



UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia “Galileo Galilei”

Corso di Laurea in Fisica

Tesi di Laurea triennale

Similitudini tra la diffusione di polimeri in un ensemble gran canonico e le fluttuazioni di prezzo degli asset finanziari

Relatore

Prof. Fulvio Baldovin

Laureando

Alice Bertelli

Anno Accademico 2022/2023

Indice

Introduzione	3
1 Moto diffusivo dei polimeri	4
1.1 Moto Browniano	4
1.1.1 Visuale Lagrangiana: l'equazione di Langevin	6
1.1.2 Visuale Euleriana: l'equazione di diffusione	7
1.2 Modello di Rouse	8
1.3 Diffusione non gaussiana	10
2 Dinamica microscopica dell'evoluzione del prezzo di un asset	13
2.1 Introduzione	13
2.2 Fatti Stilizzati	13
2.3 Diffusione non Gaussiana	14
2.4 Limit Order Book	15
2.5 Modello Markoviano dell'Order Book	16
2.6 Corrispondenze tra sistema polimerico e <i>order book</i>	18
2.7 Dipendenza dal "depth" del comportamento diffusivo del prezzo di un asset	18
Conclusioni	22
Appendici	23
Bibliografia	25

Introduzione

Studiando un sistema polimerico in contatto con un reservoir di particelle si può notare che esso presenta molte caratteristiche in comune al sistema dell'*order book* in contatto con un reservoir di ordini, che descrive le interazioni che avvengono nei mercati finanziari. In particolare, noteremo che la posizione del centro di massa di un polimero e il prezzo di un asset finanziario sono soggetti a fluttuazioni aleatorie e seguono entrambi un moto diffusivo, con un coefficiente dipendente dal numero di elementi che costituisce il sistema. In questa tesi si vogliono sottolineare le similitudini tra i due sistemi mettendoli a confronto e descrivendo un modello per la dinamica microscopica dell'evoluzione del prezzo di un asset finanziario.

Per prima cosa, studieremo il moto diffusivo di un polimero in soluzione acquosa. Per farlo si introduce il moto Browniano dei singoli monomeri: la visuale Lagrangiana del moto ci permetterà di ottenere l'equazione di Langevin e dalla visuale Euleriana ricaveremo la distribuzione della posizione per la singola particella. Utilizzando tali equazioni, studieremo il modello polimerico di Rouse. In particolare, osserveremo come il coefficiente di diffusione del centro di massa di un polimero, e quindi, la varianza della sua distribuzione, dipendano dal numero di monomeri che compongono la catena polimerica. Si vedrà, poi, che tale distribuzione del centro di massa si discosta da una gaussiana, nel caso in cui la grandezza del polimero fluttui nel tempo. Si associerà, in particolare, al processo di polimerizzazione/depolverizzazione, un processo di *birth-death*.

Si passerà, in seguito, allo studio della dinamica che governa le oscillazioni di prezzo di un asset finanziario. Noteremo come tale dinamica abbia caratteristiche empiriche (dette Fatti Stilizzati) comuni anche al moto dei polimeri, come la non gaussianità della distribuzione dei prezzi e la dipendenza della sua varianza dal numero di ordini di mercato. In particolare osserveremo i dati del DJI (Dow Jones Index), che mostreranno tale non gaussianità della distribuzione. Introduremo, quindi, il sistema dell'*order book*, che ci permetterà di costruire un paragone con il sistema polimerico, che possa giustificare le similitudini osservate empiricamente. Infine, descriveremo il modello Markoviano per l'*order book* tratto dal recente articolo in Ref. [7], dal quale ricaveremo la dipendenza tra la varianza e la *depth*, che (come vedremo) quantifica il numero di ordini nell'*order book*. Si riotterrà, dunque, la relazione tra coefficiente diffusivo e numero di elementi, che era valida anche per i modelli polimerici. Da questa relazione, ricavata teoricamente, si conclude il paragone proposto, associando il modello Markoviano descritto ad uno specifico modello polimerico.

La tesi è divisa in due capitoli. Nel capitolo 1 viene studiato il moto diffusivo dei polimeri. Nella prima sezione si introduce brevemente il moto Browniano. Nella seconda e nella terza, si studia il modello di Rouse per la catena polimerica e si osserva la non gaussianità della distribuzione della posizione del suo centro di massa, nel caso in cui la grandezza del polimero fluttui nel tempo. Nel capitolo 2 trattiamo la dinamica dei prezzi degli asset nei mercati finanziari, dalla quale si notano le similitudini con il sistema polimerico di cui sopra. Alla fine, si discuterà il modello proposto nell'articolo in Ref. [7] e si compareranno i risultati ottenuti in questo capitolo con quelli del modello polimerico proposto nel capitolo precedente.

Capitolo 1

Moto diffusivo dei polimeri

1.1 Moto Browniano

Per descrivere il moto di particelle che misurano da pochi nanometri fino a circa $1 \mu m$ in soluzione acquosa si deve introdurre un modello statistico. Infatti, studiare singolarmente il moto di queste particelle risulta complesso e poco efficace a causa delle collisioni con le molecole d'acqua [3].

Il *Random Walk* (o Passeggiata Aleatoria) è un modello matematico che si riproduce uguale a se stesso in differenti scale spaziali (modello *self-similar*). A partire da questo semplice modello si può studiare il moto Browniano, ossia il moto caotico che compie un monomero all'interno di un fluido a causa dell'azione di forze randomiche. Si tratta di un'approssimazione per la quale le particelle si muovono di passi discreti e infinitesimi, distribuiti normalmente.

Il modello unidimensionale del *random walk* (figura 1.1) si basa su spostamenti aleatori e indipendenti di incremento $\pm\Delta L$ lungo la retta, che avvengono in intervalli di tempo costanti Δt . Assumendo $t_0 = 0$ e $x_0 = 0$ e denotando con $p_N(x)$ la densità di probabilità di trovarsi in x dopo N passi, si osserva che tale funzione di massa di probabilità per $N \gg 1$ tende ad una Gaussiana per il teorema del limite centrale [3].

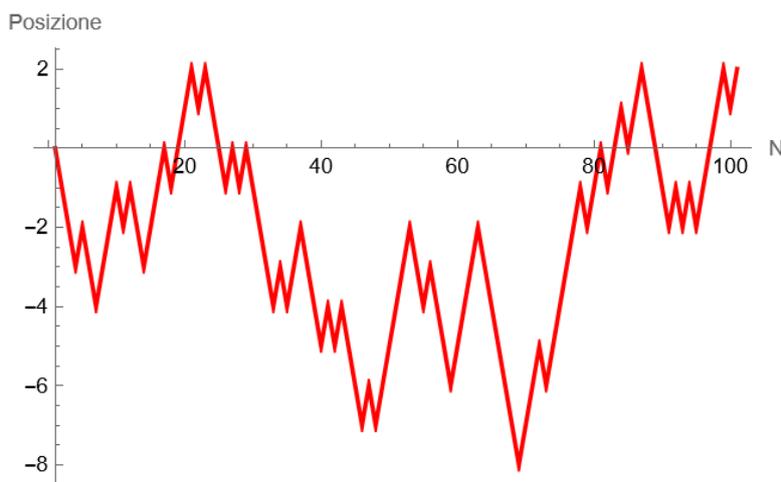


Figura 1.1: Random Walk unidimensionale con $\Delta L = 1$

Si consideri la posizione al tempo zero di una particella $x_0 = 0$. La sua posizione dopo N passi seguirà una distribuzione di probabilità con media $\langle x(N) \rangle = 0$.

Assumiamo gli spostamenti a destra e a sinistra della particella equamente probabili. Chiamiamo

$k(N) = \pm\Delta L$ lo spostamento randomico associato all' N -esimo passo, tale che $x(N) = x(N-1) + k(N)$.
Si ha

$$\langle k(N) \rangle = \frac{1}{2}\Delta L + \frac{1}{2}(-\Delta L) = 0$$

$$\langle [k(N)]^2 \rangle = \frac{1}{2}(\Delta L)^2 + \frac{1}{2}(-\Delta L)^2 = (\Delta L)^2$$

Quindi

$$\begin{aligned} \langle [x(N)]^2 \rangle &= \langle [x(N-1) + k(N)]^2 \rangle \\ &= \langle [x(N-1)]^2 \rangle + 2\langle x(N-1)k(N) \rangle + \langle [k(N)]^2 \rangle \\ &= \langle [x(N-1)]^2 \rangle + 2\langle x(N-1) \rangle \langle k(N) \rangle + (\Delta L)^2 \\ &= \langle [x(N-1)]^2 \rangle + (\Delta L)^2 \end{aligned}$$

dove nel terzo passaggio si è sfruttato il fatto che i passi $k(N)$ fossero indipendenti tra loro. Tramite un ragionamento induttivo si ottiene, dunque, la legge di diffusione

$$\langle [x(N)]^2 \rangle = N(\Delta L)^2 = \frac{t}{\Delta t}(\Delta L)^2 = 2Dt, \quad (1.1)$$

dove è definito il coefficiente di diffusione come $D = \frac{(\Delta L)^2}{2\Delta t}$.

Si può, poi, generalizzare al caso tridimensionale, che meglio descrive il moto di una particella nello spazio. Considero la *Random Walk* nelle tre dimensioni con spostamenti all'interno di un reticolo cubico come in figura 1.2. In questo caso la legge di diffusione diventa

$$\langle [\vec{r}(N)]^2 \rangle = \langle [x(N)]^2 \rangle + \langle [y(N)]^2 \rangle + \langle [z(N)]^2 \rangle = 2Dt + 2Dt + 2Dt = 6Dt. \quad (1.2)$$

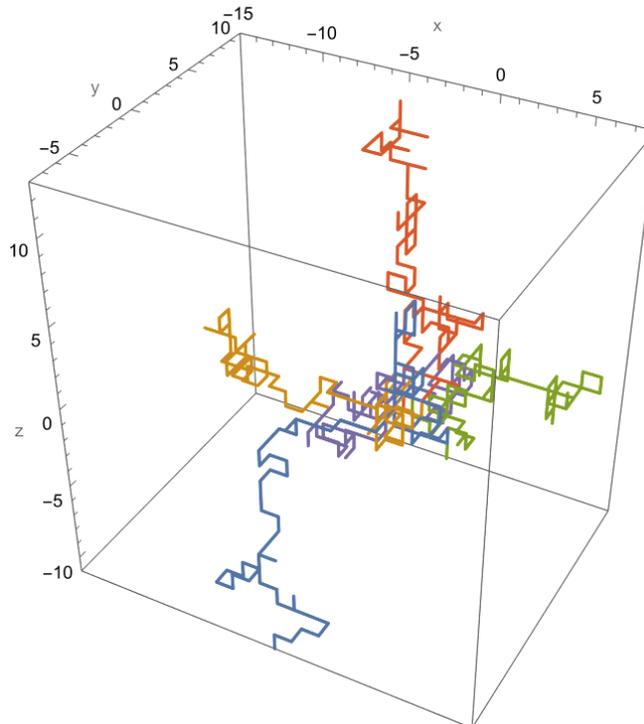


Figura 1.2: Random Walk tridimensionale di 5 particelle

1.1.1 Visuale Lagrangiana: l'equazione di Langevin

Si vuole, ora, ricavare l'equazione di Langevin. Essa risulterà utile nella trattazione del modello di Rouse per i polimeri.

Tale equazione descrive la velocità di una particella che si muove di moto Browniano all'interno di un fluido. Tale particella sarà soggetta a diverse forze:

- una forza randomica distribuita gaussianamente $\vec{f}(t)$,
- la forza di Stokes $\vec{f}_s(t) = -6\pi\eta R\vec{v}(t) = -\zeta\vec{v}(t)$,
- un eventuale campo di forze esterno costante e conservativo.

Nota: La forza di Stokes come scritta sopra si esercita su una particella sferica di raggio R che si muove a velocità \vec{v} all'interno di un fluido Newtoniano. Tale formula è valida in un regime laminare ossia a velocità \vec{v} ridotte e viscosità η elevate.

Detta $\vec{r}(t)$ la posizione della particella al tempo t , si considera ¹

$$\frac{d\vec{r}(t)}{dt} = \vec{v}(t) \quad (1.3)$$

$$\frac{d\vec{v}(t)}{dt} = -\frac{\zeta\vec{v}(t)}{m} + \frac{\vec{f}(t)}{m} - \frac{\vec{\nabla}\Phi(\vec{r})}{m} \quad (1.4)$$

Si applica ora l'approssimazione *overdamped*, per la quale si considera la cosiddetta scala temporale di Smoluchowski (Appendice A), oltre la quale si approssima l'accelerazione istantanea a zero:

$$0 = -\frac{\zeta\vec{v}(t)}{m} + \frac{\vec{f}(t)}{m} - \frac{\vec{\nabla}\Phi(\vec{r})}{m}.$$

L'equazione di Langevin diventa dunque

$$\frac{d\vec{r}(t)}{dt} = \vec{v}(t) = \frac{\vec{f}(t)}{\zeta} - \frac{\vec{\nabla}\Phi(\vec{r})}{\zeta} = \vec{\eta}(t) + \vec{v}_d(\vec{r}) \quad (1.5)$$

dove

- $\vec{\eta}(t)$ è la velocità stocastica per cui vale [3]

$$\langle \vec{\eta}(t) \rangle = 0$$

$$\langle \vec{\eta}(t') \cdot \vec{\eta}(t'') \rangle = 6D\delta(t' - t'')$$

- $\vec{v}_d(\vec{r})$ è la velocità di drift

Si può, per il momento, trascurare il campo di forze esterne applicate al singolo monomero e caratterizzare i momenti della distribuzione della posizione della particella.

Poichè nell'equazione 1.5 rimane solo la componente stocastica della forza (con autocorrelazione nel tempo nulla), per il teorema del limite centrale si ottiene una distribuzione gaussiana della posizione $\vec{r}(t)$, per la quale è sufficiente caratterizzare media e varianza.

¹A causa della forza randomica le traiettorie $\vec{r}(t)$ sono continue ma non differenziabili ovunque. Sarebbe, quindi, più corretto matematicamente scrivere le equazioni differenziali 1.3 e 1.4 come equazioni incrementali; si è scelto, tuttavia, di non applicare tale formalismo matematico, ma di utilizzare l'espressione più comune alla fisica.

$$\langle \vec{r}(t) \rangle = \vec{r}(0) + \int_0^t dt' \langle \vec{\eta}(t') \rangle = \vec{r}(0)$$

$$Var[\vec{r}(t)] = \langle (\vec{r}(t) - \vec{r}(0))^2 \rangle = \int_0^t dt' \int_0^t dt'' \langle \vec{\eta}(t') \cdot \vec{\eta}(t'') \rangle = \int_0^t dt' \int_0^t dt'' 6D\delta(t' - t'') = 6Dt.$$

1.1.2 Visuale Euleriana: l'equazione di diffusione

L'equazione deterministica che regola l'evoluzione della densità di probabilità si chiama equazione di diffusione.

Note la prima legge di Fick [3]

$$\vec{j}(\vec{r}, t) = -D \cdot \frac{\partial p(\vec{r}, t)}{\partial x},$$

dove $\vec{j}(\vec{r}, t)$ è il flusso di probabilità e $p(\vec{r}, t)$ la densità di probabilità, e l'equazione di continuità

$$\frac{\partial p(\vec{r}, t)}{\partial t} = -\vec{\nabla} \cdot \vec{j}(\vec{r}, t),$$

si ottiene l'equazione di diffusione (o seconda legge di Fick)

$$\frac{\partial p(\vec{r}, t)}{\partial t} = D\nabla^2 p(\vec{r}, t). \quad (1.6)$$

Risolvendo questa equazione si ottiene la forma della densità di probabilità. Definendo $p(\vec{r}, t)$ come antitrasformata di Fourier abbiamo:

$$p(\vec{r}, t) = \frac{1}{(2\pi)^{\frac{3}{2}}} \int_{\mathbb{R}^3} \tilde{p}(\vec{k}, t) e^{i\vec{k} \cdot \vec{r}} d^3k \quad (1.7)$$

$$\frac{\partial p(\vec{r}, t)}{\partial t} = \frac{1}{(2\pi)^{\frac{3}{2}}} \int_{\mathbb{R}^3} \frac{\partial}{\partial t} \tilde{p}(\vec{k}, t) e^{i\vec{k} \cdot \vec{r}} d^3k$$

$$\nabla^2 p(\vec{r}, t) = \frac{1}{(2\pi)^{\frac{3}{2}}} \int_{\mathbb{R}^3} \tilde{p}(\vec{k}, t) \nabla^2 e^{i\vec{k} \cdot \vec{r}} d^3k = \frac{1}{(2\pi)^{\frac{3}{2}}} \int_{\mathbb{R}^3} \tilde{p}(\vec{k}, t) (i\vec{k})^2 e^{i\vec{k} \cdot \vec{r}} d^3k$$

Quindi, dall'equazione di diffusione 1.6, si ha

$$\frac{\partial}{\partial t} \tilde{p}(\vec{k}, t) = -Dk^2 \tilde{p}(\vec{k}, t).$$

Risolvendo tale equazione ottengo

$$\tilde{p}(\vec{k}, t) = \tilde{p}(\vec{k}, 0) e^{-Dk^2 t}. \quad (1.8)$$

Assumendo $p(\vec{r}, t) = \delta^3(\vec{r} - \vec{0})$, si ha

$$\tilde{p}(\vec{k}, 0) = \frac{1}{(2\pi)^{\frac{3}{2}}} \int_{\mathbb{R}^3} d^3r e^{-i\vec{k} \cdot \vec{r}} \delta^3(\vec{r} - \vec{0}) = \frac{1}{(2\pi)^{\frac{3}{2}}},$$

dunque, sostituendo la 1.8 nella 1.7, mi restituisce un integrale gaussiano da cui

$$p(\vec{r}, t) = \frac{1}{(2\pi)^{\frac{3}{2}}} \int_{\mathbb{R}^3} \frac{1}{(2\pi)^{\frac{3}{2}}} e^{i\vec{k} \cdot \vec{r}} e^{-Dk^2 t} d^3k = \frac{1}{(4\pi Dt)^{\frac{3}{2}}} e^{-\frac{r^2}{4Dt}}.$$

Fino ad ora non si è considerata l'azione di un campo di forze esterne. Per la trattazione del modello polimerico di Rouse, tuttavia, si necessita l'equazione di diffusione nella formulazione di Fokker-Plank,

dove si considera l'azione di una forza di drift $\vec{f}_d = -\vec{\nabla}\Phi = \zeta\vec{v}_d$. In questo caso, si genera una componente di drift del flusso pari a [3]

$$\vec{j}_d(\vec{r}, t) = p(\vec{r}, t)\vec{v}_d(\vec{r}) = p(\vec{r}, t)\frac{\vec{f}_d(\vec{r})}{\zeta}.$$

Si può, quindi, generalizzare la prima legge di Fick e ottenere, così, l'equazione di diffusione nella formulazione di Fokker-Plank

$$\begin{aligned}\frac{\partial p(\vec{r}, t)}{\partial t} &= -\vec{\nabla} \cdot \vec{j}(\vec{r}, t) = \\ &= D[\nabla^2 p(\vec{r}, t) - \frac{1}{k_b T} \vec{\nabla} \cdot (\vec{f}_d(\vec{r}) p(\vec{r}, t))],\end{aligned}$$

dove si è usata la relazione di Einstein: $\zeta D = k_b T$ [3].

In condizioni di equilibrio, il flusso di probabilità attraverso la superficie è pari a 0, quindi si ha

$$\vec{j}_d(\vec{r}, t) = 0 = -\vec{\nabla} p(\vec{r}, t) + \frac{\vec{\nabla} \phi(\vec{r}) p(\vec{r}, t)}{k_b T},$$

da cui otteniamo

$$p(\vec{r}, t) = C \cdot \exp[-\beta\Phi(\vec{r})], \quad (1.9)$$

dove C è una costante.

1.2 Modello di Rouse

Un polimero è una particella composta da una catena di più particelle singole dette monomeri.

Uno dei modelli più semplici per la descrizione del moto di diffusione dei polimeri è il modello di Rouse [4] [9]. Tale modello si basa sulle seguenti assunzioni:

- Il solvente è indifferente al moto del polimero. In altre parole, il moto del polimero è soggetto alla frizione del solvente in cui è immerso, ma non genera un campo di velocità nel fluido. Questa approssimazione non è efficace quando viene applicata a soluzioni molto diluite, nelle quali la fluidodinamica del sistema è più incisiva.
- I monomeri che formano la catena polimerica interagiscono armonicamente tra loro.
- Le interazioni avvengono solo tra monomeri adiacenti, ossia, l' n -esimo monomero interagisce esclusivamente con l' $(n-1)$ -esimo e l' $(n+1)$ -esimo.

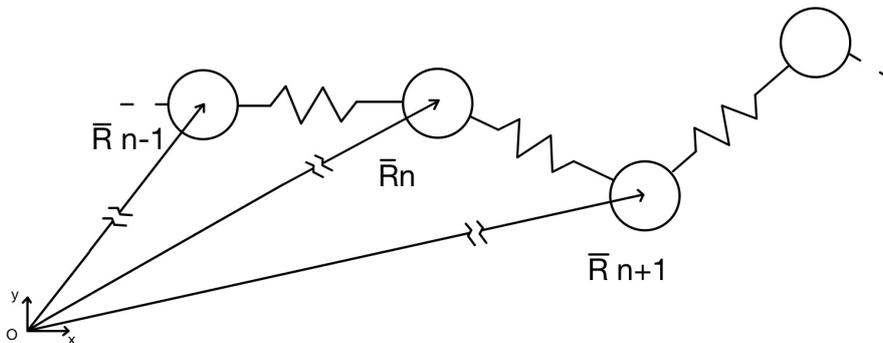


Figura 1.3: Modello di Rouse

Si vuole ricavare l'equivalente della legge di diffusione per il centro di massa del polimero a partire dal modello appena definito.

Considero la posizione del monomero n -esimo \vec{R}_n in una configurazione di equilibrio e assumo gli altri monomeri fissi. La distribuzione all'equilibrio di \vec{R}_n ha la forma dell'equazione 1.9, dove Φ viene sostituito dal valore del potenziale a cui è soggetto il monomero in questo modello [4]:

$$p_{eq}(\vec{R}_n) = C \cdot \exp\left\{-\beta\left[\frac{3K_bT}{2b^2}(\vec{R}_n - \vec{R}_{n-1})^2 - \frac{3K_bT}{2b^2}(\vec{R}_{n+1} - \vec{R}_n)^2\right]\right\}$$

dove b è un parametro chimico che rappresenta la distanza media tra i monomeri.

L'equazione di Langevin che descrive il moto del monomero sarà dunque

$$\frac{d\vec{R}_n}{dt} = -\frac{3k_bT}{\zeta b^2}(2\vec{R}_n - \vec{R}_{n-1} - \vec{R}_{n+1}) + \vec{\eta}_n, \quad (1.10)$$

dove $\vec{\eta}_n$ è la velocità randomica per la quale vale [4]

$$\langle \vec{\eta}_n(t) \cdot \vec{\eta}_m(t') \rangle = 6D\delta_{n,m}\delta(t-t').$$

L'equazione 1.10 è valida per ogni monomero fissati gli altri $N-1$, a meno dei monomeri di bordo del polimero per i quali valgono le seguenti:

$$\frac{d\vec{R}_0}{dt} = -\frac{3k_bT}{\zeta b^2}(\vec{R}_0 - \vec{R}_1) + \vec{\eta}_0$$

$$\frac{d\vec{R}_N}{dt} = -\frac{3k_bT}{\zeta b^2}(\vec{R}_N - \vec{R}_{N-1}) + \vec{\eta}_N$$

La posizione del centro di massa è data da

$$\vec{R}_{CM} = \frac{1}{N} \sum_{n=0}^N \vec{R}_n,$$

quindi l'equazione di Langevin per il centro di massa sarà

$$\frac{d\vec{R}_{CM}}{dt} = \frac{1}{N} \sum_{n=0}^N \frac{d\vec{R}_n}{dt} = \frac{1}{N} \left[-\sum_{n=1}^{N-1} \frac{3k_bT}{\zeta b^2}(2\vec{R}_n - \vec{R}_{n-1} - \vec{R}_{n+1}) - \frac{3k_bT}{\zeta b^2}(\vec{R}_0 - \vec{R}_1) - \frac{3k_bT}{\zeta b^2}(\vec{R}_N - \vec{R}_{N-1}) + \sum_{n=0}^N \vec{\eta}_n(t) \right].$$

Si nota che la prima somma e i due termini successivi si elidono tra loro:

$$\begin{aligned} & -\sum_{n=1}^{N-1} \frac{3k_bT}{\zeta b^2}(2\vec{R}_n - \vec{R}_{n-1} - \vec{R}_{n+1}) - \frac{3k_bT}{\zeta b^2}(\vec{R}_0 - \vec{R}_1) - \frac{3k_bT}{\zeta b^2}(\vec{R}_N - \vec{R}_{N-1}) = \\ & = -\frac{3k_bT}{\zeta b^2}(2\vec{R}_1 - \vec{R}_0 - \vec{R}_2 + 2\vec{R}_2 - \vec{R}_1 - \vec{R}_3 + \dots + 2\vec{R}_{N-1} - \vec{R}_{N-2} - \vec{R}_N + \vec{R}_0 - \vec{R}_1 + \vec{R}_N - \vec{R}_{N-1}). \end{aligned}$$

Rimane, dunque, solo il termine ²

$$\frac{d\vec{R}_{CM}}{dt} = \frac{1}{N} \sum_{n=0}^N \vec{\eta}_n(t) = \vec{\eta}(t). \quad (1.11)$$

²La forma matematicamente corretta della formula precedente è data dall'equazione incrementale $d\vec{R}_{CM} = \vec{\eta}(t) dt$.

Integrando l'equazione 1.11 ottengo

$$\vec{R}_{CM}(t) = \vec{R}_{CM}(0) + \int_0^t d\tau \frac{1}{N} \sum_n \vec{\eta}_n(\tau).$$

Posso, ora, calcolare la varianza come

$$\begin{aligned} \langle (\vec{R}_{CM}(t) - \vec{R}_{CM}(0))^2 \rangle &= \left\langle \int_0^t d\tau \int_0^t d\tau' \left(\frac{1}{N} \sum_n \vec{\eta}_n(\tau) \right) \cdot \left(\frac{1}{N} \sum_m \vec{\eta}_m(\tau') \right) \right\rangle = \\ &= \frac{1}{N^2} \int_0^t d\tau \int_0^t d\tau' \langle \left(\sum_{n,m} \vec{\eta}_n(\tau) \cdot \vec{\eta}_m(\tau') \right) \rangle. \end{aligned}$$

Ricordando che $\langle \vec{\eta}_n(\tau) \cdot \vec{\eta}_m(\tau') \rangle = 6D\delta(\tau - \tau')\delta_{n,m}$, si ottiene

$$\begin{aligned} \langle (\vec{R}_{CM}(t) - \vec{R}_{CM}(0))^2 \rangle &= 6D \frac{1}{N^2} \int_0^t d\tau \int_0^t d\tau' \left(\sum_{n,m} \delta(\tau - \tau') \delta_{n,m} \right) = \\ &= 6D \frac{1}{N} \int_0^t d\tau \int_0^t d\tau' (\delta(\tau - \tau')) = \\ &= 6D \frac{1}{N} \int_0^t d\tau 1 = 6D \frac{1}{N} t = 6D_{CM}t. \end{aligned}$$

Si osserva dunque che nel caso di un polimero descritto dal modello di Rouse il coefficiente diffusivo ha una dipendenza dal numero N di monomeri nella catena:

$$D_{CM} = \frac{D}{N}.$$

Più in generale la posizione del centro di massa di un polimero diffonde con un coefficiente di diffusione [10]

$$D_{CM} = \frac{D}{N^\alpha}$$

con α specifico del modello scelto per descrivere il polimero. Per il modello di Rouse si ha, dunque, $\alpha = 1$. Altri esempi importanti, di modelli polimerici sono [9] [4]:

- il modello di Zimm, in cui si introducono le interazioni idrodinamiche tra i monomeri, per cui vale $\alpha = \frac{1}{2}$;
- il modello Reptation, valido per sistemi polimerici densi ("polymer melts"), per il quale vale $\alpha = 2$.

1.3 Diffusione non gaussiana

In questa sezione consideriamo, per semplicità, un modello unidimensionale, quindi ci concentriamo solo sulla coordinata x .

Durante il processo di polimerizzazione, tramite il quale il polimero aggrega e disaggrega monomeri, il sistema è descrivibile tramite quello che in meccanica statistica si chiama ensemble gran canonico. Il modo più semplice per descrivere i processi di polimerizzazione e depolimerizzazione è tramite un processo di *birth-death*, assegnando alla polimerizzazione un rate λ e alla depolimerizzazione una frequenza μ [10].³

$$N(t) = n \quad \rightarrow \quad N(t + \tau) = \begin{cases} n + 1 & \text{con rate } \lambda \\ n - 1 & \text{con rate } \mu \end{cases} \quad \text{se } n > 1$$

$$N(t) = 1 \quad \rightarrow \quad N(t + \tau) = 2 \quad \text{con rate } \lambda$$

Questo fa sì che il numero di monomeri che costituiscono la catena polimerica sia variabile nel tempo: $N = N(t)$, con $t > \tau$, dove τ è il tempo medio che intercorre tra processi di polimerizzazione/depolimerizzazione consecutivi. Da cui si ha $D_{CM}(t) = \frac{D}{N(t)}$. R_{CM} evolve, dunque, come

$$d\vec{R}_{CM}(t) = \sqrt{2D_{CM}(N(t))} \, dW(t) \quad (1.12)$$

dove $W(t)$ è un moto Browniano standard, ossia un moto Browniano con distribuzione normale avente media 0 e varianza t ($W(t) = \mathbb{N}(0, t)$). Dalla 1.12 si osserva che ci sono due fonti di aleatorietà:

- La forza randomica, di cui si è parlato nel capitolo precedente, che genera il moto Browniano $W(t)$.
- I processi di polimerizzazione e depolimerizzazione che influenzano $W(t)$ tramite la dipendenza del coefficiente di diffusione da N . Tale seconda fonte di aleatorietà è chiamata subordinazione [5].

Si può osservare empiricamente che la distribuzione della posizione del CM di un polimero è non gaussiana. Essa, in particolare, ha delle code più lunghe ed è più piccata rispetto ad una distribuzione normale. Questa caratteristica distribuzione è dovuta proprio all'influenza che ha la subordinazione sul moto Browniano del polimero, a causa della dipendenza del coefficiente di diffusione da N .

Come visto nelle sezione 1.1.2, la distribuzione associata alla posizione di una particella con coefficiente diffusivo costante è una distribuzione normale.

Su una scala temporale breve rispetto al tempo di autocorrelazione del processo $N(t)$, posso pensare di descrivere la statistica del sistema in considerazione, come una sovrapposizione di polimeri, ciascuno con una grandezza fissata ($N = n$) e ciascuno caratterizzato da una distribuzione gaussiana con coefficiente di diffusione costante D_n . Si associa ad ogni polimero con taglia fissata, che costituisce il sistema, una probabilità $P_N(n)$, che rappresenta lo stato stazionario in cui la distribuzione di probabilità di N non varia più. Tale $P_N(n)$ dipende dalla funzione di partizione dell'ensemble gran canonico [10]. Nello specifico, definiamo $P_N(n) = (1 - p)p^{n-1}$, dove $p = \frac{\lambda}{\mu}$ [10]. Quindi, la distribuzione che descrive l'intero sistema è data da una mistura di gaussiane, che è, dunque, non gaussiana (come si vede nel grafico in figura 1.4):

$$p_{x_{CM}}(x, t) = \sum_n P_N(n) \frac{1}{\sqrt{4\pi D_n t}} \cdot e^{-\frac{x^2}{4D_n t}} \quad (1.13)$$

³I valori di λ e μ sono fissi poiché sto considerando il limite di campo medio [10].

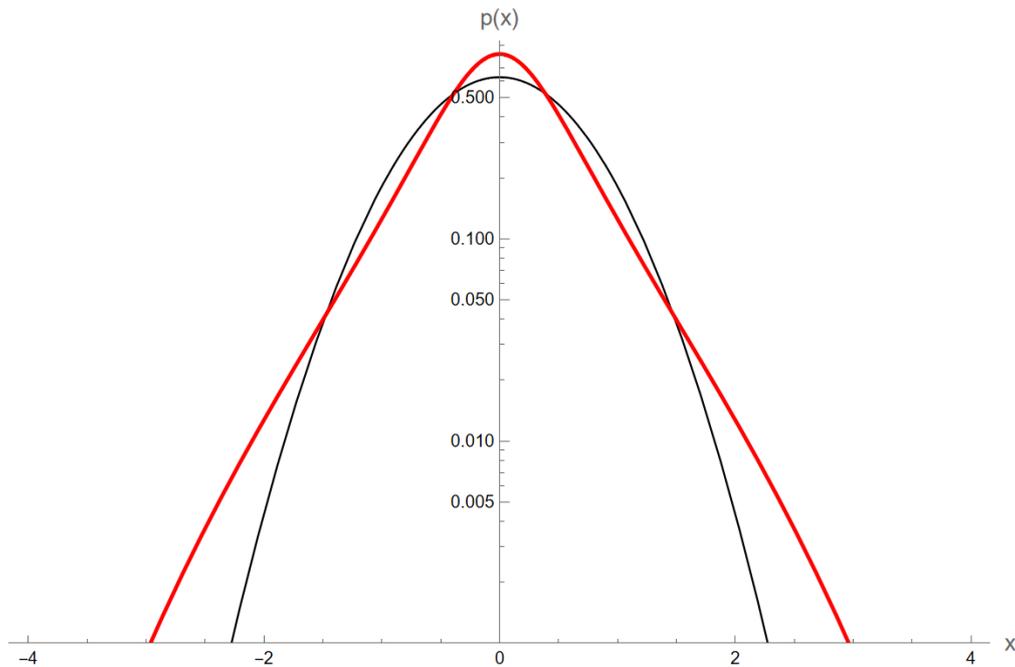


Figura 1.4: Nel grafico sono rappresentate in scala logaritmica: una distribuzione normale (in nero), con una varianza media, e una distribuzione generata da una mistura di gaussiane (in rosso), con pesi e varianze variabili come da formula 1.13 con $p = 0.8$.

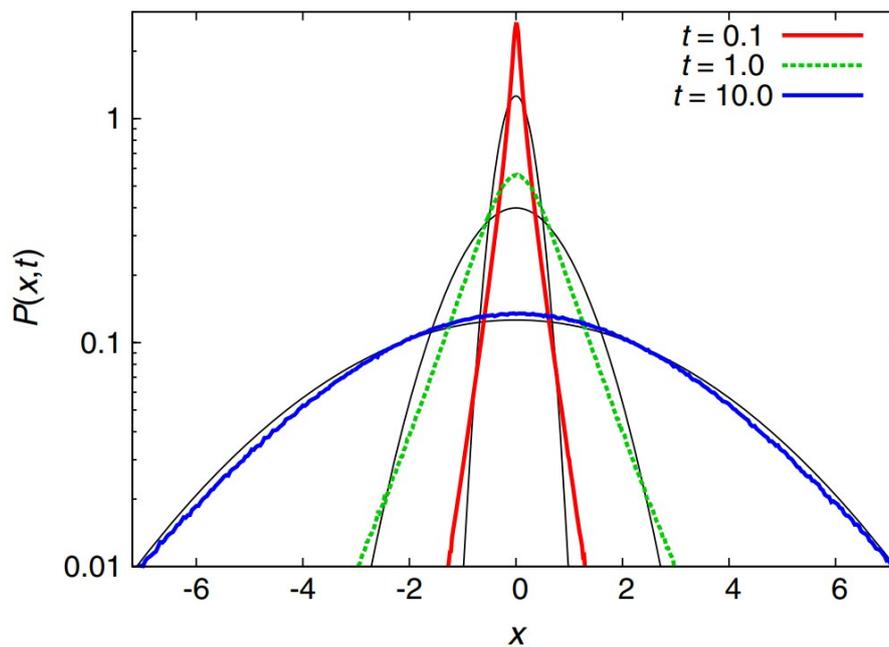


Figura 1.5: Nel grafico sono rappresentate le densità di probabilità $p_{x_{CM}}(x, t)$ in scala logaritmica, per tre diversi istanti di tempo t . Tali distribuzioni sono messe a confronto con le rispettive distribuzioni gaussiane (rappresentate in nero) [5].

Il grafico in figura 1.5, preso da Ref. [5], mostra che, per tempi lunghi, nello specifico $t \gg \tau$, la distribuzione tende sempre più alla gaussiana. Per tempi lunghi, infatti, la distribuzione totale tende alla distribuzione del singolo monomero con coefficiente diffusivo pari alla media dei coefficienti campionati [5]:

$$p_{x_{CM}}(x, t) \simeq \frac{1}{\sqrt{4\pi\langle D \rangle t}} \cdot e^{-\frac{x^2}{4\langle D \rangle t}}.$$

Capitolo 2

Dinamica microscopica dell'evoluzione del prezzo di un asset

2.1 Introduzione

In questo capitolo si discutono le caratteristiche empiriche della dinamica dei prezzi di asset finanziari, suffermandosi sulle similitudini che la accomunano alla dinamica diffusiva dei polimeri. Si propone un modello microscopico che possa giustificare tali similitudini empiriche e si costruisce un paragone tra i due sistemi.

Un asset finanziario è un bene che può essere posseduto, controllato e scambiato da un individuo, una società o un paese con lo scopo di ricavarne un eventuale profitto. Gli asset possono essere risorse materiali (come gli immobili, i beni primari o le azioni) o immateriali (come i brevetti o i copyright). I prezzi con cui tali beni vengono scambiati sono soggetti ad oscillazioni nel tempo che, come si vedrà, dipendono dalla domanda e dall'offerta.

2.2 Fatti Stilizzati

Lo studio delle variazioni di prezzo di un asset è limitato dal fatto che l'evoluzione di un mercato non sia riproducibile, inoltre i diversi mercati non sono necessariamente influenzati dagli stessi eventi. Come conseguenza, le distribuzioni e le dinamiche dei prezzi di diversi asset, provenienti da diversi mercati finanziari, possono avere diverse proprietà. Poichè non è possibile conoscere in anticipo tutte le variabili che modellano le evoluzioni di mercato, si assume che il processo di oscillazione dei prezzi sia un processo stocastico. A partire da tale ipotesi si nota una certa stabilità dei risultati (dalla ricostruzione delle probabilità empiriche e dei loro momenti), per cui mercati completamente diversi presentano caratteristiche qualitative simili tra loro. Tali caratteristiche empiriche prendono il nome di Fatti Stilizzati.

Nell'interpretazione dei risultati statistici che mi portano a definire i Fatti Stilizzati devo tener conto che si sono fatte alcune assunzioni [6] [8]:

- (Stazionarietà) Si assume la stabilità nel tempo di alcune proprietà statistiche: per ogni set $\{t_1, \dots, t_n\}$ di istanti di tempo e ogni intervallo T la distribuzione delle variazioni di prezzo $\{X_i(t_i, \tau)\}_i$ (dove τ è il tempo che intercorre prima di una variazione), è pari alla distribuzione di $\{X_i(t_i + T, \tau)\}_i$.

- (Ergodicità) La media e i momenti ricavati empiricamente da un singolo set di dati sufficientemente lungo convergono alle quantità che si otterrebbero avendo a disposizione l'ensemble che descrive il processo stocastico.

Per alcuni di questi fatti stilizzati esiste una corrispondenza nei modelli polimerici, ossia la dinamica dei polimeri obbedisce anch'essa ad alcune di queste caratteristiche empiriche. In particolare, per entrambi i sistemi si hanno le seguenti proprietà [6] [8]:

- (Distribuzione non Gaussiana) La distribuzione delle variazioni di prezzo hanno delle code più pronunciate e sono più piccate rispetto alla distribuzione gaussiana. Come già visto nel capitolo precedente, la non gaussianità si osservava anche nella distribuzione del CM polimerico.
- (Limite Gaussiano) Al crescere della scala temporale su cui sono calcolate le variazioni di prezzo la distribuzione tende alla distribuzione normale. Lo stesso accade per il sistema polimerico [10], come visto anche dal grafico in figura 1.5.
- (Dipendenza tra volatilità e volume di ordini) Le oscillazioni dei prezzi sono correlate al numero degli ordini in un mercato, così come la varianza della distribuzione polimerica era legata al numero N di monomeri nella catena.

2.3 Diffusione non Gaussiana

In questa sezione ci si vuole concentrare sulla non gaussianità. In particolare si osserva la distribuzione dei ritorni logaritmici del DJI (Dow Jones Index) [2], un indice che considera il prezzo dei principali titoli di Wall Street.

Il ritorno logaritmico è definito come

$$r(t, T) = \log \frac{s(t+T)}{s(t)},$$

in cui $s(t)$ è il prezzo al tempo t . È utilizzato poichè permette di stimare il rate Λ di crescita esponenziale del prezzo di un asset sicuro:

$$r(t, T) = \log \frac{s(t+T)}{s(t)} = \log \frac{A e^{\Lambda(t+T)}}{A e^{\Lambda t}} = \log e^{\Lambda T} = \Lambda T,$$

dove A è una costante. Poichè si assume che gli asset sicuri abbiano una crescita esponenziale, ci si aspetta che il prezzo di un asset qualsiasi abbia la forma $s(t) = A e^{\Lambda t + \eta(t)}$, dove $\eta(t)$ è un termine stocastico. Quindi, prendendo il logaritmo di tale valore compare un termine additivo che descrive un moto Browniano: $\log s(t) = B + \Lambda t + \eta(t)$, dove B è una costante.

Dal grafico in figura 2.1 è evidente che la distribuzione dei ritorni logaritmici si discosta dal fit parabolico, ossia dal fit gaussiano in scala logaritmica.

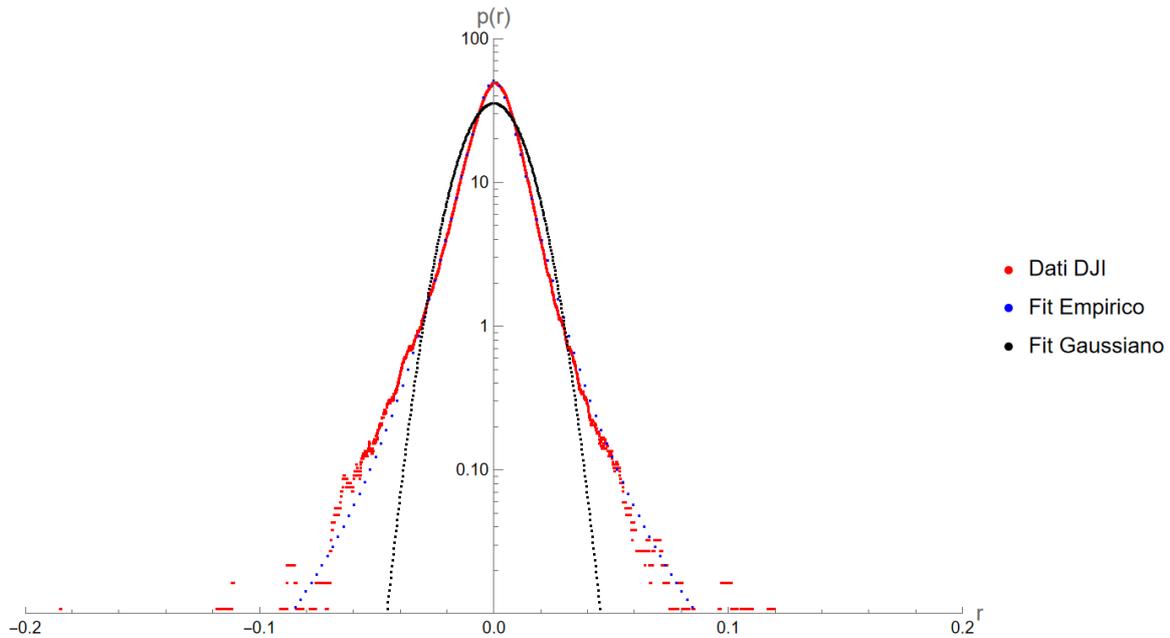


Figura 2.1: Nel grafico è rappresentata la distribuzione dei ritorni logaritmici (in rosso), comparata con un fit analitico non gaussiano e con un fit parabolico

Per evidenziare le caratteristiche empiriche che accomunano la dinamica polimerica e la dinamica dei prezzi descriviamo ora il sistema del *limit order book* per descrivere i mercati. Tale sistema, e nello specifico il modello Markoviano riportato in Ref. [7], mi permette, infatti, di costruire un paragone con il modello polimerico. Si vedrà come i due sistemi siano descritti da dinamiche microscopiche simili.

2.4 Limit Order Book

Il *Limit Order Book* è un file che tiene traccia di tutti gli ordini sul mercato con le caratteristiche di ciascuno [1]. Nello specifico gli ordini si distinguono in base:

- alla tipologia: *ask* (domanda) e *bid* (offerta)
- al prezzo
- alla quantità
- al tempo di registrazione dell'ordine

Lo studio delle interazioni che avvengono all'interno del *limit order book*, che si basano sull'accumulo e sull'esecuzione degli ordini di mercato, permette di comprendere la dinamica che guida l'oscillazione dei prezzi degli asset finanziari. I mercati finanziari che si avvalgono di tale file sono detti *order-driven market*.

I mercati si avvalgono di tre tipologie di ordini:

- I *limit orders*, ossia gli ordini che si accumulano nell'order book fissandosi ad un certo prezzo di vendita o di acquisto.
- Gli ordini di mercato, ossia gli ordini che devono essere eseguiti immediatamente al miglior prezzo di acquisto o vendita di una certa quantità di un asset.
- Gli ordini di cancellazione, che annullano un *limit order* precedente.

Gli ordini sono immagazzinati nell'*order book* fino a quando non vengono eseguiti o cancellati. Chiamo *best ask* s_t^a il miglior prezzo di vendita ad un certo tempo t (prezzo di vendita più basso) e chiamo *best bid* s_t^b il miglior prezzo di acquisto (prezzo di acquisto più alto). La differenza tra *best ask* e *best*

bid si chiama *spread* $\Delta = s_t^a - s_t^b$.

In questa tesi si considera un modello di *order-book* detto *zero-intelligence*, in cui gli agenti operano in modo randomico, senza attenersi a strategie di mercato.

Si considera, dunque, che la variazione del prezzo di un asset è un fenomeno aleatorio legato all'arrivo e all'esecuzione degli ordini nell'*order book* (eventi a loro volta randomici).

Si introducono alcuni concetti utili:

- In finanza si definisce la Volatilità come un valore che quantifica l'oscillazione dei prezzi di un determinato asset. In particolare, essa è definita come la varianza della distribuzione dei prezzi.
- Si definisce la *Depth* come una misura che quantifica il numero degli ordini accumulati nelle colonne di *best bid* e *best ask* (vedere sotto).

Nei modelli più semplici, il *limit order book* viene rappresentato tramite un asse su cui si accumulano gli ordini posizionandosi al valore di prezzo corrispondente. In questo modo si vengono a creare due lati: a prezzi inferiori si accumula l'offerta, mentre a prezzi superiori si accumula la domanda.

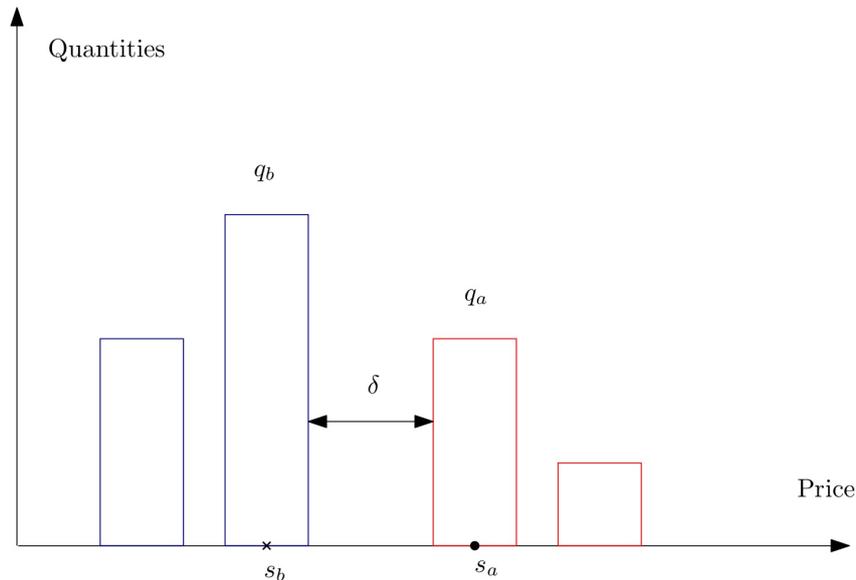


Figura 2.2: Sull'asse delle ascisse si posizionano i prezzi in ordine crescente, mentre sull'asse delle ordinate si ha la quantità di ordini. In blu è rappresentata l'offerta e in rosso la domanda [7].

L'altezza delle colonne rappresenta l'accumulo di ordini ad un determinato prezzo. In particolare chiamo q_t^a e q_t^b il numero di ordini corrispondenti a *best bid* e *best ask*.

2.5 Modello Markoviano dell'Order Book

Si introduce, in questa sezione, il modello Markoviano trattato nell'articolo in Ref. [7], che descrive le modalità di arrivo ed esecuzione degli ordini.

Il modello si basa sull'ipotesi che per analizzare le oscillazioni dei prezzi sia sufficiente concentrarsi sull'accumulo e sull'esecuzione degli ordini corrispondenti alle colonne di *best bid* e *best ask*. Considero, dunque, $q_t = (q_t^b, q_t^a)$ il numero di ordini accumulati in tali colonne. Le colonne restanti, posizionate a prezzi inferiori e superiori, rappresentano il reservoir del mio sistema. I *limit order* si accumulano nell'*order book* aumentantando il valore di q_t^b o di q_t^a in base al segno dell'ordine. L'arrivo di un

market order, invece, viene fatto corrispondere all'esecuzione di un *limit order* nella colonna di segno opposto, ossia un ordine di vendita eliminerà un ordine di acquisto e viceversa. In questo modello semplificato ignoreremo gli ordini di cancellazione.

Gli ordini, inoltre, arrivano seguendo distribuzioni di Poisson indipendenti:

- I *limit order* arrivano a tempi indipendenti ed esponenziali, con una rate λ
- I *market order* arrivano a tempi indipendenti ed esponenziali, con una rate μ

Si ha, dunque, un processo di *birth-death* per le colonne di *best bid* e *best ask*, come si era ipotizzato anche per il modello polimerico.

Chiamando V_i^a e V_i^b le variazioni di taglia delle colonne, si ha che $(V_i^a)_{i \geq 0}$ e $(V_i^b)_{i \geq 0}$ sono sequenze di variabili indipendenti e randomiche tali che

$$\mathbb{P}[V_i = 1] = \frac{\lambda}{\lambda + \mu} \quad \text{and} \quad \mathbb{P}[V_i = -1] = \frac{\mu}{\lambda + \mu} \quad (2.1)$$

Si assume, poi, che i prezzi di domanda e offerta siano discreti e multipli dello *spread* Δ , ossia le tacche sull'asse delle ascisse nel grafico 2.2 distinto tra loro Δ , cosa che si verifica nel 98% delle osservazioni [7]. Dove lo *spread* Δ è una quantità fissa tale che $s_t^a = \Delta + s_t^b$.

L'esaurirsi di una colonna viene fatto corrispondere ad una diminuzione (se $q_t^b = 0$) o all'aumentare (se $q_t^a = 0$) del prezzo di una tacca pari a Δ . La nuova colonna di *best bid* o *best ask* sarà la colonna immediatamente precedente o successiva: in corrispondenza di una diminuzione o un aumento di prezzo si ha un nuovo valore di (q_t^b, q_t^a) a cui si assegna un valore randomico che segue una distribuzione f . $f(x, y)$ è, quindi, la probabilità di avere $(q_t^b, q_t^a) = (x, y)$ dopo una diminuzione o un aumento di prezzo.

Nota: Assegnando un valore randomico alle colonne di *best bid* e *best ask*, quando il prezzo aumenta o diminuisce, si perde memoria del size iniziale, ossia stiamo assumendo l'indipendenza degli incrementi, nonostante questo non sia vero empiricamente [8]. Tale approssimazione, inoltre, discosta il sistema descritto dal sistema polimerico, in cui la memoria del size iniziale rimane per un intervallo di tempo che tende ad infinito in prossimità del punto critico.

Identifichiamo la variabile s_t senza apice come il prezzo di un asset in modo generico e non lo specifico prezzo della miglior domanda o offerta. Tale valore dipende, dal prezzo di domanda e offerta e, per convenzione, possiamo considerarlo pari ad una media dei due.

Consideriamo il processo stocastico $(s_t, t \geq 0)$, in cui il valore di t è definito e continuo da sinistra a destra per tutti i valori di t . Tale processo è costituito da "salti" che corrispondono ai momenti in cui il processo stocastico $(q_t, t \geq 0)$ tocca il valore 0. Denotiamo come $(\tau_1, \tau_2, \tau_3, \dots)$ le durate prima di successivi cambi di prezzo. Il numero di variazioni di prezzo in un intervallo di tempo t sarà dunque

$$N_t := \max\{n \geq 0; \tau_1 + \tau_2 + \dots + \tau_n \leq t\}$$

Chiamiamo $(X_1, X_2, X_3, \dots, X_n, \dots)$ le successive variazioni unitarie, che possono assumere i valori ± 1 . Tali variabili moltiplicate per Δ identificano le variazioni di prezzo. Assumendo che $f(i, j) = f(j, i) \quad \forall i, j$, si ha che le variazioni di prezzo $\{X_i\}_i$ sono variabili randomiche e indipendenti con media 0. Più in generale, si sta assumendo che $P[X_{n+1} = \Delta | X_n = \Delta] = P[X_{n+1} = -\Delta | X_n = -\Delta] = \frac{1}{2}$.

Definiamo, infine, $Z(n) = s_0 + \Delta \sum_{i=1}^n X_i$, che sarà, dunque, il prezzo corrente dopo n variazioni di prezzo. Si avrà $s_t = Z(N_t)$.

2.6 Corrispondenze tra sistema polimerico e *order book*

A partire dalle caratteristiche empiriche osservate e dal modello proposto, si può costruire un paragone tra il modello Markoviano del *limit order book* e un modello polimerico (come quello di Rouse descritto in sezione 1.2). Si osservano le seguenti corrispondenze:

- Il centro di massa del polimero viene associato al prezzo di un determinato asset finanziario.
- Entrambi (centro di massa e prezzo) seguono un processo diffusivo.
- I monomeri che costruiscono la catena polimerica, possono essere paragonati agli ordini che si accumulano nelle colonne di *best bid* e *best ask* dell'*order book*.
- Come conseguenza, l'arrivo di un ordine viene paragonato al processo di polimerizzazione e l'esecuzione di un ordine al processo di depolimerizzazione. In entrambi i sistemi, posso modellizzare tali processi con un processo di *birth-death*, assegnando una frequenza di accumulo (polimerizzazione) λ e una frequenza di esecuzione (depolimerizzazione) μ .
- Quindi si può paragonare il numero degli ordini in corrispondenza di *best bid* e *best ask* (quantificato dalla *depth*), al numero di monomeri che costituiscono la catena polimerica.

2.7 Dipendenza dal "depth" del comportamento diffusivo del prezzo di un asset

Si vuole, ora, giustificare teoricamente la dipendenza della volatilità dei prezzi dal volume di ordini all'interno dell'*order book*. In particolare, voglio stabilire se il modello Markoviano descritto possa essere associato al modello di Rouse o ad un altro modello polimerico, sulla base della relazione che vige tra varianza e numero di ordini.

A partire dal modello microscopico descritto in sezione 2.5, posso studiare la dinamica dei prezzi. Tipicamente avviene che, nell'arco di alcune decine di secondi, si verifichino migliaia di eventi all'interno dell'*order book*. Questo mi permette di studiare il modello nel suo limite diffusivo ($n \rightarrow \infty$).

Si stabilisce, dunque, un link tra la dinamica ad alte frequenze descritta dal modello microscopico e il limite diffusivo per lunghe scale di tempo, con cui si possa ricavare un'espressione teorica per la volatilità.

Introduco, prima, il concetto di convergenza in distribuzione [11].

Date delle variabili reali randomiche Y_n , $n \geq 1$ e Y , vi è convergenza in distribuzione per $n \rightarrow \infty$

$$Y_n \Rightarrow Y$$

se la funzione di distribuzione cumulata associata converge per $n \rightarrow \infty$:

$$P(Y_n \leq y) \rightarrow P(Y \leq y).$$

Considero, ora, la scala di tempo $t_n = tg(n)$. La scelta di tale scala non è arbitraria ma è guidata dalle proprietà di distribuzione degli intervalli di tempo $\{\tau_i\}_i$ tra due variazioni di prezzo successive. In particolare deve valere

$$\exists \text{ finito } \lim_{n \rightarrow \infty} \frac{\tau_1 + \dots + \tau_n}{g(n)}.$$

Eseguo, poi, un riscaldamento del processo stocastico

$$(s_t^n := \frac{s_{t_n}}{\sqrt{n}}; t \geq 0)_{n \geq 1}.$$

Tale riscaldamento fa sì che il processo verifichi un teorema funzionale del limite centrale, ossia che esso converga in distribuzione al processo $(b_t; t \geq 0)$ quando $n \rightarrow \infty$. Tale b_t è un processo diffusivo. Considero, infine, una misura della *depth*, ossia una funzione che dipenda dal numero degli ordini all'interno delle due colonne di *best bid* e *best ask*:

$$\mathcal{D}(f) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} ijf(i, j), \quad (2.2)$$

dove si ricorda che $f(i, j)$ è la probabilità che $(q_t^b, q_t^a) = (i, j)$, dopo una variazione di prezzo.¹

Si osserva che $\sqrt{\mathcal{D}(F)}$ rappresenta la media geometrica tra il numero di ordini nella colonna di *best bid* e il numero di ordini nella colonna di *best ask*.

Si analizza, adesso, il caso in cui il flusso di ordini sia bilanciato, ossia il caso $\lambda = \mu$. In questo caso, si può provare [7] che la distribuzioni dei tempi che intercorrono tra due variazioni successive di prezzo è data da

$$P[\tau > t | q_0^b = i, q_0^a = j] \sim \frac{ij}{\pi \lambda t},$$

quindi

$$P[\tau > t] \stackrel{t \rightarrow \infty}{\sim} \frac{\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} ijf(i, j)}{\pi \lambda t} = \frac{\mathcal{D}(f)}{\pi \lambda t}.$$

Dimostriamo che la serie di tempi $\{\tau_i\}_i$ è una sequenza di variabili randomiche identicamente distribuite per cui vale:

$$\frac{\tau_1 + \tau_2 + \dots + \tau_n}{n} \stackrel{n \rightarrow \infty}{\Rightarrow} \frac{\mathcal{D}(f)}{\pi \lambda e}. \quad (2.3)$$

Dimostrazione.

Considero τ una generica τ_i . La distribuzione di $\frac{\tau}{n}$ è data da:

$$p_{\frac{\tau}{n}}(t) = \int_0^{+\infty} dt' p_{\tau}(t') \delta(t - \frac{t'}{n}) = [t'' = \frac{t'}{n}, dt' = dt''n] = \int_0^{+\infty} dt'' n p_{\tau}(t''n) \delta(t - t'') = p_{\tau}(nt) n.$$

La sua trasformata è, dunque,

$$\begin{aligned} \tilde{p}_{\frac{\tau}{n}}(s) &= \int_0^{+\infty} dt e^{-st} p_{\frac{\tau}{n}}(t) = \\ &= \int_0^{+\infty} dt e^{-st} n p_{\tau}(tn) = [dt = \frac{dt'}{n}, t' = nt] = \\ &= \int_0^{+\infty} dt' e^{-\frac{st'}{n}} p_{\tau}(t') = \tilde{p}_{\tau}\left(\frac{s}{n}\right). \end{aligned}$$

Sapendo che

$$p_{\tau}(t) = \frac{d}{dt}[1 - P(\tau > t)] = -\frac{d}{dt}P(\tau > t),$$

dove $P(\tau > t)$ è la survival probability,

$$\begin{aligned} \tilde{p}_{\tau}(s) &= \int_0^{+\infty} dt e^{-st} \left[-\frac{d}{dt}P(\tau > t)\right] = \\ &= -\int_0^{+\infty} dt \left\{ \frac{d}{dt}[e^{-st}P(\tau > t)] + se^{-st}P(\tau > t) \right\} = 1 - s \tilde{p}_{P(\tau > t)}(s), \end{aligned}$$

¹Perchè la 2.2 converga deve valere $\lambda \leq \mu$

da cui

$$\tilde{p}_n(s) = \tilde{p}_\tau\left(\frac{s}{n}\right) = 1 - \frac{s}{n} \tilde{p}_{P(\tau>t)}\left(\frac{s}{n}\right).$$

La trasformata di $P(\tau > t)$ in funzione di $(\frac{s}{n})$ è data da

$$\tilde{p}_{P(\tau>t)}\left(\frac{s}{n}\right) = \int_0^{+\infty} dt e^{-\frac{s}{n}t} P(\tau > t) = \int_0^{+\infty} dt e^{-\frac{s}{n}t + \log P(\tau>t)} \stackrel{n \gg 1}{\simeq} \int_0^{t_0} dt = t_0,$$

dove t_0 è tale che

$$-\frac{s}{n}t_0 + \log\left(\frac{\mathcal{D}(f)}{\pi\lambda t_0}\right) = 1 \quad \Leftrightarrow \quad e^{-\frac{s}{n}t_0} \frac{\mathcal{D}(f)}{\pi\lambda t_0} = e.$$

Considerando il cambio di variabile $w = \frac{s}{n}t_0$, si trova

$$e^{-w} \frac{\mathcal{D}(f)s}{\pi\lambda n w} = e \quad \Leftrightarrow \quad we^w = \frac{\mathcal{D}(f)s}{\pi\lambda n e}.$$

Dall'ultima equazione si ha, dunque, $w = W_0\left(\frac{\mathcal{D}(f)s}{\pi\lambda n e}\right)$, dove W_0 è la funzione di Lambert per la quale vale $W_0(x) \stackrel{x \ll 1}{\simeq} x - x^2$. Dunque

$$t_0 \stackrel{n \gg 1}{\simeq} \frac{n}{s} \left(\frac{\mathcal{D}(f)s}{\pi\lambda n e} - \mathcal{O}\left(\frac{s^2}{n^2}\right) \right) = \frac{\mathcal{D}(f)}{\pi\lambda e} - \mathcal{O}\left(\frac{s}{n}\right),$$

da cui si ottiene

$$\tilde{p}_n(s) = 1 - \frac{s}{n} \frac{\mathcal{D}(f)}{\pi\lambda e} + \mathcal{O}\left(\left(\frac{s}{n}\right)^2\right).$$

Per la proprietà della trasformata si ha

$$\tilde{p}_{\frac{\tau_1+\dots+\tau_n}{n}}(s) = (\tilde{p}_n(s))^n = \left[1 - \frac{s}{n} \frac{\mathcal{D}(f)}{\pi\lambda e} + \mathcal{O}\left(\left(\frac{s}{n}\right)^2\right)\right]^n \stackrel{n \gg 1}{\simeq} e^{-s \frac{\mathcal{D}(f)}{\pi\lambda e}}.$$

Infine, facendo l'antitrasformata di Laplace, si ottiene

$$p_{\frac{\tau_1+\dots+\tau_n}{n}}(t) = \frac{1}{2\pi i} \int_{\mathcal{L}} ds e^{st} \tilde{p}_{\frac{\tau_1+\dots+\tau_n}{n}}(s) = \frac{1}{2\pi i} \int_{\mathcal{L}} ds e^{st} e^{-s \frac{\mathcal{D}(f)}{\pi\lambda e}} = \delta\left(t - \frac{\mathcal{D}(f)}{\pi\lambda e}\right).$$

Quindi

$$\frac{\tau_1 + \tau_2 + \dots + \tau_n}{n} \stackrel{n \rightarrow \infty}{\Rightarrow} \frac{\mathcal{D}(f)}{\pi\lambda e}.$$

□

Si usa, dunque, il riscalamento $t \rightarrow t_n = tn$, per cui N_{t_n} è il numero di variazioni di prezzo tra $[0, t_n]$. Quindi, la 2.3 stabilisce

$$\frac{t_n}{N_{t_n}} \stackrel{n \rightarrow \infty}{\simeq} \frac{\mathcal{D}(f)}{\pi\lambda e},$$

dunque vale

$$N_{t_n} \sim \frac{\pi\lambda e t}{\mathcal{D}(f)} n \quad \text{per } n \rightarrow \infty.$$

Nell'ultima equazione si osserva che il numero di variazioni di prezzo è proporzionale a $\frac{\lambda}{\mathcal{D}(f)}$. Tale relazione fa già intuire che la volatilità possa anch'essa dipendere dalla *depth*, si ricorda, infatti, che la volatilità quantifica le oscillazioni di prezzo.

Dimostriamo, che se $\lambda = \mu$, si ottiene un limite diffusivo per il processo $(\frac{st_n}{\sqrt{n}}; t \geq 0)$:

$$\left(\frac{st_n}{\sqrt{n}}; t \geq 0\right) \stackrel{n \rightarrow \infty}{\Rightarrow} \left(\Delta \sqrt{\frac{\pi\lambda}{\mathcal{D}(f)}} W_t; t \geq 0\right),$$

dove W è un moto Browniano standard, ossia un moto Browniano con media 0 e varianza pari a t .

Dimostrazione.

Sia $t_n = tn$, posso scrivere il processo $\frac{s_{t_n}}{\sqrt{n}}$ come segue:

$$\frac{s_{t_n}}{\sqrt{n}} = \frac{Z(\frac{tn\pi\lambda e}{\mathcal{D}(f)}) \Delta}{\sqrt{n}} + \left(\frac{Z(N_{t_n}) \Delta}{\sqrt{n}} - \frac{Z(\frac{tn\pi\lambda e}{\mathcal{D}(f)}) \Delta}{\sqrt{n}} \right).$$

Dato che $\{X_i\}_i$ sono variabili casuali, indipendenti e identicamente distribuite con media 0, si può applicare il teorema funzionale del limite centrale per cui

$$\left(\frac{Z(\frac{tn\pi\lambda e}{\mathcal{D}(f)}) \Delta}{\sqrt{n}}, t \geq 0 \right)_{n \geq 1} \xrightarrow{n \rightarrow \infty} (\Delta \sqrt{\frac{\pi\lambda e}{\mathcal{D}(f)}} W_t; t \geq 0).$$

Infine, ricordando che

$$N_{t_n} \sim \frac{\pi\lambda e t}{\mathcal{D}(f)} n \quad \text{per } n \rightarrow \infty$$

si osserva che

$$\left(\frac{Z(N_{t_n}) \Delta}{\sqrt{n}} - \frac{Z(\frac{tn\pi\lambda e}{\mathcal{D}(f)}) \Delta}{\sqrt{n}} \right) \xrightarrow{n \rightarrow \infty} 0.$$

Si ottiene, quindi,

$$\left(\frac{s_{t_n}}{\sqrt{n}}; t \geq 0 \right) \xrightarrow{n \rightarrow \infty} (\Delta \sqrt{\frac{\pi\lambda e}{\mathcal{D}(f)}} W_t; t \geq 0).$$

□

Il processo aleatorio considerato, quindi, converge in distribuzione ad un moto Browniano con varianza

$$\Delta^2 \left(\frac{\pi\lambda e}{\mathcal{D}(f)} \right) t. \tag{2.4}$$

Si ottiene, dunque, una stima per la volatilità che presenta una dipendenza dalla funzione $\mathcal{D}(f)$, che è di fatto una misura della *depth* dell'*order book*. Si osserva un coefficiente diffusivo della forma

$$D_s \propto \frac{1}{\mathcal{D}(f)}.$$

Si conclude che, dal modello microscopico descritto nella sezione precedente, si ottiene la dipendenza tra varianza e numero di elementi che si osservava empiricamente in finanza e che era presente anche nella formulazione del modello di Rouse e di altri modelli polimerici. In particolare, poichè il numero di ordini è rappresentato da $\sqrt{\mathcal{D}(f)}$, il modello Markoviano descritto è meglio comparabile con il modello Reptation per cui $\alpha = 2$.

Si sottolinea, infine, che la proporzione ottenuta, in formula 2.4, tra la varianza della distribuzione dei prezzi riscaldati e la frazione $\frac{\lambda}{\mathcal{D}(f)}$ è testabile empiricamente. Dai dati del DJI raccolti nel 2008, si è osservato, infatti, un andamento che poteva essere ben approssimato da un fit lineare [7].

Conclusioni

Dopo aver descritto un modello polimerico semplice, si è osservato che la dinamica del centro di massa di un polimero e la dinamica dei prezzi di un asset finanziario sono accomunate da alcune caratteristiche osservabili empiricamente. Si è, dunque, descritto il sistema dell'*order book* e si è discusso il modello Markoviano descritto nell'articolo in Ref. [7], dal quale si è potuto costruire il paragone con il sistema polimerico.

Infine, nell'ultima sezione, si è osservato come lo studio della dinamica microscopica dell'*order book* si sia rivelata di grande importanza per l'analisi delle oscillazioni dei prezzi degli asset all'interno dei mercati finanziari. Abbiamo, infatti, mostrato teoricamente la dipendenza tra volatilità e *depth*, che già si osservava empiricamente.

Più nello specifico abbiamo visto che, tra due mercati bilanciati ($\lambda = \mu$) con le stesse frequenze di accumulo ed esecuzione degli ordini, la volatilità maggiore appartiene al mercato con *depth* inferiore. Tale relazione tra volatilità e *depth* ricalca la dipendenza tra varianza della distribuzione del CM di un polimero e il numero di monomeri nella catena. In particolare, si è stabilito che il modello Markoviano descritto ricalca il modello polimerico Reptation e non il semplice modello di Rouse.

Esistono, tuttavia, una serie di limitazioni nel modello che abbiamo discusso, come può essere la perdita di memoria del size iniziale. Un possibile sviluppo futuro potrebbe essere rappresentato, dunque, da una modifica del modello che ci permetta di introdurre questa autocorrelazione temporale. Una modifica di questo tipo ci permetterebbe di inserire aspetti di non gaussianità per la distribuzione dei prezzi e di rendere, quindi, il modello più realistico e più in linea con il sistema polimerico.

Un'ulteriore prospettiva futura potrebbe essere quella di inserire nel modello un campo esterno che possa giustificare le asimmetrie empiriche della distribuzione dei prezzi e che possa essere associato, nel caso polimerico, ad un campo gravitazionale che genera sedimentazione. Si potrebbe, ad esempio, considerare uno squilibrio nel rate di accumulo λ degli ordini tra domanda e offerta.

Infine, si potrebbe tener conto del fatto che, poichè un *limit order* può posizionarsi a diversi valori di prezzo, il numero di ordini accumulati in colonne a prezzi inferiori o superiori rispetto a *best bid* e *best ask* influenzi anch'esso la dinamica.

Appendici

Appendice A: scala di tempo di Smoluchowski

Se trascuro il campo di forze esterne, l'equazione 1.4 diventa

$$\frac{d\vec{v}(t)}{dt} = -\frac{\zeta\vec{v}(t)}{m} + \frac{\vec{f}(t)}{m}.$$

Risolvendo formalmente questa equazione si ottiene

$$\vec{v}(t) = \vec{v}_0 e^{-\frac{\zeta t}{m}} + \int_0^t dt' e^{-\frac{\zeta(t-t')}{m}} \frac{\vec{f}(t')}{m},$$

da cui

$$\langle \vec{v}(t)^2 \rangle = \langle \vec{v}(t) \cdot v(t) \rangle = \vec{v}_0^2 e^{-\frac{2\zeta t}{m}} + 2 \int_0^t dt' e^{-\frac{\zeta(2t-t')}{m}} \frac{\langle \vec{f}(t') \rangle}{m} + \int_0^t dt' \int_0^t dt'' e^{-\frac{\zeta(2t-t'-t'')}{m}} \frac{\langle \vec{f}(t') \cdot f(t'') \rangle}{m^2}.$$

Sapendo che $\langle \vec{f}(t) \rangle = 0$ (per simmetria) e $\langle \vec{f}(t') \cdot \vec{f}(t'') \rangle = C_{v_0} \delta(t' - t'')$ (per indipendenza dei passi nella *random walk*) e utilizzando il teorema di Equipartizione per cui $\lim_{t \rightarrow \infty} \langle \vec{v}(t)^2 \rangle = \frac{3k_b T}{m}$, si ha

$$\langle \vec{v}(t)^2 \rangle = \vec{v}_0^2 e^{-\frac{2\zeta t}{m}} + \frac{3k_b T}{m} (1 - e^{-\frac{2\zeta t}{m}}).$$

Si nota, dunque, che in un tempo $\sim \frac{m}{\zeta}$, chiamato scala di Smoluchowsky, si perde la memoria della velocità iniziale v_0 .

Inoltre, per tale scala di tempo, vale

$$\frac{v_0 \cdot \frac{m}{\zeta}}{R} \sim 10^{-3} - 10^{-5},$$

dove $R \sim 1 - 3 nm$ è il raggio medio di una particella e $v_0 \sim \sqrt{\frac{3k_b T}{m}}$ per il teorema di Equipartizione dell'energia. Allora lo spostamento della particella durante tale scala temporale è trascurabile.

Appendice B: codici in Mathematica

Random Walk 1D

Definizione della variabile:

$$a = \{1, -1\}$$

Plot del grafico:

```
ListLinePlot[Flatten[{{0}}, Accumulate[a[[RandomInteger[{1, 2}, 100]]]]],  
  AxesLabel -> {"N", "Posizione"}, PlotStyle -> Red, AxesOrigin -> {1, 0}]
```

Random Walk 3D

Plot del grafico:

```
ListLinePlot3D [Table [Accumulate [RandomChoice[{{-1, 0, 0}, {0, -1, 0}, {0, 0, -1}, {1, 0, 0}, {0, 1, 0}, {0, 0, 1}}, 100]], 5], BoxRatios -> Automatic, AxesLabel -> {"x", "y", "z"}]
```

Mistura di gaussiane

Definizione della mistura:

```
f [N_, x_, p_] :=  
  Total [Table [((1 - p)*p^(n - 1))*  
    PDF [NormalDistribution [0, n^(1/2)], x], {n, 1, Floor [N]}]]
```

Plot del grafico:

p=0.6

```
LogPlot [{  
  PDF [NormalDistribution [0, Sqrt [Sum [1/n (1 - p) p^(n - 1), {n, 1, 50}]]], x],  
  f [50, x, 0.6]}, {x, -3, 3}, PlotRange -> {0.005, 0.95},  
  PlotStyle -> {Directive [Black, Thickness [0.002]],  
    Directive [Red, Thickness [0.004]]}, ImageSize -> Large,  
  AxesLabel -> {Style ["x", 13], Style ["p(x)", 13]}]
```

Plot DJI

Import dei dati:

```
dati1 = Import ["PrimoSet.dat"];  
dati2 = Import ["SecondoSet.dat"];  
dati3 = Import ["TerzoSet.dat"];
```

Plot del grafico:

```
ListLogPlot [{dati1, dati2, dati3}, PlotStyle -> {Green, Red, Blue},  
  PlotRange -> {{-0.2, 0.2}, {0.001, 100}},  
  PlotLegends -> {"Dati DJI", "Fit Empirico", "Fit Gaussiano"},  
  ImageSize -> Large, PlotLabel -> "",  
  AxesLabel -> {Style ["r", 13], Style ["p(r)", 13]}]
```

Bibliografia

- [1] Frédéric Abergel et al. *Limit Order Books*.
- [2] F. Baldovin e A. L. Stella. «Scaling and efficiency determine the irreversible evolution of a market». In: *PNAS* (2007).
- [3] Fulvio Baldovin. «Biological Physics Lecture Notes A.Y. 2022/2023». Dipartimento di Fisica and Sezione INFN, Università di Padova.
- [4] W.J. Briels. *Theory of Polymer Dynamics*. 1998.
- [5] Aleksei V. Chechkin et al. «Brownian yet Non-Gaussian Diffusion: From Superstatistics to Subordination of Diffusing Diffusivities». 2017.
- [6] Rama Cont. «Empirical properties of asset returns: stylized facts and statistical issues». In: *Quantitative Finance* (2001).
- [7] Rama Cont e Adrien de Larrard. «Price Dynamics in Markovian Limit Order Market». In: *SIAM Journal on Financial Mathematics* (2013).
- [8] Rama Cont e Peter Tankov. *Financial Modelling With Jump Processes*. CHAPMAN HALL/CRC FINANCIAL MATHEMATICS SERIES, 2004.
- [9] M. Doi e S. F. Edwards. *The theory of polymer dynamics*. the Clarendon Press, Oxford University Press, 1986.
- [10] Sankaran Nampoothiri et al. «Brownian non-Gaussian polymer diffusion and queuing theory in the mean-field limit». In: *New Journal of Physics* (2022).
- [11] Ward Whitt. *Stochastic-Process Limits*. Springer, 2002.