Università di Pisa

Facoltà di Scienze Matematiche Fisiche e Naturali

Laurea specialistica in Scienze Fisiche

Anno Accademico 2007/2008

Tesi specialistica

LA FISICA DEI PROCESSI DI BRANCHING

Candidato

Andrea De Luca

Relatore

Chiarissimo Prof.

Paolo Rossi

Ringraziamenti

È impossibile per me pensare questo lavoro di tesi slegandolo dal periodo in cui esso è stato prodotto; quasi per dipingere una metafora fin troppo banale, la sua conclusione coincide con quella di un periodo della mia vita. E così sono tenuto a ringraziare, oltre ad essere ben felice di poterlo fare, tutti coloro che hanno costellato questa mia esperienza a Pisa, cominciata con la timidezza di un primo anno di faticosi risvegli ogni mattina per raggiungere le lezioni già cominciate, e che si conclude quietata al passo tranquillo e rilassato di una passeggiata su corso Italia in una delle tante serate a chiacchierare per strada o in giro per il collegio. Se in questi anni sono riuscito nell'impresa di crescere (mentalmente, per carità) almeno un po', di sicuro è merito diretto della presenza, della perseveranza di chi giorno dopo giorno ha saputo svegliarmi dal torpore che verso la "flemma tarantina" conduce, anche solo per portarmi a mangiare prima che fosse troppo tardi; di chi ha saputo celarmi in un nascondiglio glottodidattico quando ancora non sapevo affrontare certe invasioni; di chi verso tali responsabilità ha saputo muovermi; di chi ha saputo accompagnarmi nelle lunghe nottate tra un'idea brillante, una risata, un pacco di biscotti ed una tazza di tè; di chi ha saputo punzecchiarmi, anche solo trascinandomi verso una partita a ping pong, per farmi affrontare ogni mia debolezza ed incertezza, anche prima che mi fossi accorto d'avercela; di chi, squillante, ha saputo istillarmi giorno dopo giorno saggezza e armonia; di chi non si è mai tirato indietro di fronte a qualunque sciocchezza gli proponessi; di chi mi ha assistito nella mia stramberia, fino a voler vedere l'alba con me.

Ancora, non posso mancare di dire grazie alla mia famiglia, per non avermi mai fatto mancare mai affetto e ristoro e per avermi insegnato ad

apprezzare le emozioni del mondo e l'equilibrio della mia testa. Grazie a tutti i miei amici di Taranto per avermi mantenuto più vicino di quanto avrei mai immaginato.

Per quanto riguarda i ringraziamenti di carattere professionale, mi rivolgo ovviamente al mio relatore, Paolo Rossi, per avermi saputo convincere fino a trovare il coraggio verso l'idea di questa tesi e per aver supportato continuamente il mio sforzo. Ringrazio anche il professor Pietro Menotti, che dopo avermi guidato nella mia tesi triennale, è stato pronto ad indirizzarmi verso il futuro.

Indice

Introduzione

In questa tesi abbiamo analizzato i processi di branching, in particolare il Galton-Watson process. Quest'ultimo fu formulato nell'analisi riguardante l'estinzione dei cognomi delle famiglie aristocratiche inglesi durante l'età vittoriana, da Francis Galton nel 1873. Insieme al reverendo Henry William Watson essi pervennero ad una soluzione nel 1874 pubblicata nell'articolo intitolato *On the probability of extinction of families* [1]. Tuttavia una medesima impostazione era stata ottenuta da Bienayme [2]. Come vedremo approfonditamente nel corso della tesi, un risultato pressoché immediato è che l'estinzione ha probabilità uno se il numero medio di figli non supera uno. Tale modello fu ignorato dopo la loro pubblicazione, fino a che venne ripreso in considerazione a partire dagli anni '30, e nel 1947 Kolmogorov e Dmitriev [12] introdussero il termine *branching process* in un'accezione più specifica. Un processo di branching può essere pensato come la modellizzazione matematica dell'evoluzione di una popolazione i cui membri si riproducono e muoiono in maniera casuale. Gli oggetti coinvolti nel processo possono essere di vario genere e la loro riproduzione può dipendere da vari fattori quali l'età, l'energia, la posizione. Caratteristica sostanziale è tuttavia il fatto che i differenti oggetti non interagiscano tra loro. Tale assunzione è verificata pienamente in alcune situazioni fisiche come nel caso dei neutroni o dei raggi cosmici, ma diviene vera solo approssimativamente e sotto le dovute ipotesi nel caso delle popolazioni biologiche. Moderne applicazioni concernono la probabilità di sopravvivenza di un gene mutante, l'avvio di una reazione nucleare a catena, o la dinamica della diffusione di malattie. Non ci occuperemo di dare una trattazione completa né troppo approfondita intorno gli sviluppi matematici concernenti tali temi che trovano nei testi di Theodore E. Harris [21] e di Krhisna Athreya [22] una soddisfacente sistematizzazione. Rimanderemo infatti a questi ogni qualvolta sarà utile un approfondimento. Obiettivo di queste pagine sarà invece tentare di sviluppare una visione concettualmente alternativa di tali problematiche, in modo da inserirle in un contesto in cui le metodologie fisiche siano direttamente applicabili. Va detto che tale lavoro non nasce dal niente, ma s'inserisce in un filone già preesistente che ha visto diversi fisici cimentarsi in quesiti analoghi. Per cominciare, il tentativo di dare una formulazione quantistica dei fenomeni di branching è stato affrontato con successo, almeno nel caso di tempo continuo [6, 7, 5] evidenziando analogie notevoli con la teoria quantistica dei campi. Parallelamente il problema della distribuzione dei cognomi, su cui torneremo nel seguito, che appare almeno approssimativamente affrontabile come processo di branching, è stato analizzato fornendo alcune parziali soluzioni [10, 11, 4, 14, 15, 16, 17, 18], ma senza riuscire ad inquadrarlo efficacemente come fenomeno di tale tipo. S'osservi che la trasmissione dei cognomi oltre ad essere interessante da un punto di vista puramente storico, può assurgere a modello semplificato anche per analisi più inerenti la genetica, essendo il cognome un gene a tutti gli effetti, detto neutro, poiché non risente delle influenze dell'ambiente. Uno spunto per un ulteriore sforzo in tale direzione viene, poi, proprio dai dati raccolti, che mostrano la possibilità di una prospettiva fino ad ora inesplorata in tale settore. Infatti, lo studio statistico in varie aree geografiche riguardo la distribuzione dei cognomi ha fatto emergere un andamento a potenza. Questo ci porta a sperare d'inserire i processi di branching in tutta una classe di fenomeni in cui non è più possibile utilizzare quello straordinario strumento di semplificazione metodologica e concettuale che risiede nell'idea riduzionista per cui ogni processo può essere scomposto in processi più elementari, che a loro volta possono essere studiati individualmente e separatamente dal contesto, ma che nonostante ciò trovano negli strumenti della meccanica statistica un'efficace fonte d'indagine. Tale classe di fenomeni comprende un miscuglio incredibilmente variegato ed a prima vista del tutto disomogeneo in cui possono inserirsi la struttura su larga scala dell'universo, l'organizzazione del sistema solare, le turbolenze nell'atmosfera, gli oceani, la metereologia, i terremoti, i meccanismi di ripiegamento delle proteine, ma anche problematiche più inerenti le scienze sociali quali la genetica delle popolazioni, l'economia dei mercati finanziari, la dinamica del traffico urbano, la sociologia della produzione scientifica, la correlazione nelle sequenze del genoma o quella presente nei testi letterari classici. Una caratteristica che accomuna e ci permette di scardinare fenomeni di tale complessità e differenziazione è l'auto-similarità. Per comprendere meglio tale caratteristica va notato che concettualmente ciò che mette insieme questo gruppo così vasto è la mancanza di una scala intrinseca. Se prendiamo, per esempio, in considerazione alcune tipi di reti sociali quali quelle realizzate dai meccanismi di diffusione delle epidemie o dalle strutture di distribuzione di alcuni beni primari, ci s'accorge che si tratta di reti che possono essere arbitrariamente piccole o grandi, senza che nessun parametro possa definirne o limitarne a priori la scala. Se si osservano prescindendo appunto dalla scala, focalizzandosi soltanto sulle loro proprietà topologiche, si scopre che queste non solo si ripetono uguali a se stesse su scale differenti di una stessa rete, ma sono sostanzialmente le stesse quando si confrontano reti di natura anche radicalmente diversa. Questo approccio al problema mostra l'eccezionale efficacia dell'invarianza di scala, un concetto già ben noto ai fisici teorici che ne hanno fatto largo uso a partire dai lavori di Kadanoff e Ken Wilson, abilissimi nello sfruttarlo per raggiungere una profonda comprensione dei fenomeni critici nella meccanica statistica.

In numerosi casi però si affrontano tali problemi senza la consapevolezza e l'esperienza che i fisici hanno acquisito e che ha insegnato a discernere gli aspetti essenziali e ad ottenere da essi risultati concreti. È quindi questo un ottimo motivo per affrontare in tale ottica innovativa i processi di branching, proprio perchè lasciano intuire una caratteristica di auto-similarità intuitivamente evidente. Abbiamo quindi seguito due direzioni differenti: per cominciare si è tentato d'inserire l'evoluzione probabilistica descritta da un processo di branching in un contesto hilbertiano tipico della teoria dei campi, partendo dall'analogia tra il processo di riproduzione governata da leggi probabilistiche e lo scattering quantistico di particelle elementari. Quindi nel primo capitolo abbiamo richiamato le caratteristiche cruciali della formalizzazione matematica del processo di Galton-Watson, mentre nel secondo ci siamo occupati di costruire un'adeguata teoria quantistica che riproducesse le medesime probabilità. Dopo la risoluzione di alcune problematiche, tale approccio si è dimostrato interessante rendendo semplici ulteriori generalizzazioni. Il terzo capitolo si muove ancora nella stessa direzione, nel caso di processi di branching continui. Abbiamo pertanto ripreso il lavoro già proficuo di Peliti [6]. Nella seconda parte della tesi, invece, ci siamo occupati di sviluppare un approccio più vicino alla fisica statistica vera e propria. Convertendo la variabile temporale discreta dell'evoluzione probabilistica in una coordinata spaziale, un espediente piuttosto familiare ai fisici, nel quarto capitolo abbiamo costruito un reticolo gerarchico di spin, in cui mediante un'opportuna scelta dell'hamiltoniana d'interazione tra i siti vicini, si sono potute riprodurre le probabilità di transizione sotto forma di funzioni di correlazione degli spin. L'efficacia di tale costruzione è emersa nella possibilità di riprodurre, seppur in una forma semplificata, un risultato già noto risalente a Richard Otter [8], seguendo una strada ed un contesto del tutto diversi. Nel quinto capitolo ci siamo dunque occupati di dare una sistematizzazione più completa di tali metodi. Abbiamo infatti riassunto l'importante ruolo del gruppo di rinormalizzazione nel caso dei reticoli gerarchici, per i quali esso permette di pervenire a soluzioni esatte. Si è visto quindi come tale strumento si mostri naturalmente efficace in un'applicazione concreta: il caso di una popolazione in cui si formino continuamente nuovi tipi d'individui, per immigrazione o per mutazione. I risultati prodotti, sotto forma di esponenti critici, si sono dimostrati interessanti sia per il loro riscontro con i dati empirici nel caso della distribuzione dei cognomi, sia perché riproducono e generalizzano i risultati ottenuti in [10] frutto di un approccio del tutto differente basato sulla *master equation*.

Capitolo 1

Galton-Watson branching process

1.1 Definizione

Vogliamo studiare degli oggetti che possono riprodursi generando oggetti dello stesso tipo. Può trattarsi in generale di esseri umani, batteri od anche neutroni nelle reazioni nucleari. Partiamo da un insieme iniziale di oggetti che corrisponde alla generazione 0. I figli di questi oggetti individuano la prima generazione e così via. Denoteremo con $X_0, X_1, \ldots$ il numero d'individui alla prima, seconda generazione, etc... Facciamo le seguenti assunzioni:

1. Se supponiamo noto il numero d'oggetti alla generazione n, la legge di probabilità per le generazioni seguenti dipenderà solo dal numero d'individui alla n-esima generazione e non da quelle precedenti. Ciò rende le $X_0, X_1, \ldots$ una catena di Markov. Inoltre assumeremo, a meno di specificazioni diverse, che le probabilità di transizione non varino nel tempo.

2. Assumeremo che i diversi oggetti abbiano evoluzioni del tutto indipendenti e non interferiscano uno con l'altro, in particolare il numero di figli di un oggetto non dipende da quanti oggetti sono presenti.

Possiamo quindi procedere con la formalizzazione matematica. Definiremo le $X_0, X_1, \ldots$ come variabili aleatorie, a valori interi di una catena di Markov. Assumeremo sempre che $X_0 = 1$ a meno di particolari specificazioni, situazione che si adatta facilmente a quella più generale per via delle proprietà prima enunciate. Indicheremo con P la misura di probabilità per tali variabili. La distribuzione di X_1 sarà uno dei dati del modello:

$$P(X_1 = k) = p_k \qquad k = 0, 1, 2, \dots \qquad \sum_k p_k = 1 \tag{1.1}$$

Per dedurre la distribuzione di X_{n+1} data quella di X_n è sufficiente tenere presente la proprietà 2, definiamo quindi:

$$P_{ij} = P(X_{n+1} = j | X_n = i) \qquad i, j, n \in \mathbb{N}$$

Chiaramente si avrà: $P_{0n} = \delta_{0n}$. Ci occuperemo nelle prossime sezioni di calcolare il valore delle P_{ij}, il loro andamento asintotico in vari casi ed alcune proprietà che da ciò potremo dedurre.

1.2 L'uso delle funzioni generatrici

Un modo molto proficuo per affrontare questo tipo di problemi è quello d'introdurre le funzioni generatrici:

Definizione 1.1 (Funzione generatrice). *Data una distribuzione di probabilità p_k, a valori interi, corrispondente alla variabile aleatoria X definiamo la funzione generatrice associata:*

$$f_X(z) = \sum_{k=0}^{\infty} p_k z^k$$

che è sicuramente convergente per $|z| \leq 1$. Osserviamo che:

$$\mathbb{E}(e^{-tX}) = f_X(e^{-t}) \equiv \phi_X(t)$$

che permette di estendere la definizione anche a variabili aleatorie continue, mentre $\phi(t)$ è detta funzione generatrice dei momenti, poiché le sue derivate in 0 corrispondono i momenti successivi.

Nel resto del capitolo useremo $f(z)$ per indicare la funzione caratteristica associata a X_1. Risulterà utile anche la seguente:

Definizione 1.2. *Data una funzione f, definiamo la sua iterata come:*

$$f^{(0)}(z) = f(z)$$

$$f^{(n+1)}(z) = f(f^{(n)}(z))$$

Nel corso del capitolo assumeremo che le p_k siano tali per cui:

1. $\forall k, \quad p_k < 1$
2. $m = f'(1) = \sum_{k=0}^{\infty} k p_k < 1$

È interessante osservare che tutta l'informazione contenuta nelle p_k viene trasferita senza perdita nella funzione caratteristica corrispondente $f(z)$, tuttavia queste ultime sono più semplici da maneggiare per determinate operazioni che possono coinvolgere la variabile aleatoria. Un utile esempio di questo fatto è costituito dal seguente:

Teorema 1.1. *Date due variabili aleatorie indipendenti X, Y e posta $Z = X + Y$, si ha:*

$$f_Z(z) = f_X(z) f_Y(z)$$

Dimostrazione. È sufficiente osservare che:

$$P(Z = n) = \sum_{k=0}^{n} P(X = k) P(Y = n - k)$$

e che quindi:

$$f_Z(z) = \sum_i P(Z = i) z^n = \sum_i \sum_{k=0}^{i} P(X = k) P(Y = i - k) z^k z^{i-k} =$$

$$= \left(\sum_m P(X = m) z^m \right) \left(\sum_n P(Y = n) z^n \right) = f_X(z) f_Y(z)$$

□

Il risultato si generalizza in modo ovvio alla somma di un qualunque numero di variabili aleatorie indipendenti che avrà pertanto come funzione generatrice il prodotto delle funzioni generatrici. In virtù di questo è facile ottenere il seguente risultato:

Teorema 1.2. *La funzione generatrice di X_n è l'iterata n-esima di $f(z)$:*

$$f_{X_n}(z) = f^{(n)}(z)$$

Dimostrazione. Per induzione, ovviamente si ha:

$$f^{(1)}(z) = f(z) = f_{X_1}(z)$$

Inoltre se supponiamo che si abbia $X_n = k$, segue dalla proprietà 2, dato che i k oggetti hanno evoluzioni indipendenti, e dal teorema (1.1), che la funzione generatrice di X_{n+1} sarà $f(z)^k$. Da cui in generale si avrà:

$$f_{X_{(n+1)}}(z) = \sum_{k=0}^{\infty} P(X_n = k) f(z)^k = f^{(n)}(f(z)) = f^{(n+1)}(z)$$

□

Per semplificare la notazione nel seguito indicheremo la funzione generatrice delle X_n semplicemente con f_n

1.3 La probabilità d'estinzione

Un interessante quesito che ci si può porre è quello di valutare la probabilità d'estinzione. Prima di tutto è necessaria una formalizzazione.

Definizione 1.3. *Chiamiamo* ***estinzione*** *l'evento E per cui la sequenza X_n è diversa da 0 solo per un numero finito di termini, cioè è definitivamente 0. Indicheremo con q la sua probabilità: $P(E) = q$.*

Osserviamo che poiché se X_n è uguale a 0, saranno 0 anche tutte le X_n con probabilità 1, si può riscrivere:

$$q = \lim_{n\to\infty} P((X_1 = 0) \cup (X_2 = 0) \dots (X_n = 0)) = \lim_{n\to\infty} P(X_n = 0) = \lim_{n\to\infty} f_n(0) \tag{1.2}$$

Questo risultato ci permette di ricavare il valore di q almeno in una forma implicita.

Teorema 1.3. *La probabilità di estinzione è la più piccola soluzione nell'intervallo $[0, 1]$ dell'equazione:*

$$f(x) = x \tag{1.3}$$

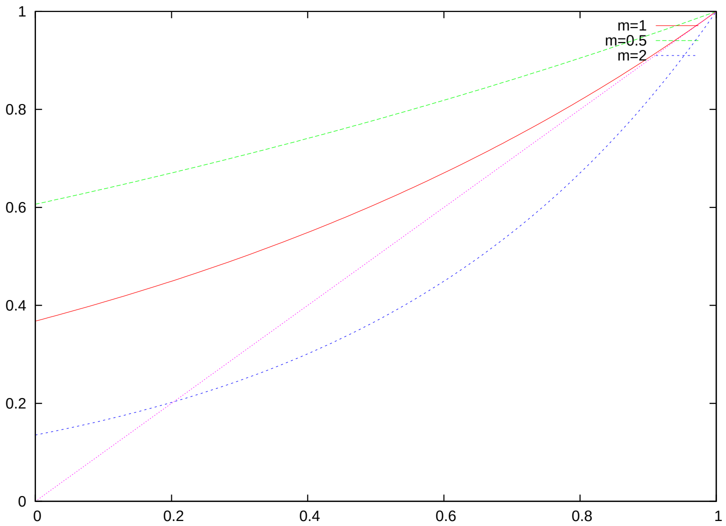


Figura 1.1: Diverse funzioni generatrici di una poissoniana al variare del numero medio di figli m

Dimostrazione. Nell'intervallo $[0, 1]$, $f(z)$ è certamente continua. Pertanto usando la (1.2):

$$f(q) = \lim_{n\to\infty} f_n(0) = q$$

Il che dimostra che q soddisfa l'equazione (1.3). Ci resta da far vedere che è proprio la minima tra le sue soluzioni. Ora chiamiamo $\tilde{q}$ la più piccola. Poiché $f(x)$ è non decrescente su $[0, 1]$ si avrà che:

$$f([0, \tilde{q}]) \subset [0, f(\tilde{q})] = [0, \tilde{q}]$$

Perciò per induzione segue che:

$$\lim_{n\to\infty} f_n(0) \leq \tilde{q}$$

e dalla (1.2) essendo $\tilde{q}$ la minima: $q = \tilde{q}$ □

Va osservato che poiché $f(1) = 1$ una soluzione esiste sempre minore di 1 com'è giusto trattandosi di una probabilità. Inoltre se $m \leq 1$ allora $f'(x) \leq 1$ e $f(x) > x$, quindi certamente $q = 1$.

Non è complesso a questo punto collegare questi risultati con l'andamento della distribuzione delle X_n per $n \to \infty$. Vale infatti:

Teorema 1.4. *Si ha:*

$$\lim_{n\to\infty} P(X_n = k) = 0 \qquad k = 1, 2, \ldots$$

Dimostrazione. Conviene guardare la probabilità che partendo da un valore k la sequenza possa ritornarvi:

$$R_k = P(\cup_{j=1}^{\infty}(X_{n+j} = k)|X_n = k)$$

Facciamo vedere che $R_k < 1$. Distinguiamo due casi:

- $p_0 = 0$. Allora
  $$\cup_{j=1}^{\infty}(X_{n+j} = k) = (X_{n+1} = k)$$
  e quindi $R_k = p_1^k < 1$ poiché ad ogni generazione ogni individuo non può morire senza riprodursi e devono perciò avere, tutti e k, un solo figlio.

- $p_0 \neq 0$. In tal caso:
  $$\cup_{j=1}^{\infty}(X_{n+j} = k) \subset (X_{n+1} = 0)^c$$
  ove la c indica l'insieme complementare. Perciò:
  $$R_k \leq P((X_{n+1} = 0)^c|X_n = k) = P_{0k} = p_0^k < 1$$

In generale otteniamo $R_k < 1$. Ciò prova che tutti gli stati $k = 1, 2, \ldots$ sono transienti e quindi, dalla teoria delle catene di Markov, è nulla la probabilità che X_n assuma infinite volte lo stesso valore k diverso da 0, perciò deve andare a 0 o a ∞. □

Osservazione 1.1. *Val la pena di notare che questo teorema ci fornisce un risultato per il valore asintotico della distribuzione delle X_n indipendente da m, cioè dal numero medio d'oggetti che un oggetto può produrre. Esso asserisce quindi che, al limite, o la popolazione si estingue (con la probabilità q già vista) oppure diviene infinita (con la probabilità complementare $1 - q$). Ovviamente se $m \leq 1$, allora $q = 1$ e quindi la popolazione s'estingue certamente. Se invece $m > 1$ la popolazione può estinguersi o diventare infinita.*

1.4 Distribuzioni asintotiche

Vogliamo in questa sezione analizzare sinteticamente se esiste un qualche modo in cui si può dare un significato alla nozione di distribuzione asintotica o stazionaria per le X_n. Tralasceremo i dettagli matematici delle varie dimostrazioni ed in alcuni punti ci limiteremo ad enunciare dei risultati. Merita d'essere ricordato che esistono più nozioni di convergenza per una successione di variabili aleatorie Z_n, qui ricapitoliamo le principali:

1. Convergenza in probabilità
$$\forall \epsilon > 0 \quad \Rightarrow \quad \lim_{n\to\infty} P(|Z_n - Z| > \epsilon) = 0 \tag{1.4}$$
2. Convergenza in norma L^2
$$\lim \mathbb{E}(|Z_n - Z|^2) = 0 \tag{1.5}$$
3. Convergenza con probabilità 1 [1]
$$\exists A \subset \Omega,\ P(A) = 1 \quad t.c. \quad \forall \omega \in A \quad \lim_{n\to\infty} Z_n(\omega) = Z \tag{1.6}$$

Ciascuna di queste nozioni impone la convergenza della distribuzione delle Z_n a quella di Z in ogni punto di continuità di quest'ultima. Detto questo, osserviamo che i risultati della precedente sezione hanno mostrato come la distribuzione asintotica per le X_n sia poco interessante, in quanto mostra un andamento banale. Tuttavia a seconda del valore di m è possibile individuare una opportuna normalizzazione delle X_n, che corrisponde alla giusta scala a cui guardare il fenomeno, per cui si ottiene una distribuzione stazionaria interessante. Per determinare il fattore correttivo conviene guardare il valor medio delle X_n:
$$\mathbb{E}(X_n) = f_n'(1)$$
Mediante la derivazione a catena otteniamo immediatamente:
$$f_{n+1}'(1) = f'(f_n(1))f_n'(1) = f'(1)f_n'(1)$$
che fornisce per induzione
$$f_n'(1) = m^n$$

[1] con Ω s'indica lo spazio di probabilità su cui sono definite le variabili aleatorie.

Questo c'induce a considerare anziché la successione delle X_n quella delle $X_n m^{-n}$.

1.4.1 Caso supercritico: $m > 1$

Poniamo $W_n = X_n m^{-n}$. Senza entrare nel merito della dimostrazione è possibile far vedere che le W_n convergono sia secondo la (1.6) che secondo la (1.5) ad una variabile aleatoria W con $\mathbb{E}(W) = 1$. Il seguente risultato ci fornisce un'equazione funzionale per W:

Teorema 1.5. *Detta $\phi(t)$ la funzione generatrice dei momenti di W: $\phi(t) = \mathbb{E}(e^{-tW})$, allora è l'unica soluzione di:*

$$\Phi(mt) = f(\Phi(t)) \tag{1.7}$$

con la condizione $\Phi'(0) = -1$

Dimostrazione. Dato che $W_n = X_n m^{-n}$ la sua funzione generatrice dei momenti ϕ_n soddisferà:

$$\phi_n(t) = \mathbb{E}(e^{-tW_n}) = f_n\left(\exp\left(-\frac{t}{m^n}\right)\right)$$

Usando il teorema 1.2 otteniamo:

$$\phi_{n+1}(mt) = f(\phi_n(t))$$

e prendendo il limite si trova che $\phi(t)$ deve soddisfare la (1.7). Inoltre la convergenza in norma L^2 ci assicura che:

$$1 = \mathbb{E}(W_n) \to \mathbb{E}(W) \quad \Rightarrow \quad \phi'(0) = -1$$

Omettiamo per brevità la prova dell'unicità. □

Un risultato ancora interessante riguarda la velocità con cui $P(X_n = 0)$ tende a q. Una risposta è fornita dal seguente risultato:

Teorema 1.6. *Se $m > 1$ e $q > 0$, allora si ha:*

$$f_n(0) = q - d[f'(q)]^n + O(f'(q)^{2n})$$

Si tratta sostanzialmente di una stima sulle iterate di una funzione. Si noti che può essere generalizzata al caso in cui $m < 1$ e quindi $q = 1$, fornendo:

$$1 - f_n(0) \simeq Km^n \tag{1.8}$$

1.4.2 Caso subcritico: $m < 1$

Come abbiamo visto in questo caso le X_n tendono a 0 e perciò la distribuzione asintotica è priva d'interesse. Tuttavia ciò che si può fare è guardare la distribuzione condizionale di X_n posto che $X_n \neq 0$ cioè:

$$P(X_n = j | X_n \neq 0) = \frac{P(X_n = j)}{\sum_{k=1}^{\infty} P(X_n = k)}$$

Dal secondo membro si nota che in quest'espressione sia il numeratore che il denominatore tendono a 0 nel limite $n \to \infty$, ma il rapporto può essere finito, come asserisce il seguente:

Teorema 1.7. *Supponiamo che $m < 1$ e che $\mathbb{E}(X_1^2) < \infty$, allora esiste il limite:*

$$b_j = \lim_{n \to \infty} P(X_n = j | X_n \neq 0)$$

con $\sum_{j=1}^{\infty} b_j = 1$ e posto

$$g(z) = \sum_j b_j z^j$$

si ha che $g(1) = 1$, $g'(1) < \infty$ e che g è l'unica soluzione di:

$$G(f(z)) = mG(z) + 1 - m$$

che soddisfi tali condizioni.

Per brevità omettiamo la dimostrazione un po' tecnica di questo fatto.

1.4.3 Caso critico: $m = 1$

Abbiamo visto che nei casi in cui $m \neq 1$, è possibile individuare una distribuzione asintotica ed essa può assumere una forma pressoché arbitraria, dato che le equazioni funzionali per $\phi(t)$ e $g(t)$ possono essere invertite determinando la $f(z)$ corrispondente. Nel caso critico vedremo invece che si giunge ad una forma indipendente da $f(z)$, una distribuzione di tipo esponenziale. Ovviamente abbiamo già visto per $m \leq 1$ si ha che X_n tende a 0, quindi è preliminarmente necessaria una normalizzazione; per individuarla, vale il seguente risultato dovuto a Kolmogorov (1938):

$$P(X_n > 0) \simeq \frac{2}{n f''(1)} \tag{1.9}$$

Da quest'ultima otteniamo subito che:

$$\mathbb{E}(X_n|X_n \neq 0) \simeq \frac{nf''(1)}{2}$$

Quindi saremo interessati alla distribuzione condizionata di $\frac{2X_n}{nf''(1)}$ posto che $X_n \neq 0$. Vale allora il seguente risultato:

Teorema 1.8. *Supposto che $m = 1$ e $f'''(1) < \infty$ allora*

$$\lim_{n\to\infty} P\left(\frac{2X_n}{nf''(1)} > u\big|X_n \neq 0\right) = e^{-u}$$

1.4.4 Rapidità della convergenza delle f_n

Abbiamo già visto che le $f_n(x)$ tendono per ogni $x < 1$ a q. Inoltre nel caso critico s'è già visto che $1 - f_n(x) \simeq n$. Nei casi in cui $m \neq 1$, la situazione è però diversa. Per comodità scriviamo $\gamma = f'(q)$. Allora ponendo:

$$Q_n(x) = \gamma^{-n}[f_n(x) - q]$$

si avrà che:

$$Q'_n(x) = \frac{f'_n(x)}{\gamma^n} = \prod_{j=0}^{n-1} \frac{f'(f_j(x))}{\gamma} = \prod_{j=0}^{n-1} \left(1 - \left(\frac{f'(f_j(x))}{\gamma}\right)\right)$$

e con tale scrittura si può mostrare il seguente teorema:

Teorema 1.9. *Vale la seguente:*

$$\forall x \in [0,1) \quad \exists \lim_{n\to\infty} Q'_n(x) \equiv U(x)$$

Inoltre:

1. $\forall x \in [0,1)$ $G(x) = 0 \Leftrightarrow m < 1, \sum_j p_j j \log_j = \infty$
2. *Negli altri casi:* $x \in [0,1), G(x) > 0$ *e* $\lim_{x\to q} G(x) = 1$

e quindi ponendo:

$$\forall x \in [0,1) \quad Q(x) = \int_q^x G(x)dx$$

si ha nello stesso intervallo:

$$\lim_{n\to\infty} Q_n(x) = Q(x)$$

Per la forma di $Q(x)$ ci viene in aiuto il seguente:

Teorema 1.10. *Data l'equazione funzionale:*

$$\mathcal{Q}(f(x)) = \gamma\mathcal{Q}(x) \qquad x \in [0, 1)$$

si ha che $Q(x)$ è l'unica soluzione che soddisfi:

$$\mathcal{Q}(q) = 0 \qquad \lim_{x \to q} Q'(x) = 1$$

Omettiamo le dimostrazioni un po' tecniche che sono presenti su [22].

1.5 Galton-Watson a più tipi

Abbiamo fin ora considerato processi in cui gli oggetti sono tutti identici e quindi indistinguibili. Vogliamo ora generalizzare il problema al caso in cui vi siano più tipi di oggetti aventi un diverso comportamento probabilistico. Supporremo che i tipi di oggetti siano un numero finito k, ma è possibile generalizzare la trattazione al caso in cui i tipi di oggetti siano un insieme astratto. Senza entrare nei dettagli, per cui rimandiamo a [21][22], vedremo come il problema può essere formalizzato ancora in termini delle funzioni generatrici. Inoltre si trova una forte analogia con il Galton-Watson già considerato che espliciteremo nella probabilità d'estinzione. Può essere interessante notare che tale analogia prosegue nel senso che si possono trovare delle distribuzioni limite, determinate da una nuova grandezza che gioca il ruolo del numero medio di figli.

1.5.1 Definizione

Visto che in questo caso abbiamo k tipi di oggetti diversi, dovremo avere altrettante funzioni generatrici. Ciascuna di esse ci darà informazioni sulla prole di ciascun tipo. Notiamo che ogni tipo potrà avere tra i suoi figli oggetti di tipo diverso, altrimenti il problema si ridurrebbe al Galton-Watson trattato prima; abbiamo perciò:

$$f^I(s_1, \ldots, s_k) = \sum_{j_1, \ldots, j_k} p^I(j_1, \ldots, j_k) s_1^{j_1} \cdots s_k^{j_k} \qquad 0 \leq s_I \leq 1 \qquad (1.10)$$

dove l'indice maiuscolo scorre su tutti i tipi. Nel seguito indicheremo i vettori di $\mathbb{R}^k$ col grassetto e le loro componenti con un indice in basso maiuscolo. Per semplificare le espressioni vettoriali useremo la seguente notazione:

- $\mathbf{u} \leq \mathbf{v} \Leftrightarrow u_I \leq v_I \quad \forall I \in \{1, \ldots, k\}$
- $\mathbf{0} = (0, \ldots, 0) \qquad \mathbf{1} = (1, \ldots 1)$
- $e_I = (0, \ldots, 1, \ldots, 0)$ con l'1 all'i-esima componente
- $x^y = \prod x_I^{y_I}$
- $|x| = (|x_1|, \ldots, |x_k|)$
- $||x|| = \max(|x_1|, \ldots, |x_k|)$
- Se M è una matrice: $||M|| = \sup(|m_{IJ}|)$
- $\mathbf{p}(\mathbf{j}) = (p^1(\mathbf{j}), \ldots, p^k(\mathbf{j})), \qquad \mathbf{f}(\mathbf{s}) = (f^1(\mathbf{j}), \ldots, f^k(\mathbf{s}))$
- Sia C_k il cubo unitario: $= \{x | \mathbf{0} \leq \mathbf{x} < \mathbf{1}\}$.

Con queste convenzioni, la (1.10) diventa:

$$\mathbf{f}(\mathbf{s}) = \sum_{j \in \mathbb{N}^k} \mathbf{p}(\mathbf{j}) \mathbf{s}^{\mathbf{j}}$$

Possiamo quindi dare la seguente:

Definizione 1.4. *Un Galton-Watson a k tipi è una catena di Markov $\{\mathbf{X}_n, n = 0, 1, 2, \ldots\}$ su $\mathbb{N}^k$, in cui le probabilità di transizione sono date da:*

$$P(\mathbf{i}, \mathbf{j}) = P(\mathbf{X_{n+1}} = \mathbf{j} | \mathbf{X_n} = \mathbf{i}) = \mathbf{s}^{\mathbf{j}} : [\mathbf{f}(\mathbf{s})^{\mathbf{i}}]$$

ove con l'ultimo membro a destra s'intende il coefficiente di $\mathbf{s}^{\mathbf{j}}$ nello sviluppo di $\mathbf{f}(\mathbf{s})^{\mathbf{i}}$.

Indicheremo con $\mathbf{X}_{n,\mathbf{i}}$ il vettore contenente il numero di oggetti di ciascun tipo alla generazione n-esima partendo da $\mathbf{X}_0 = \mathbf{i}$. Con tale notazione s'ottiene una **proprietà d'additività**:

$$\mathbf{Z}_{n,\mathbf{i}} = \sum_I i_I \mathbf{Z}_{n,\mathbf{e_I}}$$

Anche in questo caso sono molto utili le funzioni generatrici e indicheremo con $\mathbf{f}_{n,\mathbf{i}}(\mathbf{s})$ la funzione generatrice associata a $\mathbf{X}_{n,\mathbf{i}}$; ometteremo la $\mathbf{i}$ nei casi in cui non sarà necessario specificare la condizione iniziale. In analogia col teorema 1.2, si ottiene:

$$\mathbf{f}_{n+m}(\mathbf{s}) = \mathbf{f}_n[\mathbf{f}_m(\mathbf{s})]$$

Nelle considerazioni che seguiranno è da ritenersi escluso il caso in cui $\mathbf{f}(\mathbf{s})$ sia una funzione lineare (caso singolare), che corrisponde alla situazione in cui ciascun oggetto fa esattamente un figlio e si riduce ad una catena di Markov finita tipica.

1.5.2 Estinzione e stati transienti

Abbiamo visto che nel caso in cui è presente un tipo solo, è determinante per il valore della probabilità d'estinzione, il valore del numero medio di figli. In questo caso è possibile un'analogia. Per cominciare si definisce la matrice $M = \{m_{IJ}\}$ come:

$$m_{IJ} = \frac{\partial f^I}{\partial s_J}(\mathbf{1})$$

che rappresenta il numero medio di figli di tipo J da parte di un individuo di tipo I. Chiaramente M sarà una matrice non negativa. A determinare i diversi regimi per la probabilità d'estinzione e per le distribuzioni asintotiche interviene in questo caso ρ, l'autovalore massimo di M. Per chiarire la cosa abbiamo:

Teorema 1.11. *Tutti gli stati escluso lo* $\mathbf{0}$ *sono transienti:*

$$P(\mathbf{Z_n} = \mathbf{j} \textit{ infinite volte}) = 0$$

Teorema 1.12. *Sia q^I la probabilità d'estinzione del processo iniziato con l'oggetto I e $\mathbf{q} = (q^1, \ldots, q^k)$. Valgono allora:*

- $\rho \leq 1 \;\Leftrightarrow \mathbf{q} = \mathbf{1}$
- $\rho > 1 \;\Leftrightarrow \mathbf{q} < \mathbf{1}$
- $\forall \mathbf{s} \in C_k \quad \lim_{n\to\infty} \mathbf{f}(\mathbf{s}) = \mathbf{q}$
- $\mathbf{q}$ *è l'unica soluzione di* $\mathbf{f}(\mathbf{s}) = \mathbf{s}$

Come si vede quest'ultimo teorema richiama il 1.3 e l'unica differenza consiste nel fatto che le grandezze qui in gioco sono tutte vettoriali. Questo cambiamento costituisce una complicazione, poiché il valor medio dei figli diviene una matrice. Come si è già detto, esistono anche in questo contesto delle distribuzioni asintotiche e gli elementi cruciali sono ρ, l'autovalore massimo di M e i corrispondenti autovettori sinistro e destro. Per un approfondimento sulla questione rimandiamo ai testi già citati.

Capitolo 2

Riformulazione quantistica

In questo capitolo ci occuperemo di far vedere come il Galton-Watson process si presti ad essere trattato mediante il formalismo della teoria dei campi. Accadrà che le probabilità di transizione tra diverse generazione potranno essere espresse, esattamente come accade nei processi di scattering, attraverso dei diagrammi di Feynman. Sarà comunque necessaria qualche modifica al formalismo per via del fatto che:

- il tempo è discreto;
- gli oggetti che trattiamo benché identici tra loro non si comportano esattamente come particelle indistinguibili;
- non vi è un analogo dell'interferenza quantistica.

Vedremo nelle prossime sezioni come queste difficoltà si presentano e in che modo potranno essere superate. Il risultato finale sarà comunque confortante, come vedremo, e potrà estendersi agevolmente al Galton-Watson a più tipi.

2.1 Definizione dello spazio e della dinamica

Il primo passo che dobbiamo compiere è quello di costruire uno spazio di Hilbert adatto per il nostro problema, in cui l'evoluzione probabilistica del processo di branching si interpreti come una evoluzione quantistica. La dinamica quantistica che andremo ad introdurre farà passare da una generazione a quelle successive perciò nello spazio in cui opereriamo dobbiamo ammettere l'esistenza di operatori di creazione e distruzione per oggetti a tutti i tempi. Poniamo perciò:

Definizione 2.1. *Siano a_t e $a_t^\dagger$, $t \in \mathbb{N}$, gli operatori di distruzione e creazione di un individuo al tempo t. Avremo le regole di commutazione:*

$$[a_k, a_h] = 0 \tag{2.1}$$

$$[a_k^\dagger, a_h^\dagger] = 0 \tag{2.2}$$

$$[a_k, a_h^\dagger] = \delta_{kh} \tag{2.3}$$

A questo punto lo spazio di Hilbert si costruisce nel modo solito, cioè completando lo spazio che s'ottiene dall'applicazione dei polinomi negli operatori di creazione:

Definizione 2.2. *Sia $|0\rangle$ il vuoto e sia $\mathcal{H}$ lo spazio di Fock:*

$$\mathcal{H} = \overline{\{P(a_{t_1}^\dagger, \ldots, a_{t_n}^\dagger)|0\rangle, P \in \mathbb{C}[x_1, x_2, \ldots]\}}$$

ove la barra indica la chiusura dello spazio e $\mathbf{C}[x_1, x_2, \ldots]$ i polinomi a coefficienti complessi nelle variabili $x_1, x_2, \ldots$.

Con questa scelta abbiamo formalizzato la configurazione ad n oggetti come uno stato di n particelle bosoniche indistinguibili. Come avevamo preannunciato questo conduce a delle differenze, come risulterà più evidente in seguito, a meno di scegliere opportunamente la normalizzazione degli stati con un numero di particelle definito:

$$|n, t\rangle = (a_t^\dagger)^n |0\rangle \qquad \langle n, t| = \langle 0|(a_t)^n$$

per la quale si ha:

$$\langle m, t_1 | n, t_2\rangle = \langle 0|(a_{t_1})^m (a_{t_2}^\dagger)^n |0\rangle = m! \delta_{t_1, t_2} \delta_{mn}$$

Osservazione 2.1. *Gli stati $|n, t\rangle$ costituiscono una base per $\mathcal{F}$:*

$$\mathcal{F} = \mathrm{Span}\{|n, t\rangle,\ (n, t) \in \mathbb{N}^2\} \tag{2.4}$$

La costruzione che abbiamo così effettuato fa sì che siano possibili stati in cui coesistono oggetti a tempi diversi, ad esempio: $a_0^\dagger a_1^\dagger |0\rangle$. Conviene allora distinguere un sottospazio:

Definizione 2.3 (Stati fisici)**.** *Detti* ***stati a tempi definiti*** *i seguenti:*

$$\mathcal{F}_t = \{P(a_t^\dagger)|0\rangle,\ P \in \mathbb{C}^1[x]\} \subset \mathcal{H}$$

chiameremo ***stati fisici*** *la loro somma diretta al variare del tempo:*

$$\mathcal{F} = \bigoplus_{t=0}^{\infty} \mathcal{F}_t$$

Quindi denotiamo con stati fisici gli stati in cui non possono coesistere oggetti a tempi diversi. Al suo interno tuttavia sono presenti le sovrapposizioni lineari di oggetti a tempi diversi, ad esempio: $(a_0^\dagger + a_1^\dagger)|0\rangle$. Ovviamente nel calcolare le ampiezze non prenderemo in considerazione stati di questo tipo, ma solo gli stati a tempi definiti. Vedremo comunque che sarà utile avere a che fare con tutto il sottospazio degli stati fisici. Richiediamo che gli operatori che andremo a costruire connettano stati a tempi definiti. Perciò:

Definizione 2.4 (Operatori fisici)**.** *Dato un operatore A su $\mathcal{H}$, diremo che è* ***fisico*** *se:*

$$\forall\ |\phi\rangle \in \mathcal{F}_t \quad \Rightarrow \quad A|\phi\rangle \in \mathcal{F}_{t'}$$

Questo implica in più che connetteranno stati fisici. Sarà per tanto utile la seguente:

Definizione 2.5 (Proiettore sul vuoto)**.** *Denotiamo con $\mathcal{O}_t$ il proiettore*[1] *sul vuoto al tempo t-esimo. Cioè:*

$$\mathcal{O}_t|0\rangle = |0\rangle \qquad \mathcal{O}_t(a_t^\dagger)^n|0\rangle = 0$$

$$\forall h \neq t \quad [\mathcal{O}_t, a_h^\dagger] = 0$$

Il proiettore sul vuoto al tempo t elimina gli stati al tempo t e lascia intatti quelli a tempi diversi. Vedremo che sarà utile per assicurarci che gli operatori connettano stati a tempi definiti. Cerchiamo ora di costruire l'hamiltoniana del nostro sistema. Sappiamo che essa corrisponde al generatore delle traslazioni temporali. Nel nostro caso il tempo è discreto e quindi

[1] una forma esplicita di $\mathcal{O}_t$ è data da:

$$\mathcal{O}_t = \frac{\sin(\pi a_t^\dagger a_t)}{\pi a_t^\dagger a_t}$$

l'hamiltoniana è collegata all'evoluzione tra una generazione e quella direttamente successiva. Partendo da un oggetto singolo ad una generazione, dovremo avere al tempo successivo una sovrapposizione dei vari stati aventi tanti individui quanti i figli possibili:

$$|1,t\rangle \quad \xrightarrow{H} \quad \sum_n b_n |n, t+1\rangle$$

Questo c'induce a definire:

Definizione 2.6. *Sia:*

$$H(t) = \Phi(a_{t+1}^\dagger) a_t \tag{2.5}$$

l'hamiltoniana al tempo t ove si è posto:

$$\Phi(z) = \sum_n b_n z^n \tag{2.6}$$

Chiaramente ci aspettiamo che i b_n siano esprimibili in termini dei p_n definiti nella (1.1); cerchiamo ora di trovare tale legame. Il fatto che i p_n non dipendano dal tempo implica che il sistema sia invariante per traslazioni temporali e ciò si traduce qui nel fatto che nemmeno b_n non dipendono dal tempo.

Saremmo ora portati a definire l'evolutore temporale a partire dall'hamiltoniana nella maniera solita:$U(t) = e^{H(t)}$. In questo caso tuttavia non otterremmo un operatore fisico. Proviamone l'azione su uno stato a due individui:

$$U(t)(a_t^\dagger)^2|0\rangle = \left[1 + \Phi(a_{t+1}^\dagger)a_t + \frac{(\Phi(a_{t+1}^\dagger)a_t)^2}{2}\right](a_t^\dagger)^2|0\rangle =$$

$$= \left[(a_t^\dagger)^2 + 2\Phi(a_{t+1}^\dagger)a_t^\dagger + \Phi^2(a_{t+1}^\dagger)\right]|0\rangle$$

Come si vede lo stato risultante contiene ancora termini al tempo precedente e quindi non è uno stato fisico. Per far funzionare le cose correttamente si pone:

Definizione 2.7. *Sia $U(t)$ l'evolutore temporale elementare al tempo t:*

$$U(t) = \mathcal{O}_t e^{H(t)} \tag{2.7}$$

Il proiettore che abbiamo aggiunto elimina i termini aggiuntivi e fa sì che resti un operatore fisico. Nel caso in cui $b_n = \delta_{1,n}$, l'hamiltoniana si riduce a:

$$H_0(t) = a^\dagger_{t+1} a_t$$

che chiameremo **hamiltoniana** libera, di conseguenza:

Definizione 2.8. *Sia:*

$$U_0(t) = \mathcal{O}_t e^{a^\dagger_{t+1} a_t}$$

l'evolutore temporale libero.

È facile verificare che l'evolutore temporale libero trasla semplicemente il tempo in avanti senza modificare lo stato.

Per capire meglio l'azione di un evolutore di questa forma sugli stati ci viene in aiuto il seguente

Teorema 2.1. *Siano $b^\dagger, b$ e $c^\dagger, c$ operazioni di creazione e distruzione indipendenti:*

$$[b, c] = 0 \quad [b^\dagger, c^\dagger] = 0 \qquad [b, b^\dagger] = [c, c^\dagger] = 1 \tag{2.8}$$

Posto:

$$H = f(c^\dagger)b \qquad U = \mathcal{O}e^H$$

ove $\mathcal{O}$ è il proiettore sul vuoto delle $b^\dagger$, si ha:

$$\mathcal{O}e^H (b^\dagger)^n |0\rangle = (f(c^\dagger))^n |0\rangle \tag{2.9}$$

Dimostrazione.

$$\mathcal{O} \sum_{k=0}^{\infty} \frac{(f(c^\dagger)b)^k}{k!} (b^\dagger)^n |0\rangle$$

ora dalla (2.8):

$$(f(c^\dagger)b)^k = (f(c^\dagger))^k b^k$$

Inoltre:

$$\forall k < n \quad \mathcal{O} b^k (c^\dagger)^n |0\rangle = 0$$

$$\forall k > n \quad b^k (c^\dagger)^n |0\rangle = 0$$

$$b^n (c^\dagger)^n |0\rangle = n! |0\rangle$$

Perciò si ottiene:

$$U(b^\dagger)^n|0\rangle = (\Phi(c^\dagger))^n|0\rangle$$

□

Corollario 2.2.

$$U(t)|n,t\rangle = U(t)(a_t^\dagger)^n|0\rangle = (\Phi(a_{t+1}^\dagger))^n|0\rangle$$

Corollario 2.3. *Sia $h(z)$ una funzione analitica. Allora si ha:*

$$U(t)h(a_t^\dagger)|0\rangle = h(\Phi(a_{t+1}^\dagger))|0\rangle$$

Dimostrazione. Se h è un polinomio, è una conseguenza della linearità e della (2.9). Si conclude allora per continuità. □

Possiamo notare che con questi risultati è facile scrivere l'inverso dell'evolutore libero:

Proposizione 2.4. *Sugli stati fisici l'inverso dell'evolutore libero è:*

$$U_0^{-1}(t) = \mathcal{O}_{t+1}e^{a_t^\dagger a_{t+1}}$$

Dimostrazione. Prendiamo la base data dalla (2.4) e verifichiamolo per essa:

$$\mathcal{O}_t e^{a_{t+1}^\dagger a_t}|n,h\rangle = \begin{cases} |n,h\rangle & h \neq t \\ |n,t+1\rangle & h = t \end{cases}$$

$$\mathcal{O}_{t+1} e^{a_t^\dagger a_{t+1}}|n,h\rangle = \begin{cases} |n,h\rangle & h \neq t+1 \\ |n,t\rangle & h = t+1 \end{cases}$$

Da cui si vede che gli operatori sono uno l'inverso dell'altro. □

Ora diamo qualche definizione:

Definizione 2.9. *Poniamo:*

$$\forall t,t_0 \in \mathbf{N}, t \geq t_0 \qquad \mathcal{U}(t,t_0) \equiv \begin{cases} 1 & t = t_0 \\ \prod_{k=0}^{t-t_0-1} U(t_0+k) & t > t_0 \end{cases}$$

cioè l'operatore che evolve dal tempo t_0 al tempo t.

$$\rho_{mn}^{(k)} \equiv \frac{1}{m!}\langle m,t+k|\mathcal{U}(t+k,t)|n,t\rangle$$

$$\rho_{mn} \equiv \rho_{mn}^{(1)}$$

C'interessa a questo punto trovare una forma più esplicita per $\mathcal{U}$. Faremo nuovamente uso delle iterate delle funzioni riutilizzando la notazione della 1.2 A questo punto possiamo enunciare:

Teorema 2.5. *Vale la seguente uguaglianza:*

$$k \geq 1 \qquad \mathcal{U}(t+k,t) = \mathcal{O}_t e^{H(t+k,t)}$$

ove:

$$H(t+k,t) = \Phi^{(k)}(a^\dagger_{t+k})a_t$$

Dimostrazione. Per induzione su k: se $k = 1$ è ovvio perché $H(t+1,t) = H(t)$. Supposto vero per k si ha:

$$\mathcal{U}(t+k+1,t)|n,t\rangle = U(t+k)\mathcal{U}(t+k,t)|n,t\rangle = U(t+k)\mathcal{O}_t e^{H(t+k,t)}|n,t\rangle$$

Dal teorema 2.1, s'ottiene:

$$\mathcal{O}_t e^{H(t+k,t)}|n,t\rangle = \left(\Phi^{(k)}(a^\dagger_{t+k})\right)^n|0\rangle$$

e dal corollario 2.3:

$$U(t+k)\mathcal{O}_t e^{H(t+k,t)}|n,t\rangle = U(t+k)\left(\Phi^{(k)}(a^\dagger_{t+k})\right)^n|0\rangle =$$

$$= \left(\Phi^{(k+1)}(a^\dagger_{t+k+1})\right)^n|0\rangle = \mathcal{O}_t e^{H(t+k+1,t)}|n,t\rangle$$

□

2.2 Determinazione dell'hamiltoniana

Ci occupiamo ora di trovare la forma corretta dell'hamiltoniana per riprodurre le probabilità del Galton-Watson process. Vogliamo determinare i b_n, introdotti nella (2.6), da cui poi la funzione Φ; per farlo dobbiamo collegare le probabilità calcolate quantisticamente con le p_n. In una teoria quantistica la probabilità di transizione si ottiene prendendo il modulo quadro dell'elemento di matrice dell'operatore di evoluzione tra lo stato iniziale e quello finale:

$$p = \frac{|\langle f|U|i\rangle|^2}{|\langle f|f\rangle\langle i|i\rangle|} = \frac{\langle i|U^\dagger|f\rangle\langle f|U|i\rangle}{|\langle f|f\rangle\langle i|i\rangle|} = \frac{\langle i|U^{-1}|f\rangle\langle f|U|i\rangle}{|\langle f|f\rangle\langle i|i\rangle|} \tag{2.10}$$

Nel nostro caso questa scelta si dimostra poco adeguata. Infatti, come s'era detto all'inizio del capitolo, implicherebbe la comparsa di effetti puramente quantistici. Vedremo che risulterà corretto per i nostri fini porre

$$\frac{1}{n!}\left|\langle n, t+1|1, t\rangle\right| = P[|1, t\rangle \to |n, t+1\rangle] \tag{2.11}$$

ove il secondo membro indica la probabilità di passare dallo stato ad un individuo a quello ad n. Analizzeremo nel dettaglio le differenze concettuali derivanti da questo differente approccio. Cominciamo col determinare i coefficienti b_n, imponendo che la probabilità della (2.11) eguagli le p_n, cioè la probabilità che ha un individuo di avere n figli:

$$\begin{aligned} p_n &= \frac{1}{n!}\left|\langle n, t+1|U(t)|1, t\rangle\right| = \frac{1}{n!}\left|\langle 0|(a_{t+1})^n \Phi(a^\dagger_{t+1}))|0\rangle\right| = \\ &= \frac{1}{n!}\left|\langle 0|(a_{t+1})^n \sum_{k=0}^{\infty} b_k (a^\dagger_{t+1})^k|0\rangle\right| = \frac{b_n}{n!}\left|\langle 0|(a_{t+1})^n (a^\dagger_{t+1})^n|0\rangle\right| = b_n \end{aligned} \tag{2.12}$$

Da cui otteniamo proprio:

$$b_n = p_n \tag{2.13}$$

e quindi:

$$\Phi(z) = \sum_{k=0}^{\infty} p_k z^k = f(z) \tag{2.14}$$

e cioè la $\Phi(z)$ introdotta nell'hamiltoniana eguaglia esattamente la funzione generatrice introdotta nella 1.1. Osserviamo che la nostra scelta ci ha permesso di ottenere un risultato particolarmente semplice, che ci incoraggia a proseguire verso lo sviluppo di tale formalismo. Se, invece, per definire le probabilità di transizione avessimo usato, in piena analogia con quanto accade in meccanica quantistica, la (2.10) avremmo ottenuto: $b_n = \sqrt{p_n}$. Ciò non mostra alcuna incongruenza a questo livello, tuttavia nel momento in cui vogliamo calcolare la probabilità P_{mn} che ci siano m individui all'istante $t+1$ partendo da n all'istante t, abbiamo:

$$P_{mn} = \left|\langle m, t+1|\mathcal{U}(t+1, t)|n, t\rangle\right|^2 = |\rho_{mn}|^2 \tag{2.15}$$

Note le p_n possiamo calcolare P_{mn} per via puramente probabilistica:

$$P_{mn} = \sum_{\substack{k_i \geq 0 \\ k_1+\ldots+k_n=m}} \prod_{i=1}^{n} p_{k_i} \tag{2.16}$$

C'interessa confrontare le due. Semplifichiamo le cose verificando per un caso semplice: $m = 2, n = 2$. Allora la (2.16) diventa:

$$P_{22} = 2p_0p_2 + p_1^2$$

La (2.15):

$$P_{22} = |\rho_{22}|^2 = \frac{1}{4}\left|\langle 0|a_{t+1}^2\Phi(a_{t+1}^\dagger)^2|0\rangle\right|^2 = |2\sqrt{p_0p_2} + p_1|^2$$

Le due sono chiaramente diverse. Ciò è dovuto ad una profonda differenza concettuale. Nello sviluppo della teoria quantistica dei campi, i diagrammi di Feynman corrispondono alle ampiezze. Ad ogni vertice corrisponde una costante d'accoppiamento nel computo finale dell'ampiezza di transizione. Invece, nel processo di Galton-Watson ad ogni vertice corrisponde la probabilità di quel processo e sono proprio le probabilità associate ai differenti eventi a sommarsi. Risulta chiaro come avevamo preannunciato che non è possibile utilizzare la definizione solita di probabilità, dato che il processo di branching ha ben poco di quantistico di per se. Tuttavia resta il fatto che la teoria dei campi può servirci come ottimo ausilio a costo di utilizzare la (2.11), apportando alcune modifiche interpretative. Sarà necessario pertanto considerare le ampiezze, piuttosto che i moduli quadri di queste, come probabilità. Al fine di garantire la coerenza della nostra costruzione anche i valori medi degli operatori necessitano di una ridefinizione. Poniamo perciò:

$$\langle A\rangle_\phi = \langle \mathbb{1}|A|\phi\rangle \tag{2.17}$$

dove s'intende:

$$|\mathbb{1}\rangle \equiv \sum_{k,t}\frac{|k,t\rangle}{k!} \qquad \langle 1| = \sum_{k,t}\frac{\langle k,t|}{k!}$$

2.3 La probabilità d'estinzione

Abbiamo già definito nella 1.3 la probabilità d'estinzione $q = P(E)$. Vediamo in che modo è possibile ridefinire tale grandezza in termini quantistici:

$$q = \lim_{t\to\infty}\langle 0|\mathcal{U}(t,1)|1\rangle$$

Usando il teorema (2.5) si ottiene:

$$q = \langle 0|\mathcal{O}_1 \exp(\Phi^{(t)}(a_t^\dagger)a_1)|1\rangle = \lim_{t\to\infty}\Phi^{(t)}(0)$$

Poiché Φ ha tutte le derivate positive ed è quindi convessa possiamo dire che p_E è la più piccola soluzione dell'equazione:

$$x = \Phi(x) \tag{2.18}$$

come s'era già trovato nella (1.3)

2.4 Il numero d'individui e le distribuzioni asintotiche

É facile introdurre l'osservabile collegato al numero d'Individui al tempo t:

$$N(t) = a_t^\dagger a_t$$

e più in generale l'operatore numero dato da:

$$N = \sum_t N(t) = \sum_t a_t^\dagger a_t$$

In accordo con la (2.17) si avrà pertanto che il numero medio d'individui nello stato ϕ_t, al tempo t, sarà dato da:

$$\langle N \rangle_{\phi_t} = \langle \mathbb{1}|N(t)|\phi\rangle = \sum_k \langle \mathbb{1}|N(t)|k,t\rangle\langle k,t|\phi_t\rangle = \sum_k k\langle k,t|\phi_t\rangle$$

Se prendiamo $|\phi_t\rangle = U(t-1)|1,t-1\rangle$ otteniamo:

$$\langle N \rangle = \sum_k k\langle k,t|U(t)|1,t-1\rangle = \sum_k k\langle k,t|\Phi(a_t^\dagger)|0\rangle = \sum_k kp_k = m$$

coerentemente con la sezione 1.2. Cerchiamo di capire in che modo possiamo affrontare il problema delle distribuzioni asintotiche sfruttando questo punto di vista. Usiamo la seguente:

Definizione 2.10. *Chiamiamo $|\phi(t)\rangle$* ***stato asintotico*** *se si ha:*

$$|\phi(t+1)\rangle \equiv U(t)|\phi(t)\rangle = U_0(t)|\phi(t)\rangle \qquad \forall t$$

e cioè se l'evoluzione temporale su di esso eguaglia l'evoluzione libera.

Sappiamo già che in questo caso però l'unico stato che soddisfa tale espressione è lo stato vuoto. Per vederlo scriviamo:

$$|\phi(t)\rangle = \sum_n k_n |n,t\rangle = \sum_n k_n (a_t^\dagger)^n |0\rangle$$

osservando che la definizione impone che i k_n non dipendano dal tempo. Da cui:

$$U(t)|\phi(t)\rangle = \sum_n k_n f(a_{t+1}^\dagger)^n |0\rangle = \sum_n k_n (a_{t+1}^\dagger)^n |0\rangle = U_0(t)|\phi(t)\rangle$$

Applicando a sinistra $\frac{\langle m,t+1|}{m!}$ otteniamo:

$$k_m = \sum_n \frac{1}{n!} \langle m,t| k_n f^n(a_{t+1}^\dagger)|0\rangle = \frac{1}{m!} \sum_n k_n \left. \frac{\partial^m f^n(x)}{\partial x^m} \right|_{x=0} \tag{2.19}$$

che, definendo la matrice:

$$\mathbf{P}_{mn} = \frac{1}{m!} \left. \frac{\partial^m f^n(x)}{\partial x^m} \right|_{x=0} = P(X_{t+1} = m | X_t = n) \tag{2.20}$$

si trasforma nel sistema lineare:

$$(\mathbf{P} - \mathbf{I})\mathbf{k} = \mathbf{0} \tag{2.21}$$

ove $\mathbf{k} = (k_0, k_1, \ldots)$ e $\mathbf{I}$ indica l'identità. È facile vedere che

Teorema 2.6. *L'unica soluzione di (2.21) è data da* $\mathbf{e}_0 = (1, 0, 0, \ldots)$.

Dimostrazione. Osserviamo che se $\tilde{k}$ è una soluzione di (2.21), allora vale anche:

$$\mathbf{P}^N \tilde{\mathbf{k}} = \tilde{\mathbf{k}}$$

È facile verificare $\mathbf{P}^N = \mathbf{P}(f^{(N)})$ ove $\mathbf{P}$ è espressa dalla (2.20) e allora dato che $f^{(N)}(z) \to q$ uniformemente con tutte le derivate in ogni compatto in $[0, 1)$:

$$\tilde{\mathbf{k}} = \lim_{N\to\infty} \mathbf{P}^N \tilde{\mathbf{k}} = \tilde{k}_0 \mathbf{e}_0$$

che mostra l'asserto. □

Rimane comunque interessante domandarsi se si può dare un significato in questo contesto alle distribuzioni asintotiche normalizzate opportunamente che abbiamo introdotto nella sezione 1.4.

2.5 Estensione al caso a più tipi

Abbiamo visto nello scorso capitolo come è possibile estendere il Galton-Watson al caso in cui siano presenti diversi tipi di oggetti. Anche nell'approccio quantistico tale estensione è possibile ed anzi, come vedremo, è diretta. Infatti nelle precedenti sezioni abbiamo visto come si può creare lo spazio in cui vivono gli stati del sistema e su di essi abbiamo definito un'hamiltoniana. A questo punto semplicemente cambiandone la forma si possono ottenere sistemi più complessi. Indichiamo con Γ l'insieme dei tipi di oggetti distinti. Dobbiamo quindi ammettere l'esistenza di differenti operatori di creazione e costruzione per ciascun elemento $\gamma \in \Gamma$. Ovviamente operatori inerenti a tipi distinti commutano tra loro. Quindi:

$$\forall \gamma, \gamma' \in \Gamma,\ t, t' \in \mathbb{N} \qquad \begin{array}{rc} [a_{t,\gamma}, a_{t',\gamma'}] = & 0 \\ [a_{t,\gamma}, a^\dagger_{t',\gamma'}] = & \delta_{tt'}\delta_{\gamma\gamma'} \end{array}$$

ove gli operatori $a^\dagger_{t,\gamma}$ e $a_{t,\gamma}$ rispettivamente creano e distruggono un oggetto di tipo γ al tempo t. Lo spazio che s'ottiene prendendo i polinomi in tutti gli operatori di creazione (di tipi diversi e a tempi diversi) sarà il prodotto tensore degli spazi a singolo tipo introdotti nella 2.2:

$$\mathcal{H} = \bigotimes_{\gamma \in \Gamma} \mathcal{H}_\gamma$$

ed allo stesso modo s'estende la definizione di **stati fisici**, prendendo il prodotto tensore degli stati fisici di ciascun oggetto. Abbiamo quindi creato uno spazio complessivo in cui può coesistere un arbitrario numero di oggetti che possono essere di tipo diverso. Estendiamo nel seguente modo la definizione della base di stati a numero d'oggetti definito:

$$|n_1, n_2, \ldots, n_k, t\rangle = \left(\prod_{i=1}^{k} (a^\dagger_{t,\gamma_i})^{n_i} \right) |0\rangle$$

ove $k = \#\Gamma$. Ne deriva l'ortogonalità nella forma:

$$\langle m_1, m_2, \ldots, m_k | n_1, n_2, \ldots, n_k \rangle = \prod_{i=1}^{k} n_i! \delta_{n_i,m_i}$$

Definito lo stato a più tipi come il prodotto tensore di quelli ad un singolo tipo, dobbiamo trovare l'hamiltoniana corrispondente. Esattamente come accade in un sistema puramente quantistico, se nella nostra costruzione prendiamo un'hamiltoniana che è semplicemente la somma delle hamiltoniane di singolo oggetto, otterremo un modello in cui i tipi diversi non interagiscono tra loro e seguono ciascuno la sua propria evoluzione. Verifichiamolo con un esempio. Poniamo perciò:

$$H(t) = \sum_{\gamma \in \Gamma} f_\gamma(a^\dagger_{t+1,\gamma}) a_{t,\gamma}$$

L'evolutore temporale sarà sempre dato dalla (2.7), ove il proiettore sul vuoto al tempo t sarà chiaramente il prodotto dei proiettori sul vuoto al tempo t per ciascun tipo di oggetto. Prendendo lo stato in cui sono presenti due tipi distinti $|\phi, t\rangle = a^\dagger_{t,\gamma_1} a^\dagger_{t,\gamma_2} |0\rangle$; applicando tale evolutore s'ottiene:

$$\begin{aligned} U(t)|\phi, t\rangle &= \mathcal{O}_t \sum_{n=0}^{\infty} \frac{H(t)^n}{n!} = \\ &= \frac{1}{2!} \cdot 2 f_1(a^\dagger_{t+1,\gamma_1}) f_2(a^\dagger_{t+1,\gamma_2}) a_{t,\gamma_1} a_{t,\gamma_2} |\phi, t\rangle = f_1(a^\dagger_{t+1,\gamma_1}) f_2(a^\dagger_{t+1,\gamma_2}) |0\rangle \end{aligned} \tag{2.22}$$

ove nella seconda uguaglianza s'è tenuto conto che l'unico termine non nullo nello sviluppo in serie dell'evolutore temporale è quello in cui vengono distrutti i due oggetti di tipo γ_1 e γ_2. La cosa si fa più interessante se ammettiamo che nell'hamiltoniana siano presenti dei termini interagenti e che quindi mischiano gli operatori di campo di oggetti di tipo diverso. In analogia con quanto fatto nella sezione 1.5, possiamo supporre che ciascun tipo possa produrre oggetti di tipo diverso dal suo e che la funzione $f^\gamma(x_{\gamma_1}, x_{\gamma_2}, \ldots, x_{\gamma_k})$ sia la funzione generatrice della distribuzione dei figli di un oggetto di tipo $\gamma \in \Gamma$, come nella (1.10). Basterà allora prendere come hamiltoniana:

$$H(t) = \sum_{\gamma \in \Gamma} f^\gamma(a^\dagger_{t,\gamma_1}, a^\dagger_{t,\gamma_2}, \ldots, a^\dagger_{t,\gamma_k}) a_{t,\gamma}$$

per riprodurre l'evoluzione del Galton-Watson a più tipi. Infatti, la (2.22) continua a valere ed in generale l'evoluzione indipendente di ciascun individuo nel branching process, ci assicura che l'hamiltoniana data riprodurrà le corrette probabilità.

Capitolo 3

Branching process continui

3.1 Introduzione

Nei precedenti capitoli ci siamo occupati di introdurre il Galton-Watson process, come viene normalmente trattato attraverso il formalismo delle funzioni generatrici e, successivamente, di come è possibile affrontare tale trattazione in un contesto diverso qual è quello della teoria dei campi. Abbiamo visto che tale adattamento, che a priori può apparire strano e forzato, rivela alcuni pregi, visto che si presta facilmente a generalizzazioni successive. In questo capitolo sposteremo lo sguardo sui processi di branching continui, che sono un tipo particolare di processo stocastico. I processi stocastici, inclusi quelli Markoviani, hanno trovato numerose applicazioni nella biologia e nella fisica. Il motivo per definire e studiare poi un Galton-Watson process continuo è duplice: da un lato essi permettono lo studio di alcuni problemi mediante una diversa impostazione (che permette di rendere anche la durata di una generazione una variabile aleatoria), dall'altro, passando al caso continuo, è possibile raggiungere risultati analitici più agevoli, come spesso accade. Nella seconda parte del capitolo mostreremo come anche nel caso continuo sia possibile tentare un approccio basato sulla teoria dei campi. In particolare, in questo caso le analogie con quanto si fa comunemente in fisica, divengono più evidenti e come vedremo sarà possibile sviluppare un vero e proprio formalismo di path-integral.

3.2 Markov branching process

Cominciamo introducendo il problema dal punto di vista puramente probabilistico. Si tratta di studiare un processo stocastico $Z(t)$ che assume valori interi non negativi e che rappresenta il numero d'individui. Il processo è Markoviano nel senso che se supponiamo nota con certezza la popolazione ad un tempo t_1: $Z(t_1) = i$, la successiva evoluzione dipenderà da i e dal tempo t_2, ma non da qualunque proprietà degl'individui che dipenda dalla storia precedente t_1. Inoltre, il nostro processo in particolare sarà assimilabile al Galton-Watson process che abbiamo analizzato nel precedente capitolo poiché assumeremo che le differenti famiglie sviluppatesi dagli i individui al tempo t_1 non s'influenzino l'un l'altra.

Per studiare $Z(t)$ è utile guardare le probabilità di transizione:

Definizione 3.1. *Dato un processo stocastico $Z(t)$ definiamo le probabilità di transizione come:*

$$P_{ij}(t_1, t_2) = P\{Z(t_2) = j | Z(t_1) = i\} \tag{3.1}$$

Le proprietà che abbiamo prima enunciato si riassumeranno in due sistemi di equazioni differenziali che determinano le $P_{ij}(t_1, t_2)$. Tutto il processo è in realtà descritto da un gruppo di funzioni: $b_i(t)$ e $p_{ij}(t)$. Esse rappresentano:

- $b(t)$: la probabilità per unità di tempo che avvenga un cambiamento di stato a partire da $Z(t) = i$. Cioè la probabilità di un cambio di stato nell'intervallo $(t, t + \delta t)$ sarà data al prim'ordine da: $b_i(t)\delta t$;

- $p_{ij}(t)$: è la probabilità che il nuovo stato sia j, supposto che avvenga un cambio di stato al tempo t a partire dallo stato i.

In realtà a noi interessa collegare queste funzioni alla probabilità che un individuo scompaia e alla probabilità che l'individuo abbia un determinato numero di figli. Perciò:

Definizione 3.2. *Data $b(t)$ funzione continua non negativa diciamo che: $b(t)\delta t$ è la probabilità che ha un individuo di morire nell'intervallo $(t, t+\delta t)$. Dopo la morte, la probabilità di essere sostituito da n individui sarà $\rho_n(t)$.*

Possiamo assumere nella nostra trattazione che $p_1(t) = 0$. Infatti questo evento corrisponde ad un individuo che muore e viene sostituito da un altro individuo. Ma poiché nel nostro approccio Markoviano gl'individui non hanno età, questo non corrisponderebbe a nessun cambiamento reale. Vediamo in che modo è possibile legare $b_i(t)$ e $p_{ij}(t)$ a $b(t)$ e $\rho_n(t)$. Se abbiamo i individui al tempo t, la probabilità $d_k(t, t+\delta t)$ di avere k morti nell'intervallo $(t, t+\delta t)$ si ottiene dalla distribuzione binomiale e al prim'ordine in δt vale:

$$d_0(t, t+\delta t) = 1 - ib(t)\delta t \qquad d_1(t, t+\delta t) = ib(t)\delta t \qquad d_{k\geq 2}(t, t+\delta t) = o(\delta t)$$

Tale probabilità coincide con quella che vi sia un cambiamento di stato, perciò siamo portati a porre: $b_i(t) = ib(t)$. Inoltre la probabilità che dopo questa morte vi siano j individui eguaglia la probabilità che l'individuo morto venga sostituito da $j - i + 1$ nuovi individui. Perciò: $p_{ij}(t) = \rho_{j-i+1}(t)$. Questo mostra che le grandezze che determinano l'evoluzione sono le funzioni: $b(t)$ e $\rho_n(t)$.

3.3 Le equazioni per le $P_{ij}(t_1, t_2)$

In questa sezione otterremo le equazioni soddisfatte dalle $P_{ik}(t)$ e alcune delle proprietà delle soluzioni. Cominciamo con l'osservare cosa accade se guardiamo la differenza: $P_{ik}(t_0, t+\delta t) - P_{ik}(t_0, t)$. La probabilità di transizione nell'intervallino infinitesimo δt può variare per due motivi:

- diminuisce nei casi in cui lo stato era già diventato k, ma è avvenuto un cambiamento di stato durante δt;
- aumenta nei casi in cui gli individui erano diventati j, ma durante δt un individuo è morto ed è stato sostituito da $k - j + 1$ individui.

Questo si traduce nella seguente:

$$P_{ik}(t_0, t+\delta t) - P_{ik}(t_0, t) = -(kb(t)\delta t)P_{ik}(t_0, t) + \sum_j P_{ij}(t_0, t)(jb(t)\delta t)\rho_{k-j+1}$$

Gli estremi nella somma su j si determinano in questo modo: perché possa esserci una morte deve essere $j \geq 1$ e affinché lo stato finale possa essere k deve essere $j \leq k + 1$, dato che l'individuo che muore non può essere sostituito da un numero negativo d'individui. Nel limite $\delta t \to 0$ s'ottiene il sistema d'equazioni differenziali:

$$\frac{\partial P_{ik}(t_0,t)}{\partial t} = -kb(t)P_{ik}(t_0,t) + b(t)\sum_{j=1}^{k+1} j\rho_{k-j+1}(t)P_{ij}(t_0,t) \tag{3.2}$$

con la condizione iniziale: $P_{ik}(t_0,t_0) = \delta_{ik}$.

In maniera analoga si può analizzare l'incremento: $P_{ik}(t+\delta t,t_1)-P_{ik}(t,t_1)$. In questo caso avremo che la probabilità di transizione:

- aumenta nei casi in cui nell'intervallo δt vi era un cambio di stato (che ora non è più incluso);
- diminuisce nei casi in cui nell'intervallo δt uno degli i individui muore, viene sostituito da $j-i+1$ individui e nell'intervallo $(t+\delta t,t_1)$ rimanente vi è una transizione da j a k.

Questo conduce a:

$$P_{ik}(t+\delta t,t_1)-P_{ik}(t,t_1) = ib(t)P_{ik}(t+\delta t,t_1)\delta t - \sum_j ib(t)\rho_{j-i+1}(t)P_{jk}(t+\delta t,t_1)\delta t$$

e per gli estremi della sommatoria abbiamo che: $j > i-1$ perché l'individuo morto nell'intervallo δt può al più non essere sostituito da nessuno, ma non v'è limite superiore perché la transizione da j a k avviene in un tempo finito. Nel limite $\delta t \to 0$ otteniamo:

$$\frac{\partial P_{ik}(t,t_1)}{\partial t} = -kb(t)P_{ik}(t,t_1) + ib(t)\sum_{j=i-1}^{\infty} \rho_{j-i+1}(t)P_{jk}(t,t_1) \tag{3.3}$$

con la condizione iniziale: $P_{ik}(t_1,t_1) = \delta_{ik}$. Si può verificare che esiste sempre una soluzione P_{ik} comune per le (3.2) e le (3.3). La seguente osservazione poi assicura l'indipendenza nell'evoluzione delle differenti famiglie come avevamo richiesto al Markov branching process:

Osservazione 3.1. *Ogni soluzione della (3.2) soddisfa:*

$$P_{ik}(t_0,t_1) = \sum_{r_1+r_2+\ldots+r_i=k} P_{1r_1}(t_0,t_1)\ldots P_{1r_i}(t_0,t_1)$$

Osservazione 3.2 (Unicità). *Le (3.2) hanno sempre un'unica soluzione per la quale in generale vale la condizione:*

$$\sum_k P_{ik}(t_0, t_1) \leq 1 \qquad i = 0, 1, \ldots \tag{3.4}$$

Nei casi in cui vale l'uguaglianza anche le (3.3) hanno un'unica soluzione. Nei casi invece di disuguaglianza stretta le (3.3) hanno infinite soluzioni che esplodono (infatti la (3.4) ci dice che la probabilità di rimanere finita non è 1), ma di queste ve n'è sempre solo una che soddisfi la proprietà caratterizzante di un branching process, cioè l'osservazione 3.1.

Un caso particolare, che analizzeremo nella prossima sezione, è costituito dal seguente:

Definizione 3.3. *E' detto Markov branching process omogeneo quello in cui:*

$$b(t) = b \qquad p_n(t) = p_n$$

Dalle (3.2) è evidente che in questo caso s'avrà:

$$P_{ij}(t_1, t_2) = P_{ij}(t_2 - t_1)$$

3.4 Approccio quantistico

Come nel precedente capitolo, ci occuperemo di introdurre una struttura quantistica per studiare l'evoluzione di un Markov branching process. D'ora in poi ci riferiremo al caso omogeneo descritto nella osservazione 3.3. Avremo per tanto gli operatori $a, a^\dagger$ per cui vale:

$$[a, a^\dagger] = 1 \tag{3.5}$$

Definizione 3.4. *Sia $|0\rangle$ il vuoto e sia $\mathcal{H}$ lo spazio di Fock:*

$$\mathcal{H} = \overline{\{P(a^\dagger)|0\rangle, P \in \mathbb{C}[x]\}}$$

ove la barra indica la chiusura dello spazio e $\mathbf{C}[x]$ *i polinomi a coefficienti complessi nella variabile x.*

Definizione 3.5. *Useremo come base la seguente:*

$$|n\rangle = (a^\dagger)^n|0\rangle \qquad \langle n| = \langle 0|a^n$$

La base è ortonormale, perciò s'avrà: $\langle n|m\rangle = m!\delta_{mn}$. *Inoltre:*

$$a^\dagger|n\rangle = |n{+}1\rangle \qquad a|n\rangle = n|n\rangle \qquad \langle n|a = \langle n{+}1| \qquad \langle n|a^\dagger = \langle n{-}1|n \quad (3.6)$$

La relazione di completezza è:

$$\sum_{n\geq 0} \frac{|n\rangle\langle n|}{n!} = \mathbf{1} \tag{3.7}$$

Definizione 3.6. *Dato lo stato* $|\phi(t)\rangle \in \mathcal{H}$*, la probabilità di avere n individui al tempo t sarà data da:*

$$p_n(t) = \frac{\langle n|\phi(t)\rangle}{\langle n|n\rangle} = \frac{\langle n|\phi(t)\rangle}{n!}$$

Postuliamo per tutti gli stati un'evoluzione governata da un'hamiltoniana H [1]:

$$\frac{d}{dt}|\phi(t)\rangle = H|\phi(t)\rangle \tag{3.8}$$

Vogliamo esprimere l'operatore H usando le funzioni che determinano il branching process: $b(t)$ e $\rho_n(t)$; per farlo sfrutteremo le (3.2). Prendiamo in considerazione gli stati $|n(t)\rangle$ che soddisfino la (3.8) e per cui si abbia:

$$\langle m|n(0)\rangle = m!\delta_{nm} \quad \Rightarrow \quad |n(0)\rangle = |n\rangle$$

In base alla (3.1) otteniamo che:

$$n_m(t) \equiv \frac{\langle m|n(t)\rangle}{m!} = P_{nm}(0,t) = P_{nm}(t)$$

derivando quest'uguaglianza e usando le (3.2) otteniamo:

$$\frac{1}{m!}\frac{d}{dt}\langle m|n(t)\rangle = \frac{1}{m!}\frac{d}{dt}P_{nm}(t) = -mbP_{nm}(t) + b\sum_{j=1}^{m+1} j\rho_{m-j+1}P_{nj}(t) =$$

$$= -mb\frac{1}{m!}\langle m|n(t)\rangle + b\sum_{j=1}^{m+1} j\rho_{m-j+1}\frac{1}{j!}\langle j|n(t)\rangle$$

[1] Poiché ci stiamo occupando dei processi omogenei possiamo ammettere che H sia indipendente dal tempo.

Ora prendendo la (3.8) e proiettando sullo stato $|m\rangle$ e inserendo la (3.7):

$$\frac{d}{dt}\langle m|n(t)\rangle = \sum_{k\geq 0}\frac{\langle m|H|k\rangle\langle k|n(t)\rangle}{k!} = -mb\langle m|n(t)\rangle + b\sum_{j=1}^{m+1} j\rho_{n-j+1}\frac{m!}{j!}\langle j|n(t)\rangle =$$

$$= \sum_{k\geq 0}\left[-mb\langle m|k\rangle + b\sum_{j=1}^{m+1} j\rho_{m-j+1}\frac{m!}{j!}\langle j|k\rangle\right]\langle k|n(t)\rangle =$$

$$= \sum_{k\geq 0} m!\left[-mb\delta_{mk} + b\sum_{j=1}^{m+1} j\rho_{m-j+1}\delta_{jk}\right]\langle k|n(t)\rangle \quad (3.9)$$

Ponendo in aggiunta $\rho_{\{i<0\}} = 0$, possiamo scrivere:

$$\frac{1}{m!}\langle m|H|k\rangle = -mb\delta_{mk} + bk\rho_{m-k+1}$$

A questo punto c'interessa poter trovare un'espressione di H in termini degli operatori $a, a^\dagger$. Per far ciò è utile riscrivere quest'ultima uguaglianza come:

$$\frac{H_{mk}}{m!} = -kb\delta_{mk} + b\sum_n k\rho_n\delta_{k,m-n+1}$$

Osserviamo allora che:

$$\langle m|a^\dagger a|k\rangle = km!\delta_{mk}$$

$$\langle m|(a^\dagger)^n a|k\rangle = k\langle m|(a^\dagger)^n|k-1\rangle = k\langle m|n+k-1\rangle = km!\delta_{m,n+k-1} = km!\delta_{k,m-n+1}$$

Questo ci permette di scrivere un'espressione per H:

$$H = -ba^\dagger a + b\sum_n \rho_n (a^\dagger)^n a \quad (3.10)$$

Definizione 3.7. *Poniamo:*

$$h(s) = \sum_{n=0}^{\infty} p_n s^n \qquad |s| \leq 1$$

Si scrive:

$$H = b[h(a^\dagger) - a^\dagger]a$$

Abbiamo quindi ottenuto una forma esplicita per l'hamiltoniana del sistema. Una soluzione formale per la (3.8) si ottiene immediatamente introducendo l'evolutore temporale $U = e^{Ht}$, per cui si ha:

$$\phi(t) = U(t)\phi(0) \quad (3.11)$$

Nella prossima sezione vedremo in che modo è possibile riscrivere U utilizzando un formalismo di path-integral.

3.5 L'integrale sui cammini

Il primo passo per questa riscrittura è notare che è possibile stabilire un omomorfismo tra $\mathcal{H}$ e lo spazio delle funzioni olomorfe sul disco unitario:

$$\mathcal{H} \ni |\phi\rangle = \sum_n \phi_n |n\rangle \to \phi(z) = \sum_n \phi_n z^n$$

Osserviamo che nel nostro contesto, in cui le ϕ_n rappresentano le probabilità che ci siano n individui, $\phi(z)$ è la funzione generatrice. Al fine di preservare i prodotti scalari possiamo associare agli stati duali una distribuzione nella maniera seguente:

$$\langle n| \to (-1)^m \delta^{(m)}(z) = \left(-\frac{d}{dz}\right)^m \delta(z)$$

Infatti integrando per parti è facile verificare che si ha:

$$\langle m|n\rangle = n!\delta_{mn} = \int dz z^n (-1)^m \delta^{(m)}(z)$$

Osserviamo che quest'ultima eguaglia:

$$= \frac{1}{2\pi} \int dz dw z^n (\imath w)^m e^{-\imath z w}$$

Questo permette di completare la definizione dell'omomorfismo:

$$|n\rangle \to z^n \qquad \langle m| \to (\imath w)^m$$

e il prodotto scalare è dato dall'integrazione con la misura: $\frac{e^{-\imath z w} dz dw}{2\pi}$.

In maniera naturale possiamo trovare il corrispondente dell'operatore A. È sufficiente utilizzare le relazioni di completezza (3.7) per ottenere:

$$A = \sum_{mn} \frac{\langle m|A|n\rangle}{m!n!} |m\rangle\langle n| \to A(z, \imath w) = \sum_{mn} \frac{z^m (\imath w)^n}{m!n!} A_{mn}$$

ove si sono scritti gli elementi di matrice: $A_{mn} = \langle m|A|n\rangle$. La funzione $A(z, \imath w)$ associata all'operatore A è detto kernel. Perciò:

$$|\phi\rangle = A|\phi\rangle = \sum_{mn} \frac{\langle m|A|n\rangle}{m!n!} |m\rangle\langle n|\phi\rangle \to \phi(z) = \frac{1}{2\pi} \int A(z, \imath w) \phi(z') e^{-\imath w z'} dz' dw \tag{3.12}$$

In maniera diretta è possibile trasportare la composizione di due operatori:

$$[AB](z, \imath w) = \frac{1}{2\pi} \int dz' dw' A(z, \imath w') B(z', \imath w) e^{-\imath z' w'} = \tag{3.13}$$

È importante ottenere una relazione che permetta di collegare il Kernel di un operatore con la sua rappresentazione in forma normale. Infatti ogni operatore si può mettere nella forma:

$$A = \sum_{mn} \alpha_{mn} (a^\dagger)^m a^n$$

Qual è il Kernel corrispondente?

Definizione 3.8 (Normal Kernel)**.** *Dato l'operatore A, che si scrive in forma normale come:*

$$A = \sum_{mn} \alpha_{mn} (a^\dagger)^m a^n$$

definiamo il normal kernel di A come:

$$\mathcal{N}_A(z, \imath w) = \sum_{mn} \alpha_{mn} z^m (\imath w)^n$$

In pratica il normal kernel s'ottiene sostituendo nell'espressione dell'operatore in forma normale gli $a^\dagger$ con z e gli a con $\imath w$.

Allora vale il seguente:

Teorema 3.1. *Sia dato un operatore A. Allora vale la seguente relazione:*

$$A(z, w) = e^{zw} \mathcal{N}_A(z, w)$$

La dimostrazione può essere trovata su [13]. Ora abbiamo tutti gli strumenti necessari per poter riscrivere l'evolutore temporale come un path integral. Ci è utile la formula di Trotter:

$$U(t) = \lim_{N \to \infty} \left(1 + \frac{tH}{N}\right)^N$$

C'interessa trovare un'espressione per il kernel di $U(t)$. Dato che sappiamo che il kernel del prodotto di due operatori si scrive attraverso la (3.13), ci è sufficiente trovare il kernel di uno dei fattori al secondo membro. Vale allora:

$$\left(1 + \frac{tH}{N}\right)(z, w) = e^{zw} \left(1 + \frac{\mathcal{N}_H(z, w)t}{N}\right)$$

che si ottiene facilmente dal teorema 3.1 e osservando che $\mathcal{N}_{\mathbf{1}}(z,w) = 1$. Perciò il kernel di $U(t)$ si scrive come:

$$U_t(z,\imath w) = \lim_{N\to\infty} \prod_{j=1}^{N-1} \int \frac{e^{-\imath\zeta_k\omega_k} d\zeta_k d\omega_k}{2\pi} \prod_{k=1}^{N} e^{\imath\zeta_{k-1}\omega_k} \left[1 + \frac{t\mathcal{N}_H(\zeta_{k-1},\imath\omega_k)}{N}\right]$$

dove si pone: $\zeta_0 = z$ e $\omega_N = w$. E' possibile riscrivere quest'ultima sfruttando il fatto che c'interessa il limite $N \to \infty$ per cui si ha:

$$1 + \frac{t\mathcal{N}_H(\zeta_{k-1},\imath\omega_k)}{N} = \exp\left(\frac{t\mathcal{N}_H(\zeta_{k-1},\imath\omega_k)}{N}\right) + o\left(\frac{1}{N}\right)$$

Mettendo insieme i pezzi all'esponente:

$$U_t(z,\imath w) = \lim_{N\to\infty} \prod_{j=1}^{N-1} \int \frac{e^{-\imath\zeta_k\omega_k} d\zeta_k d\omega_k}{2\pi}$$
$$\exp\left\{\sum_{k=1}^{N-1}\left[-\imath\omega_k(\zeta_k - \zeta_{k-1}) + \frac{t\mathcal{N}_H(\zeta_{k-1},\imath\omega_k)}{N}\right] + \frac{t\mathcal{N}_H(\zeta_{N-1},\imath w)}{N} + \imath w\zeta_{N-1}\right\}$$

Prendendone il limite per $N \to \infty$ con $t' = \frac{kt}{N}$, $\zeta_k \to \zeta(t')$ e $\omega_k \to \omega(t')$ la possiamo riscrivere come integrale sui cammini nella forma seguente:

$$U_t(z,\imath w) = \int \mathcal{D}\zeta\mathcal{D}\omega \exp\left\{-\int_0^t dt'[\imath\omega(t')\dot{\zeta}(t') - \mathcal{N}_H(\zeta,\imath\omega)] + \imath w\zeta(t)\right\} \tag{3.14}$$

con le condizioni al contorno: $\zeta(0) = z$ e $\omega(t) = w$. É conveniente eseguire alcuni passaggi per arrivare ad una forma più comoda. Integriamo per parti:

$$= \int \mathcal{D}\zeta\mathcal{D}\omega \exp\left\{-\int_0^t dt'[-\imath\dot{\omega}(t')\zeta(t') - \mathcal{N}_H(\zeta,\imath\omega)] + \imath\omega(0)z\right\}$$

e cambiamo il verso del tempo:

$$t' \to -t' \Rightarrow \begin{cases} \zeta(t') \to \zeta(1-t') \\ \omega(t') \to \omega(1-t') \end{cases}$$

con cui otteniamo l'espressione:

$$U_t(z,\imath w) = \int \mathcal{D}\zeta\mathcal{D}\omega \exp\left\{-\int_0^t dt'[\imath\dot{\omega}(t')\zeta(t') - \mathcal{N}_H(\zeta,\imath\omega)] + \imath z\omega(t)\right\} \tag{3.15}$$

ove l'integrale funzionale è da intendersi sulle funzioni reali $\zeta(\tau)$ e $\omega(\tau)$ per cui vale: $\zeta(t) = z$ e $\omega(0) = w$. S'osservi che attraverso la (3.12) è possibile trasporre la (3.11):

$$\phi_t(z) = \frac{1}{2\pi} \int dz' dw e^{-\imath z' w} U_t(z, \imath w) \phi_0(z') \qquad (3.16)$$

Vedremo nel prossimo paragrafo un'applicazione di questa scrittura.

3.6 Processi di decadimento

Uno dei processi più semplici da studiare che si dimostra completamente risolubile mediante il path integral è il caso in cui nella (3.10) poniamo $h(x) = 1$. Ciò significa che l'unico processo che può avvenire è quello in cui un individuo muore; non è perciò ammesso che venga sostituito da altri individui. Si ha perciò:

$$H = b(1 - a^\dagger)a$$

Dato che questa espressione è già ordinata normalmente possiamo ottenere subito:

$$\mathcal{N}_H = b(1 - z)\imath w$$

Inserendolo nella (3.15):

$$U_t(z, \imath w) = \int \mathcal{D}\zeta \mathcal{D}\omega \exp\left\{ -\int_0^t dt' [\imath \dot{\omega}(t')\zeta(t') + b(\zeta(t') - 1)\imath\omega(t')] + \imath z \omega(t) \right\}$$

É utile effettuare il seguente cambiamento di variabile: $(\zeta(t') - 1) \rightarrow \zeta(t')$

$$= \int \mathcal{D}\zeta \mathcal{D}\omega \exp\left\{ -\int_0^t dt' [\imath \dot{\omega}(t')(1 + \zeta(t')) + b\zeta(t')\imath\omega(t')] + \imath z \omega(t) \right\} =$$
$$= \int \mathcal{D}\zeta \mathcal{D}\omega \exp\left\{ -\int_0^t dt' [\imath \dot{\omega}(t')\zeta(t') + b\zeta(t')\imath\omega(t')] + \imath(z - 1)\omega(t) + \imath w \right\} =$$
$$= \int \mathcal{D}\zeta \mathcal{D}\omega \exp\left\{ -\int_0^t dt' [\imath\zeta(t')(\partial_{t'} + b)\omega(t')] + \imath(z - 1)\omega(t) + \imath w \right\} \qquad (3.17)$$

Per il calcolo, consideriamo dapprima:

$$\int \mathcal{D}\zeta \exp\left[-\int_0^t dt' \imath\zeta(t')[\partial_{t'} + b]\omega(t') \right] = \int \mathcal{D}\zeta \prod_{0<s\leq t} \exp\left(-\imath\zeta(t')[\partial_{t'} + b]\omega(t') \right)$$

ricordando che: $\int dk e^{ikz} = 2\pi\delta(z)$, otteniamo che:

$$= \text{const.} \times \prod_{0<s\leq t} \delta(\dot{\omega}(s) + b\omega) = \text{const.} \times \prod_{0<s\leq t} \delta(\omega(s) - we^{-bs})$$

dove l'ultima uguaglianza si ottiene imponendo l'equazione data dalla δ:

$$\frac{d\omega(s)}{ds} = -b\omega(s)$$

con la condizione $\omega(0) = w$. Sostituendo nella (3.17):

$$U_t(z, \imath w) = K \exp\left\{\imath w(1 + (z-1)e^{-bt})\right\}$$

Sostituendo nella (3.16) si ottiene infine:

$$\begin{aligned} \phi_t(z) = \frac{K}{2\pi} \int dz' dw \exp\left\{\imath w \left(1 + (z-1)e^{-bt} - z'\right)\right\} \phi_0(z') = \\ = K \int dz' \delta(1 + (z-1)e^{-bt} - z')\phi_0(z') = \\ = \phi_0\left(1 + (z-1)e^{-bt}\right) \end{aligned}$$

dove $K = 1$ si determina imponendo la normalizzazione.

Capitolo 4

L'uso della fisica statistica

4.1 Introduzione

In questo capitolo e nel prossimo ci occuperemo di mettere in luce alcuni legami, non evidenti a priori che hanno una base profonda, tra i fenomeni di branching e la fisica statistica su particolari tipi di reticoli. Vedremo gradualmente, partendo da un caso semplice, che è possibile studiare il processo di Galton-Watson proprio come un sistema di spin su un reticolo che si costruisce bene attraverso un procedimento di ricorsione, a costo di reinterpretare il tempo d'evoluzione del processo di branching come una coordinata spaziale, o come si chiarirà meglio nel capitolo che segue come il grado di magnificazione. Sarà possibile quindi ottenere una hamiltoniana per il sistema che avrà una forma sorprendentemente semplice. La struttura che si ottiene si presta alla derivazione di relazioni di ricorsione per la funzione di partizione, che potrà essere ottenuta in maniera quasi esplicita in quanto soluzione di una relazione funzionale. Vedremo che tale approccio permetterà di ritrovare per altra via dei risultati noti per via puramente probabilistica. Sarà importante osservare che tale approccio non dà un'interpretazione diretta per il concetto di temperatura; si potranno comunque definire delle differenti grandezze, a seconda dello scopo che ci prefiggiamo di volta in volta, che talvolta hanno un valore puramente formale, ma che ne riproducono alcune proprietà, senza tuttavia dar luogo ad una piena analogia. Cominciamo, perciò, da una situazione particolare, che rende la modellizzazione più semplice e permette di chiarire la struttura su cui si basa la costruzione nel caso generale: prendiamo come distribuzione per la prole di un individuo una binomiale con due figli al massimo.

4.2 Caso semplice: distribuzione binomiale di due figli

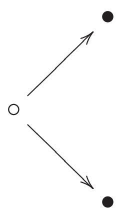
Ci mettiamo nella situazione in cui ogni padre può avere al massimo due discendenti. Inoltre, la distribuzione dei figli è la distribuzione binomiale di parametro p; perciò indicando con P_k la probabilità di avere k figli:

$$P_k = \binom{2}{k} p^k (1-p)^{2-k} \qquad k = 0, 1, 2 \tag{4.1}$$

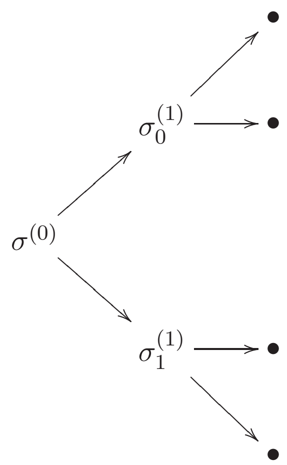
Vediamo in che modo sarà possibile costruire un reticolo su cui studiare questo processo. Ci serviremo di un procedimento iterativo. Affinché tutti gli stati probabilisticamente possibili siano ammissibili, il reticolo deve prevedere che ad ogni individuo, che andrà a corrispondere ad un sito del reticolo, siano collegati, e quindi possano interagire con esso, i due siti corrispondenti agli eventuali figli. Si tratterà poi, una volta realizzata tale geometria, di costruire l'interazione attraverso l'hamiltoniana che riproduca le opportune probabilità. Il reticolo si può ottenere facilmente ragionando nel seguente modo: si parte dal passo iniziale costituito semplicemente da

$$\bullet$$

e ad ogni passo sostituiremo $\bullet$ con il seguente:



Dopo due iterazioni il risultato sarà pertanto:



Iterando il procedimento perciò avremo: 2^k vertici di livello k. Introduciamo ora un campo che in ciascuno dei siti potrà assumere il valore 0 o 1[1], mentre è fissato ad 1 il valore nel sito iniziale:

$$\sigma^* = \sigma^{(0)} = 1 \qquad \sigma_j^{(k)} = 0, 1 \qquad k = 1, 2 \ldots \quad j = 0, \ldots, 2^k - 1$$

Vogliamo aggiungere ora un'interazione tra i siti di questo reticolo che ricalchi le probabilità (4.1). Ci limitiamo ad una hamiltoniana che coinvolga al più 2 siti, che come vedremo sarà efficace in virtù della struttura markoviana del fenomeno; poiché il valore del campo in ciascun sito può essere solo di questi due valori, la più generale espressione possibile per l'hamiltoniana è la seguente[2]:

$$H(\sigma_1, \sigma_2) = c + a\sigma_1 + b\sigma_2 + d\sigma_1\sigma_2 \tag{4.2}$$

Per passare dall'hamiltoniana alla funzione di partizione sappiamo che dobbiamo sommare il peso statistico associato a ciascuna configurazione, su

[1] Questo è evidentemente differente rispetto a quanto si fa normalmente nel caso dei reticoli di spin, in cui il valore in ciascun sito può essere -1 o 1. Evidentemente il passaggio da una scelta all'altra muta semplicemente l'espressione dell'hamiltoniana, ma non coinvolge la struttura di base. Con tale scelta in questo caso, l'hamiltoniana assume tuttavia un'espressione più comoda ed è questo il motivo per cui facciamo tale scelta.

[2] Poiché $\sigma = 0, 1$, per qualunque n:

$$\sigma^n = \sigma$$

Quindi nell'espressione dell'hamiltoniana possiamo limitarci ai termini che contengono termini al massimo di primo grado nel campo in ciascun sito.

tutte le configurazioni possibili. In tale contesto viene generalmente introdotta la temperatura che parametrizza la distribuzione sulle configurazioni. In questo caso tale definizione di temperatura si rivela sostanzialmente priva di significato. I valori dei coefficienti nell'hamiltoniana verranno determinati affinché siano riprodotte le corrette probabilità di transizione, e risulteranno quindi funzioni di p. Per il momento, evitiamo d'introdurre la temperatura nella forma solita: come già detto una nozione di temperatura non del tutto analoga a quella canonica, sarà ottenibile ricollegandosi proprio al parametro p. Poniamo perciò:

$$Z \equiv \sum_{\{\sigma\}} e^{-\mathbf{H}} \tag{4.3}$$

La nostra idea è di determinare i coefficienti di H proprio in modo tale che i pesi statistici in tale somma corrispondano alle effettive probabilità corrispondenti alle varie configurazioni. Ciò si traduce nella scrittura seguente:

$$P(\sigma_2 = k|\sigma_1 = h) = \sum_{\substack{\sigma_1 = h \\ \sigma_2 = k}} e^{-\mathbf{H}}$$

Da tale scrittura è facile determinare i coefficienti a, b, c, d. Infatti è sufficiente imporre:

$$P(\sigma_2 = 0|\sigma_1 = 0) = 1 = e^{-c} \quad \Rightarrow \quad c = 0$$

$$P(\sigma_2 = 1|\sigma_1 = 0) = 0 = e^{-b}$$

$$P(\sigma_2 = 0|\sigma_1 = 1) = 1 - p = e^{-a}$$

$$P(\sigma_2 = 1|\sigma_1 = 1) = p = e^{-a-b-d}$$

da cui determiniamo:

$$c = 0 \quad b = -\log\epsilon \quad a = -\log(1-p) \quad d = \log\left(\frac{(1-p)\epsilon}{p}\right) \tag{4.4}$$

ove si è introdotto il parametro ϵ che viene mandato a 0 nel momento opportuno e che fa si che abbiano peso nullo le configurazioni probabilisticamente vietate. Si noti che l'hamiltoniana così scritta non è simmetrica nei due campi in punti diversi da cui dipende; ciò è dovuto al fatto che essi corrispondono rispettivamente al sito del padre e a quello del figlio. Otteniamo l'hamiltoniana completa sommando H su tutti i legami orientati. Perciò detto:

$$L_n = \{(\sigma_i^{(k)}, \sigma_j^{(k+1)}) | k \geq n,\ 0 \leq i \leq 2^k - 1,\ j \in \{2i, 2i+1\}\}$$

l'insieme delle coppie tra cui è presente un legame da padre a figlio a partire dalla generazione n-esima, si potrà scrivere l'hamiltoniana complessiva come:

$$\mathbf{H}(\sigma) = \sum_{(\sigma_1,\sigma_2)\in L_0} H(\sigma_1, \sigma_2) \tag{4.5}$$

Con queste considerazioni, può essere riordinata la complessa scrittura di tale hamiltoniana: ogni sito compare tre volte nell'hamiltoniana completa, una volta come seconda variabile quando è un figlio e due volte come prima variabile quando è un padre; cioè per il generico sito σ avremo i tre termini:

$$H(\sigma_p, \sigma) + H(\sigma, \sigma_{f1}) + H(\sigma, \sigma_{f2})$$

Ciò permette di riscrivere l'hamiltoniana complessiva come:

$$\mathbf{H}(\sigma) = 2a + d\sigma_0^{(1)} + d\sigma_1^{(1)} + (2a+b) \sum_{\sigma\ \in V_1} \sigma + d \sum_{(\sigma_1,\sigma_2)\in L_1} \sigma_1\sigma_2 \tag{4.6}$$

ove

$$V_k = \{\sigma_j^{(n)} | n \geq k\}$$

è l'insieme dei vertici a partire dalla generazione k-esima. Vedremo nella prossima sezione che è possibile calcolare esplicitamente Z attraverso una relazione di punto fisso, che deriva dalla forma particolare che ha il reticolo. Per farlo tuttavia sarà necessario rivedere meglio il limite che sottintende la costruzione del reticolo.

4.2.1 Calcolo di Z

Vogliamo calcolare la seguente espressione:

$$Z(p) = \lim_{\epsilon\to 0} \sum_{\{\sigma\}} e^{-\mathbf{H}(\sigma)} \tag{4.7}$$

Per farlo osserviamo che è utile spezzare la somma al secondo membro nel seguente modo:

$$\sum_{\{\sigma\}} = \sum_{\{00\}} + \sum_{\{01\}} + \sum_{\{10\}} + \sum_{\{11\}}$$

dove abbiamo usato una notazione abbreviata:

$$\sum_{\{\sigma,\sigma_0^{(1)}=k,\sigma_1^{(1)}=h\}} = \sum_{\{kh\}}$$

in tal modo, vengono messe in mostra le proprietà di ricorsività dell'hamiltoniana e del reticolo stesso. Sostituiamo questa espressione nel calcolo di Z e calcoliamo i vari termini separatamente. Osserviamo che nel limite $\epsilon \to 0$ danno contributo nullo tutte le configurazioni in cui non compare la somma $b + d$ (in cui la ϵ s'elide) bensì i due termini separatamente. Perciò per il primo termine danno contributo nullo tutte le configurazioni eccetto quella che ha tutti i $\sigma = 0$ per la quale si ha:

$$e^{-\mathbf{H}(\sigma^*=1,\sigma=0)} = (1-p)^2$$

Il secondo ed il terzo termine sono analoghi e danno:

$$e^{-\mathbf{H}(\sigma_0^{(1)}=1,\sigma_1^{(1)}=0)} =$$

$$= \exp\left(-2a - \delta - 2a - (2a+b)\sum_{\sigma\in V_2}\sigma - d\sigma_0^{(1)} - d\sigma_1^{(1)} - d\sum_{(\sigma_1,\sigma_2)\in L_2}\sigma_1\sigma_2\right) =$$

$$= \exp(-2a-\delta)Z = (1-p)pZ$$

ed infine l'ultimo termine con un ragionamento analogo conduce a:

$$e^{-\mathbf{H}(\sigma_0^{(1)}=1,\sigma_1^{(1)}=1)} = \exp(-2a-2\delta)Z^2 = p^2 Z$$

In totale tutto ciò conduce all'equazione:

$$Z = (1-p)^2 + 2p(1-p)Z + p^2 Z^2 \tag{4.8}$$

Essa ha due soluzioni: $\mathcal{Z}_1 = 1$ e $\mathcal{Z}_2 = \left(\frac{1-p}{p}\right)^2$. A quale di queste due corrisponde il reale valore di Z? Per rispondere a questa domanda bisogna indagare la somma che definisce Z. Infatti nel modo in cui l'abbiamo definita essa concerne una somma su infiniti termini, corrispondenti agli infiniti siti del reticolo. Poiché viene anche effettuato il limite su ϵ, sono presenti due limiti distinti ed a seconda del modo in cui vengono effettuati si ottengono i differenti valori di Z. Per addentrarci nel problema consideriamo il reticolo troncato alla generazione T:

$$R_T = \{(\sigma_i^{(k)}, \sigma_j^{(k+1)}) | 0 \le k \le T,\ 0 \le i \le 2^k - 1,\ j \in \{2i, 2i+1\}\}$$

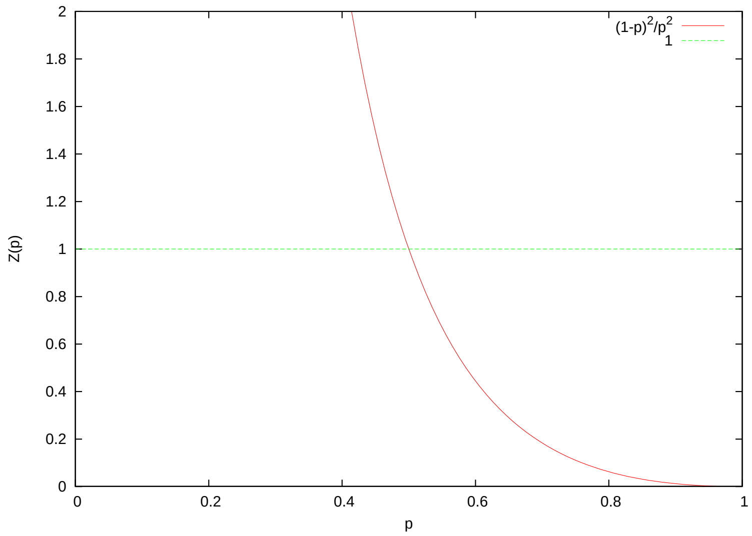


Figura 4.1: I due valori possibili per la funzione di partizione in base all'equazione di punto fisso (4.8)

Possiamo quindi costruire su di esso un'hamiltoniana troncata H_T nel solito modo:

$$H_T(\sigma) = \sum_{(\sigma_1,\sigma_2)\in R_T} H(\sigma_1, \sigma_2)$$

dove ora la somma concerne un numero finito di pezzi. Maneggiando l'hamiltoniana come nella (4.6) dobbiamo tenere conto che i termini alla generazione T-esima non possono avere figli, poiché questi ultimi non sono inclusi nel reticolo stesso:

$$H_T(\sigma) = 2a+d\sigma_0^{(1)}+d\sigma_1^{(1)}+(2a+b)\sum_{\sigma\ \in V_{1,T-1}} \sigma+d\sum_{(\sigma_1,\sigma_2)\in L_{1,T}} \sigma_1\sigma_2+b\sum_{j=0}^{2^T-1} \sigma_j^{(T)} \tag{4.9}$$

che presenta rispetto alla (4.6) un termine di bordo aggiuntivo. Un modo per eliminarlo è il seguente: possiamo considerare le sole configurazioni del sistema per cui alla T-esima generazione non sono presenti più individui:

$$\{\sigma\}_T = \{\sigma_j^{(i)} = \pm 1, 0 \le i < T, 0 \le j \le 2^i - 1; \sigma^{(t)} = 0\}$$

dove stiamo definendo:

$$V_{k,t} = \{\sigma_j^{(i)} | k \le i \le t\} \qquad L_{k,t} = L_k \cap (V_{k,t} \times V_{k,t})$$

Possiamo ora porre:

$$Z = \lim_{T\to\infty} \lim_{\epsilon\to 0} \sum_{\{\sigma_T\}} e^{-H_T(\sigma)} = \lim_{T\to\infty} Z_T$$

Spezzando nuovamente la somma in Z_T, otteniamo questa volta la relazione ricorsiva:

$$Z_T = (1-p)^2 + 2p(1-p)Z_{T-1} + p^2 Z_{T-1}^2 \tag{4.10}$$

ove prendendo il limite per $T \to \infty$ di ambo i membri vediamo che chiaramente Z così definita soddisfa ancora la (4.8). Inoltre:

$$Z_1 = (1-p)^2$$

Perciò se chiamiamo $h(z) = (1-p)^2 + 2p(1-p)z + p^2z^2$ si avrà:

$$Z = \lim_{n\to\infty} h^{(n)}(1-p)$$

In analogia con la (2.18), osserviamo che Z così ottenuta è proprio la probabilità d'estinzione, dato che per $p \in (0,1)$, tale limite converge alla più piccola delle soluzioni della (4.8). Infatti la condizione al bordo che abbiamo inserito seleziona le configurazioni per cui, da una certa generazione in poi, non ci sono più individui; quindi per la convessità di h:

$$Z = \min\left\{1, \frac{(1-p)^2}{p^2}\right\} = q$$

Cosa accade con altre scelte delle condizioni al bordo? Se non mettiamo alcun vincolo su ciò che accade alla generazione T-esima si ha che:

$$Z_1 = (1-p)^2 + 2p(1-p) + p^2 = (1-p+p)^2 = 1$$

e resta ancora valida la (4.10), per cui:

$$Z = 1$$

che è consistente col fatto che, ammettendo tutte le configurazioni possibili (incluse quelle in cui non v'è estinzione), Z eguaglia la probabilità totale. Bisogna tenere presente che quest'ultima condizione al bordo, aggiunge un termine divergente all'hamiltoniana limite che è quindi mal definita; dovremo quindi fare una certa attenzione nel maneggiarla.

È interessante notare che al di sotto del valore critico $p_c = \frac{1}{2}$, il valore limite di Z è sempre 1 e quindi questa scelta non influisce sulle nostre considerazioni. Al di sopra, la situazione, come vedremo, è più delicata come vedremo. Val la pena d'osservare a questo punto, che come s'era detto, un ruolo simile a quello normalmente svolto dalla temperatura è qui interpretato dal parametro p o meglio dal parametro $a = -\log(1-p)$. Infatti se $p \to 0$, anche $a \to 0$ e il reticolo è completamente spento, cioè si concentra nello stato che minimizza l'hamiltoniana, come accade tipicamente quando la temperatura va a 0. Se $p \to 1$, allora $a \to \infty$ ed il reticolo tende ad essere tutto acceso, come accade per temperature elevate. Inoltre il valore $p_c = \frac{1}{2}$ distingue tra due differenti fasi del sistema: infatti al di sotto di tale valore sono ammissibili solo configurazioni finite. Può quindi ritenersi un punto critico anche in un senso "termodinamico".

4.2.2 Il numero di individui complessivo

Siamo giunti alla forma esplicita dell'hamiltoniana e del reticolo ed abbiamo trovato il valore esplicito della funzione di partizione, per la quale si è notata la dipendenza esplicita dalle condizioni al bordo. In questa sezione metteremo in luce alcune caratteristiche del numero di siti attivi. Esso è una variabile aleatoria che indicheremo con N. Indicheremo invece con N_t il numero di siti attivi alla generazione t-esima. Sarà possibile collegare i momenti della distribuzione di N alle derivate della funzione di partizione, prenderemo in considerazione la (4.6) senza esplicitare la dipendenza da p di a e $\delta = b + d$. Valgono ovviamente le considerazioni fatte nella sezione precedente riguardanti il fatto che l'hamiltoniana limite non è definita rigorosamente, ma è comunque definita Z con le opportune condizioni al bordo. Vedremo quando questo fatto risulterà evidente. Ragionando come prima e quindi spezzando la somma che coinvolge Z s'avrà questa volta l'equazione:

$$Z = e^{-2a} + 2e^{-2a-\delta}Z + e^{-2a-2\delta}Z^2 \tag{4.11}$$

Possiamo sfruttare la (4.11) nel seguente modo. Per prima cosa:

Definizione 4.1. *Poniamo:*

$$\mathcal{N}_t = \langle N_t \rangle$$

cioè il numero medio d'individui alla generazione t. Inoltre:

$$\mathcal{N} = \sum_t \langle N_t \rangle$$

cioè il numero medio complessivo d'individui.

Possiamo allora utilizzare la (4.7) e la (4.6) per scrivere:

$$\mathcal{N} = \sum_{\{\sigma\}} \sum_{j,k} \sigma_j^{(k)} e^{-\mathbf{H}} = - \left. \frac{\partial Z}{2 \partial a} \right|_{a=-\log(1-p)}$$

Perciò, moltiplichiamo la (4.11) per e^{2a} e poi derivando la (4.11) rispetto ad a otteniamo:

$$2e^{2a} Z + e^{2a} \frac{\partial Z}{\partial a} = 2e^{-\delta} \frac{\partial Z}{\partial a} + e^{-2\delta} 2Z \frac{\partial Z}{\partial a}$$

Tenendo conto che per $a = -\log(1-p)$ e $\delta = -\log\left(\frac{p}{1-p}\right)$ si ha $Z = 1$, otteniamo:

$$\mathcal{N} = \frac{1}{1-2p}$$

per $p < \frac{1}{2}$ questa può essere vista come:

$$\sum_{k=0}^{\infty} (2p)^k$$

ove $2p$ è il numero medio di figli di un individuo. Come avevamo detto in precedenza, al di sopra del valore critico $p = \frac{1}{2}$, hanno probabilità non nulla le configurazioni con un numero infinito di siti attivi. Se vogliamo tenere conto delle configurazioni infinite, l'hamiltoniana limite non è più ben definita ed il valore medio è in realtà infinito. Vedremo nel seguito come questo calcolo può essere effettuato correttamente con un limite sui reticoli finiti.

4.2.3 Momenti successivi e proprietà asintotiche

Abbiamo visto come è possibile calcolare nel formalismo della meccanica statistica, quindi partendo dalla funzione di partizione, il valore medio del numero d'individui. Vedremo in che modo è possibile ottenere altre informazioni riguardo la distribuzione di N. In particolare posto:

$$P_n = P(N = n)$$

C'interessa l'andamento di P_n per n grande. Bisogna osservare che le proprietà asintotiche dei P_n emergono nei punti di non analiticità di Z. Infatti nei punti in cui tutte le derivate sono finite, il valore di Z risente dei valori dei P_n al finito. Invece nei punti in cui qualche derivata diverge, il modo in cui tende verso l'infinito ci dà delle informazioni sull'andamento dei P_n. Tratteremo per cominciare il caso subcritico ($p < \frac{1}{2}$) per il quale sono finiti i momenti. Poi seguiremo un approccio perturbativo che, come vedremo nel seguito, permette d'ottenere informazioni più generali.

4.2.4 Caso subcritico

In questo caso la scelta delle condizioni al bordo è ininfluente, dato che comunque hanno peso non nullo solo le configurazioni con un numero finito d'individui. Perciò possiamo calcolare le derivate di Z rispetto ad a. Questo permette di collegare i momenti successivi di N alla funzione di partizione. Si avrà:

$$M_s(N) = \sum_{\{\sigma\}} \left(\sum_{j,k} \sigma_j^{(k)} \right)^s e^{-\mathbf{H}}$$

ove con $M_s(N)$ si indicherà il momento s-esimo della variabile aleatoria N. Sfruttando la (4.7) possiamo quindi scrivere:

$$M_s = -\left(\frac{1}{2}\right)^s \frac{\partial^s Z}{(\partial a)^s}\bigg|_{a=-\log(1-p)} \tag{4.12}$$

Questa espressione richiama la funzione generatrice dei momenti. Ricordiamone la definizione:

Definizione 4.2. *Data una variabile aleatoria X definiamo la funzione generatrice dei momenti:*

$$M_X(t) = \mathbb{E}(e^{tX}) = \sum_{k=0}^{\infty} \frac{t^k}{k!}\mathbb{E}(X^k) = \sum_{k=0}^{\infty} \frac{t^k M_k(X)}{k!}$$

Nel nostro caso sarà:

$$M_N(t) = \sum_{n=0}^{\infty} P_n e^{tn} \tag{4.13}$$

ove con P_n indichiamo la probabilità che nel reticolo siano attivi n siti. Possiamo vedere ora in che modo le due sono collegate. S'ottiene subito che:

$$M_k(X) = \left. \frac{\partial^k M_X(t)}{\partial t^k} \right|_{t=0}$$

Confrontando quest'ultima equazione con la (4.12) otteniamo immediatamente che:

$$M_N(t) = Z\left(-\frac{t}{2} - \log(1-p)\right)$$

ove si sta indicando esclusivamente per Z la dipendenza dal parametro a. Il valore di Z lo otteniamo da (4.11) che ci fornisce due valori:

$$\mathcal{Z}_{1,2} = \frac{(e^{2a+2\delta} - 2e^{\delta}) \pm \sqrt{e^{4a+4\delta} - 4e^{2a+3\delta}}}{2}$$

Essendo al di sotto del valore critico, la scelta corretta è la più piccola delle due. Sostituendo:

$$\delta = \log\left(\frac{1-p}{p}\right) \qquad a = -\log(1-p) - \frac{t}{2} \tag{4.14}$$

otteniamo:

$$Z = \frac{e^{-t} - 2p + 2p^2 - \sqrt{e^{-2t} - 4e^{-t}(p-p^2)}}{2p^2}$$

Vediamo subito che non è analitica per:

$$t_0 = -2\log 2 - \log(p(1-p))$$

Sviluppando intorno a t_0:

$$Z = \left(\frac{1-p}{p}\right)[1 + 2(t_0 - t) - 2\sqrt{t_0 - t}]$$

La parte non analitica va perciò come: $\sqrt{t_0 - t}$. Usando la (4.13) possiamo ottenere l'andamento di P_n per $n \to \infty$ con un'antitrasformata di Laplace:

$$P_n \approx e^{-t_0 n} n^{-\frac{3}{2}} = (4(1-p)p)^n n^{-\frac{3}{2}} \tag{4.15}$$

Tale espressione è in realtà valida anche per valori diversi di p. Vediamo in che modo si può affrontare in casi più generali.

4.2.5 Calcolo perturbativo

Per $p = \frac{1}{2}$ ci aspettiamo un andamento a potenza essendo in un punto critico. Ciò è consistente con quanto ottenuto nella (4.15) infatti per $p \to \frac{1}{2}$, il termine esponenziale scompare. Per verificare questo andamento, prendiamo in considerazione il reticolo troncato alla T-esima generazione R_T. Come abbiamo visto è possibile ottenere una relazione tra le funzioni di partizione per due reticoli successivi. Varrà, usando la (4.9):

$$Z_0 = 1$$

$$Z_T = e^{-2a} + 2e^{-2a-\delta} Z_{T-1} + e^{-2a-2\delta} Z_{T-1}^2$$

Sostituendo con le (4.14):

$$Z_T = e^t[(1-p)^2 + 2(1-p)pZ_{T-1} + p^2 Z_{T-1}^2] = e^t h(Z_{T-1}) \tag{4.16}$$

Supponiamo che t sia piccolo e limitiamoci al primo ordine. Osservando che, posto:

$$Z_{T-1} = 1 + z_{T-1}t$$

s'ottiene:

$$Z_T = 1 + (2z_{T-1}p + 1)t = 1 + z_T t$$

Che conduce ad un equazione ricorsiva per z_T, termine lineare in t di Z_T:

$$z_T = 2z_{T-1}p + 1$$

che è risolta da:

$$z_T = \sum_{k=0}^{T-1} (2p)^k$$

Questo corregge il risultato errato per $p \geq \frac{1}{2}$ della sezione 4.2.2 e permette una via alternativa per individuare il valore critico di $p = \frac{1}{2}$. Ci aspettiamo un andamento a potenza per tale valore: $P_n \approx n^{-\alpha}$ che in $Z(t)$ corrisponde ad un termine che va come $t^{\alpha-1}$ come evidente mediante la trasformata di Laplace. Il valore di α può essere trovato andando al secondo ordine in t. Infatti, benché le derivate di $Z(t)$ siano divergenti in 0, se è corretta l'ipotesi di legge a potenza che abbiamo fatto per $Z(t)$, deve valere che:

$$0 < \lim_{t \to 0} \frac{Z'(t)^{\frac{\alpha-3}{\alpha-2}}}{Z''(t)} < \infty$$

Con questo sistema, posto:

$$Z_T = 1 + z_T t + v_T t^2$$

ed essendo ν l'esponente per cui:

$$0 < \lim_{T\to\infty} \frac{z_T^\nu}{v_T} < \infty$$

si avrà:

$$\alpha = \frac{2\nu - 3}{\nu - 1}$$

Quindi dalla (4.16) s'ottiene per $p = \frac{1}{2}$:

$$Z_T = \frac{e^t}{4}(Z_{T-1} + 1)^2 = 1 + (z_{T-1} + 1)t + \frac{(z_{T-1}^2 + 4(v_{T-1} + z_{T-1}) + 2)t^2}{4}$$

da cui otteniamo le relazioni ricorsive:

$$\begin{cases} z_T = z_{T-1} + 1 \\ v_T = \frac{(z_{T-1}^2 + 4(v_{T-1} + z_{T-1}) + 2)}{4} \end{cases}$$

La prima si risolve immediatamente: $z_T = T$. Inserendola nella seconda:

$$v_T = \frac{T(T+1)(2T+1)}{24} + \frac{T(T+1)}{4} - \frac{T}{4} = \frac{T^3}{12} + o\left(T^3\right)$$

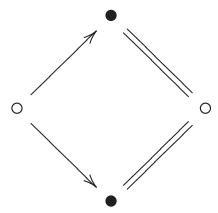
Questo fornisce immediatamente: $\nu = 3$ e quindi $\alpha = \frac{3}{2}$ come ci aspettavamo. Vedremo nel seguito come questa procedura può essere generalizzata ad una distribuzione qualunque della prole di un singolo individuo.

4.3 Distinguere le generazioni

Nella precedente sezione abbiamo visto come questo formalismo ci abbia dato un metodo diretto per ottenere delle informazioni sul numero totale degli individui. Tuttavia può essere interessante riuscire ad ottenere alcune proprietà del numero d'individui ad una particolare generazione. Per far ciò conviene introdurre alcune modifiche al formalismo precedente. Per semplicità può essere utile osservare che nel caso analizzato, e cioè in cui ad ogni sito è associata la medesima probabilità (vedremo come si affronta il caso generale nella prossima sezione) con una lieve modifica nella costruzione del reticolo, si può ottenere una semplificazione alla relazione ricorsiva soddisfatta da Z. Partiamo da

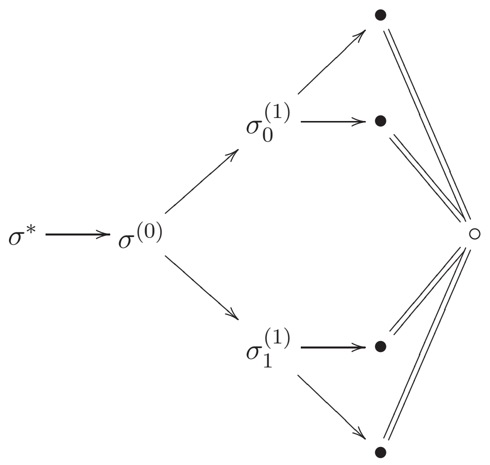
$$\circ \longrightarrow \bullet = \circ \tag{4.17}$$

Ad ogni passo sostituiremo la parte doppia con il seguente:



(4.18)

E dopo due iterazioni:



Analogamente a prima porremo: $\sigma^* = 1$ ed introdurremo un'hamiltoniana complessiva mediante la somma su tutti i link della (4.2). Tuttavia questa volta, per differenziare le varie generazioni, faremo la modifica seguente nella definizione dell'hamiltoniana completa:

$$\mathbf{H} = \sum_g \sum_{j=0}^{2^g-1} H(\sigma_j^{(g)}, \sigma_{2j}^{(g+1)}, a_g) + H(\sigma_j^{(g)}, \sigma_{2j+1}^{(g+1)}, a_g)$$

dove s'è lasciata esplicita la dipendenza di H dal parametro a e s'è quindi opportunamente sommato su parametri diversi per le varie generazioni. È chiaro che alla fine tutti gli a_g saranno presi uguali a $-\log(1-p)$. Tuttavia nei calcoli intermedi conviene mantenerli diversi. Indicata con:

$$\{a_0\} = \{a_0, a_1, \ldots\}$$

potremo scrivere:

$$Z(\{a_0\}) = \sum_{\{\sigma\}} e^{-\mathbf{H}(\sigma,\{a_0\})}$$

Con la costruzione effettuata per il reticolo, è comunque possibile ottenere una relazione di ricorsione tra le funzioni di partizione. In particolare, con la scelta effettuata abbiamo eliminato il termine di primo grado in Z che era presente nella (4.11)[3]. Essa assume la forma:

$$Z(\{a_i\}) = e^{-a_i} + e^{-a_i-\delta} Z^2(\{a_{i+1}\})$$

Ponendo:

$$f(a,x) \equiv e^{-a} + e^{-a-\delta} x^2$$

Otteniamo:

$$Z(\{a_i\}) = f(a_i, Z(\{a_{i+1}\})) \tag{4.19}$$

Sappiamo poi che:

$$\mathcal{N}_t = -\frac{\partial Z(\{a_0\})}{\partial a_t}$$

Quindi siamo interessati al calcolo delle derivate di Z rispetto ai diversi a_i. Osserviamo allora che:

$$\frac{\partial Z(\{a_i\})}{\partial a_j} = \begin{cases} \frac{\partial f}{\partial x} \frac{\partial Z(\{a_{i+1}\})}{\partial a_j} & i < j \\ 0 & i > j \\ \frac{\partial f}{\partial a} & i = j \end{cases}$$

[3]La modifica che abbiamo fatto al reticolo si traduce in una modifica nella relazione di ricorsione per Z. Infatti la (4.8) si può scrivere come: $Z = ((1-p) + pZ)^2$. Il nuovo reticolo equivale a considerare:

$$Z' = pZ + 1 - p$$

da cui s'ottiene la nuova relazione:

$$Z' = (1-p) + pZ'^2$$

Risostituendo le (4.4), ponendo tutti gli $a_i = a$, otteniamo:

$$\mathcal{N}_t = -\frac{\partial Z(\{a_0\})}{\partial a_t} = \left(\frac{\partial f(a,1)}{\partial x}\right)^t \frac{\partial f}{\partial a}$$

dove si ha:

$$\frac{\partial f(-\log(1-p),1)}{\partial a} = 1$$
$$\frac{\partial f(-\log(1-p),1)}{\partial x} = 2p$$

Perciò come ci aspettavamo si ha:

$$\mathcal{N}_t = (2p)^t$$

4.4 Caso generale

4.4.1 Finiti figli

Abbiamo visto cosa si può dire nel caso semplice in cui ogni individuo può avere al massimo due figli, entrambi con probabilità p. Vogliamo ora trovare una generalizzazione al caso in cui la distribuzione dei figli di un individuo sia data da dei generici $\rho_n = P(X_1 = n|X_0 = 1)$. Cominciamo dapprima con il seguente caso:

Definizione 4.3. *Chiamiamo fattorizzabile la distribuzione di probabilità* ρ_n *se:*

$$\exists M \quad t.c. \quad \forall k > M \quad \rho_k = 0 \tag{4.20}$$

ed inoltre il polinomio:

$$P(x) = \sum_k \rho_k x^k \tag{4.21}$$

ha solo radici reali.

Supporremo nel seguito che M indichi il minimo valore per cui la (4.20) vale. Indicheremo con x_k l'opposto delle sue radici e si avrà $x_k \geq 0$ perché chiaramente si tratta di un polinomio a coefficienti positivi. Allora:

$$P(x) = \rho_N \prod_{k=0}^{M} (x + x_k) \tag{4.22}$$

Ora ponendo:

$$x_k = \frac{1-p_k}{p_k} \quad \Leftrightarrow \quad p_k = \frac{1}{x_k+1}$$

ove $0 < p_k \leq 1$, si può riscrivere:

$$P(x) = \left[\rho_M \prod_{m=0}^{M}(1+x_k)\right] \prod_{k=0}^{M}((1-p_k)+p_k x)$$

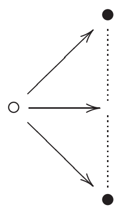
Ma usando la (4.22):

$$1 = \sum_k \rho_k = P(1) = \rho_M \prod_{m=0}^{M}(1+x_k)$$

Quindi possiamo scrivere:

$$P(x) = \prod_{k=0}^{M}((1-p_k)+p_k x) \tag{4.23}$$

Dato che ora il numero massimo di figli è M il reticolo dovrà essere modificato nel seguente modo: ad ogni passo sostituiremo $\bullet$ con



dove i puntini rappresentano le M frecce che connettono $\circ$ ai $\bullet$. Dobbiamo ora differenziare i vari punti alla stessa generazione. In caso contrario otterremmo nuovamente una distribuzione binomiale. Prendiamo in considerazione allora un'hamiltoniana a due punti diversa per ciascun p_k. Perciò chiameremo $H_k(\sigma_1, \sigma_2)$ l'hamiltoniana della forma (4.2) in cui il valore dei parametri è determinato dalla (4.4) con $p = p_k$. Possiamo quindi scrivere l'hamiltoniana complessiva come fatto nella (4.5) introducendo, per $m \in \{1, \ldots, M\}$:

$$L_n^m = \left\{(\sigma_i^{(k)}, \sigma_{m-1+Mi}^{(k+1)}) | k \geq n,\ 0 \leq i \leq M^k - 1\right\}$$

$$V_k^m = \left\{\sigma_j^{(n)} | n \geq k\,,\ j \equiv m-1(\mod M)\right\}$$

e perciò:

$$\mathbf{H} = \sum_{m=1}^{M} \sum_{(\sigma_1,\sigma_2)\in L_0^m} H_m(\sigma_1,\sigma_2)$$

Possiamo riordinare come nella (4.6), tenendo conto questa volta che per ciascun individuo abbiamo M contributi diversi ogniqualvolta è padre:

$$\mathbf{H}(\sigma) = \sum_m \left[a_m + d_m\sigma_m^{(1)} + \left(a_m + \frac{b}{M}\right) \sum_{\sigma\ \in V_1} \sigma + d_m \sum_{(\sigma_1,\sigma_2)\in L_1^m} \sigma_1\sigma_2 \right] + B = \\ = \sum_m H^m + B \quad (4.24)$$

ove abbiamo inserito una B ad indicare i termini di bordo all'infinito. Come visto nella 4.2.1, per una trattazione corretta anziché prendere l'hamiltoniana totale, dovremmo considerare le hamiltoniane del reticolo troncato. Con tale approccio siamo giunti, nel caso semplificato trattato in precedenza, ad una relazione di ricorsione per le funzioni di partizione. Vogliamo ripercorrere i medesimi passi in questo caso generale. Di nuovo indicheremo con un pedice T le grandezze relative al reticolo troncato alla T-esima generazione. Spezziamo la somma nella (4.7) nella somma sulla prima generazione e in quella sulla restante parte del reticolo:

$$Z_T = \lim_{\epsilon\to 0} \sum_{\sigma^{(1)}} \sum_{\sigma^{(k)},k>1} e^{-\mathbf{H}_T} \quad (4.25)$$

La prima somma è su tutti i valori della prima generazione; ogni sito può assumere due valori: $0, 1$. Fissiamo una configurazione $\{\sigma\}$. Chiamiamo $Y \subset \{1, \ldots, M\}$ l'insieme degli indici dei siti della prima generazione attivi e y il loro numero. Possiamo riscrivere la $\mathbf{H}$ per tale $\{\sigma\}$:

$$\mathbf{H}_T(\{\sigma\}) = A + \sum_{m\in A} \delta_m + \\ + \sum_{j\in Y}\sum_m \left[a_m + d_m\sigma_{m+jM}^{(2)} + \left(a_m + \frac{b}{M}\right) \sum_{\sigma\ \in V_2} \sigma + d_m \sum_{(\sigma_1,\sigma_2)\in L_2^m} \sigma_1\sigma_2 \right]$$

ove $\delta_m = d_m + b$, $A = \sum_{m=1}^{M} a_m$ e si sono trascurati i termini che danno un contributo nullo a Z nel limite $\epsilon \to 0$. Immettendo questa espressione nella

(4.25) e sfruttando la seconda sommatoria otteniamo:

$$Z_T = e^{-A} \sum_{\sigma^{(1)}} Z_{T-1}^{y} \prod_{m \in Y} e^{-\delta_m} \tag{4.26}$$

Ricordando che:

$$a_m = -\log(1 - p_m) \qquad \delta_m = \log\left(\frac{1 - p_m}{p_m}\right)$$

$$\begin{aligned} Z_T(p_1, \dots, p_M) = \sum_{\sigma^{(1)}} Z_{T-1} \prod_{m \in Y} p_m \prod_{n \in Y^c} (1 - p_n) = \\ = \prod_m (1 - p_m + p_m Z_{T-1}) = P(Z_{T-1}) \end{aligned} \tag{4.27}$$

che è analoga alla (4.10).

Possiamo nuovamente determinare alcune informazioni riguardo il numero totale d'individui. Dalla hamiltoniana nella (4.24) deduciamo subito che:

$$\mathcal{N} = -\lim_{T \to \infty} \sum_m \left.\frac{\partial Z_T(a_1, \dots, a_M)}{\partial a_m}\right|_{a_m = -\log(1-p_m)} = -\lim_{T \to \infty} \left.\frac{\partial Z_T(A)}{\partial A}\right|_{A = -\log(\rho_0)}$$

dove si è usata la (4.23) con $x = 0$ per ottenere:

$$\rho_0 = P(0) = \prod_m (1 - p_m)$$

È ovvia la generalizzazione ai momenti successivi:

$$M_s = -\lim_{T \to \infty} \left.\frac{\partial^s Z_T(A)}{\partial A^s}\right|_{A = -\log(\rho_0)}$$

Come nella 4.2.4 nel dominio in cui Z_T converge con le sue derivate ad una Z:

$$M_N(t) = Z(-t - \log(\rho_0))$$

È utile eseguire la sostituzione $A = -t - \log(\rho_0)$ nella (4.26):

$$Z_T = e^t \rho_0 \sum_{\sigma^{(1)}} Z_{T-1}^{y} \prod_{m \in Y} \frac{p_m}{1 - p_m} = e^t P(Z_{T-1}) \tag{4.28}$$

che generalizza la (4.16). Con la notazione della 4.2.5:

$$Z_T = 1 + z_T t = e^t P(1 + z_{T-1} t) = 1 + (1 + z_{T-1} P'(1)) t$$

da cui:

$$z_T = \sum_{k=0}^{T-1} (P'(1))^k = \sum_{k=0}^{T-1} \left(\sum_m p_m \right)^k$$

come ci aspettavamo, il valore critico si ha per $P'(1) = 1$, cioè quando il numero medio di figli di un individuo è 1. Sviluppiamo al secondo ordine in t in corrispondenza di tale valore:

$$\begin{aligned} Z_T = 1 + z_T t + v_T t^2 = e^t P(1 + z_{T-1} t + v_{T-1} t^2) = \\ = 1 + (1 + z_{T-1})t + \frac{(1 + (z_{T-1} + v_{T-1}) + z_{T-1}^2 P''(1))}{2} \end{aligned}$$

indicando con $\zeta = P''(1)$, si ricavano le relazioni ricorsive:

$$\begin{cases} z_T = z_{T-1} + 1 \\ v_T = \frac{(z_{T-1}^2 \zeta + 2(v_{T-1} + z_{T-1}) + 1)}{2} \end{cases}$$

Perciò $z_T = T$ e:

$$v_T = \frac{T(T+1)(2T+1)\zeta}{12} + \frac{(1-\zeta)T(T+1)}{2} + \frac{T(\zeta - 1)}{2} = \frac{\zeta T^3}{6} + o\left(T^3\right)$$

che è consistente col caso semplificato per cui $\zeta = \frac{1}{2}$ e quindi anche in questo caso $\alpha = \frac{3}{2}$.

4.4.2 Radici complesse e funzioni analitiche

Abbiamo generalizzato il modello introdotto per il caso binomiale a due figli, ad un caso più generale e cioè quando la funzione caratteristica è un polinomio a radici reali. Abbiamo voluto distinguere il caso di polinomio generico per un motivo concettuale. In generale, un polinomio a coefficienti tutti positivi avrà radici reali negative, come nel caso già visto, ma può anche avere delle coppie di radici complesse coniugate. La (4.22) diventa:

$$P(x) = \rho_N \prod_{k=0}^{M_1} (x + x_k) \prod_{h=0}^{M_2} ((x - a_h)^2 + b_h^2)$$

In realtà, il motivo per cui avevamo avuto bisogno di tale fattorizzazione è che a noi interessava poter ottenere un reticolo di siti che assumono il valore 0 o 1 ed inserire un'hamiltoniana che contenesse l'interazione tra due punti. Tuttavia non è sempre detto che ciò sia fattibile; non è detto che un polinomio a coefficienti positivi sia completamente fattorizzabile in termini di primo grado. Nel nostro contesto, ciò significa che se vogliamo continuare a parlare di probabilità, dovremmo ammette che nel reticolo siano presenti siti che possono ammettere valori più grandi di 1. In particolare, una volta fattorizzato il polinomio nei polinomi di grado minimo a coefficienti tutti positivi, dovremmo costruire un reticolo tale che a ciascun vertice siano collegati tanti vertici quanti il numero di questi polinomi e ciascuno dovrebbe poter assumere un valore che va da 0 al grado del corrispondente fattore. Tutto ciò complica notevolmente il nostro approccio alla teoria, anche perché la forma (4.2) non è più la più generale, dovendo ora tenere conto di tutti i valori possibili per il campo nei vari punti. Ma per i nostri scopi non è necessario questo: anche se le p_k non sono numeri reali e tra 0 e 1, ma numeri complessi, sebbene non siano interpretabili come probabilità, possiamo comunque inserirle nel nostro modello. Infatti, in tutte le grandezze fisiche, si combineranno dando valori reali dato che in qualunque espressione entreranno combinazioni delle ρ_k che sono reali. Quindi possiamo comunque costruire un reticolo sfruttando tale fattorizzazione, tenendo presente che l'espressione risultante per la ricorsione sarà comunque nella forma (4.28) in cui sono presenti solo le ρ_k.

Tuttavia noi vogliamo arrivare a generalizzare le nostre considerazioni ad una distribuzione generica che ammetta un qualunque numero di figli. Questo conduce come abbiamo visto nella sezione 1.2 ad una funzione generatrice analitica almeno nell'intervallo $[0,1)$, mentre fino ad ora abbiamo considerato funzioni polinomiali. I risultati possono però essere generalizzati a tale classe di funzioni almeno nel caso in cui $f''(1) < \infty$. Come abbiamo visto, infatti, tale derivata appare nella (4.27). Possiamo allora sostituire le ρ_k (che in questo caso non saranno più definitivamente 0), con una successione approssimante ma finita:

$$\tilde{\rho}_k = \rho_k, \quad k \in 0, 1, \ldots, n \qquad \tilde{\rho}_{s+1} = 1 - \sum_{k=0}^{n} \rho_k \qquad \tilde{\rho}_m = 0, \quad m > n+1$$

Capitolo 5

Rinormalizzazione e modelli gerarchici

5.1 Contesto generale

Prima di entrare nei dettagli di alcuni modelli vogliamo considerare da una prospettiva generale il gruppo di rinormalizzazione (**RG**). L'analisi attraverso il RG è stata introdotta nel contesto della teoria dei campi e delle transizioni di fase, ma è uno strumento matematico, oltre che concettuale, molto generale che permette di decomporre il problema d'individuare il comportamento macroscopico di molti oggetti interagenti in problemi più semplici con un numero più piccolo di oggetti interagenti, le cui proprietà variano con la scala d'osservazione. È particolarmente proficuo nei fenomeni critici e più in generale in sistemi che dimostrano un'invarianza di scala. Tecnicamente, si basa sull'individuazione di una mappa tra la scala d'osservazione e la distanza $|T - T_c|$ dal punto critico. Per scala d'osservazione normalmente s'intende la scala fisica d'osservazione. Ad esempio, nel caso degli spin, essa corrisponderà alla dimensione del gruppo di spin che uno analizza all'interno del sistema. La comodità del RG si basa sul fatto che, in prossimità del punto critico, un sistema manifesta invarianza di scala e self-similarity. Cerchiamo di mostrarne in concreto il funzionamento. Prendiamo in considerazione l'energia libera F di un sistema di spin. Si tratta di una quantità comoda perché caratterizza l'organizzazione del sistema di spin e nel contempo può essere misurata sperimentalmente. L'assunzione del RG si traduce nell'assumere che esista un legame tra F a due differenti temperature (a diverse distanze dal punto critico). Posta $x = |T_c - T|$, ciò è espresso dalle trasformazioni:

$$x' = \phi(x)$$

$$F(x) = g(x) + \frac{1}{\mu} F(\phi(x)) \tag{5.1}$$

La funzione ϕ è normalmente detta **flusso del RG**. Supporremo nel seguito che $F(x = 0) = 0$ nel senso che stiamo ridefinendo $F(x)$ come:

$$F(x) = \tilde{F}(x + T_c) - \tilde{F}(T_c)$$

ove $\tilde{F}$ indica l'energia libera ordinaria, mentre μ è una costante che descrive come l'energia libera riscala mentre riscaliamo la distanza dal punto critico. La funzione $g(x)$ è la parte non singolare di $F(x)$; assumeremo comunque che $F(x)$ sia continua e $\phi(x)$ sia derivabile. Per punti critici s'intendono i punti in cui $F(x)$ diviene singolare, cioè una sua derivata diviene infinita. Si può ottenere una soluzione formale della (5.1). Poniamo:

$$F_0(x) \equiv g(x) \tag{5.2}$$

$$F_{n+1}(x) = g(x) + \frac{1}{\mu} F_n(\phi(x)) \tag{5.3}$$

Per induzione è facile vedere che:

$$F_n(x) = \sum_{i=0}^{n} \frac{1}{\mu^i} g\left[\phi^{(i)}(x)\right] \tag{5.4}$$

e che: $\lim_{n\to\infty} F_n(x) = F(x)$, nei punti in cui tale limite esiste[1]. La potenza del RG sta nel ricostruire la natura della singolarità critica dalla conoscenza della parte non-singolare $g(x)$ e del flusso $\phi(x)$ che descrive il cambiamento di scala. La connessione tra questo formalismo e i punti critici viene dal fatto che questi ultimi corrispondono a punti fissi instabili del flusso $\phi(x)$. La singolarità nel punto fisso x_c emerge se:

$$\lambda = \left.\frac{d\phi}{dx}\right|_{x_c=\phi(x_c)}$$

[1] Una trattazione matematica più accurata si può trovare in [9]

è maggiore di 1 cioè se la mappa ϕ è instabile per le iterazioni in corrispondenza di tale punto fisso. Infatti dalla (5.4) vediamo che la derivata k si esprime come una serie il cui termine n-esimo è $(\lambda^k/\mu)^n$, che può superare 1 se k è abbastanza grande visto che $\lambda > 1$, da cui ha origine l'andamento singolare. Il comportamento qualitativo di $F(x)$ in prossimità del punto critico può essere dedotto semplicemente dalla struttura di $\phi(x)$: se $x = 0$ denota un punto fisso instabile e $\phi(x) = \lambda x + \ldots$ è la trasformazione linearizzata corrispondente, in prossimità del punto critico si avrà:

$$F(x) \simeq x^\alpha \tag{5.5}$$

ove α è una soluzione di:

$$\frac{\lambda^\alpha}{\mu} = 1$$

che conduce a:

$$\alpha = \frac{\log \mu}{\log \lambda} \tag{5.6}$$

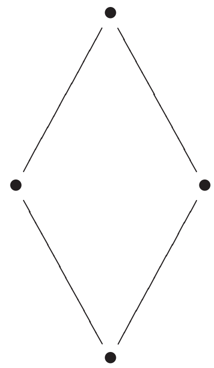
L'esponente e il comportamento di F sono quindi determinati dal modo in cui riscala l'energia libera μ e dalla velocità λ con cui ci allontaniamo dal punto critico al variare della scala d'osservazione.

5.2 Un esempio esplicito: spin su un reticolo gerarchico

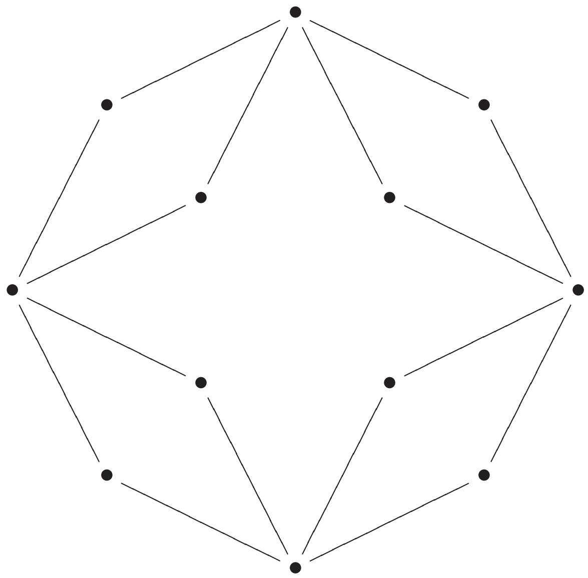
5.2.1 Le equazioni del gruppo di rinormalizzazione

Illustreremo qui un caso in cui l'applicazione del RG permette una soluzione esatta. Costruiamo il reticolo attraverso una regola iterativa: partiamo da un segmento con due siti agli estremi:

e ad ogni passo sostituiremo ciascun segmento con il seguente:



Dopo due iterazioni s'ottiene il seguente:



Il reticolo che s'ottiene è detto **reticolo a diamante**. Nel seguito chiameremo **grado di magnificazione** il numero di iterazioni n. Supponiamo che ad una data magnificazione si abbia la seguente energia d'interazione:

$$E = -J \sum_{ij} \delta(\sigma_i - \sigma_j)$$

ove la δ fa 1 solo se l'argomento è 0 ed è nulla altrimenti. Tutte le funzioni termodinamiche sono determinate una volta calcolata la funzione di partizione:

$$Z(\beta) = \sum_{\{\sigma\}} e^{-\beta E}$$

Tale calcolo è a priori molto complesso. Tuttavia, dato che tutti gli spin sono posti ai vertici di un reticolo frattale, vedremo che sarà possibile scrivere delle equazioni esatte per il RG e quindi risolvere il problema. Indichiamo con Z_n la funzione di partizione al grado di magnificazione n. Anziché effettuare la somma completa, sommeremo soltanto gli spin aggiunti all'ultimo passo e che interagiscono solo con un altro spin. Consideriamo un singolo diamante e indichiamo con $\sigma_{1,2}$ i due spin interni e con $s_{1,2}$ quelli esterni. Il contributo di tale diamante a Z è:

$$\exp\left(\beta J(\delta(\sigma_1 - s_1) + \delta(\sigma_2 - s_1) + \delta(\sigma_1 - s_2) + \delta(\sigma_2 - s_2)\right)$$

sommando su $\sigma_{1,2}$ e con un po' d'algebra s'ottiene:

$$(2K + Q - 2)^2 K'^{\delta(\sigma_2 - \sigma_1)} \tag{5.7}$$

ove s'è posto $K = \exp(\beta J)$, Q è il numero di valori distinti che può assumere ciascun sito e:

$$K' \equiv \left(\frac{K^2 + Q - 1}{2K + Q - 2}\right) = \phi(K) \tag{5.8}$$

Osserviamo che quando eseguiamo tale somma su tutti i diamanti, il fattore a moltiplicare nella (5.7) compare tante volte quanti sono i diamanti all'iterazione n-esima cioè 2^{2n+1}. Ciò conduce a:

$$Z_n(K) = (2K + Q - 2)^{2n+1} Z_{n-1}(\phi(K))$$

Possiamo ricavare un'equazione per l'energia libera per sito prendendone il logaritmo e dividendo per il numero di siti 4^{n+1}:

$$F_n(K) = \frac{1}{2}\log(2K + Q - 2) + \frac{1}{4}F_{n-1}(\phi(K))$$

che realizza in un esempio concreto la (5.3). Questo calcolo esplicito mostra chiaramente l'origine delle relazioni a scale diverse dell'energia libera: a scale diverse la funzione di partizione assume la stessa forma funzionale a meno di un fattore moltiplicativo che tiene conto dei gradi di libertà sommati, passando da una magnificazione all'altra. Da tali gradi di libertà s'origina la funzione $g(K)$. In aggiunta a questo aspetto geometrico, che dipende dal fatto che a magnificazioni differenti è presente un numero diverso di siti, c'è anche un aspetto più "fisico": la costante d'accoppiamento appare

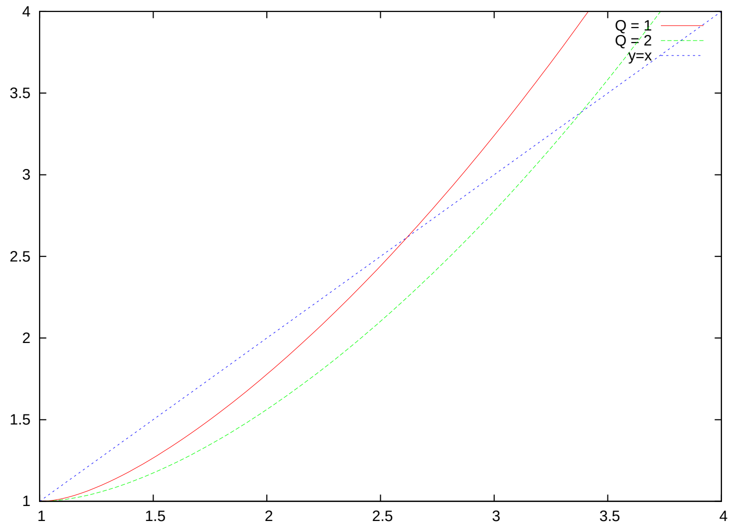


Figura 5.1: La funzione di scala: $K' = \phi(K)$

differente a scale differenti. Una differenza importante tra il contributo fisico e quello geometrico viene dal fatto che mentre il numero di siti di spin varia esponenzialmente con la magnificazione, la costante d'accoppiamento ha una variazione più complessa. Nella prossima sezione analizzeremo gli effetti della funzione che lega le costanti d'accoppiamento a scale diverse sugli esponenti critici.

5.2.2 Punti fissi, fasi stabili e punti critici

Prendiamo ora in considerazione la mappa $K' = \phi(K)$, in figura 5.1, che lega le costanti d'accoppiamento a due magnificazioni successive. Essa ha tre punti fissi, cioè punti per cui $\phi(K) = K$: $K = 1, K_c, \infty$. Il valore di K_c può essere determinato numericamente, ad esempio: $K_c \equiv 3.38$ per $Q = 2$ e $K_c = 2.62$ per $Q = 1$. Prendiamo dapprima in considerazione i due punti critici $K = 1, \infty$. Si può intuire l'esistenza di tali punti critici nel seguente modo. Ricordiamo che $K = e^{\beta J}$. Perciò i due punti critici corrispondono a $J = 0, \infty$. Nel primo caso gli spin sono non interagenti e perciò a qualunque scala li andiamo a guardare essi appariranno completamente disordinati e ciò si manifesta in un punto fisso di $\phi(K)$. Invece per $J = \infty$ tutti gli spin sono allineati e quindi di nuovo a qualunque scala lo guardiamo appare allo stesso modo. Entrambi sono punti fissi attrattivi. Ciò significa che se partiamo con un $K > K_c$ ad una magnificazione molto elevata, diminuendo la magnificazione, guardando a scale macroscopiche, il sistema apparirà sempre più correlato e gli spin tutti paralleli ed infatti K tenderà ad andare al punto fisso attrattivo a ∞. Allo stesso modo se partiamo con $K < K_c$ diminuendo la magnificazione otterremo via via K più piccoli e vicini a 1 e cioè spin completamente scorrelati. In pratica i due differenti regimi corrispondono a temperature elevate ($K < K_c$) e temperature basse ($K > K_c$) e seguono il tipico comportamento di un sistema ferromagnetico.

In sostanza i punti fissi attrattivi del flusso del RG corrispondono alle fasi termodinamiche stabili. Il punto intermedio K_c è invece un punto fisso repulsivo e svolge un ruolo molto diverso. Esso non descrive una fase termodinamica stabile, ma piuttosto il passaggio da una fase all'altra. Il fatto che si tratti di un punto fisso repulsivo, indica il fatto che tale transizione di fase avviene per un valore particolare dei parametri del sistema (in questo caso della temperatura o della costante d'accoppiamento). Una volta che abbiamo un sistema di spin per cui $K = K_c$ ad una scala microscopica e quindi magnificazione elevata, tale valore rimarrà fissato anche se guardiamo il sistema ad un livello macroscopico. Ciò spiega la lunghezza di correlazione infinita tipica dei punti critici, infatti se così non fosse essa dovrebbe diminuire mano a mano che diminuiamo la magnificazione. Il fatto che si tratti di un punto fisso instabile è espresso matematicamente dalla condizione:

$$\lambda = \frac{d\phi(K_c)}{dK} > 1$$

linearizzando:

$$K' - K_c \equiv \lambda(K - K_c)$$

cioè quando siamo in prossimità del punto critico la costante d'accoppiamento cambia riscalando la sua distanza da quella critica. Osserviamo che, se ci troviamo ad un valore di poco inferiore a K_c, il sistema a scale macroscopiche apparirà sempre più distante dal valore critico, poiché infatti si trova in una fase stabile. Il punto critico è importante perché tutte queste scale diverse che sono sempre meno "critiche" diventano improvvisamente tutte egualmente critiche quando $K = K_c$. In sintesi il RG é basato su un duplice processo di scala: cambiando la scala spaziale, cambia anche la distanza dal punto critico nello spazio delle costanti d'accoppiamento e della temperatura (o del tempo nella QFT). Vedremo nel seguito in che modo gl'ingredienti del RG ricompaiono nell'analisi fenomenologica dei processi di branching.

5.3 Un'interessante applicazione

Una situazione che, come vedremo in seguito, può avere un'interessante applicazione è la seguente: consideriamo degli individui che si riproducono attraverso il meccanismo del Galton-Watson con la stessa funzione generatrice $f(x)$. Immaginiamo inoltre che in questa popolazione vi siano due modalità con cui possono formarsi nuove **famiglie**[2], ove s'intende con ciò semplicemente che gl'individui di una famiglia sono distinguibili da quella di un'altra, ma si riproducono tutti allo stesso modo:

- per immigrazione;
- per mutazione.

Per immigrazione s'intende che dall'esterno giungano nuove tipologie d'individui, e perciò tale meccanismo è del tutto indipendente dalla popolazione preesistente. Invece per mutazione intendiamo che un individuo della popolazione già presente può subire un qualche cambiamento che faccia sì che esso non appartenga più alla sua famiglia d'origine, ma ne formi una nuova. In entrambi i casi cerchiamo per tempi grandi, quante sono le famiglie $F(k,t)$ composte da k individui. Converrà introdurre una funzione generatrice[3]:

$$\mathcal{F}(x,t) = \sum_{k=1}^{\infty} F(k,t)x^k \tag{5.9}$$

Consideriamo i due casi separatamente.

[2] per evitare confusione, puntualizziamo che nel seguito chiameremo famiglie i diversi tipi d'individui, piuttosto che, come si potrebbe essere portati a pensare, la discendenza di un individuo. Questo perché in un'applicazione concreta i vari tipi corrisponderanno proprio ai cognomi, cioè i nomi delle famiglie.

[3] Il termine di grado 0 non ha significato in questo caso perché corrisponderebbe alle famiglie composte da 0 individui.

5.3.1 Immigrazione

Supporremo per semplicità che l'immigrazione sia costante nel tempo cioè ad ogni generazione si aggiungano sempre lo stesso numero R d'individui che andranno a formare altrettante nuove famiglie composte da una sola persona; supponiamo per comodità che al tempo $t = 0$ siano presenti R individui. Ciò significa che avremo:

$$\mathcal{F}(x, 0) = Rx$$

Ad ogni generazione gl'individui già presenti si riproducono e ciò corrisponde alla sostituzione:

$$x \to f(x)$$

Inoltre vengono aggiunte R famiglie costituite da un solo individuo che andranno a contribuire al termine di primo grado. Complessivamente s'ottiene la seguente espressione ricorsiva:

$$\mathcal{F}(x, t+1) = \mathcal{F}(f(x), t) + Rx \tag{5.10}$$

che è risolta con la notazione delle precedenti sezioni, da:

$$\mathcal{F}(x, t) = R \sum_{n=0}^{t} f_n(x)$$

dove $f_0(x) = x$. Si noti che tale scrittura ha un valore solo formale, essendo evidentemente divergente in tutti i punti poiché $f_n(x) \xrightarrow{n \to \infty} q$; ciò non ci preoccupa perché siamo in realtà interessati ai coefficienti dello sviluppo in serie di $\mathcal{F}(x)$ e possiamo pensare di risottrarre il termine di grado 0 divergente che non ha alcun significato (non hanno infatti senso le famiglie costituite da 0 individui). È facile osservare che l'equazione (5.10) è analoga alla (5.1) che abbiamo introdotto parlando del RG. Possiamo quindi usare gli strumenti che il gruppo di rinormalizzazione fornisce per ottenere alcune informazioni. Siamo chiaramente interessati al limite:

$$\mathcal{F}(x) = \lim_{t \to \infty} \mathcal{F}(x, t)$$

Vogliamo far vedere che:

$$F(k, t) \xrightarrow{t \to \infty} Ak^{-\gamma} \quad \Rightarrow \quad F(k, t)Ak^{-\gamma}(1 + o(1)$$

dove γ dovrà essere determinato e A è una opportuna costante di proporzionalità irrilevante per i nostri scopi, cioé la distribuzione delle differenti famiglie segue una legge a potenza. Osserviamo che poiché $F(1,t) \to \infty$ dovremo avere $\gamma \leq 1$. Il valore di γ si collega a quello dell'esponente critico introdotto nella (5.5). Vediamo come.

Il ruolo del flusso del RG è svolto dalla funzione generatrice del processo di Galton-Watson: $f(x)$. Per comprendere l'andamento del sistema dobbiamo guardare allora i suoi punti fissi, cioé le soluzioni dell'equazione:

$$f(x) = x$$

sappiamo giá che la piú piccola di tali soluzioni é la probabilitá d'estinzione q introdotta nella definizione 1.3. Poiché $f(x)$ é convessa e vale 1 in 1 deve aversi $f'(q) < 1$ (figura 1.1). Quindi si tratta di un punto fisso attrattivo. Abbiamo poi che $f(1) = 1$ ed inoltre $f'(1) = m$. Se ammettiamo che la popolazione stia crescendo esponenzialmente, cioè $m > 1$, avremo che in 1 è presente un punto fisso repulsivo. Un ulteriore punto fisso attrattivo é l'infinito. Quindi, usando il linguaggio termodinamico, possiamo dire che il punto fisso in 1 é un punto fisso che separa le due fasi stabili corrispondenti ai due punti fissi $x = q$ e $x = \infty$. Osserviamo che in questa trasposizione la funzione $\mathcal{F}(x)$ assume il ruolo di energia libera, il tempo diviene il grado di magnificazione, mentre la x, che nel Galton-Watson è semplicemente una variabile formale, assume il ruolo di una costante d'accoppiamento o di una temperatura. Da quanto visto nell'analisi generale del RG sappiamo che in prossimità del punto critico dovremo avere:

$$\mathcal{F}(x) \propto (1-x)^{\alpha}$$

il valore di α si determina mediante la (5.6). Poichè nel nostro caso abbiamo $\mu = 1$ avremo però:

$$\alpha = 0 \tag{5.11}$$

Questo ci fa intuire che in realtá abbiamo un punto critico particolare. Infatti come abbiamo visto nel calcolo esplicito nella sezione 5.2, la costante μ é associata al rapporto tra il numero di gradi di libertá in gioco a due magnificazioni successive. In questo caso il valore $\mu = 1$, indica la situazione anomala in cui a due magnificazioni successive, pur allontanandoci dal punto critico, poichè $m > 1$, il numero dei gradi di libertà non sta cambiando. Il risultato é che la F continua a divergere, ma lo fa piú lentamente di qualunque potenza come asserisce la (5.11); essa diverge infatti logaritmicamente. Per verificarlo, è sufficiente osservare che:

$$\lim_{x\to 1} \frac{F'(x)}{\exp[\log(m)F(x)]} = \lim_{x\to 1}\lim_{T\to\infty} \frac{F'(x,T)}{\exp[\log(m)F(x,T)]} =$$
$$= \lim_{T\to\infty}\lim_{t\to 1} \frac{F'(x,T)}{\exp[\log(m)F(x,T)]} = \lim_{T\to\infty} \frac{1}{(m-1)} \frac{m^{T+1}-m}{m^T} = \frac{m}{m-1} \tag{5.12}$$

Ma l'equazione seguente:

$$\mathfrak{F}'(x) = K\exp(a\mathfrak{F}(x))$$

è risolta da:

$$\mathfrak{F}(x) = -\frac{1}{a}\log(c - aKx)$$

ove c è una costante arbitraria. Scegliendo $a = \log m$ e $K = \frac{m}{m-1}$ riotteniamo la (5.12) la cui soluzione in prossimità di 1 è:

$$\mathcal{F}(x) \simeq -\frac{\log(1-x)}{\log m} \tag{5.13}$$

C'interessa ora collegare la forma ottenuta per $F(x)$ in prossimità del punto critico, con il valore dell'esponente γ che regola il modo in cui vanno a 0 le $F(k,T)$ per T molto grande. Introduciamo allora:

$$Li_s(z) = \sum_{k=1}^{\infty} \frac{z^n}{n^s}$$

sono delle funzioni note dette **polilogaritmi**, per via del fatto che per $s = 1$ si ottiene proprio $-\log(1-z)$. Come osserviamo, i coefficienti dello sviluppo in serie vanno a coincidere con la forma asintotica che cerchiamo per gli $F(k,T)$. Possiamo scrivere allora:

$$\mathcal{F}(x) = CLi_\gamma(x)(1 + o(1))$$

ove C è una opportuna costante di proporzionalità ininfluente per i nostri fini. È facile verificare che per i polilogaritmi vale:

$$Li_s(z) \xrightarrow{z\to 1} \begin{cases} \Gamma(-s+1)(1-z)^{s-1} & s < 1 \\ -\log(1-z) & s = 1 \end{cases} \tag{5.14}$$

dalla forma che abbiamo ottenuto nella (5.13), otteniamo immediatamente, confrontandolo con tali espressioni, che $\gamma = 1$ e $C = 1/\log(m)$.

5.3.2 Mutazioni

Assumeremo per semplicità che la probabilità che un individuo subisca una mutazione sia ρ uguale per tutti. Cerchiamo anche in questo caso di ottenere una regola ricorsiva per $\mathcal{F}(x,t)$. Partiamo dalla condizione iniziale $\mathcal{F}(x,0) = x$. Al passare di una generazione una parte proporzionale a ρ dei figli di ciascun individuo andrà a formare nuove famiglie composte da un individuo mentre la parte restante continuerà a far parte della sua famiglia; ciò equivale a:

$$\mathcal{F}(x,t+1) = \mathcal{F}\left(f\left(x^{1-\rho}\right),t\right) + \rho m \mathcal{F}'(1,t)x$$

dove come nel caso precedente di immigrazione la scrittura è da ritenersi solo formale in quanto va sottratto il contributo divergente e privo di senso costituito dalle famiglie composte da 0 individui. Osserviamo che $\mathcal{F}'(1,t)$ corrisponde al numero totale d'individui al tempo t e perciò, visto che l'effetto di mutazione agisce solo sulle famiglie e non muta il numero totale degli individui, si ha che:

$$\mathcal{F}'(1,t) = m^t$$

Indicando con $r(x) = f\left(x^{1-\rho}\right)$ e con $r_n(x)$ le sue iterate avremo:

$$\mathcal{F}(x,t) = \rho \sum_{n=0}^{t-1} m^{t-n} r_n(x) + r_t(x)$$

Qui però la situazione è diversa da quanto visto in precedenza: basta utilizzare il teorema 1.9 con cui possiamo vedere facilmente che non solo il termine di grado 0, che come già detto era divergente anche nel caso d'immigrazione e non ha alcun significato fisico, ma anche tutti gli altri termini nello sviluppo in serie sono divergenti. Ciò significa che tale meccanismo di formazione delle famiglie, per cui il numero di nuove famiglie cresce proporzionalmente al numero totale d'individui, fa si che diverga il numero di famiglie di qualunque dimensione. Tuttavia è possibile trovare ancora una distribuzione asintotica interessante, perchè tutti questi coefficienti divergono allo stesso modo. Anziché interessarci quindi al numero assoluto di famiglie composte da un certo numero d'individui, possiamo considerare il valore relativo al numero totale di famiglie. Infatti in questo caso divergono non solo il numero di famiglie, cioè $\mathcal{F}(1,t)$, (che divergeva anche nel caso precedente), ma anche il numero di famiglie composte da k individui $F(k,t)$, mentre è finito il loro rapporto. Lavoriamo perciò su:

$$\Phi(x,t) = \sum_{k=1}^{\infty} \frac{F(k,t)}{\mathcal{F}(1,t)} x^k = \sum_{k=1}^{\infty} \phi(k,t) x^k$$

dove abbiamo: $\Phi(1,t) = \sum_k \phi(k,t) = 1$. Per semplificare le cose osserviamo che $\mathcal{F}(1,t) \propto m^t$, quindi:

$$\phi(k,t) \propto F(k,t) m^{-t}$$

e, visto che siamo interessati all'andamento delle $\phi(k,t)$ e non ad eventuali costanti moltiplicative, possiamo sfruttare l'espressione al secondo membro, per la quale s'ottiene la ricorsione nella semplice forma:

$$\Phi(x,t+1) = \rho x + \frac{\Phi(r(x),t)}{m} \tag{5.15}$$

È sorprendente notare che si tratta ancora di un'equazione ricorsiva del genere della (5.1) che abbiamo visto per il RG. Anche in questo caso la funzione generatrice $\Phi(x)$ svolge il ruolo dell'energia libera, la x diviene la temperatura o la costante d'accoppiamento, mentre il ruolo del flusso del RG è svolto in questo caso dalla funzione $r(x)$. Dobbiamo quindi analizzarne i punti fissi per capire l'andamento della funzione Φ. Per valori piccoli di ρ la funzione $r(x)$ ha un andamento simile a quello di $f(x)$ di cui dovrà continuare ad avere lo stesso valore in 0 e in 1; continua ad essere monotona, tuttavia non lo sono più le sue derivate. Infatti $r(x)$ ha la derivata singolare in 0, mentre in 1 la sua derivata vale $(1-\rho)m$. Perciò la derivata prima non è più monotona e la funzione presenterà un punto di flesso. Assume in questo caso un ruolo determinante il prodotto:

$$h \equiv (1-\rho)m$$

per l'equazione dei punti fissi di $r(x)$:

$$r(x) = x$$

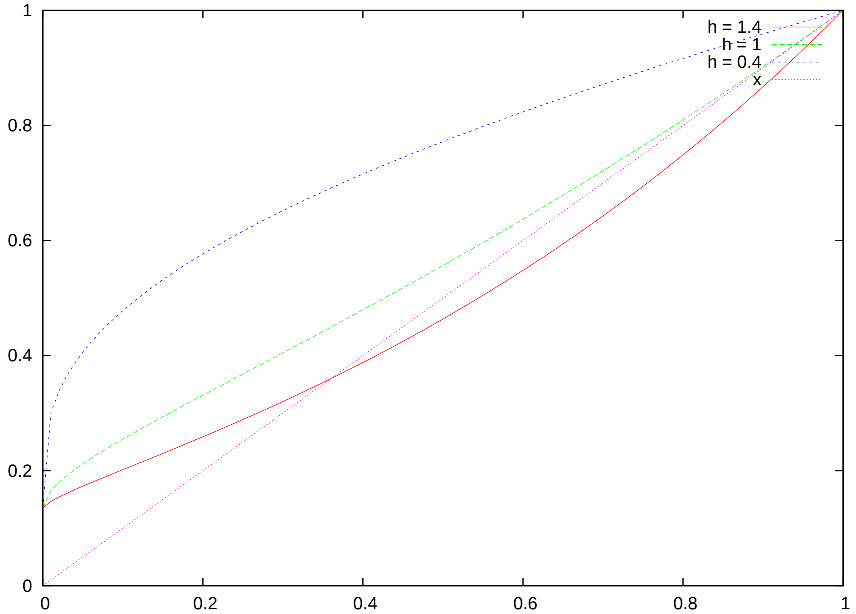


Figura 5.2: Le funzioni $r(x)$ per una poissoniana con $m = 2$ al variare del parametro h.

Infatti come si vede nella figura 5.2, se $h > 1$ allora $r(x)$ dovendo arrivare con una pendenza maggiore di quella della bisettrice avrà una intersezione $q' > q$. In tali condizioni, cioè, la situazione è del tutto analoga a quanto accadeva per la funzione $f(x)$, con $m > 1$. Se ρ cresce, h diviene minore o uguale a 1 ed in tali condizioni non vi è più un punto fisso minore di 1 e quest'ultimo non è più repulsivo. Possiamo quindi dire che $r(x)$ si comporta dal punto di vista dei punti fissi proprio come la funzione $f(x)$ studiata nella sezione 1.2 a costo di sostituire m con h. Mettiamoci pertanto nel caso $h > 1$ e cioè:

$$\rho < \frac{m-1}{m}$$

In tal caso in 1 abbiamo un punto fisso repulsivo, cioè un punto critico. Asseriamo di nuovo che si ha una legge a potenza per i $\phi(k) = \lim_{t\to\infty} \phi(k,t)$, cioè per t grande:

$$\phi(k) \simeq Ak^{-\gamma}(1 + o(1))$$

ove come prima γ è da determinare e A è una costante di proporzionalità irrilevante. Per determinare γ utilizzeremo il valore dell'esponente critico

α introdotto nella (5.6). Mediante tale espressione tenendo conto che nel nostro caso $\mu = m$ e $\lambda = h$ s'ottiene:

$$\alpha = \frac{\log(m)}{\log(h)} = \frac{\log(m)}{\log(m) + \log(1-\rho)}$$

Per connettere α con γ utilizziamo le (5.14) da cui s'ottiene:

$$\gamma = \alpha + 1 = 2 + \frac{\log(1-\rho)}{\log(m) + \log(1-\rho)} \simeq 2 - \frac{\rho}{\log(m)}$$

dove nell'ultima espressione abbiamo sviluppato per valori piccoli di ρ.

5.3.3 I casi rimanenti: $m > 1$ e ρ grande

Ci siamo fino ad ora limitati a considerare i casi in cui $m > 1$ e quindi la popolazione cresceva esponenzialmente e susseguentemente nel caso di mutazioni abbiamo analizzato il caso in cui ρ era abbastanza piccolo da far sì che fossimo ancora in tale regime. Ora vogliamo analizzare quali differenze compaiono nei casi in cui la popolazione non ha più una crescita esponenziale. Cominciamo con delle considerazioni generali a partire dalle equazioni ricorsive del tipo della (5.4), in particolare le (5.10) e (5.15); ci riferiremo alle soluzioni limite con $F(x)$, mentre indicheremo con $f(x)$ il flusso del RG, cioè la funzione di cui sono importanti i punti fissi. Cominciamo col caso in cui $f'(1) < 1$. Il fatto essenziale è che la $f(x)$, in tali condizioni, non ha un punto fisso repulsivo in 1. Ammetteremo per semplicità che $f(x)$ sia definita e C^∞ su tutta la retta reale. Allora l'equazione di punto fisso (1.3) avrà una soluzione $x_c > 1$ e $f'(x_c) > 1$, questo perché $f(x)$ dovrà essere convessa, visto che tutte le derivate in 0 sono positive. Si avrà che 1 e ∞ sono due punti fissi attrattivi mentre x_c è un punto fisso repulsivo che separa le due fasi stabili, quindi un punto critico. Perciò per quanto visto in precedenza, per valori prossimi a x_c:

$$F(x) \propto (x_c - x)^\alpha \propto \left(1 - \frac{x}{x_c}\right)^\alpha \simeq Li_{\alpha+1}\left(\frac{x}{x_c}\right)$$

da cui s'ottiene che i coefficienti dello sviluppo in serie di $F(x)$ vanno asintoticamente come $x_c^{-n} n^{\alpha+1}$: quindi all'andamento a potenza, si aggiunge un esponenziale che domina.

Vediamo il caso in cui $f'(1) = 1$. Ora la situazione dei punti fissi è particolare. Infatti il punto fisso attrattivo, normalmente collocato in $q < 1$ è andato ad identificarsi col punto fisso repulsivo. Pertanto, iterando $f(x)$, tutti i punti al di sotto di 1 convergono a 1, mentre quelli al di sopra di 1 crescono indefinitamente verso il punto fisso all'infinito. Per affrontare tale caso conviene riprendere in considerazione il teorema 1.8. Esso ci permette di scrivere una stima per le iterate di $f(x)$:

$$f_n(x) \simeq \left(1 - \frac{2}{n\zeta}\right) + \frac{2}{n\zeta}\left(\frac{2}{2 - n\zeta \log x}\right)$$

ove con ζ indichiamo $f''(1)$. Inserendo tale espressione nella (5.4), dobbiamo distinguere i due casi: $\mu = 1$ e $\mu > 1$.

- per $\mu = 1$ otteniamo:

$$\sum_{n=0}^{N} f_n(x) = N + \sum_{n=0}^{N}\left(\frac{2\log x}{2 - n\zeta\log x}\right) \simeq N - \frac{2}{\zeta}\int_{-\frac{\zeta\log x}{2}}^{N} \frac{dy}{y} \simeq$$

$$\simeq N - \frac{2}{\zeta}(\log N + \log(-\log x))$$

  ove abbiamo trascurato dei termini costanti. Quindi a parte delle costanti divergenti, rimane in prossimità di 1 un andamento analogo alla (5.13), quindi $\gamma = 1$.

- per $\mu > 1$ la situazione è più complessa poiché ora:

$$F(x) = \sum_{n=0}^{\infty} \frac{f_n(x)}{\mu^n}$$

  è C^∞ fino ad 1, ove ammette tutte le derivate sinistre. Invece per $x > 1$ la somma non converge più, poiché le iterate crescono molto più velocemente della progressione geometrica al denominatore. Ne consegue che non possiamo più determinare l'andamento asintotico dello sviluppo di Taylor di $F(x)$ in 0 guardando i punti in cui non è analitica (di solito in 1), poiché o essa ammette tutte le derivate o non è proprio convergente. Per farlo andrebbe quindi analizzato proprio l'andamento intorno allo 0 ma ciò non è semplice.

5.3.4 La distribuzione dei cognomi

Un interessante contesto empirico in cui il modello che abbiamo visto in questa sezione trova una applicazione diretta è nello studio della distribuzione dei cognomi. Con quest'ultima s'intende, fissata una popolazione, il numero di cognomi che hanno una data dimensione. Risulta quindi evidente che le $F(k,t)$ introdotte nella (5.9) fanno proprio al caso nostro. I dati sperimentali([10, 4]), riportati nella tabella 5.1, mostrano effettivamente una legge a potenza per le $F(k,t)$. L'esponente corrispondente sarà nuovamente indicato con γ . Con N_c s'indica invece il numero totale di cognomi presenti che risulta interessante se messo in relazione con N. Proprio da tale relazione, si nota che i dati disponibili si prestano molto bene ad essere divisi in due gruppi: quelli per cui $N_c \simeq \log N$ (Cina e Corea) e quelli per cui $N_c \simeq N^\nu$, ove ν è un opportuno esponente (gli altri paesi). Collegandoci a quanto visto prima, i due gruppi lasciano intendere una differente dinamica nella formazione dei cognomi. Abbiamo infatti visto che se la formazione di nuovi cognomi è legata all'immigrazione, allora il numero di cognomi sarà lineare nel tempo, mentre il numero d'individui cresce esponenzialmente. Quindi i due saranno legati da una legge logaritmica. A conferma del fatto che alla base dell'esponente della distribuzione sta proprio il meccanismo con cui s'originano i nuovi cognomi, viene il fatto che l'esponente da noi ottenuto ricalca quello di questo gruppo di paesi. Il motivo per cui in Cina e Corea dominano le immigrazioni rispetto alle mutazioni è che vi è una forte tradizione che ha teso nel corso dei secoli a mantenere intatti i nomi delle famiglie, tanto da rendere il loro numero straordinariamente ridotto[4]. Nel secondo gruppo la situazione è diversa e il numero di cognomi è legato da una legge a potenza a quello degli individui. Ciò ci porta verso il secondo meccanismo di formazione di nuove famiglie che abbiamo analizzato: le mutazioni. Il valore ρ da noi inserito nel modello, corrispondente alla probabilità di mutazione dei cognomi ad ogni generazione, è in generale piuttosto piccolo. Ad esempio i dati sulla popolazione svedese tra il 2004 e il 2006 forniscono un incremento dello 0.456%. Nello stesso periodo, circa 100 nuovi cognomi si

[4]In Cina i primi 100 cognomi più diffusi coprono quasi il 90% della popolazione. In Corea sono presenti circa 250 cognomi attualmente in uso [24, 20, 19]

Tabella 5.1: Il coefficiente della distribuzione dei cognomi e la relazione tra numero di cognomi e di abitanti per diverse regioni

Regione	γ	N_c
Cina	1	$\log N$
Corea	1	$\log N$
Argentina	2	$N^{0.84}$
Austria	2	$N^{0.83}$
Berlino	2	$N^{0.90}$
Francia	1.5	$N^{0.90}$
Germania	1.5	$N^{0.77}$
Isle of Man	1.5	$N^{0.75}$
Italia	1.75	$N^{0.75}$
Giappone	1.75	$N^{0.65}$
Olanda	2.16	$N^{0.69}$
Norvegia	2.16	$N^{1.0}$
Sicilia	0.46-1.83	$N^{1.0}$
Spagna	1.9	$N^{0.81}$
Svizzera	1.9	$N^{0.73}$
Taiwan	1.9	$N^{0.69}$
Usa	1.94	$N^{0.69}$
Venezuela	1.43	$N^{0.69}$
Vietnam	1.43	$N^{0.27}$

sono formati ogni mese, corrispondenti ad un incremento relativo al numero di abitanti, annuale pari allo 0.015%. S'ottiene una stima pertanto di:

$$\frac{\rho}{\log m} \simeq \frac{0.015}{0.456} \simeq 0.03$$

Pertanto l'esponente da noi ottenuto, è approssimabile a 2. Inoltre il numero di cognomi cresce linearmente col numero d'individui. Evidentemente questo secondo modello risulta più vicino ai dati in questione, ma non ne spiega la varietà né per i vari esponenti λ, né per l'esponente ν che governa la relazione tra N_c e N. Chiaramente si tratta di un modello semplificato, che non può tenere conto del contesto storico differente nelle varie situazioni geografiche riportate. Si può comunque ritenere un buon risultato la parziale spiegazione di questi numeri, oltre che del mero andamento a potenza sulla base del Galton-Watson process, tramite l'ausilio del RG. Inoltre a supportare i nostri risultati s'inserisce il fatto che il lavoro presente in [10], basato sulla master equation, perviene a risultati analoghi basandosi su un modello semplificato, con tempo continuo e completamente risolubile. Ciò vale come ulteriore prova dell'universalità tipica dei problemi in cui interviene l'invarianza di scala.

Conclusioni

A consuntivo del lavoro di tesi effettuato, può essere interessante delineare alcune direzioni per un ulteriore approfondimento che per ragioni di tempo non si sono potute seguire.

Per cominciare osserviamo che l'inquadramento concettuale che s'è dato nel capitolo 2 lascia alcune domande:

- è possibile in maniera semplice trovare un'hamiltoniana con un'opportuna interazione, in cui si formalizzi l'accoppiamento, ad esempio il Galton-Watson bisessuale?

- le distribuzioni asintotiche per il Galton-Watson, individuate nei vari casi nella sezione 1.4, possono essere viste come problemi agli autovalori relativi ad opportuni operatori?

Inoltre, uno sguardo complessivo al lavoro effettuato pone naturalmente una questione: è possibile collegare la trattazione quantistica con quella più vicina alla fisica statistica e al gruppo di rinormalizzazione degli ultimi capitoli? Osserviamo infatti che gli strumenti del RG emergono non appena si sia raggiunta l'equazione ricorsiva che normalmente riproduce la variazione del sistema al variare della scala d'osservazione, da cui è possibile dedurre le informazioni sui punti fissi e quindi le fasi e i punti critici. Nel nostro caso, l'equazione ricorsiva racconta un'evoluzione temporale e la struttura gerarchica è frutto del fatto che l'evoluzione segue un fenomeno di branching. Sarebbe pertanto auspicabile poter far emergere la struttura ricorsiva in maniera diretta dalla dinamica e pertanto, inserendolo nel contesto quantistico da noi introdotto, dall'hamiltoniana. In tal maniera, dato che l'hamiltoniana può essere complicata semplicemente aggiungendo interazioni differenti, è possibile applicare gli strumenti della fisica statistica e quindi del gruppo del rinormalizzazione, anche a processi di branching, quali ad esempio il Galton-Watson bisessuale, che non è possibile, o quantomeno evidente inserire in un reticolo che ne riproduca le probabilità di transizione. Ciò permetterebbe, per esempio, d'ottenere informazioni riguardo l'andamento delle distribuzioni asintotiche anche in tali casi.

Anche la parte conclusiva della tesi, in cui abbiamo visto da più vicino un'applicazione concreta quale quella della distribuzione dei cognomi, lascia aperte alcune questioni. Infatti analizzando i dati, s'osserva che se da un lato troviamo alcune regioni in cui la distribuzione dei cognomi rivela un esponente direttamente collegabile con i modelli da noi esposti, rimangono altri casi in cui si mostrano esponenti differenti. È vero che esponenti notevolmente discosti dai valori da noi ottenuti, oltre dagli esponenti delle altre regioni, sono dovuti a situazioni contingenti, interpretabili come regioni che non hanno ancora raggiunto la condizione stazionaria, ma rimangono parallelamente valori degli esponenti simili, o intermedi tra 1 e 2 che si può auspicare di spiegare tenendo conto di ulteriori meccanismi. Infatti sappiamo bene che dei due regimi, quello delle mutazioni è immediatamente dominante sull'altro e perciò un valore intermedio non può spiegarsi come una coesistenza dei due, ma richiede una modifica del modello stesso. Varie modifiche e approfondimenti sono interessanti per verificare se riescono a modificare basilarmente la dinamica, fino a mutare gli esponenti critici:

- si sono considerate velocità d'immigrazione, di mutazione e di riproduzione costanti nel tempo; s'intuisce tuttavia che a determinare l'esponente è il differente andamento tra la velocità di riproduzione degli individui e quella di produzione dei cognomi. Ciò si può capire perché una variazione dei due rate, che introduca la medesima dipendenza temporale, lasciandone invariato il rapporto, viene eliminata variando la scala dei tempi e non può quindi generare modifiche al valore asintotico per tempi grandi.

- nel caso delle mutazioni, s'è ammesso che tutti i cognomi mutino allo stesso modo, quando esistono due argomenti che si muovono in direzioni opposte per dubitare di questo fatto:

1. storicamente i cognomi nascono come soprannomi poi attribuiti alla discendenza; quando essi divenivano troppo diffusi non erano più comodi per identificare un individuo e subivano quindi una modifica. Perciò per tale motivo i cognomi più diffusi in passato mutavano maggiormente.
2. un altro meccanismo che produce nuovi cognomi per mutazioni è proprio dovuto ad errori di trascrizione. Evidentemente un cognome più diffuso è più difficile che venga sbagliato perché più noto. Tale meccanismo fa sì che i cognomi più diffusi mutino meno.

Varrebbe pertanto la pena di tentare d'adattare e complicare il modello tenendo conto, ove possibile, di tali fattori aggiuntivi, al fine di comprendere se essi hanno una effettiva influenza sugli esponenti critici, che si collegano all'andamento asintotico della distribuzione.

Bibliografia

[1] Watson, H. W. and Galton, F., *On the Probability of the Extinction of Families*, Journal of the Anthropological Institute of Great Britain, volume 4, pages 138–144, 1875

[2] Heyde, C. C. and Seneta, E., *I.J. Bienayme: Statistical Theory Anticipated*, Berlin, Germany, 1977

[3] Kendall, D. G., *Branching process since 1873*, Journal of the London Mathematical Society, 41: 385–406, 1966

[4] Manrubia, S.C. and Zanette Damian, H., *At the boundary between biological and cultural evolution: The origin of surname distribution*, arXiv:cond-mat/0201559 30 Jan 2002

[5] Jarvis, P.D., Bashford, J.D., Sumner, J.G., *Path integral formulation of phylogenetic branching processes*, arXiv:q-bio/0411047v1 [q-bio.PE] 27 Nov 2004

[6] Peliti, L., *Path integral approach to birth-death processis on a lattice*, J. Physique 46 (1985) 1469-1483

[7] Dickman, R. and Vidigal, R., *Path integrals and Perturbation Theory for Stochastic Processes*, Brazilian Journal of Physics, vol. 33, no. 1, March, 2003

[8] Otter, R., *The Multiplicative Process*, The Annals of Mathematical Statistics, Vol. 20, No. 2 (Jun. 1949) pp. 206-224)

[9] Eckmann, J.P. and Wittwer, P., (1986) *Multiplicative and Additive Renormalization*, in Osterwalder, K. and Stora, R. eds., Critical Phenome-

na, Random Systems, Gauge Theories. Les Houches, Session XLIII, 1984, Elsevier Science Publishers B.V., 455–465.

[10] Seung Ki Baek, Hoang Ahn Tuan Kiet, Beom Jun Kim, *Family name distributions: Master equations approach*, Physical Review E76, 046113 (2007)

[11] Miyazima Sansuke, Lee Youngki, Tomomasa Nagamine, Hiroaki Miyajima, *Power-law distribution of family names in japanese society*, Physica A 278 (2000) 282-288

[12] Kolmogorov, A. and Dmitriev, N. A., *Branching stochastic processes*, Docklady **56**, 5–8

[13] Itzykson, C. and Zuber, J. B., *Quantum field theory*, McGraw-Hill New York, 1980, p. 435

[14] Derrida, B., Manrubia, S.C. and Zanette, D.H., *Statistical properties of genealogical trees*, Phys. Rev. Lett. 82 (1999) 1987

[15] Derrida, B., Manrubia, S.C., and Zanette, D.H., *Distribution of repetitions of ancestors in genealogical trees*, Physica A 281 (2000) 1

[16] Zanette, B. H. and Manrubia, S.C., *Vertical transmission of culture and the distribution of family names*, Physica A 295 (2001) 1

[17] Reed, W.J. and Hughes, B.D., *From gene families to incomes and internet file sizes: Why power laws are so common in nature*, Phys. Rev. E 66 (2002) 067103

[18] Reed, W.J. and Hughes, B.D., *On the distribution of family names*, Physica A 319 (2003) 579

[19] Kiet, H.A.T., Baek, S.K., Kim, B.J., Jeung, H., *Koran Family Name Distribution in the Past*, J. Korean Phys. Soc. 51 (2007) 1812

[20] Kim, B.J. and Park, S. M., *Distribution of Korean Family Names*, Physica A 347 (2005) 683

Libri

[21] Harris, T.E., *The theory of branching process*

[22] Athreya, K. B., *Branching process*

[23] Sornette, D., *Critical Phenomena in Natural Sciences*, Springer (2006)

[24] Lee, Ki-baek ,*A new history of Korea*, Seoul: Ilchokak (1984)