|_L \gamma^\mu \begin{vmatrix} u \\ d \end{vmatrix}_L + \frac{1}{2} |\bar{\nu}_e \bar{e}|_L \gamma^\mu \begin{vmatrix} \nu_e \\ e \end{vmatrix}_L +$$

$$+ \frac{2}{3} \bar{u}_R \gamma^\mu u_R - \frac{1}{3} \bar{d}_R \gamma^\mu d_R - \bar{e}_R \gamma^\mu e_R,$$

$$J_J^\mu = \frac{1}{2} |\bar{u} \bar{d}|_L \gamma^\mu \sigma_J \begin{vmatrix} u \\ d \end{vmatrix}_L + \frac{1}{2} |\bar{\nu}_e \bar{e}|_L \gamma^\mu \sigma_J \begin{vmatrix} \nu_e \\ e \end{vmatrix}_L$$

(A4.2)

dove: i coefficienti (numerici) degli addendi in  $J^\mu$  sono le «cariche  $Y$ »;  $g_1$  e  $g_2$  sono, rispettivamente, le costanti di interazione per le forze portate dai campi di gauge legati a  $U(1)$  e  $SU(2)$ ; il campo di gauge (abeliano) di  $U(1)$  è  $Y^\mu$  e il campo di gauge (non abeliano) di  $SU(2)$  è  $W_J^\mu$ .

Dopo la rottura, lo stesso potenziale va riorganizzato nella forma

$$V = \frac{ie}{\sqrt{8} \sin \theta_W} (W_+^\mu J_-^\mu + W_-^\mu J_+^\mu) +$$

$$+ \frac{ie}{2 \sin \theta_W \cos \theta_W} J_{NC}^\mu Z^\mu + \quad (\text{potenziale delle forze deboli})$$

$$+ ie J_{em}^\mu A^\mu \quad (\text{potenziale delle forze elettromagnetiche})$$

con  $J_\pm^\mu = 2(J_1^\mu \mp J_2^\mu)$ ,  $J_{NC}^\mu = 2(J_3^\mu - \sin^2 \theta_W J_{em}^\mu)$ ,  $(\text{correnti deboli})$

$$J_{em}^\mu = J_3^\mu + J^\mu \quad (\text{corrente elettrica})$$

Le espressioni del potenziale delle forze elettromagnetiche e della corrente elettrica, vista l'espressione (A4.2) di  $J^\mu$ , spiegano anche l'origine della relazione (A4.1).

Perché  $V$  coincida complessivamente con l'espressione (A4.2), deve essere

$$W_\pm^\mu = \frac{1}{\sqrt{2}} (W_1^\mu \mp W_2^\mu), \quad Z^\mu = \cos \theta_W W_3^\mu - \sin \theta_W Y^\mu$$

$$A^\mu = \sin \theta_W W_3^\mu + \cos \theta_W Y^\mu$$

(A4.3)

sicché deve anche essere

$$e = g_1 \cos \theta_W = g_2 \sin \theta_W,$$

(A4.4)

stabilendosi così una precisa relazione tra le intensità delle forze deboli e elettromagnetiche.

La fenomenologia delle forze deboli si basa sul potenziale di Fermi, in cui risultano direttamente accoppiate le correnti, mentre non appaiono i campi di gauge. Ciò è dovuto all'elevata massa ( $M_W, M_Z \sim 80 \text{ GeV}$ ) che essi acquisiscono alla rottura di simmetria, che li rende inosservabili se non in esperimenti a energie estremamente alte, prossime a  $100 \text{ GeV}$ .

Se si tiene conto dello scambio di particelle  $W_\pm$  e  $Z$ , ma ci si limita a studiare effetti di bassa energia, tutti i risultati quantitativi della teoria completa coincidono con quelli ottenibili dal

potenziale «alla Fermi»

$$(A4.5) \quad V_{eff} = \left( \frac{e}{\sqrt{8} \sin \theta_W M_W} \right)^2 J_+^\mu J_-^\mu + \left( \frac{e}{\sqrt{8} \sin \theta_W \cos \theta_W M_Z} \right)^2 J_{NC}^\mu J_{NC}^\mu$$

in cui non compaiono più i campi  $W^\mu$  e  $Z^\mu$ , ma solo le masse delle particelle  $W$  e  $Z$ . Se si confronta l'espressione di  $V_{eff}$  con l'espressione del potenziale fenomenologico alla Fermi

$$V_F = \frac{G_F}{\sqrt{2}} (J_+^\mu J_-^\mu + \rho J_{NC}^\mu J_{NC}^\mu),$$

in cui  $G_F = 1.1664 \times 10^{-5} GeV^{-2}$  e  $\rho = 1.022$ , si vede che deve essere

$$(A4.6) \quad M_W = \frac{37.4 \text{ GeV}}{\sin \theta_W}, \quad \rho = \left( \frac{M_W}{M_Z \cos \theta_W} \right)^2.$$

Le due relazioni (A4.6) contengono quantità che sono tutte sperimentalmente misurabili, a parte  $\theta_W$ . Concordemente quindi le due relazioni portano allo stesso valore per

$$\sin^2 \theta_W \simeq 0.2319$$

essendo sperimentalmente noti anche i valori delle masse  $M_W = 80.22 \text{ GeV}$ ,  $M_Z = 91.19 \text{ GeV}$ .

# Appendice A5

## La rottura spontanea delle simmetrie

### 1. Il modello di Goldstone

Il meccanismo che porta alla rottura spontanea delle simmetrie, in teoria dei campi, va visto anzitutto nel contesto più semplice, con un solo campo scalare complesso  $\phi$ , la cui lagrangiana

$$\mathcal{L} = \partial^\mu \phi^* \partial^\mu \phi - V(|\phi|),$$

contenente il potenziale

$$(A5.1) \quad V(|\phi|) = \lambda(|\phi|^2 - \phi_o^2)^2 = (-2\lambda\phi_o^2)|\phi|^2 + \lambda|\phi|^4 + \phi_o^4,$$

è invariante per trasformazioni di fase:

$$(A5.2) \quad \phi \rightarrow \phi' = e^{ix}\phi, \quad \phi^* \rightarrow \phi'^* = e^{-ix}\phi^*.$$

Tralasciando il termine costante  $\phi_o^4$ , che non ha effetti dinamici, e ponendo

$$(A5.3) \quad m^2 = -2\lambda\phi_o^2 \quad (\text{con questi segni!})$$

ci troviamo di fronte a un'ordinaria teoria di campo scalare, con quanti di massa  $m$ , auto-interagente tramite un termine «potenziale»  $\lambda|\phi|^4$ . Tuttavia, è chiaro che tutto si gioca sul segno di  $m^2$ ; se prendiamo  $\phi_o$  reale, per cui risulta  $m^2 < 0$ , ci si trova di fronte a una teoria tachionica. Infatti, sostituendo una soluzione di onda piana di argomento  $\pm ip^\mu x^\mu$  nelle equazioni del moto

$$(A5.4) \quad \partial^\alpha \partial^\alpha \phi^{(*)} - \nabla^2 \phi^{(*)} - m^2 \phi^{(*)} = 0,$$

dedotte dalla lagrangiana (A5.1) tralasciando il termine d'interazione, troviamo

$$(p^0)^2 = p^2 + m^2 \quad \text{con } p = |\mathbf{p}|$$

e, quindi,  $p/p^0 = v < c (> c)$  se  $m^2 > 0 (< 0)$ .

Nonostante le apparenze, tuttavia, anche nel caso  $m^2 < 0$  la teoria è perfettamente lecita. A questo proposito si consideri il grafico del potenziale  $V(|\phi|)$  in figura X.8. Per  $\phi_o^2 > 0$  ( $m^2 < 0$ ), il suo andamento è quello qualitativamente rappresentato dalle curve 1 o 2. Infatti la derivata

$$(A5.5) \quad \frac{\partial V}{\partial |\phi|} = 4|\phi|(|\phi|^2 - \phi_o^2)$$

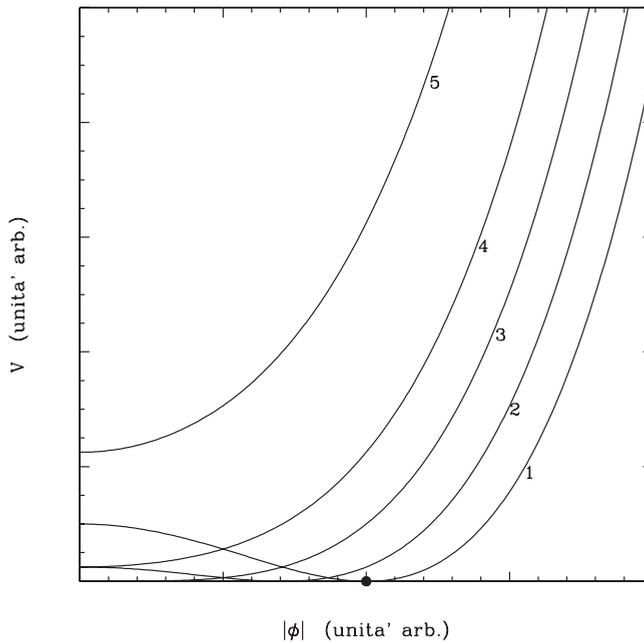
si annulla comunque per  $|\phi| = 0$  e per  $|\phi| = \phi_o$  se  $\phi_o^2 > 0$ . Avremo poi rispettivamente  $V(0) = \lambda\phi_o^4$ , mentre è  $V(\phi_o) = 0$ .

Se il secondo estremante esiste, esso è il minimo assoluto, mentre nell'origine c'è un massimo (casi 1 e 2 in figura). Se invece il secondo estremante non esiste, l'unico estremante, in  $\phi = 0$ , diviene un minimo; questo si verifica nei casi 4 e 5 in figura. Il caso 3 corrisponde a  $\phi_o = 0$ .

La densità d'energia del campo, conformemente a eq. (A1.25), sarà

$$(A5.6) \quad T^{00} = \rho = \frac{\partial \mathcal{L}}{\partial^0 \phi} \partial^0 \phi + \frac{\partial \mathcal{L}}{\partial^0 \phi^*} \partial^0 \phi^* - \mathcal{L} = |\dot{\phi}|^2 + \nabla \phi^* \cdot \nabla \phi + V(|\phi|).$$

Il minimo del (la densità d'energia) potenziale  $V$ , per  $|\phi| = \phi_o$  (o  $\phi = 0$ ), è anche il minimo di  $\rho$ .

**FIGURA X.8**

Andamento del potenziale di Goldstone per valori decrescenti di  $\phi_0^2$  (positivi e negativi). La curva 4 corrisponde a  $\phi_0^2 = 0$ .

Qui ci troviamo, in realtà, di fronte a un fatto nuovo nella relazione tra campi classici e quantistici, la relazione tra **configurazione** di minima energia e **stato** di vuoto. Ogni configurazione classica è completamente nota se si hanno i valori assunti al tempo  $t$  dalle coordinate lagrangiane  $q_i(t)$ . Una tale configurazione è anche un preciso stato quantistico. Esistono però stati quantistici in cui le coordinate lagrangiane non hanno valori definiti. Ad esempio, nel caso di un punto materiale, per un autostato di  $\mathbf{p}$  (impulso), esiste soltanto il «valore di aspettazione» di  $\mathbf{r}$ . Ciò segue al fatto che  $[r_i, p_j] = i\hbar\delta_{ij}$  sicché gli operatori  $\mathbf{r}$  e  $\mathbf{p}$  non commutano. Lo stato  $|\mathbf{p}\rangle$ , perfettamente lecito, non corrisponde dunque a una precisa configurazione classica.

In teoria dei campi, le funzioni di campo sono le coordinate lagrangiane. La teoria classica dei campi ci dirà facilmente – come abbiamo appena visto – i valori di tali coordinate lagrangiane nella configurazione di minima energia.

A livello quantistico, invece, lo stato di vuoto è certamente quello di minima energia, ma è precisamente definito come autostato di autovalore 0 dell'operatore  $n$ , numero di particelle,

$$n|0\rangle = 0,$$

e va ricordato che gli operatori  $n$  e  $\phi$  non commutano. Il valore classico di un campo  $\phi$ , per la configurazione di minima energia, diviene allora il **valore di aspettazione nel vuoto** (VEV) del campo quantistico  $\phi$ .

Tutto ciò non crea problemi, se la configurazione di minima energia è unica. È qui il punto critico di tutto il modello: lo stato definito da  $|\phi| = 0$  **non è unico**; sul piano  $Re\phi - Im\phi$ , tutti i punti sulla circonferenza di centro nell'origine e di raggio  $\phi_0$  sono egualmente minimi. Ciò che li distingue è una fase  $\alpha$ .



FIGURA X.9

Analogo meccanico: la verga in figura è compressa in direzione orizzontale da una morsa schematizzata dalle superfici circolari agli estremi. Il sistema, comprese le forze agenti su di esso, è del tutto simmetrico per rotazioni (come indicate dalla tratteggiata). Ciononostante, aumentando la pressione, la verga si piegherà in una direzione a priori indeterminata. Una piccolissima perturbazione, dovuta a impurità del metallo, genera in questo modo una rottura grossolana della simmetria.

Di fatto, tuttavia, il sistema «sceglierà» una delle fasi  $\alpha$  come quella che caratterizza il «vuoto». In altri termini, esistono  $\infty$  configurazioni classiche di minima energia, per  $|\phi| = \phi_0$  e  $\alpha$  arbitrario. Lo stato quantistico di vuoto è invece unico e esisterà un VEV (*vacuum expectation value*: valore di aspettazione nel vuoto) di  $\phi$ , il cui modulo è  $\phi_0$ , ma la cui fase è anche precisamente *determinata*. Valori di  $\phi$  diversi corrispondono a stati  $|\Psi\rangle \neq |0\rangle$ .

Conviene quindi definire tutte le configurazioni del sistema con riferimento al VEV di  $\phi$ , che rompe la simmetria di fase.

È utile considerare un analogo, in meccanica classica dei solidi, di questa situazione quantistica e relativistica. Il sistema in figura X.9 è costituito da una verga sottile, compressa alle estremità. Esso è simmetrico per rotazioni attorno all'asse della verga. Quando la compressione supera un determinato limite, la verga s'incurva e sarà la reazione elastica all'incurvamento a compensare le forze agenti agli estremi.

La direzione in cui la verga si incurva, a priori, è indeterminata. Nella realtà fisica essa dipenderà da piccolissime asimmetrie nelle strutture cristalline interne al metallo. Queste perturbazioni, di ampiezza trascurabile, ingenerano una rottura grossolana della simmetria.

Tornando al modello di Goldstone, senza perdita di generalità possiamo far passare l'asse  $Re\phi$  per la direzione di rottura della simmetria e definire

$$(A5.7) \quad Re\phi = \phi_0 + \phi_1/\sqrt{2}, \quad Im\phi = \phi_2/\sqrt{2}.$$

La lagrangiana (A5.1), rispetto a queste variabili di campo, ha allora espressione

$$(A5.8) \quad \begin{aligned} \mathcal{L} &= \frac{1}{2}\partial^\mu\phi_1\partial_\mu\phi_1 + \frac{1}{2}\partial^\mu\phi_2\partial_\mu\phi_2 - V \\ \text{con } V &= \lambda \left[ \left( \phi_0 + \frac{1}{\sqrt{2}}\phi_1 \right)^2 + \frac{1}{2}\phi_2^2 - \phi_0^2 \right]^2 = \lambda \left[ \sqrt{2}\phi_0\phi_1 + \frac{1}{2}\phi_1^2 + \frac{1}{2}\phi_2^2 \right]^2 = \\ &= 2\lambda\phi_0^2\phi_1^2 + \sqrt{2}\lambda\phi_0\phi_1(\phi_1^2 + \phi_2^2) + \frac{1}{4}\lambda(\phi_1^2 + \phi_2^2)^2. \end{aligned}$$

Questa espressione è ordinata secondo le potenze dei campi ( $\phi_0$  non è una «variabile» di campo). Il primo termine, quadratico, è un «termine di massa». I termini cubico e quartico sono «termini di interazione». Ponendo

$$(A5.9) \quad \mu^2/2 = 2\lambda\phi_0^2$$

si vede che i quanti del campo  $\phi_1$  avranno massa  $\mu$ , mentre quelli del campo  $\phi_2$  hanno massa nulla. I segni di  $\phi_0^2$  e  $\mu^2$  qui sono concordi; nella relazione (A5.3), invece, i segni di  $\phi_0^2$  e  $m^2$  erano discordi. La rottura di simmetria ha dunque portato a campi assolutamente non tachionici.

## 2. Il modello di Higgs in teoria di gauge abeliana

È relativamente semplice, sul piano formale, passare dal modello di Goldstone al modello di Higgs. Se le trasformazioni (A5.2) divengono

$$(A5.10) \quad \phi \rightarrow \phi' = e^{ig\alpha}\phi, \quad \phi^* \rightarrow \phi'^* = e^{-ig\alpha}\phi^*.$$

con  $\alpha$  locali,  $g$  essendo una «costante di interazione», l'invarianza è preservata se si introduce un campo di gauge  $A^\mu$ , con la lagrangiana (A5.1) che diviene

$$(A5.11) \quad \begin{aligned} \mathcal{L} &= (\partial^\mu - igA^\mu)\phi^* (\partial^\mu + igA^\mu)\phi - V = \\ &= \partial^\mu\phi^* \partial^\mu\phi + ig(\partial^\mu\phi^*\phi - \phi^*\partial^\mu\phi)A^\mu + g^2|\phi|^2 A^\mu A^\mu - V \end{aligned}$$

cui va aggiunta la «lagrangiana libera» del campo di gauge

$$(A5.12) \quad \mathcal{L}_g = (1/4)F^{\mu\nu}F^{\mu\nu} \quad \text{con} \quad F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$$

A questo livello il campo di gauge non ha massa. Un «termine di massa», inevitabilmente della forma  $m^2 A^\mu A^\mu$ , distruggerebbe l'invarianza di gauge.

La rottura spontanea della simmetria, che avviene come per Goldstone, genera invece tale termine. Ciò che allora viola l'invarianza è lo stato di minima energia. Quindi, anche dopo la rottura spontanea, la lagrangiana è invariante e la parte di spin=0 del campo vettoriale non entra nel gioco dinamico.

Tuttavia, usando  $\phi_0, \phi_1, \phi_2$ , avremo:

$$(A5.13) \quad g^2|\phi|^2 A^\mu A^\mu = g^2 \left[ (\phi_0 + \phi_1/\sqrt{2})^2 + \phi_2^2/2 \right]^2 A^\mu A^\mu = g^2\phi_0^2 A^\mu A^\mu + \dots\dots\dots$$

I termini indicati con ..... sono termini «di interazione», di grado  $> 2$  nei campi. Il termine  $g^2\phi_0^2 A^\mu A^\mu \equiv m^2 A^\mu A^\mu$ , invece, viene a costituire un «termine di massa» per il campo vettoriale di gauge.

Se il campo di gauge si accoppia dunque a un campo scalare complesso, con VEV non nullo, i quanti del campo acquisiscono massa.

## 3. La rottura della simmetria $SU(2) \otimes U(1)$

Lo schema qui introdotto si estende senza problemi al caso non abeliano, introducendo 4 nuove particelle scalari, che sono i quanti di un doppietto di campi scalari complessi  $\phi_a$ , appartenenti alla rappresentazione fondamentale di  $SU(2)_L$ . Il valore di  $Y$  assegnato al doppietto è  $-1/2$ , sicché le due componenti avranno, rispettivamente, carica elettrica 0 e  $-1$ .

Il potenziale che porta alla rottura della simmetria sarà della forma

$$(A5.14) \quad V(\phi) = \lambda \left( |\phi|^2 - \phi_0^2 \right)^2 \quad \text{con VEV} \quad \phi_0 = \left| \begin{array}{c} v/\sqrt{2} \\ 0 \end{array} \right|$$

sicché la massa delle particelle di Higgs, dopo la rottura di simmetria, sarà  $\sqrt{2\lambda}\phi_0$ . La stessa rottura genera, similmente a quanto visto in eq. (A5.13), i termini di massa:

$$(A5.15) \quad \mathcal{L}_{\text{massa}} = \left( \frac{g_1 v}{2} \right)^2 W_+^\mu W_-^\mu + \frac{v^2}{8} |W_3^\mu Y^\mu| \left| \begin{array}{cc} g_1^2 & -g_1 g_2 \\ -g_1 g_2 & g_2^2 \end{array} \right| \left| \begin{array}{c} W_3^\mu \\ Y^\mu \end{array} \right|.$$

La matrice  $2 \otimes 2$  viene poi diagonalizzata passando ai campi  $Z^\mu$  e  $A^\mu$ , sicché alla fine avremo

le masse

$$(A5.16) \quad m_W = g_1 V/2, \quad m_Z = \sqrt{g_1^2 + g_2^2} V/2, \quad m_A = 0$$

conformemente a quanto visto in appendice A4.

Il campo di Higgs viene anche «usato» per produrre la massa dei fermioni, causando così sia la relazione tra parte  $L$  e parte  $R$  che l'accoppiamento tra generazioni di particelle.

Quest'ultimo «uso» del campo di Higgs aumenta il livello di consistenza del modello, ma non è molto predittivo: a ogni parametro osservativo corrisponde una costante immessa nel modello. Rimandiamo ai trattati specializzati per questa parte dell'analisi e, in particolare, per l'accoppiamento inter-generazionale.

Val la pena comunque di notare che ci sono tre tipi di connotati per le particelle elementari:

- 1) Quelli legati alle forze elettrodeboli, determinati dagli autovalori delle matrici diagonali in  $SU(2) \otimes U(1)$ , che abbiamo qui discusso.
- 2) Quelli legati alle forze forti, determinati dagli autovalori delle 2 matrici diagonali di  $SU(3)$  di colore.
- 3) Quelli legati alla «famiglia».

La gerarchia di famiglie ci appare come un cascame non compreso di una dinamica dormiente, probabilmente attiva a temperature molto maggiori di quelle osservative, dove potrebbe anche essere legata a forze ulteriori, rispetto alle 4 forze che oggi si osservano in natura. A questo proposito va anche notato che gli accoppiamenti di Higgs non si estrinsecano con forze; essi non determinano una «dinamica» ma una «statica»: i dati sperimentali riflettono la sistemazione nelle configurazioni di minima energia rispetto ai relativi potenziali.

## Appendice A6

### Formazione dei residui di particelle fredde

L'evoluzione temporale della densità in numero ( $n_\chi$ ) di particelle  $\chi$  non relativistiche (non necessariamente neutralini), che si possono annichilire per interazione diretta, si ricava dall'equazione

$$(A6.1) \quad \frac{dn_\chi}{dt} = \Psi - 3\frac{\dot{a}}{a}n_\chi - \langle\sigma_a v_T\rangle n_\chi^2$$

dove:

- 1)  $\Psi$  è un termine di sorgente, che rende conto di un'eventuale input di particelle  $\chi$ , a causa di processi esterni;
- 2) il termine  $(\dot{a}/a)n_\chi$  tiene conto della diluizione dovuta all'espansione del cosmo;
- 3) il termine  $n_\chi^2\langle\sigma_a v_T\rangle$  tiene conto del processo di annichilazione  $\chi + \chi \rightarrow$  particelle (ordinarie) più leggere.

Se i termini 1) e 3) sono trascurabili, si ricava la legge  $n_\chi \propto a^{-3}$ , da usare dopo il «disaccoppiamento». In effetti, posto  $\tau = 1/n_\chi\langle\sigma_a v_T\rangle$ , gli ultimi due termini dell'eq. (A6.1) sono della forma  $n_\chi/2t$  (consideriamo un'espansione radiativa) e  $n_\chi/\tau$ . Il tempo  $t$  a cui  $2t = \tau$  segna il prevalere della diluizione da espansione rispetto all'annichilazione. La prevalenza, tuttavia, non ha effetti istantanei ed è questa la ragione per cui occorre studiare l'eq. (A6.1) se si vuole ricostruire quantitativamente il fondo residuo di  $\chi$ .

Convien allora definire la densità comovente  $n_c = n_\chi a^3$ . Risultando

$$\frac{d}{dt}n_c = \frac{d}{dt}(n_\chi a^3) + 3\frac{\dot{a}}{a}(n_\chi a^3)$$

e tralasciando il termine di input, avremo

$$(A6.2) \quad \frac{d}{dt}n_c + n_c^2 \frac{\langle\sigma_a v_T\rangle}{a^3} = 0$$

La dipendenza dal tempo di  $\langle\sigma_a v_T\rangle$  è critica nel fissare la soluzione di questa equazione. A titolo di esempio, se fosse  $\langle\sigma_a v_T\rangle = \text{const.}$ , l'eq. (A6.2), nell'era radiativa, acquisisce la forma

$$\frac{dn_c}{n_c^2} + \kappa \frac{dt}{t^{3/2}} = 0 \quad \text{con} \quad \kappa = \langle\sigma_a v_T\rangle t_{\chi,dg}^{3/2} / a_{\chi,dg}^3 \quad \text{costante,}$$

ed ha l'integrale

$$n_c(t) = \frac{n_c(t_{\chi,dg})}{1 + 2t_{\chi,dg}\langle\sigma_a v_T\rangle \frac{n_c(t_{\chi,dg})}{a_{\chi,dg}^3} [1 - (t_{\chi,dg}/t)^{1/2}]}$$

ovvero

$$(A6.3) \quad n_\chi(t) = \frac{n_\chi(t_{\chi,dg})}{1 + 2(t_{\chi,dg}/\tau)[1 - (t_{\chi,dg}/t)^{1/2}]} \left[ \frac{a_{\chi,dg}}{a(t)} \right]^3$$

Pertanto, se si fosse scelto come istante di disaccoppiamento il tempo  $t_{\chi,dg} = \tau$ , assumendo l'assenza di annichilazioni da allora in poi, si troverebbe un valore di  $n_{\chi}(t)$ , a  $t \gg t_{\chi,dg}$ , che è circa 3 volte maggiore di quello calcolato da eq. (A6.3) passando al limite  $t \rightarrow \infty$ .

Indipendentemente dalla validità dell'ipotesi  $\kappa = \text{const.}$  ciò mostra che la correzione dovuta alle annichilazioni a  $t > t_{\chi,dg}$  deve comunque essere tenuta presente in valutazioni quantitative.

## CAPITOLO XI

# L'universo inflazionario

### 1. Il laboratorio cosmologico

Alla fine degli anni Settanta, la fisica fondamentale si trova di fronte a un problema del tutto nuovo. La crisi si apre con un grande successo, di cui abbiamo parlato nel precedente capitolo, la scoperta del modello standard delle forze elettrodeboli. Non era allora chiara l'enorme quantità di dati che esso avrebbe permesso di catalogare, ma c'era la netta sensazione di avere in mano il bandolo della matassa. Il modello standard, tuttavia, dipende ancora da una ventina di parametri. Troppi, per essere un punto d'arrivo.

Si apre così la caccia alla *fisica oltre il modello standard*. Per un paio di decenni, sperimentali e teorici cercano effetti di altissima energia che generino violazioni del modello standard a energie minori, tecnicamente accessibili. Un po' come la struttura molecolare, su scale  $\sim 10^{-6}$  cm, limita la validità del modello di *fluido perfetto*, già su scale dell'ordine del millimetro, causando conduzione termica, viscosità, tensione superficiale, ecc.

Ebbene, questi effetti ci sono! La materia oscura è fatta di particelle che non rientrano nel modello standard; i neutrini hanno una piccolissima massa, a differenza di quanto previsto dal modello standard; più recentemente, poi, si è scoperta l'energia oscura, la cui natura fa probabilmente riferimento a fisica oltre il modello standard. Si tratta peraltro di dati provenienti dal mondo astrofisico e che già fanno presagire l'importanza che osservabili astrofisici potranno avere per la fisica fondamentale. Nessuna ricaduta di bassa energia, attribuibile alle supersimmetrie, o alla grande unificazione, viene rilevata nei dati di laboratorio.

Occorre dunque esplorare ambiti energetici ancora più elevati di quelli permessi dal LEP. Nasce così il progetto LHC, basato sull'aspettativa di «fisica oltre il modello standard» a energie fra 1 e 10 TeV. A queste energie potrebbero esserci segnali diretti delle supersimmetrie (SUSY); oppure si potrebbero scoprire le particelle di una quarta famiglia, con ulteriori *sapori* di quark; oppure potrebbero vedersi finalmente effetti *secondari* attribuibili indirettamente alle GUT. Ma nulla esclude che, anche a energie 100 volte quelle già sondate, possa esserci ancora e solo fisica consistente col modello standard.

In effetti, l'ispezione critica, quella da cui ci si potrebbe aspettare conoscenza radicalmente nuova, era stata ben focalizzata proprio dagli anni Settanta. Occorre accedere alle energie della *grande unificazione* (GUT), attesa a temperature  $T_{GUT} \sim 10^{15}$  GeV, quando le costanti di interazione delle forze elettromagnetiche, deboli e forti raggiungono valori simili. Si tratta purtroppo di un'energia sciaguratamente elevata, appena 3 o 4 ordini di grandezza al di sotto della temperatura di Planck  $T_p \simeq 1.22 \times 10^{19}$  GeV, che segna il passaggio al regime di gravità quantistica, e 10-12 ordini di grandezza al di sopra delle energie esplorabili da LHC.

È in questa luce che va vista la teoria dell'inflazione: il laboratorio cosmologico in sostituzione dell'acceleratore. La sua prima versione venne proposta, proprio al finire degli anni Settanta, da Starobinsky, Kazanas e Guth. La novità era il tentativo di risolvere assieme problemi di fisica delle interazioni fondamentali e di cosmologia. Lasciando in disparte la trattazione storica, in questo capitolo ci proponiamo di riassumere gli elementi portanti di quanto si ritiene sia accaduto poco dopo il *Big-bang*, lasciando tracce riconoscibili nella fisica del mondo odierno. La conclusione a cui si perviene è che precisi osservabili astrofisici, esplorabili tramite esperimenti, magari costosi, ma di imminente fattibilità, potrebbero produrre conoscenza attinente quella fisica del microcosmo che si stenta sempre più a approcciare tramite esperimenti di laboratorio.

## 2. Perché l'inflazione

Il successo dello scenario dell'instabilità gravitazionale mostra che tutta la complessa fenomenologia del cosmo attuale può essere derivata da «condizioni iniziali» di (quasi-)omogeneità e isotropia.

L'omogeneità, l'isotropia, nonché l'ampiezza e la distribuzione delle loro modeste violazioni, sono poi verificate, quasi per ispezione diretta, nei dati su anisotropia e polarizzazione del CMB. Ricordiamo che la temperatura attuale del CMB è  $\sim 2.73$  K, salvo fluttuazioni di  $1:10^5$  tra le diverse posizioni sulla volta celeste. La sorgente del CMB si trova a un *redshift*  $z_{rec} \sim 1000$  ed è il plasma primordiale in corso di ricombinazione, a una temperatura 1000 volte maggiore dell'attuale. La temperatura che misuriamo, eguale in tutte le direzioni, è frutto di una temperatura eguale in tutti i siti di emissione e di un *redshift* che ha agito allo stesso modo in tutte le direzioni. L'età del cosmo a  $z_{rec}$  è l'età attuale divisa  $\sim 1000^{3/2}$  (all'incirca; il dato sarebbe esatto in una cosmologia SCDM); di conseguenza l'orizzonte comovente di allora è (all'incirca) quello attuale diviso per  $\sim 1000^{1/2}$ . Sulla volta celeste, pertanto, vediamo radiazione proveniente da almeno  $10^4$  siti tra cui non è mai apparentemente esistito un contatto causale, ciascuno esterno all'orizzonte di tutti gli altri. Eppure, la temperatura a  $z_{rec}$  e la dinamica successiva sono le stesse in tutte le direzioni.

Le stesse tenui disomogeneità primordiali, necessarie al successo dello scenario di instabilità gravitazionale e misurate dalle modeste fluttuazioni di temperatura del CMB, soddisfano anch'esse al criterio di omogeneità e isotropia: le direzioni  $\mathbf{n}$  delle

onde di compressione che le generano sono isotropicamente distribuite e le loro ampiezze sono distribuite (quasi) Gaussianamente attorno valori medi omogenei nello spazio.

Essendo così passati *ex pluribus ad unum* (dall'articolazione del cosmo attuale a un quadro primordiale unitario), resta da capire come è possibile che, tra le molte possibili condizioni iniziali, in cui il tensore metrico  $g^{\mu\nu}(x^\mu)$  è una generica funzione delle coordinate spazio-temporali, si riscontrino quelle caratterizzate da (quasi-)omogeneità e isotropia.

Si tratta di un quesito dai contorni assai delicati. Se il *mondo osservabile* ha queste caratteristiche, ha senso discutere sulla loro *probabilità*? Per farlo bisogna riferirsi a una realtà fisica più ampia, per definizione *inosservabile*.

Apparentemente, diventa inevitabile assumere l'esistenza di un infinito meta-cosmo, al di là dell'orizzonte. Ma allora, in una tale infinità, una qualunque condizione sarà comunque realizzata in un qualche orizzonte. E se l'esistenza di un osservatore, all'interno di un orizzonte, è condizionata da *opportuni* limiti su  $g_{\mu\nu}(x_\mu)$ , alla fin fine il problema non esiste: l'osservatore non può che osservare una  $g_{\mu\nu}(x_\mu)$  *opportuna*.

È questa una forma piuttosto blanda del cosiddetto *principio antropico*, una congettura che ipotizza che condizioni diverse da quelle osservate siano inosservabili, per la semplice carenza di realtà cosmiche (pianeti, stelle, galassie, ammassi, ecc.) che possano ospitare un osservatore. Per questa congettura, il meta-cosmo è desolatamente disabitato; l'osservatore non si colloca in un sito *generico*: solo in siti estremamente specializzati un osservatore può esistere.

### ***Inflazione generica e modelli inflazionari***

La prima motivazione dell'*inflazione cosmologica* è quella di rendere più «generica» la collocazione meta-cosmica dell'osservatore. Si tratta di un obiettivo filosoficamente interessante, ma povero di risvolti fisici. La ragione del successo dell'inflazione è un'altra: una volta ottenuto questo risultato, segue una serie di ulteriori previsioni. Alcune sono univoche, valgono per qualsiasi modello di inflazione. Altre variano di modello in modello.

Ebbene, le prime sono *vere*: i dati le confermano. Le seconde, invece, serviranno a delimitare la fisica delle altissime energie sulla base di dati.

Nel procedere della ricerca sull'inflazione si incontrano quindi due scuole di pensiero: una scuola «astrofisica» che cerca di formulare i vincoli ai possibili quadri teorici; una scuola «microfisica» che parte da opportuni modelli fisici altrimenti motivati e cerca di verificare:

- 1) se essi abbiano impatto sull'inflazione,
- 2) gli eventuali limiti che la cosmologia impone ai parametri dei loro modelli.

Alla base dello schema sta comunque un insieme di «previsioni» assolutamente generiche, di cui facciamo una lista:

- 1) *Lo stesso meccanismo che ingenera omogeneità e isotropia spinge il parametro totale di densità,  $\Omega_o$ , su valori assai prossimi all'unità.* Al momento in cui venne proposta l'inflazione, si fronteggiavano due alternative osservative:  $\Omega_o \simeq 1$  e  $\Omega_o \simeq 0.2-0.3$ . La prima coincideva con la previsione «generica» dell'inflazione. Ma anche la seconda alternativa diveniva accettabile, senza orribili *fine tunings*, solo grazie a opportuni adattamenti del meccanismo inflazionario. Dati successivi hanno mutato lo scenario sperimentale. Il valore  $\Omega_o = 1$  è oggi al centro dell'intervallo osservativo. Resta solo da puntare su esperimenti che riducano ulteriormente l'ampiezza di tale intervallo.
- 2) *La dinamica delle transizioni di fase primordiali provoca mostri topologici, come monopoli magnetici, muri di dominio, stringhe cosmiche. È difficile evitare la previsione che ce ne siano in numero significativo nel nostro orizzonte osservativo. Tali anomalie sono state accuratamente ricercate nei campioni profondi, ma non ce n'è traccia. L'inflazione elimina i mostri topologici dall'orizzonte osservativo.*
- 3) *Nell'ambito dei modelli SUGRA, che estendono le supersimmetrie (SUSY) alla gravità e sono considerati attendibili, si prevede l'esistenza di partner supersimmetrici del gravitone, denominati gravitini. In contesti teorici analoghi, si prevedono anche quanti dei campi di modulo. A ridosso della scala di Planck quantità abnormi di queste particelle dovrebbero essere state create sicché, in ambito SUSY, il cosmo attuale dovrebbe essere composto essenzialmente di gravitini. Ciò evidentemente non corrisponde alla realtà osservativa che, anzi, fornisce limiti molto stringenti alla loro abbondanza. L'inflazione diluisce radicalmente i gravitini primordiali e gli altri quanti analoghi, rendendo SUGRA e SUSY del tutto coerenti coi dati.*
- 4) *Altre previsioni riguardano la natura e la statistica delle fluttuazioni di densità primordiali. Qui bisogna distinguere tra previsioni generiche e violazioni di tali previsioni generiche che caratterizzano specifiche classi di modelli di inflazione. I dati osservativi attuali concordano con le seguenti previsioni generiche sulle fluttuazioni primordiali: (a) Esse debbono essere adiabatiche. (b) La distribuzione delle loro ampiezze attorno al valor medio, scala per scala, deve essere Gaussiana. (c) Il loro spettro primordiale*

$$(XI.1) \quad P(k) = A k^{n-4}$$

deve avere un *indice spettrale*  $n \simeq 1$  o di poco inferiore all'unità. Particolarmente efficaci, in questo campo, sono state le misure sulle anisotropie del CMB, che colgono le fluttuazioni ancora a livello lineare. I dati attuali sono del tutto coerenti con le previsioni (a), (b), sebbene resti ancora un discreto spazio per lievi violazioni. Per l'indice spettrale  $n$ , i dati sono coerenti con valori nell'intervallo 0.95-1.

### **Vincoli osservativi ai modelli. Le onde gravitazionali.**

Vincoli più specifici ai modelli inflazionari possono invece derivare da osservazioni più avanzate nei seguenti ambiti:

- 1) Forma e natura delle fluttuazioni primordiali: Le lievi deviazioni di  $n$  dall'unità, le tenui non-Gaussianità e l'eventuale presenza di una modesta componente di fluttuazioni primordiali non-adiabatiche possono portare a forti vincoli sulla microfisica responsabile dell'inflazione.
- 2) L'ampiezza media delle fluttuazioni «scalari», cioè il valore della costante  $A$  nello spettro primordiale (XI.1). Si tratta di un punto critico dell'inflazione: Modelli generici tendono a «produrre» fluttuazioni di ampiezza eccessiva. Ottenere un'ampiezza coerente coi dati osservativi è una delle sfide che i modelli specifici debbono superare.
- 3) Le onde gravitazionali (GW: *gravitational waves*): Accanto al modo scalare, che dà luogo allo spettro (XI.1), è possibile un modo «tensoriale», che si estrinseca in GW primordiali. Si possono istituire vincoli abbastanza generici tra presenza di GW e valore di  $n$ . Si possono altresì istituire vincoli tra l'intensità del modo tensoriale e la scala energetica a cui avviene l'inflazione. Modelli inflazionari specifici possono anche portare alla violazione di questi vincoli.

Tra i veicoli fisici che portano informazione dall'Universo profondo, le GW sono il veicolo che reca i messaggi più ancestrali. L'Universo è opaco ai fotoni a partire da  $z_{rec} \simeq 10^3$ : il CMB reca informazione sul cosmo 400 mila anni dopo il *Big-bang*. L'Universo è opaco ai neutrini a partire da  $z_\nu \simeq 10^{10}$ , quando l'età del cosmo si misurava in minuti. Il neutrini del fondo cosmico ci porterebbero informazione sul cosmo alla vigilia della BBN, ma nessuna tecnologia è attualmente adeguata a osservarli. L'Universo, infine, è trasparente a GW fin dalla scala di Planck. Un'analisi delle GW ci porterebbe veramente informazione su microfisica inaccessibile in laboratorio.

L'osservazione di GW primordiali tramite antenne gravitazionali è assai problematica. Ma sembra esserci la possibilità di «vedere» GW tramite i loro effetti sulla polarizzazione del CMB.

### **3. La dinamica inflazionaria**

La cinematica del cosmo è determinata dal *fattore di scala*  $a(t)$ ;  $a(t)$  crescente significa espansione. Il ritmo di espansione, peraltro, può crescere o affievolirsi nel tempo: l'espansione può essere *accelerata*, o *decelerata*.

L'inflazione è un periodo primordiale, sufficientemente lungo, di espansione accelerata. Il paradigma inflazionario spiega che un'espansione accelerata durata  $\sim 10^{-35}$ - $10^{-40}$  sec., a ridosso del tempo di Planck, può avere messo in contatto causale zone del cosmo talmente vaste da includere tutto l'orizzonte oggi osservabile, causando gli altri effetti sopra enumerati.

Ci sono specifici vincoli all'*equazione di stato* dei contenuti cosmici, perché si produca espansione accelerata, che comunque è un effetto general-relativistico. In gravità newtoniana ciò che causa gravità è la *massa*. Nelle equazioni di Einstein, invece, la sorgente è data dalle componenti del tensore energia-impulso  $T_{\mu\nu}$ , che comprendono sia la densità d'energia  $\rho$  che la pressione  $p$ . È la pressione che modifica le cose.

Per la materia ordinaria, fatta essenzialmente di particelle non relativistiche, è  $\rho \gg |p|$ . Questa è la ragione del successo di Newton: se si studia il moto dei pianeti o delle stelle, le pressioni che entrano in gioco sono trascurabili e la gravitazione universale non ha bisogno di correzioni. L'azione della massa è sempre attrattiva, per Newton l'antigravità non esiste e la massa del cosmo non può che rallentare il ritmo di espansione. Come l'attrazione della Terra su un corpo spedito verso l'alto ne rallenta l'ascesa.

Per questo è clamoroso il *dato* che l'espansione del cosmo sia accelerata. Ciò che a Newton è impossibile, può però essere concesso a Einstein. In effetti, le equazioni cosmologiche di FRW, caso particolare delle equazioni di Einstein, mostrano che l'espansione può essere accelerata in presenza di *pressioni negative* che verifichino la condizione

$$(XI.2) \quad p < -\rho/3$$

(vedi appendice A1). Ciò che è un *fatto* nel cosmo odierno, come discuteremo meglio nel prossimo capitolo dedicato all'*energia oscura*, potrebbe essere avvenuto anche nei primi istanti dell'espansione cosmica.

Vediamo quali potrebbero essere le sostanze che soddisfano eq. (XI.2):

**Il falso vuoto.** Ai connotati del falso vuoto si può arrivare da due punti di vista. Un cultore di Relatività generale, che vede come sorgente dell'espansione il *tensore energia-impulso*  $T^{\mu\nu}$ , può ipotizzare una sorgente della forma

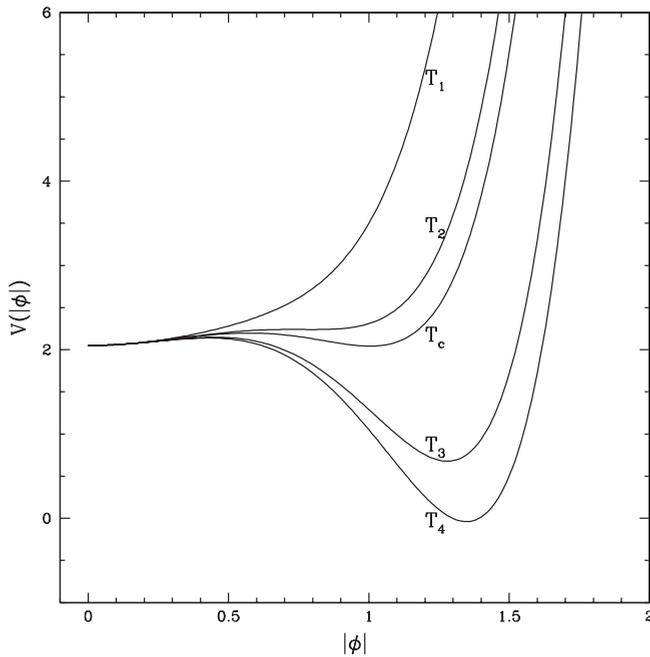
$$(XI.3) \quad T^{\mu\nu} = \Lambda g^{\mu\nu} \quad (g^{\mu\nu}: \text{tensore}, \Lambda: \text{costante scalare positivo}).$$

Se abbiamo  $g^{00} = 1$ ,  $g^{ii} = -1$  (metrica di background), risulta

$$(XI.4) \quad T^{00} = \rho = \Lambda, \quad T^{ii} = p = -\Lambda \quad (\text{per } i = 1, 2, 3)$$

e quindi  $p = -\rho$ . Quando Einstein introdusse la sua *costante cosmologica*, salvo poi considerarla «il più grosso svarione della mia vita», fece essenzialmente questo.

Ma al falso vuoto si arriva anche da un contesto particellare. Abbiamo visto, nel capitolo precedente, il ruolo dei campi (scalari) di Higgs, che portano alla *rottura spontanea* delle simmetrie e la dipendenza dell'andamento della loro densità d'energia potenziale (o, semplicemente, *potenziale*)  $V(\phi)$  da  $T$  (temperatura). È uno dei risultati chiave della meccanica quantistica a temperatura finita, quello che ha anche permesso la costruzione del modello elettrodebole e la formulazione dell'ipotesi delle GUT.

**FIGURA XI.1**

Andamento variabile del potenziale al cambiare della temperatura. A temperature  $\geq T_1$  il potenziale presenta un minimo nell'origine. Quando la temperatura scende fino a  $T_2$ , si assiste alla creazione di un secondo estremante (flesso). A temperature comprese tra  $T_2$  e  $T_c$ , tale estremante si suddivide in un massimo e un minimo. Il potenziale ha quindi 2 minimi, ma il minimo assoluto permane nell'origine. Alla temperatura «di transizione»  $T_c$ , i due minimi hanno la stessa altezza. Quando la temperatura scende ulteriormente, il minimo a  $|\phi| \neq 0$  diviene il minimo assoluto, sicché si produce una rottura spontanea di simmetria. Alla temperatura  $T_4 \simeq 0$ , il minimo assoluto è sempre a  $|\phi| \neq 0$ . Nel caso qui illustrato permane ancora un minimo relativo nell'origine; in altri casi, esiste una temperatura sotto la quale nell'origine non c'è più un minimo locale, ma un semplice punto di massimo.

In figura XI.1 diamo un esempio di come può variare  $V(|\phi|)$  al calare di  $T$  dal valore  $T_1$  al valore  $T_4$ , passando per  $T_c$ . Per chiarezza grafica, le curve sono state normalizzate in modo da coincidere nel punto  $\phi = 0$ , mentre l'origine nelle ordinate è stata posta al livello del minimo di  $V(|\phi|)$  alla temperatura più bassa,  $T_4$ . L'ordinata zero, invece, dovrebbe cambiare con  $T$  e trovarsi al livello del minimo assoluto a ciascuna temperatura. L'energia potenziale che il grafico sembra indicare anche in corrispondenza del minimo assoluto, dunque, è solo un artificio grafico.

Come indicato, il potenziale dipende solo dal modulo di  $\phi$ : le forze sono simmetriche per rotazioni di fase. Peraltro, a alte temperature, il potenziale esibisce un solo minimo, per  $\phi = 0$ , un punto che non cambia per rotazioni di fase. Al calare di  $T$  si manifesta una struttura di massimi e minimi e, a temperature  $T < T_c$ , il minimo assoluto si colloca in un punto con  $|\phi| = \bar{\phi} \neq 0$ . La configurazione classica che corrisponde al vuoto quantistico si troverà a  $|\phi| = \bar{\phi}$ , ma avrà

necessariamente *una fase ben determinata*. Una volta fissato lo stato di vuoto, quindi, cessa l'invarianza per rotazioni di fase (rottura spontanea della simmetria, vedi capitolo precedente).

In genere il sistema fisico non è «vuoto». A parte l'energia dei quanti di altri campi, all'energia del vuoto va assommata l'energia cinetica dei quanti di  $\phi$  che, a alte temperature, è assai elevata. Giunti alla temperatura  $T_c$ , il problema è se l'energia cinetica sia sufficiente a superare la *barriera di potenziale* che si frappone tra i due minimi. In genere essa non è affatto adeguata.

Quindi, anche a  $T < T_c$ , il sistema rimane attorno alla configurazione  $\phi = 0$ . Via via che il cosmo si espande, l'energia cinetica continua a diluirsi, in modo  $\propto a^{-3}$  o  $a^{-4}$ , a seconda che i quanti del campo siano massivi o a massa nulla. Oltre all'energia cinetica, il sistema possiede ora l'energia potenziale che corrisponde alla differenza

$$\Delta V = V(0) - V(\bar{\phi}) : \quad \text{potenziale nell'origine} - \text{potenziale del minimo.}$$

Questa energia cresce al calare di  $V(\bar{\phi})$ , per stabilizzarsi alla temperatura  $T_4$ .

L'evoluzione di massimi e minimi è rappresentata ancora in figura XI.2 dove si pone anche in rilievo la necessità che la transizione da  $\phi = 0$  a  $\phi = \bar{\phi}$  sia un processo quantistico di *tunneling*, che avviene dopo un periodo di sovraraffreddamento (*supercooling*).

La densità d'energia  $\rho = \Delta V$  va dunque attribuita al *falso vuoto* in cui il sistema permane in fase di *supercooling*.  $\rho$  dapprima cresce lentamente, poi permane co-

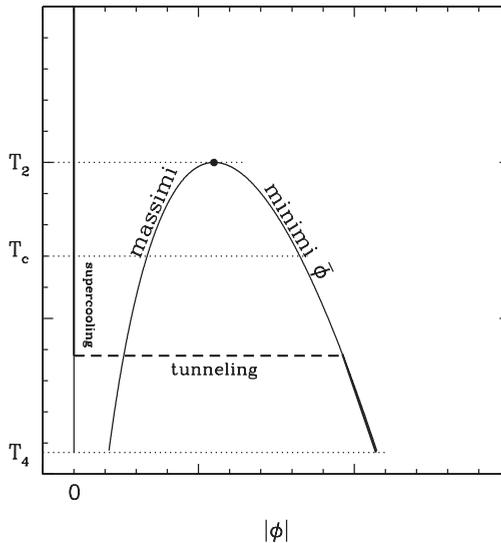


FIGURA XI.2

Andamento con la temperatura dei minimi (e massimi) di  $V(|\phi|)$ . Per  $T > T_2$  esiste solo il minimo nell'origine. Il punto in grassetto, a  $T = T_2$ , è la posizione del flesso. A  $T < T_2$  vi sono due minimi e un massimo. Il minimo per  $|\phi| = \bar{\phi}$  diviene minimo assoluto a  $T < T_c$ . Nel corso del raffreddamento del cosmo, la configurazione fisica di vuoto resta nell'origine anche a  $T < T_c$ , in una fase di *supercooling*, finché un rapido *tunneling* porta il vuoto a posizionarsi in un punto con  $|\phi| = \bar{\phi}$ .

stante, nel corso dell'espansione. L'energia del *falso vuoto* non è quindi soggetta a diluizione, nel corso dell'espansione cosmica che accompagna il calo della temperatura.

La variazione di ogni densità d'energia  $\rho$ , al crescere di  $a$ , è determinata dall'equazione di Friedmann

$$(XI.5) \quad \Delta\rho = -3\Delta a(\rho + p)$$

e, se è  $\Delta\rho = 0$  con  $\Delta a \neq 0$ , dev'essere  $p + \rho = 0$ . Il falso vuoto delle transizioni di fase cosmologiche ha dunque equazione di stato  $p = -\rho$ .

Le sue caratteristiche si riconducono quindi a quelle di una *costante cosmologica*, anche se va segnalato che alcuni relativisti insistono sulla possibilità di una costante cosmologica di natura *geometrica*, svincolata da un'origine fisica.

Nei primi modelli di inflazione, si proponeva che fosse il falso vuoto, dovuto a supercooling, a produrre l'espansione accelerata.

**Campi scalari.** Anche a prescindere dalle transizioni di fase, un campo scalare  $\phi$  omogeneo ha comunque densità d'energia e pressione

$$(XI.6) \quad \begin{aligned} \rho &= \frac{1}{2}\dot{\phi}^2 + V(\phi) = \rho_{kin} + V(\phi), \\ p &= \frac{1}{2}\dot{\phi}^2 - V(\phi) = \rho_{kin} - V(\phi) \end{aligned}$$

e la sua equazione di stato sarà tanto più prossima a  $p = -\rho$ , quanto più trascurabile è il contributo della (densità di) energia cinetica  $\rho_{kin}$ . Per verificare la condizione (XI.2) occorre sia

$$(XI.7) \quad \rho_{kin} \ll V/2.$$

Questa è l'alternativa cui fanno attualmente riferimento i modelli inflazionari, sia che il campo scalare abbia a che fare con rotture di simmetria, sia che l'inflazione si produca a ridosso della scala di Planck.

#### 4. Leggi di espansione e orizzonti

In appendice A1 viene rivisitata la relazione tra il rapporto  $w = p/\rho$  e l'esponente  $\alpha$  in un'espansione a *legge di potenza*,

$$(XI.8) \quad a \propto t^\alpha, \quad \text{trovando che} \quad \alpha = \frac{2}{3(w+1)},$$

finché è  $w > -1$ ;  $w = -1/3$  segna la transizione tra  $\alpha$  minore o maggiore dell'unità. Il parametro di Hubble, per ogni espansione a legge di potenza, è

$$(XI.9) \quad H = \alpha/t,$$

mentre la lunghezza dell'orizzonte ha le espressioni

$$(XI.10) \quad \begin{aligned} L_h &= \frac{\alpha}{1-\alpha} \frac{c}{H} = \frac{ct}{1-\alpha} && \text{per } \alpha < 1 \\ L_h &= \frac{\alpha}{\alpha-1} \frac{c}{H} = \frac{ct}{\alpha-1} && \text{per } \alpha > 1 \end{aligned}$$

(vedi appendice A1). Nei due casi la natura dell'orizzonte è però diversa. Per  $\alpha < 1$  abbiamo un *orizzonte di particella*: esso è la massima distanza da cui possiamo ricevere segnali. Per  $\alpha > 1$  abbiamo un *orizzonte degli eventi*: esso è la massima distanza a cui possiamo inviare segnali. In questo secondo caso, *la distanza tra due punti fisici, che cresce col fattore di scala  $a \propto t^\alpha$ , aumenta più rapidamente della scala dell'orizzonte, che cresce  $\propto t$* . Punti precedentemente in contatto causale sono condotti al di fuori dell'orizzonte reciproco.

Il paradigma inflazionario stabilisce che questo è accaduto in un passato remoto. Poi il ritmo di espansione si è «normalizzato», si è tornati a un'espansione decelerata, e in queste ere successive ciò che era uscito da contatto causale vi rientra. Non deve quindi meravigliare di trovare «omogeneizzate» zone di cosmo apparentemente esterne all'orizzonte.

Ciò che però rende l'inflazione così attraente è che questo risultato è ottenibile con un intervallo brevissimo di espansione accelerata.

Oltre all'espansione accelerata con legge di potenza di tipo (XI.8) con  $\alpha > 1$ , va considerata la possibilità di espansione ancora più rapida, con ritmo esponenziale:

$$(XI.11) \quad a = a_i \exp [H(t - t_i)]$$

( $H$  è il parametro di Hubble nel corso dell'espansione; si assume qui che l'espansione abbia avuto inizio a un tempo  $t_i$ , quando il fattore di scala era  $a_i$ ). Anzi, questa espansione è considerata emblematica dell'inflazione e si verifica nel caso  $w = -1$  (cioè quando è esattamente  $p = -\rho$ ).

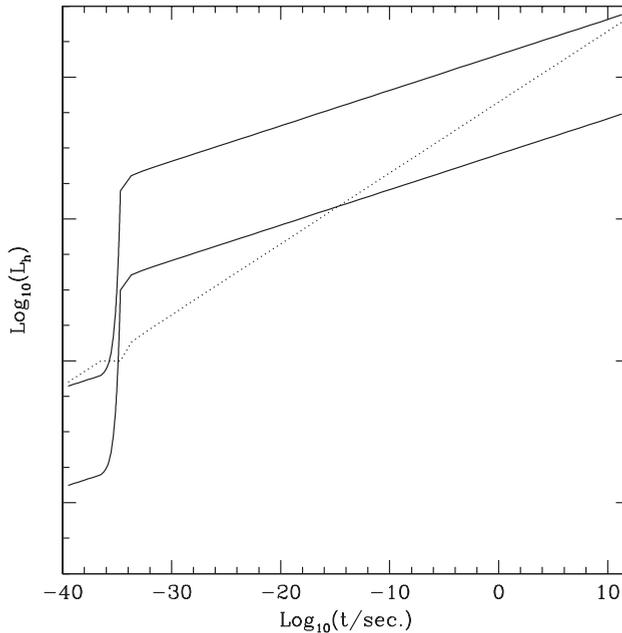
In un'espansione esponenziale l'orizzonte è costante e vale  $c/H$  (appendice A1) mentre, appunto, due punti fisici qualunque si allontanano con velocità crescente in modo esponenziale. Una sorgente di «falso vuoto» dà naturalmente luogo a espansione esponenziale. In appendice A2 si discute sotto quali condizioni anche un campo scalare (*campo inflatonico*) possa dar luogo a espansione (quasi) esponenziale. Il tempo  $1/H$ , in generale, a meno di un fattore  $\sim \mathcal{O}(1)$ , coincide con il tempo iniziale dell'espansione inflazionaria ( $t_i$ ).

Nel caso di espansione esponenziale, la durata di una «buona inflazione» si aggira sui 60-80 *e-folding*. Con ciò s'intende che l'argomento dell'esponenziale, al tempo  $t_e$  quando l'inflazione ha termine, sarà

$$(XI.12) \quad H(t_e - t_i) \simeq H t_e \simeq 60 - 80.$$

Il numero di e-folding

$$(XI.13) \quad N_e \simeq H t_e \sim t_e/t_i$$

**FIGURA XI.3**

Andamento dell'orizzonte (linea tratteggiata). L'inflazione causa una deviazione, brevissima sulla stessa scala logaritmica dei tempi, da un andamento a legge di potenza. Il simultaneo andamento delle distanze tra punti è rappresentato dalle linee continue. Si considerano due coppie di punti, la cui distanza prima e all'inizio dell'inflazione è decisamente inferiore alla scala dell'orizzonte. La crescita esponenziale delle distanze, nell'intervallo temporale in cui l'orizzonte permane costante, porta le distanze fuori dell'orizzonte. I grafici sono compilati per  $N_e = 60$  e-folding e si estendono grosso modo fino all'equivalenza materia-radiazione. La maggiore delle distanze considerate non è ancora rientrata sotto l'orizzonte all'equivalenza, pur non trattandosi della maggior distanza inflazionata.

è quindi, in sostanza, il rapporto tra tempo finale e iniziale dell'inflazione. Se l'inflazione ha inizio, ad esempio, al tempo  $t_i \sim 10^{-40}$  sec., la sua fine può essere avvenuta a  $t_e \sim 10^{-38}$  sec.

Si noti che l'orizzonte  $c/H \simeq ct_i \sim 10^{-30}$  cm, per questi tempi, è al di sotto del raggio d'azione di ogni tipo di forza elementare.

Accanto alla lunghezza fisica dell'orizzonte è significativo considerarne la *lunghezza comoviente*, in pratica la lunghezza riportata alle distanze attuali. In generale sarà

$$(XI.14) \quad L_{h,c}(t) = L_h(t)/a(t)$$

e si può supporre il fattore di scala normalizzato a 1 al tempo attuale ( $a_o = 1$ ).

In questo contesto vanno considerate figura XI.3 e figura XI.4, che descrivono il caso di inflazione esponenziale per 60 e-folding. L'asse logaritmico dei tempi giunge approssimativamente dell'epoca in cui le densità di radiazione e materiali non relativistici

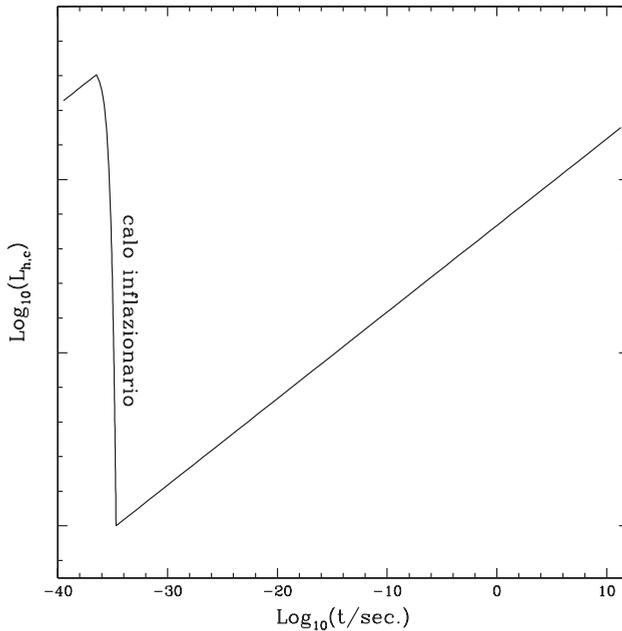


FIGURA XI.4

Andamento dell'orizzonte *comovente*. In questo grafico le distanze comoventi sono rappresentate dalle rette a ordinata costante. Nel corso dell'inflazione, l'orizzonte comovente  $L_{h,c} = L_h/a$  ha un calo esponenziale dovuto alla crescita del fattore di scala. Questo grafico rende chiaro come 60 e-folding portino l'orizzonte comovente a superare, all'inizio dell'inflazione, l'orizzonte all'equivalenza materia-radiazione.

diventano eguali (*equivalenza*). I grafici mettono in evidenza che 60 e-folding sono marginalmente sufficienti a produrre una «buona inflazione», tale che le scale che oggi entrano nell'orizzonte siano state già in contatto causale all'inizio o prima dell'inflazione.

Dal punto di vista analitico si giunge a questo risultato richiedendo che la scala fisica dell'orizzonte all'inizio dell'inflazione sia superiore alla scala fisica dell'orizzonte di oggi. L'ordine di grandezza dei due orizzonti è  $ct_i \simeq c/H$  (inizio inflazione) e  $ct_o$  (oggi). L'espansione della scala  $ct_i$  avviene esponenzialmente durante l'inflazione fino al tempo  $t_e$ . Per l'espansione successiva possiamo non curarci dei dettagli, richiedendo però che non ci siano stati significativi *input* di entropia. Ciò porta alla condizione

$$(XI.15) \quad \begin{array}{ccccccc} ct_i & \times & e^{Ht_e} & \times & a_o/a_e & > & ct_o \\ \text{oriz. iniz.} & & \text{cresc. infl.} & & \text{cresc. postinfl.} & & \text{oriz. oggi} \end{array}$$

che, usando della conservazione dell'entropia tra la fine dell'inflazione, quando l'Universo viene riscaldato fino alla temperatura  $T_{rh}$  dal decadimento del campo inflatonico, e oggi

$$(XI.16) \quad g_{rh} a_e^3 T_{rh}^3 = g_o a_o^3 T_o^3 \quad \text{da cui} \quad a_e/a_o = (g_o/g_{rh})^{1/3} (T_o/T_{rh})$$

( $g_{rh}$  : numero di stati di spin bosonici + (7/8) numero stati di spin fermionici, alla fine dell'inflazione, quando la temperatura è  $T_{rh}$ ;  $g_o$  : idem per oggi, quando la temperatura è  $T_o$ , tenendo opportunamente conto dei neutrini), si ricava

$$(XI.17) \quad e^{N_e} > (t_o/t_i)(g_o/g_{rh})^{1/3}(T_o/T_{rh}) = (t_o/t_p)(T_o/T_p) \times (g_o/g_{rh})^{1/3}(t_p/t_i)(T_p/T_{rh})$$

( $t_p$ ,  $T_p$ : tempo, temperatura di Planck). Per il primo fattore dell'ultimo termine si calcola  $(t_o/t_p)(T_o/T_p) \simeq 10^{28}$ . Il secondo fattore invece dipende dal modello dettagliato di inflazione. Peraltro  $t_p/t_i$  sarà  $\leq 1$ , mentre  $T_p/T_{rh}$  dipende dall'efficienza del meccanismo di *reheating* (vedi appendice A4), conseguente al decadimento del campo scalare che ha guidato l'inflazione, e sarà  $\sim \mathcal{O}(10^4)$  o minore. Valori molto maggiori renderebbero problematica la *bariosintesi* a  $T_{rh}$ . Infine, è certamente  $g_o/g_{rh} < 1$ .

Riassumeremo l'ignoranza sul reheating in un unico parametro, ponendo

$$(XI.18) \quad \frac{T_{rh}}{T_p} \left( \frac{g_{rh}}{g_o} \right)^{1/3} = 10^{-f_{rh}},$$

sapendo comunque che  $f_{rh} \sim \mathcal{O}(4)$ ; deve allora essere

$$(XI.19) \quad e^{N_e} > 10^{28+f_{rh}}(t_p/t_i)$$

Quindi, richiedendo che sia

$$(XI.20) \quad e^{N_e} > 10^{32} \quad \text{ovvero} \quad N_e \gtrsim 70$$

si pone una condizione sufficiente, che potrebbe essere non necessaria. La stima  $N_e \sim 60$  implica un valore alto per  $T_{rh}$  associato, ad esempio, a un tempo  $t_i \sim 10^4 t_p$ . Una stima  $N_e \sim 80$  vuole portare l'orizzonte pre-inflazionario sicuramente al di sopra dell'orizzonte attuale.

## 5. Diluizione

Una conseguenza immediata del meccanismo generico su cui ci siamo basati è l'aspettativa della *diluizione*.

L'inflazione generica è dovuta all'azione di un campo scalare; se, nel corso dell'inflazione, esso sta procedendo dalla configurazione corrispondente al vuoto di alta energia a quella corrispondente al vuoto di bassa energia, può significare che una simmetria viene rotta spontaneamente, come nel meccanismo di Higgs. La configurazione in cui finalmente il campo si assesta potrebbe essere diversa nei diversi orizzonti. Quando gli orizzonti venissero in comunicazione, nel futuro, si assisterebbe quindi a un progressivo riarrangiamento dello stato di vuoto. Il processo dovrebbe continuare fino a oggi; ma ciò che è peggio, su volumi molto maggiori dell'orizzonte alla transizione, potrebbero instaurarsi dei *pattern* di rottura non riconducibili a un unico vuoto. Un esempio è dato in figura XI.5, per il caso della rottura di  $U(1)$ .

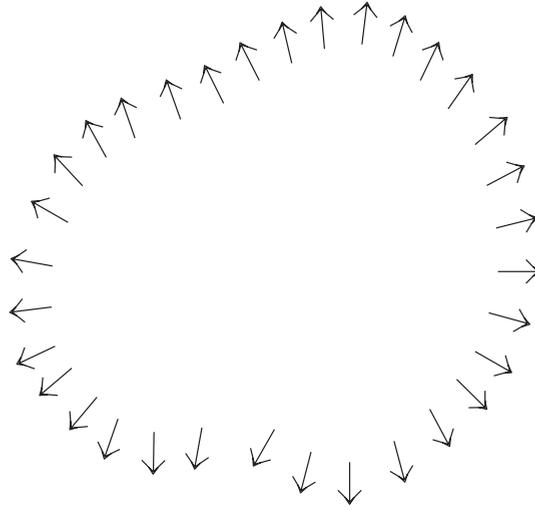


FIGURA XI.5

Esempio di direzioni di rottura spontanea di una simmetria  $U(1)$  in orizzonti inizialmente disgiunti. Quando gli orizzonti venissero in contatto causale, si rende evidente la presenza di una rotazione di  $2\pi$  dell'angolo di rottura lungo il pattern chiuso evidenziato nel disegno. Non esiste quindi un'unica direzione su cui tutte le direzioni di rottura possano gradualmente convergere.

In questo caso, al centro del pattern dovrebbe necessariamente permanere una zona in cui la transizione non è avvenuta.

La topologia diviene più complessa nei casi di gruppi meno semplici. In generale, la messa a contatto di vari orizzonti può generare 3 tipi di strutture in cui la transizione non è avvenuta. Se le strutture sono bidimensionali avremo *muri di dominio*, nel caso unidimensionale avremo *stringhe* e, infine, si possono avere *monopoli* 0-dimensionali.

Ovviamente, se l'orizzonte attuale è tutto contenuto nell'orizzonte pre-inflazione il problema non si pone. Ma, se nell'orizzonte pre-inflazione erano contenuti  $N$  mostri topologici derivanti da precedenti transizioni, essi possono essere oggi ricollocati su un volume molto maggiore dell'attuale orizzonte (ad esempio se  $N_e \sim 80$ ), per cui nel nostro orizzonte possono non essercene o, al più, essercene uno solo.

L'inflazione, quindi, favorisce la sparizione o la diluizione dei *mostri topologici*.

Alla fine dell'inflazione e dopo il *reheating*, una parte sostanziale della densità d'energia del campo scalare  $\rho_\phi \sim (3/8\pi)m_p^2 H^2$  si trasforma in energia *termica*, dando luogo a una densità d'entropia

$$(XI.21) \quad \sigma_{rh} \sim \rho_\phi^{3/4} \sim (m_p H)^{3/2}$$

non molto inferiore alla densità d'entropia pre-inflazionaria. La stessa densità si trova in un volume che si è espanso di un fattore  $\exp(3N_e) \sim 10^{95} \sim (e^{80})^3$ . Se, prima dell'inflazione, un determinato tipo di particella aveva una densità in numero  $n_p$ , il rapporto  $n_p/\sigma$  è calato di poco meno di  $10^{95}$ .

Ad esempio, se si pensasse che il *numero barionico* esiste da prima dell'inflazione, per avere oggi un rapporto  $n_b/\sigma \sim 10^{-8}$ , si sarebbe dovuti partire da un rapporto  $n_b/\sigma \sim 10^{85}$ , che non ha senso fisico, specialmente a temperature e  $\sigma$  elevate. Lo stesso meccanismo di diluizione risolve i problemi di sovrabbondanza dei *gravitini* e dei campi di *modulo*.

## 6. Piattezza e parametro di densità

Vedremo ora come  $\sim 60$ -80 e-folding avvicino a soluzione anche il *problema della piattezza*, dando un senso al dato che il parametro di densità ( $\Omega_o$ ) è oggi prossimo all'unità.

Riprendiamo brevemente alcuni sviluppi già discussi nel capitolo IV, ripartendo dall'equazione di Friedmann e dalla definizione di *densità critica* ( $\rho_{cr}$ ), che si possono associare nelle equazioni

$$(XI.22) \quad H^2 \equiv \frac{8\pi}{3m_p^2} \rho_{cr} = \frac{8\pi}{3m_p^2} \rho - \frac{\kappa}{a^2},$$

dove  $m_p = G^{-1/2}$  è la *massa di Planck* ( $\hbar = c = 1$ ) e  $\rho$  è la densità del cosmo. Dividendo secondo e terzo membro della relazione (XI.22) per  $8\pi\rho/3m_p^2$ , otteniamo

$$(XI.23) \quad 1 - \Omega^{-1} = \frac{\kappa}{(8\pi/3m_p^2)\rho a^2}.$$

dove  $\Omega = \rho/\rho_{cr}$  è il *parametro di densità* a un tempo generico ( $\Omega_o$  è il parametro di densità oggi). Se la curvatura  $\kappa$  è nulla, risulta  $\Omega \equiv 1$ : in questo caso la sezione spaziale della metrica è sempre e rigorosamente piatta. Se invece è  $\Omega \neq 1$ , il valore di  $\Omega$  sarebbe costante solo se fosse costante il prodotto  $\rho a^2$ .

Useremo la quantità  $|1 - \Omega^{-1}|$  come indicatore della distanza del parametro di densità dall'unità. L'eq. (XI.23) mostra allora che, se è  $\rho \propto a^{-s}$  e  $s > 2$ ,  $\Omega$  tende a allontanarsi dall'unità al crescere di  $a$ . Questa è la situazione durante un'espansione dominata dalla radiazione o dalla materia non relativistica, come ampiamente discusso nel capitolo IV.

In un'espansione esponenziale, invece, in cui  $\rho$  permane costante, da qualsiasi valore iniziale parta,  $\Omega$  tende a avvicinarsi all'unità.

Nella storia di ogni modello cosmologico con inflazione si succedono quindi due fasi.

- 1) Espansione inflazionaria:  $\Omega$ , partendo da un valore «iniziale» generico, si avvicina sempre più all'unità.
- 2) Espansione radiativa (o dominata dalla materia):  $\Omega$  riprende a allontanarsi dall'unità.

Ad esempio, in una inflazione esponenziale che duri per 80 e-folding, il fattore di scala cresce di  $e^{80} \sim 10^{35}$ , sicché la differenza di  $\Omega$  dall'unità viene ridotta da un drastico fattore  $\sim 10^{-70}$ . In questo modo, anzitutto, un valore iniziale generico di  $\Omega$  di-

viene assai prossimo all'unità alla fine dell'inflazione. Ancora più importante: anche nella fase 2), che pure dura un tempo enormemente più lungo dell'inflazione, ma in cui la variazione del fattore di scala è assai meno rapida, difficilmente  $\Omega$  si allontana dall'unità in modo significativo.

Vediamo tutto ciò in termini quantitativi: Per  $N_e$  e-folding, usando eq. (XI.23) con  $\rho \equiv \text{const.}$ , vediamo che risulta

$$(XI.24) \quad 1 - \Omega_e^{-1} = [1 - \Omega_i^{-1}] (a_i/a_e)^2 = [1 - \Omega_i^{-1}] e^{-2N_e}.$$

(il suffisso  $_e$  si riferisce alla fine dell'inflazione). Assumiamo per semplicità che tutta l'espansione successiva sia *radiativa* ( $\rho a^2 \propto a^{-2}$ ). Oggi (suffisso  $_o$ ) sarà

$$(XI.25) \quad 1 - \Omega_o^{-1} = [1 - \Omega_e^{-1}] (a_o/a_e)^2 = [1 - \Omega_i^{-1}] e^{-2N_e} (a_o/a_e)^2.$$

Bisogna dunque valutare il fattore  $(a_o/a_e)^2 \exp(-2N_e)$  per sapere quale *range* di valori iniziali  $\Omega_i$  sia mappato in un intorno di  $\Omega_o = 1$ , opportunamente piccolo. Ciò che eq. (XI.25) rende evidente è che tale mappa è tanto più concentrata, quanto minore è il rapporto  $a_o/a_e$ , cioè tanto minore è la temperatura  $T_{rh}$ .

Parametrizzando tale temperatura tramite  $f_{rh}$ , come in eq. (XI.19), avremo allora che la discrepanza da 1 del parametro di densità pre-inflazione viene ridotta di un fattore

$$(XI.26) \quad f_\Omega \equiv \frac{1 - \Omega_o^{-1}}{1 - \Omega_i^{-1}} = e^{-2N_e} 10^{64 - 2f_{rh}} = \exp[-2N_e + (64 - 2f_{rh}) \ln 10]$$

che è  $< 1$  non appena  $N_e > 74 - 2.3 \times f_{rh}$ . Ad esempio, per  $N_e \simeq 60$ , sarà  $f_\Omega < 1$  solo per  $f_{rh} > 6$ . In generale, quindi, è *più facile approssimare la piattezza in presenza di un cattivo reheating*. Si tratta di un'opzione su cui non si può molto elaborare, a pena di entrare in conflitto con il problema della *bariosintesi*.

Non va però dimenticato che, in assenza di inflazione, perché oggi sia  $\Omega_o \sim \mathcal{O}(1-10^{-1})$ , occorre un *fine tuning* di  $\sim 1 : 10^{60}$  al tempo di Planck. Un periodo inflazionario che risolva il problema dell'omogeneità elimina del tutto questi numeri paradossali riportando, alla peggio, il valore attuale di  $\Omega_o$  a una buona scelta del valore al tempo di Planck.

*Valori di  $\Omega_o$  prossimi all'unità, anche se non  $\Omega_o \equiv 1$  (ad esempio  $\Omega_o = 0.1$ ), sono dunque previsti da un modello generico di inflazione.*

*Per  $N_e > 72-73$  cominciano a diventare ragionevoli valori di  $\Omega_o$  assai prossimi all'unità.*

*Per  $N_e \gtrsim 80-100$  si può affermare che, partendo da una densità generica, saranno comunque raggiunti valori  $\Omega_o \sim 1$ .*

È infine opportuno rammentare che, anche se il meccanismo inflazionario porta da una generica densità a  $\Omega_o \simeq 1$ , non per questo la geometria dello spazio diviene piatta. È invece vero che il suo raggio di curvatura può divenire così grande da essere osservativamente indistinguibile da  $\infty$ .

## 7. Le fasi dell'inflazione e i modelli d'inflazione

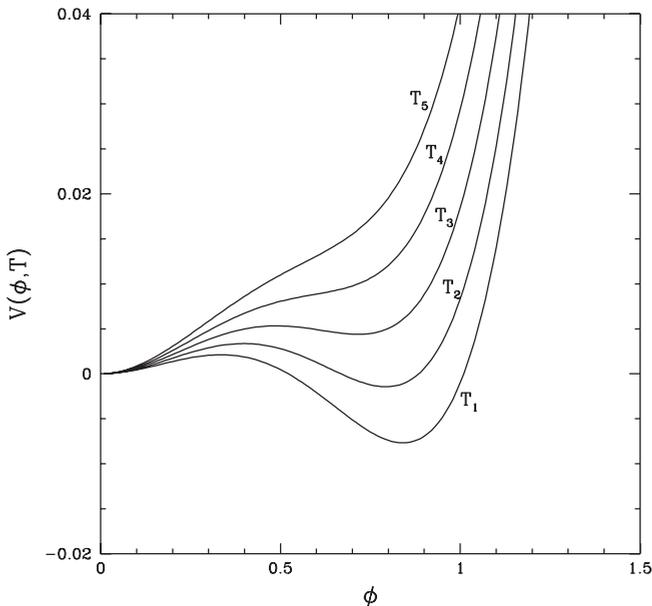
Seguendo approssimativamente la linea dello sviluppo storico, possiamo suddividere i modelli inflazionari in tre tipi:

- 1) a piccolo campo;
- 2) a grande campo;
- 3) ibridi.

### **Modelli a piccolo campo**

Il primo modello a piccolo campo proposto si basava su un potenziale alla *Coleman-Weinberg*, riportato in figura XI.6, quale nasce dalla regolarizzazione e rinormalizzazione dell'interazione dovuta a un campo di Higgs. Ne risultava un quadro assai interessante: il campo inflatonico era lo stesso campo che portava alla rottura della GUT. La successione degli andamenti in figura XI.6 segue il calo delle temperature, dalla più elevata ( $T_5$ ) a quella praticamente nulla ( $T_1$ ).

Purtroppo i valori dei parametri necessari alla rottura della GUT non coincidono con quelli richiesti per una buona inflazione. Un'inflazione dovuta a un potenziale con l'andamento raggiunto alla temperatura  $T_1$  resta una possibilità. Ma il campo  $\phi$  non è il



**FIGURA XI.6**

Andamento del potenziale di *Coleman-Weinberg* al variare della temperatura; è manifesta la simiglianza qualitativa tra questi andamenti e quelli mostrati in figura XI.1.

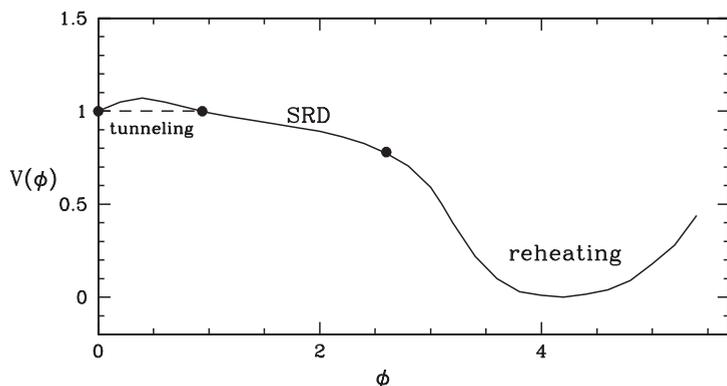


FIGURA XI.7

Andamento del potenziale in un modello di inflazione a campo «piccolo». Partendo dall'origine, il campo  $\phi$  1) «passa» per effetto tunnel sotto la barriera di potenziale, 2) compie la fase di lento rotolamento (SRD), che è la vera fase inflazionaria, 3) inizia una discesa più rapida seguita da oscillazioni attorno al minimo di potenziale, riscaldando con la propria energia alla componente radiativa (*reheating*).

campo di Higgs. Conviene allora schematizzare il potenziale come in figura XI.7, così da porre ben in evidenza le diverse fasi dell'inflazione:

- 1) Tunneling.
- 2) SRD, (*slow-rolling-down*: lento rotolamento verso il basso).
- 3) Reheating.

Ciò che differenzia tra loro questi modelli è il tunneling, che dà inizio all'inflazione. Le fasi successive sono sostanzialmente comuni a tutti i tipi di inflazione.

L'inflazione a piccolo campo inizia col raffreddamento del cosmo. La forma del potenziale diviene così  $V(\phi, T_1)$  (fig. XI.6), ma il campo è rimasto nel minimo relativo a  $\phi = 0$ , nonostante il minimo assoluto a  $\phi = \bar{\phi}$ . Da questa configurazione metastabile esso uscirà solo grazie a un effetto quantistico. Data l'espressione analitica del potenziale, è possibile calcolare il tempo medio necessario perché il tunneling si produca. In un buon modello di inflazione, tale tempo deve essere al più confrontabile col tempo necessario a raggiungere la temperatura  $T_1$ .

Anche nella configurazione di falso vuoto, il campo  $\phi$  produce inflazione. Va notato che l'ampiezza delle fluttuazioni di densità che si ingenerano in tale configurazione è peculiare. Questo fenomeno è stato sottolineato quando si riteneva che i dati osservativi propendessero per  $\Omega_o \sim 0.2-0.3$ .

Lo studio dei modelli a piccolo campo segue una transizione di fase mentre essa si verifica. In laboratorio, una transizione di fase del primo ordine, che inizi dopo un periodo di *supercooling*, come quello che trascorre qui in attesa del *tunneling*, è un evento esplosivo, che si propaga dal punto in cui nuclea una prima bolla, tramite un fronte d'onda che può procedere a velocità subsonica (detonazione) o supersonica

(deflagrazione). Altre bolle nucleano poi in altri siti, i fronti d'onda collidono, si genera un vasto ammontare di entropia e, alla fine, si recupera una quasi omogeneità nella nuova fase, con una temperatura risalita a un valore prossimo alla temperatura di transizione  $T_c$ .

Questi fenomeni sono attesi anche in un'inflazione a piccolo campo. Essi però avvengono in ambito cosmologico, in cui energia e pressione della bolla sono sorgente dell'espansione cosmica. Inoltre i tempi cosmologici sono brevi e non vale quindi più l'approssimazione di considerare istantaneo il fenomeno di nucleazione. Ciò porta a modifiche sostanziali nel quadro.

Se vogliamo restare aderenti alla descrizione «a bolle», potremo dire che tutto l'orizzonte oggi osservato e anche oltre saranno compresi in una singola bolla. Se poi ci confrontiamo con figura XI.2, troviamo anche lì indicato il tunneling che si conclude col campo che raggiunge direttamente il valore  $\bar{\phi}$  («vero vuoto»). Un raffronto tra figura XI.2 e figura XI.7 deve però tener conto che la prima è solo una descrizione *statica*, che pone in rilievo solo le configurazioni di equilibrio, la seconda una descrizione *dinamica*, che segue il campo  $\phi$  nella sua evoluzione dall'uno all'altro minimo. Questo diventa il periodo essenziale dell'inflazione.

Una volta passato sotto la barriera di potenziale, il campo inizia una fase di lento rotolamento (SRD) lungo la discesa del potenziale. Lo SRD si conclude quando il progressivo aumento di pendenza provoca la fuoriuscita dal ritmo stazionario di crescita di  $\phi$ . Non c'è ragione per cui le lunghezze dei diversi tratti del grafico in figura XI.7 debbano corrispondere ai tempi impiegati a percorrerli. Lo SRD, in genere, è molto lento. Il reheating, invece, avviene con veloci oscillazioni attorno al minimo e, specie se il modello dà luogo a un *buon reheating*, si conclude rapidamente. Saranno SRD e reheating a portare  $\phi$  dal valore nullo iniziale a un valore grande. Questa classe di modelli, che partono da valori piccoli di  $\phi$ , è detta «a piccolo campo». In realtà, i valori tipici di  $\phi$ , nel corso del processo, in questo e in altri tipi di modelli, sono comunque  $\sim m_p$  (massa di Planck). Invece che di «piccolo campo» sarebbe forse più corretto parlare di «campo crescente».

Perché l'inflazione abbia luogo, il campo  $\phi$  deve avere un'energia cinetica  $\rho_{kin} \ll V(\phi)$ . Così, da un lato  $V(\phi)$  (praticamente) esaurisce la densità d'energia, dall'altro  $\phi$  è (quasi) costante e quindi (approssimativamente) lo è  $V(\phi)$  stesso. Ne segue che il parametro di Hubble

$$(XI.27) \quad H^2 = \frac{8\pi V}{3m_p^2}$$

è (quasi) costante e l'espansione del cosmo è (quasi) esponenziale:

$$(XI.28) \quad a(t) = a_i \exp \left[ \frac{\sqrt{8\pi V/3}}{m_p} (t - t_i) \right].$$

Per una trattazione analitica più dettagliata di questa fase si veda l'appendice A2. Essa mostra come il potenziale possa indurre il «moto» di  $\phi$ , concedendogli un minimo di

energia cinetica, entro limiti che garantiscono tuttavia la validità di eq. (XI.27) e (XI.28). La pendenza del potenziale viene espressa dalla sua «derivata prima»  $V'(\phi)$ , la sua curvatura dalla sua «derivata seconda»  $V''(\phi)$ . Queste derivate entrano nella definizione dei parametri di *pendenza* e di *curvatura*:

$$(XI.29) \quad \varepsilon = \frac{m_p^2}{16\pi} \left( \frac{V'}{V} \right)^2, \quad \eta = \frac{m_p^2}{8\pi} \frac{V''}{V}.$$

In appendice A2 si dimostra che il parametro di pendenza esprime il rapporto energia cinetica/energia potenziale,

$$(XI.30) \quad \varepsilon = 3\rho_{kin}/V,$$

sicché i requisiti cui deve sottostare il potenziale sono:

- 1)  $\varepsilon \ll 1$ ; piccola energia cinetica rispetto all'energia potenziale;
- 2)  $\eta \ll 1$ ; piccola accelerazione nei tempi fissati dal parametro di Hubble.

Su questo secondo punto si veda ancora appendice A2, dove sono anche date ulteriori espressioni di  $\varepsilon$  e  $\eta$ , nonché un esempio analitico di potenziale per inflazione a piccolo campo.

I parametri  $\varepsilon$  e  $\eta$  sono anche connessi allo spettro primordiale di fluttuazioni. Della loro genesi e del reheating, nei vari modelli di inflazione, tratteremo brevemente in un prossimo paragrafo e nelle Appendici A4 e A5.

### **Modelli a grande campo**

I modelli a grande campo nascono con la proposta dell'*inflazione caotica*. In questo caso il potenziale del campo  $\phi$  è della forma

$$(XI.31) \quad V_{(4)} = \lambda\phi^4 \quad \text{ma anche, semplicemente,} \quad V_{(2)} = \frac{m^2}{2}\phi^2,$$

o comunque una somma di più termini contenenti potenze diverse di  $\phi$ .

L'andamento di  $V_{(2)}$  è illustrato in figura XI.8. Il minimo dei potenziali (XI.31) è unico e si trova comunque a  $\phi = 0$ .

Nell'inflazione caotica, il fatto che  $\phi$  non si trovi «inizialmente» nel punto di minimo non ha più nulla a che fare con le transizioni di fase. Il punto è invece che, all'emergere dal mondo pre-planckiano, le *relazioni di indeterminazione* non permettono di fissare  $\phi$  a valori inferiori a  $m_p$ .

Volendo esplicitamente ricavare il valore di  $\phi$  dalle relazioni di indeterminazione, si parte dalla condizione  $\Delta\varepsilon\Delta t \sim \hbar$ , dove l'energia  $\Delta\varepsilon$  si esprime come  $\rho(c\Delta t)^3$ . La densità d'energia  $\rho$  ha le componenti cinetica e potenziale

$$\rho_{kin} = \dot{\phi}^2/2 \simeq (\phi/\Delta t)^2, \quad \rho_{pot} = V(\phi)$$

$\Delta t$  è chiaramente il tempo di Planck, entro cui andrebbe operata la «misurazione» di  $\varepsilon$ . La componente potenziale dell'energia dipende dal modello.

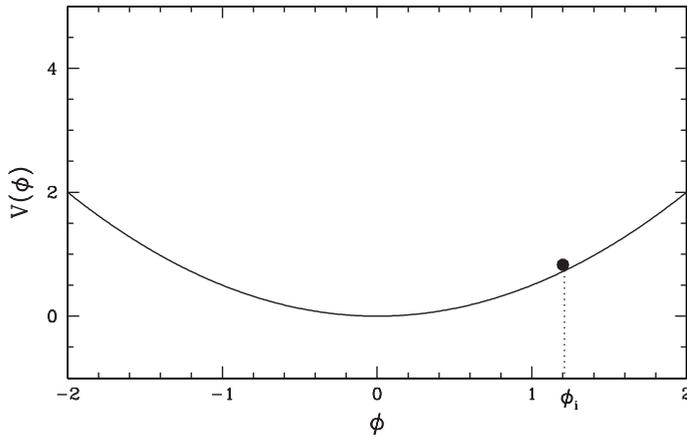


FIGURA XI.8

Andamento del potenziale in un modello di inflazione a campo «piccolo». Il valore «iniziale» del campo, quello con cui esso emerge dall'era pre-planckiana, è indicato da una pallina. Successivamente il campo tende a evolvere verso la configurazione di equilibrio  $\phi = 0$ .

Ad esempio, per  $V_{(4)}(\phi) = \lambda\phi^4$ , le relazioni di indeterminazione daranno

$$\lambda(\phi t_p)^4 \sim \hbar/c^3 \quad \text{e, quindi } \lambda\phi^4 \sim m_p^4 c^5/\hbar^3.$$

In unità naturali, ciò significa  $\phi \sim m_p \lambda^{-1/4}$ . In questi modelli, per ragioni concernenti l'ampiezza delle fluttuazioni generate dall'inflazione, si richiede che  $\lambda$  sia, di svariati ordini di grandezza, al di sotto dell'unità. Di conseguenza, la richiesta  $\phi \sim m_p$  è necessaria, ma potrebbe essere ampiamente insufficiente.

In modo analogo, per  $V_{(2)}(\phi) = m^2\phi^2/2$  si trova, in unità naturali,  $\phi \sim m^2/m_p$  e, a meno dell'irragionevole caso  $m \gg m_p$ , si giunge ancora a  $\phi \sim m_p$ .

Se si considera poi il termine cinetico, la richiesta  $(\phi/t_p)^2 \sim 1$  (in unità naturali), porta direttamente alla condizione  $\phi \sim m_p$ .

In definitiva, la componente cinetica garantisce che debba essere almeno  $\phi \sim m_p$ , mentre i potenziali considerati danno, come tipici, valori iniziali maggiori di  $m_p$  per un fattore anche superiore a 10.

La scelta del punto segnato in grassetto in figura XI.8 è dunque genericamente restrittiva; partendo da quel punto, il campo evolve verso  $\phi = 0$  e, nei modelli «caotici», detti «a campo grande» sulla base dei valori assunti inizialmente da  $\phi$ , l'inflazione è dovuta a questa necessaria evoluzione.

Il caso di  $V_{(2)}$  è interessante anzitutto perché un potenziale quadratico non descrive forze, sicché abbiamo semplicemente un campo libero con quanti massivi. Inoltre esso permette l'integrazione analitica (vedi appendice A3). In particolare, l'andamento del campo, a partire dal tempo  $t_i$  quando il campo vale  $\phi_i$ , può essere calcolato ed è

$$(XI.32) \quad \phi = \phi_i - \frac{m m_p}{\sqrt{12} \pi} (t - t_i),$$

mentre il fattore di scala ha l'andamento

$$(XI.33) \quad a = a_i \exp \left\{ \left( \frac{4\pi}{3} \right)^{1/2} \frac{m}{m_p} \left[ \frac{\phi_i}{m_p} \frac{t - t_i}{t_p} - \frac{1}{\sqrt{48\pi}} \frac{m}{m_p} \left( \frac{t - t_i}{t_p} \right)^2 \right] \right\}.$$

Secondo eq. (XI.32), il campo  $\phi$  conclude la sua discesa raggiungendo il minimo in  $\phi = 0$  al tempo  $t_o = t_i + \sqrt{12\pi} t_p (\phi_i/m)$ . Il secondo addendo in parentesi quadra in eq. (XI.33) vale allora  $(\phi_i/2m_p)(t_o - t_i)/t_p$  e rappresenta ancora solo una correzione del 50% rispetto al primo termine. Se restiamo a valori  $\phi$  ragionevolmente al di sopra dello 0, dunque, l'andamento di  $a(t)$  può essere ben approssimato dall'espressione

$$(XI.34) \quad a \simeq a_i \exp \left[ \left( \frac{4\pi}{3} \right)^{1/2} \frac{m}{m_p} \phi_i (t - t_i) \right].$$

quale si otterrebbe assumendo  $a = a_i \exp[H(t - t_i)]$  e attribuendo a  $H$  il suo valore iniziale  $H_i = (8\pi/3m_p^2)^{1/2} V_{(2)}(\phi_i) = (4\pi/3)^{1/2} \phi_i m/m_p$ , che rimarrebbe quindi costante nel tempo. Il numero di e-folding al tempo  $t_e$  risulta quindi

$$(XI.35) \quad N_e \simeq \left( \frac{4\pi}{3} \right)^{1/2} \frac{m}{m_p} \phi_i t_e$$

e si possono anche ricavare valori esatti di

$$(XI.36) \quad \varepsilon = \eta = \frac{1}{4\pi} \left( \frac{m_p}{\phi} \right)^2$$

(appendice A3). Assumiamo che lo SRD termini se diviene  $\varepsilon, \eta \simeq 0.1$ , cioè se si ha  $\phi_e/m_p \simeq \sqrt{4\pi/10}$ . Questo valore di  $\phi_e$  si può sostituire in eq. (XI.32) ricavando

$$(XI.37) \quad t_e \simeq (\sqrt{3}/m)(2\sqrt{\pi} \phi_i/m_p - 1).$$

Sostituito questo valore di  $t_e$  in eq. (XI.35) si ricava

$$(XI.38) \quad N_e \simeq \sqrt{4\pi} \frac{\phi_i}{m_p} \left( \sqrt{4\pi} \frac{\phi_i}{m_p} - 1 \right) \simeq 4\pi \left( \frac{\phi_i}{m_p} \right)^2.$$

Per avere  $N_e \sim 60-80$  occorre quindi che sia  $\phi_i/m_p \sim 3-4$ . Usando allora il valore  $\phi_i \simeq 4m_p$  in eq. (XI.37), sarà

$$(XI.39) \quad t_e \simeq \frac{1}{m} \left( 4\sqrt{12\pi} - \sqrt{3} \right)$$

e quindi, sulla base di eq. (XI.32),

$$(XI.40) \quad \phi_e \simeq 4m_p - \frac{m_p}{\sqrt{12\pi}} \left( 4\sqrt{12\pi} - \sqrt{3} \right) = \frac{m_p}{2\sqrt{\pi}}$$

indipendentemente dal valore di  $N_e$ .

COMMENTO. Nella discussione dell'inflazione, ci troviamo di fronte al trasferimento dell'indeterminazione quantistica su proprietà di grande scala. Le relazioni di indeterminazione richiedono sia «almeno»  $\phi_i \sim m_p$ ; per avere la distribuzione statistica dei possibili valori iniziali di  $\phi$  occorrerebbe, in realtà, la *quantum gravity*. Anche le considerazioni basate sull'espressione dei potenziali debbono essere prese con molta cautela e dobbiamo riconoscere di non sapere se un valore  $\phi_i \simeq 3 m_p$  sia «probabile». Possiamo però affermare con sicurezza che esso rientra tra i valori «possibili». Nell'ambito di questo modello, quindi, tutti gli orizzonti planckiani, di dimensioni  $\sim ct_p$ , in cui sia  $\phi_i > 3-4 m_p$ , sono destinati a ampliarsi fino a dimensioni superiori a quelle del nostro attuale orizzonte.

Potremmo dire che la nostra esistenza implica una misura del valore iniziale, nel nostro orizzonte, di  $\phi_i (> 3-4 m_p)$ . Si tratta ancora di una blanda applicazione del «principio antropico».

Va anche aggiunto che la stessa applicazione della metrica FRW e delle equazioni di Friedmann implica ancora il principio antropico. Tra gli  $\infty$  orizzonti a cui tali equazioni non possono forse essere applicate, in carenza di omogeneità e isotropia, seppure su piccolissima scala, il principio antropico seleziona quelli adeguatamente isotropi e omogenei, da magnificare fino a includere tutto l'attuale cosmo osservabile.

### Modelli ibridi

Il modello basato sul potenziale  $V_{(2)}$  non ha un buon reheating: le condizioni  $\varepsilon, \eta \ll 1$ , necessarie allo SRD, vengono violate non perché cresce l'energia cinetica, ma perché il potenziale si approssima allo zero (vedi Fig. XI.8 e eq. XI.30).

Al fine di produrre un valido reheating con potenziali di questi tipo, venne proposto un potenziale dipendente da due campi  $\phi$  e  $\chi$ , quale quello in figura XI.9.

Una forma analitica che dà il comportamento in figura è la seguente:

$$(XI.41) \quad V(\phi, \chi) = \frac{m^2}{2} \phi^2 + V_o(\phi, \chi).$$

Il primo addendo è simile a  $V_{(2)}$ . Se il secondo addendo fosse costante, il grafico del potenziale coinciderebbe con quello in figura XI.8, salvo uno spostamento dell'origine delle ordinate verso il basso, di una quantità pari a  $V_o$ . Altrimenti  $V_o$  può dipendere da un secondo campo  $\chi$ , ma in modo diverso, a seconda dei valori di  $\phi$ . Ad esempio, per  $\phi$  grande, il minimo di  $V_o$  rispetto a  $\chi$  può trovarsi in  $\chi = 0$ ; invece, quando  $\phi$  scende sotto un valore critico  $\phi_c$ ,  $\chi = 0$  può diventare un massimo, mentre di producono due minimi in  $\pm \mu$ .

Ciò avviene, ad esempio, se

$$(XI.42) \quad V_o(\phi, \chi) = \lambda_1 [\chi^2 - \mu^2(\phi)]^2 \quad \text{e} \quad \mu^2(\phi) = \lambda_2 (\phi_c^2 - \phi^2)$$

Si noti che il segno di  $\mu^2$  cambia quando il valore di  $\phi$  passa per  $\phi_c$ . Per  $\mu^2 > 0$ ,  $V_o$  è sempre positivo, tranne che per  $\chi = \pm \mu$ , dove si annulla. Questi punti sono quindi dei

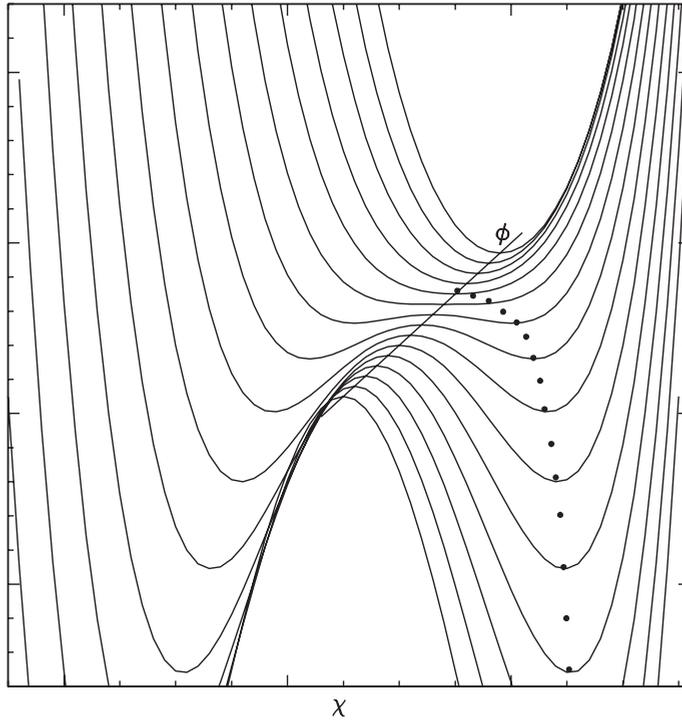


FIGURA XI.9

Andamento del potenziale in un modello ibrido. Quando  $\phi$  è ancora lontano dalla zona dei minimi assoluti, il minimo rispetto a  $\chi$  si trova in  $\chi = 0$  (situazione descritta al fondo della figura). Quando  $\phi$  scende verso valori minori, lungo la linea continua, si producono due minimi simmetrici sull'asse  $\chi$ ; uno di essi è indicato dalla linea punteggiata, al variare di  $\phi$ . Quando il minimo rispetto a  $\chi$  si sdoppia, il sistema abbandona la linea del dosso. Per opportune scelte di parametri, si possono innescare oscillazioni smorzate sui fianchi del «canalone» marcato dalla punteggiata, dando origine a *reheating*; oppure può continuare lo SRD, con  $\chi$  che gradualmente sostituisce  $\phi$  come campo inflatonico.

minimi. Per  $\mu^2 < 0$ , invece, il valore minimo di  $V_0$  si ha quando la quantità in parentesi quadra è minima, cioè per  $\chi = 0$ .

Nell'espressione di  $V_0$  entrano dunque 3 parametri:  $\lambda_1$ ,  $\lambda_2$ ,  $\phi_c$ . L'evoluzione nel piano  $\phi$ - $\chi$  dipende da questi parametri. È possibile che la discesa «iniziale» di  $\phi$  lungo l'asse  $\chi = 0$  implichi un'elevata «inerzia». Quando l'asse  $\chi = 0$  non è più il luogo dei minimi rispetto a  $\chi$ , il moto non s'indirizzerà allora, immediatamente, lungo la curva tratteggiata (vedi Fig. XI.9), ma proseguirà con oscillazioni da un lato all'altro del canalone, il cui fondo è marcato dai puntini. Queste oscillazioni possono essere responsabili del reheating.

Questo comportamento è tipico di una situazione in cui l'energia cinetica è già elevata, prima della «biforcazione» dei minimi. Essa si verifica quindi se, ancor prima della biforcazione, la fase di SRD era prossima al termine o finita.

Altrimenti, giunto al punto di biforcazione, il sistema semplicemente «sceglie» uno dei due canali; la discesa prosegue lungo la curva dei minimi e il ruolo di campo inflatonico è gradualmente preso dal campo  $\chi$ .

Rispetto al campo  $\phi$  la dinamica è simile a quella di un modello «a grande campo». Viceversa, l'evoluzione del potenziale rispetto a  $\chi$  riproduce gli effetti della variazione della temperatura in un modello «a piccolo campo». In questo caso, peraltro, ciò che causa l'evoluzione dell'andamento del potenziale in  $\chi$  è la variazione di  $\phi$ . Alla fine i valori di  $\chi$  tenderanno comunque a allontanarsi dallo zero, come nei modelli «a piccolo campo». Del tutto legittimo, per questi modelli, è quindi l'appellativo di «modelli ibridi».

## 8. Il reheating

L'inflazione, in genere, ha inizio a una temperatura  $T_i = f_i T_p$ . Un'inflazione a grande campo o ibrida, partendo a ridosso della scala di Planck, sarà in genere caratterizzata da  $f_i \sim \mathcal{O}(1)$ ; un'inflazione a piccolo campo, invece, legata alla transizione GUT, sarà in genere caratterizzata da  $f_i \sim 10^{-3}$ - $10^{-5}$ .

Durante lo SRD l'energia cosmica prevalente deriva dalla (densità di) energia potenziale:  $\rho_{SRD} \simeq V(\phi)$ . Ciò è conseguenza dell'equazione di Friedmann che, in assenza di scambi energetici, ci dà la (densità di) potenza che ciascuna componente dissipa a causa dell'espansione cosmica:

$$(XI.43) \quad W = -3H(\rho + p),$$

$p$  e  $\rho$  sono pressione e densità d'energia della componente. L'eq. (XI.43) dice che le componenti più «resistenti» sono quelle per cui  $\rho + p$  è nullo o prossimo a annullarsi. Un eventuale residuo termico, ad esempio, dissipa rapidamente la sua energia; infatti, con  $p = \rho/3$ , risulta  $W \propto \rho$ ; ne consegue che  $\rho \propto T^4 \propto a^{-4}$  e, per il residuo termico, la temperatura  $T$  ( $\propto a^{-1}$ ) cala (quasi) esponenzialmente. Un campo inflatonico, invece, con  $\rho = \rho_{kin} + V$  e  $p = \rho_{kin} - V$ , a causa dell'espansione dissipa una potenza

$$(XI.44) \quad W = -6H\rho_{kin},$$

proporzionale alla sola energia cinetica del campo; la stessa condizione  $\rho_{kin} \ll V$  ci garantisce allora la (quasi) costanza di

$$(XI.45) \quad V(\phi) \equiv \mu^4;$$

$\mu$ , nel corso dello SRD, varierà moderatamente, tipicamente per un fattore 1/2, un nonnulla in confronto all'andamento esponenziale di  $T$ . Inizialmente sarà

$$(XI.46) \quad \mu \sim f_i T_p$$

e, quindi, per inflazione legata a una transizione di fase (con temperatura critica  $T_c$ ) sarà anche  $\mu \sim T_c$ .

Se, a un certo punto dello SRD, si riaccende l'interazione tra campo inflatonico e

altre componenti, di modo che l'energia potenziale si *termalizza* (quasi) istantaneamente, la temperatura risale a un valore dato dall'equazione

$$(XI.47) \quad \mu^4 = \frac{\pi^2}{30} g_{rh} T_{rh}^4 \quad \text{con} \quad g_{rh} = \frac{7}{8} \mathcal{N}_{fer} + \mathcal{N}_{bos}$$

dove  $\mathcal{N}_{fer}$  e  $\mathcal{N}_{bos}$  sono i numeri di stati di spin fermionici e bosonici, rispettivamente, alla temperatura  $T_{rh}$ , nel *brodo termico* così prodotto.

Ciò accade se la microfisica del modello permette un *buon reheating*, cioè un valido trasferimento dell'energia potenziale del campo inflatonico in nuova componente radiativa, senza significative dispersioni.

Ci sono varie ragioni per privilegiare modelli che diano un buon reheating (e qualche ragione che spinge nella direzione opposta). Tra le ragioni per cui un'elevata  $T_{rh}$  è auspicabile, c'è soprattutto il problema di permettere un'adeguata *bariogenesi*.

### **Reheating perturbativo**

Nelle analisi iniziali del meccanismo di reheating, il processo veniva senz'altro trattato in modo perturbativo: il reheating avviene per decadimento «individuale» dei quanti del campo inflatonico in particelle del futuro brodo termico. Esse possono appartenere a campi bosonici  $\eta_r$  (per semplicità scalari) e campi fermionici  $\psi_r$  (l'indice  $r$  serve a enumerare i campi). Al potenziale efficace delle forze responsabili per i decadimenti

$$(XI.48) \quad \phi \rightarrow \eta_r + \eta_r \quad \phi \rightarrow \bar{\psi}_r + \psi_r$$

viene attribuita un'espressione del tipo

$$(XI.49) \quad V_{int} = - \sum_r (m C_{b,r} \phi \eta_r^2 + C_{f,r} \phi \bar{\psi}_r \psi_r)$$

dove  $C_{b,r}$  e  $C_{f,r}$  sono costanti di accoppiamento adimensionali  $\ll 1$  (ciò rende possibile una trattazione perturbativa);  $m$  è la massa dell'inflatone. Si ottiene così un'espressione per il tasso di decadimento dell'inflatone

$$(XI.50) \quad \Gamma = g_{rh} \langle C^2 \rangle \frac{m}{8\pi},$$

dove si fattorizza il numero degli stati di spin del (futuro) brodo termico ( $g_{rh}$ ) per mettere in evidenza la molteplicità dei canali di decadimento; il valor medio  $\langle C^2 \rangle$  sarà poi valutato in modo che sia

$$(XI.51) \quad g_{rh} \langle C^2 \rangle = \sum_r (C_{b,r}^2 + C_{f,r}^2).$$

Si prova allora che decadimenti dissipano una potenza (vedi appendice A5)

$$(XI.52) \quad W_d = -2\Gamma \rho_{kin}$$

che si confronta facilmente con la dissipazione dovuta all'espansione cosmica (XI.44). Se

poi, quando  $\Gamma$  raggiunge  $3H$ , l'accoppiamento sale ancora e i decadimenti rendono (praticamente) istantaneo il reheating, avremo

$$(XI.53) \quad \Gamma^2 = (3H)^2 = 24\pi\mu^4/m_p^2 \quad \text{e, quindi,} \quad \mu^4 = \Gamma^2 m_p^2 / 24\pi.$$

Usando questo valore in eq. (XI.47) si ricava

$$(XI.54) \quad T_{rh} = R(\Gamma m_p)^{1/2} = T_p R \sqrt{\Gamma t_p} \quad \text{con} \quad R \simeq \left( \frac{5}{4\pi^3 g_{rh}} \right)^{1/4} \simeq 0.14 \left( \frac{100}{g_{rh}} \right)^{1/4}$$

che è la temperatura massima raggiungibile ( $g_{rh}$  è stato normalizzato a 100, valore ragionevole per il numero di stati di spin poco sotto  $T_{GUT}$ ).

Questi limiti superiori al reheating perturbativo si possono ulteriormente precisare tenendo conto che la validità stessa dell'espressione perturbativa (XI.50) richiede sia  $\langle C^2 \rangle \sim \mathcal{O}(10^{-2})$ . Pertanto, la stessa eq. (XI.50) implica che, per un'inflazione caotica con potenziale  $V_{(2)}$ , deve essere

$$(XI.55) \quad \Gamma \gtrsim m/8\pi$$

mentre, al fine di generare fluttuazioni di ampiezza ragionevole (vedi prossimo paragrafo), deve essere

$$(XI.56) \quad m \gtrsim 10^{-5} m_p.$$

L'assieme dei vincoli (XI.55) e (XI.56) implica allora che sia

$$(XI.57) \quad \Gamma/m_p = \Gamma t_p < 10^{-5}/8\pi \quad \text{ovvero} \quad \sqrt{\Gamma t_p} < 6 \times 10^{-4}.$$

Sarà quindi  $T_{rh} < 10^{-5} T_p$ .

Se ne conclude che è comunque difficile superare una temperatura di reheating  $\sim 10^{14}$  GeV, circa un ordine di grandezza al di sotto della temperatura delle GUT. La conclusione è che, se questo è il reheating, c'è un problema riguardante la bariogenesi che deve essere affrontato nell'ambito di modelli specifici.

Tra gli elementi da tenere presenti, comunque, c'è la scelta della forma dell'interazione tra campo inflatonico e campi quantistici (XI.49), che può essere opportunamente modificata. Ad esempio, anche se ciò crea difficoltà quantitative, non è detto che  $\langle C^2 \rangle$  debba necessariamente essere  $\sim \mathcal{O}(10^{-2})$ . Ma la forma stessa dell'accoppiamento può essere diversa e portare al cosiddetto *preheating* e/o al *reheating risonante*.

### **Dinamica del reheating**

Restando ancora nell'ambito di una accoppiamento con espressione (XI.49), è opportuno sottolineare che l'equazione obbedita dal campo inflatonico, prima e durante il reheating, può essere analoga a quella di un oscillatore armonico smorzato, come per i circuiti RCL. Lo smorzamento è dovuto alle perdite di potenza  $W$  (eq. XI.43) e  $W_d$  (eq. XI.52). Quando le oscillazioni hanno luogo, esse avvengono attorno al minimo di  $V$ , per potenziali delle forme mostrate sia in figura (XI.7) sia in figura (XI.8). Come per i circuiti RCL esiste una condizione di smorzamento critico che dà la condizione ge-

nerale per la presenza di oscillazioni:

$$(XI.58) \quad 3H + \Gamma < 2\sqrt{V'(\phi)/\phi};$$

nel caso di un potenziale  $V_{(2)} = m^2\phi^2/2$ , essa diviene

$$(XI.59) \quad 3H + \Gamma < m$$

e, se è esattamente  $3H + \Gamma = m$ , il reheating è ottimale. Se allora poniamo l'inizio del reheating al momento in cui  $\Gamma = 3H$  e utilizziamo la relazione  $H^2 = 8\pi V/3m_p^2$ , la condizione (XI.58) diviene

$$(XI.60) \quad \phi V/V' < (m_p/6\pi)^2$$

sicché, nel caso di un potenziale  $V_{(2)}$ , si ricava

$$(XI.61) \quad \phi < m_p/6\pi.$$

Possiamo confrontare  $m_p/6\pi \simeq 0.053 \times \phi_e$  col valore di  $\phi_e \simeq 0.28 \times m_p$  calcolato in eq. (XI.40) che, dunque, viola la condizione (XI.61). Nel caso del potenziale  $V_{(2)}$ , dunque, non si hanno oscillazioni attorno al minimo, ma un avvicinamento significativamente rallentato rispetto alle condizioni di smorzamento critico.

Questo risultato vale se si assume che la condizione  $\Gamma = 3H$  è raggiunta (quasi) simultaneamente alla condizione  $\varepsilon$ ,  $\eta \simeq 0.1$ . È invece possibile che, per  $\varepsilon$ ,  $\eta \lesssim 0.1$  sia ancora  $\Gamma \ll 3H$  e si debba andare a valori decisamente inferiori di  $\phi$  prima che lo smorzamento prevalente provenga dal reheating. Se il valore raggiunto da  $\phi$  è  $\gtrsim 0.2\phi_e$  (eq. XI.40), si innescheranno allora delle oscillazioni smorzate (vedi appendice A4), secondo la legge

$$(XI.62) \quad \phi = \Phi(t) \exp(imt) \quad \text{con} \quad \Phi(t) = \frac{m_p}{m\sqrt{3\pi t}} \exp(-\Gamma t/2),$$

nel caso di inflazione caotica da potenziale  $V_{(2)}$ . Andamenti analoghi possono essere peraltro attesi in tutti i casi in cui si raggiunga il regime oscillatorio.

Se il regime oscillatorio è prolungato e il potenziale, attorno al minimo, ha un andamento quadratico, la densità d'energia prevalente sarà quella tipica di un moto armonico, in cui  $\langle \rho_{kin} \rangle \simeq \langle V \rangle$ . Il valore medio della pressione sarà allora  $\langle p \rangle = \langle (\rho_{kin} - V) \rangle \simeq 0$  e il contributo del campo inflatonico all'espansione, nel corso del reheating, è analogo a quello della materia *non relativistica*. Fin tanto che il reheating non porta la nuova componente radiativa a superare la densità d'energia del campo inflatonico, avremo quindi un periodo di espansione in regime *dominato dalla materia*.

### **Reheating risonante**

La maggiore riserva che va fatta, su queste conclusioni, concerne la forma assunta per l'interazione tra campo inflatonico e campi quantistici (XI.49) e l'ipotesi di decadimento individuale dei quanti del campo in quanti del brodo termico.

Già nel corso degli anni Novanta è stato notato che il reheating può essere com-

pletamente diverso se l'accoppiamento inflatone-scalari ha la forma

$$(XI.63) \quad V_{rh} = (\lambda/2) \phi^2 \eta^2,$$

ove le particelle generate dal reheating sono i quanti del campo  $\eta$ . Un simile potenziale, anche per valori piccoli della costante di accoppiamento  $\lambda$ , dà origine alla sequenza di tre processi:

- 1) Preheating. Viene così chiamato un processo, che può iniziare anche prima della fine dello SRD, in cui avviene una produzione risonante di quanti  $\eta$ .
- 2) Reheating perturbativo. Anche col potenziale (XI.63) c'è la possibilità di produzione individuale di particelle  $\eta$ , da calcolare in approssimazione di Born.
- 3) Termalizzazione. Gran parte della produzione risonante non distribuisce termicamente i prodotti. È quindi necessario pensare a una loro successiva termalizzazione con redistribuzione dell'entropia, oltre che sui quanti  $\eta$ , sui quanti di altri campi interagenti con le particelle  $\eta$  prodotte nel decadimento dell'inflatone.

In appendice A4 tratteremo brevemente il modello che si ottiene assommando i potenziali  $V_{(2)}$  e  $V_{rh}$ . La potenza dissipata per produzione di particelle  $\eta$ , in questo caso, è

$$(XI.64) \quad W_d \simeq \lambda \langle \eta^2 \rangle \sqrt{\rho_{kin}}$$

e la difficoltà di una trattazione quantitativa sta nella valutazione di  $\langle \eta^2 \rangle$ . Il valore massimo dell'impulso delle particelle prodotte è dato dalla relazione

$$(XI.65) \quad k \gtrsim \lambda^{1/4} (m\Phi/2)^{1/2}$$

dove l'ampiezza  $\Phi$  si ricollega al campo  $\phi$  tramite la prima delle eq. (XI.62). Una buona produzione risonante di particelle  $\eta$  si ha per  $\lambda \gtrsim 10^{-5}$ . Maggiori dettagli sono forniti in appendice A4. È opportuno sottolineare, comunque, che sensibili modifiche a questi risultati ci sono in altri modelli considerati in letteratura.

### **Bariogenesi**

Anzitutto, *la bariogenesi deve avvenire alla fine dell'inflazione o dopo l'inflazione*. Un'eventuale densità di numero barionico  $B$ , precedente l'inflazione, viene diluita dal processo inflazionario di un fattore  $(a_i/a_e)^3 \simeq \exp(-3N_e) \sim 10^{-100}$ . In questo essa segue la stessa sorte di eventuali gravitini primordiali o «mostri topologici», il cui numero va al di sotto di 1 per orizzonte, nella nostra epoca.

Le *condizioni di Sacharov* per la bariogenesi sono state enunciate nel Par. 1 del precedente capitolo. Le GUT sono certamente B-violanti e tutto il processo inflazionario è una rincorsa al ristabilimento dell'equilibrio statistico. La violazione di  $\mathcal{CP}$  può essere espressa tramite un angolo «piccolo»  $\varepsilon_{CP}$ . Calcoli dettagliati forniscono allora un valore

$$(XI.66) \quad \frac{B}{S} \simeq SF \times \varepsilon_{CP}$$

in cui  $SF \sim \mathcal{O}(10^{-3}-10^{-4})$  è un fattore statistico derivante dall'integrazione sullo spazio delle fasi. L'ordine di grandezza atteso per  $\varepsilon_{CP}$  è  $\sim 10^{-5}-10^{-6}$ , sicché valori di  $B/S \sim 10^{-9}$  non sono difficili da ottenere. Il valore preciso dipende peraltro dal modello specifico.

## 9. Fluttuazioni e onde gravitazionali

Il capitolo concettualmente più delicato e, probabilmente, il più significativo, nello studio dell'inflazione riguarda la nascita delle fluttuazioni. Gran parte dell'articolazione dei modelli specifici proposti per l'inflazione ha avuto proprio lo scopo di verificare la possibilità di avere fluttuazioni di diverso tipo. È bene quindi distinguere anche qui tra previsioni generiche e risultati ottenibili da specifici modelli, dettagliando ulteriormente quanto già visto nel Par. 2.

### Previsioni generiche

L'inflazione provoca fluttuazioni scalari e, *possibilmente*, fluttuazioni tensorali. Dalle prime si ricaveranno i «semi» delle disomogenità oggi presenti nel cosmo. Le seconde, se presenti, si manifesteranno come *onde gravitazionali* (GW: *gravitational waves*).

INDICE SPETTRALE. Ci sono specifiche previsioni sulla forma dello *spettro di potenza* e sulla *funzione spettrale* delle fluttuazioni scalari che, fino all'entrata nell'orizzonte e nella gauge *newtoniana*, hanno rispettivamente la forma

$$(XI.67) \quad P(k) = \langle |\delta(k)|^2 \rangle = A k^{n_s}, \quad \Delta^2(k) = \frac{k^3}{2\pi^2} P(k).$$

Esse riguardano, in primo luogo, l'indice spettrale che ha espressione

$$(XI.68) \quad n_s = 1 - 6\varepsilon + 2\eta$$

e, per  $\varepsilon, \eta \simeq 10^{-2}$ , è quindi previsto essere  $\sim 0.96$ . Più in generale, valori lievemente sotto l'unità sono preferiti.

Ricordiamo che, per  $n_s = 1$ , abbiamo il cosiddetto *spettro di Harrison-Zel'dovich*, cui corrispondono fluttuazioni uguali su tutte le scale al momento dell'entrata nell'orizzonte. Spettri con  $n_s > 1$  vengono detti *blu*. Spettri con  $n_s$  lievemente sotto l'unità vengono invece detti *tilted*. Non si usa la denominazione *spettri rossi*, che letteralmente presupporrebbe un valore negativo di  $n_s$ .

ADIABATICITÀ. In un generico modello inflazionario, inoltre, le fluttuazioni scalari sono *adiabatiche*, riguardando le componenti materiale (indice  $m$ : DM e barioni) e radiativa (indice  $r$ :  $\gamma$  e  $\nu$ ) nelle proporzioni

$$(XI.69) \quad \delta_m = (3/4)\delta_r$$

sicché, in particolare, sarà  $\delta_b = (3/4)\delta_\gamma$  (per barioni e fotoni) con conservazione del-

*l'entropia-per-barione*, punto per punto. Nei modelli con DE dinamica, le fluttuazioni riguardano anche la DE, fino all'entrata nell'orizzonte e poco dopo.

**GAUSSIANITÀ.** Mentre  $P(k)$  si ricava dal valore medio di  $|\delta(k)|^2$ , è ovvio che sarà  $\langle \delta(k) \rangle = 0$ . A ogni realizzazione del modello cosmologico, in un determinato *patch* di cielo, corrisponde peraltro un diverso valore di  $\delta(k)$  e, in un generico modello di inflazione, si prevede che la distribuzione di  $\delta(k)$ , attorno allo zero, sia *Gaussiana*.

**ONDE GRAVITAZIONALI.** Uno spettro di potenza e una funzione spettrale ( $P_T$  e  $\Delta_T^2$ ) possono essere analogamente definite anche per le fluttuazioni tensoriali. In un modello generico di inflazione, l'indice spettrale dello spettro tensoriale è previsto essere

$$(XI.70) \quad n_T = -2\varepsilon$$

e, assumendo  $\varepsilon$  e  $\eta$  dello stesso ordine di grandezza, ne seguirebbe che

$$(XI.71) \quad n_T \simeq n_s - 1,$$

relazione che peraltro è *facilmente violata*, anche in modelli non *ad-hoc*.

**AMPIEZZA DEGLI SPETTRI.** Si possono dare espressioni generiche per le ampiezze dei modi scalare e tensoriale, in funzione del valore del parametro di Hubble  $H$ , al momento in cui una determinata scala esce dall'orizzonte durante lo SRD, e di  $H' \simeq H/\phi$ , allo stesso istante, risultando

$$(XI.72) \quad \Delta(k) = \frac{2}{m_p^2} \frac{H^2}{|H'|}, \quad \Delta_T(k) = \frac{4}{\sqrt{\pi}} \frac{H}{m_p}.$$

Della prima di queste equazioni daremo una prova semiquantitativa nel prosieguo del paragrafo.

Facendo il rapporto tra le funzioni spettrali si ricava

$$(XI.73) \quad \frac{\Delta_T}{\Delta} = \frac{2}{\sqrt{\pi}} \frac{|H'| m_p}{H} \simeq \frac{2}{\sqrt{\pi}} \frac{m_p}{\phi}$$

e il rapporto tra ampiezze sarà il quadrato di questo valore.

### **Potenziali $V_{(2)}$ e $V_{(4)}$**

In modelli caotici a grande campo da potenziali  $V_{(2)} = m^2 \phi^2/2$  o  $V_{(4)} = \lambda \phi^4$  le eq. (XI.72) e (XI.73) divengono

$$(XI.74) \quad \Delta^{(2)}(k) = 4\sqrt{\frac{\pi}{3}} \frac{m\phi^2}{m_p^3}, \quad \Delta_T^{(2)}(k) = \frac{8}{\sqrt{3}} \frac{m\phi}{m_p^2}, \quad \frac{\Delta_T^{(2)}}{\Delta^{(2)}} = \frac{2}{\sqrt{\pi}} \frac{m_p}{\phi}$$

$$(XI.75) \quad \Delta^{(4)}(k) = 2\sqrt{\frac{\pi}{3}} (2\lambda)^{1/2} \frac{\phi^3}{m_p^3}, \quad \Delta_T^{(4)}(k) = \frac{8}{\sqrt{3}} (2\lambda)^{1/2} \frac{\phi^2}{m_p^2}, \quad \frac{\Delta_T^{(4)}}{\Delta^{(4)}} = \frac{4}{\sqrt{\pi}} \frac{m_p}{\phi}$$

Queste relazioni permettono di ottenere la stima di  $m$ , molte volte citata. Trascurando i coefficienti numerici e ponendo  $\dot{\phi} \simeq m_p$ , per avere  $\Delta \simeq 10^{-5}$  occorre sia  $m \simeq 10^{-5} m_p$ . Se si tiene conto dei coefficienti numerici e del fatto che  $\dot{\phi} \sim 2-3m_p$ , la stima «peggiore» di un fattore  $\sim 10$ . Analogamente, nel caso di  $V_{(4)}$ , si ottiene  $\lambda \sim 10^{-11}$ . Questo dato da solo testimonia del fatto che *il campo di Higgs delle GUT non può essere il campo inflatonico*.

### Significato della stima di $\Delta$

Si può stimare il valore di  $\Delta$  in modo semi-quantitativo, tenendo conto di due fatti noti e due dati che qui non dimostreremo. È infatti noto che, durante lo SRD, è  $-\dot{\phi} = V'/3H$  mentre  $\rho \simeq V$ . Occorre viceversa accettare che, fuori dall'orizzonte, la quantità  $\delta\rho/(\rho + p)$  rimanga costante nel tempo. La si stima quindi all'uscita dall'orizzonte degli eventi e il valore ottenuto vale al momento del «recente» rientro nell'orizzonte di particella. Inoltre, il valore del campo inflatonico, sull'orizzonte degli eventi, è soggetto a indeterminazione quantistica entro un intervallo di ampiezza  $\delta\phi \simeq H/4\pi$ .

Tenendo conto che è  $H^2 \simeq 8\pi V/3m_p^2$  e, quindi,  $V' \simeq 3m_p^2 HH'/4\pi$ , avremo:

$$(XI.76) \quad \frac{\delta\rho}{\rho + p} = \frac{V'(\phi)\delta\phi}{\dot{\phi}^2} = \frac{9H^2}{V'} \delta\phi = \frac{9}{4\pi} \frac{H^3}{V'} = \frac{3}{m_p^2} \frac{H^2}{|H'|}.$$

Questo risultato, a meno di un piccolo fattore numerico, coincide con l'espressione (XI.72).

Esso mette peraltro in evidenza un fatto di grande portata concettuale: le fluttuazioni prodotte dall'inflazione hanno natura quantistica. Una discussione più dettagliata sulla loro natura, nonché del procedimento per ricavare l'ampiezza (XI.72), si trovano in appendice A5.

## 10. Da fluttuazioni quantistiche alle distribuzioni classiche osservate

È allora lecito e importante porsi la domanda: *quando e come è che queste fluttuazioni quantistiche si trasformano in fluttuazioni «classiche»?* Un osservatore vede gli oggetti cosmici in posizioni ben determinate e la domanda si può riformulare come segue: *perché una determinata galassia, con la sua specifica morfologia, si trova in un sito determinato?* Una condizione necessaria per far emergere fluttuazioni classiche dalle fluttuazioni quantistiche è la distruzione delle loro proprietà iniziali di *coerenza* e, su questo punto, c'è stato molto lavoro di ricerca. Non si tratta tuttavia di una condizione sufficiente: ogni stato quantistico è una sovrapposizione di molti diversi autostati dell'operatore «osservatore». Un simile stato, ad esempio, è invariante per traslazione: tutti i siti sono equivalenti. Quando e perché uno di questi autostati, dotati di eguali proprietà statistiche, è stato selezionato?

Si tratta di una situazione analoga a quella del famoso *gatto di Schrödinger*, salvo che nell'esperienza di Schrödinger c'è tutto un apparato *ad-hoc*, nel cui intervento si può celare, in vari modi, il passaggio da quantistico a classico. In questo caso, invece, si

tratta di un'evoluzione spontanea, su cui né l'osservatore né altri mette mano, se non altro perché avvenuta al di fuori dell'orizzonte.

Va aggiunto – e questo potrebbe essere il punto determinante – che il collegamento tra un determinato autostato e la geografia dei siti cosmici non è essa stessa ancora determinata. Da questo punto di vista ci sono differenze fondamentali tra la geografia tridimensionale della materia negli spazi e la geografia bidimensionale di fluttuazioni e polarizzazione del CMB.

Nel primo caso, una determinata distribuzione attuale delle masse può essere figlia di uno qualunque tra innumerevoli possibili autostati, anche molto diversi tra loro. Né si può affermare che un autostato di questa famiglia avrebbe necessariamente portato alle condizioni osservate. In effetti, il passaggio dalla linearità alla non-linearità è regolato da fisica *caotica*. Ciò è vero fin dalla scala degli ammassi, ma al di sotto della scala delle galassie si può affermare che ogni pretesa di determinismo è insostenibile. Nella distribuzione dei grandi ammassi c'è un residuo di selezione tra possibili autostati. Non ha invece assolutamente senso pensare che la posizione di Betelgeuse, Sirio o della Crab Nebula derivi dalla scelta di un determinato autostato. Tutto ciò, in qualche modo, crea un velo opaco tra la scelta dell'autostato, al di sotto di una determinata scala, e il risultato dell'osservazione.

In altri termini: un determinato quadro osservativo della distribuzione di materia è in sé consistente non con una singola autofunzione, né con un singolo pacchetto d'onda, ma con autofunzioni distribuite in modo difficile da codificare.

Molto diversa è la situazione per ciò che concerne il CMB. Tra le osservazioni di massimi e minimi di temperatura e dei campi di polarizzazione, da un lato, e il set di stati che l'inflazione ha generato, dall'altro, non c'è nessun velo caotico. A nostro avviso qui c'è un problema aperto. Se ci atteniamo ai dettami della *scuola di Copenhagen* dovremmo affermare che le *osservazioni* di COBE hanno selezionato tra i possibili stati del campo di fluttuazioni del CMB, e che le ulteriori osservazioni successive, di *Boomerang*, *WMAP*, ecc., hanno creato ulteriori selezioni tra gli autostati che a priori erano possibili. È come se il cosmo fosse rimasto, in questa sua componente, a lungo indeterminato, e abbia assunto una geografia determinata solo quando osservato.

È opportuno richiamare l'attenzione sul fatto che le osservazioni non hanno interferito con le proprietà statistiche, che esse erano e restano prevedibili una volta assegnato il modello. Ma una cosa sono gli andamenti degli spettri  $C_l$  e altra sono le distribuzioni «geografiche» delle temperature. Ad esempio, se tutto il quadro osservato delle temperature viene traslato in blocco, come avverrebbe se si modifica la direzione dell'asse polare, le proprietà statistiche restano esattamente le stesse, mentre la temperatura di ciascun sito è cambiata.

In entrambi i casi, comunque, i quesiti circa le modalità della selezione non hanno una risposta fisica; la loro soluzione è demandata all'interpretazione *filosofica* delle basi concettuali della meccanica quantistica. Si tratta peraltro di una situazione *unica*: una dinamica quantistica, in questo caso, provoca indeterminazione su scale ultramacroscopiche.

# Appendice A1

## Cinematica dell'inflazione

1. Espansione accelerata secondo legge di potenza.

L'andamento del fattore di scala  $a(t)$  è regolato dalle equazioni di Friedmann. In particolare, l'equazione

$$(A1.1) \quad -\frac{\ddot{a}}{a} = \frac{8\pi}{3}G(\rho + 3p)$$

ci mostra direttamente che l'espansione sarà accelerata ( $\ddot{a} > 0$ ) se e solo se

$$(A1.2) \quad \rho + 3p < 0 \quad \text{vale a dire:} \quad w \equiv p/\rho < -1/3.$$

Se  $w$  è costante, il fattore di scala varia nel tempo con la legge di potenza

$$(A1.3) \quad a/a_i = (t/t_i)^\alpha, \quad \text{con } \alpha = 2/[3(w+1)],$$

ove  $a_i$  è il valore del fattore di scala a un tempo di riferimento  $t_i$ . È allora facile vedere che parametro di Hubble e accelerazione sono

$$(A1.4) \quad H = \frac{\alpha}{t} = \frac{2}{3(w+1)} \frac{a}{t}, \quad \ddot{a} = \alpha(\alpha-1) \frac{a}{t^2},$$

mentre l'eq. (A1.3) ci conferma che si avrà  $\alpha > 1$ , e quindi accelerazione, se è  $w < -1/3$ .

Gli stessi valori,  $\alpha = 1$  e  $w = -1/3$ , delimitano anche due regimi diversi, per quanto concerne l'orizzonte causale: per espansione con  $\alpha > 1$  ( $< 1$ ) (da tempi e verso tempi indefiniti) c'è orizzonte degli eventi (di particella) e non c'è orizzonte di particella (degli eventi). Infatti i due orizzonti sono

$$(A1.5) \quad \begin{aligned} l_p(t) &= c a(t) \int_0^t \frac{dt'}{a(t')} = c t^\alpha \int_0^t \frac{dt'}{t'^\alpha} = c t^\alpha \frac{1}{1-\alpha} \left\{ t'^{1-\alpha} \Big|_0^t \right\} && \text{(particella)} \\ l_{ev}(t) &= c a(t) \int_t^\infty \frac{dt'}{a(t')} = c t^\alpha \int_t^\infty \frac{dt'}{t'^\alpha} = c t^\alpha \frac{1}{1-\alpha} \left\{ t'^{1-\alpha} \Big|_t^\infty \right\} && \text{(eventi)} \end{aligned}$$

La quantità in parentesi grafa diverge se  $\alpha > 1$  ( $\alpha < 1$ ) nel caso di orizzonte di particella (degli eventi). Viceversa, per  $\alpha < 1$ , l'orizzonte di particella risulta

$$(A1.6) \quad l_p(t) = ct/(1-\alpha),$$

mentre, per  $\alpha > 1$  (espansione accelerata), l'orizzonte degli eventi risulta

$$(A1.7) \quad l_{ev}(t) = ct/(\alpha-1).$$

Se è  $\alpha > 1$ , le distanze, essendo  $\propto t^\alpha$ , crescono più di  $l_{ev} \propto t$ : anche se inizialmente era possibile mandare un segnale da  $A$  a  $B$ , dopo un certo tempo  $B$  esce comunque dall'orizzonte degli eventi di  $A$ . Così, in un periodo inflazionario primordiale, si possono realizzare processi fisici (ad esempio omogeneizzazione) su scale che poi escono dall'orizzonte e sono ancor oggi superiori all'orizzonte o solo di recente divenute inferiori all'orizzonte.

L'espansione accelerata secondo legge di potenza viene talora denominata **subinflazione**.

## 1. Espansione esponenziale

I risultati del precedente paragrafo escludono il caso  $w = -1$ . Ad esempio, eq. (A1.3) darebbe  $\alpha \rightarrow \infty$  per  $w = -1$ . Il caso  $w = -1$  è però fisicamente significativo, ad esempio per una sorgente di «falso vuoto», quando  $T_{ij} = Ag_{ij}$  sicché risulta  $p = -\rho = -A$ . Per  $w = -1$ , l'equazione di Friedmann

$$d(\rho a^3) = \rho d(a^3)$$

implica sia  $\rho \equiv \text{const.}$ : il lavoro compiuto per espandere la sostanza cosmica produce esattamente l'energia necessaria a mantenerne costante la densità.

Viceversa, l'equazione di Friedmann

$$(A1.8) \quad \frac{\dot{a}}{a} = \left( \frac{8\pi\rho}{3m_p^2} \right)^{1/2} \equiv H \quad \text{ha integrale} \quad a(t) = a_i \exp[H(t - t_i)],$$

se  $\rho$  e  $H$  sono costanti, sicché è anche

$$(A1.9) \quad \ddot{a} = H^2 a > 0.$$

Questo regime espansivo è quello comunemente denominato **inflazione**.

In questo caso risultano finiti sia l'orizzonte di particella

$$l_p = \frac{1}{H} (e^{Ht} - 1)$$

che quello degli eventi

$$(A1.10) \quad l_{ev} = \frac{1}{H},$$

che è costante: la distanza cui è possibile inviare un segnale permane sempre eguale, mentre invece la distanza tra punti cresce esponenzialmente.

Il processo inflazionario avviene per un intervallo di tempo finito. Denominiamo allora  $t_i$  e  $a_i$  il tempo e il fattore di scala quando l'inflazione ha inizio. Supponiamo poi che essa abbia fine a un tempo  $t_e \gg t_i$ , quando il fattore di scala ha raggiunto un valore

$$(A1.11) \quad a_e = a_i \exp[H(t_e - t_i)] \simeq a_i \exp[H t_e] \equiv a_i \exp N_e.$$

$N_e$  è il «numero di e-folding» del processo inflazionario.

Si può anche immaginare un regime di **superinflazione**, che qui non discuteremo, in cui sia  $a \propto \exp[(Kt)^v]$ , con  $v$  costante (o crescente) nel tempo.

Del tutto in generale, si può denominare inflazione ogni regime espansivo che avvenga con legge  $a \propto \exp[F(t)]$ , con  $F(t)$  crescente. La subinflazione rientra in questo caso, assumendo che  $F(t)$  abbia un andamento logaritmico.

## Appendice A2

### Dinamica inflazionaria

Nel caso di inflazione dovuta a un campo scalare  $\phi$ , partendo dalla lagrangiana

$$(A2.1) \quad \mathcal{L} = (1/2)g_{\mu\nu}\partial^\mu\phi\partial^\nu\phi - V(\phi),$$

si ricavano le componenti del tensore energia-impulso

$$(A2.2) \quad T^{\mu\nu} = \partial^\mu\phi\partial^\nu\phi - [(1/2)g_{\alpha\beta}\partial^\alpha\phi\partial^\beta\phi - V(\phi)]g^{\mu\nu}$$

che, al limite di omogeneità ( $\phi$  indipendente da  $x^i$ ,  $i = 1, 2, 3$ ), danno

$$(A2.3) \quad T^{00} = \rho_\phi = \dot{\phi}^2/2 + V(\phi), \quad -T^{ii} = p_\phi = \dot{\phi}^2/2 - V(\phi)$$

(in  $T^{ii}$  non si somma sugli indici ripetuti). In una metrica FRW spazialmente piatta, la lagrangiana (A2.1) dà luogo all'integrale variazionale

$$(A2.4) \quad S = \int dt a^3 x a^3(t) [(1/2)g_{\mu\nu}\partial^\mu\phi\partial^\nu\phi - V(\phi)]$$

da cui, al limite di omogeneità, si deriva l'equazione del moto

$$(A2.5) \quad \frac{\partial}{\partial t} \left[ a^3(t) \frac{\partial\phi}{\partial t} \right] + a^3(t)V'(\phi) = 0 \quad \text{da cui:} \quad \ddot{\phi} + 3\frac{\dot{a}}{a}\dot{\phi} + V'(\phi) = 0.$$

Questa equazione è equivalente all'equazione di Friedmann

$$(A2.6) \quad \dot{\rho} + 3\frac{\dot{a}}{a}(\rho + p) = 0.$$

Difatti, tramite le espressioni (A2.3), si ricava

$$\frac{\partial}{\partial t} \left[ \frac{1}{2}\dot{\phi}^2 + V(\phi) \right] + 3\frac{\dot{a}}{a}\dot{\phi}^2 = 0 \quad \text{da cui} \quad \left\{ \ddot{\phi} + 3\frac{\dot{a}}{a}\dot{\phi} + V'(\phi) \right\} \dot{\phi} = 0,$$

purché, nel cosmo, sia presente il solo campo  $\phi$ . Se, accanto a esso, c'è una componente con equazione di stato  $p_r = \rho_r/3$ , l'equazione di Friedmann (A2.6) dà

$$(A2.7) \quad \left\{ \ddot{\phi} + 3H\dot{\phi} + V'(\phi) \right\} \dot{\phi} = \dot{\rho}_r + 4H\rho_r$$

Se il campo  $\phi$  entra solo nella lagrangiana (A1.1), vale l'eq. del moto (A2.5) e i due membri dell'eq. (A2.7) si annullano separatamente. Se invece c'è interazione tra campo e componenti radiative, con  $\mathcal{F}$  opportuna, sarà

$$(A2.8) \quad \ddot{\phi} + 3H\dot{\phi} + V'(\phi) = \mathcal{F}, \quad \dot{\rho}_r + 4H\rho_r = -\mathcal{F}\dot{\phi}.$$

Sotto opportune condizioni (appendice A5) sarà  $\mathcal{F} = -\Gamma\dot{\phi}$  e quindi

$$(A2.9) \quad \ddot{\phi} + 3H\dot{\phi} + \Gamma\dot{\phi} + V'(\phi) = 0, \quad \dot{\rho}_r + 4H\rho_r = \Gamma\dot{\phi}^2 :$$

l'interazione tra campo e radiazione costituisce una specie di «attrito» che muta energia ordinata nel campo  $\phi$  in componente termica. Queste equazioni danno anche le (densità di) potenza

dissipate o ricevute

$$(A2.10) \quad W_{\dot{\phi}} = \dot{\rho}_{\dot{\phi}} = -(3H + \Gamma)2\rho_{kin}, \quad W_r = \dot{\rho}_r = -4H\rho_r + \Gamma 2\rho_{kin}$$

da ciascuna componente ( $\rho_{kin} = \dot{\phi}^2/2$ ).

L'espansione sarà (quasi-)esponenziale se  $p \simeq -\rho$  e, quindi, viste le eq. (A2.3), la componente radiativa deve essere trascurabile, e deve essere

$$(A2.11) \quad \dot{\phi}^2 \ll V(\phi)$$

(energia cinetica trascurabile rispetto all'energia potenziale), sicché risulti

$$(A2.12) \quad H^2 = 8\pi V(\phi)/3m_p^2.$$

Lo SRD è assicurato solo se la variazione di  $\dot{\phi}$  è trascurabile; deve perciò essere

$$(A2.13) \quad \ddot{\phi} \ll 3H\dot{\phi}, \quad V'(\phi) \quad \text{e} \quad 3H\dot{\phi} \simeq -V'(\phi),$$

sicché sarà anche trascurabile la potenza dissipata in radiazione. L'inevitabile variazione di  $V(\phi)$  deve quindi essere talmente diluita nel tempo da rendere  $H$  (quasi) costante, di modo che l'espansione sia (quasi-)esponenziale. Le blande deviazioni dal ritmo esponenziale saranno poi quantificate tramite i parametri

$$(A2.14a) \quad \varepsilon = (3/2)(\dot{\phi}^2/V), \quad \eta = -\ddot{\phi}/H\dot{\phi}$$

Lo SRD termina quando  $\dot{\phi}^2/2$  approssima  $V/3$  (cioè  $\varepsilon \sim 1$ ) e  $\ddot{\phi}$  approssima  $H\dot{\phi}$  (cioè  $\eta \sim 1$ ). Sia pure lentamente, comunque, durante l'espansione inflazionaria  $\phi$  varia, sicché il numero di e-folding va espresso come

$$(A2.15) \quad N_e \equiv \frac{a_f}{a_i} = \int_{t_i}^{t_f} dt H[\phi(t)] = \int_{\phi(t_i)}^{\phi(t_f)} d\phi \frac{H(\phi)}{\dot{\phi}} = \frac{8\pi}{m_p^2} \int_{\phi(t_i)}^{\phi(t_f)} d\phi \frac{V(\phi)}{V'(\phi)}$$

con  $H(\phi)$  dato da eq. (A2.11). Si può poi facilmente verificare che le espressioni

$$(A2.14b) \quad \varepsilon = (m_p^2/16\pi)(V'/V)^2, \quad \eta = (m_p^2/8\pi)(V''/V)$$

e le espressioni

$$(A2.14c) \quad \varepsilon = (m_p^2/4\pi)(H'/H)^2, \quad -\eta = \varepsilon + (m_p^2/4\pi)(H''/H)$$

(in cui ' indica la derivata rispetto a  $\phi$ ) sono equivalenti alle eq. (A2.14a). Le eq. (A2.14b) e (A2.14c) danno i parametri di SRD in funzione della dipendenza da  $\phi$ , rispettivamente, del potenziale o del parametro di Hubble. A titolo d'esempio verificheremo le espressioni (A2.14b). Le espressioni (A2.14c) si provano utilizzando passaggi analoghi.

**Verifica.** L'espressione di  $\dot{\phi}$  nello SRD (eq. A2.13), sostituita nell'espressione di  $\varepsilon$  (A2.14a), essendo  $H^2 \simeq 8\pi V/3m_p^2$ , permette di ricavare

$$\varepsilon = \frac{V'^2}{6H^2V} = \frac{m_p^2}{16\pi} \left( \frac{V'}{V} \right)^2;$$

allo stesso modo, ricavando  $\ddot{\phi}$  dall'espressione (A2.13) di  $\dot{\phi}$ , si ha

$$\eta = \frac{V''\dot{\phi}}{3H} \frac{1}{H\dot{\phi}} = \frac{m_p^2 V''}{8\pi V}.$$

# Appendice A3

## Modelli di inflazione

### 1. Inflazione con campo «grande». Inflazione caotica

L'inflazione caotica deriva in genere da un «potenziale» di tipo

$$(A3.1) \quad V_{(2n)} = \lambda m^2 m_p^2 (\phi/m_p)^{2n}$$

di cui esempi sono  $V_{(2)}$  e  $V_{(4)}$  (eq. XI.31). All'emergere dal mondo pre-planckiano, sarà  $\phi \sim \mathcal{O}(m_p)$  e alla successiva evoluzione verso  $\phi = 0$  è dovuta l'inflazione.

Consideriamo più in dettaglio il caso di  $V_{(2)}$ , illustrato in figura XI.8, in cui

$$(A3.2) \quad H^2 \simeq \frac{4\pi m^2 \phi^2}{3m_p^2}.$$

Il valore iniziale  $H_i = (4\pi/3)^{1/2} m \phi_i / m_p$  può essere utilizzato per dare l'andamento approssimato del fattore di scala:

$$(A3.3) \quad a \simeq a_i \exp \left[ \left( \frac{4\pi}{3} \right)^{1/2} \frac{m}{m_p} \phi_i t \right].$$

Se si trascura l'energia cinetica di  $\phi$ , le equazioni di moto del sistema,

$$\frac{\dot{a}}{a} = \left( \frac{4\pi}{3} \right)^{1/2} \frac{m}{m_p} \phi \quad (= H),$$

$$(A3.4) \quad \ddot{\phi} + 3 \frac{\dot{a}}{a} \dot{\phi} + m^2 \phi = 0,$$

hanno comunque un integrale esatto. In esso  $\phi$  ha l'andamento

$$(A3.5) \quad \phi = \phi_i - \frac{m m_p}{\sqrt{12} \pi} (t - t_i),$$

da cui risulta  $\dot{\phi} = m m_p / \sqrt{12} \pi$  (e  $\ddot{\phi} = 0$ ) coincidente con la relazione

$$\dot{\phi} = -m^2 \phi / 3H;$$

la condizione di lento rotolamento coincide quindi col soddisfacimento della seconda delle eq. (A3.4). La prima delle eq. (A3.4) diviene invece

$$\frac{\dot{a}}{a} \simeq \left( \frac{4\pi}{3} \right)^{1/2} \frac{m}{m_p} \left( \phi_i - \frac{m m_p}{\sqrt{12} \pi} t \right)$$

e si integra immediatamente dando

$$a \simeq a_i \exp \left\{ \left( \frac{4\pi}{3} \right)^{1/2} \frac{m}{m_p} \left[ \frac{\phi_i}{m_p} t - \frac{1}{\sqrt{48} \pi} \frac{m}{m_p} \left( \frac{t}{t_p} \right)^2 \right] \right\}.$$

Si noti che, secondo eq. (A3.5), risulterebbe  $\phi = 0$  al tempo  $t_0 = \sqrt{12} \pi t_p (\phi_i / m)$ . Il secondo ad-

dendo in parentesi quadra, al tempo  $t_0$ , vale dunque  $(\phi_i/4m_p)(t_0/t_p)$  e rappresenta al più una correzione del 25% rispetto al primo termine. L'andamento di  $a(t)$  è quindi sempre ben approssimato da eq. (A3.3).

Essendo poi  $(V'/V)^2 = 2V''/V = 2/\phi^2$  si vede poi che i parametri di lento rotolamento hanno i valori

$$(A3.6) \quad \varepsilon = \eta = \frac{1}{4\pi} \left( \frac{m_p}{\phi} \right)^2,$$

sicché avremo un normale SRD solo se  $\phi_i \gg m_p/\sqrt{4\pi}$ . In generale, il numero di e-folding risulta

$$(A3.7) \quad N_e \simeq \left( \frac{4\pi}{3} \right)^{1/2} \frac{m}{m_p} \phi_i t_f$$

e, utilizzando  $t_0$  come  $t_f$ , si ottiene

$$N_e \simeq 4\pi \left( \frac{\phi_i}{m_p} \right)^2;$$

$\phi_i \simeq 3m_p$ , ad esempio, dà  $N_e \sim 100$ . Il valore di  $m$  si lega all'ampiezza delle fluttuazioni di densità, di cui parleremo al prossimo paragrafo, e deve essere  $\sim 10^{-5}m_p$ , un valore non assurdo per un campo scalare correlato al processo di grande unificazione.

## 2. Inflazione da campo «piccolo»

L'inflazione caotica avverrebbe a ridosso della scala di Planck. Attorno alla scala di grande unificazione (GUT), invece, potrebbe avvenire inflazione legata a un campo pseudo-Nambu-Goldstone (PNG), dovuto alla rottura spontanea di una simmetria globale. Un bosone di questo

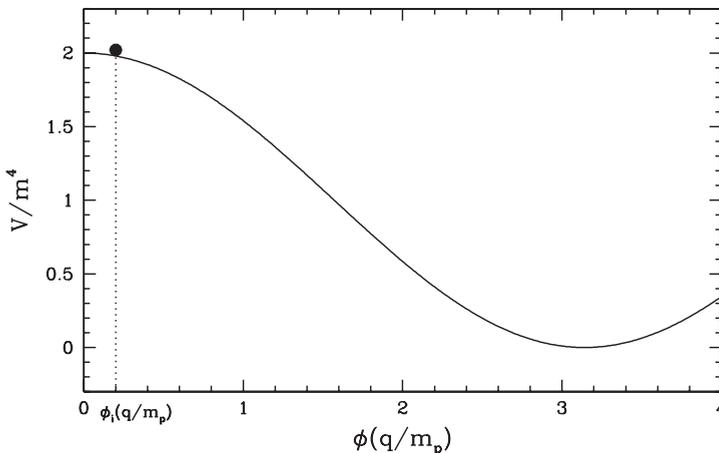


FIGURA XI.10

Andamento del potenziale in un modello di inflazione a campo «piccolo». Un valore tipico «iniziale» del campo è indicato dalla pallina. Successivamente il campo tende a evolvere verso la configurazione di equilibrio.

tipo, a bassa energia, è il mesone  $\pi$ . Alle energie tipiche del processo inflazionario, il potenziale potrebbe avere una forma approssimabile tramite l'espressione

$$(A3.8) \quad V(\phi) = m^4 \left[ 1 + \cos \left( q \frac{\phi}{m_p} \right) \right]$$

con  $m$  sulla scala delle GUT e  $q \sim \mathcal{O}(1)$  (vedi figura XI.10).

Il valore iniziale del campo è  $\sim 0$  o, comunque,  $\ll m_p(\pi/q)$ . L'inflazione avviene durante l'evoluzione di  $\phi$  verso  $\phi_m = m_p(\pi/q)$  dove il potenziale ha un minimo. Successivamente  $\phi$  oscilla attorno a  $\phi_m$ , nella fase di reheating. I parametri di SRD,

$$(A3.9) \quad \varepsilon = \frac{q^2}{16\pi} \left[ \frac{\sin(\phi q/m_p)}{1 + \cos(\phi q/m_p)} \right]^2, \quad \eta = -\frac{q^2}{8\pi} \frac{\cos(\phi q/m_p)}{1 + \cos(\phi q/m_p)}$$

dipendono da  $q$  ma non da  $m$ . Inizialmente è senz'altro  $\varepsilon \ll 1$ , mentre  $-\eta \sim q/16\pi$  è anch'esso  $\ll 1$ , sebbene in modo meno pronunciato. Possiamo quindi considerare periodo di SRD quello che si conclude quando  $\varepsilon$  raggiunge un valore  $\simeq q/16\pi$  cioè  $\phi \simeq m_p(\pi/2q)$ .

Nel corso dello SRD sarà

$$H = \left( \frac{8\pi}{3} \right)^{1/2} \frac{m^2}{m_p} \left[ 1 + \cos \left( q \frac{\phi}{m_p} \right) \right]^{1/2}$$

$$\dot{\phi} = qm^2 \frac{\sin(q\phi/m_p)}{[1 + \cos(q\phi/m_p)]^{1/2}}$$

sicché il numero di e-folding

$$(A3.10) \quad N_e = \int_{\phi_i}^{m_p(\pi/2q)} d\phi \frac{H}{\dot{\phi}}$$

dipende essenzialmente da quanto piccolo è  $\phi_i$  e divergerebbe per  $\phi_i = 0$ . Si può verificare che, con  $q \simeq 1$ , si ha  $N_e \sim 100$  per  $\phi_i \sim 10^{-2}m_p$ .

### 3. Inflazione ibrida

Consideriamo un potenziale della forma

$$(A3.11) \quad V(\phi, \chi) = \frac{m^2}{2} \phi^2 + V_o(\phi, \chi),$$

simile a quello dell'inflazione da campo grande, più un termine in prima approssimazione costante. Il suo grafico non si discosta molto da quello di figura XI.8, salvo che, al minimo,  $V$  non si annulla ma vale  $V_o$ .

La legge con cui  $V_o$  dipende da  $\chi$ , tuttavia, può essere diversa, a seconda dei valori di  $\phi$ . Supponiamo, in particolare, che, per  $\phi$  grande, il minimo rispetto a  $\chi$  si trovi a  $\chi = 0$ ; ma che, quando  $\phi$  scende al di sotto di un valore critico  $\phi_c$ ,  $\chi = 0$  divenga un massimo locale, mentre si producono due minimi in  $\pm \mu$ .

Ciò avviene, ad esempio, se

$$V_o(\phi, \chi) = \lambda_1 [\chi^2 - \mu^2(\phi)]^2 \quad \text{e} \quad \mu^2(\phi) = \lambda_2 (\phi_c^2 - \phi^2)$$

Infatti, per  $\phi < \phi_c$  ( $\phi > \phi_c$ ), risulta  $\mu^2 > 0$  ( $\mu^2 < 0$ ). Viceversa risulta

$$\partial V_o / \partial \chi = 4\chi\lambda_1[\chi^2 - \mu^2(\phi)]$$

e, se  $\mu^2 > 0$ , cioè se  $\phi < \phi_c$ ,  $V_o$  ha minimi per  $\chi = \pm \mu$  (invece, se  $\mu^2$  è negativo, l'unico minimo è in  $\chi = 0$ ).

Complessivamente il potenziale si può scrivere

$$V(\phi, \chi) = \frac{m^2}{2}\phi^2 + \lambda_1[\chi^2 - \lambda_2(\phi_c^2 - \phi^2)]^2$$

e viene rappresentato in figura XI.9; al suo minimo assoluto ( $\phi = 0$ ,  $\chi = \sqrt{\lambda_2}\phi_c$ ), si annulla. Rispetto a  $\phi$  ( $\chi$ ) si tratta dunque di inflazione da campo «piccolo» («grande»). Nella versione originale di Linde l'inflazione ibrida era essenzialmente dovuta a  $\phi$ . Il ruolo di  $\chi$  era fornire un buon reheating che, per un potenziale (A3.1), non è assicurato.

In questo caso, sulla base delle eq. (A3.5)-(A3.7), è facile vedere che il numero di e-folding è

$$(A3.12) \quad N = 4\pi(\phi_i/m_p)^2(1 - f_c) \quad \text{con} \quad f_c = \phi_c/\phi_i$$

e può essere soddisfacente anche se  $f_c \sim 0.2-0.3$ , purché sia  $\phi_i \gtrsim 4m_p$ . L'inflazione tuttavia può avere un seguito, dopo il raggiungimento di  $\phi_c$ , e in questo caso il numero totale di e-folding può ulteriormente aumentare, sicché si può rilassare il limite  $\phi_i > 3m_p$ .

# Appendice A4

## Il reheating

L'inflazione termina con il reheating, in cui ci si attende che l'energia del campo inflatonico  $\phi$  si termalizzi. Si tratta di un'aspettativa ragionevole, ma le modalità sono fortemente legate al modello e, in particolare, alla forma dell'interazione tra il campo inflatonico e i campi quantistici i cui quanti entreranno a far parte del «brodo termico» post-reheating.

### 1. Reheating perturbativo

Lo SRD termina quando  $\varepsilon, \eta \sim 1$ , sicché l'equazione (A2.5) va considerata per intero, confrontando la potenza dissipata dall'espansione

$$(A4.1) \quad W = \frac{d}{dt} \left( \frac{\dot{\phi}^2}{2} + V \right) = -3H\dot{\phi}^2$$

con quella dissipata in particelle del futuro brodo termico (la cui massa è quindi trascurabile rispetto a quella dell'inflatone). Una prima ipotesi di lavoro, capace di facilitare il raffronto tra le due potenze dissipate, si basa sul potenziale

$$(XI.49-A4.2) \quad V_{int} = - \sum_r (m C_{b,r} \phi \eta^2 + C_{f,r} \phi \bar{\psi} \psi)$$

che rende conto del possibile decadimento dei quanti di  $\phi$  in coppie di bosoni (scalari  $\chi_r$ ) e di fermioni (spinori  $\psi_r$ ) con costanti d'accoppiamento efficaci dipendenti dalla specie di particelle del brodo termico.

Se si includono allora le correzioni quantistiche all'eq. (A4.1), si ricava

$$(A4.4) \quad \ddot{\phi} + 3H\dot{\phi} + [V'(\phi) + \Pi(m)] = 0,$$

dove  $\Pi(m)$  è l'operatore di polarizzazione dell'inflatone. Essendo trascurabili le masse dei campi  $\chi$  e  $\psi$  rispetto a  $m$ ,  $Re \Pi(m)$ , in genere, sarà trascurabile rispetto a  $V'$ ; ad esempio, per un potenziale inflatonico  $V_{(2)}$ , è esattamente  $V'_{(2)} = m^2 \phi$ . Avremo invece che

$$(A4.5) \quad Im \Pi(m) = -m \Gamma \dot{\phi}$$

con

$$(XI.50-A4.6) \quad \Gamma = g_{rh} \langle C^2 \rangle \frac{m}{8\pi} \quad \text{dove} \quad g_{rh} \langle C^2 \rangle = \sum_r (C_{b,r}^2 + C_{f,r}^2).$$

La soluzione dell'eq. (A4.4), per un potenziale inflatonico  $V_{(2)}$ , sarà

$$(A4.7) \quad \phi = \Phi(t) e^{imt} \quad \text{con} \quad \Phi(t) = \frac{m_p}{m\sqrt{3\pi t}} e^{-\Gamma t}$$

come è possibile verificare per ispezione diretta. La stessa soluzione si ottiene anche aggiungendo un termine di potenza dissipata

$$(A4.8) \quad W_d = -\Gamma \dot{\phi}^2$$

al termine (A4.1), come se l'eq. (A4.4) avesse la forma

$$(A4.9) \quad \ddot{\phi} + (3H + \Gamma)\dot{\phi} + V'(\phi) = 0,$$

cui si fa riferimento, in genere, anche per potenziali diversi da  $V_{(2)}$ . Il reheating è dunque un «buon reheating» se la condizione  $\Gamma > 3H$  si realizza rapidamente alla fine dello SRD, di modo che il grosso della potenza venga subito devoluto alla creazione di entropia.

## 2. Un modello con reheating non perturbativo.

Vedremo ora come si articola il reheating, sempre in un modello di inflazione con potenziale  $V_{(2)}$ , se l'interazione coi campi quantistici prende invece la forma

$$(A4.10) \quad V(\phi, \eta) = (\lambda/2) \phi^2 \eta^2.$$

Sviluppando l'operatore di campo  $\eta$  in serie di Fourier,

$$(A4.11) \quad \eta(t, \mathbf{x}) = (2\pi)^{-3/2} \int d^3k \left[ a(\mathbf{k}) \tilde{\eta}(t, \mathbf{k}) e^{-ik \cdot \mathbf{x}} + a^\dagger(\mathbf{k}) \tilde{\eta}^*(t, \mathbf{k}) e^{ik \cdot \mathbf{x}} \right],$$

( $a^{(\pm)}(\mathbf{k})$  sono gli operatori di creazione e distruzione di quanti del campo  $\eta$ ) si ricavano le equazioni per le ampiezze  $\eta^{(\pm)}(t, \mathbf{k})$ , che dipendono solo da  $k = |\mathbf{k}|$  e si possono scrivere

$$(A4.12) \quad \ddot{\eta}_k + 3H\dot{\eta}_k + \left[ \lambda\phi^2 + (k/a)^2 \right] \eta_k = 0$$

( $\eta_k \equiv \eta(t, \mathbf{k})$ ). Se si definisce

$$(A4.13) \quad X_k = \eta_k a^{3/2}$$

l'eq. (A4.12) assume la semplice forma

$$(A4.14) \quad \ddot{X}_k + \omega_k^2(\phi, t) X_k = 0,$$

con

$$(A4.15) \quad \omega_k^2(\phi, t) = \lambda\phi^2 + (k/a)^2 - (3/4)[(\dot{a}/a)^2 + 2\ddot{a}/a].$$

Il senso di questa equazione si capisce trascurando la dipendenza di  $\omega_k$  dalle derivate di  $a$ , di modo che essa diviene

$$(A4.16) \quad X_k'' + (A_k - 2q \cos 2u) X_k = 0$$

in cui ' indica derivazione rispetto a  $u = mt$  mentre

$$(A4.17) \quad A_k = (k/ma)^2 + 2q, \quad q = (\lambda/4)(\Phi/m)^2 \quad (\text{e } \Phi = \phi e^{-iu}).$$

L'eq. (A4.16) è un'equazione di Mathieu e la sua soluzione presenta un complesso quadro di comportamenti risonanti, con regioni di instabilità dove  $X_k$  cresce esponenzialmente, di modo che sono prodotti in grande numero quanti di impulso  $k$  del campo  $\eta$ . Nel caso fisico in esame, l'eq. (A4.17) stabilisce il vincolo

$$(A4.18) \quad A_k > 2q$$

e la differenza  $A_k - 2q$  stabilisce l'impulso delle particelle che possono essere create. In figura XI.11, le regioni di instabilità consistenti coi vincoli fisici sono i triangoloidi marcati I, II, III. Il I triangoloide, detto «risonanza stretta» ammette solo valori di  $q \gtrsim 1$  e solo pochi valori di  $k$  sono

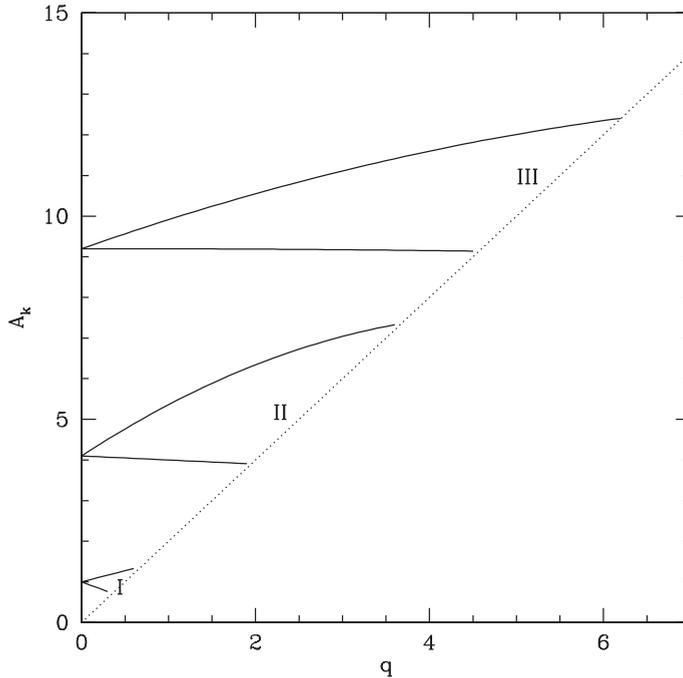


FIGURA XI.11

Gli interni dei triangoloidi I, II, III costituiscono le bande di instabilità dell'equazione di Mathieu per valori fisici delle coordinate.

possibili entro la sua area. Nei triangoloidi II e III, invece, sono ammessi valori di  $q$  superiori e quindi una banda più ampia di  $k$  («risonanze ampie»).

Il tasso di crescita della produzione di particelle  $\eta$  cresce con  $q$  e, quindi, le risonanze ampie sono più significative della risonanza stretta. Ciò che determina la natura della risonanza sono l'ampiezza iniziale del campo inflatonico e l'accoppiamento  $\lambda$ . Valendo il vincolo  $m/m_p \lesssim 10^{-5}$  (ampiezza fenomenologicamente corretta per le fluttuazioni), per avere valori di  $q$  significativamente  $> 1$  occorre sia  $\lambda \lesssim 10^{-8}$  e  $\Phi(t_i) \sim 0.1-0.2 m_p$ . Il valore massimo dell'impulso delle particelle  $\eta$  prodotte può essere stimato dalla condizione

$$(A4.19) \quad k \gtrsim \lambda^{1/4} (m\Phi/2)^{1/2}$$

e, quindi, particelle di elevato impulso si hanno per elevati valori di  $\lambda$  e  $\Phi$ .

Per valori di  $q \gg 1$ , la risonanza su ciascun modo non dà luogo a produzione coerente di particelle. Infatti, a causa dell'aumento del fattore di scala,  $\omega_k$  varia in modo significativo nell'ambito di ciascuna oscillazione di  $\Phi$ , di modo che le fasi del campo  $\eta$ , quando è  $\dot{\phi} = 0$ , sono tra loro scorrelate.

Negli stadi iniziali della produzione risonante di particelle,  $\eta$  attraversa rapidamente tutte le bande di instabilità. Quando  $q$  cala e l'espansione cosmica rallenta,  $\eta$  rimane più a lungo in ciascuna banda. La produzione risonante finisce quando  $q$  e  $A_k$  vanno a collocarsi al di sotto della risonanza stretta.

Oltre a ciò si deve tener conto della retro-reazione sul campo  $\phi$  della produzione di particelle  $\eta$  che, ovviamente, drena energia dal campo inflatonico. Dalle eq. (A4.1), aggiungendo a  $V_{(2)}$  il

potenziale (A4.10), si ricava che la potenza dissipata è

$$(X1.64-A4.20) \quad W_d \simeq \lambda \langle \eta^2 \rangle \dot{\phi},$$

in cui

$$(A4.21) \quad \langle \eta^2 \rangle = (2\pi^2)^{-1} \int dk k^2 |\eta_k|^2$$

e può assumere valori significativi ancor prima della fine dello SRD. In questo caso si dice che l'interazione (A4.10) causa «preheating».

# Appendice A5

## L'origine delle fluttuazioni

### 1. Metrica e equazioni dinamiche

In questa appendice ci limiteremo al caso di inflazione dovuta a un singolo campo. Rimandiamo a trattati sull'inflazione sia per casi più generali sia per i passaggi, spesso piuttosto complessi, che qui non sono riportati.

Partendo dalla forma più generale della metrica che includa fluttuazioni scalari ci porremo nella «gauge newtoniana» ove, in assenza di sforzi anisotropi, la metrica ha le espressioni

$$(A5.1) \quad ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta = a^2(\tau) [(1 + 2\Phi)d\tau^2 - (1 - 2\Phi)\delta_{ij} dr^i dr^j],$$

(a differenza del capitolo VII, in cui la gauge newtoniana conforme è stata introdotta, qui usiamo  $\tau$  e  $r^i$  in luogo di  $\hat{\tau}$  e  $\hat{r}^i$ , non essendo possibili confusioni con la gauge sincrona) valide in sistemi di riferimento opportunamente definiti.  $\Phi$  è il potenziale gravitazionale da non confondere col campo «inflatonico»

$$(A5.2) \quad \phi \equiv \phi_o(\tau) + \varphi(\tau, r^i)$$

in cui  $\phi_o$  è la parte «di background», cioè il campo  $\phi$  usato nei precedenti paragrafi, mentre le fluttuazioni  $\delta\phi \equiv \varphi$  sono dello stesso ordine del potenziale  $\Phi$ .

In un generico sistema di riferimento, la metrica ha ovviamente un'espressione più complessa di (A5.1): le variabili temporale e spaziali qui introdotte sono quindi «privilegiate»; a sua volta, quindi, sarà legata alla scelta di questo sistema di riferimento privilegiato anche la legge con cui  $\phi$  (o  $\varphi$ ) dipende dalle coordinate spazio-temporali; anche il potenziale  $\Phi$ , che determina la componente  $g_{00}$  del tensore metrico, è ovviamente legato alla scelta del sistema.

Cambiando sistema di riferimento si passa a un'altra gauge gravitazionale, con diverse componenti del tensore metrico e diversa dipendenza di  $\varphi$  da  $x^\mu$ . Esiste peraltro una combinazione di  $\Phi$  e  $\varphi$  che si dimostra essere «gauge-invariante»: essa ha valori e andamento indipendenti dalla scelta della gauge. Pur omettendo ogni dimostrazione, quando arriveremo a definire tale combinazione richiameremo l'attenzione del lettore sulla sua essenziale proprietà di «gauge-invarianza».

Quando non ci siano ambiguità, porremo:  $d/d\tau \equiv ' .$  Definiamo inoltre

$$(A5.3) \quad \mathcal{H} \equiv \frac{1}{a} \frac{da}{d\tau} = \dot{a}$$

e, utilizzando il tempo conforme, le equazioni di Friedmann si possono anche scrivere

$$(A5.4) \quad \mathcal{H}^2 = \frac{8\pi}{3m_p^2} a^2 \rho, \quad -\mathcal{H}' = \frac{4\pi}{3m_p^2} a^2 (\rho + 3p)$$

Anche la dinamica delle fluttuazioni deve essere ricavata dalle equazioni di Einstein. Da esse si può ricavare una equazione di «pseudo»-continuità ( $T_{\nu;\mu}^\mu = 0$ ) per le componenti del tensore energia-impulso, in cui non compare quindi la costante gravitazionale  $G = m_p^{-2}$ . Questa equazione si esplicita nella forma

$$(A5.5) \quad \varphi'' + 2\mathcal{H}\varphi' - \nabla^2\varphi + V''(\phi)a^2\varphi = 4\phi'_o\Phi' - 2V'(\phi)a^2\Phi,$$

e appare quindi come un'equazione del moto per la fluttuazione  $\varphi$  del campo inflatonico. Accanto

a essa vanno considerate le equazioni

$$(A5.6) \quad u'' - \nabla^2 u - \frac{\zeta''}{\zeta} u = 0, \quad \text{dove } u = \frac{m_p^2}{4\pi} \frac{\Phi}{\dot{\phi}_o}, \quad \zeta = \frac{H}{a\dot{\phi}_o} = \frac{\mathcal{H}}{a\dot{\phi}_o},$$

$$(A5.7) \quad \dot{\Phi} + H\Phi = \frac{4\pi}{m_p^2} \phi \dot{\phi}_o.$$

L'eq. (A5.6) riguarda i gradi di libertà del campo gravitazionale che, però, l'eq. (A5.7) pone in relazione coi gradi di libertà del campo inflatonico.

## 2. Campi classici

Il problema si risolve tramite una ridefinizione delle funzioni che descrivono i gradi di libertà del sistema. In luogo di  $\varphi$  dovrà essere quantizzata una variabile di campo che non è la pura fluttuazione del campo inflatonico, ma compone  $\varphi$  con  $\Phi$ . Sarà proprio l'eq. (A5.7) a permettere di definire tale quantità composta, dopo che l'integrale dell'eq. (A5.6) avrà permesso di calcolare il primo membro di eq. (A5.7).

L'integrale di eq. (A5.6) dipenderà da due costanti arbitrarie, da determinare sulla base di opportune «condizioni iniziali». Tra le soluzioni possibili c'è quindi anche quella identicamente nulla: nessuna violazione dell'omogeneità. L'importanza della parte quantistica dell'analisi sta proprio nel forzare le costanti di integrazione a assumere valori non nulli. In un certo senso, la trattazione preliminare dei campi classici serve a definire il contenitore che verrà poi riempito quantisticamente.

Si noti anche che  $G = m_p^{-2}$  compare nell'eq. (A5.6) tramite il termine  $z''/z$ .

Tale equazione viene ora considerata al limite  $k \ll z''/z$ , cioè quando le fluttuazioni sono ben al di fuori del raggio di Hubble. Essa allora si semplifica in  $u''/u = \zeta''/\zeta$  e ha l'integrale generale

$$(A5.8) \quad u = C_1 \zeta + C_2 \zeta \int d\tau / \zeta^2$$

in cui, a tempi grandi, domina il termine in  $C_2$ . Sulla base delle eq. (A5.4), risulta infatti

$$\zeta = \frac{\mathcal{H}}{a\dot{\phi}_o} = \left( \frac{8\pi}{3m_p^2} \right)^{1/2} \frac{1}{a} \left( \frac{\rho}{\rho+p} \right)^{1/2} = \left( \frac{4\pi}{m_p^2} \right)^{1/2} \frac{1}{a} \left( 1 - \frac{\mathcal{H}'}{\mathcal{H}^2} \right)^{-1/2}$$

per cui risulta

$$\begin{aligned} \int \frac{d\tau}{\zeta^2} &= \frac{m_p^2}{4\pi} \int d\tau a^2 \left[ 1 + \left( \frac{1}{\mathcal{H}} \right)' \right] = \frac{m_p^2}{4\pi} \left( \frac{a^2}{\mathcal{H}} - \int d\tau a^2 \right) = \\ &= \frac{m_p^2}{4\pi} \frac{a}{H} \left( 1 - \frac{H}{a} \int dt a \right) = \frac{m_p^2}{4\pi} \frac{a}{H} \frac{d}{dt} \left( \frac{1}{a} \int dt a \right) \end{aligned}$$

e il primo addendo nell'espressione (A5.8) è, grosso modo,  $\propto a^{-1}$ , mentre il secondo è costante. Tenendo conto del solo modo dominante, sarà allora

$$(A5.9) \quad \Phi = \dot{\phi}_o \frac{4\pi^2}{m^2} u = \dot{\phi}_o C_2 \zeta \frac{a}{H} \frac{d}{dt} \left( \frac{1}{a} \int dt a \right) = C_2 \frac{d}{dt} \left( \frac{1}{a} \int dt a \right)$$

siché, nell'approssimazione in cui  $a \propto \exp(Ht)$  e  $H = \text{const.}$ , sarebbe  $\Phi = 0$ . Le espressioni

$$(A5.10) \quad \Phi = C_2 \left( 1 - \frac{H}{a} \int dt a \right), \quad \varphi = \frac{m_p^2}{4\pi} \frac{1}{\dot{\phi}_0} (\dot{\Phi} + H\Phi) = C_2 \dot{\phi}_0 \frac{1}{a} \int dt a,$$

tenendo conto della variazione di  $H$  all'ordine più basso, diventano invece

$$(A5.11) \quad \Phi \simeq -C_2 \frac{\dot{H}}{H^2}, \quad \varphi \simeq C_2 \frac{\dot{\phi}_0}{H}.$$

Si possono anche evitare approssimazioni e eliminare l'integrale dalle due espressioni (A5.10) definendo la quantità

$$(A5.12) \quad \mathcal{R} = \Phi + \frac{H}{\dot{\phi}_0} \varphi = \Phi + \frac{8\pi}{3m_p^2} \frac{\rho\varphi}{\rho+p} = \Phi + \frac{8\pi}{3m_p^2} \frac{\delta\rho}{\rho+p}.$$

Questa è la quantità gauge-invariante preannunciata all'inizio di questa appendice. Si prova che  $\mathcal{R}$  esprime la fluttuazioni (comoventi) della curvatura. Al limite di grandissime lunghezze d'onda, fuori dall'orizzonte, oltre a essere gauge-invariante  $\mathcal{R}$  è dunque costante, coincidendo allora con  $C_2$ .

Nei limiti in cui il contributo di  $\Phi$  a  $\mathcal{R}$  possa essere trascurato o ritenuto costante,  $\delta\rho/(\rho+p)$  è esso stesso una quantità indipendente dal tempo e dalla scelta della gauge gravitazionale. Tale rimane anche al mutare della «sostanza cosmica», alla fine dello SRD, durante il reheating e in tutte le fasi in cui l'evoluzione avviene fuori dell'orizzonte.

### 3. Quantizzazione

I gradi di libertà da quantizzare possono essere direttamente collegati a  $\mathcal{R}$ , introducendo la variabile

$$(A5.13) \quad v = \mathcal{R}/\zeta = \mathcal{R}a\dot{\phi}_0/H = a \left[ \varphi + \Phi(\dot{\phi}_0/H) \right]$$

che, tramite eq. (A5.5) e noto l'andamento di  $\Phi$ , si prova soddisfare l'equazione

$$(A5.14) \quad v'' - \nabla^2 v - (z''/z)v = 0 \quad \text{con} \quad z = 1/\zeta,$$

formalmente ricavabile dall'integrale variazionale

$$(A5.15) \quad S = \int d^4x \mathcal{L} = \int d\tau d^3x \frac{1}{2} [v'' - \partial^i v \partial_i v + (z''/z)v^2]$$

contenente una (densità) lagrangiana  $\mathcal{L}$  formalmente identica a quella di un campo scalare con massa  $m^2 = -z''/z$ . Ci aspettiamo quindi, per il campo  $v$ , una dinamica analoga a quella di un campo scalare nello spazio minkowskiano, in presenza di un campo esterno variabile.

È anche facile verificare che, durante lo SRD, al limite  $\varepsilon, \eta \rightarrow 0$ , risulta

$$(A5.16) \quad z''/z = 2\alpha^2 H^2$$

mentre, all'ordine successivo, avremo

$$(A5.17) \quad z''/z = 2\alpha^2 H^2 (1 + \varepsilon - 3\eta/2).$$

**Verifica dell'eq. (A5.16) a partire dalle eq. (A5.4).** Nel corso dello SRD trascuriamo le variazioni di  $H$  e  $\dot{\phi}_0$  con  $\tau$  e, allo stesso modo, prendiamo  $p \equiv -\rho$ ; risulta allora

$$\frac{z''}{z} \sim \frac{\alpha''}{\alpha} = \mathcal{H}' + \mathcal{H}^2 = -\frac{4\pi}{3m_p^2} a^2 (\rho + 3p) + \mathcal{H}^2 \sim 2\mathcal{H}^2 = 2\alpha^2 H^2.$$

Quindi, al di sotto del raggio di Hubble, il termine  $z''/z$  sarà trascurabile.

Il momento coniugato alla variabile operatoriale  $v$  sarà

$$(A5.18) \quad \pi = \partial \mathcal{L} / \partial v'$$

e dovranno valere le relazioni di commutazione (a  $\tau$  costante)

$$(A5.19) \quad \begin{aligned} [\pi(\tau, \mathbf{x}), v(\tau, \mathbf{y})] &= -i \delta^{(3)}(\mathbf{x} - \mathbf{y}), \\ [v(\tau, \mathbf{x}), v(\tau, \mathbf{y})] &= [\pi(\tau, \mathbf{x}), \pi(\tau, \mathbf{y})] = 0 \end{aligned}$$

Sviluppando la dipendenza spaziale in serie di Fourier si ricava

$$(A5.20) \quad v(\tau, \mathbf{x}) = (2\pi)^{-3/2} \int d^3k (a_k v_k(\tau) e^{-k \cdot \mathbf{x}} + a_k^+ v_k^*(\tau) e^{-k \cdot \mathbf{x}})$$

e l'eq. (A5.14) prescrive che le ampiezze  $v_k^{(*)}$  soddisfino l'equazione

$$(A5.21) \quad v'' + k^2 v - (z''/z)v = 0.$$

Imponendo alle sue soluzioni di soddisfare la condizione (wronskiana)

$$v_k^* v_k' - v_k v_k'^* = -i,$$

le relazioni di commutazione (A5.19) si traducono nelle usuali relazioni di commutazione tra operatori di creazione e distruzione  $a^{(*)}$ :

$$(A5.22) \quad [a_k, a_k^+] = \delta^{(3)}(\mathbf{k} - \mathbf{k}')$$

(gli altri commutatori si annullano).

Il vuoto quantistico si definisce tramite gli operatori di distruzione:

$$(A5.23) \quad a_k |0\rangle = 0.$$

Su scale per cui  $k/a \gg H$ , per cui  $v_k(\tau) \simeq \exp(-ik\tau)$ , il vuoto sarà quindi univocamente definito e tutto avviene come per un campo scalare senza massa in uno spazio di Minkowski.

Sulle grandi lunghezze d'onda, invece, il modo crescente avrà ora comportamento  $v \propto z = \zeta^{-1}$ . Gli operatori di creazione e distruzione, a partire da  $k \simeq aH$  vanno quindi variando nel tempo e lo stato di vuoto al tempo  $\tau$  non è più vuoto al tempo  $\tau + \Delta\tau$ .

L'evoluzione delle ampiezze si ottiene risolvendo allora l'equazione

$$(A5.24) \quad v_k'' + \left( k^2 - \frac{v^2 - 1/4}{\tau^2} \right) v_k = 0$$

che si ottiene da eq. (A5.21) sostituendovi eq. (A5.17), con

$$(A5.25) \quad v = (3/2)[1 + 2\varepsilon - 2\eta/3].$$

Le soluzioni dell'equazione di Bessel (A5.24) vengono convenientemente scritte in termini di funzioni di Hänkel, ottenendo così

$$(A5.26) \quad v_k(\tau) = \frac{\sqrt{\pi}}{2} e^{i(v+1/2)\pi/2} \sqrt{-\tau} H^{(1)}(-k\tau) \quad (\tau < 0, \text{ durante lo SRD})$$

che ridiviene un esponenziale complesso a piccole scale, mentre, a grandi scale, risulta

$$(A5.27) \quad v_k(\tau) \sim e^{i(v+1/2)\pi/2} 2^{v-2} [\Gamma(v)/\Gamma(3/2)] \sqrt{-\tau} (-k\tau)^{-v}.$$

L'elemento critico, in questa soluzione, sta nel fatto che si è trattata  $v$  come costante mentre, in generale,  $\varepsilon$  e  $\eta$  dipendono (blandamente) da  $\tau$ .

#### 4. Funzione spettrale

Dalle componenti  $v_k$  si può risalire ora alle analoghe componenti in uno sviluppo

$$(A5.28) \quad \mathcal{R} = (2\pi)^{-3/2} \int d^3k \mathcal{R}_k e^{-ik \cdot x}$$

da cui, vista la definizione di  $\mathcal{R}$  (eq. A5.12-A5.13), si ricava che deve essere

$$(A5.29) \quad \Delta^2(k) = \frac{k^3}{2\pi^2} |\mathcal{R}_k|^2 = \frac{k^3}{2\pi^2} |v_k \zeta|^2.$$

È immediato quindi vedere che

$$(A5.30) \quad \Delta^2(k) \propto k^{3-2\nu} = k^{-6\epsilon+2\eta}$$

mentre calcoli più dettagliati portano all'espressione

$$(A5.31) \quad \Delta^2(k) = \frac{4}{m_p^4} \frac{H^4}{|H'|^2} \Big|_{k/a=H} = \frac{1}{\pi} \frac{H^2}{m_p^2 \epsilon^2} \Big|_{k/a=H}$$

che qui non ricaveremo.

## CAPITOLO XII

# L'energia oscura

### 1. Espansione accelerata oggi: i problemi

Nel capitolo V abbiamo visto le ragioni osservative che spingono a ritenere che l'Universo sia ben descritto da un modello con parametro di densità totale  $\Omega_o \simeq 1$  e parametro di densità per la materia  $\Omega_{o,m} \simeq 0.2-0.3$ . Ricordiamo che nella «materia» rientrano materia oscura (DM: *dark matter*), di densità  $\rho_c$ , e materia barionica, di densità  $\rho_b$ . La materia barionica è la materia ordinaria, fatta essenzialmente di protoni, neutroni e elettroni, ma che comprende tutte le particelle massive del modello standard delle interazioni elettrodeboli (EW).

Ricordiamo che i valori di  $\Omega_o$  e  $\Omega_{o,m}$  – così come quello del parametro di densità per la componente barionica  $\Omega_{o,b}$  – si ricavano da dati su osservabili astrofisici. Essi richiedono anche che il *gap*  $\Omega_o - \Omega_{o,m}$  sia coperto da una sostanza con parametro di stato  $p/\rho = w \simeq -1$  che si denomina *energia oscura* (DE: *dark energy*). Un discorso a parte meritano i dati sulle SNIa, testimonianza diretta dell'accelerazione nell'espansione, come previsto in presenza di una componente tanto importante con pressione fortemente negativa.

Su questa base, nel capitolo V, abbiamo definito il *modello delle concordanze* o  $\Lambda$ CDM, in cui la DE, dovuta a una sostanza il cui parametro di stato è esattamente  $w \equiv -1$ , copre il 70% dell'energia del cosmo.

Nell'ambito di  $\Lambda$ CDM trovano sistemazione una quantità di dati osservativi. Esso è l'equivalente cosmologico del *modello standard* delle interazioni fondamentali: brutto ma efficiente. Ci sono infatti due problemi concettuali riguardanti la componente di DE:

- 1) Coincidenza.
- 2) *Fine tuning*.

### ***Il problema della coincidenza***

Il problema della coincidenza è illustrato in figura XII.1, dove si confrontano gli andamenti di  $\rho_m$  e  $\rho_A$ , rispettivamente le densità di materia e DE.

In figura XII.1 si vede che la condizione  $\rho_m \sim \rho_A$ , valida oggi a meno di un fattore  $\sim \mathcal{O}(2-3)$ , è tipica della nostra era, in cui si sono da poco formate strutture cosmiche

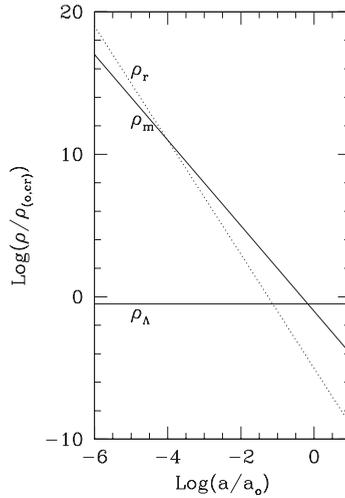


FIGURA XII.1

Andamento delle densità della materia, della DE e della radiazione (rispettivamente:  $\rho_m$ ,  $\rho_A$ ,  $\rho_r$ ) su scale logaritmiche. Già a  $z \sim 10^6$ , la DE è una frazione  $\sim 10^{-18}$  della materia (e  $\sim 10^{-20}$  della radiazione).

come galassie e ammassi. A  $z = a_0/a - 1 \sim 10$ , era  $\rho_A/\rho_m \ll 100$  e a  $z \sim 10^{-6}$ , era  $\rho_A/\rho_m \sim 10^{-18}$ . Se poi andiamo nel «futuro», quando  $a/a_0$  varrà  $\sim 10$ , la densità  $\rho_m$  diviene trascurabile rispetto alla densità della DE che, inesorabilmente, appare destinata a divenire l'unica componente significativa nel cosmo. Tutto ciò deriva dalla legge  $\rho \propto a^{-3(w+1)}$  che, per  $w = 0$  e  $-1$  dà, rispettivamente,

$$(XII.1) \quad \rho_m \propto a^{-3} \quad \text{mentre} \quad \rho_A = \text{const.}$$

Oltre al paradosso che l'osservatore esista *proprio quando*  $\rho_m \sim \rho_A$ , c'è un «drammatico» quesito riguardante lo sviluppo delle disomogeneità.

La crescita lineare delle fluttuazioni nella materia, infatti, è avvenuta in modo  $\propto a$ , solo fin tanto che era  $\rho_m \gg \rho_A$ . Col prevalere di  $\rho_A$ , la crescita «lineare» si è bloccata: le fluttuazioni che oggi hanno ancora un'ampiezza  $\ll 1$ , sono destinate a restare *congelate*, per sempre. Il congelamento non riguarda le fluttuazioni già entrate nelle fasi «non-lineari» di evoluzione. Se una fluttuazione ha ampiezza  $\delta\rho/\rho \sim 1$  oggi, o a  $z \sim 0.1-0.2$  quando la DE era già «dominante», il suo sviluppo continua.

Dunque, nel modello delle concordanze, strutture come gli ammassi di galassie hanno fatto *giusto in tempo* a formarsi. Dopo  $z \sim 0.5$ , ben pochi ammassi si sono potuti staccare dal flusso di Hubble e si sono potute formare galassie quasi solo entro ammassi, già distaccati dal flusso di Hubble, non galassie «di campo», in zone precedentemente buie.

A questo punto resta da domandarsi cosa «sarebbe accaduto» se l'ampiezza delle fluttuazioni primordiali fosse stata  $\sim 1/10$  di quella osservata. Tutto l'orologio dello sviluppo delle strutture, fin tanto che  $\rho_m$  domina, si sposterebbe indietro di un fattore

10 sull'asse  $a$ . In un modello cosmologico in cui non ci fosse DE, tutto ciò significherebbe «solo» un ritardo di un fattore 10 nello sviluppo delle strutture. Un mondo del genere, sia pure nel futuro, sarebbe assai simile al nostro; l'unica differenza significativa riguarderebbe la temperatura del CMB: per un ipotetico osservatore di quel mondo essa sarebbe  $\simeq 0.27$  K.

In presenza della DE, invece, c'è un secondo orologio in competizione con quello che scandisce lo sviluppo delle fluttuazioni, sicché il «rischio» è che  $\rho_A$  superi  $\rho_m$  quando il primo orologio non segna ancora il tempo dell'innesco della non-linearità. Con ampiezze al di sotto dell'unità, in piena «fase lineare», le fluttuazioni si sarebbero congelate per sempre. In un tale modello, il cosmo apparirebbe vuoto e buio, salvo poche isolate stelline, più distanti tra loro di quanto non lo siano oggi gli ammassi di galassie, generate da fluttuazioni ultra-eccezionali.

Il fatto che, al contrario, il cosmo sia abbastanza popolato, sembrerebbe dunque da addebitare a una fortuita *coincidenza*, tra ampiezza delle fluttuazioni primordiali e densità  $\rho_A$ , sebbene quest'ultima, andando all'indietro nel tempo, ci appaia come una presenza trascurabile.

### ***Il problema del fine-tuning***

Non abbiamo fin qui toccato la questione della natura della DE. La dizione  $\Lambda$ CDM si riferisce alla *costante cosmologica*  $\Lambda$ , introdotta da Einstein, per giustificare l'apparente staticità del cosmo pregalattico. Vista in questi termini, la costante cosmologica è una *non*-soluzione di un *non*-problema.

Per Einstein,  $\Lambda$  era una costante attinente la geometria spazio-temporale e questa è ancora la tesi di qualche fisico matematico. La maggior parte dei ricercatori, invece, ritiene si debba individuare la natura fisica della costante  $\Lambda$ , che entra nell'espressione

$$(XII.2) \quad T^{\mu\nu} = \Lambda g^{\mu\nu}$$

del tensore energia-impulso, di una componente cosmica con  $w = -1$ . Come abbiamo visto nel precedente capitolo dedicato all'Inflazione, la presenza di un *falso vuoto* con densità d'energia  $\rho$  avrebbe gli effetti gravitazionali di una costante cosmologica  $\Lambda$ .

È dunque possibile pensare che il vuoto, nel cosmo attuale, non corrisponda al *punto zero* dell'energia, ma a una densità d'energia lievemente positiva?

Per rendersi conto delle difficoltà di una simile ipotesi, si consideri il cosmo  $\sim 10^{-12}$ sec. dopo il *Big-bang*, alla vigilia e durante la transizione EW. Tutto fa ritenere che essa sia stata una transizione di fase del primo ordine e, dunque, per un certo periodo di tempo, magari assai breve, un falso vuoto è veramente esistito, con densità d'energia  $\rho_{EW} \sim T_{EW}^4$ . Passato un breve periodo di *supercooling*, l'energia di tale vuoto si termalizzò e avvenne la transizione al vuoto attuale. Se, tuttavia, anche il vuoto attuale è, in qualche misura, un «falso vuoto», ci si può chiedere quale sia il rapporto  $\rho_A/\rho_{EW}$ .

Figura XII.1 mostra chiaramente che  $\rho_A = \rho_r$  per  $T \simeq 10 T_o$ . Risulta quindi

$$(XII.3) \quad \frac{\rho_A}{\rho_{EW}} \sim \frac{(10 T_o)^4}{T_{EW}^4} \sim \left( \frac{10^{-3} \text{eV}}{10^{11} \text{eV}} \right)^4 \sim 1 : 10^{56} :$$

alla fine della transizione EW, invece di raggiungere un vero vuoto, sarebbe avvenuta una transizione su un livello energetico che è questa minuscola e peculiare frazione del livello energetico di partenza.

Si può ben dire che il modello  $\Lambda$ CDM sembra vanificare in parte i risultati dell'inflazione, in quanto paradigma capace di evitare inspiegabili *fine tuning*.

C'è però una curiosa coincidenza numerica. Il rapporto  $10 T_o/T_{EW}$  è praticamente uguale al rapporto  $T_{EW}/T_{GUT}$ . Un «cosmologo», operante nel periodo di supercooling antecedente la transizione EW, si sarebbe chiesto come mai, invece di azzerarsi dopo la transizione GUT, la densità d'energia del vuoto si fosse fermata a un valore  $\sim (1 : 10^{56})\rho_{GUT}$ . Un *fine tuning* del vuoto, esattamente della stessa entità osservata oggi, non è dunque una novità nella storia del cosmo. Tra questa osservazione e l'inferire che il cosmo oggi sia alla vigilia di una nuova transizione di fase, con rottura di qualche simmetria residua, il passo è però lungo e presumibilmente avventato. Questa «coincidenza» è invece la prima ragione numerica del successo di modelli di DE dinamica con potenziale a *legge di potenza*, nei quali la densità attuale d'energia della DE si riconduce a una scala energetica  $\Lambda \simeq T_{EW}$  (attenzione: questa scala  $\Lambda$  è una quantità del tutto diversa dalla costante cosmologica!)

Sulla base di considerazioni di quantum gravity, si arriva anche a affermare che esistono due *valori naturali* per la densità del vuoto: 0 e  $m_p^4$ . È allora facile calcolare che  $\rho_A/m_p^4 \sim (10 T_o/m_p)^4 = 1 : 10^{128}$ . Taluni autori, nel descrivere il *fine tuning* della costante cosmologica, si riferiscono a questo valore.

## 2. DE dinamica

Come si è passati dalla preistoria dell'inflazione, attribuita al falso vuoto delle GUT, al moderno paradigma inflazionario, basato sulla dinamica di campi scalari, si può pensare anche a DE dovuta a un campo scalare, che non concorre alla formazione di strutture, e chiedersi se tale alternativa alleggerisca i problemi di coincidenza e fine tuning.

Ci sono diverse analogie inflazione e *DE dinamica*, dovuta a un campo scalare  $\phi$ , anche se le due ipotesi si collegano a problematiche assai distanti.

Iniziamo comunque dagli elementi di similarità, che riguardano anzitutto densità d'energia e pressione

$$(XII.4) \quad \rho_{de} = \rho_k + V(\phi), \quad p_{de} = \rho_k - V(\phi)$$

in cui  $\rho_k$  è la (densità d')energia cinetica di  $\phi$  (per la sua espressione differenziale vedi appendice A1) e  $V(\phi)$  è la (densità d')energia potenziale del campo, risultante dalla sua

auto-interazione. Ne risulta, in analogia con quanto detto nel caso dell'inflazione, che il parametro di stato della DE,  $w_{de} = p_{de}/\rho_{de}$ , garantirà un'espansione accelerata se  $\Omega_{de} w_{de} < -1/3$  ( $\Omega_{de}$ : parametro di densità della DE) e quindi

$$(XII.5) \quad \frac{\rho_k}{V} < \frac{3\Omega_{de} - 1}{3\Omega_{de} + 1}.$$

Ad esempio, per  $\Omega_{de} = 0.72$ , deve essere  $\rho_k < 0.367 V$ . Valori  $w_{de} \sim -0.9$ , o minori, si possono poi raggiungere solo se  $\rho_k \gtrsim 0.1 V$ .

Di fatto gli elementi di similarità finiscono qui. Per la DE dinamica, infatti, si pone subito un problema di *condizioni iniziali* che non ha gran peso nel caso del campo inflatonico. Il fatto è che il campo  $\phi$  della DE si è evoluto per decine di miliardi d'anni, mentre il campo inflatonico evolve a ridosso del tempo di Planck o di una transizione di fase, che fissano il suo valore iniziale. Per questa ragione, nello studio della DE dinamica, grande peso ha avuto la scoperta dei *potenziali traccianti* (*tracker/tracking potentials*).

Se  $V(\phi)$  è un potenziale tracciante, valore e andamento attuali del campo  $\phi$  sono indipendenti dalle condizioni iniziali: praticamente *ogni scelta di condizioni iniziali porta a uguali condizioni dinamiche* in tempi recenti. Questa proprietà si può esprimere in termini fisico matematici affermando che, per determinati potenziali  $V(\phi)$ , ci sono soluzioni delle equazioni dinamiche che costituiscono degli *attrattori*. Se il campo  $\phi$  viene posto al di fuori dell'attrattore, esso è comunque destinato a riconvergere su di esso in tempi brevi.

Considereremo qui due possibili potenziali traccianti, simili a alto  $z$ , ma che, a basso *redshift*, forniscono comportamenti ai limiti opposti:

$$\begin{aligned} \text{Ratra - Peebles (RP):} \quad & V(\phi) = \frac{A^{4+\alpha}}{\phi^\alpha} \\ \text{SUGRA:} \quad & V(\phi) = \frac{A^{4+\alpha}}{\phi^\alpha} \exp\left(\frac{4\pi\phi^2}{m_p^2}\right) \end{aligned}$$

$A$  è una costante che ha le dimensioni di un'energia e *non va confusa con la costante cosmologica*, mentre l'esponente  $\alpha$  è una costante numerica. Nel primo caso, la denominazione del potenziale risale agli autori che l'hanno proposto; nel secondo, il potenziale è legato a teorie di supergravità. In figura XII.2 sono riportati i grafici dei due potenziali, per il caso  $A = 10^{-15} m_p$ ,  $\alpha = 4$ . I grafici pongono in rilievo come i due potenziali appaiano identici fin tanto che  $\phi$  non approssima  $m_p$  e il termine esponenziale assuma valori significativi.

Per il potenziale RP, si dimostra che il rapporto tra  $\rho$  e  $V$  sulla soluzione tracciante è

$$(XII.6) \quad \frac{\rho}{V} = 1 + \frac{1}{2} \frac{\alpha}{1 + (2u - 1)/\beta}$$

in cui  $\alpha$  viene dall'espressione del potenziale, mentre  $\beta$  e  $u$  dipendono anche dal para-

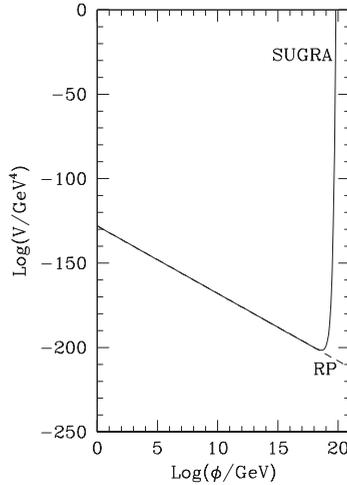


FIGURA XII.2

Andamento dei potenziali traccianti SUGRA (linea continua) e RP (linea tratteggiata), per  $\Lambda = 10^{-15}m_p$ ,  $\alpha = 4$ .

metro di stato della sostanza dominante e dal conseguente regime espansivo:

	$w$	regime espansivo	$\beta$	$u$
(XII.7)	$1/3$	$\rho a^4 = \text{const.}$	$4/(2 + \alpha)$	1
	0	$\rho a^3 = \text{const.}$	$6/(2 + \alpha)$	2
	$-2/3$	$\rho a = \text{const.}$	$-1/(2 + \alpha)$	-2

Un'espressione interpolatoria tra regime espansivo dominato dalla radiazione e regime dominato dalla materia, sarà fornita più avanti. Come discusso più oltre, queste relazioni non sono direttamente applicabili quando la sostanza dominante diviene la DE stessa. La stessa relazione vale ovviamente anche per un potenziale SUGRA, fin tanto che il suo andamento non si distanzia da RP.

Queste relazioni confermano che, assumendo proporzionalità tra  $\rho$  e  $V$ , con un fattore di proporzionalità  $\sim \mathcal{O}(1)$ , non si commette un errore nell'ordine di grandezza (per la dimostrazione di eq. (XII.6)-(XII.7) vedi appendice A1 e, in particolare eq. A1.14). Come ordine di grandezza, possiamo anzi assumere che la densità d'energia della DE coincida con  $V$ .

Abbiamo visto, alla fine del paragrafo precedente (eq. XII.3), che è

$$(XII.8) \quad \rho_{o,de}/T_{EW}^4 \sim T_{EW}^4/T_{GUT}^4.$$

Tenendo conto che  $m_p$  è di poco superiore a  $T_{GUT}$ , è allora evidente che il valore di  $\alpha$  che soddisfa l'equazione

$$(XII.9) \quad V(\phi) \sim \rho_{o,de} = T_{EW}^{4+\alpha}/m_p^2$$

non sarà molto lontano da 4. Chiaramente, da un potenziale RP, con  $\alpha \sim 4$  e  $\Lambda \sim 10^2\text{-}10^4\text{GeV}$ , si ottiene una densità  $\rho_{o,de} \sim \rho_\Lambda$  se  $\phi \sim m_p$  (vedi comunque la relazione (XII.22) qui sotto).

Viceversa, se un potenziale RP, con simili valori per l'energia  $\Lambda$  e l'esponente  $\alpha$ , deve rendere conto della DE, occorre che la *soluzione tracciante*, alla fine, porti  $\phi$  a valere  $\sim m_p$  per l'epoca attuale. Il caso di un potenziale SUGRA non è molto diverso, visto che il termine esponenziale implica una correzione che si estrinseca solo quando  $\phi$  ha effettivamente raggiunto valori  $\sim m_p$ .

La conclusione è che, almento in apparenza, la DE dinamica permette di evitare il *fine tuning*: nel potenziale entra solo una scala energetica che potrebbe essere quella EW o quella a cui avviene la rottura delle SUSY; l'esponente  $\alpha$ , a sua volta, avrà valori ragionevolmente compresi tra 2 e 6.

### 3. Evoluzione della DE dinamica

Passando da valutazioni di ordine di grandezza a calcoli precisi, possiamo dare un quadro delle difficoltà che la DE dinamica introduce, dal punto di vista analitico, rispetto a cosmologie senza DE o modelli in cui la DE avesse equazione di stato

$$(XII.10) \quad p = w\rho \quad \text{con } w \simeq -1, \text{ ma } \neq 1 \text{ e costante.}$$

A queste conclusioni si giunge facilmente osservando il sistema di equazioni differenziali da risolvere, senza bisogno di affrontarne l'integrazione.

Le equazioni da risolvere, per trovare l'andamento del fattore di scala col tempo, sono in sostanza le seguenti:

$$(XII.11) \quad d\rho/da = -3(\rho + p)/a,$$

$$(XII.12) \quad [(da/dt)/a]^2 = (8\pi G/3)\rho.$$

Le incognite di queste equazioni differenziali sono le funzioni  $\rho(a)$  e  $a(t)$  che esprimono rispettivamente l'andamento della densità (di ciascuna sostanza cosmica) rispetto al fattore di scala  $a$  e di quest'ultimo rispetto al tempo.  $d\rho/dt$  e  $da/dt$  sono derivate di queste funzioni, ma non ci interessa qui sapere come si esegue l'operazione che fornisce la derivata di una funzione.

Se risulta  $p = w\rho$ , per tutte le sostanze presenti nel cosmo, l'eq. (XII.11) riguarda le sole variabili  $\rho$  e  $a$  (densità e fattore di scala) e si può integrare ottenendo l'andamento delle funzioni  $\rho(a)$ , per ciascuna sostanza cosmica. Va qui opportunamente sottolineato che c'è un'eq. (XII.11) per ciascuna sostanza presente nel cosmo; in generale le diverse sostanze differiscono per il valore di  $w$ .

A questo punto, a secondo membro di eq. (XII.12), si inseriscono gli andamenti di  $\rho$  con  $a$ , per le varie sostanze. L'eq. (XII.12) contiene così le sole variabili  $a$  e  $t$  e se ne può quindi ricavare l'andamento di  $a$  con  $t$ .

Note le funzioni  $a(t)$  e (dalla precedente equazione)  $\rho(a)$ , si ricavano anche gli andamenti della densità di ciascuna sostanza in funzione del tempo.

In presenza di DE dinamica la situazione si complica perché l'eq. (XII.11) per la DE è equivalente all'equazione di propagazione del campo  $\phi$  :

$$(XII.13) \quad \frac{d^2\phi}{dt^2} + \frac{3}{a} \frac{da}{dt} \frac{d\phi}{dt} + V'(\phi) = 0$$

(vedi appendice A2 del precedente capitolo sull'inflazione). Qui intervengono derivate seconde, si deve considerare la derivata del potenziale  $V$ , ecc. Tutte queste complicazioni sarebbero superabili, se non avvenisse un fatto determinante: nell'eq. (XII.13) compaiono simultaneamente le 3 variabili  $\phi$ ,  $a$ ,  $t$ : non è possibile isolare in essa  $\phi$  e  $a$ , come si faceva per  $\rho$  e  $a$  nell'eq. (XII.11).

Né, d'altro canto, è possibile isolare in essa le variabili  $\phi$  e  $t$ . Ciò che osta è la necessità di conoscere, per scrivere il secondo addendo, la dipendenza di  $a$  da  $t$ . Viceversa, per sapere come dipende  $a$  da  $t$  occorre integrare l'eq. (XII.12), ciò che presupporrebbe la conoscenza di come  $\rho_{de}$  e  $p_{de}$  (vedi eq. XII.4) dipendono dal tempo tramite  $\phi$ .

In un solo caso è possibile trovare una via d'uscita: quando  $\rho_{de}$  e  $p_{de}$  sono trascurabili, rispetto a materia o radiazione, nello stabilire la legge d'espansione del cosmo. Allora l'eq. (XII.13) si può integrare per ultima, quando la connessione tra  $a$  e  $t$  è già nota.

Altrimenti le due equazioni (XII.11)-(XII.12) costituiscono un sistema che va trattato nel suo insieme con tecniche numeriche.

### Soluzioni traccianti

Quando invece la radiazione o la materia sono dominanti, si ottengono le *soluzioni traccianti* per il campo  $\phi$ . In appendice A1 si ricava che l'andamento di  $\phi$  col tempo «conforme»  $\tau$  si può esprimere tramite leggi di potenza, finché la DE è sottodominante.

Nelle ere dominate dalla radiazione e dalla materia non-relativistica è, rispettivamente,  $t \propto \tau^2$  e  $t \propto \tau^3$ ; in tali ere risulta:

$$(XII.14) \quad \phi = \phi_i (\tau/\tau_i)^\beta$$

con  $\beta$  dato da eq. (XII.7); il tempo conforme  $\tau_i$  è un tempo di riferimento, a cui il campo ha valore

$$(XII.15) \quad \phi_i = A \left[ \frac{\alpha(a_i \tau_i A)^2}{\beta^2 + \beta(2u - 1)} \right]^{1/(\alpha+2)}$$

(vedi eq. (A1.7);  $a_i$  è il fattore di scala al tempo  $\tau_i$ ). Queste soluzioni valgono per entrambi i potenziali RP che SUGRA.

In eq. (XII.7), oltre ai parametri di stato  $w = 0$  e  $w = 1/3$  (materia ordinaria e radiazione) per la sostanza dominante, si considera anche il valore  $w = -2/3$ , cui

corrisponde l'andamento  $\rho \propto a^{-1}$ . Discuteremo tra breve il significato euristico di questo caso anomalo.

Noto  $\beta$ , la (densità d')energia cinetica del campo  $\phi$  sarà

$$(XII.16) \quad \rho_k = \frac{\dot{\phi}^2}{2} \left( \frac{\beta}{\tau} \right)^2$$

e si può quindi ricavare il parametro di stato

$$(XII.17) \quad w_{de} = \frac{\rho_k - V}{\rho_k + V}$$

della DE che dipende da  $\beta$  e, quindi, dal regime di espansione. Si incontra quindi qui una prima anomalia: mentre le equazioni di stato di componenti relativistiche o non-relativistiche dipendono solo dalle proprietà della sostanza, l'equazione di stato della DE cambia, se cambia la legge di espansione.

In figura XII.3 si mostrano i valori di  $w$ , per la DE, nei casi di espansione dominata da sostanze la cui densità dipende dal fattore di scala con leggi  $\rho \propto a^{-c}$  ( $c = 4, 3, 1$ ). In ascissa abbiamo i valori dell'esponente  $\alpha$  del potenziale.

In generale si vede che i valori di  $w$  risultano comunque negativi. Valori  $w \sim -1$ , tuttavia, si hanno solo per  $c = 1$ , quando la sorgente dell'espansione ha parametro di stato  $w = -2/3$ .

Supponiamo allora, a titolo d'esempio, la presenza di due componenti con:  $w = 0$  e  $\Omega_m = 0.27$ ;  $w = -0.9$  e  $\Omega_{de} = 0.73$ . La pressione totale sarà allora

$$(XII.18) \quad p_{tot} = -0.9 \Omega_{de} \rho_{tot} = -0.667 \rho_{tot} \simeq -(2/3) \rho_{tot}$$

In apparenza si ha quindi una situazione autoconsistente: se la DE contribuisce per

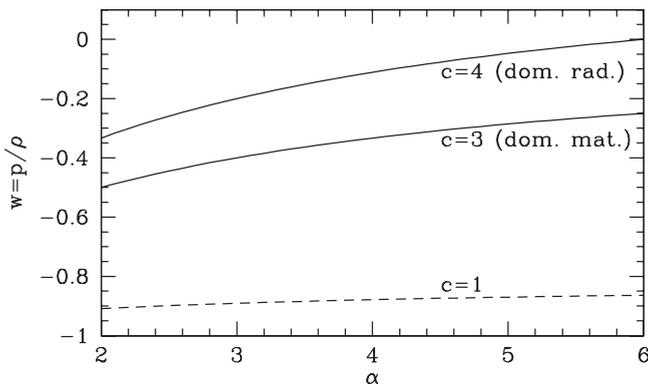


FIGURA XII.3

Parametro di stato della DE dinamica, dovuta a un campo scalare  $\phi$  autointeragente tramite un potenziale  $\propto \phi^{-\alpha}$ , nel corso di espansione dominata da radiazione, materia non relativistica e da una sostanza con parametro di stato  $w = -2/3$ .

73% alla sostanza cosmica e il resto è materia non relativistica, il parametro di stato complessivo è  $-2/3$  e questo parametro di stato dà una DE con parametro di stato  $-0.9$ .

Questa coincidenza vale solo da un punto di vista euristico: la coincidenza riscontrata è solo un transiente. Se anche, a un determinato istante, il rapporto  $\rho_m/\rho_{de}$  permette di fare questa estrapolazione, con l'evolvere del fattore di scala il rapporto  $\rho_m/\rho_{de}$  cambia. In un intorno del punto considerato, anche se continuiamo a supporre che sia  $w_{de} = -0.9$ , si ha

$$(XII.19) \quad \rho_m \propto a^{-3}, \quad \rho_{de} \propto a^{-0.3}, \quad \rho_{de}/\rho_m = (0.27/0.73) \times a^{2.7}$$

e il rapporto tra le densità cambia rapidamente. Non si tratta solo di aggiornare il ragionamento variando, in funzione del parametro di scala, il valore di  $w_{de}$  e del corrispondente esponente  $c$ . Il punto è che la relazione tra  $c$  e  $w_{de}$  è corretta solo per  $w_{de}$  costante, mentre abbiamo constatato che il rapporto  $\rho_{de}/\rho_m$  varia.

Di fatto, comunque, non esiste un procedimento analitico per la soluzione complessiva del problema definito dalle equazioni differenziali (XII.11), (XII.12), (XII.13), quando la DE contribuisce significativamente alla sorgente dell'espansione cosmica. In questo caso il problema si risolve solo per via numerica. In figura XII.4 mostriamo l'evoluzione delle densità delle diverse componenti, ottenuta numericamente, per un potenziale RP con  $\Lambda = 1000$  GeV e  $\Omega_{de} = 0.73$ . Vengono forniti separatamente energia cinetica e potenziale della DE che, come si vede, sono molto prossime tra loro durante la maggior parte dell'evoluzione, per distanziarsi solo dopo l'equivalenza materia-radiazione, quando la sostanza cosmica dominante evolve in modo  $\propto a^{-3}$ , e soprattutto

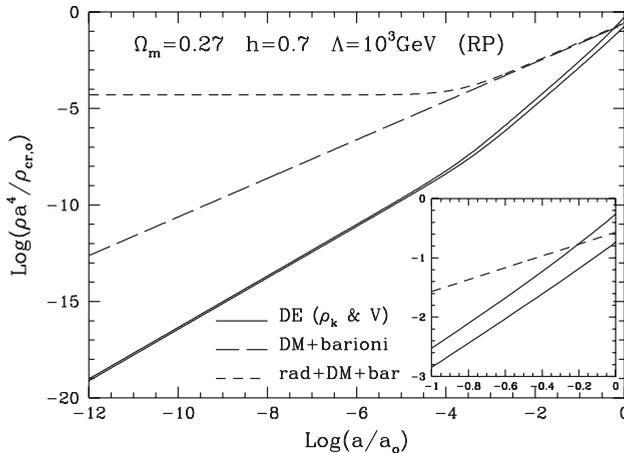


FIGURA XII.4

Evoluzione delle densità delle diverse componenti cosmiche in un modello con DE dinamica ottenuta da un potenziale RP. Nel riquadro viene magnificato l'andamento all'avvicinarsi dell'epoca attuale. Tra le due curve continue, quella superiore descrive l'andamento dell'energia potenziale.

quando da DE diviene una componente significativa della sorgente dell'espansione. Naturalmente, la curva più elevata delle due curve continue rappresenta la (densità d')energia potenziale e ciò fa sì che  $w$  divenga fortemente negativo.

Grafici analoghi descrivono l'evoluzione delle diverse componenti al variare dei parametri cosmologici e in altri modelli con DE dinamica.

A questo proposito è importante notare che, dei 3 parametri  $\Omega_{o,de}$ ,  $\alpha$  e  $A$ , solo 2 possono essere arbitrariamente fissati. In effetti, finché la DE è subdominante l'andamento di  $\phi$  è dato con ottima approssimazione dall'espressione

$$(XII.20) \quad \phi = \phi_i (\tau a / \tau_i a_i)^b \quad \text{con} \quad b = 2/(2 + a).$$

che interpola efficacemente le soluzioni di traccia nelle ere dominate dalla radiazione e dalla materia non relativistica (vedi eq. A1.16 in appendice A1);  $\phi_i$  è dato da eq. (XII.15) con  $\beta = 2b$ , mentre l'espressione

$$(XII.21) \quad u = \frac{1 + (\sqrt{2} - 1)\tau/\tau_e}{1 + (\sqrt{2} - 1)(\tau/\tau_e)^2}$$

ridà i valori in eq. (XII.7) quando una delle due sostanze è trascurabile ( $\tau_e$  è il valore del tempo conforme all'equivalenza). Quindi, una volta assegnati  $\alpha$  e  $A$ ,  $\rho_k$  e  $V$  sono ricavabili analiticamente dall'andamento (XII.20) di  $\phi$ .

La novità, quando la DE diviene dominante, è l'impossibilità di continuare a seguire l'evoluzione di  $\phi$  per via analitica, ma l'andamento di  $\rho_k$  e  $V$  resta univocamente fissato, come esemplificato in figura XII.4. Ciò fissa definitivamente il valore di  $\Omega_{o,de} = (\rho_{o,k} + V_o)/\rho_{o,cr}$  nell'epoca attuale.

Per modificare  $\Omega_{o,de}$  bisogna perciò cambiare  $\alpha$  o  $A$ . Quindi, se si vuole scegliere a priori  $\Omega_{o,de}$ , si può scegliere arbitrariamente solo uno dei 2 parametri del potenziale.

Con ottima approssimazione, per i modelli SUGRA che sono i meglio aderenti ai dati osservativi, e per  $h = 0.7$ ,  $\Omega_{o,m} = 0.27$ , vale la relazione lineare

$$(XII.22) \quad \alpha = 3.2 + 0.6 \log (A/10^2 \text{ GeV})$$

per  $\log (A/10^2 \text{ GeV})$  compreso tra  $-1$  e  $5$ . Variando  $\Omega_{o,de}$  e  $h$  questa relazione va opportunamente adeguata.

#### 4. Modelli cosmologici con DE dinamica: confronto coi dati

Una volta noti gli andamenti delle densità di *background* l'analisi di un modello cosmologico richiede anche lo studio delle fluttuazioni di densità. Su questo rimandiamo alla letteratura specializzata. Forniremo qui alcune informazioni sul successo che i modelli incontrano rispetto ai diversi osservabili astrofisici.

Un'importante differenza tra modelli  $\Lambda$ CDM e con DE dinamica è la possibile presenza di fluttuazioni nella DE dinamica. La differenza ha origine da termini in cui è presente a fattore la somma  $p + \rho$ , il cui contributo è nullo nel caso  $w = -1$  (e solo in esso).

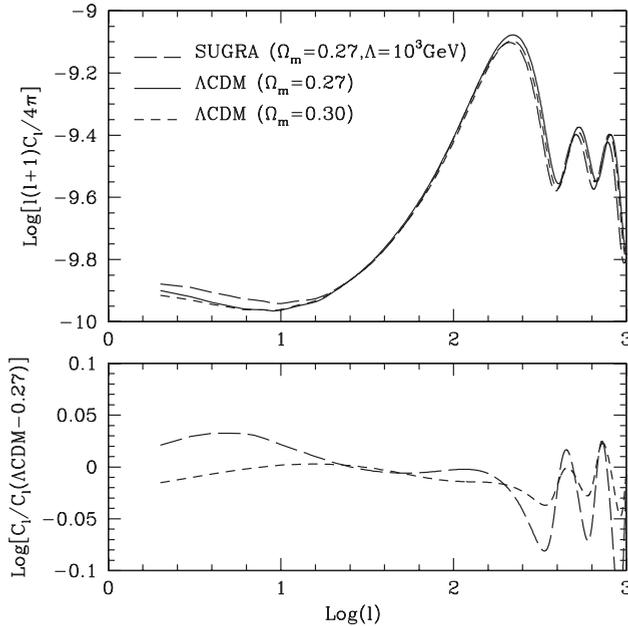


FIGURA XII.5

Spettro delle anisotropie del CMB, in modelli  $\Lambda$ CDM con diversi valori di  $\Omega_{de}$  e in un modello SUGRA. Nel riquadro inferiore sono riportate le discrepanze tra  $C_l$ .

*Anisotropie del CMB.* Nello studio delle anisotropie del CMB, per esempio, è essenziale tenere conto delle fluttuazioni nella DE. Fuori dall'orizzonte, esse sono presenti su tutte le scale. Una volta che le fluttuazioni sono entrate nell'orizzonte, il valore negativo di  $w_{de}$  fa sì che siano rapidamente soppresse. Ciò implica peraltro modifiche sulle ampiezze all'orizzonte su tutte le scale, particolarmente visibili su quelle che entrano nell'orizzonte dopo la ricombinazione. In figura XII.5 confrontiamo lo spettro angolare delle anisotropie in modelli  $\Lambda$ CDM e SUGRA.

Le discrepanze tra modelli permettono di raggiungere alcuni risultati specifici:

- 1) I modelli RP hanno notevoli difficoltà a rendere conto dei dati CMB e possono adattarsi solo per valori di  $\Lambda$  molto al di sotto del GeV, non più in relazione con alcuna scala microfisica.
- 2) I modelli SUGRA, invece, sono addirittura (lievemente) in miglior accordo coi dati, rispetto a  $\Lambda$ CDM. Anche in questo caso i risultati migliori si ottengono con bassi valori di  $\Lambda$ , sebbene non sia necessario scendere al di sotto del GeV.

*Struttura su grande scala.* I modelli di DE dinamica non offrono, rispetto a osservabili concernenti la struttura su grande scala, né vantaggi né svantaggi. In generale, poiché la DE raggiunge valori apprezzabili prima che in modelli  $\Lambda$ CDM, le strutture che si osservano dovrebbero essere più antiche. Ricordiamo infatti che, dopo che la DE di-

viene prevalente, lo sviluppo lineare delle fluttuazioni si congela. Occorre quindi che le fluttuazioni primordiali abbiano un'ampiezza (lievemente) maggiore. Inoltre, il tempo più lungo a disposizione fa sì che le concentrazioni dei sistemi legati siano (lievemente) più elevate.

SN Ia. I dati più recenti sulle SN Ia sembrano favorire i modelli  $\Lambda$ CDM rispetto a modelli con DE dinamica. Si tratta però di un vantaggio marginale ed entrambi i modelli sono ampiamente compatibili coi dati.

### **Parametri osservabili**

Il confronto coi dati mette tuttavia in luce quello che è forse il problema più significativo nello studio della DE: la difficoltà a porre limiti osservativi veramente stringenti alla forma del potenziale.

Per RP il parametro di stato  $w_{de}$  varia lentamente e il suo valor medio dipende da  $A/\text{GeV}$ ; i valori di  $A/\text{GeV} \ll 1$  sono preferiti proprio perché danno  $w_{de}$  più prossimo a  $-1$ . Per SUGRA si ha un andamento opposto: quando  $\phi$  inizia a «scalare» la salita determinata dal fattore esponenziale,  $w_{de}$  varia rapidamente, calando da valori  $\sim -0.6$  a valori  $\sim -0.9$ .

Ma, in generale, dato un generico andamento di  $w_{de}(a)$ , è possibile invertire il sistema di equazioni e determinare un potenziale  $V(\phi)$  che lo causi. Né vale il vincolo che il potenziale sia tracciante. Infatti, la forma «osservata» del potenziale potrà poi essere interpolata con un potenziale tracciante a *redshift* maggiore, dove  $\Omega_{de}$  è così basso che non esistano osservabili accessibili.

In questo contesto la ricerca sulla DE si è concentrata sulle tecniche che permettano di ricavare  $w_{de}(a)$  dai dati. Vari *ansatz* sulla forma di  $w_{de}(a)$  sono stati proposti, a partire da  $w_{de} = \text{const}$ . Vi sono attualmente pochi osservabili astrofisici che permettano di porre vincoli a  $w_{de}$  e nessun osservabile in grado, per esempio, di valutarne la pendenza. Ma la situazione è in evoluzione e molti esperimenti oggi in preparazione potrebbero porre vincoli significativi.

Ci pare opportuno affermare che, se la DE è di natura dinamica e soddisfa equazioni del tipo qui considerato, non sarà possibile determinarne in modo definitivo l'equazione di stato per via osservativa. Le osservazioni, tuttavia, potranno porre limiti cogenti a opportuni modelli motivati dalla microfisica.

## **5. Modelli cosmologici con DE dinamica: problemi concettuali aperti**

Abbiamo visto come, nel caso dei modelli inflazionari, si assista al confronto tra una scuola astrofisica e una scuola particellare. La prima mantiene lo studio al più elevato livello di generalità, mirando a porre vincoli astrofisici al *range* di modelli ammissibili. La seconda usa il paradigma inflazionario per verifiche di specifiche ipotesi fisiche. Nel caso della DE, il precedente paragrafo ha messo in luce le difficoltà di una prospettiva

esclusivamente astrofisica. In questo paragrafo si vogliono circoscrivere le difficoltà concettuali di una DE-non- $\Lambda$ .

In primo luogo dobbiamo constatare che la DE dinamica sembra risolvere il problema del *fine-tuning*, certo non quello della *coincidenza*: Il tempo attuale resta sempre «privilegiato». Solo alla vigilia di oggi il DE diviene una componente significativa del cosmo. A proposito basta rivedere figura XII.4. Tra gli altri problemi resta quindi aperto quello del «tuning» dell'ampiezza delle fluttuazioni che, altrimenti, rischiano di non produrre strutture osservabili.

Nel tentativo di superare anche il problema della coincidenza sono state proposte teorie che prevedono un accoppiamento tra DE e DM. In effetti, tra i paradossi dei modelli cosmologici, c'è la necessità di due componenti oscure con caratteristiche diverse: non potrebbero esse essere l'espressione di un'unica realtà microfisica? Il cosmo sarebbe così dotato di un *lato oscuro* entro cui agirebbero forze, in analogia a quelle che intervengono nel *lato luminoso*.

In questo caso sarebbe possibile un flusso continuo di energia dalla DM alla DE. La densità della DM calerebbe più rapidamente; la densità della DE, soggetta a input continuo, avrebbe un andamento prossimo a quello della DM.

Si noti che i vincoli all'interazione DM-DE sono tutti di natura astrofisica. Le forze ipotizzate sono di intensità non indagabile in laboratorio. Il confronto con osservabili astrofisici ha portato a importanti restrizioni sul livello e sull'andamento dell'accoppiamento che, sebbene non escluso, resta permesso a un livello tale da non risolvere (appieno) il problema della coincidenza.

Un'ultima questione concettuale riguarda la scelta di descrivere il campo  $\phi$  come... campo. Ad esempio, il campo *pionico*, che ha come quanti i mesoni  $\pi$ , è anch'esso un campo scalare; la sua pressione e la sua densità d'energia sono date dalle espressioni (XII.4). La sua realtà fenomenologica è tuttavia fatta di quanti e una distribuzione di mesoni  $\pi$  in equilibrio statistico ha una pressione comunque positiva (anche se difficilmente raggiungerà un regime  $p \simeq \rho/3$ , in quanto, a temperature  $T \sim m_\pi$ , i quark del  $\pi$  diventano particelle autonome).

Per la DE, invece, non si passa a una descrizione in termini di quanti. Perché ciò sia fisicamente ragionevole occorre che la massa dei quanti del campo  $\phi$  sia talmente piccola da renderne difficile la *localizzazione*, al confronto con le dimensioni dell'orizzonte. Il *fine-tuning*, cacciato dalla porta, rientra quindi dalla finestra della massa infinitesima – ma non nulla – per i campi del campo.

L'insieme di queste considerazioni porta alla conclusione che, tra i problemi di fisica cosmologica oggi aperti, quello della natura della DE è uno dei più lontani dalla soluzione. È probabile che la nuova fisica messa in evidenza dalla presenza di DE resterà ancora a lungo oggetto di indagine.

# Appendice A1

## Soluzioni traccianti con potenziali a legge di potenza

Nello studio della DE conviene utilizzare il «tempo conforme»  $\tau$ . Se, in una cosmologia a sezione spaziale piatta, un'unica componente di densità  $\rho$  è sorgente dell'espansione, l'equazione di Friedmann diviene

$$(A1.1a) \quad (a'/a)^2 = (8\pi/3m_p^2)\rho a^2 = H_i^2 a_i^2 (a_i/a)^{c-2}$$

dove  $' = d/d\tau$ ,  $a_i$  è il fattore di scala a un tempo  $\tau_i$  «di riferimento» e si è posto

$$\rho = \rho_i (a_i/a)^c \quad H_i^2 \equiv (8\pi/3m_p^2)\rho_i;$$

sarà, per esempio,  $c = 3$  (4) se  $\rho$  è densità di materia (radiazione).

Posto  $R = a/a_i$ , l'eq. (A1.1) diviene:  $dR R^{c/2-2} = H_i a_i d\tau$  e ha integrale

$$(A1.2) \quad a/a_i = [(c/2 - 1)(H_i a_i \tau_i)]^{1/(c/2-1)} (\tau/\tau_i)^{1/(c/2-1)}.$$

Accanto a materia e radiazione conviene considerare un caso con  $c = 1$ . L'equazione di Friedmann

$$(A1.1b) \quad d\rho/\rho = -3(1 + \bar{w})da/a$$

(valida per  $\bar{w}$  costante), ci dice che  $\rho \propto a^{-3(1+\bar{w})}$ ; quindi,  $c = 1$  significherebbe che la sostanza dominante ha parametro di stato  $\bar{w} = -2/3$ . Consideriamo questo caso senza annettervi alcun significato fisico, ai fini della discussione.

Utilizzando eq. (A1.2) è facile vedere che risulta:

$$(A1.3) \quad \begin{aligned} c = 4 \text{ (radiazione):} & \quad (a/a_i) = (H_i a_i \tau_i)(\tau/\tau_i), & a'/a = u/\tau \text{ con } u = 1; \\ c = 3 \text{ (materia):} & \quad (a/a_i) = (H_i a_i \tau_i/2)^2 (\tau/\tau_i)^2, & a'/a = u/\tau \text{ con } u = 2. \\ c = 1: & \quad (a/a_i) = (-H_i a_i \tau_i/2)^2 (\tau/\tau_i)^{-2}, & a'/a = u/\tau \text{ con } u = -2. \end{aligned}$$

Queste soluzioni vanno essere inserite nell'eq. del moto del campo scalare  $\phi$ :

$$(A1.4) \quad \phi'' + 2(u/\tau)\phi' + a^2 V'_\phi = 0$$

(vedi appendice A2 del precedente capitolo sull'inflazione) che considereremo per il solo caso

$$(A1.5) \quad V = A^{\alpha+4}/\phi^\alpha$$

in quanto il fattore esponenziale del potenziale SUGRA è  $\sim 1$  finché  $\phi/m_p \ll 1$  e non ha rilievo nelle fasi iniziali dell'espansione.

Non daremo qui alcuna dimostrazione del fatto che la soluzione tracciante deve avere la forma

$$(A1.6) \quad \phi = \phi_i (\tau/\tau_i)^\beta$$

e verificheremo invece che

$$(A1.7) \quad \beta(2 + \alpha) = 2c/(c - 2), \quad \phi_i = A \left[ \frac{\alpha(a_i \tau_i A)^2}{\beta^2 + \beta(2u - 1)} \right]^{1/(\alpha+2)}$$

per cui, se

$$(A1.8) \quad \begin{array}{ll} c = 4 \text{ (radiazione)} & \beta = 4/(2 + \alpha) \\ c = 3 \text{ (materia)} & \beta = 6/(2 + \alpha) \\ c = 1 & \beta = -1/(2 + \alpha) \end{array}$$

In effetti, l'eq. (A1.4) per il potenziale (A1.5), utilizzando l'espressione (A1.6) di  $\phi$ , diviene

$$(A1.9) \quad (\beta^2 - \beta + 2u\beta)\phi/\tau^2 = \alpha a^2 A^{4+\alpha}/\phi^{1+\alpha}$$

ovvero, sostituendo l'espressione (A1.6) di  $\phi$ , si ricava

$$(A1.10) \quad \begin{aligned} \beta^2 + \beta(2u - 1) &= \alpha (a\tau)^2 A^{4+\alpha}/\phi^{2+\alpha}, \\ \phi_i^{2+\alpha}(\tau/\tau_i)^{\beta(2+\alpha)} &= \alpha (a\tau)^2 A^{4+\alpha}/[\beta^2 + \beta(2u - 1)]. \end{aligned}$$

Ponendo  $\tau = \tau_i$ ,  $a = a_i$ , si ricava

$$(A1.11) \quad \phi_i^{2+\alpha} = A^{2+\alpha}(a_i\tau_i A)^2 \alpha / [\beta^2 + \beta(2u - 1)]$$

e, di qui, subito la seconda delle (A1.7). Tenendo poi conto che l'eq. (A1.2) stabilisce la legge  $a \propto \tau^{1/(c/2-1)}$ , è immediato ricavare la prima delle (A1.7).

Nei precedenti passaggi si è già utilizzato il fatto che  $\phi' = \phi\beta/\tau$ . Possiamo ora riutilizzare questa equazione per calcolare il rapporto  $r = \rho_k/V$ , ricordando che la (densità d')energia cinetica ha espressione

$$(A1.12) \quad \rho_k = \phi'^2/2a^2 = (\beta\phi)^2/2(a\tau)^2.$$

Dal momento che  $\beta$  e la relazione tra  $a$  e  $\tau$  dipendono da  $c$ , anche  $r$  dipende da  $c$ . Ne segue che il parametro di stato della DE,

$$(A1.13) \quad w = p/\rho = (\rho_k - V)/(\rho_k + V) = (r - 1)/(r + 1),$$

dipenderà anch'esso da qual'è la componente dominante nell'espansione.

Sostituendo in eq. (A1.12) l'espressione che eq. (A1.10) dà per  $(a\tau)^2$ , sarà

$$(A1.14) \quad \rho_k = \frac{1}{21 + (2u - 1)/\beta} \frac{\alpha}{\phi^z} A^{4+\alpha} \quad \text{e quindi} \quad r = \frac{1}{21 + (2u - 1)/\beta} \frac{\alpha}{\phi^z}.$$

Questo valore di  $r$  è utilizzato per ricavare figura XII.3. Essa mette anche in evidenza la dipendenza del parametro di stato della DE dal regime di espansione.

In particolare, il caso  $c = 1$  mostra che la DE potrebbe avere un parametro di stato  $w \simeq -0.9$  se l'espansione è dovuta a una miscela di materia ( $w = 0$ ) e DE ( $w = -0.9$ ) con parametro di stato «medio»  $\Omega_{de}w = \bar{w} \simeq -2/3$ , facilmente ottenibile con  $\Omega_{de} \simeq 0.73$ .

Questo risultato ha solo valore euristico. Se, a un certo istante nell'evoluzione del cosmo, risulta  $\Omega_{de} = 0.73$ , si tratta di un transiente. Infatti, per ciascuna componente risulta  $\rho \propto a^{-3(1+\bar{w})}$  e, quindi, sarà

$$(A1.15) \quad \rho_m/\rho_{de} = a^{3w_{de}} \quad (\text{con } w_{de} < 0).$$

Siamo quindi di fronte a un caso in cui il valore complessivo di  $\bar{w}$  cambia rapidamente. Ciò contraddice l'assunto  $\bar{w} = \text{const.}$ , che ha reso possibile l'integrazione dell'equazione di Friedmann (A1.1b)

## CAPITOLO XIII

# Il fondo cosmico microonde

### 1. Introduzione

Uno dei fondamenti della cosmologia è la scoperta del fondo cosmico microonde (CMB: *cosmic microwave background*). Negli anni Quaranta, partendo dai dati sull'abbondanza di  ${}^4\text{He}$ , George Gamow l'aveva previsto e ne aveva stimato la temperatura, tra 4 e 5 K. Non si era però reso conto che, cadendo nella banda micro-onde (MW: *microwaves*), il CMB sarebbe stato perfettamente visibile da terra. Così, dopo che parecchi dati, specialmente sulle molecole interstellari, avevano messo in evidenza una «temperatura anomala» negli spazi galattici, Penzias & Wilson, nel 1965, lo scoprirono praticamente per caso.

Attualmente la temperatura del CMB è stimata essere

$$(XIII.1) \quad T_o = 2.726 (\pm 0.004) \text{ K} = 2.346 (\pm 0.003) \times 10^{-4} \text{ eV}.$$

Tale temperatura designa la nostra epoca cosmica. Perché essa si sposti di 1/1000 di grado, peraltro, occorrono 70 milioni di anni: il tempo cosmico ha altre scale, non solo rispetto ai tempi umani, ma anche rispetto ai tempi della paleontologia.

Il CMB è dunque una specie di fossile arcaico e tuttora vivente. Come tale, riesce a comunicarci informazioni su uno stato veramente ancestrale del cosmo, lungo due strade: le sue (eventuali) distorsioni rispetto al corpo nero e le sue tenui anisotropie e debolissima polarizzazione.

### 2. Lo spettro del CMB

Dopo il disaccoppiamento dei neutrini e il decadimento degli elettroni, l'interazione tra materia barionica e fotoni avviene tramite 3 processi principali:

- 1) urto compton;
- 2) doppio urto compton;
- 3) emissione (e assorbimento) *free-free*.

Tra questi processi, 1) conserva il numero dei fotoni. Il doppio urto compton,

$$\gamma + e \rightarrow \gamma + \gamma + e$$

e l'effetto *free-free* o *bremstrahlung*,

$$e + \text{ione} \rightarrow e + \text{ione} + \gamma$$

(che possono procedere anche alla rovescia) danno luogo invece all'emissione di nuovi fotoni.

Di per sé, fin tanto che il cosmo è ionizzato, l'urto compton mantiene la stessa temperatura per barioni e radiazione.

In generale, quando la distribuzione dei fotoni è della forma

$$(XIII.2) \quad f = \frac{2}{(2\pi)^3} \frac{1}{e^{\varepsilon/T} - 1},$$

( $\varepsilon$ : energia dei fotoni) essa rimane invariante nel corso dell'espansione cosmica. Perché ciò accada non occorrono collisioni di alcun tipo, ma i processi 1), 2), 3) non alterano comunque tale distribuzione. In particolare, tale distribuzione fissa la densità in numero di fotoni

$$(XIII.3) \quad n_\gamma = \frac{2\zeta(3)}{\pi^2} T^3$$

(ricordiamo che  $\zeta(3) \simeq 1.20..$  è un *numero di Riemann*).

Se, per qualche ragione, il cosmo si trovasse a una determinata temperatura  $T$ , ma il numero di fotoni fosse diverso da (XIII.3) e i processi per produrre nuovi fotoni fossero troppo lenti, la distribuzione di equilibrio termodinamico dei fotoni effettivamente presenti sarebbe distorta nella forma

$$(XIII.4) \quad f = \frac{1}{(2\pi)^3} \frac{1}{e^{(\varepsilon+\mu)/T} - 1},$$

che, di per sé, non è invariante nel corso dell'espansione, ma viene mantenuta invariante dagli urti compton, finché il cosmo è adeguatamente ionizzato. Se invece doppio scattering compton e/o urti *free-free* danno luogo a un'adeguata produzione (o assorbimento) di  $\gamma$ , il *potenziale chimico*  $\mu$  viene riportato a zero.

Nella storia dell'espansione cosmologica, tali processi sono efficaci a *redshift*  $z > 10^7$ ; eventuali fotoni da decadimento di particelle pesanti o provocati da altri eventi, vengono ridistribuiti e termalizzati solo se si originano a  $z > 10^7$ ; processi che causassero ritorsioni spettrali a *redshift* minori lascerebbero traccia.

Dal punto di vista osservativo, nella zona spettrale attorno al picco, il CMB è il più perfetto corpo nero esistente in natura, come è stato accertato dall'esperimento *FIRAS* del satellite *COBE*. Distorsioni spettrali nella zona di Wien sono molto più difficili da determinare, soprattutto a causa dei *foreground* e in particolare della polvere interstellare.

È poi noto che i processi di formazione stellare e l'azione dei nuclei galattici attivi causano un ampio *input* energetico. I fotoni derivanti da tali processi vengono

in gran parte termalizzati dalle polveri interstellari e riemessi nel lontano infrarosso, ai confini con lo spettro MW. La conseguente distorsione spettrale è stata in parte già rilevata.

### 3. L'effetto SZ

Di grande importanza, in relazione alla presenza di sorgenti, è poi l'effetto SZ (Sunyaev-Zel'dovich); esso avviene quando i fotoni del CMB attraversano un gas caldo a temperatura  $T_{el}$ , anche se mediamente il gas è trasparente alla radiazione stessa. L'evento si verifica tipicamente quando la radiazione del CMB attraversa un ammasso di galassie. In questo caso una piccola percentuale dei  $\gamma$  del CMB, provenienti dalla zona di Raleigh-Jeans, urtati dagli elettroni del gas caldo, acquisisce energia e viene portata su frequenze superiori al picco dello spettro. Poiché il numero di  $\gamma$  di alta frequenza, nel CMB, è assai minore del numero a bassa frequenza, oltre a un apparente calo di temperatura nella zona di Raleigh-Jeans, ne risulta una distorsione spettrale potenzialmente ben visibile oltre il picco.

La distorsione di bassa frequenza è ben descritta dal parametro

$$(XIII.5) \quad y = \int dt \sigma_T n_e(t) T_{el} / m_e$$

( $m_e$ : massa dell'elettrone) direttamente legato al calo di temperatura, che risulta essere  $\delta T/T = -2y$ .

I valori attesi sono  $y \sim \mathcal{O}(10^{-4}-10^{-5})$ . Il diametro dei volumi contenenti gas caldi, negli ammassi, è  $\sim 1 h^{-1} \text{Mpc}$ . Ammassi situati a una distanza  $D_{cl}$  tra 100 e  $1000 h^{-1} \text{Mpc}$  danno luogo a anisotropie del CMB per  $l \simeq 2\pi D_{cl} h / \text{Mpc} \sim 600-6000$ . Questo *range* si sovrappone marginalmente con quello ove sono attesi segnali cosmologici dal profondo, che possono avere ampiezza analoga ma hanno certamente una diversa dipendenza dalle frequenze, e va quindi tenuto presente nell'analisi dei dati.

Negli esperimenti in corso e in progetto, basati sull'effetto SZ, si mira comunque a determinare direttamente l'immagine «in negativo» degli ammassi lontani, più che gli effetti statistici della presenza di molti ammassi.

### 4. Anisotropie e polarizzazione: definizioni

L'analisi di anisotropie e polarizzazione del CMB, originatesi nel cosmo primordiale, ci fornisce informazioni di prima mano sui parametri cosmologici.

La definizione degli osservabili astrofisici parte dai versori  $\mathbf{n}$  (o dalle coordinate polari  $\theta$  e  $\phi$ ) che fissano i punti della volta celeste. Si prendano poi 2 altri versori  $\mathbf{e}_1, \mathbf{e}_2$ , a formare con  $\mathbf{n}$  una terna ortogonale sinistrorsa.

### Il tensore di polarizzazione

Le proprietà di una radiazione proveniente dalla direzione  $\mathbf{n}$  sono allora espresse dal «tensore»

$$(XIII.6) \quad I_{ij} = \frac{1}{4\pi c} \langle E_i E_j \rangle \quad (i, j = 1, 2),$$

dove le  $E_i$  sono le componenti del campo elettrico nelle direzioni  $\mathbf{e}_i$  ( $i = 1, 2$ ). La media va fatta su intervalli temporali sufficientemente lunghi da coprire le discontinuità quantiche della radiazione. La traccia (invariante per rotazioni di asse  $\mathbf{n}$ ) del tensore (XIII.6),

$$(XIII.7) \quad I = I_{11} + I_{22},$$

è l'intensità della radiazione nel punto  $\mathbf{n}$  (della volta celeste). Mediando ulteriormente  $I(\mathbf{n})$  sulle direzioni  $\mathbf{n}$ , si ottiene  $\bar{I}$ . Sottraendo  $\bar{I}/2$  dalle componenti di  $I_{ij}$ , si evidenziano le caratteristiche della radiazione che dipendono dalla direzione di provenienza. Per questo risulta più utile riferirsi al tensore

$$(XIII.8) \quad \Delta_{ij}(\mathbf{n}) = \frac{I_{ij}(\mathbf{n}) - \delta_{ij}\bar{I}/2}{\bar{I}},$$

le cui componenti si possono tutte trattare come infinitesime, invece che a  $I_{ij}(\mathbf{n})$ .

### I parametri di Stockes

$\Delta_{ij}$  ha traccia nulla e le 3 combinazioni

$$\begin{aligned} T(\mathbf{n}) &= [\Delta_{11}(\mathbf{n}) + \Delta_{22}(\mathbf{n})]/4 && \text{fluttuazione in temperatura} \\ Q(\mathbf{n}) &= [\Delta_{11}(\mathbf{n}) - \Delta_{22}(\mathbf{n})]/4 && \text{parametro di Stockes tra direzioni } \mathbf{e}_1, \mathbf{e}_2 \\ U(\mathbf{n}) &= \Delta_{12}(\mathbf{n})/2 && \text{parametro di Stockes tra direzioni a } 45^\circ \end{aligned}$$

esauriscono l'informazione che esso convoglia.

Avendo a che fare con una radiazione che ha spettro esattamente di corpo nero, il passaggio dalle fluttuazioni d'intensità alle fluttuazioni di temperatura, espresse da  $T(\mathbf{n})$ , è assicurato dalla divisione per 4, essendo  $I \propto T^4$ . Di questa semplice proprietà differenziale diamo qui sotto una verifica diretta.

Si faccia peraltro attenzione all'uso del simbolo  $T$ , in luogo dell'espressione  $\delta T/T$ , più consona a esprimere una *fluttuazione*, ma che appesantirebbe la scrittura. Ad esempio, nella relazione  $I \propto T^4$ ,  $T$  è la *temperatura*, non la sua fluttuazione. Quando questo avviene lo faremo notare.

In particolare avremo ora a che fare con la *temperatura* nel verificare che la divisione per 4 assicura veramente il passaggio da fluttuazioni di intensità  $\delta I$  a flut-

tuazioni di temperatura  $\delta T$ . Dette  $I' = I + \delta I$  e  $T' = T + \delta T$ , sarà

$$(XIII.9) \quad \frac{\delta I}{I} = \frac{I' - I}{I} = \frac{T'^4 - T^4}{T^4} = \left(1 + \frac{\delta T}{T}\right)^4 - 1 = \\ = 1 + 4 \frac{\delta T}{T} + 6 \left(\frac{\delta T}{T}\right)^2 + 4 \left(\frac{\delta T}{T}\right)^3 + \left(\frac{\delta T}{T}\right)^4 - 1 \simeq 4 \frac{\delta T}{T},$$

in quanto  $\delta T/T$  è  $\sim \mathcal{O}(10^{-5})$  e le potenze di  $\delta T/T$  superiori alla prima sono legittimamente trascurate. In effetti stiamo descrivendo una situazione fisica cui possono essere applicate tutte le approssimazioni legate alla *linearità*, trascurando, sia a livello sperimentale che analitico, tutte le potenze delle fluttuazioni superiori alla prima.

Le definizioni dei parametri di Stokes  $Q$  e  $U$  mostrano che essi dipendono dalla scelta della base  $\mathbf{e}_1, \mathbf{e}_2$ . Se si sceglie un'altra base ruotata di un angolo  $\psi$ ,

$$(XIII.10) \quad \hat{\mathbf{e}}_1 = \mathbf{e}_1 \cos \psi - \mathbf{e}_2 \sin \psi, \\ \hat{\mathbf{e}}_2 = \mathbf{e}_1 \sin \psi + \mathbf{e}_2 \cos \psi,$$

i parametri di Stokes diventano

$$(XIII.11) \quad \hat{Q} = Q \cos 2\psi - U \sin 2\psi, \\ \hat{U} = Q \sin 2\psi + U \cos 2\psi,$$

come è facile verificare. Il fattore 2 davanti a  $\psi$  esprime il fatto che i parametri di Stokes sono un campo di *spin 2*.

### Modi di polarizzazione

$T, Q, U$  si possono espandere in armoniche sferiche.  $T$  è uno scalare e avremo semplicemente

$$(XIII.12) \quad T(\mathbf{n}) = \sum_{l,m} a_{T,lm} Y_{lm}(\mathbf{n});$$

le  $Y_{lm}(\mathbf{n})$  sono le funzioni associate di Legendre (usate, ad esempio, nello studio quantistico dell'atomo di idrogeno). Per lo sviluppo di  $Q$  e  $U$  si richiedono invece le armoniche sferiche di spin 2,  ${}_{\pm 2}Y_{lm}(\mathbf{n})$ , con cui si ottiene:

$$(XIII.13) \quad [Q + iU](\mathbf{n}) = \sum_{l,m} a_{+,lm} {}_{+2}Y_{lm}(\mathbf{n}), \\ [Q - iU](\mathbf{n}) = \sum_{l,m} a_{-,lm} {}_{-2}Y_{lm}(\mathbf{n}).$$

Le funzioni  ${}_{\pm 2}Y_{lm}(\mathbf{n})$  si trovano in proutuari matematici; non le riportiamo, anche

perché non saranno più utilizzate. Si dimostra infatti che le combinazioni

$$(XIII.14) \quad a_{E,lm} = -(a_{+,lm} + a_{-,lm}), \quad a_{B,lm} = -(a_{+,lm} - a_{-,lm})$$

definiscono due campi o *modi* di polarizzazione scalari

$$(XIII.15) \quad E(\mathbf{n}) = \sum_{l,m} a_{E,lm} Y_{lm}(\mathbf{n}), \quad B(\mathbf{n}) = \sum_{l,m} a_{B,lm} Y_{lm}(\mathbf{n}).$$

Essendo però definiti attraverso lo spazio di Fourier,  $E(\mathbf{n})$  e  $B(\mathbf{n})$  si ricavano pesando opportunamente i dati di tutto il cielo (non solo nel punto  $\mathbf{n}$ ). In appendice A1 esplicitiamo le espressioni di  $a_{T,lm}$ ,  $a_{+,lm}$ ,  $a_{-,lm}$ , ma mostriamo anche come, al limite in cui un tratto di cielo si potesse considerare piatto,  $E$  e  $B$  sarebbero «locali». Quindi, al crescere di  $l$ ,  $E$  e  $B$  esprimono sempre più le proprietà di polarizzazione in un ristretto intorno di  $\mathbf{n}$ .

Il vantaggio di usare  $E$  e  $B$  non è solo operativo.  $E$  e  $B$  convogliano informazioni diverse sull'origine del cosmo. I due campi hanno inoltre diverse proprietà topologiche:  $E$  ha divergenza nulla, mentre il rotore di  $B$  è nullo, in stretta analogia coi campi elettrico e d'induzione magnetica da cui prendono il nome, ma con cui... non hanno nulla a che fare (!) Ancora in analogia con tali campi,  $E$  e  $B$  differiscono per il comportamento sotto l'operazione di parità. Se si effettua la riflessione  $\mathbf{n} \rightarrow -\mathbf{n}$ ,  $E$  è pari,  $B$  è dispari, come si può facilmente verificare a partire dalle definizioni (XIII.13)-(XIII.15). Allo stesso modo si verifica che  $T$  è pari.  $T$  e  $B$  non possono quindi essere correlati. Mentre, invece, la correlazione tra fluttuazioni di temperatura  $T$  e modo  $E$  di polarizzazione può avere grande rilevanza fisica.

### Spettri delle anisotropie

I coefficienti complessi (discreti)  $a_{T,lm}$ ,  $a_{E,lm}$ ,  $a_{B,lm}$  ( $l = 0, 1, \dots, n, \dots$ ;  $-l \leq m \leq +l$ ) contengono rispettivamente le stesse informazioni dei campi  $T(\mathbf{n})$ ,  $E(\mathbf{n})$ ,  $B(\mathbf{n})$ ; allo stesso modo in cui lo spettro complesso  $\delta(\mathbf{k})$  contiene tutte le informazioni contenute nel campo delle fluttuazioni di densità  $\varepsilon(\mathbf{x})$ . Ci sono due differenze importanti, a parte la bidimensionalità dello spazio angolare rispetto alla tridimensionalità dello spazio fisico:

- 1) La volta celeste ha un'estensione angolare finita ( $4\pi$ ) e quindi  $T(\mathbf{n})$ ,  $E(\mathbf{n})$ ,  $B(\mathbf{n})$  sono funzioni periodiche, sicché il loro spettro è effettivamente *discreto*.
- 2) La distribuzione delle  $\varepsilon(\mathbf{x})$  è sicuramente non-Gaussiana sulle piccole scale, per effetto della non-linearità; la distribuzione delle fluttuazioni del CMB, invece, se è non-Gaussiana lo è per ragioni *primordiali* e, comunque, i dati sinora disponibili pongono robusti limiti superiori alla non-Gaussianità. (Ma, se futuri esperimenti a maggiore sensibilità accertassero anche una tenue non-Gaussianità primordiale, si tratterebbe di un fatto fisico di estrema importanza.)

In assenza di non-Gaussianità, gli spettri

$$(XIII.16) \quad \begin{aligned} C_l^{(T)} &= \langle |a_{T,lm}|^2 \rangle, \\ C_l^{(E)} &= \langle |a_{E,lm}|^2 \rangle, \\ C_l^{(B)} &= \langle |a_{B,lm}|^2 \rangle, \\ C_l^{(TE)} &= \langle a_{T,lm}^* a_{E,lm} \rangle \end{aligned}$$

esauriscono l'informazione sul processo di cui le fluttuazioni osservate sono realizzazione.

Le medie in eq. (XIII.16) sono idealmente effettuate su diverse realizzazioni. Ad esempio, si potrebbe pensare (!) a una media tra un gran numero di osservatori appartenenti a orizzonti disgiunti.

Nell'impossibilità di effettuarla, si esamina il nostro (unico) cielo da diversi riferimenti polari, mediando tra i risultati ottenuti. A questa procedura e all'assunto di isotropia segue che gli spettri non dipendono da  $m$ : la procedura infatti comprende medie effettuate mantenendo invariato l'asse polare, ma cambiando l'origine della coordinata angolare  $\phi$ .

Sul dominio bidimensionale finito costituito dal cielo si replicano tutte le relazioni ottenute a partire dalle fluttuazioni  $\varepsilon(\mathbf{x})$ , nel dominio tridimensionale costituito dallo spazio. Nel caso bidimensionale ci sono solo le complicazioni dovute alla topologia sferica del cielo. Ad esempio, l'espressione della funzione di correlazione a 2 punti, in termini dello spettro, ha la forma

$$(XIII.17) \quad \langle T(\mathbf{n}_1)T(\mathbf{n}_2) \rangle = \sum_{l=0}^{\infty} \frac{2l+1}{4\pi} C_l^{(T)} P_l(\mathbf{n}_1 \cdot \mathbf{n}_2)$$

in cui i  $P_l$  sono gli ordinari *polinomi di Legendre*. Relazioni analoghe possono darsi per la correlazione tra polarizzazione, dell'uno o dell'altro modo, in due diverse direzioni  $\mathbf{n}_{1,2}$ , o per la correlazione tra fluttuazioni di temperatura e modo  $E$  di polarizzazione. La difficoltà nell'esprimerle sta nelle funzioni non banali che rimpiazzano in questi casi i polinomi di Legendre. Rinviamo quindi per questo a lavori specialistici.

## 5. Mappe del cielo MW

Concentriamoci ora sulle anisotropie, per una discussione che potrebbe essere poi estesa, con qualche complicazione non banale, anche alla polarizzazione. Le prospettive osservative e i potenziali risultati derivanti dall'analisi della polarizzazione sono tra gli argomenti attualmente in elaborazione sul fronte della ricerca. Dedicheremo in questo testo uno spazio limitato a tali temi, rimandando a articoli specializzati.

Le anisotropie possono essere descritte tramite la funzione  $T(\mathbf{n})$ , oppure tramite lo spettro  $C_l^{(T)}$ , ricavato dalle componenti  $a_{T,lm}$ . Conformemente all'eq. (XIII.12), l'in-

formazione contenuta nelle componenti  $a_{T,lm}$  è del tutto equivalente a quella contenuta nella funzione  $T(\mathbf{n})$ . La differenza apparente sta nel fatto che nel primo caso abbiamo una infinità *numerabile* di componenti, nel secondo caso una infinità *continua* di valori; ciò avviene sempre quando si fa una trasformata di Fourier su una funzione periodica.

In questo caso, si tratta comunque di una differenza solo concettuale. Gli esperimenti che misurano  $T(\mathbf{n})$  non ne determinano valori per precise direzioni  $\mathbf{n}$ , ma una media tra i suoi valori in un fascio di direzioni  $\mathbf{n}'$  attorno a  $\mathbf{n}$ , pesati da un «filtro» o «funzione finestra», con un profilo caratteristico dell'esperimento e in genere prossimo a una Gaussiana, di tipo

$$(XIII.18) \quad W_\sigma(\Delta\theta) = \frac{1}{2\pi\sigma} \exp\left\{-\frac{1}{2} \left[\frac{\Delta\theta(\mathbf{n}, \mathbf{n}')}{\sigma}\right]^2\right\},$$

in cui  $\Delta\theta(\mathbf{n}, \mathbf{n}')$  è l'angolo tra le direzioni  $\mathbf{n}$ ,  $\mathbf{n}'$  e  $\sigma$  esprime la *risoluzione* dell'esperimento. Lo *scanning* del cielo effettuato dal radiometro, o da altro apparato ricevitore, viene pertanto espresso tramite i valori  $\bar{T}_i = \bar{T}(\mathbf{n}_i)$  per un insieme finito di direzioni  $\mathbf{n}_i$ , ottenuti mediando attorno a esse tramite un filtro di tipo (XIII.18). Il procedimento è denominato *pixelizzazione* e ciascuna temperatura  $\bar{T}_i$  è attribuita al corrispondente *pixel* attorno alla direzione  $\mathbf{n}_i$  di raggio  $\sim \sigma$ .

Il miglior utilizzo delle risultanze sperimentali si ha distribuendo le direzioni  $\mathbf{n}_i$  a distanze angolari  $\sim \sigma/2$ . Gli esperimenti recenti e quelli futuri utilizzano, di norma, le  $\sim 16\pi/\sigma^2$  direzioni fissate dall'algoritmo *healpix*, che fornisce la distribuzione ottimale tenendo conto della curvatura del cielo. Evidentemente i pixel si sovrappongono ampiamente tra loro.

La risoluzione  $\sigma$  limita anche il numero di componenti  $a_{T,lm}$  che si possono apprezzare. Si può infatti giungere fino a  $l \sim 2\pi/(\sigma/2)$ ; per  $l$  più elevati non c'è segnale apprezzabile. Il numero delle componenti significative, all'incirca, sarà

$$(XIII.19) \quad \sum_{l=0}^{4\pi/\sigma} 2l + 1 \simeq (4\pi/\sigma + 1)^2,$$

di poco superiore al numero di pixel.

Il numero di componenti apprezzabili, per  $C_l^{(T)}$ , è invece nettamente inferiore. In effetti, mentre le quantità  $\bar{T}_i$  o  $a_{T,lm}$  esprimono la geografia dettagliata del segnale, le componenti dello spettro ne danno solo proprietà statistiche.

In figura XIII.7 sono rappresentate immagini del cielo MW, ottenute dall'esperimento COBE. Molto maggiore è la risoluzione nell'immagine in figura XIII.8, ottenuta dall'esperimento WMAP. I colori artificiali rappresentano col blu il deficit di temperatura e col rosso l'eccesso di temperatura. I punti sono dati in coordinate galattiche; la fascia rossa che attraversa la figura XIII.7 è dovuta alla radiazione, prevalentemente sincrotrone, da parte della galassia. Essa è stata tentativamente sottratta nell'immagine XIII.8.

## 6. Dalle fluttuazioni di densità primordiali alle anisotropie del CMB

Le anisotropie e la polarizzazione osservate nel CMB traggono origine dalle stesse fluttuazioni primordiali, responsabili anche della successiva formazione delle strutture cosmiche. Il fatto che lo stesso spettro primordiale di fluttuazioni renda conto sia dei dati CMB che della struttura su grande scala, anzi, è un grande successo della cosmologia; della stessa portata del successo della BBN.

Ci proponiamo ora di seguire la procedura che permette di ricavare gli spettri osservati dai parametri dei modelli. Lo facciamo prima di esibire le risultanze delle osservazioni, anche per rendere condivisibile la «meraviglia» che si prova nel trovare riprodotte, nell'articolazione dei dati reali, le complesse previsioni elaborate sulla base di una fisica completamente estrapolabile da risultanze di laboratorio.

In parte questa procedura è stata già focalizzata nel capitolo VIII, quando si sono studiate le onde di compressione nel plasma fotoni-barioni. Il problema è però più complesso e le previsioni teoriche tengono conto di altri effetti. Tra di essi, ad esempio, c'è «banalmente» la presenza della componente neutrinica. Ma le problematiche più significative da mettere in luce riguardano:

- 1) Le azioni gravitazionali sui fotoni;
- 2) la graduale perdita delle caratteristiche di fluido, da parte della componente fotonica.

Il punto di snodo di tutta la trattazione è la distribuzione nello spazio delle fasi dei fotoni,  $f(\mathbf{x}, p, \mathbf{n}, t)$ . Analoghe distribuzioni si possono scrivere per ogni tipo di particelle. Il numero di particelle in un volume  $d^3x$  attorno al sito  $\mathbf{x}$  e con impulso compreso entro  $d^3p$  attorno al valore  $p\mathbf{n}$ , al tempo  $t$ , è allora

$$(XIII.20) \quad dN = f(\mathbf{x}, p, \mathbf{n}, t) d^3x d^3p.$$

La distribuzione  $f$  descrive comunque i fotoni, sia prima che durante o dopo la ricombinazione, a differenza delle variabili  $\delta$  e  $v$  (fluttuazione di densità e campo di velocità) nel modello di fluido, che si possono utilizzare solo quando fotoni e barioni sono strettamente accoppiati;  $\delta$  e  $v$  possono comunque essere ricavati dalla distribuzione  $f$ . Anche le componenti del tensore di polarizzazione si possono ricavare da  $f$ , che soddisfa poi opportune equazioni dinamiche. Nell'insieme, dunque, passando attraverso queste equazioni, si riesce a collegare la fenomenologia al modello.

In generale, la distribuzione  $f$  sarà esprimibile come somma di un termine di *background* e un termine di fluttuazione:

$$(XIII.21) \quad f(\mathbf{x}, p, \mathbf{n}, t) = f_o(p, t) + \tilde{\delta}f(\mathbf{x}, p, \mathbf{n}, t)$$

dove

$$(XIII.22) \quad f_o(p, t) = \frac{2}{(2\pi)^3} \frac{1}{e^{p/T(t)} - 1}$$

è la distribuzione di Bose in cui l'unica dipendenza dal tempo si ha attraverso la temperatura, mentre non ci sono dipendenze dal sito o dalla direzione dell'impulso (omogeneità e isotropia).

Le fluttuazioni dipendono invece da  $\mathbf{x}$  e  $\mathbf{n}$ . In linea di principio, poi, ci potrebbero essere fluttuazioni diverse per ciascun diverso valore di  $p$ . Tuttavia, in regime non relativistico, la dinamica dei fotoni non dipende dal loro impulso, in quanto la sezione d'urto angolare Thomson è indipendente dall'energia. Ha quindi senso considerare in ciascun punto un'intensità integrata

$$(XIII.23) \quad I(t) = (\pi^2/15)T^4(t)$$

e fluttuazioni d'intensità

$$(XIII.24) \quad \delta I(\mathbf{x}, \mathbf{n}, t) = \int dp \delta \tilde{f}(\mathbf{x}, p, \mathbf{n}, t)$$

che si ricavano sommando nuovamente, nel continuo, su tutti i possibili valori dell'impulso dei fotoni.

Il rapporto  $\delta I(\mathbf{x}_o, \mathbf{n}, t_o)/\langle I(t) \rangle$ , dove  $\mathbf{x}_o$  e  $t_o$  sono il sito e il tempo dell'osservazione, coincide allora con  $4T(\mathbf{n}) = Tr \Delta_{ij}$  (vedi eq. XIII.8).

Le relazioni che permettono di ricavare le altre componenti di  $\Delta_{ij}$  e quindi i campi di polarizzazione E e B non verranno fornite in questo testo, ma si trovano in molti articoli specializzati. Lo scopo della trattazione che riporteremo in appendice è quello di spiegare i problemi e le tecniche seguite per superarli, quando si vogliono prevedere le caratteristiche del CMB per un determinato modello. I passaggi forniti sono altresì autoconsistenti e potrebbero essere anche utilizzati a scrivere un programma numerico che porti a prevedere le anisotropie. Questo tuttavia non è affatto lo scopo per cui vengono presentati.

I programmi disponibili pubblicamente, che seguono linee analoghe, ma sono giustamente più sofisticati, comprendono svariate alternative e forniscono adeguate previsioni anche sulla polarizzazione. Chi fosse interessato al loro utilizzo o a ulteriori sviluppi è rimandato quindi alla letteratura specializzata.

Vediamo quindi, in generale, cosa occorre fare per prevedere le anisotropie e la polarizzazione del CMB.

In primo luogo servono le equazioni dinamiche soddisfatte dalla distribuzione  $\tilde{f}(\mathbf{x}, p, \mathbf{n}, t)$ , che terranno conto anzitutto dell'espansione cosmica; gli effetti essenziali, peraltro, sono le interazioni tra le fluttuazioni  $\tilde{f}(\mathbf{x}, p, \mathbf{n}, t)$  e le fluttuazioni nelle altre componenti, che avvengono via urto Thomson e tramite forze gravitazionali. Queste equazioni dinamiche, assieme a quelle soddisfatte dalle altre componenti nella fluttuazione, sono fornite in appendice A1.

Al di là dei dettagli analitici, che si trovano in appendice, possiamo riportare alcune informazioni di carattere specifico:

- 1) Come in altri casi, essendo il problema lineare, conviene studiare separatamente il comportamento di ciascuna onda di compressione, di numero d'onda  $\mathbf{k}$ , e sovrapporre i risultati, tenendo conto del peso statistico del  $\mathbf{k}$  considerato.

- 2) La sezione d'urto fotone-elettrone non dipende dall'energia, mentre il rapporto energia/impulso dei singoli fotoni, che non hanno massa, è ovviamente costante. La dinamica dei fotoni è così indipendente dal loro impulso  $p$  e, invece di studiare le fluttuazioni della fluttuazione  $\delta f(\mathbf{k}, p, \mathbf{n}, \tau)$ , si studia l'andamento della funzione  $\Theta(\mathbf{k}, \mathbf{n}, \tau)$ , definita in eq. (A1.3), per ciascuna onda di compressione, in cui si è opportunamente «sommato» sugli impulsi  $p$ .  $\tau$  è il tempo conforme.
- 3) Conviene studiare il problema nella *gauge newtoniana conforme*, in cui l'azione delle fluttuazioni nella gravità si esplica tramite i potenziali  $\Psi$  e  $\Phi$ .
- 4) La funzione  $\Theta(\mathbf{k}, \mathbf{n}, \tau)$  viene espansa in armoniche sferiche, ovvero in serie di multipoli. Se i fotoni fossero un fluido (e fin tanto che essi possono essere trattati come un fluido), solo il termine di monopolo e di dipolo ( $l = 0, 1$ ; eq. A1.7) sarebbero  $\neq 0$ . Essi sono allora riconducibili alle fluttuazioni di densità e al campo di velocità nel «fluido fotonico».

Quando l'accoppiamento fotoni-barioni si attenua, oltre alle componenti  $\Theta_0$  e  $\Theta_1$  si accendono gradualmente le componenti successive, a  $l$  via via crescenti. Le «condizioni iniziali» sono peraltro stabilite dal modello di fluido.

Lo studio dettagliato dell'accensione dei multipoli, a  $l > 2$ , serve comunque principalmente a determinare la progressiva attenuazione dell'ampiezza delle onde di compressione, nel corso del disaccoppiamento.

Dal momento che anche le componenti  $a_{T,lm}$  si definiscono tramite espansione in armoniche sferiche (eq. XIII.12), esiste una relazione diretta tra le componenti dello spettro  $C_l^{(T)}$  e le  $\Theta_l$  al tempo attuale  $\tau_0$ . Tale relazione è l'eq. (A1.10). Si potrebbe quindi procedere a integrare il sistema dinamico fino al tempo  $\tau_0$ , seguendo la progressiva accensione dei multipoli successivi, ricavandone poi le componenti  $C_l^{(T)}$ . Questa procedura veniva seguita in alcuni dei primi algoritmi dedicati a questo studio, ma è molto dispendiosa.

La procedura attualmente seguita si basa invece sugli sviluppi esposti in appendice A2; si mostra che è possibile scrivere un *termine di sorgente*  $S(k, \tau)$  che dipende solo dalle prime armoniche dello sviluppo in multipoli. Noto  $S(k, \tau)$ , l'eq. (A2.4) fornisce le armoniche sferiche delle *fluttuazioni integrate*,  $\Theta_l$ , a qualunque tempo e, in particolare, al tempo attuale. Le  $\Theta_l$  vanno poi «pesate» tramite lo spettro primordiale  $P(k)$ , di modo da stabilire come contribuisca l'ampiezza primordiale su ciascun  $k$  al formarsi del segnale su un determinato  $l$ . Abbiamo visto sin dall'inizio del capitolo VIII una relazione diretta tra  $k$  e  $l$ . Si tratta tuttavia di una relazione che passa attraverso due successive espansioni in serie. Una determinata componente  $C_l^{(T)}$  riceve «contributo» da tutto un pennello di  $k$ , centrati su un particolare valore, ma abbastanza ampio. Dal punto di vista analitico, le componenti dello spettro  $C_l^{(T)}$  si ricavano tramite l'integrale in eq. (A1.10), «sommando» quindi nel continuo tra i contributi corrispondenti ai  $k$  diversi, opportunamente «pesati».

In questo contesto conviene sottolineare che sono state sviluppate tecniche che permettono di integrare il sistema di equazioni dinamiche (A1.10)-(A1.11)-

(A1.12), senza seguire tutte le  $\Theta_l$ , via via che esse si «accendono». Queste tecniche mirano a dare valori esatti per le  $\Theta_l$  con  $l < 3$ , mentre le armoniche a  $l$  maggiore vengono opportunamente approssimate.

## 7. Forma dello spettro angolare e sua dipendenza dai parametri cosmologici

L'espressione della sorgente  $\mathcal{S}$ , fornita per intero in eq. (A2.3), contiene termini di diversa natura, come dettagliato in appendice A2, eq. (A2.3)

Un primo insieme di termini è «acceso» nella *banda dell'ultimo scattering* e la *funzione di visibilità* ne determina il contributo allo spettro delle anisotropie. Tra questi termini troviamo:

- 1) Termini proporzionali alla fluttuazione di densità integrata dei  $\gamma$ ,  $\propto \Theta_o(k, \tau_o)$  (eccessi-deficit di temperatura)
- 2) Termini dipendenti dalla velocità dei barioni  $v_b$  (effetto Doppler)
- 3) Termini dipendenti dal campo gravitazionale dovuto alla concentrazione di densità (effetto SW: *Sachs & Wolfe*).
- 4) Termini dovuti all'anisotropia dello scattering Thomson, contenenti  $\Theta_2(k, \tau)$

Accanto ai termini suelencati, ne abbiamo altri dovuti alla variazione dei potenziali gravitazionali, che esprimono l'effetto ISW (*Sachs & Wolfe* integrato); questi ultimi termini sono importanti soprattutto in relazione al cambio di ritmo di espansione dovuto al contributo «recente» della DE.

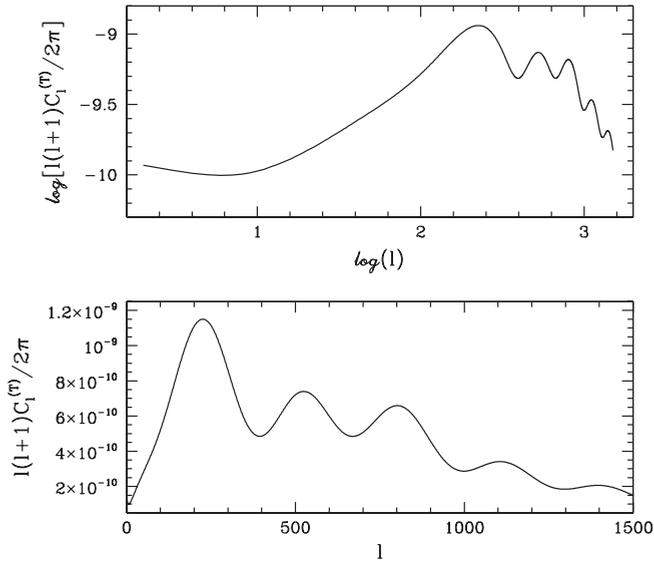
In figura XIII.1 riportiamo l'andamento previsto per lo spettro  $C_l^{(T)}$ , per la cosmologia nella didascalìa. Richiamiamo l'attenzione del lettore sul fatto che gli spettri sono proporzionali al quadrato dell'ampiezza delle fluttuazioni. A un valore  $C_l \sim \mathcal{O}(4 \times 10^{-10})$  corrispondono quindi fluttuazioni in temperatura  $\delta T/T \sim \mathcal{O}(2 \times 10^{-5})$  e fluttuazioni di densità  $\delta\rho/\rho \sim \mathcal{O}(10^{-4})$ .

Interessa ora vedere come varia l'andamento dello spettro, in funzione delle variazioni dei diversi parametri. Per taluni di essi l'effetto è diretto e chiaramente distinguibile; per altri ci sono fatti di *degenerazione*: variazioni analoghe nello spettro seguono la variazione di parametri diversi.

Per questo, nell'analisi reale dei dati, si utilizzano tecniche avanzate, capaci di tenere conto delle variazioni nel complesso spazio parametrico. Senza entrare in questi aspetti operativi, distingueremo anzitutto tra i parametri collegati alle armoniche sferiche per  $l$  elevati, tipicamente  $l > 200$ , e quelli che influiscono più nettamente sulle basse armoniche.

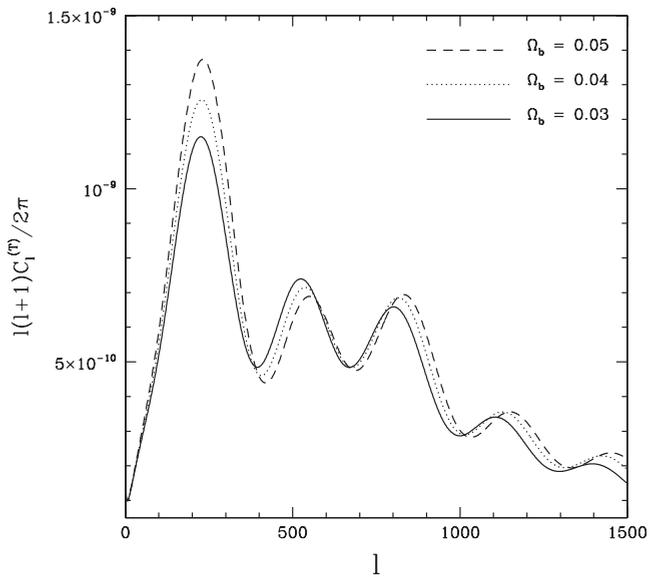
Tra i primi vanno annoverati l'abbondanza di barioni e di DM. Tra i secondi, i parametri concernenti la DE e l'opacità del cosmo tra la banda dell'ultimo scattering e l'osservatore. Gli effetti di una variazione dell'opacità, se considerati solo per le anisotropie, sono degeneri con quelli causati da una variazione dell'indice spettrale  $n_s$ . Anche il valore di  $H_o$  ha effetti diretti sullo spettro  $C_l^{(T)}$ , alla zona di confine tra alti e bassi  $l$ .

Passiamo ora in esame i diversi parametri più in dettaglio.



**FIGURA XIII.1**

Spettro delle anisotropie del CMB, per un modello  $\Lambda$ CDM con  $\Omega_m = 0,28$ ,  $\Omega_b = 0,04$ ,  $h = 0,70$ ,  $n_s = 1$ . Lo spettro è mostrato sia su scale lineari che logaritmiche.



**FIGURA XIII.2**

Spettro delle anisotropie del CMB al variare di  $\Omega_b$ . A parte quest'ultimo parametro, il modello è lo stesso della precedente figura.

### Variazioni del contenuto barionico del modello

In figura XIII.2 è riportato lo spettro  $C_l^{(T)}$  per 3 diversi valori di  $\Omega_b$ .

Il cambio di  $\Omega_b$  provoca una variazione della velocità del suono alla ricombinazione ( $c_s$ ). Conformemente a quanto discusso nel capitolo VIII ci aspettiamo quindi una dislocazione progressiva, lungo l'ascissa  $l$ , dei massimi successivi al primo. L'effetto è chiaramente visibile per il primo e il secondo massimo; al crescere di  $k$ , l'espressione completa di  $c_s$  contiene anche fattori legati allo smorzamento delle onde, che non abbiamo considerato in questo testo; essi inquinano progressivamente l'effetto.

Un altro effetto caratteristico, dovuto al cambio di  $\Omega_b$ , altrettanto vistoso, è la crescita del primo picco, simultanea al calo dello spettro in corrispondenza del primo minimo. L'origine dell'effetto è chiara: variando il contenuto barionico, varia l'ampiezza delle onde soniche.

### Variazioni del contenuto di materia nel modello

In figura XIII.3, si vede come, al variare di  $\Omega_m$ , la variazione dello spettro sia molto differente da quella che si aveva al variare del contenuto barionico. Siccome  $\Omega_b$  è tenuto costante, il rapporto  $\Omega_c/\Omega_b$  varia, ma nella direzione opposta rispetto a quanto illustrato nella precedente figura. La velocità del suono e, quindi, la posizione dei minimi e dei massimi si sposta allora nel verso opposto.

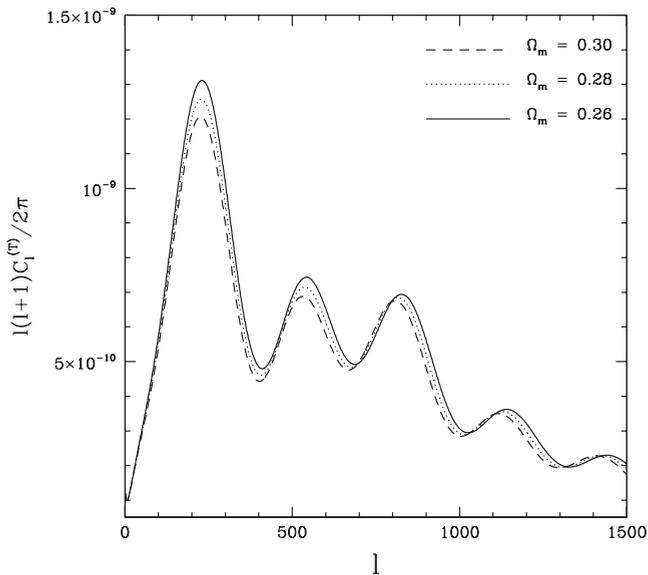


FIGURA XIII.3

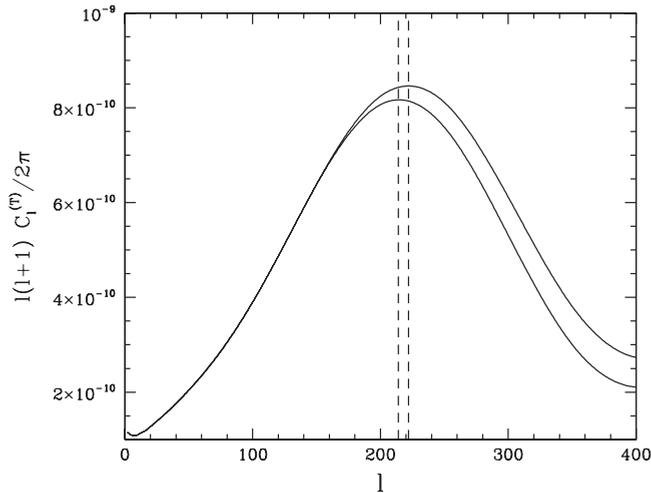
Spettro delle anisotropie del CMB al variare di  $\Omega_m$ . A parte quest'ultimo parametro, il modello è lo stesso di figura XIII.1.

Anche l'ampiezza delle fluttuazioni subisce gli effetti del cambio di  $\Omega_m$ . In questo caso, tuttavia, gli spostamenti al primo picco e al primo minimo sono entrambi nella stessa direzione, verso l'alto o verso il basso: non essendo variato  $\Omega_b$ , l'ampiezza delle onde soniche è rimasta sostanzialmente la stessa.

Crescendo o calando  $\Omega_m$ , invece, tutti gli effetti crescono o calano in modo grosso modo proporzionale.

### ***Variazioni nel valore di $H_0$***

La posizione del primo picco dei grafici nelle figure XIII.1, XIII.2, XIII.3 è fissata dalle fluttuazioni che entrano nell'orizzonte acustico alla ricombinazione. In figura XIII.4 si studia lo spostamento del tale primo picco al variare di  $H_0$ . In effetti, per la distanza dell'orizzonte acustico dall'osservatore valgono le relazioni di proporzionalità  $R_H \propto 300\,000 \text{ (km/s)}/100 h \text{ [(km/s)/Mpc]} \propto h^{-1}$ . Poichè il valore di  $l_{\text{picco}}$  si deduce dal rapporto tra  $2\pi R_H$  e l'orizzonte acustico alla ricombinazione, ci aspetteremo una variazione  $\Delta l_{\text{picco}}/l_{\text{picco}} \simeq \Delta R_H/R_H = \Delta h/h$ .



**FIGURA XIII.4**

Spettro delle anisotropie del CMB al variare di  $H_0$  da 0.65 a 0.75. Le linee tratteggiate sono tracciate in corrispondenza del primo picco.

Per le curve mostrate in figura, risulta  $\Delta h/h \simeq 0.14$ , mentre si osserva una variazione  $\Delta l_{\text{picco}}/l_{\text{picco}} \simeq 0.04$ . La discrasia è dovuta al fatto che  $h$  interviene anche su molti altri fattori, ad esempio sul *redshift* di equivalenza, che cancellano in parte l'effetto atteso.

### Varianza cosmica

A  $l$  inferiori si ha a che fare con scale che, alla ricombinazione, sono al di fuori dell'orizzonte acustico e, ben presto, fuori dall'orizzonte di particella *tout-court*.

Quando si tratta con queste scale occorre tener conto molto seriamente del fatto che, via via che  $l$  cala, i dati riguardano un numero sempre più ridotto di realizzazioni del processo che genera il cosmo. Possiamo considerare i  $2l + 1$  valori dei coefficienti  $a_{lm}$  come la realizzazione del processo di cui  $C_l$  ci dà una misura statistica. In altri termini, abbiamo  $\sim 2l + 1$  misure di ciascun  $C_l$  e la sua incertezza, a livello di 1 variazione standard, sarà

$$(XIII.25) \quad \frac{\Delta C_l}{C_l} \simeq \left( \frac{2}{2l+1} \right)^{1/2}.$$

Si passa dunque da  $\Delta C_l/C_l \sim 0.03$  per  $l \sim 1000$ , a un errore del 30% per  $l \sim 10$ .

Quest'incertezza intrinseca sui segnali previsti è chiamata *varianza cosmica*. In realtà, anche i segnali a basso  $l$  hanno un contenuto di informazione non insignificante, purché si consideri un ampio insieme di  $C_l$ .

Questi elementi sono ancora più significativi ove si pensi che a bassi  $l$  si misurano anisotropie ottenute dalla condizione primordiale delle fluttuazioni, prima che esse siano elaborate dall'idrodinamica.

### Effetti gravitazionali

In particolare, per  $2 < l \lesssim 30$ , l'effetto prevalente nel forgiare le anisotropie è il cosiddetto *effetto SW*, cioè il *redshift* o *blueshift* gravitazionale dei fotoni all'uscita dal campo delle fluttuazioni.

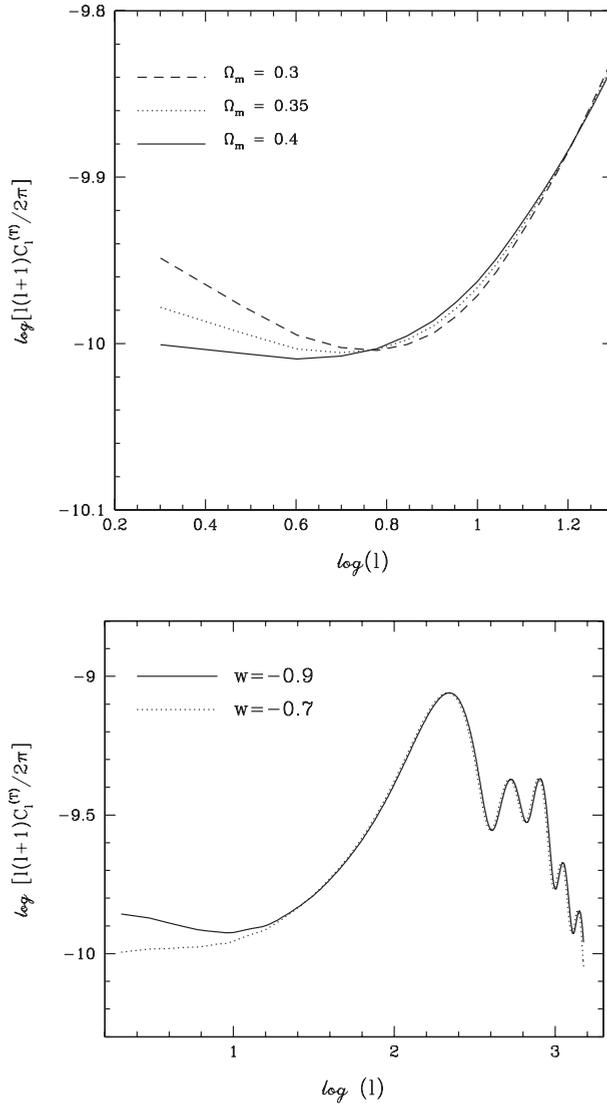
In cosmologie  $\Lambda$ CDM o con altro tipo di DE è anche ben visibile l'effetto ISW, dovuto al cambio nel ritmo di espansione quando la DE diviene una componente significativa della sostanza cosmica.

Su queste scale si delinea un'interessante corripendenza tra ciascun valore di  $l$ , il numero d'onda  $k$  e la lunghezza d'onda comovente  $\lambda_c$  corrispondenti, da un lato, e il *redshift* a cui  $\lambda_c$  è la scala comovente dell'orizzonte, dall'altro.

Difatti, ciascun valore di  $\lambda_c$  contrassegna direttamente un tempo e quindi un *redshift*  $z$ . Se al *redshift*  $z$  il ritmo dell'espansione era già significativamente diverso da quello di una cosmologia SCDM, allora sull' $l$  corrispondente appariranno segnali di effetto ISW.

Nelle figure XIII.5 è magnificata la parte dello spettro delle anisotropie a bassi  $l$ , per mostrare le variazioni nello spettro dovute all'effetto ISW,

- 1) al variare di  $\Omega_m$  e quindi della costante cosmologica, in cosmologie  $\Lambda$ CDM spazialmente piatte;
- 2) al variare del parametro di stato in modelli con  $w = p/\rho$  costante, ma  $\neq 1$ .



**FIGURA XIII.5**

Effetto ISW a bassi  $l$ . Si considerano modelli con diverso  $\Omega_m$ , ( $\Omega_b = 0,04$ ,  $h = 0,7$ ,  $n_s = 1$ ; pannello in alto) o con diverso valore del parametro di stato  $w = p/\rho$  (pannello in basso); in questo secondo caso gli altri parametri cosmologici sono  $\Omega_m = 0,30$ ,  $\Omega_b = 0,05$ ,  $h = 0,7$ ,  $n_s = 1$ .

**Valutazione di parametri tramite la polarizzazione**

In figura XIII.6 riportiamo l'andamento previsto per i 3 spettri T, TE e E.

Sono anche indicate le distorsioni agli spettri TE e E che derivano dall'esistenza di opacità, tra la banda dell'ultimo scattering e  $z = 0$ . Tale opacità è associata alla for-

mazione delle prime stelle e galassie, in un intervallo  $\Delta z_g$  compreso tra  $\sim 20$  e  $3-4$ , ed effettivamente le distorsioni spettrali riguardano gli  $l$  corrispondenti a scale che entrano nell'orizzonte durante  $\Delta z_g$ . La genesi dell'opacità costituisce un ulteriore capitolo di cosmologia, riguardante il cosmo più recente, su cui si sono avuti recenti importanti sviluppi, che in questo testo non viene trattato.

L'opacità ha un modesto effetto sullo spettro  $T$ , che può essere compensato con una variazione di  $n_s$ ; è questo un manifesto caso di degenerazione che si può rompere utilizzando i dati sulla polarizzazione.

La figura XIII.6 ha anche un impatto sulle prospettive sperimentali. Un'opacità relativamente modesta ( $\tau_{op} = 0.04$ ) fa salire lo spettro  $C_l^{(TE)}$ , su piccoli  $l$ , di un fattore  $\sim 100$ . La polarizzazione diviene così misurabile con uno strumento di sensibilità  $\sim 10$  volte superiore agli strumenti che misurano le anisotropie.

Un esperimento mirato alla misura della polarizzazione a basso  $l$ , in vista della possibilità che l'opacità fosse significativa, progettato da ASI, per volare sulla stazione spaziale in ambito ESA, venne tuttavia inopinatamente fermato (dopo che una parte significativa del *budget* era stata spesa), poco prima che WMAP accertasse valori di  $\tau_{op}$  ben oltre 0.04. Citiamo questo esempio per far assaporare al lettore la complessa dinamica, non solo scientifica, che vige nella preparazione degli esperimenti sul CMB, che sono altamente dispendiosi e coinvolgono ampi *teams* di ricercatori.

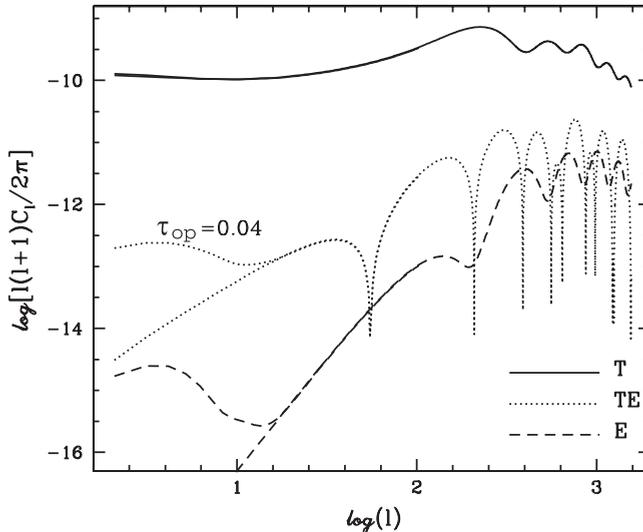
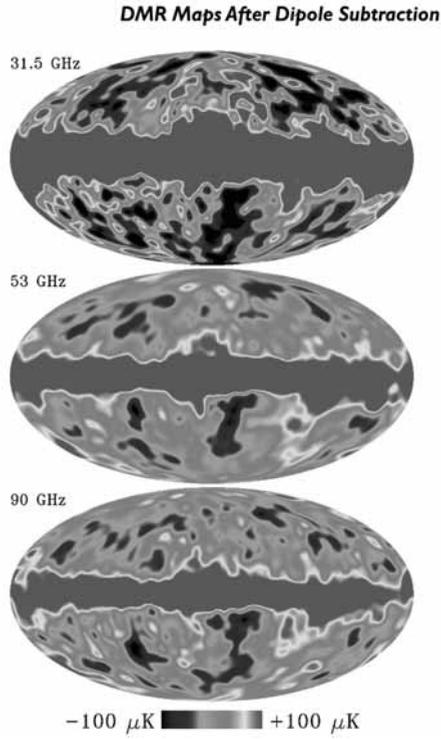


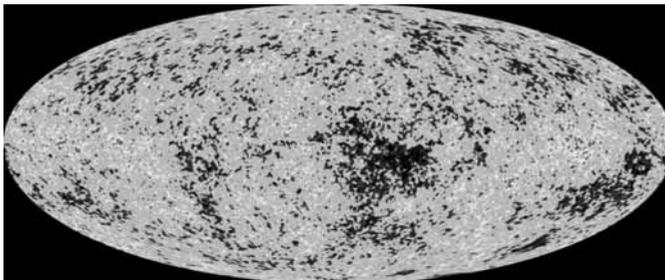
FIGURA XIII.6

Spettri T, TE e E per un modello con  $\Omega_b = 0,03$ ,  $\Omega_m$ ,  $\Omega_b = 0,04$ ,  $h = 0,7$ ,  $n_s = 1$ . Per l'opacità  $\tau_{op}$ , dalla banda dell'ultimo scattering, si considera sia il valore nullo che il valore  $\tau_{op} = 0,04$ , che provoca la «gobba» a bassi  $l$  negli spettri TE e E.



**FIGURA XIII.7**

Immagine del cielo MW visto da COBE a 3 diverse frequenze. La radiazione proveniente dal piano galattico ha un diverso peso in funzione della frequenza. (si veda l'appendice TAVOLE A COLORI).



**FIGURA XIII.8**

Immagine del cielo MW visto da WMAP, da cui è stata sottratta l'immagine della galassia. (si veda l'appendice TAVOLE A COLORI).

## 8. Conclusioni

Abbiamo infatti lasciato sin qui un po' in disparte i dati osservativi. Il dato clamoroso, a questo proposito, al di là dei dettagli degli spettri osservati, è che il loro andamento ricalca chiaramente quanto previsto dalle previsioni teoriche, che sono oggettivamente complesse. In particolare, nei dati sono ormai visibili fino al quarto picco e al terzo minimo delle anisotropie.

In questo testo abbiamo lasciato un po' in disparte gli sviluppi concernenti la polarizzazione; di fatto uno studio esauriente del CMB potrebbe da solo riempire uno spazio equivalente a tutto questo testo. Teniamo peraltro a precisare che anche lo spettro delle correlazioni tra temperatura e modo  $E$  di polarizzazione è stato osservato, mentre ci sono segnali che riguardano anche il puro modo  $E$ .

Abbiamo altresì tralasciato in questo testo tutte le questioni aperte sull'effetto dei *foregrounds* e sul *lensing*, da parte della distribuzione non lineare della materia oggi presente nel cosmo, sui segnali provenienti dal profondo.

Ciò ci collega a uno dei temi più importanti legati allo studio del CMB, la verifica delle previsioni dell'inflazione. Previsioni generiche dell'inflazione sono che le fluttuazioni primordiali sono adiabatiche e hanno natura Gaussiana. La maggior parte dei segnali provenienti dai *foregrounds*, dovuti al *lensing*, e anche all'effetto SZ, invece, introducono elementi di non-Gaussianità.

La sfida primaria in questo campo è già quella di definire un osservabile operativo che misuri direttamente la Gaussianità dei segnali. A ciò segue la necessità di ottimizzare la statistica usata in modo da separare un'eventuale non-Gaussianità primordiale dalla non-Gaussianità introdotta dalla fisica succitata.

Un discorso analogo riguarda la presenza di una componente non-adiabatica nelle fluttuazioni primordiali. Si tratta di due previsioni generiche dell'inflazione e, ove esse fossero «falsificate», si aprirebbe la caccia ai modelli specifici di inflazione che permettono la nascita di anisotropie non-Gaussiane e/o non adiabatiche.

Un'ulteriore previsione generica dei modelli inflazionari è  $n_s \sim 1$ . In questo caso tuttavia, deviazioni di  $n_s$  dall'unità non sono né abnormi né impreviste. Anzi, il comportamento dettagliato di  $n_s$  in funzione della scala è potenzialmente uno strumento per ricostruire la forma del potenziale inflatonico. A livello osservativo, deviazioni dell'indice spettrale da 1 sono già apprezzabili, a livello di 2 deviazioni standard, nei dati osservativi.

Va infine citato, per la sua straordinaria importanza, che va al di là della pura indagine scientifica, l'obiettivo di raggiungere una sensibilità adeguata a misurare il modo  $B$  di polarizzazione. In questo testo non abbiamo dettagliato gli andamenti previsti per la polarizzazione  $B$ , salvo segnalare come l'ampiezza del segnale sia collegata all'intensità delle onde gravitazionali primordiali.

Si tratta di una aspirazione strategica, che aprirebbe un campo di indagine altamente innovativo, nella prospettiva di valutare segnali sulla scala di Planck, all'uscita dal *realm* della *quantum gravity*, nell'ambito di una realtà che oggi descriviamo tentativamente con algoritmi sempre più lontani da quelli cui ci hanno abituato l'osservazione astronomica e gli esperimenti di laboratorio.

# Appendice A1

## La dinamica della distribuzione di fotoni a partire dall'equazione di Boltzmann

La distribuzione nello spazio delle fasi dipenderà dal tempo conforme  $\tau$  secondo l'equazione (di Boltzmann)

$$(A1.1) \quad \frac{\partial f}{\partial \tau} + \frac{\partial f}{\partial x^i} \frac{dx^i}{d\tau} + \frac{\partial f}{\partial p} \frac{dp}{d\tau} + \frac{\partial f}{\partial n_i} \frac{dn_i}{d\tau} = \left( \frac{\partial f}{\partial \tau} \right)_{coll}$$

dove il termine di destra esprime gli effetti degli urti, mentre il termine di sinistra è una specie di «derivata totale», che ci dice come varia complessivamente la distribuzione, sia per la sua dipendenza diretta dal tempo sia per il variare delle coordinate (nello spazio delle fasi) delle singole particelle.

Partendo dall'espressione (XIII.21), che distingue la distribuzione di background  $f_0(p, \tau)$  dalla sua fluttuazione  $\delta f(\mathbf{k}, p, \mathbf{n}, \tau)$ , conviene anzitutto considerare la trasformata

$$(A1.2) \quad \delta f(\mathbf{k}, p, \mathbf{n}, \tau) = \int d^3x e^{i\mathbf{k}\cdot\mathbf{x}} \tilde{\delta f}(\mathbf{x}, p, \mathbf{n}, \tau)$$

che esprime la fluttuazione in numero di fotoni di impulso  $p$  (frequenza  $pc/2\pi\hbar$ ), provenienti dalla direzione  $\mathbf{n}$ , dovuti a un'onda di compressione di numero d'onda  $\mathbf{k}$ , mentre la quantità  $p \times \delta f(\mathbf{k}, p, \mathbf{n}, \tau)$  sarà allora la fluttuazione di intensità della radiazione. Tramite essa si definisce la fluttuazione relativa di temperatura

$$(A1.3) \quad \Theta(\mathbf{k}, \mathbf{n}, \tau) = \frac{1}{4} \frac{\int dp p^3 \delta f(\mathbf{k}, p, \mathbf{n}, \tau)}{\int dp p^3 f_0(p, \tau)},$$

integrando su  $p$ , cosa dettata dal fatto che lo scattering Thomson dei fotoni ha sezione d'urto

$$(A1.4) \quad \sigma(\theta) = \frac{3}{16\pi} \sigma_T (1 + \cos^2 \theta)$$

indipendente dal loro impulso ( $\theta$  è l'angolo di diffusione nel riferimento dell'elettrone a riposo).

Conviene ora porsi nella gauge newtoniana conforme, sicché la metrica è

$$(A1.5) \quad ds^2 = a^2(\tau) [(1 + 2\Psi)d\tau^2 - (1 - 2\Phi)dx_i dx^i];$$

e la variazione dell'impulso delle particelle si ricava dall'equazione del moto

$$(A1.6) \quad p' = p\Phi' - pn_i \partial_i \Psi.$$

Conviene procedere poi all'espansione in armoniche sferiche dell'anisotropia di temperatura, usando  $\mathbf{n}$  come asse polare:

$$(A1.7) \quad \Theta(\mathbf{k}, \mathbf{n}, \tau) = \sum_{l=0}^{\infty} (-i)^l \Theta_l(k, \tau) P_l(\mu);$$

$\mu = \cos \theta$ : coseno dell'angolo  $\theta$  tra  $\mathbf{n}$  e  $\mathbf{k}$ ,  $P_l(\mu)$ : polinomi di Legendre.

Si può ora integrare l'eq. (A1.1) su  $p$  ricavando l'equazione

$$(A1.8) \quad \frac{\partial \Theta}{\partial \tau} + ik\mu\Theta = (\Phi' - ik\mu\Psi) + a n_e \sigma_T [-\Theta + \Theta_0 + i\mu v_b + P_2(\mu)\Theta_2/10],$$

in cui  $v_b$  è il campo di velocità nel fluido barionico e lo scattering Thomson dà luogo al secondo addendo a destra, che contiene il prodotto  $n_e \sigma_T$  (densità in numero degli elettroni liberi  $\times$  sezione d'urto Thomson), mentre è  $\mu = \cos \mathcal{S}$ ,  $\mathcal{S}$  essendo l'angolo tra  $\mathbf{k}$  e  $\mathbf{n}$ .

I passaggi, semplici ma articolati, che portano dall'eq. (A1.1) all'eq. (A1.8) tengono anche conto del fatto che la derivata della direzione,  $d\mathbf{n}/dt$  è dello stesso ordine della fluttuazione, di modo che il termine  $(\partial f / \partial n_i)(dn_i / d\tau)$ , prodotto di due «infinitesimi», non dà contributi al primo ordine.

La scelta di procedere tramite lo sviluppo (A1.7) nasce dal fatto che, fin tanto che i fotoni sono strettamente legati ai barioni, solo  $\Theta_0(k, \tau)$  e  $\Theta_1(k, \tau)$  sono significativamente  $\neq 0$ . L'affievolirsi dell'accoppiamento fa «accendere» gradualmente le altre componenti  $\Theta_l(k, \tau)$ , a partire dagli  $l$  bassi.

$\Theta_0(k, \tau)$  e  $\Theta_1(k, \tau)$  esprimono i termini di monopolo e di dipolo dell'anisotropia di temperatura dovuta alle onde con numero d'onda  $k$ . Lo sviluppo delle anisotropie avviene dunque dopo il disaccoppiamento barioni-radiazione e, fin tanto che una sostanza è un fluido perfetto, le prime due armoniche sono riconducibili alla fluttuazione di densità e al campo di velocità nel fluido, secondo le relazioni

$$(A1.9) \quad \delta(k, \tau) = 4 \Theta_0(k, \tau), \quad v(k, \tau) = \Theta_1(k, \tau).$$

Lo spettro delle anisotropie di temperatura, al tempo attuale  $\tau_0$ , è direttamente collegato alle armoniche sferiche  $\Theta_l$ , tenendo conto del peso statistico delle onde di compressione rispetto al numero d'onda  $k$ , tramite la relazione

$$(A1.10) \quad C_l^{(T)} = (4\pi)^2 \int dk k^2 P(k) |\Theta_l(k, \tau_0)|^2,$$

che richiede lo studio dello sviluppo delle armoniche sferiche fino al tempo attuale. L'eq. (A1.10) non è però di uso corrente, sebbene sia comunque essenziale ricavare le equazioni cui soddisfano le componenti armoniche.

A questo scopo si usano le proprietà di ortonormalità dei polinomi di Legendre, si moltiplica l'eq. (A1.8) per ciascun  $P_l(\mu)$ , si integra su  $\mu$ , e si ottiene la successione di equazioni:

$$(A1.11) \quad \begin{aligned} \Theta'_l &= -an_e \sigma_T \Theta_l + k(u_{l-1} \Theta_{l-1} - u_{l+1} \Theta_{l+1}) \quad (\text{per } l > 2), \\ \text{dove} \quad u_{l-1} &= l/(2l-1), \quad u_{l+1} = (l+1)/(2l+3); \end{aligned}$$

le prime due armoniche soddisfano invece alle equazioni

$$(A1.12) \quad \delta' = -(4/3)kv + 4\Phi', \quad v' = -an_e \sigma_T (v - v_b) + (3/4)kc_s^2 \delta + k\Psi - (k/10)\Theta_2$$

( $c_s = c/\sqrt{3}$ : velocità del suono nella componente fotonica) che differiscono da quelle cui soddisfa un fluido solo per la presenza dell'ultimo termine in  $\Theta_2$ . Infine la terza armonica soddisfa l'equazione

$$(A1.13) \quad \Theta'_2 = -(9/10)an_e \sigma_T \Theta_2 + k[(2/3)v - (3/7)\Theta_3].$$

Si noti che, quando per il cammino libero medio (tempo di collisione) Thomson è  $\lambda_c (= \tau_c) = (n_e \sigma_T)^{-1} \ll (ka)^{-1}$  ( $\sim$  lunghezza d'onda della componente di Fourier della fluttuazione), le eq. (A1.10) si riducono a  $d\Theta_l/\Theta_l = -dt/\tau_c$  e significano il decadimento esponenziale, con tempo caratteristico  $\tau_c$ , di ogni termine di anisotropia oltre il secondo. Lo stesso coefficiente, con analoghi effetti, è presente nell'equazione per  $\Theta_2$ . Infine, esso è presente anche in eq. (A1.11), dove però significa il decadimento esponenziale di ogni differenza tra il campo di velocità nella radiazione e  $v_b$ , campo di velocità nella componente barionica.

## 1. Equazioni del moto per barioni, DM e neutrini

Le equazioni del moto della componente barionica, che si comporta comunque come un fluido, sono assai simili alle eq. (A1.12):

$$(A1.14) \quad \delta'_b = -k v_b + 3\Phi', \quad v'_b = -an_e \sigma_T (4\rho_\gamma / 3\rho_b)(v_b - v) - (a'/a)v_b + k\Psi.$$

Il termine  $a'/a \sim 1/\tau$ , tuttavia, non era presente nel caso dei fotoni. Molte differenze nei fattori numerici nascono invece dal fatto che, nelle onde adiabatiche, è  $\delta_b = (3/4)\delta$ .

Consideriamo queste equazioni, assieme alle eq. (A1.11), quando è  $\tau_c \ll \tau, \lambda$ , riferendoci alla fluttuazione di temperatura  $\delta_T = \delta/4 = \delta_b/3$  e, imponendo che sia  $v = v_b$ . Si ricava allora, sia dalle (A1.12) che dalle (A1.14),

$$(A1.15) \quad \delta'_T = -(k/3)v + 3\Phi', \quad v' = k\Psi + 3kc_s^2\delta_T$$

e, derivando ulteriormente la prima equazione per poi sostituire in essa la seconda, si ottiene

$$(A1.16) \quad \delta''_T + c_s^2 k^2 \delta_T = -(k^2/3)\Psi + \Phi''$$

che è l'equazione di propagazione delle onde nel fluido fotoni-barioni, scritta nella gauge newtoniana conforme. Essa è strettamente analoga all'equazione già usata nel capitolo VIII (vedi, a es., eq. A3.7) e da cui si sono ricavate, in modo euristico, le distribuzioni di picchi e minimi nello spettro delle anisotropie. La novità più importante consiste nel lato destro dell'equazione, che esprime le «forze gravitazionali» nella gauge newtoniana, invece che in quella sincrona.

Le equazioni per la componente di DM sono assai più semplici; si scrivono

$$(A1.17) \quad \delta'_c = -k v_c + 3\dot{\Phi}, \quad v'_c = -(a'/a)v_c + k\Psi$$

e coincidono con quelle dei barioni, tranne il termine di interazione.

Le equazioni per la componente di neutrini, supposti privi di massa, sono allo stesso modo analoghe a quelle della componente fotonica, salvo la soppressione dei termini di interazione coi barioni.

Nel caso di neutrini con massa, invece, la situazione è complicata dalla necessità di distinguere tra diversi impulsi  $p$ . Non entreremo nel merito di queste complicazioni rimandando alla letteratura specializzata.

## 2. Equazioni per il campo gravitazionale

La differenza tra i potenziali gravitazionali  $\Phi$  e  $\Psi$  sussiste solo se sono presenti sforzi anisotropi, vale a dire il termine  $\Theta_2$  nello sviluppo delle anisotropie dei fotoni (e dei neutrini). Risulta infatti

$$(A1.18) \quad k^2(\Psi - \Phi) = -(8\pi/5)G a^2 (\rho_\gamma \Theta_2 + \rho_\nu \Theta_2^{(\nu)}).$$

La dinamica dei campi si può allora determinare tramite l'equazione

$$(A1.19) \quad k[\Phi' + (a'/a)\Psi] = 4\pi a^2 \Sigma_j (\rho_j + p_j) v_j$$

dove la somma è estesa alle varie componenti (fotoni, barioni, DM, neutrini) di densità e pressione (di background)  $\rho_j$  e  $p_j$ .

## Appendice A2

### Calcolo delle anisotropie di temperatura

La soluzione dell'equazione (A1.8) è stata riportata alla soluzione del sistema di equazioni differenziali (A1.11)-(A1.12)-(A1.13). Essa è invece anche equivalente all'equazione integrale

$$(A2.1) \quad \Theta(k, \mu, \tau_0) = \int_0^{\tau_0} d\tau e^{ik\mu(\tau-\tau_0)} \left[ e^{-\tau_{op}(\tau)} a n_e \sigma_T \left( \Theta_0 + i\mu v_b + P_2(\mu)\Theta_2/10 \right) + \dot{\phi} - ik\mu\Psi \right]$$

$$\text{con} \quad \tau_{op}(\tau) = \sigma_T \int_{\tau}^{\tau_0} d\bar{\tau} a(\bar{\tau}) n_e(\bar{\tau})$$

La verifica si effettua derivando rispetto a  $\tau_0$  e tenendo conto che la derivazione deve tenere conto sia della dipendenza da  $\tau_0$  all'estremo superiore dell'integrale che nell'argomento del primo esponenziale.

L'eq. (A2.1) si può anche scrivere

$$(A2.2) \quad \Theta(k, \mu, \tau_0) = \int_0^{\tau_0} d\tau e^{ik\mu(\tau-\tau_0)} \mathcal{S}(k, \tau)$$

$$\text{con} \quad \mathcal{S}(k, \tau) = V(\tau) \left( \Theta_0 - v_b'/k + \Psi - \Theta_2/20 - 3\Theta_2'/4k^2 \right) + e^{-\tau_{op}(\tau)} (\Phi' + \Psi') -$$

$$(A2.3) \quad -V'(\tau) (v_b/k + 3\Theta_2'/20 k^2) - V''(\tau) 3\Theta_2/20 k^2,$$

$$\text{dove} \quad V(\tau) = a(\tau)n_e(\tau)\sigma_T e^{-\tau_{op}(\tau)}$$

costituisce la «funzione di visibilità», che coincide con la «banda dell'ultimo scattering» definita nel capitolo VIII e ivi studiata in dettaglio in appendice A2. Il passaggio da eq. (A2.1) a eq. (A2.2)-(A2.3), semplice ma laborioso, si fa integrando per parti.

Alcuni termini della «sorgente»  $\mathcal{S}(k, \tau)$  hanno un senso trasparente:

- 1) I primi due termini,  $\Theta_0$  e  $v_b'/k$ , si riportano alla fluttuazione di densità nell'onda acustica (la derivata della velocità ha un comportamento simile all'ampiezza); il campo di velocità entra invece direttamente nel termine  $V'v_b/k = Vv_b/(k\sigma_x)$  nella seconda riga, dove  $\sigma_x \equiv V/V'$  esprime (approssimativamente) la profondità lungo  $\tau$  della banda dell'ultimo scattering, come si verifica attribuendole un andamento simile alla Gaussiana (vedi appendice A2 del capitolo VIII). In sostanza, l'effetto Doppler deve agire coerentemente lungo la banda dell'ultimo scattering: per  $k$  grandi (piccole lunghezze d'onda) è possibile un'alternanza della direzione della velocità nella banda, che cancella gran parte del segnale da effetto Doppler.
- 2) Il terzo termine esprime invece l'effetto del potenziale gravitazionale sui fotoni, nella banda dell'ultimo scattering, cui abbiamo accennato.
- 3) Il secondo addendo sulla prima riga, in cui compare il fattore  $e^{-\tau_{op}}$ , esprime il cosiddetto *effetto Sachs & Wolfe integrato* (ISW: *integrated Sachs & Wolfe*). Esso agisce durante la ricombinazione, ma ha effetti assai più significativi a *redshift* bassi, quando cambia il ritmo di espansione del cosmo, per l'intervento della DE.

- 4) Tutti gli altri termini contengono  $\Theta_2$ . Essi sarebbero nulli se i fotoni costituissero un fluido. Il loro contributo è invece essenziale, non solo a fornire un calcolo esatto delle anisotropie del CMB, ma soprattutto perché da essi, in un modo che non andremo a dettagliare, trae origine la polarizzazione del CMB.
- 5) Nel termine di sorgente  $S$  compaiono le armoniche sferiche  $\Theta_l$  per  $l = 0$  e  $2$ ; queste sono le incognite del sistema di equazioni differenziali (A1.11)-(A1.12) che occorre conoscere, per conoscere la «sorgente».

Si noti ora un fatto essenziale: il passaggio da eq. (A2.1) a eq. (A2.2)-(A2.3) porta la dipendenza da  $\mu$  solo all'esponenziale. Volendo quindi riprendere l'espansione di  $\Theta(k, \mu, \tau_o)$  in armoniche sferiche, basta espandere l'esponenziale, che è ben noto essere esprimibile come  $\sum_l j_l P_l$ , dove le  $j_l$  sono funzioni di Bessel. Sarà quindi

$$(A2.4) \quad \Theta_l(k, \tau_o) = \int_0^{\tau_o} d\tau j_l[k(\tau - \tau_o)] S(k, \tau)$$

Ne consegue che, note le prime 3 armoniche sferiche lungo tutto il cammino ottico, da questa equazione si possono ricavare tutte le armoniche sferiche, al tempo  $\tau_o$  (che potrebbe essere un tempo generico, ma operativamente sarà il tempo attuale, in cui esse vengono osservate).

## Il tempo e l'entropia

### 1. Topologia e metrica del tempo

Per sua stessa natura la cosmologia estrapola il nostro tempo, segnato dalla vita e dalla storia, a domini estranei, in cui esso è segnato da dati fisici di altra stazza, in primo luogo dalla temperatura della radiazione di fondo microonde (CMB), che per noi è invece definitivamente  $\sim 2.73$  K. Quando il plasma primordiale diventa gas atomico essa era  $\sim 3000$  K, quando si accendono le prime stelle essa era  $\sim 100$  K, quando il Sole inizia a splendere essa era  $\sim 6$  K.

Oltre la vita e la storia, possiamo apprezzare la modestia degli stessi tempi paleontologici, rispetto alla cosmologia: un *tyrannosaurus rex* che avesse puntato il suo radiometro verso il cielo microonde, avrebbe misurato  $\sim 2.73$  K, più o meno come noi. Viceversa, quando tutte le terre emerse erano riunite in *Pangea*, donde sarebbe partita la deriva dei continenti, la temperatura del CMB era di poco superiore a 2.8 K: anche i tempi geologici sono quasi un nulla!

La cosmologia si immerge poi in temperature ben oltre i 3000 K della ricombinazione; scopre le ere primordiali, dove la fisica nota è chiamata a agire su scale di tempo brevissime; affronta il *range* energetico che supera i confini della sperimentabilità diretta, su scale di tempo ancora più brevi. Su tutte le scale temporali essa mantiene tuttavia l'assunto che la fisica in azione è la stessa che agisce oggi sulle scale dei Giganni che trascorrono.

Di conseguenza, in tutto il *background* fisico, il tempo mantiene lo *status* di «variabile indipendente», come nelle applicazioni di laboratorio; ma la cosmologia non si può fermarsi a questa definizione, non può rinunciare a approfondire la natura del tempo, una volta che i suoi dati indicano che tutto scaturisce da un *tempo zero*, al *Big-bang*; qui il riferimento alla «fisica nota» non giova oltre, non si sperimenta un inizio dei tempi in laboratorio.

Va anche sottolineato che la natura del *tempo* è stata spesso considerata *cosa nostra* dai filosofi, da Anassimandro a Heidegger, e il ricercatore deve distinguere ciò che ha valenza scientifica da valutazioni psicologiche o metafisiche (senza nulla togliere a filosofi come Agostino e alle loro profonde intuizioni).

In una trattazione scientifica della natura del tempo, il punto di partenza sta nel

confronto tra variabile temporale e variabili spaziali. In particolare, nella descrizione dello spazio siamo abituati a distinguere la topologia dalla metrica e, in effetti, se soggetto a misura, lo spazio diviene distanza. Ma lo spazio ha una flessibilità che il tempo non conosce: lungo una coordinata spaziale possiamo muoverci, tornando poi sui nostri passi e la distanza non altera la nostra natura; il tempo invece ci costringe, agisce sul corpo e sull'anima, e un libero *viaggio nel tempo* resta tra i sogni della fantascienza. Così, l'usuale concetto di tempo racchiude in sé, senza che siamo avvezzi a distinguerle, sia una nozione topologica che una nozione metrica.

La topologia del tempo è semplice, è la topologia della retta; o della semiretta, se il tempo include un'origine. È grazie alla semplicità di questa topologia che il tempo ha un *verso*, una nozione topologica applicabile ai soli domini unidimensionali. Se il tempo ha un'origine, ci dobbiamo poi chiedere se essa sia inclusa nel tempo, vale a dire se la semiretta sia *chiusa* o *aperta*, se la sua estremità faccia parte o meno della semiretta stessa. In altri termini: l'istante del *Big-bang* ha realtà fisica?

Di fatto, come vedremo nel prossimo paragrafo, i contorni di questa domanda sono molto sfocati. Inclusione o esclusione dell'origine, tuttavia, hanno un risvolto immediato: Una semiretta aperta si può mappare su una retta, mentre ciò è impossibile per una semiretta chiusa. Per maggiore concretezza, supponiamo di usare sull'asse temporale la metrica definita dall'usuale tempo  $t$  e denominiamo  $t_P$  il tempo di Planck. Se, come suggerito da Harrison, passiamo alla variabile logaritmica

$$(XIV.1) \quad s = \log(t/t_P),$$

l'ordinamento degli eventi lungo l'asse temporale non è cambiato. Il cambio di metrica non ha ovviamente influito sulla topologia. Gli eventi catalogati da  $t$  sono nella stessa sequenza anche se catalogati da  $s$ .

Ma vediamo più in dettaglio cosa implica questa mappa. Ricordiamo i valori del tempo attuale,  $t_0 \sim 10^{17}$ sec., e del tempo di Planck,  $t_P \sim 10^{-43}$ sec.. A partire da  $t_P$  il rapporto  $t/t_P$  varia dunque da 1 a  $10^{60}$ . Il suo logaritmo varia allora da 0 a 60. Se estendiamo la mappa a  $t < t_P$ , cioè al di sotto di  $10^{-43}$ sec., avremo  $t/t_P < 1$  e quindi  $s < 0$ . Al tendere del rapporto  $t/t_P$  a 0,  $s$  tende a  $-\infty$ . Una mappa così definita si può fare solo se l'istante  $t = 0$  non fa parte dei tempi *fisici* (non esiste il logaritmo dello zero!).

Quindi, l'asse  $s$  appare diviso in due parti distinte: per  $s$  positivi, siamo dopo il tempo di Planck e *oggi* è marcato dal valore  $s_0 \simeq 60$ ; prima, invece, siamo nel dominio preplanckiano e i valori ammessi sono  $-\infty < s \leq 0$ . La zona positiva è finita, addirittura «piccola», 0-60; il *dominio preplanckiano*, invece, ha un'estensione illimitata. Questa e simili mappe spostano il *Big-bang* a  $-\infty$ ; non ci dobbiamo più «preoccupare» di capire come i tempi hanno avuto origine: il cosmo esiste *da sempre*.

Tra le varie metriche che possiamo istituire sull'asse temporale, limitato o illimitato, il tempo  $t$  segnato dai nostri *orologi* ha peraltro molte ragioni a suo favore. La madre di tutti gli orologi è il ciclo giorno-notte. A esso si aggiungono poi il ciclo annuale, col succedersi delle stagioni, e il ciclo delle fasi lunari. Il numero di cicli diurni in

un ciclo annuale è costante; con discreta approssimazione lo stesso si può dire per il ciclo lunare. Con buona approssimazione è anche costante il numero di battiti cardiaci al dì e la *battuta di polso* galileiana preconizza l'unità di misura c.g.s., il secondo!

Abbiamo quindi una serie di fenomeni ciclici *co-periodici*, se usiamo la metrica ordinaria del tempo. I cicli astronomici nascono tutti da moti rotatori, mentre gli orologi moderni fanno riferimento alla periodicità di fenomeni oscillatori. Ma, in definitiva, tutti gli orologi, fanno riferimento a uno di questi fenomeni ciclici, rotazione o moto armonico.

Non è quindi un caso che i soli problemi che si riesce a risolvere esattamente, in meccanica classica o quantistica, siano quelli del rotatore (atomo di idrogeno) e dell'oscillatore armonico. È la scelta stessa della metrica sull'asse temporale che li favorisce.

Ci sono però altri dati cosmologici che non favoriscono direttamente questa scelta metrica. Abbiamo fatto riferimento numerose volte, in questo testo, al *tempo conforme*. A esso ci si può utilmente riferire, ad esempio, per studiare le onde soniche prima della ricombinazione. Esse sono (quasi-)periodiche rispetto al tempo conforme che, grosso modo, è proporzionale a  $t^{1/3}$ , mentre la durata di un'oscillazione è variabile rispetto al tempo ordinario. Non si tratta di una nozione teorica, da ciò dipende la posizione dei picchi e dei minimi osservati nello spettro delle anisotropie angolari del CMB.

Un «osservatore» operante all'epoca della ricombinazione, avrebbe dunque privilegiato una metrica, sull'asse temporale, basata sul tempo conforme. Rispetto a tale metrica, il ciclo stagionale non è periodico e gli anni hanno durata crescente; ma per un tale «osservatore» l'articolazione planetaria, da cui una tale nozione dipende, sarebbe stata una realtà futura difficilmente prevedibile.

C'è dunque una larga arbitrarietà nella scelta della metrica del tempo. Alla fin fine è necessario rendersi conto che l'idea che l'istante del *Big-bang* abbia realtà fisica è strettamente legata all'uso di una metrica figlia di dati della nostra era cosmica. Sono questi dati che ci spingono a riproporci la domanda antica «cosa c'era prima della Creazione del Mondo», rfrasata in «cosa c'era prima del *Big-bang*».

L'uso abituale e privilegiato del tempo  $t$  è invece dettato dalle condizioni tipiche della nostra era, e il riferimento a un *tempo-zero*, nonché tutte le fantasie su ciò che era a monte di questo tempo, sono conseguenza delle condizioni fisiche della nostra era cosmica. Con altre scelte metriche non si pone neppure il problema della fisicità del *Big-bang*; esso si sposta all'infinito, oltre cui non si concepisce nessuna realtà temporale.

Non si può dunque evitare di concludere che *l'idea che il Big-bang implichi un'origine dei tempi è una nozione strettamente antropocentrica*.

## 2. Il dominio quantistico

Nel precedente paragrafo si è usata una certa cautela nel riferirsi a valori di  $t < t_p$  o di  $s < 0$ , e con ragione! Di fatto, all'avvicinarsi di  $t$  a  $t_p$  si entra nel dominio della

gravità quantistica e, ancor prima, possono entrare in gioco dimensioni spaziali oggi *compattificate*, sicché si modifica la topologia dello spazio, e in parallelo potrebbe modificarsi la stessa topologia del tempo. Se essa non fosse più unidimensionale, potrebbe essere impossibile porre gli eventi cosmici in sequenza: addio verso, *freccia* del tempo!

Stiamo parlando di un dominio fisico che è stato sinora esplorato solo a livello teorico, soprattutto nell'ambito delle *teorie di corda (string)*. La nostra nozione di spazio implica che il suo *elemento* è il punto; in una teoria di corda, l'elemento è la corda stessa. Si tratta di una nozione fortemente anti-intuitiva e come tale difficile da convogliare se non in termini matematici. Va aggiunto che la visuale prevalente delle teorie di corda fa riferimento a uno spazio 11-dimensionale con una sola coordinata temporale.

Vi sono stati numerosi tentativi di articolare una cosmologia alla luce di queste elaborazioni teoriche. In questo testo abbiamo deliberatamente lasciato questo argomento in disparte, ma su di esso c'è molto interesse e vengono spese significative risorse.

Le attuali possibilità di trovare riscontri sperimentali a simili elaborazioni teoriche si fondano sull'esistenza di onde gravitazionali primordiali. Il cosmo è trasparente alle onde gravitazionali sin dal tempo di Planck e una loro rivelazione porterebbe informazione su eventi a ridosso della transizione quantistico-classico.

L'osservazione diretta di tali onde gravitazionali non è fattibile nell'ambito di tecnologie oggi disponibili, mentre una loro rilevazione indiretta potrebbe essere basata su dati riguardanti il *modo B* di polarizzazione del CMB.

Abbiamo già trattato questo tema nel precedente capitolo. Possiamo qui ricordare che ci sono comunque grosse difficoltà a muoversi anche su questo terreno. Anzitutto perché è difficile prevedere l'intensità stessa di tali onde e del segnale che ne potrebbe derivare: nella migliore delle ipotesi, per percepire un segnale del genere occorrerebbe acquisire una sensibilità  $\sim 10^{-3}$ - $10^{-2}$   $\mu$ K, mentre le anisotropie danno luogo a un segnale medio  $\sim \mathcal{O}(10 \mu$ K). Ma è possibile che il segnale sia anche 1000 volte meno intenso. Inoltre si deve depurare il modo-*B* da segnali sovrapposti, di origine remota ma non primordiale. È certo, ad esempio, che il *lensing gravitazionale* trasforma una quota della polarizzazione *E* in polarizzazione *B*. Infine c'è oggi un grave problema di *foregrounds*, di contaminazione da segnali galattici o comunque «locali».

A fronte di queste difficoltà, l'atteggiamento nei riguardi del dominio quantistico può risultare influenzato dal *bias* ideologico del ricercatore.

In più d'una visuale teorica si supera il «muro» del *Big-bang* e si va a descrivere una fisica *pre-Big-bang*. È inevitabile notare che queste elaborazioni sono ancorate alla persistenza di una metrica del tempo simile a quella stabilita nella nostra era cosmica, e portano alla conclusione che tale tempo *t non ha inizio*. Sarebbe fortemente diminutivo valutarle per questo come esposizioni in chiave fisico-matematica dei pregiudizi ideologici di taluni autori; siamo qui in un contesto che ha fondamenti fisici

di ben altro spessore, rispetto alle elaborazioni legate ai modelli di *stato stazionario* negli anni Trenta-Cinquanta.

Non c'è dubbio, invece, che lo spazio e il credito che queste visuali hanno spesso trovato nei *media* è solo giustificato dalla diffusione dei pregiudizi *scientisti*, motivati da contingenze del tutto estranee alla ricerca, al di fuori del mondo degli «addetti-ai-lavori». C'è invece assoluto accordo, nel mondo scientifico, che queste elaborazioni hanno un grande valore a livello di sviluppo teorico, ma che non si pone neppure il problema di una loro attendibilità finché non se ne dedurranno conseguenze fisiche sperimentabili.

C'è una sorta di simmetria tra questa accentuazione e quella dettata dall'opposto *bias* ideologico, di chi vede nel *dominio pre-planckiano* la natura incorrotta, allo stato assolutamente potenziale, che solo l'*Adam Kadmon* costituito dall'osservatore traduce in mondo fisico. Questa visuale è assai meno pubblicizzata della precedente ed è stata indipendentemente suggerita da studiosi di fisica con cognizioni concernenti diversi filoni tradizionali.

Uno di essi è quello *giudaico-cabbalistico* e, indubbiamente, la *Cabbalàh* è il più importante filone esoterico sviluppatosi nel mondo occidentale nei due millenni dell'*era volgare*. Rimandando chi fosse veramente interessato agli studi specifici, val la pena di citare il lavoro di fisici come *Joel Primack* e altri, che hanno visto nell'espansione cosmologica il riflesso sul mondo fisico dello *Tsinsum* luriano, il ritrarsi, l'*esilio* di Dio, all'uscita dal mondo pre-planckiano. L'Adamo primigenio (*Adàm Kadmon*), essere super-umano antecedente la stessa creazione del mondo in cui viviamo (non l'unica creazione, secondo alcuni cabbalisti, ma quella *riuscita*), ha un ruolo essenziale nel provocare lo *Tsinsum*, con la sua ambizione a conoscere, che viene riletta come *osservazione* del cosmo.

Un altro filone che va citato, in termini oggettivamente paralleli, è quello *vedico*. In esso la natura incontaminata, *Prakrti*, che contiene miscele indifferenziate dei tre *guna* (cfr. capitolo I), viene forgiata dall'intervento di *Purusha*, pur restando incontaminata (*vergine*) in sé, a generare realtà fisiche in cui la miscela dei *guna* (luce, azione, gravità; cfr. *c*,  $\hbar$ , *G*) è specificamente definita. Il parallelo tra l'intervento di *Purusha* e la fuoriuscita dal mondo preplanckiano è stato oggetto di varie speculazioni, tra cui citiamo quelle di *Nicolò Dallaporta*.

La storia della fisica è piena di avventure ideologiche basate sulle cognizioni scientifiche cui si era giunti in varie fasi dello sviluppo della ricerca fondamentale. A distanza di qualche decennio la maggior parte di queste elaborazioni si dimostrano superate dall'ulteriore *input* scientifico. Questa ci pare la riserva essenziale da formulare a fronte di queste e altre elaborazioni che, a nostro avviso, mettono soprattutto in luce i profondi legami tra le costruzioni esoteriche nate nell'ambito di tradizioni tra loro remote. Legami nati nelle profondità universali della  $\psi\upsilon\chi\eta$  umana.

L'autore tiene peraltro a non nascondere il fascino profondo che questo secondo tipo di elaborazioni ha per lui. Ciò è tra le ragioni specifiche per cui alcune parti di questo libro sono state scritte.

### 3. La microfisica e il tempo

C'è in realtà un ulteriore elemento che sembra rendere peculiare la metrica  $t$  sull'asse temporale, quello dell'invarianza delle forze microscopiche rispetto all'operazione  $\mathcal{T}$  (*time reversal*: inversione del tempo).

Le lagrangiane elettromagnetica a forte sono invarianti per  $\mathcal{CP}$  ( $\mathcal{C}$ : coniugazione di carica;  $\mathcal{P}$ : parità) e quindi, assumendo solo la «località» dell'interazione, se ne deduce anche l'invarianza rispetto a  $\mathcal{T}$ . Si potrebbe allora dubitare che, a fronte di una deformazione della metrica sull'asse temporale, la simmetria possa cadere.

In realtà le equazioni del moto in ogni teoria di gauge con accoppiamento minimale, in assenza di termini di massa fermionici o bosonici, sono invarianti rispetto al cosiddetto *gruppo conforme*. La metrica FRW si può comunque scrivere

$$(XIV.2) \quad ds^2 = a^2(\tau) [d\tau^2 - d\lambda^2]$$

e l'invarianza rispetto al gruppo conforme significa che deformazioni sostanzialmente arbitrarie sono lecite sull'asse  $\tau$  e sugli assi spaziali contenuti in  $\lambda$ . La rottura della simmetria conforme va addebitata ai «termini di massa»; nell'ignoranza di quale sia il giusto modello inflazionario non è allora del tutto chiaro se ciò avvenga già in prossimità del tempo di Planck o alla rottura della simmetria  $GUT$ .

L'invarianza per  $\mathcal{T}$  può essere rifrasata usando la coordinata  $\tau$  e dunque, fino all'entrata in gioco di termini di massa, non c'è alcuna coordinata temporale privilegiata.

A fronte di questi dati microscopici sta il fatto che il macrocosmo è manifestamente non-invariante rispetto a  $\mathcal{T}$ . La distinzione corrente tra passato e futuro è legata all'esistenza di sistemi macroscopici in *disequilibrio termodinamico*. In effetti riconosciamo come *futuro* la direzione verso cui l'equilibrio s'accresce. Ma, se tutti i sistemi macroscopici fossero in equilibrio termodinamico, non ci sarebbe questa via per distinguere tra passato e futuro. L'asimmetria però permarrrebbe, legata alla difficile misurazione della temperatura calante del CMB.

Da questa disamina nasce un'idea non del tutto chiarita, che il meccanismo che presiede alla nascita di sistemi in disequilibrio sia associato alla rottura della simmetria conforme e quindi alla possibilità di *misurare*  $T$ , nonché alla conseguente selezione di particolari metriche sull'asse temporale.

È peraltro certo che lo spazio fisico e termodinamico per l'esistenza di osservatori nel cosmo è legato alla presenza di disequilibrio. La configurazione di equilibrio, per un sistema macroscopico, è quella di *massima entropia*. In un sistema in equilibrio non esiste quindi alcuna *informazione* e tanto meno lo spazio termodinamico per l'esistenza di sistemi organizzati e vitali.

In un prossimo paragrafo discuteremo come i sistemi cosmici hanno acquisito *informazione* nel corso dell'espansione cosmologica. Si tratta di una comprensione basata sull'accettazione dei dati sulla dinamica espansiva e che rimanda quindi ancora una volta alle ragioni di fondo per cui tale dinamica sussiste.

#### 4. Sistemi chiusi e aperti

Il secondo principio della termodinamica, durante alcuni decenni a cavallo tra l'Ottocento e il Novecento, sottopose a tensione la fisicità del metabolismo degli organismi viventi, per i quali non si assiste a un sostanziale degrado entropico, se non quando «sono abbandonati dall'anima». Va subito affermato che la termodinamica della vita è oggi del tutto compresa, mentre ci sta a cuore rimarcare subito che *l'esistenza stessa della vita è frutto dell'espansione cosmica*.

Ricordiamo anzitutto la definizione dell'entropia di un macrostato di un sistema,

$$(XIV.3) \quad S = k_B \ln N + \text{const.};$$

$k_B$ , al solito, è la costante di Boltzmann;  $N$  è il *numero dei microstati* in cui il sistema potrebbe trovarsi, per dar luogo al *macrostato* osservato. Ad esempio, se scambiamo un atomo di carbonio nella mascella di un animale con un atomo di carbonio nel suo cervello, otteniamo un altro microstato dell'animale stesso, ma l'animale non si sarebbe accorto di nulla, essendo sempre nello stesso macrostato.

Ciononostante il numero  $N$  di microstati dell'animale, pur enorme, è un nulla rispetto al numero dei microstati possibili per i suoi componenti atomici. Finché l'animale è in vita, si assiste a una sua evoluzione tra svariati macrostati, ma tutti caratterizzati da un valore di  $N$  *relativamente* piccolo, sempre dello stesso ordine. La morte, invece, porta il corpo dell'animale verso macrostati cui corrisponde un numero  $N$  sempre crescente. La crescita dell'entropia  $S$  è la crescita di  $N$ , sinché gli stessi elementi microscopici si disperdono nell'ambiente e non ha neppure senso ricercarne l'organizzazione.

Accanto all'entropia  $S$  possiamo definire il suo opposto, la *neghentropia*  $-S$ ; quando  $S$  cresce la neghentropia cala, perché  $S$  diminuisca occorre che ci sia un *input* di neghentropia.

La maggior parte delle attività fisiche dell'animale, finché esso è in vita, causano una crescita di  $S$ : muoversi, camminare o correre, guardare, annusare, accoppiarsi, sono tutti processi irreversibili che causano consumo di neghentropia. Se il valore complessivo di  $S$  (ed  $N$ ), nel corpo dell'animale, resta sostanzialmente costante finché l'animale vive, ciò è dovuto a altre specifiche attività, che implicano l'*input* di neghentropia dall'ambiente, *in primis* l'alimentazione.

Vale la pena qui di riferirsi brevemente alla catena alimentare, che porta dai carnivori agli erbivori, dagli erbivori alle piante, e dall'informazione ricavata tramite l'alimentazione alla sorgente primaria di tutta l'informazione dei viventi terrestri, la sintesi clorofilliana; essa sfrutta l'informazione presente nella luce solare, corpo nero estremamente *diluito* e quindi ricchissimo di neghentropia.

Viene spesso sottolineato l'afflusso di *energia* dal Sole alla Terra. Ma quale luogo è più ricco di energia e quale luogo è meno ospitale per la vita, della superficie o dell'interno solare? Lo spazio termodinamico per la vita terrestre non nasce dall'energia presente nello spettro solare, ma dalla distribuzione dell'energia sulle frequenze. Lo

spettro solare ha un picco d'intensità legato alla temperatura superficiale di  $\sim 6000$  K; una volta termalizzata, la luce solare, alla superficie terrestre, ha un picco attorno ai 300 K. È nel processo di termalizzazione, se attuato tramite la sintesi clorofilliana, che avviene la cessione di informazione dalla luce ai viventi.

La vitalità di un animale o di una pianta, come sistema organizzato, si lega dunque alla sua capacità di sfruttare la neghentropia che gli giunge dall'esterno, per compensare il degrado entropico che nasce dalle sue funzioni vitali.

Ciò fa sì che l'animale o la pianta sia un sistema *aperto*. La termodinamica dell'Ottocento pensava che nessun sistema fosse individuabile, univocamente determinato, se i suoi componenti elementari non erano fissati. Un essere vivente ha invece un'individualità che discende dalla sua organizzazione, non dall'elencazione dei suoi componenti elementari, che sono in continua evoluzione e sostituzione.

Le funzioni vitali sono dunque processi fisici. Col cessare delle funzioni vitali, ineluttabilmente, cesseranno i processi fisici che sono di supporto allo psichismo, più o meno avanzato, a seconda della specie animale. Se così vogliamo pensare, cessa di esistere il supporto fisico dell'*anima*.

## 5. L'informazione

Non c'è molta differenza tra

$$(XIV.4) \quad \begin{array}{ll} \text{neghentropia:} & -S = \text{const.} - k_B \ln N \\ \text{e informazione:} & I = \log_2(N_{max}/N) = \log_2 N_{max} - \log_2 N. \end{array}$$

Al di là dell'omissione della costante di Boltzmann  $k_B$  e dell'uso del *logaritmo in base 2* in luogo del *logaritmo in base e*, la differenza sostanziale dell'informazione, rispetto all'entropia, sta nella fissazione della costante additiva che è ora il *logaritmo* del numero massimo ( $N_{max}$ ) di microstati che un macrostato del sistema possa avere.

Il rapporto  $N/N_{max}$ , in effetti, si dice anche *probabilità termodinamica* del sistema. Essa vale 1 quando il sistema si trova nella configurazione più probabile. L'informazione  $I$  vale allora 0: nella configurazione più probabile il sistema non contiene informazione.

Il secondo principio della termodinamica, dunque, si traduce facilmente da entropia a informazione; l'ineluttabile crescita dell'entropia pare tradursi nel calo, comunque ineluttabile, dell'informazione, per ogni sistema limitato.

Su questo punto si è avuto un contributo fondamentale di *Alfonso Cavaliere*, il quale ha reso esplicita un'osservazione che molti avevano compiuto circa il contesto cosmologico: se in laboratorio  $N_{max}$  è oggettivamente costante, è ancora vero che, nell'espansione cosmica, esso non sia destinato a aumentare?

Si tratta di un punto di importanza fondamentale: se l'espansione può causare una crescita di  $N_{max}$ , ci si può chiedere se sarà più rapida tale crescita espansiva o la crescita di  $N$ , legata all'irreversibilità dei fenomeni naturali. In effetti, se  $N_{max}$  batte  $N$ , l'informazione naturalmente aumenta.

In altri termini, al di là delle sottigliezze sulla definizione della costante additiva nell'eq. (XIV.4), sarebbe come avessimo individuato un meccanismo che fa decrescere l'entropia.

Che  $N_{max}$  possa crescere, nell'espansione cosmologica, appare del tutto intuitivo, come conseguenza dell'aumento nei volumi a disposizione, in cui divengono possibili sempre più numerose distribuzioni degli elementi microscopici. Il punto sta nel raffronto tra questo ritmo di aumento e il ritmo con cui, naturalmente, va crescendo  $N$ .

I termodinamici dell'Ottocento facevano riferimento all'*Universo* come sistema *chiuso*, che conteneva e assommava tutti i sistemi fisici. Molti libri di termodinamica recano ancora la conclusione classica che i termodinamici di allora ne traevano: *l'entropia dell'Universo è destinata a crescere*, comunque, sempre. Il cosmo doveva quindi essere stato dotato, *ab initio*, di un'elevata dose di neghentropia e era destinato ineluttabilmente alla *morte entropica*, quando tale dose si fosse esaurita e tutto avesse raggiunto l'equilibrio nella configurazione più probabile.

La cosmologia ha permesso di superare questa concezione, come abbiamo già motivato e vedremo più in dettaglio nel prossimo paragrafo.

## 6. L'abbandono degli equilibri

Dopo l'iniezione d'entropia susseguente all'inflazione, il brodo termico è in sostanziale equilibrio. Il verso del tempo è comunque segnalato dal calo delle temperature e dalla conseguente progressiva scomparsa di particelle dal brodo termico, quando  $T$  va al di sotto della loro massa  $m$ . Un secondo elemento può però inquinare la condizione di equilibrio del brodo termico: la fuoriuscita di qualche specie di particella dal brodo stesso, se il suo tempo di collisione con ogni altra particella, incluse quelle della sua stessa specie,  $\tau$ , diviene maggiore del tempo cosmico  $t$ . Vi sono naturalmente altre possibili interazioni oltre agli urti, ma nel cosmo primordiale il ruolo degli urti è essenziale.

Emblematica, in questo senso, è la fuoriuscita dal brodo termico dei neutrini, quando  $T \simeq 900$  keV. Ma, a una temperatura significativamente maggiore, dell'ordine del TeV, un'altra componente ha preso la sua strada autonoma, quella composta dalle particelle  $\chi$ , i neutralini, probabili componenti della DM odierna.

In realtà non c'è processo elementare che non sia destinato a uscire dall'equilibrio. Per ogni processo di collisione possiamo definire il tempo caratteristico

$$(XIV.5) \quad \tau = \frac{1}{\sigma n v}$$

dove  $\sigma$  è la sezione d'urto del processo,  $n$  è la densità in numero dei «bersagli»,  $v$  la velocità di collisione legata alla temperatura. Il ritmo di «diluizione» di  $n$  è almeno  $\propto a^{-3}$  ( $a$ : fattore di scala), mentre  $v$  è costante (urti relativistici) o  $\propto T^{1/2} \propto a^{-1/2}$

Nel caso più «favorevole» avremo dunque  $\tau \propto a^3/\sigma$ . Quanto alla sezione d'urto  $\sigma$ , la microfisica prescrive che un andamento  $\sigma \propto \ln T/T$  è il limite estremo oltre il quale

non si preserva l'unitarietà della *matrice*  $S$ . Di conseguenza,  $\tau \propto a^2 \ln a$  è il limite assoluto per la crescita dei tempi di collisione; per la maggior parte dei processi la crescita sarà più rapida, ma non si può scendere al di sotto di questo ritmo.

A fronte di ciò sta la dipendenza del tempo  $t$  da  $a$ : finché domina la radiazione, risulta  $t \propto a^2$ , successivamente  $t \propto a^{3/2}$ .

Il tempo di collisione  $\tau$ , in questo contesto, esprime l'inverso del ritmo con cui la distribuzione delle particelle si adatta alle mutate condizioni cosmiche. L'inverso del tempo  $t$  o, per maggior precisione, il parametro di Hubble  $H$ , esprime il ritmo con cui le condizioni cosmiche si modificano. In definitiva, l'andamento del rapporto  $t/\tau$ , tra due tempi crescenti, ci ragguaglia sull'andamento del rapporto  $N_{max}/N$ , tra il numero crescente delle configurazioni possibili e il numero crescente del numero dei microstati che corrisponde al macrostato dei sistemi cosmici, via via che essi cercano di adattarsi alle mutate condizioni segnate dalla crescita di  $N_{max}$ .

Ebbene, le valutazioni che precedono dimostrano che, anche nelle condizioni più «favorevoli»,  $t/\tau$  ha una decrescita logaritmica mentre, nella maggior parte dei processi, il calo sarà assai più robusto. Inevitabilmente, dunque, ogni processo elementare è destinato a uscire dall'equilibrio, a causa dell'espansione cosmica.

Va peraltro segnalato che l'uscita dall'equilibrio non implica la nascita di configurazioni diverse da quelle di massima entropia.

Ricordiamo due esempi già considerati in questo testo:

- 1) I fotoni del CMB. Essi si disaccoppiano dalle altre componenti a  $\sim 3000$  K e, al momento del disaccoppiamento hanno una distribuzione di Bose-Einstein

$$(XIV.6) \quad f_\gamma(\varepsilon, T) = \frac{2}{(2\pi\hbar)^3} \frac{1}{\exp(\varepsilon/k_B T) - 1} .$$

I fotoni sono particelle a massa zero e quindi la loro energia individuale è totalmente soggetta a *redshift*; al crescere del fattore di scala da  $a$  a  $a'$ , risulta dunque

$$(XIV.7) \quad \varepsilon/\varepsilon' = a'/a ,$$

con ovvio significato dei simboli, allo stesso modo in cui risulta

$$(XIV.8) \quad T/T' = a'/a .$$

Ne segue che l'espressione (XIV.6), al crescere di  $a$ , rimane invariata e i fotoni del CMB sono sempre nella configurazione di equilibrio termodinamico. È del resto ben noto che i fotoni del CMB hanno uno spettro esattamente di corpo nero.

- 2) Neutrini con massa  $m_\nu \ll T_{\nu,dg}$  ( $T_{\nu,dg} \simeq 900$  keV; ragionevolmente  $m_\nu \sim 0.1$  eV). La distribuzione dei neutrini, al disaccoppiamento, è analoga a quella del  $\gamma$ :

$$(XIV.9) \quad f_\nu(\varepsilon, T) = \frac{2}{(2\pi\hbar)^3} \frac{1}{\exp(\varepsilon/k_B T) + 1} .$$

In questo caso, il *redshift*, agendo sulla lunghezza d'onda e quindi sull'impulso, fa sì che

$$(XIV.10) \quad \frac{\varepsilon}{\varepsilon'} = \frac{a'}{a} \sqrt{\frac{(p/m_v c)^2 + 1}{(p/m_v c)^2 + (a'/a)^2}},$$

e la forma (XIV.9), non appena è  $pc \sim k_B T \sim m_v c^2$ , non è più preservata. In realtà, siccome al disaccoppiamento la distribuzione si può scrivere

$$(XIV.11) \quad f_v(\varepsilon, T) = \frac{2}{(2\pi\hbar)^3} \frac{1}{\exp(pc/k_B T) + 1},$$

senza commettere un errore apprezzabile, ed è esattamente  $p/p' = a'/a$ , l'espressione (XIV.11) è la distribuzione permanente dei neutrini: inizialmente coincidente con la distribuzione di equilibrio statistico, ma destinata a distanziarsene quando  $k_B T \sim m_v c^2$  e la distribuzione di equilibrio dovrebbe essere una distribuzione di Boltzmann.

## 7. Il ruolo della massa

I casi 1) e 2) del precedente paragrafo valgono come esempi; in particolare, è chiaro che l'informazione immessa nella distribuzione neutrinica dall'espansione, lì è congelata. Il caso dei neutrini è però emblematico per il ruolo della massa; è la sua presenza a permettere l'accumulo di informazione.

In effetti, su un piano generale, la massa è il *jolly* di tutta la partita:

- 1) Nei modelli inflazionari essa permette la formazione di orizzonti *quasi* omogeneo-isotropi, arricchiti da un opportuno spettro di fluttuazioni.
- 2) Rompendo l'invarianza conforme essa differenzia le metriche sull'asse temporale.
- 3) Nel brodo primordiale, in cui tutto altrimenti apparirebbe immutabile, in un eterno equilibrio termodinamico, essa permette di fissare il verso del tempo, rendendo rilevabile il verso in cui  $T$  discende tramite la graduale scomparsa delle specie pesanti.
- 4) La massa caratterizza altresì i residui di CDM (probabilmente neutralini) e barioni, che costituiscono componenti non-relativistiche. Questo senza dubbio è il suo ruolo più importante, verso la formazione dei grandi disequilibri della nostra era. Sarà tramite queste componenti che decolleranno i grossi accumuli di informazione su cui si basa l'articolazione dell'Universo che oggi osserviamo.

In effetti, finché le componenti non-relativistiche non prendono il sopravvento, la massa di Jeans è oltre l'orizzonte. Le fluttuazioni di densità, anche quando la loro scala è raggiunta dall'orizzonte, si estrinsecano in onde soniche per le componenti barionica e fotonica, si dissipano nella componente neutrinica, mentre restano congelate nella

componente di DM. Oltre all'equilibrio nei processi elementari, abbiamo così un equilibrio nei processi idrodinamici, nel quale le diverse scale potrebbero apparire scarsamente differenziate.

### ***Equilibrio viriale e equilibrio termodinamico***

In realtà, facendo crescere il tempo di collisione fotone-elettrone, l'espansione lo renderebbe superiore al tempo caratteristico di evoluzione delle fluttuazioni, su scale via via crescenti. Se null'altro accadesse, questo meccanismo condannerebbe alla dissipazione le fluttuazioni nei fotoni. Per le fluttuazioni nei barioni resterebbe invece una differenziazione dipendente dalla scala: sulle scale dove le onde soniche hanno avuto modo di svilupparsi, le fluttuazioni barioniche svanirebbero, in concomitanza alla dissipazione delle onde nei fotoni; sulle scale raggiunte dall'orizzonte quando il tempo di collisione supera ormai il tempo cosmico, le fluttuazioni nei barioni si congelerebbero, similmente a quanto accade per la DM.

Questa ipotetica sequenza di eventi non si realizza perché, al *redshift* di equivalenza, l'energia di massa prende il sopravvento. Le fluttuazioni nel CDM «scoprono» così di essere instabili e iniziano a crescere. Poco dopo l'idrogeno ricombina e anche la distribuzione della componente barionica si viene a trovare in condizioni di improvvisa instabilità.

A questo proposito è essenziale sottolineare la differenza tra equilibrio viriale e equilibrio termodinamico. L'equilibrio viriale, prescrivendo che sia  $2E_k + V = 0$ , fissa la velocità quadratica media di una qualsiasi componente massiva, inclusa in una fluttuazione di massa complessiva  $M$  e raggio  $R$ , al valore

$$(XIV.12) \quad v^2 = GM/R,$$

indipendentemente dalla massa.

In un fluido non relativistico in equilibrio termodinamico, in cui siano presenti particelle di masse diverse,  $m_1, m_2, m_3 \dots$ , invece, sono le energie cinetiche a essere eguali. L'energia di ciascuna specie è fissata dalla temperatura, sicché deve essere

$$(XIV.13) \quad m_1 v_1^2 = m_2 v_2^2 = m_3 v_3^2 = \dots,$$

dove  $v_i$  è la velocità quadratica media della particella di massa  $m_i$ .

Il materiale delle onde soniche pre-ricombinazione ha la temperatura del brodo termico residuale, formato di fotoni. Dopo il disaccoppiamento, la componente fotonica non è più un fluido e anche la componente barionica viene a trovarsi lontanissima dall'equilibrio termodinamico, in particolare per quanto riguarda la dinamica delle sue disomogeneità. Se concentriamo l'attenzione sui volumi in cui la fluttuazione nella DM è di segno positivo e nei quali i barioni si dispongono a accrescere sulla DM, andremo a seguire le zone di Universo deputate a vedere il formarsi di strutture. Se l'ampiezza della fluttuazione è qui sufficiente, essa crescerà fino al livello di massima espansione, per poi ricollassare su se stessa e raggiungere l'equilibrio viriale, in un tempo inferiore o pari all'età attuale del cosmo.

L'equilibrio viriale, per la componente di DM, è uno stato di quasi stazionarietà; la componente barionica, invece, evolve ancora: essa può raffreddarsi per emissione di  $\gamma$ , ciò che è impossibile alla DM, che non sente la forza elettromagnetica; inoltre, essa tende a raggiungere l'equilibrio termodinamico.

### ***Presente e futuro delle fluttuazioni***

Questa sequela di eventi, logicamente, riguarda i volumi inizialmente adeguatamente sovradensi. Altre zone cosmiche, sottodense, sono destinate a evolvere nei grandi vuoti; zone sovradense, ma con contrasto di densità non adeguato, sono suscettibili di evolvere verso la formazione di strutture nel futuro.

A questo proposito due punti sono da sottolineare.

Nel riferirsi a zone sovra- o sotto-dense non ci si riferisce necessariamente a volumi disgiunti. Al crescere della scala, una zona adeguatamente sovradensa, che vediamo evolvere – ad esempio – in una galassia, può essere inclusa in una zona adeguatamente sovradensa che vediamo evolvere in un ammasso di galassie. Allo stesso modo, il volume del protoammasso può essere incluso in un volume ancor maggiore, ancora sovradenso ma non sufficientemente, quindi non evoluto in alcuna struttura nell'epoca attuale. In questa struttura a scatole cinesi, si possono anche alternare sovradensità e zone sottodense; esse possono risultare, ad esempio, nella formazione di piccole galassie isolate all'interno di grandi vuoti.

Il secondo punto, invece, riguarda il futuro delle sovradensità che oggi non siano ancora uscite dalla fase di crescita lineare: un modello  $\Lambda$ CDM (standard-cold-dark-matter), in cui la densità critica è assicurata dalla DM, ha una sorta di invarianza per traslazioni temporali. In un tale modello, ogni epoca può essere contrassegnata dalla scala di massa che mediamente virializza a quel tempo. Se vivessimo in una cosmologia  $\Lambda$ CDM sapremmo che, in futuro remoto, strutture che oggi vediamo al più come *superclusters* sarebbero destinate a raggiungere conformazioni simili a quelle degli odierni ammassi di galassie. In tale cosmologia, gli ammassi di galassie sono le più grandi strutture legate, oggi.

In una cosmologia  $\Lambda$ CDM o comunque includente una componente preponderante di DE, ci si prospetta un futuro ben diverso. Le fluttuazioni oggi lineari rimarranno congelate per sempre, non ci saranno nel futuro strutture più ampie degli ammassi odierni; la DE prenderà sempre più il sopravvento, le zone in cui si sono coagulate le strutture diverranno sempre più lontane tra loro e tra di esse si estenderà il vuoto impalpabile del cosmo riempito essenzialmente di energia oscura.

### ***Condensazioni barioniche***

Tornando allo stato attuale del cosmo, notiamo come la distribuzione delle disomogeneità, su masse  $M > M_g$  (massa tipica delle galassie) sia essenzialmente creata dalla DM. I dati ottici e radio, invece, provengono dalla componente barionica. Sulla scala di ammassi e gruppi di galassie, i barioni sostanzialmente tracciano la DM e, anche se i meccanismi di raffreddamento sono all'opera, il loro risultato dinamico si potrà ri-

scontrare solo in un remoto futuro. La scala galattica segna il confine al di sotto del quale i meccanismi di raffreddamento hanno effetti dinamici: raffreddandosi, i barioni scivolano sempre più all'interno delle buche di potenziale gravitazionale e così le immagini ottiche delle galassie sono in genere assai più concentrate dei loro aloni di DM.

I dati sulle galassie a spirale, ad esempio, ci danno aloni il cui raggio supera la zona ottica di un fattore  $\sim 4-5$ . Dal momento che la DM non si concentra, il rapporto  $\rho_b/\rho_{DM}$  sale, passando da valori «cosmologici»  $\sim 0.2$  a valori  $\sim 10-100$  volte maggiori.

Nel disco delle galassie a spirale la componente di DM ha dunque una densità modesta, rispetto ai barioni, e non segue assolutamente i barioni nelle loro ulteriori condensazioni stellari, alla cui formazione presiedono ancora meccanismi di raffreddamento da emissione elettromagnetica.

Gli sconvolgimenti creati dal prevalere dell'energia di massa e dal disaccoppiamento barioni-radiazione creano una situazione difficile da ricondurre nell'alveo di un generale equilibrio termodinamico. Potremmo anche speculare sui meccanismi e sui tempi necessari perché esso si ricrei, quando le stelle terminano il combustibile nucleare, il loro *core* degenera, in molti casi prendono la strada verso il *bucò nero*, che a sua volta è destinato a evaporare radiazione, ecc. Al di là della complicazione e dell'articolazione di questi meccanismi, è chiaro che un riavvicinamento all'equilibrio sarà solo locale, nell'ambito creato dalle fluttuazioni che sono ricollassate e hanno creato sistemi non-lineari. Esso non riguarderà il cosmo nel suo assieme, che viaggia sempre più verso un completo disequilibrio.

### ***Un cosmo senza massa***

In assenza della componente di DM e senza la violazione della simmetria  $B-\bar{B}$  il cosmo si sarebbe ridotto, a un certo punto, a soli fotoni e neutrini. Se poi anche i neutrini fossero stati senza massa, esso avrebbe raggiunto uno stato di apparente equilibrio termodinamico, in cui le distribuzioni delle particelle si mantenevano eguali in forma a dispetto della variazione delle scale. In effetti, sia la distribuzione di Bose che quella di Fermi sono invarianti nel corso dell'espansione cosmica.

L'equilibrio termodinamico sarebbe *apparente* perché non garantito dalle interazioni, ma dall'immutabilità in forma delle distribuzioni iniziali. È pur vero che tale immutabilità implica una graduale riduzione dell'energia di ogni singolo fotone e la progressiva diluizione della densità d'energia nel cosmo. Ma, per apprezzare l'uno o l'altro di questi fatti, occorrerebbe una scala di riferimento. Sparite tutte le particelle con massa, non ci sarebbe un *regolo* cui riportare energie o temperature.

L'apparente situazione di equilibrio termodinamico sarebbe quindi destinata a permanere, apparentemente immutata e immutabile, sicché la stessa freccia del tempo, inosservabile, non avrebbe avuto alcun rilievo fisico.

Ancora una volta è chiaro che sono le violazioni dell'equilibrio termodinamico, la creazione dei disequilibri – prima – e la loro progressiva riduzione – poi – a rendere apprezzabile il verso in cui trascorre il tempo.

## CAPITOLO XV

# Cosmologia & Cosmologie

### 1. Cosmologia scientifica e cosmologie tradizionali

L'illustrazione delle sintesi e dei temi attualmente in discussione nella moderna cosmologia scientifica apre nuove prospettive anche all'analisi delle cosmologie tradizionali.

Non c'è dubbio che, in questo e in altri campi, la scienza moderna abbia cassato il ruolo di supplenza giocato da filosofia e religione. Che di supplenza si trattasse era noto, seppure non a tutti. Nel Vangelo gnostico, detto *di Tommaso*, Gesù di Nazareth irride a chi colloca fisicamente il soprannaturale in cielo: «Se coloro che vi guidano dicono: Ecco il Regno è nei cieli, allora gli uccelli vi arriveranno prima di voi.» Per poi concludere: «... il Regno è dentro a Voi e fuori di Voi.» Ma questi distinguo non erano per molti e secoli di agiografia e mito testimoniano a favore degli astronauti sovietici che ironizzavano comunicando a terra: «Voliamo nell'alto dei cieli, non vi abbiamo trovato gli angeli».

Infatti, se anche il dotto scartava il cielo come sede del soprannaturale, ciò non impediva al pensiero corrente di attribuire il tuono a Giove pluvio. La scienza, dunque, fa strage nel pollaio degli exoterismi da cortile. Ma anche riscatta le componenti tradizionali da sempre estranee al baccano.

Viene allora naturale chiedersi a che livello si collochino le cosmologie presenti in tutte le culture, se esse siano puro mito o se posseggano aspetti ancora rilevanti, oggi che la cosmologia scientifica ha ridefinito il loro ruolo. Essendo comunque e definitivamente chiaro che non è nelle scritture, di qualsiasi tradizione, che troveremo la soluzione a quesiti di natura scientifica.

A nostro avviso ci sono due spazi di relazione, tra cosmologie e cosmologia, a livello soggettivo e oggettivo. A dispetto di ogni forma di irrisione, l'idea che il «Fiat Lux» ricalchi il lampo del *Big-bang* è oltremodo diffusa. Forse più tra i detrattori delle scritture che tra i loro supporters.

Ciò che vale allora la pena di chiedersi, in primo luogo, è se ciò abbia avuto influenza nello sviluppo del pensiero scientifico e/o se altri elementi delle cosmologie pre-scientifiche siano stati travasati nella dottrina scientifica o possano esserlo in futuro. A simili possibilità abbiamo già accennato, sporadicamente, in vari punti del testo; ci

proponiamo qui di addivenire a una visuale d'assieme effettuando un'esplorazione su ciò che chiameremo nesso soggettivo, stabilendo da subito che, se questa influenza è esistita o può sussistere, essa non altera il meccanismo di acquisizione scientifica, che matura solo dopo che il quadro teorico e quello sperimentale hanno raggiunto la mutua saturazione.

Il nesso sarebbe invece presente nella motivazione individuale del ricercatore, quando egli decide su quale canale convogliare la sua capacità speculativa, o con strumenti matematici che tramite la costruzione di apparati osservativi e sperimentali. Tale nesso, a volte, raggiungerebbe anche un ambito interpersonale, garantendo un *temporaneo* consenso capace di indurre ampie comunanze d'intenti, nel mondo scientifico, anche prima che il ciclo della «scientificizzazione» (ci si perdoni questo neologismo) si sia concluso.

Ancora più complesso è il discorso sul nesso oggettivo. Esso si diparte dall'individuazione di un sostrato comune alle diverse tradizioni, anche se accompagnato da elementi mitici distinti. Tale sostrato sembra indicare che è errore attribuire, *tout-court*, il testo di una «scrittura» al suo estensore. Non solo perché è corretto presumere che le sue aggiunte alla precedente tradizione orale siano state modeste, sicché il suo ruolo equivale a quello del monaco che trascriveva testi classici. Ma soprattutto perché la presenza di una tematica comune richiede una spiegazione. Il credente la troverà nell'unicità della Rivelazione, che viene poi adattandosi ai diversi contesti culturali. Quello che certamente si rileva è l'unicità della natura umana, che si estende dunque ai livelli più profondi della  $\psi\upsilon\chi\eta$  e del pensiero, quando essi vengono avvicinati tramite la meditazione e le pratiche ascetiche.

A nostro avviso, non occorre aderire a una tradizione religiosa per focalizzare l'esistenza oggettiva di questo ambito e cogliere le relazioni che esso segna tra elementi del pensiero. Quando poi tali relazioni ricalcassero taluni stralci delle relazioni fisiche, si aprirebbe una questione imbarazzante: se la dinamica del mondo fisico possa rappresentare un riflesso o una componente di dinamiche più estese. Perché allora la posizione di taluni metafisici tradizionali (non del greco Aristotele, per esempio), irridente nei riguardi del «vile» mondo della natura, andrebbe capovolta. Se la conoscenza filosofica o *mistica* non ha caratteri di oggettività scientifica, è invece possibile che la scoperta dei rapporti tra entità naturali, dotata di tutti i crismi della scientificità, permetta di apprezzare gradualmente le relazioni tra «entità» di un mondo di cui la Natura è solo parte.

Alla denominazione di scienza ambiscono molte dottrine di costruzione talora recente, dalla sociologia alla psicologia del profondo. A molte di tali ambizioni siamo fortemente ostili, mentre siamo portati a riconoscere l'esistenza di conoscenze al di fuori della scienza. Nel trattare del nesso oggettivo avremo in mente questo genere di conoscenza *a-scientifica*, di cui tratteremo per la sua contiguità con la cosmologia scientifica. Tutto ciò arricchisce il contesto della ricerca cosmologica che è fondamentalmente scienza della Natura, della  $\phi\acute{o}\sigma\iota\varsigma$ , ma ha valenze che legittimamente si coniugano con un mondo oltre la fisica.

## 2. Il nesso soggettivo

Ci sono due archetipi che ricorrono, nel pensiero prescientifico sull'origine del mondo, quello della Creazione e quello dell'Eterno ritorno. Essi si ritrovano, in forma quasi esclusiva, alla base di due grandi tradizioni culturali, quella giudaica e quella induista. In altre tradizioni vi sono spesso misture dei due approcci.

### *Creazione e Eterno ritorno: Genesi biblica e Veda*

La parte più nota della Genesi biblica consiste di un racconto mirato a illustrare la centralità dell'Uomo, coronamento della Creazione. L'Uomo dunque giunge per ultimo, quando il Creato è pronto per lui. Orbene, che la specie umana sia recente e che la sintesi evoluzionistica abbia dovuto passare per esseri semplici prima di realizzare l'essere complesso, è un fatto scientificamente oggettivo. Nella Genesi la successione di esseri via via più complessi ha il fine di realizzare l'ambiente dell'Uomo; la sequenza evoluzionistica, invece, è scevra da ogni elemento finalistico. Ma le due prospettive portano a conclusioni parallele e, quindi, inevitabilmente, c'è uno straordinario parallelismo tra la storia degli eventi cosmici e biologici, ricostruita da planetologia, geologia, paleontologia, genetica, e la sequenza degli atti creativi descritti nella Genesi. (La modesta discrepanza sulla creazione degli uccelli non guasta il quadro d'assieme.)

In questo quadro l'elemento caratteristico è l'attribuzione di un senso «sovranaturale» al tempo, un'ascissa che porta dalla *terra informe e vuota*, all'Essere di caratura divina, *fatto a Sua immagine e simiglianza*. È vero che tra il tempo della Genesi e la contemporaneità c'è la cesura del settimo giorno e l'inizio del riposo di Dio, ma la quotidianità sarà inevitabilmente percepita come estensione di una sequenza sovranaturale.

Il tempo della cultura occidentale, prima o dopo Kant, Maxwell o Einstein, è il tempo giudaico. Nella cultura classica la freccia del tempo era assai più confusa. Invece, Passione, Morte e Resurrezione, inserite nel contesto provvidenziale dell'Impero, ribadiscono la verticalità del tempo e la sua natura metafisica. È quest'idea di tempo che viene incontro alla scienza, dalla termodinamica dell'Ottocento, alla cosmologia del Novecento.

Vale la pena di confrontare questo contesto con quello della tradizione induista. È forse superfluo ricordare che essa si basa su scritture del tutto estranee alla tradizione biblica, i Veda (*visione, sapienza*), un vasto corpus redatto in lingua sanscrita. I testi più antichi riguardano la ritualità sacerdotale. Più recenti sono gli *Upanisad*. Vi sono diverse interpretazioni dei Veda, i cosiddetti *darshana*, il più profondo dei quali è sicuramente il *Vedanta* (fine dei Veda). A sua volta, la dottrina del Vedanta è coordinata e riassunta in una serie di aforismi chiamati *Brahma-Sûtra*, che hanno dato luogo a ulteriori commentari, tra cui il *Ramânuya*, di carattere sostanzialmente panteistico, e il *Shankarachârya*, cui faremo soprattutto riferimento.

Il modo in cui, in questi testi, si coniugano mistica e filosofia, è oggettivamente estraneo alla cultura occidentale, nella quale la delimitazione religione/filosofia si situa al confine tra eredità giudaica e eredità classica e quindi, nonostante le faticose elaborazioni tomiste, separa due campi ben delineati e reciprocamente impermeabili. Nello iato tra filosofia e religione si è potuta istituire, come disciplina a sé stante, la teologia, che genera dialetticamente eresie e scismi, storicamente spesso basi ideologiche delle articolazioni del potere. Nella cultura indiana, invece, assente lo iato tra sovrannaturale e filosofia, non esiste la teologia e sono quindi inconcepibili eresie o scismi, sicché i conflitti tra potenti hanno divoto cercare le loro motivazioni altrove. In questo contesto, l'ascesi mistica perde senso senza il riferimento filosofico, mentre le basi della filosofia sono le «scritture» che, anzi, sono esse stesse in gran parte filosofia.

Il tempo nei Veda: il cosmo in sé è eterno, sebbene si trovi in una condizione di costante mutamento, in quanto i numerosi mondi che lo compongono (*brahmanda*, uova di Brahma) passano ciclicamente attraverso nascita, crescita, corruzione, morte. Il divenire è dunque pura apparenza, non c'è né fine né un fine per la ruota cosmica. Va da sé allora che il tempo perde l'intensa virulenza metafisica del tempo occidentale.

Come il mondo è caduco, così lo sono, individualmente, gli stessi dei. Tra di essi quello che vive più a lungo è *Brahma* – da non confondere con l'entità assoluta denominata *Brahman* – e la sua vita durerà 100 anni di Brahma. Seguendo i dettami dei Veda si calcola trattarsi di  $3.11 \times 10^5$  miliardi di anni, circa 20 000 volte il tempo trascorso dal *Big-bang*. Tutti gli esseri viventi, «dal dio Brahma dai quattro volti, fino allo stelo d'erba», possiedono un'anima immortale, *ab aeterno*, migrante da un corpo materiale a un'altro, lungo una scala stabilita dai meriti conseguiti nella precedente vita. Le stesse anime degli dèi, che vivono meno del supremo Brahma, entrano nel ciclo della metempsicosi, fatta eccezione per *Isvara*, Dio, che si situa in qualche modo all'esterno del sistema e che tutte le scuole riconoscono. Quando l'attuale Brahma cesserà, cadrà con esso tutto il sistema cosmico. Dopo un tempo altrettanto lungo di assenza, il mondo rinascerà con un nuovo Brahma.

Le tradizioni biblica e vedica si caratterizzano dunque per l'accentuazione di uno degli opposti archetipi, Creazione o Eterno ritorno. È una contrapposizione netta che tuttavia non esclude, nell'ambito di entrambe, quelle che si possono definire «esperienze» dell'archetipo opposto.

Quella che, probabilmente, è la più importante corrente mistica della cultura occidentale, non è nata nell'ambito cristiano, ma nelle comunità giudaiche della diaspora; si tratta della Cabbalàh. Non è certo questo il luogo per affrontare storia e sviluppo del pensiero cabbalistico, che esalta il ruolo del numero e, in special modo, il calcolo combinatorio, ma è soprattutto legato all'esame totale della *Toràh*. Essa viene recepita nella sua forma originale, scritta in ebraico arcaico, vocaboli senza vocali e interspazi, al di là delle interpretazioni correnti, alla ricerca delle Verità più nascoste.

Questa lettura del testo biblico ha svelato ai cabbalisti significati altrimenti sfuggenti. In talune concezioni del processo dell'emergere di Dio dalle profondità di

Se stesso, nella creazione, un ruolo vitale viene assunto dalla leggenda midrashica dei mondi creati e scomparsi prima o a lato del mondo presente. Una variazione di questa idea si trova nello *Zohar* (Spendore), principale collezione dei testi cabbalistici. In esso si descrive come Dio entrò nella forma dell'*Adàm Kadmòn* (Uomo Primordiale) o nelle varie configurazioni di questa forma. L'esatta natura delle altre creazioni *non-riuscite*, che lo *Zohar* denomina *Malkhei Edom* (Re di Edom), non viene approfondita. Si tiene invece a sottolineare che solo quando la figura dell'Uomo Primordiale venne forgiata correttamente, con esatto bilanciamento (*matkela*) delle componenti maschile e femminile, la creazione fu in grado di sostenersi e il Creatore se ne compiacque («ed Egli vide che ciò era buono...»). Rispetto alla molteplicità dei mondi vedici qui è molto più sfumata la sequenza temporale.

### ***Monoteismo giudaico e politeismo nelle religioni ariane***

Tutto l'Antico Testamento è incentrato sull'affermazione dell'Unicità, mirabilmente enfatizzata dalla preghiera mosaica: «Shemà Israel, Adonai elohinu, Adonai echad» (Ascolta Israele, l'Eterno è il Signore, l'Eterno è Uno). La Trinità cristiana, conseguenza dell'attribuzione a Cristo di una Natura Divina, ineluttabilmente indebolisce questo richiamo unitario. Ma la stessa tradizione Cabbalistica, cui abbiamo già accennato, ha ampiamente elaborato sugli *aspetti di Dio* (dottrina delle *Sefirot*). Strettamente unitaria, invece, la tradizione coranica.

La tradizione vedica, al contrario, contempla un florilegio di divinità maggiori o minori, maschili o femminili, e il culto indù molto spesso si riduce alla devozione per *Siva*, *Kali*, o *Agni*, o per divinità ancor minori. Ma non dobbiamo scordare il culto dei Santi di matrice cattolica, la cui realtà non si discosta molto dalle devozioni indù.

Probabilmente di derivazione vedica sono gli dei portati dagli arii invasori, sulle sponde del Mediterraneo. Il sistema greco o romano, richiama in molti aspetti quello induista, quasi ne fosse una derivazione superficiale. In effetti, Roma imperiale avrebbe avuto poche difficoltà a accomodare nel Pantheon di Adriano i *deva* indù: *Agni* sarebbe stato Efesto, *Skanda* sarebbe stato Ares, i due *Asvin* sarebbero stati i Dioscuri, le *Apsaras* sarebbero state le ninfe, e così via.

Se la religione olimpica è derivata in qualche modo dalla tradizione vedica, ciò è certamente avvenuto tramite l'oblio dei Veda stessi e di tutto il sostrato esoterico, che porterà dagli Upanisad al Vedanta, ma si trova già implicito nei libri più antichi dei Veda, in cui è chiaro il concetto di Unità dell'Essere. Denominato *Brahman*, esso racchiude tutto il Reale; Essere e anche non-Essere, perché la negazione dell'Essere ha senso solo in presenza di Esso.

Di qui nasce tutta l'articolazione della realtà, descritta lasciando uno spazio assai limitato a elementi mitici. Su questo aspetto essenziale dell'induismo, che rappresenta forse l'esposizione più analitica delle tematiche comuni alle grandi tradizioni, avremo modo di tornare ancora e diffusamente nel prossimo paragrafo.

### **La cultura dell'Occidente e il decollo della cosmologia scientifica**

Sul tempo, potremmo essere tentati di segnare un punto a favore della cultura occidentale: il «suo» tempo ha vinto. Ma non è questo il punto. Ci dobbiamo invece chiedere se – e fino a che punto – gli scopritori della cosmologia scientifica trovarono la strada segnata dall'idea di tempo cui erano educati, tenendo anche conto che il tempo «metafisico», della tradizione giudaica, si coniuga direttamente col monoteismo. Ciò che vogliamo sottolineare è che creazione e monoteismo costruiscono comunque un riferimento culturale, indipendentemente dall'adesione dello scienziato a una «religione»: anche l'agnostico o l'ateo occidentali, per esempio, tendono a considerare il monoteismo come un valore rispetto al politeismo «pagano». È allora legittimo chiedersi se questo *background* abbia avuto un peso nello sviluppo della cosmologia del Novecento.

Molti fatti nella recente storia scientifica portano a una risposta positiva. L'idea giudaico-cristiana del tempo e le sue conseguenze, tradotte in ipotesi scientifiche, hanno subito raccolto un consenso corale, spesso inizialmente ingiustificato dai dati scientificamente acquisiti.

Questa affermazione rischia di essere fuorviante, se non viene ribadito che a questa fase è sempre seguita una completa *scientificizzazione*; il punto è che essa è spesso successiva non solo alle proposizioni ma alla loro stessa accettazione.

Questa dinamica intellettuale non riguarda elementi minori o collaterali. Un primo esempio giunge assieme all'avvio della cosmologia moderna, quando viene formulata l'ipotesi dell'espansione dell'Universo. Sin dal primo capitolo abbiamo visto che le galassie scoperte da Hubble sono le più vicine alla Via Lattea e si situano tutte nel *Gruppo Locale*, sicché le loro velocità rispetto alla Via Lattea sono dovute a *moti peculiari*, disordinati. Il Gruppo Locale s'incentra su due grandi galassie, Via Lattea e Andromeda, addirittura in mutuo avvicinamento. Dai dati di Hubble, non è possibile tracciare una linea sul piano distanza-velocità che mostri veramente una proporzionalità tra le due grandezze. Tanto meno è possibile ricavarne il rapporto

$$H = \frac{\text{velocità}}{\text{distanza}} .$$

Di fatto, però, l'idea della proporzionalità portava con sé la possibilità di far girare all'inverso le lancette dell'orologio fino a un punto zero, con le galassie in mutuo contatto, e prima ancora, con tutta la materia concentrata in un solo punto. Ciò faceva parte dell'*inespresso* hubbiano, ma non ci fu astrofisico o fisico cui questa estrapolazione non giungesse immediata.

Ci sono due aspetti distinti nelle scoperte di Edwin Powell Hubble. Nel Dicembre 1924 Hubble comunica i risultati dell'analisi degli spettri della nebulosa in Andromeda (Messier 31) e di altre nebulose, mettendo in evidenza la loro natura extragalattica e fissandone le distanze sulla scala dei Mpc.

Nel 1929, invece, Hubble presenta i suoi risultati sulla proporzionalità tra distanza e velocità di allontanamento. La portata delle scoperte del 1924 giustifica in parte la comunità scientifica, quasi stordita dalla loro portata e pronta a accettare ulteriori

stravolgimenti. Le obiezioni si concentrarono, semmai, sull'inespresso. L'analisi dei dati fu accettata anche da chi, con spirito luciferino, pretese di collocarli in un contesto capace di escludere quel tempo zero percepito come simulacro della Creazione.

Si pensi a quanto diversa sarebbe stata la storia della cosmologia se una immediata rianalisi dei dati di Hubble avesse mostrato che l'espansione del cosmo, a quel punto, era una fantastica illusione. L'impatto dell'espansione sul grande pubblico fu molto efficace a convogliare risorse sulla ricerca astrofisica. Tali risorse andarono anche alla costruzione di telescopi più grandi del 2.40 m di Mount Wilson e fu tramite gli specchi da 4 metri che si riuscì a raggiungere finalmente il *range* di distanze dove l'espansione è un fatto osservativo. A lato delle scoperte di Hubble abbiamo dunque visto il ruolo dei relativisti cosmologi, come Robertson, Walker, Lemaître e Friedmann. Il lettore di questo libro è al corrente del fatto che il «problema cosmologico», risolto da Friedmann e della cui soluzione Hubble forse tenne conto, attiniva una realtà assolutamente virtuale, per nulla aderente alcuna risultanza osservativa dell'epoca. Ciò non toglie che lo studio della dinamica di un continuo, isotropo e omogeneo, era definito «problema cosmologico» sin dai tempi di Isaac Newton: nonostante non ci fosse prova alcuna dell'esistenza di una scala sopra la quale le disomogeneità osservate e sperimentali potessero essere assimilate alla molecolarità di un fluido. In sostanza, scoperte le galassie, Hubble intuì di essere finalmente vicino alla scala di omogeneità. Ci fu una forzatura, per omogeneità e relative conseguenze osservative ci voleva almeno un altro fattore 10, ma le ricadute della forzatura furono quanto mai positive.

Va finalmente aggiunto che il *bluff* di Hubble non venne e non viene affatto percepito come scandaloso. Per esempio, a Hubble è stato intitolato il recente telescopio spaziale. Similmente, il moto coerente delle galassie nell'espansione cosmica, cui si sovrappongono poi le velocità peculiari – le sole che Hubble poté misurare –, è denominato *flusso di Hubble*. Come se il mondo scientifico, di fronte alla «scoperta» dell'espansione del cosmo, si fosse liberato di un peso, vedendo finalmente ri-convergere i risultati della ricerca sulle aspettative culturali. Di questo passo, in qualche modo *profetico*, la comunità scientifica è tuttora grata a Erwin Hubble.

### ***L'evoluzione del cosmo e la crisi dello Steady-State***

Del resto questa vicenda è tutt'altro che isolata. All'inizio degli anni Sessanta la dialettica tra sostenitori del *Big-bang* e dello Stato-Stazionario (Steady-State) era ancora caldissima, anche se stavano gradualmente attenuandosi le sue basi ideologiche. Cinque anni più tardi la ricerca cosmologica compie la grande virata e lo stato stazionario viene relegato nelle illusioni della storia della scienza.

Il mutamento di prospettiva viene correntemente attribuito a due fatti osservativi:

- 1) la scoperta della radiazione di fondo a 2.7 K (CMB);
- 2) la scoperta delle radiosorgenti, i cui numeri ( $N$ ) e luminosità apparenti ( $S$ ) risultavano distribuiti in modo che la curva  $\log N$ - $\log S$  aveva una pendenza diversa da 1.5. Tale pendenza era attesa in modelli «non-evolutivi».

A proposito del CMB va ricordato che la forma dello spettro di questa radiazione divenne chiara solo dai dati dell'esperimento *FIRAS*, all'inizio degli anni Novanta. Per molto tempo svariate osservazioni sembrarono rilevare pesanti distorsioni. Sarebbe stata una grossa sfida per i teorici conciliare tali distorsioni con l'origine «cosmologica» del CMB.

Ancor più peculiare è la storia delle radiosorgenti. Una delle più importanti centrali per il loro rilevamento si situava nella campagna dell'*East Anglia*, vicino a Cambridge, e i dati convergevano al *Cavendish Laboratory*, situato allora al centro della cittadina, poco lontano dai suoi famosi *colleges*. Martin Ryle e collaboratori affermarono ben presto che i dati escludevano una pendenza 1.5 della curva  $\log N - \log S$ . Ma Martin Ryle rifiutava di far conoscere a Fred Hoyle i dati dettagliati. Li avesse avuti in mano, Sir Fred sarebbe certamente riuscito in pochi giorni a dimostrare che essi erano ancora troppo scarni per falsificare la pendenza critica della  $\log N - \log S$ . In effetti, ogni punto della curva così definita non è statisticamente indipendente dai successivi e una buona analisi, basata su dati differenziali e non integrali, avrebbe posto in evidenza che l'errore sulla pendenza era ancora sufficientemente ampio da includere il valore 1.5.

Ciononostante la comunità scientifica riconobbe subito le conclusioni del Cavendish Laboratory e Fred Hoyle abbandonò la direzione della sua creatura, il neonato *Institute of Theoretical Astronomy*.

A distanza di diversi decenni, abbiamo oggi una tale ricchezza di dati da poterci permettere una precisa disamina in termini differenziali. Essa conferma la precognizione di Martin Ryle e collaboratori, dando ragione al consenso precocemente ottenuto sulla conclusione che essi certificassero l'evolutivezza del cosmo.

### ***Il cosmo preinflazionario e il multiverso. Una rivincita politeistica?***

Verso la fine degli anni Settanta il quadro cosmologico era giunto a una prima sintesi sostanziale. Lo *scenario dell'instabilità gravitazionale* permetteva una generale sistemazione, chiarendo altresì quali elementi osservativi erano legati alle condizioni iniziali e quali altri discendessero dalla casualità locale. Restavano da sistemare vari «dettagli», ma il successo dello scenario d'assieme appariva totale. Tutto era dunque disceso da un quadro unitario, da uno stato primevo in cui il cosmo era «quasi» omogeneo e isotropo, ma persino omogeneo e isotropo nelle piccole violazioni di omogeneità e isotropia. Quasi il riflesso fisico del monoteismo più integrale.

Eppure, ciò che appariva un successo, nascondeva i sintomi della crisi. Dopo avere mostrato che tutta la varietà delle morfologie cosmiche si era evoluta dall'uniformità primeva, i cosmologi si trovarono a dover spiegare *come mai* il cosmo primevo fosse stato così uniforme. In termini fisico-matematici il quesito è il seguente: se partiamo da condizioni iniziali *generiche*, da una metrica spazio-temporale

$$ds^2 = g_{\mu\nu}(x_\rho) dx^\mu dx^\nu,$$

in cui le componenti del tensore  $g_{\mu\nu}(x_\rho)$  sono funzioni sostanzialmente arbitrarie, possiamo affermare che le eventuali asperità del tensore metrico si smusseranno co-

munque, e un sito generico del *metacosmo* evolverà verso un cosmo omogeneo e isotropo? Vale poi la pena di aggiungere che anche il numero di dimensioni spaziali e temporali andrebbe lasciato libero, a priori. Ma non è ancora abbastanza: quali altre fantastiche realtà matematiche contengono come sotto-alternativa gli spazi metrici?

Per un breve periodo parve che la risposta potesse essere positiva. Venne formulata la *congettura* che andava sotto il nome di *mix-master*, in cui si pretendeva che isotropizzazione e omogeneizzazione fossero dovute alle forze gravitazionali legate alle disomogeneità metriche. Metaforicamente, dopo aver mostrato che l'Olimpo contemporaneo era mera apparenza e che tutto discendeva dall'Unico, si tentava di provare che l'Unico era conseguenza necessaria del primordiale Caos.

Così formulato, il progetto del *mix-master* fallì rapidamente. Questo progetto è stato oggi rimpiazzato dai modelli inflazionari, in cui le interazioni non-gravitazionali e, in particolare, uno o più *campi scalari* autointeragenti, entrano nella dinamica primava. Non sorprende affatto che, quando l'orizzonte era piccolissimo, interazioni non gravitazionali cooperassero con la gravitazione a stabilire la dinamica cosmica. Il programma iniziale dell'inflazione era tuttavia quello di individuare queste interazioni nell'estrapolazione di dati sperimentali di alta energia. Questa parte del programma, attualmente, non si è realizzata.

I modelli inflazionari, dunque, introducono un'interazione *a hoc*, accanto all'ipotesi che esista a priori, nell'ambito di un *metacosmo* sconfinato, una piccola oasi omogenea e isotropa. L'interazione riesce allora a gonfiare (*inflare*, *inflazionare*) quest'area fino a portarla a comprendere tutto l'orizzonte oggi osservabile.

Di fatto, per definizione, l'unica parte del cosmo osservabile è quella compresa nel nostro orizzonte di particella. L'economia di pensiero garantita dal meccanismo inflazionario sta nell'affermare che orizzonti «come il nostro» sono «frequenti», ogni *piccola area* isotropa-omogenea si gonfia e produce condizioni cosmiche simili a quelle che noi osserviamo. Se il meccanismo non funziona, orizzonti come il nostro sono «infrequenti». Ne esistono comunque – questo è chiaro – vista la totale genericità delle condizioni iniziali.

Non può sfuggire che sin qui, nel formulare l'ipotesi inflazionaria, si siano applicate categorie abituali, normalmente in uso quando sussiste una molteplicità di eventi sperimentabili, a un evento – la nascita dell'Universo – che, per definizione, è l'unico osservabile.

La storia tuttavia non si conclude a questo punto. Volendo un meccanismo che generi un ambito isotopo e omogeneo, si ottengono dei risultati che vanno al di là della richiesta iniziale. Il successo principale dell'inflazione non sta nel dimostrare che le qualifiche del mondo osservabile costituirebbero un dato generico, ma nel collogare le proprietà osservative di omogeneità e isotropia su grande scala a una serie di altri parametri, in primo luogo quelli che caratterizzano lo spettro delle fluttuazioni primordiali. Le previsioni dell'inflazione, in questo caso, sono pienamente consistenti coi dati sinora disponibili che, anzi, cominciano ormai a essere sufficientemente restrittivi da permettere di selezionare tra diversi tipi di dinamica inflazionaria.

Un ruolo chiave è qui giocato da anisotropie e polarizzazione del CMB, che permettono di studiare il campo di fluttuazioni prima dell'inserimento di fattori di non-linearità. Ciò è di estrema importanza: la dinamica non-lineare non è solo difficile da seguire sul piano analitico; il punto è che essa implica l'intervento di una fisica caotica, che svincola gran parte degli osservabili dalle condizioni lineari di partenza.

La censura stabilita dalle biforcazioni caotiche non impedisce tuttavia verifiche statistiche e così molti aspetti delle osservazioni sul CMB possono essere implementati da osservazioni più vicine a noi. Ciò permette una verifica incrociata ed è anche una prova diretta dello scenario dell'instabilità gravitazionale. A titolo d'esempio ricordiamo che il CMB stesso, con le componenti a basso  $l$  dei suoi spettri angolari  $C_l^{(T)}$ ,  $C_l^{(TE)}$ ,  $C_l^{(E)}$  e  $C_l^{(B)}$ , fornisce informazioni sulla fase di re-ionizzazione e quindi sulla distribuzione delle sorgenti a altissimo *redshift* ( $z \sim 5-20$ ). Esse si incrociano con l'osservazione delle *Ly $\alpha$  clouds*, che offrono dati a *redshift* intermedi ( $z \sim 4-10$ ). Campioni profondi di galassie e SNIa, usate come candele standard, permettono infine di esplorare il cosmo più prossimo.

Tra le previsioni di alcuni modelli inflazionari c'è anche l'esistenza di un fondo primordiale di onde gravitazionali. Abbiamo però visto che la loro osservazione diretta è molto difficile, per non dire impossibile: il «rumore» di altre sorgenti copre il segnale. Queste onde, tuttavia, potrebbero essere rilevate in modo indiretto, tramite l'esame dello spettro  $C_l^{(B)}$  della polarizzazione della radiazione di fondo microonde (CMB). Abbiamo visto che si tratta ancora di un segnale estremamente tenue, con cui interferisce poi il campo gravitazionale degli ammassi, che potrebbe alimentare lo spettro  $C_l^{(B)}$  a spese di altre componenti della polarizzazione. Ma, in questo caso, la scienza e la tecnologia necessarie sono dietro l'angolo. La rilevazione di onde gravitazionali primordiali, selezionando tra i modelli inflazionari, permetterebbe di determinare con maggior precisione altri parametri cosmologici.

Ma c'è molto di più: la selezione tra modelli inflazionari potrebbe allora dare, come sottoprodotto, informazioni sul cosmo al di là del limite osservativo stabilito dall'orizzonte. Informazioni su una realtà che esula dall'Universo osservabile, che riguardano qualcosa di enormemente più esteso, per cui sono state usate le denominazioni di *multi-verso* o *pluri-verso*. Si tratterebbe, in un certo senso, di informazioni «statistiche», che ci direbbero, per esempio, a quale *distanza media* si situano due Universi simili al nostro. Dove ci sarebbe molto da discutere sul significato metrico di «distanza media».

Qui, comunque, la fantasia di molti cosmologi si è sbizzarita nel formalizzare contesti teorici, più o meno falsificabili sulla base di queste potenziali informazioni «statistiche». Essi descrivono successioni di mondi figliati da mondi parenti, o unità cosmiche parallele. Gli straordinari pregi estetici di taluni di questi modelli non debbono farci scordare che oggi siamo in un campo assolutamente ipotetico. Comunque, anche in futuro, quando qualifiche e parametri del modello inflazionario ci fossero completamente disponibili, ci sarebbero non pochi problemi, concettuali o «filosofici», a affermare la realtà fisica di ciò che è definitivamente inosservabile.

A fronte di queste prospettive fisiche è inevitabile il richiamo a idee e formulazioni delle cosmologie tradizionali, alle «uova di Brahma», o ai «Re di Edom». Ma si può andare oltre e annoverare, tra le motivazioni di non pochi sviluppi teorici, il fascino innegabile di formulazioni tradizionali. Una tale valutazione non mira affatto a sminuire la portata delle ipotesi fisiche. Specialmente alla luce del successo riportato dai padri nobili della cosmologia, alle cui ipotesi non sembrano affatto estranee analoghe influenze.

Metaforicamente, il successo di uno di questi approcci starebbe a significare la rivincita di un fondamentale politeismo primevo, tradottosi poi in monoteismo osservabile, ma completamente distinto dalle articolazioni politeistiche che il quadro Unitario ha generato nel mondo osservato.

Ciò che va comunque rimarcato è la totale insussistenza *attuale* di elementi di verifica per la quasi totalità di queste formulazioni ipotetiche. Sarebbe sbagliato affermare che si tratta «banalmente» di riformulazioni, in chiave fisico matematica, di ipotesi pre-scientifiche: a queste ipotesi si arriva da una strada costellata di verifiche sperimentali, esse fanno parte del meccanismo di sviluppo della conoscenza fisica. Ma proprio per il fatto che il meccanismo di acquisizione di nuova fisica è divenuto tanto complesso e oggettivamente lento, sussiste un pesante rischio che il meccanismo delle comunicazioni di massa distorca il significato operativo di questi approcci teorici, finendo col dare la sensazione di una scienza autoreferenziale; con tutti i rischi che ciò implicherebbe, per esempio nel rapporto con le tentazioni rinascenti del *magico*.

### 3. Il nesso oggettivo

Ancora più delicato e complesso è il discorso sul nesso oggettivo. Ciò che cercheremo di mostrare, innanzitutto, è la presenza, nelle cosmologie tradizionali, di tematiche sovrapponibili, una volta che esse siano sfrondate del bagaglio mitico. Va subito rilevato che queste tematiche pretendono di estendersi ben oltre il mondo fisico, sebbene a esse debba spesso essere attribuito anche un significato strettamente «materiale». Ne segue che, quando si individuasse nel mondo fisico un «format» sovrapponibile a queste tematiche, si rischia di dover trarne la disturbante conclusione che la  $\phi\upsilon\sigma\iota\varsigma$  è un aspetto della  $\psi\nu\chi\eta$ .

Queste conclusioni vanno comunque subito attenuate da due punti di vista concomitanti. Il fatto basilare è che non c'è nessuna ipotesi metafisica che permetta previsioni fisiche. Semmai la sovrapponibilità può essere osservata *a posteriori*, in un modo di cui daremo alcuni esempi. Questi esempi metteranno poi in luce che qui ci muoviamo in un campo *a*-scientifico, non ci sono teoremi matematici e neppure concomitanze conclusive di teoria e esperimento. La condivisione delle conclusioni dipende quindi dal convincimento individuale.

Esistono tuttavia altre acquisizioni *a*-scientifiche, proprio riguardanti il profondo della psiche umana, che vengono correntemente accettate, sia pure in diverse prospettive, e utilizzate quanto meno come supporto terapeutico. In questo campo, ci

potremmo trovare a contatto con dati almeno altrettanto significanti, ma che taluni potrebbero essere portati a sottovalutare o negare, sulla base delle proprie convinzioni *a-religiose* o *anti-religiose*. Il nostro punto è che la condivisione di talune conclusioni non va pensata, *tout-court*, come condivisione del contesto tradizionale da cui esse giungono. Non più di quanto l'accettazione del complesso di Edipo non porti a accettare la religione olimpica.

Tra le cosmologie tradizionali, quella che offre una descrizione più articolata delle origini è probabilmente quella indù. Rispetto a altre tradizioni essa ha anche il «pregio» di limitare l'elemento mitico alla scelta di «nomi» per le entità primeve. Proporremo quindi un discorso tentativo sul nesso oggettivo, vale a dire sugli elementi comuni alle diverse tradizioni, partendo dal darshana denominato Vedanta e attendoci al Shankaracharya.

### **Induismo e origine del cosmo**

Visiteremo dunque anzitutto i temi di partenza delle cosmologie induista. Ci troveremo qui su un terreno estraneo alla ricerca fisica e cosmologica, ma è bene compiere questa ispezione nell'ambito dei linguaggi tradizionali e senza la presunzione di poter tradurre tutto e subito in tematiche scientifiche.

Nella cultura indù, l'«essere» di ogni cosa, materiale, vivente o divina, è *atma* (stesso etimo di *anima*). Nel *Vedanta-sara* (Essenza del Vedanta) di Sadànanda, leggiamo che *l'Io ha rifugio nel Sé, l'Indivisibile, Essenza-Coscienza-Bene Assoluto, irraggiungibile dalle parole e dal pensiero, sostrato del tutto, per il raggiungimento del massimo desiderio*. Quando l'individualità dell'Io si dissolve nell'Essere, nel *Sé*, esprimibile solo apofaticamente, *atma* si confonde con *paramatma*, in essenza identico al *Brahman* (che, come già ricordato, non va confuso col dio *Brahma*). Altrove, il rapporto tra *atma* e *Brahman* viene raffrontato a una goccia d'acqua marina, che possiamo vedere isolata o ricongiunta all'abisso oceanico. Il *Brahman* va però oltre l'essenza, includendo persino la propria negazione: l'alternativa del non-Essere può essere concepita solo grazie all'esistenza. Quando riconosciamo di esistere, constatiamo il *Brahman*.

L'Universo nella sua interezza è attraverso *Brahman*, che costituisce sostrato di tutta la creazione, del mondo manifestato e fenomenico, sebbene vada molto oltre essa. Esso si manifesta come *Brahman Nirguna*, Bene assoluto, indescrivibile tramite attributi, o come *Brahman Saguna*. A partire da quest'ultimo, il Vedanta descrive un'evoluzione cosmica che forgia ogni entità. Essa, però, non si costituisce di eventi, essendo estranea all'ascissa temporale, intesa sia in senso topologico che metrico, sicché sussiste *ab aeterno*.

In questo processo, il *Brahman* si polarizza in *Purusha* e *Prakrti*, cui i Veda associano caratteristiche maschile e femminile. In sé, *Prakrti* è una «miscela indifferenziata» dei tre *Guna* (stesso etimo del latino «genus»), denominati *Sattwa*, *Rajas*, *Tamas*. *Sattwa* rappresenta la tendenza alla spiritualità, la *serenità* e la *luce*; *Rajas* è

la passione e la tendenza all'operosità, l'*azione*; Tamas è la tendenza alla rigenerazione, anche attraverso la morte, l'*inerzia* e la *pesantezza* insita nella materia. Le loro personificazioni saranno poi i tre dei della Trimurti indù, Brahma, Visnù, Siva. La transizione dai guna alle divinità personali avviene tramite *Buddhi*, che costituisce l'Intelletto, il fattore di percezione e scelta nel costituirsi della psiche, il *λόγος*, nato dall'*azione* di Purusha su Prakrti. Egli è la *sapienza*, la *conoscenza*, l'*Uomo Primordiale* che dà la forma osservata alle entità del mondo.

Grazie alla complessa *azione* di Purusha, dunque, l'indifferenziazione di Prakrti si articola in una realtà poliforme e differenziata, che si estrinseca negli «esseri» del mondo, dagli dei, alle entità astratte, fino alla più umile pietruzza.

L'induismo, ovviamente, non è solo una cosmologia e qui ne abbiamo appena sfiorato alcuni elementi, sia pure assai salienti. Alla luce del raffronto con altre culture, è anche opportuno sottolineare la ricchezza delle fonti su questa che è la più antica tradizione tuttora vitale. Si pensi a cosa ci rimarrebbe di tutta la sua sofisticazione, se l'induismo fosse stato sopraffatto da altre culture, che ne avessero cancellato i testi, salvo tramandare qualche riassunto elaborato magari in chiave buddista o islamica.

Dobbiamo tenerlo presente quando incontreremo tradizioni per cui le fonti sono povere e mediate da alte culture.

### ***Origine del cosmo nelle tradizioni occidentali***

Altrove abbiamo già rivisto l'incipit della Genesi, primo libro della Bibbia giudaico-cristiana: *All'inizio Dio creò il cielo e la terra, la terra era informe e vuota e le tenebre regnavano sull'abisso, ma sulle acque aleggiava lo Spirito. E Dio disse: Sia la luce... .*

Ci sono tematiche comuni tra le diverse culture che sono ben note. Un esempio «banale» è il *diluvio universale* (Genesi 6-9). Non sorprende che lo si ritrovi nelle leggende babilonesi, certamente tra le fonti della Genesi. Esse narrano del re Utnapistim, il quale apprende dal dio Ea che gli dei – e, in particolare, il perfido Enlil – hanno deciso di annientare la vita sulla terra tramite un diluvio. Utnapistim costruisce un'arca per la sua famiglia e per un certo numero di animali. Il diluvio babilonese dura sette giorni, poi la colomba viene mandata a esplorare se vi siano terre emerse e il tutto si conclude con la transizione di Utnapistim allo stato di semidio.

Sorprende un po' di più ritrovare il diluvio anche nella tradizione olimpica e in quella germanica, che lo illustrano con protagonisti diversi e, forse, con un'enfasi minore.

Ancora più sorprendente è trovarlo nelle leggende di popoli lontanissimi dal bacino mediterraneo, come gli Aztechi, secondo cui l'attuale periodo o *sole* è stato preceduto da altri quattro periodi: il primo, *sole d'acqua*, si conclude col diluvio. Gli Aztechi non lesinavano in disastri; anche gli altri tre periodi si sono conclusi con catastrofi; un «sole» viene addirittura divorato dai giaguari e si sa già che l'attuale «sole» finirà con un immane terremoto.

L'elenco delle culture in cui si ritrova il diluvio sarebbe ancora lungo e sarebbe

semmai più breve citare le tradizioni che non ne fanno cenno. Una catastrofe naturale potrebbe essere all'origine di queste leggende. Si è parlato della possibilità che il bacino del Mar Nero fosse originariamente separato dal Mediterraneo e le acque del Mar Nero, allora molto più piccolo, fossero al fondo di una depressione, come quella del Mar Morto. Sulle rive del proto-Mar-Nero si sarebbero trovate popolazioni di contadini e allevatori. Il diluvio è causato dalla tracimazione del Mediterraneo attraverso gli stretti, dove ora c'è il Bosforo.

Non necessariamente alternativa è l'idea che il diluvio faccia riferimento a eventi perinatali, impressi nel profondo di ciascun individuo. Questa seconda spiegazione lega questa tematica alle altre che vorremmo ora sottolineare. Esse colpiscono meno la fantasia, ma forse sono più significative del diluvio stesso.

Riandando dunque all'incipit della Genesi, notiamo due fasi nella creazione. La seconda è quella più nota e avviene tramite la *parola*, il *λόγος*, («e Dio disse...»); si tratta della sequenza di eventi finalizzati all'arrivo dell'Uomo, coronamento della creazione. La fase precedente è descritta più brevemente ed è creazione diretta di Dio: Egli crea *Cielo e Terra*, entità maschile e femminile; la Terra (sostanza primordiale, *Prakṛti*) è informe e vuota e le tenebre regnano sull'abisso. Poi lo Spirito (*Purusha*) soffia sulle acque (altra raffigurazione della sostanza primordiale) e ciò dà inizio all'azione creatrice del *λόγος*.

Ci proponiamo ora di portare l'attenzione sui vari avvicinamenti, tra entità legate a diverse tradizioni, proposte da vari studiosi. L'ambito in cui queste connessioni sono state introdotte è fortemente contrassegnato dall'aderenza a filoni filosofico-religiosi. La problematica sentita dagli autori riguarda la molteplicità dei credi e ciò spiega lo sforzo per trovare un substrato che accomuni le diverse tradizioni. Vorremmo qui astrarci da queste motivazioni, compiendo un'ispezione del materiale tradizionale.

Tra gli avvicinamenti più immediati, su cui non ci pare valga molto la pena di insistere, sta la relazione tra *JHVH-Elohim*, e il Brahman. Citeremo un passo non molto noto, perché proveniente da un vangelo gnostico, il cosiddetto *Vangelo della verità*, il cui inizio ricorda le parole del Vedanta-sara: *L'Universo si è dato alla ricerca di Lui, dal Quale è uscito. L'Universo si trovava dentro a Lui, l'inafferrabile, l'impensabile, al di sopra di ogni concetto.*

Il mondo occidentale è familiare con un'altra tradizione, quella olimpica. La fonte più sistematica di notizie è la Teogonia di Esiodo, un poema di circa 1000 esametri, presumibilmente redatto attorno al 700 a.C. La narrazione inizia dal regno del Caos e termina quando Zeus diviene re degli dei.

Sin dall'inizio Caos ci appare un Brahman dell'uomo semplice, che non s'è affatto preoccupato di «capire la lezione». Del fatto che Brahman tutto comprenda, in modo «inizialmente» informe, è rimasta solo l'idea del massimo di confusione, di... caos. Da Caos si originano Erebo e Etere, ma anche Uranos e Thalatta (Cielo e Mare), che potrebbero essere in relazione con Purusha e Prakṛti.

Di fronte alla tradizione olimpica si ha la sensazione di trovarsi in un contesto vedico degradato. Gli studiosi, per esempio, hanno individuato un legame tra Buddhi e

Ercole, ma i connotati di uomo forzuto hanno sopraffatto ogni tematica «metafisica». La «teologia» olimpica tutto riconduce a relazioni familiari, coniugali e, magari, extraconiugali, fatte salve forse due tematiche che meritano una sottolineatura: quella dei giganti, i *Titani*, visti come pericolo per l'ordine cosmico, e il ruolo di Chronos.

Dei giganti parla anche la Genesi (6-4: *In quel tempo sulla Terra vivevano i giganti...*), anche se si tratta di un passo misteriosamente slegato dal contesto. I giganti si ritrovano in molte altre cosmologie, come quelle cinese e germanica. Si tratta di un tema in qualche modo analogo a quello del diluvio; molti critici oscillano tra il riferimento a eventi preistorici (chi era più alto, l'Uomo di Neandertal o quello di Cromagnon?), un'interpretazione legata alla psicologia del profondo (gli adulti visti dal bambino), ma anche ipotesi più fantasiose.

L'ultimo dei Titani della Teogonia è Chronos, il tempo, che divora i suoi figli. Finché la madre, con uno stratagemma, salva Zeus, che sostituirà Chronos segnando il passaggio dalla violenza a ordine e giustizia. Ma anche qui, i drammatici rapporti familiari fanno aggio sulla transizione dalla dinamica primeva, *a-temporale*, a *Isvara-Zeus*, dio nel tempo e nel mondo manifestato.

### ***L'origine del cosmo in altre culture. La cosmologia germanica***

Questa rapida disamina di culture disparate mira, come dovrebbe essere ormai chiaro, a segnalare la presenza di uno «schema», cui si aggiungono risvolti di varia importanza e soprattutto tematiche mitiche.

Il pensiero pre-scientifico pone in rilievo, innanzitutto, la meraviglia dell'«*esistere*». Prima di interrogarsi sulla dinamica che porta il mondo fisico a assumere le forme osservate, viene l'identificazione di Dio nell'esistenza stessa. Si può dire che il programma successivo consiste nel provare che, se c'è esistenza, essa deve necessariamente articolarsi nei termini che ci sono noti. Nell'implementare questo programma, sembra lecito affermare che in tutte le tradizioni compaia quella che, nell'induismo, è la polarizzazione *Purusha-Prakrti*, da cui trae origine il *λόγος*, *Buddhi*, l'*Adam Kadmon*, che è l'Uomo primordiale, in quanto capacità di percezione e scelta, ma anche l'Adamo che fissa col nome l'esistenza stessa di ogni elemento del mondo fenomenico e naturale. L'Uomo primordiale rovescia, in qualche modo, i termini della creazione e, con la sua capacità di percezione, rende definito l'esistente.

L'azione *a-temporale* di Purusha su Prakrti ne risolve l'indefinitezza, portando a entità con un bilancio definito di guna. La trinità dei guna, e della Trimurti che li personifica, non sembra trovare riscontro in altre cosmologie salvo forse in quella germanica, cui accenneremo tra breve; dopo aver visitato, ancor più brevemente, la cosmologia cinese.

Al suo vertice sta, ancora una volta, un'unità indifferenziata, l'altissimo *Tao*, il *Tai-dji*, anticamente *Wang*. Il Tao è *legge*, ma anche realtà personale. Dal Tao traggono origine le forze primigenie denominate *Yang* e *Yin*. Yang è maschile, luminoso, rappresentato in bianco nel *Tai-ki*; Yin è femminile, oscuro, rappresentato in nero, nello

stesso simbolo. Yang è sferico, solare, estivo, ed è rappresentato anche dal *Cielo*, Yin è quadrato, invernale, gelido, ed è rappresentato anche dalla *Terra*. Yang e Yin sono contrapposti, ma nulla esiste senza la loro mutua azione, e da essi nasce il gigante *P'ancku* ed è attraverso a esso che i prodotti di Yang e Yin assumono un contorno di realtà.

Tentare di riassumere un pensiero articolato come quello cinese è comunque impresa disperata. Qui non andremo oltre la citazione di questi elementi, il cui rilievo andrebbe attentamente valutato, in termini storici e nel confronto tra scuole di pensiero. Abbiamo persino ommesso sin qui la citazione dei due grandi filosofi, Lao-Tse (604-517 a.C.) e Confucio (551-479 a.C.), le cui dottrine, miranti a armonizzare il mondo celeste, il mondo terreno, la struttura sociale e l'esistenza umana, facendo rientrare ogni elemento della vita individuale in uno schema cosmico, stanno alla base delle articolazioni della cultura tradizionale cinese.

Alla fine di questa ricostruzione merita ancora attenzione la cosmologia antica dei popoli germanici; non solo abitanti la Germania attuale, ma anche i paesi nordici e parte delle isole britanniche. Della loro cosmologia parla *Steve Weinberg*, nel suo libro «I primi tre minuti», da cui molti hanno appreso i rudimenti della cosmologia scientifica. Per Weinberg si tratta comunque di un tema incidentale, in cui risalta la dotta citazione della *Vacca Audumla*; lo scopo è quello di contrapporre l'inconsistenza concettuale del mito alla solida credibilità della scienza.

Della cosmologia germanica, in realtà, sappiamo veramente poco. La maggior fonte di informazione è la cosiddetta *Edda prosastica* – curiosa contraddizione *in terminis*, poiché «Edda» significa «poetica» – dovuta all'irlandese *Snorri Sturluson* (1176-1241), che la destinò agli *Scaldi*, i poeti cantori di quel profondo Medio Evo. Accanto a essa abbiamo anche l'*Edda poetica*, un'antologia di canti, raccolti ancora in Irlanda, in un periodo lievemente antecedente; tra essi, alcuni sono di grande importanza storica. Ci sono poi le informazioni che derivano dai contatti tra Roma imperiale e le tribù germaniche, nelle opere di Cesare, Strabone, Tacito, Ammiano Marcellino e Iordanes. Fonti dirette ma molto problematiche sono i testi runici.

Lo Sturluson era cristiano e, sebbene nutrisse un grande rispetto per le leggende antiche, non ci possiamo attendere dall'Edda una qualsiasi profondità «metafisica». Egli vede la cultura degli avi come un insieme di credenze *pagane*, da cui il popolo si è liberato grazie alla predicazione cristiana. Per Sturluson, il cristianesimo è come la scienza per Weinberg.

Diametralmente opposto l'atteggiamento dei Romani. Essi miravano a capire la cultura dei barbari, dando per scontato che il loro pensiero fosse definitivamente «inferiore», pensando di nobilitarne gli dei, adattandoli a quelli del loro pantheon antropomorfo. Così *Odino*, dio dei guerrieri, venne identificato con Marte, il robusto *Thor* con Ercole, e così via. Lungo questa strada, della cultura dei popoli germanici, non si capirebbe nulla.

La cultura germanica antica viene riscoperta nell'Ottocento, a partire dall'opera monumentale di Jacob Grimm (*Deutsche Mythologie*, 1834), per passare alla popola-

rizzazione dei miti guerrieri in chiave wagneriana e finire con l'ideologia nazista. I temi che meritano qui di essere sottolineati sono però assai distanti dai miti de *L'illusione di Re Gylfís* o delle *Walkirie*, ma anch'essi risuonano di quella stessa atmosfera fiabesca, che rende dolorosa la percezione di quanto sia andato perduto con quella grande cultura.

Secondo Snorri il cosmo ha inizio con la creazione di *Niflheim* e *Muspelheim*, due siti tra cui si apre l'abisso *Ginnungagap*. In esso si riversano, provenendo da Niflheim, correnti fredde che si consolidano in ghiaccio e nebbie velenose che si depositano sui ghiacci come rugiada. Da Muspelheim cade allora sui ghiacci la scintilla della vita, generando il gigante ermafrodito *Ymir*. Tramite una complessa genealogia, da Ymir nascono i *giganti del gelo*. Ma sui ghiacci troviamo anche la vacca *Audumla*, attratta dal sale di cui sono ricoperti, sicché li lecca per tre giorni, liberando da essi *Buri*, l'Uomo primigenio dalle sembianze divine. Buri ha un figlio, *Bor*, che si unisce alla gigantessa *Bestla* generando tre figli, *Wotan*, *Vile* e *Va*. Essi sono i primogeniti degli *Asi*, gli dei germanici, che uccidono i giganti del gelo tramutando i loro corpi e il loro sangue nella terra, nel mare, nelle alte montagne.

Questa sequenza mitica contiene la successione degli eventi vedici, arricchendoli di temi paralleli, ma privandoli, in apparenza, della «razionalità» filosofica. Niflheim è la componente passiva, femminile, sui cui ghiacci si riversa la scintilla vitale di Muspelheim. Ne nascono due sequenze parallele, una con capostite Ymir e fatta dei *giganti del gelo*, l'altra per il tramite della Vacca Audumla, con capostipite Buri. L'unione tra le due genealogie porta alla trimurti germanica, costituita da Wotan, Vile e Va. Ciò che si aggiunge, alla teogonia vedica, è il ramo dei giganti del gelo. Non un'aggiunta da poco, visto il ruolo dei giganti in molte cosmologie tradizionali.

Del resto, rispetto ai Veda, Niflheim è il sito di Prakrti, Muspelheim è il sito di Purusha. L'azione che parte da Muspel su Nifl, attraverso Buri, porta agli dei della Trimurti. Qui il parallelo tra Buri, Wotan, Vile e Va, da un lato, Buddhi, Brahma, Visnù e Siva, dall'altro, si arricchisce persino di assonanze linguistiche. Gli unici etimi a essere sicuramente disgiunti sono quelli di Brahma e Wotan-Odino. Le somiglianze sono talmente smaccate da rendere legittimo il quesito se esse non nascano dalle migrazioni di popoli dello stesso ceppo, piuttosto che da ispirazione del profondo.

A nostro avviso le due componenti possono essersi sovrapposte, laddove la permanenza dei temi ricchi di valenze «metafisiche», a fronte dell'ampliamento mitico, si sostiene tramite la pulsione di ciò che nasce nel profondo di ogni spirito umano.

#### 4. Dal nesso oggettivo alla cosmologia scientifica

Ci resta ancora da chiudere il cerchio, tentando di ricucire i temi che ricorrono nelle cosmologie tradizionali con quelli della cosmologia scientifica. È chiaro che non miriamo a una ricucitura superficiale, come se in una cultura o in qualche cultura si parlasse della nucleosintesi dell'elio o del CMB.

Ci sono però delle concordanze e delle prospettive che meritano di essere enu-

cleate, dopo aver ulteriormente ribadito due fatti in qualche misura contrapposti:

- 1) Non può esistere un valore di «preveggenza scientifica» in alcun testo tradizionale.
- 2) Se la realtà scientifica trovasse tuttavia collocazione nell'ambito di dinamiche «metafisiche», di ispirazione prescientifica, si aprirebbe un preoccupante squarcio sull'autonomia del mondo fisico. Le sue dinamiche finirebbero per apparirci come un riflesso di rapporti tra entità allargate; tali rapporti, quantificabili nel mondo materiale, sarebbero segnale di realtà profonde, ben oltre il mondo materiale stesso.

Al momento di cercare una collocazione «metafisica» alle dinamiche fisiche, tuttavia, siamo già pesantemente in campo *a*-scientifico, con conseguente perdita di oggettività. Anche il rapporto di apparenza più convincente potrà quindi essere sempre negato, poiché la convergenza tra dati e loro interpretazione persuade solo in un contesto *scientificizzabile*, a fronte di una assoluta ripetibilità dell'esperienza o dell'osservazione.

Un esempio che vogliamo riportare, riguarda i tre *guna*. Ogni entità del mondo manifestato, secondo il Vedanta, ne contiene una miscela definita. Ciò si può confrontare col fatto che ogni grandezza fisica ha *dimensioni* fisiche ottenibili componendo tre grandezze fondamentali, normalmente scelte essere lunghezza (*L*), tempo (*T*) e massa (*M*). Questo fatto era noto sin dai tempi dei grandi meccanici secenteschi. Tutti i sistemi di unità di misura ne tengono conto. Nel sistema c.g.s., per esempio, tutte le grandezze fisiche sono esprimibili in centimetri, grammi e secondi. Nel sistema MKS, le unità sono metro, kilogrammo e secondo. Si possono aggiungere ulteriori unità, arbitrariamente; per esempio si può aggiungere il grado Kelvin o il Coulomb; ma non di può scendere al di sotto di tre unità.

La ragione di questo fatto, oscura fino alla fine dell'Ottocento, si è chiarita grazie alle due grandi rivoluzioni della fisica che hanno segnato l'inizio del Novecento: la scoperta della Relatività e della Meccanica Quantistica. Ciò ha posto in chiaro che in fisica esistono tre (e solo tre) costanti fisiche *dimensionali*. La prima a essere scoperta è stata la costante di gravità, *G*; le altre due sono *c*, la velocità della luce, e *ħ*, l'azione elementare. Queste due si legano, rispettivamente, a relatività e meccanica quantistica.

Le «grandezze fisiche fondamentali», di cui si compongono tutte le altre grandezze fisiche, non sono quindi lunghezze, tempi e masse; ma piuttosto *gravità*, *azione elementare* e *velocità della luce*. La ragione della scelta storica sta nella piccolezza o enormità di queste costanti, espresse nelle unità abituali. Nel sistema c.g.s.,  $G \simeq 10^{-7}$ ,  $c \simeq 3 \times 10^{10}$ ,  $\hbar \simeq 10^{-27}$ . Questi numeri non sono casualmente grandi o piccoli, le loro enormità o piccolezza segnano la distanza del dominio in cui non vale la «fisica classica» dal mondo dell'esperienza ordinaria. La scienza dovette procedere a lungo, giovandosi a sua volta dei prodotti tecnologici che essa stessa suscitava, prima di attingere a queste conoscenze.

La conclusione che ogni grandezza fisica sia esprimibile come opportuna composizione di *velocità della luce*, *azione elementare* e *gravità*, va dunque raffrontata con l'assunto vedico che ogni realtà manifestata si compone di *sattwa* (luce), *rajas* (azione) e *tamas* (gravità).

Ancora una volta, qui non c'è nulla di scientifico, tanto meno di ripetibile. Chi dissentisse con qualsiasi argomento da questo parallelismo sostiene un'opinione senza dubbio valida.

Avendo così chiarito la portata e i limiti di ciò che ci si può aspettare da questo genere di rapporti, possiamo disporci a cercare, nelle conoscenze cosmologiche acquisite, un senso per le entità del sostrato comune che risulta dallo studio delle basi cosmologiche delle diverse culture. Qui ci muoviamo ancor più su un terreno scivoloso, col rischio di dare per acquisiti elementi che si riveleranno validi, ma parziali, all'indagine dei prossimi decenni o secoli. Esercizi di questo tipo tentati nel passato, se riportati oggi, danno spesso una penosa sensazione di ridicolo.

Ci limiteremo dunque a affermare che le culture tradizionali tendono a non esaurire la nascita del «mondo manifestato» nell'atto creativo proveniente dall'*Essere*. Un ruolo determinante è recitato dall'*Adàm Kadmôn*, dal *Buddhi* o *Buri*, dal *λόγος*. È il rapporto tra l'Uomo-primordiale – l'essenza dell'umano – e la creazione a dare equilibrio alla realtà, a sospingerla a assumere i contorni fisici noti.

Ciò che ci attendiamo, dalla cosmologia e dalla scienza dei prossimi decenni e/o secoli, è una precisazione di questo ruolo, che oggi è ancora confuso nel mito midrashico dei *Re di Edom* o, forse, nel contestato ruolo dell'*osservatore*, nell'evoluzione da natura quantica a realtà fisica delle protostrutture figliate in una fase cosmologica, che oggi indichiamo problematicamente col nome di inflazione.