



UNIVERSITÀ DEGLI STUDI DI MILANO
FACOLTÀ DI SCIENZE E TECNOLOGIE

Dipartimento di Fisica
Corso di Laurea Triennale in Fisica

**Matrici di Wishart e fluttuazioni della
conduttanza in mezzi mesoscopici**

Relatore:
Prof. Luca Guido Molinari

Candidato:
Gianluca Roggero

Matricola: **867580**

Anno accademico 2019-2020

Indice

1	Introduzione	3
2	Matrici di Wishart-Laguerre	6
2.1	Un po' di storia	6
2.2	Cenni di statistica multivariata	9
2.3	Distribuzione normale multivariata	9
2.4	Misure su insiemi di matrici	10
2.5	La funzione ipergeometrica generalizzata	14
2.6	Matrici di Wishart	16
2.6.1	Autovalori delle matrici di Wishart	19
2.6.2	Legame con i polinomi di Laguerre	21
2.7	Teorema di Marchenko-Pastur	24
2.7.1	Nozioni preliminari	24
2.7.2	Dimostrazione del teorema	26
3	Teoria di Landauer	30
3.1	Cenni di elettromagnetismo e meccanica quantistica	30
3.2	L'approccio di Landauer	36
3.2.1	Approssimazioni	37
3.2.2	Condizioni di scattering al contorno	38
3.2.3	Probabilità di trasmissione e riflessione	40
3.2.4	Corrente totale	43
3.2.5	Limite di bias nullo	44
3.2.6	Conduttanza di sonda doppia	45
4	Fluttuazioni della conduttanza	46
4.1	Matrici di scattering	46
4.1.1	Relazione tra matrice di trasferimento e conduttanza	48
4.2	Varianza della conduttanza	49
4.2.1	Equazione integrale per la densità degli autovalori	50
4.2.2	Formula per la varianza di una statistica lineare	55
4.2.3	Formula della varianza	56
4.3	Teoria di Dorokhov-Mello-Pereyra-Kumar	56
4.3.1	Equazione di diffusione	58
4.3.2	Soluzione generale dell'equazione DMPK	61

Capitolo 1

Introduzione

Il presente elaborato si propone di raccogliere i principali risultati riguardanti le matrici di Wishart e la loro applicazione nell'ambito della fisica dei mezzi mesoscopici disordinati.

Nel capitolo 1, dopo aver introdotto nozioni di base della statistica multivariata, vengono presentate brevemente le proprietà più notevoli della distribuzione normale multivariata, della misura di Haar e delle funzione ipergeometrica generalizzata, tutte nozioni necessarie per ricavare alcuni risultati fondamentali riguardanti le matrici di Wishart, con un focus sulla distribuzione congiunta degli autovalori, seguendo l'esposizione di [27]. Il capitolo si conclude infine con la dimostrazione del teorema di Marchenko-Pastur presentata in [38].

Il capitolo 2 presenta l'approccio di Landauer al problema del trasporto di elettroni nei conduttori, nel caso particolare di un sistema chiuso, come esposto in [10]. In particolare, viene studiato il caso di un conduttore quasi 1-dimensionale, ovvero avente lunghezza molto maggiore della sua sezione trasversa, al quale viene applicata ai capi una differenza di potenziale chimico, considerando l'equazione di Schrödinger associata e calcolando i coefficienti di trasmissione agli estremi del conduttore. Infine, viene ricavata la formula di Landauer per la quantizzazione della conduttanza, prendendo in considerazione il limite per la differenza di potenziale tendente a zero.

Infine, nel terzo capitolo viene calcolata la varianza delle fluttuazioni della conduttanza nei sistemi mesoscopici disordinati, seguendo la tecnica di [4], presentata la teoria di Dorokhov-Mello-Pereyra-Kumar come generalizzazione dello studio delle fluttuazioni della conduttanza nel caso di N canali, ricavando l'equazione di DMPK nel caso trattato in [26] e infine, sfruttando i modelli di Calogero-Sutherland e alcuni argomenti di teoria dei gruppi di Lie, fornita la soluzione generale dell'equazione DMPK per indice di Dyson $\beta = 1, 2, 4$, mostrando, in particolare come quest'ultima possa essere scritta in forma esatta per $\beta = 2$ e semplificata nei due regimi metallico e isolante per ogni $\beta = 1, 2, 4$, seguendo la trattazione di [9].

Lo scopo di questa tesi, come accennato in precedenza, è quello di riassumere i risultati fondamentali degli argomenti sopracitati, senza avere la pretesa di fornire risultati o tecniche originali. In particolare, l'intenzione principale era quella di rendere la tesi presente il più autoesautiva possibile, fornendo talvolta la dimostrazione di alcuni tra gli enunciati presentati, cercando però di non appesantire la trattazione e di fornire solo

le nozioni strettamente necessarie. Spero di essere riuscito nell'intento e di non far annoiare i miei 24 lettori.

La mia speranza è che questo documento possa dare l'idea dell'incredibile legame tra le più disparate aree della fisica e della matematica e del profondo fascino che questo stretto rapporto esercita su di me.

Ringraziamenti

Capitolo 2

Matrici di Wishart-Laguerre

2.1 Un po' di storia

La nascita della teoria delle matrici random, o RMT dall'acronimo di *Random Matrix Theory*, viene fatta coincidere convenzionalmente con l'articolo sulle matrici di covarianza pubblicato dal matematico scozzese John Wishart nel 1928 [37]; tuttavia, già nel 1897, il tedesco Adolf Hurwitz [21] aveva introdotto, nell'ambito della *teoria degli invarianti*, la nozione di misura invariante per i gruppi ortogonale $O(N)$ e unitario $U(N)$, calcolando il volume di questi ultimi fattorizzando la loro misura in termini degli *angoli di Eulero*. I risultati di Hurwitz si legano alla RMT in quanto gli angoli di Eulero di una matrice distribuita uniformemente in $SO(N)$ o $U(N)$ sono indipendenti e sono distribuiti secondo la funzione Beta di Eulero. I risultati di Hurwitz vennero successivamente generalizzati da Hermann Weyl ([11]) al gruppo simplettico, ovvero ai quaternioni, $Sp(2N)$.

I risultati di Wishart del 1928 vennero poi ampliati dai lavori indipendenti di Ronald Fisher [15], del cinese Leetsch Charles Hsu [20], del sovietico Meyer Abraham Girshick [18] e dell'indiano Samarendra Nath Roy [29] sulla distribuzione congiunta degli autovallori delle matrici di covarianza, tutti pubblicati nel 1939.

La fama della RMT crebbe esponenzialmente tra gli anni '50 e gli anni '60, nel momento in cui i fisici si resero conto delle sue potenziali applicazioni.

Il padre dell'approccio RMT alla fisica teorica fu l'ungherese Eugene Wigner, che in un lavoro del 1950 [35] introdusse il concetto di distribuzione statistica dei livelli di energia nucleare e nel 1955 [36] quello di ensemble di matrici random. Nel 1956, Wigner diede uno dei contributi più celebri della RMT, derivando quella che viene definita *Wigner surmise* dalla distribuzione dello spacing dei livelli energetici degli spettri dei nuclei di atomi pesanti. In particolare, Wigner ricavò che in una sequenza semplice, ovvero quando i livelli energetici hanno stesso spin e parità, la densità di probabilità per lo spacing è data da:

$$p_w(s) = \frac{\pi s}{2} e^{-\frac{\pi s^2}{4}} \quad (2.1)$$

per le matrici reali simmetriche 2×2 con elementi indipendenti e distribuiti identicamente secondo una distribuzione gaussiana, con $s = S/D$, dove S è lo spacing tra due livelli e D è la distanza media tra due livelli consecutivi.



Adolf Hurwitz (Hildesheim, Regno di Hannover, 1859 – Zurigo, Svizzera 1919) studiò matematica a Monaco, Berlino, dove frequentò i corsi di Karl Weierstrass e Leopold Kronecker, e Lipsia, dove ottenne il dottorato sotto la guida di Felix Klein. Insegnò a Gottinga, Königsberg, dove ebbe come allievi David Hilbert e Hermann Minkowski, e infine al Politecnico di Zurigo. I contributi principali di Hurwitz furono principalmente in teoria delle superfici riemanniane e in teoria dei numeri.

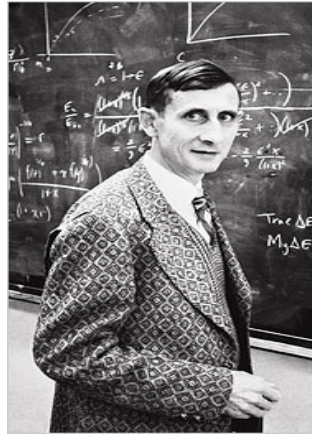
John Wishart (Montrose, Scozia, 1898 – Acapulco, Messico, 1956) studiò matematica a Edimburgo, sotto la guida di Edmund Taylor Whittaker, e a Cambridge, dove successivamente fu collaboratore di Ronald Fisher. I suoi lavori si concentrarono principalmente nell'ambito della statistica multivariata e dell'applicazione della statistica al settore agrario. Partecipò ad entrambe i conflitti mondiali: nel primo, come soldato semplice in Francia, nel secondo, come membro dell'intelligence britannica.

Altro grande protagonista dello sviluppo della RMT fu l'inglese Freeman John Dyson, che introdusse la classificazione degli ensemble di matrici random in base alle loro proprietà di invarianza rispetto all'inversione temporale [13]:

- (i) se un sistema non è invariante rispetto all'inversione temporale, gli elementi della matrice random associata al sistema sono numeri complessi e l'ensemble corrispondente è l'*ensemble unitario gaussiano* o *GUE*;
- (ii) se un sistema è invariante rispetto all'inversione temporale con modulo quadro dell'operatore di inversione temporale uguale a 1 o -1 , gli elementi della matrice random associata al sistema sono, rispettivamente, reali o quaternionici e gli ensemble corrispondenti sono l'*ensemble ortogonale gaussiano* o *GOE* e l'*ensemble simplettico gaussiano* o *GSE*.

I primi contributi principali alla RMT vennero raccolti dal fisico indiano Madan Lal Mehta [25] nel 1967 nella prima edizione di quello che ormai è diventato un testo classico della RMT. Dagli anni '60 in poi, la RMT cominciò ad insinuarsi in svariati ambiti della matematica e della fisica, che citiamo brevemente (vedi [17] per una presentazione più completa):

- *Statistica multivariata*;
- *Teoria della misura su ensemble di matrici random invarianti*;



Pál Jeno Wigner (Budapest, Austria-Ungheria, 1902 – Princeton, Stati Uniti, 1995), conosciuto maggiormente con il nome Eugene Paul una volta conseguita la cittadinanza statunitense, studiò ingegneria chimica a Budapest e Berlino, dove conobbe Leó Szilárd, e ottenne un dottorato in chimica fisica sotto la guida di Michael Polanyi. Fu in seguito assistente di Hilbert a Gottinga e lavorò al Progetto Manhattan. Vinse il Nobel per la fisica nel 1963 per i suoi contributi alla fisica atomica e alla fisica delle particelle elementari. Wigner diede dei contributi fondamentali in meccanica quantistica, ambito nel quale fu tra i pionieri dell'applicazione della teoria dei gruppi a sistemi quantistici, in fisica della materia e nella teoria delle matrici random.

John Freeman Dyson (Crowthorne, Inghilterra, 1923 – Princeton, Stati Uniti, 2020) studiò matematica al Trinity College di Cambridge, studi che vennero temporaneamente interrotti durante il secondo conflitto mondiale, nel quale Dyson lavorò nel reparto ricerca della RAF. Fu in seguito studente di Hans Bethe presso la Cornell University, dove conobbe Richard Feynman, nome al quale viene spesso legato in quanto Dyson dimostrò l'equivalenza delle formulazioni diagrammatica di Feynman e operatoriale di Schwinger e Tomonaga della QED. Dyson diede contributi notevoli negli ambiti più disparati: dalla teoria dei numeri all'astrofisica, dalla teoria delle matrici random all'ingegneria nucleare.

- *Teoria analitica dei numeri*, in relazione allo studio degli zeri della ζ di Riemann;
- *Fisica nucleare*;
- *Teoria dei sistemi disordinati*, nell'ambito della localizzazione di Anderson;
- *Caos quantistico*;
- *Cromodinamica quantistica*;
- *Gravità quantistica*;

oltre a trovare applicazioni anche in ambito finanziario [7].

2.2 Cenni di statistica multivariata

Definiamo innanzitutto alcuni momenti di un vettore random.

Il valore atteso $\mathbb{E}$ di un vettore random $m \times 1$ $\mathbf{X} = (X_1, \dots, X_m)^T$ è il vettore delle medie:

$$\mathbb{E}(\mathbf{X}) = \mathbb{E} \begin{pmatrix} \mathbb{E}(X_1) \\ \vdots \\ \mathbb{E}(X_m) \end{pmatrix}$$

La nozione di valore atteso si può estendere in maniera naturale a una matrice $m \times n$ generica Z come la matrice dei valori delle singole componenti:

$$\mathbb{E}(\mathbf{Z}) = \mathbb{E} \begin{pmatrix} \mathbb{E}(z_{11}) & \cdots & \mathbb{E}(z_{1n}) \\ \vdots & \cdots & \vdots \\ \mathbb{E}(z_{m1}) & \cdots & \mathbb{E}(z_{mn}) \end{pmatrix} = \mathbb{E}(z_{ij})$$

Se $\mathbf{X}$ ha valore atteso $\boldsymbol{\mu}$, la matrice di covarianza di $\mathbf{X}$ è definita come la matrice $m \times m$:

$$\Sigma \equiv \text{Cov}(\mathbf{X}) = \mathbb{E} \left[(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})^T \right]$$

con elemento $(\Sigma)_{ij} = \sigma_{ij} = \mathbb{E}[(X_i - \mu_i)(X_j - \mu_j)]$, che costituisce la covarianza tra X_i e X_j , mentre i termini diagonali $\sigma_{ii} = \mathbb{E}[(X_i - \mu_i)^2]$ costituiscono la varianza di ciascuna X_i . Osserviamo inoltre che le matrici di covarianza sono simmetriche.

Enunciamo ora una proprietà fondamentale delle matrici di covarianza.

Teorema 2.2.1. *Una matrice $m \times m$ Σ è una matrice di covarianza se e solo è definita non negativa.*

2.3 Distribuzione normale multivariata

Definizione 2.3.1. *Si dice che un vettore $m \times 1$ random $\mathbf{X}$ ha una distribuzione m -variata normale se, per ogni $\boldsymbol{\alpha} \in \mathbb{R}^m$, la distribuzione di $\boldsymbol{\alpha}'\mathbf{X}$ è una normale univariata.*

La distribuzione normale m -variata del vettore random $\mathbf{X}$ viene denotata con $N_m(\boldsymbol{\mu}, \Sigma)$.

Teorema 2.3.1. *Se $\mathbf{X}$ è $N_m(\boldsymbol{\mu}, \Sigma)$ e Σ è definita positiva allora la funzione densità di $\mathbf{X}$ è:*

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-m/2} (\det \Sigma)^{-1/2} \exp \left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu}) \right]$$

La funzione densità è costante sull'ellissoide, avente centro in $\boldsymbol{\mu}$ e forma ed orientazione legati a Σ , definito da:

$$(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x} - \boldsymbol{\mu}) = k$$

in $\mathbb{R}^m$, per ogni $k > 0$.

Vediamo ora qualche proprietà notevole della distribuzione normale m -variata.

Teorema 2.3.2. Se $\mathbf{X}_1, \dots, \mathbf{X}_n$ sono vettori random indipendenti e sono $N_m(\boldsymbol{\mu}_i, \Sigma_i)$ per $i = 1, \dots, N$, allora, date $\alpha_1, \dots, \alpha_N$ costanti fissate, allora $\sum_{i=1}^N \alpha_i \mathbf{X}_i$ è $N_m\left(\sum_{i=1}^N \alpha_i \boldsymbol{\mu}_i, \sum_{i=1}^N \alpha_i^2 \Sigma_i\right)$.

Corollario 2.3.2.1. Se $\mathbf{X}_1, \dots, \mathbf{X}_n$ sono vettori random indipendenti, ciascuna avente distribuzione $N_m(\boldsymbol{\mu}, \Sigma)$, allora la distribuzione del vettore **media campionaria** $\bar{\mathbf{X}} = \frac{1}{N} \sum_{i=1}^N \mathbf{X}_i$ è $N_m(\boldsymbol{\mu}, (1/N)\Sigma)$.

Il corollario precedente afferma che la distribuzione di:

$$N^{1/2} (\bar{\mathbf{X}} - \boldsymbol{\mu}) = N^{-1/2} \sum_{i=1}^N (\mathbf{X}_i - \boldsymbol{\mu})$$

è $N_m(\mathbf{0}, \Sigma)$.

Per completezza di esposizione, forniamo anche il *teorema del limite centrale multivariato*, che tratta il caso generale di vettori indipendenti distribuiti identicamente, ma senza il vincolo che la loro distribuzione sia normale.

Teorema 2.3.3. Sia $\mathbf{X}_1, \mathbf{X}_2, \dots$ una successione di vettori random indipendenti distribuiti identicamente con media $\boldsymbol{\mu}$ e matrice di covarianza Σ e sia $\bar{\mathbf{X}}_N = \frac{1}{N} \sum_{i=1}^N \mathbf{X}_i$, con $N \geq 1$. Allora, per $N \rightarrow +\infty$, la distribuzione asintotica di $N^{1/2}(\bar{\mathbf{X}}_N - \boldsymbol{\mu})$ è $N_m(\mathbf{0}, \Sigma)$.

2.4 Misure su insiemi di matrici

Apriamo ora una digressione sulle misure che possono essere introdotte sullo spazio di particolari matrici.

Per una data matrice X , denotiamo con dX la matrice dei differenziali (dx_{ij}) . Data X una matrice $n \times m$ qualsiasi, denotiamo con (dX) il prodotto esterno degli $n \times m$ elementi di dX :

$$(dX) \equiv \bigwedge_{j=1}^m \bigwedge_{i=1}^n dx_{ij}$$

Nei casi particolari nei quali X sia una matrice $m \times m$ simmetrica, (dX) indica il prodotto esterno degli $\frac{m}{2}(m+1)$ elementi distinti di dX , ovvero:

$$(dX) = \bigwedge_{1 \leq i \leq j \leq m} dx_{ij}$$

Teorema 2.4.1. Se $X = BY$ dove X e Y sono matrici $n \times m$ e B è una matrice $n \times n$ non singolare, allora vale:

$$(dX) = (\det B)^m (dY)$$

dove quindi $J(X \rightarrow Y) = (\det B)^m$.

Nella definizione delle misure definite su gruppi di matrici, svolge un ruolo fondamentale la *funzione Gamma di Eulero multivariata*.

Definizione 2.4.1. La funzione Gamma multivariata viene definita come:

$$\Gamma_m(a) = \int_{A>0} \exp \{ \text{Tr}(-A) \} (\det A)^{a-(m+1)/2} (dA) \quad (2.2)$$

dove $\Re(a) > \frac{(m-1)}{2}$ e l'integrale viene effettuato sullo spazio delle matrici $m \times m$ definite positive (denotato con $A > 0$), avente misura di Lebesgue: $(dA) = da_{11} \wedge da_{12} \wedge \dots \wedge da_{mm} \equiv da_{11} da_{12} \dots da_{mm}$.

La funzione Gamma multivariata può essere fattorizzata in termini di un prodotto di funzioni Gamma unidimensionali, infatti vale il seguente risultato:

$$\Gamma_m(a) = \pi^{m(m-1)/4} \prod_{i=1}^m \Gamma \left[a - \frac{1}{2}(i-1) \right]$$

Tornando alle matrici simmetriche, osserviamo che tali matrici possiedono $m \times m$ ha $\frac{m}{2}(m+1)$ elementi indipendenti e quindi l'insieme di tutte queste matrici forma un cono aperto in $\mathbb{R}^{m^2}$ definito dalla disequazione:

$$\begin{aligned} \det A = a_{11} &> 0 && \text{in } \mathbb{R} \\ \det A = a_{11}a_{22} - a_{12}a_{12} &> 0 && \text{in } \mathbb{R}^4 \\ &&& \text{e così via per ogni } m. \end{aligned}$$

Teorema 2.4.2. Se $\Re(a) > \frac{m-1}{2}$ e Σ è una matrice simmetrica avente $\Re(\Sigma) > 0$, allora:

$$\int_{A>0} \exp \text{Tr} \left(-\frac{1}{2} \Sigma^{-1} A \right) (\det A)^{a-(m+1)/2} (dA) = \Gamma_m(a) (\det \Sigma)^{a/2} 2^{ma}$$

con il solito vincolo sulla parte reale di a .

Forniamo ora vari risultati dei quali ci serviremo in seguito.

Teorema 2.4.3. Sia Z una matrice $n \times m$, con $m \leq n$, di rango m e tale che $Z = H_1 T$, dove H_1 è una matrice $n \times n$ con $H_1^T H_1 = \mathbb{I}_n$ e T è una matrice triangolare superiore $m \times m$ con elementi diagonali positivi. Sia H_2 una matrice $n \times (n-m)$ tale che $H = [H_1 : H_2]$ è una matrice ortogonale $n \times n$, dove $H = [\mathbf{h}_1, \dots, \mathbf{h}_m : \mathbf{h}_{m+1}, \dots, \mathbf{h}_n]$, con $\mathbf{h}_1, \dots, \mathbf{h}_m$ colonne di H_1 e $\mathbf{h}_{m+1}, \dots, \mathbf{h}_n$ colonne di H_2 . Allora:

$$(dZ) = \prod_{i=1}^m t_{ii}^{n-i} (dT) (H_1^T dH_1)$$

con

$$(H_1^T dH_1) \equiv \bigwedge_{i=1}^m \bigwedge_{j=i+1}^n \mathbf{h}_j^T d\mathbf{h}_i$$

Come conseguenza del teorema precedente, si ottiene il seguente teorema:

Teorema 2.4.4. Con l'assunzione del Teorema 2.4.3,

$$(dZ) = 2^{-m}(\det A)^{(n-m-1)/2}(dA)(H_1^T dH_1) \quad (2.3)$$

dove $A = Z'Z$.

Le matrici H_1 introdotte in precedenza formano quella che viene detta *varietà di Stiefel*, dal nome del matematico svizzero Eduard Stiefel:

$$V_{m,n} = \{H_1(n \times m) : H_1^T H_1 = \mathbb{I}_m\}$$

La proprietà soddisfatta dalle matrici della varietà di Stiefel ci porta ad individuare $\frac{m}{2}(m+1)$ componenti indipendenti per ciascuna matrice: di conseguenza, gli elementi di H_1 possono essere considerati come le coordinate di un punto su una varietà $mn - \frac{m}{2}(m+1)$ -dimensionale immersa nello spazio euclideo mn -dimensionale. Se $H_1 = (h_{ij})$, allora, siccome $\sum_{i=1}^n \sum_{j=1}^m h_{ij}^2 = m$, questa varietà è un sottoinsieme della sfera di raggio $\sqrt{m}$ nello spazio mn -dimensionale.

Citiamo due casi notevoli della varietà di Stiefel:

- **m=n**: la varietà di Stiefel coincide con un insieme di matrici quadrate che, dotate di un'operazione di prodotto tra matrici, formano un *gruppo di Lie*; in particolare, quello che è definito come *gruppo ortogonale* $O(m)$

$$V_{m,m} \equiv O(m) = \{H(m \times m) : H^T H = \mathbb{I}_m\}$$

- **m=1**: la varietà di Stiefel coincide con la *sfera unitaria* S^n :

$$V_{1,n} \equiv S^n = \{\mathbf{h}(n \times 1) : \mathbf{h}^T \mathbf{h} = 1\}$$

Consideriamo ora il caso di $O(m)$: per $H \in O(m)$

$$(H^T dH) \equiv \bigwedge_{i < j}^m \mathbf{h}_j^T d\mathbf{h}_i$$

Questa forma differenziale è il prodotto esterno degli elementi sottodisegnali della matrice antisimmetrica $H^T dH$. Osserviamo che questo è invariante sotto traslazioni *sinistra* e *destra*: $H \rightarrow QH$ e $H \rightarrow HQ^T$, con $Q \in O(m)$. Quindi valgono: $H^T dH \rightarrow H^T Q^T Q dH = H^T dH \Rightarrow (H' dH) \rightarrow (H^T dH)$, per l'invarianza sotto moltiplicazione a sinistra, e $H^T dH \rightarrow QH^T dHQ^T \Rightarrow (H' dH) \rightarrow (QH^T dHQ^T) = (\det Q)^{m-1}(H^T dH) = (H^T dH)$, per quella a destra. Questa *forma differenziale invariante* definisce una misura μ su $O(m)$, data da:

$$\mu(\mathcal{D}) = \int_{\mathcal{D}} (H^T dH) \quad \mathcal{D} \subset O(m)$$

dove $\mu(\mathcal{D})$ rappresenta il *volume* della regione $\mathcal{D}$ sulla varietà ortogonale. L'invarianza della forma differenziale implica l'invarianza della misura μ , ovvero:

$$\mu(Q\mathcal{D}) = \mu(\mathcal{D}Q) = \mu(\mathcal{D}) \quad \forall Q \in O(m)$$

La misura μ è chiamata anche *misura invariante su $O(m)$* o *misura di Haar*, dal nome del matematico ungherese Alfréd Haar, che dimostrò l'esistenza di una misura invariante su un qualsiasi gruppo topologico localmente compatto [19]. Il volume di $O(m)$ è dato da:

$$\text{Vol}[O(m)] = \mu[O(m)] = \int_{O(m)} (H' dH)$$

Consideriamo ora il caso generale di $V_{m,n}$. Abbiamo quindi, per $H_1 \in V_{m,n}$:

$$(H_1^T dH_1) \equiv \bigwedge_{i=1}^m \bigwedge_{j=i+1}^n \mathbf{h}_j^T d\mathbf{h}_i$$

dove $[H_1 : H_2] = [\mathbf{h}_1 \dots \mathbf{h}_m : \mathbf{h}_{m+1} \dots \mathbf{h}_n] \in O(n)$ è una funzione di H_1 . Si può dimostrare che la forma differenziale non dipende dalla scelta di H_2 ed è invariante sotto le trasformazioni: $H_1 \rightarrow QH_1$, per $[Q \in O(n)]$, e $H_1 \rightarrow H_1 P$, con $[P \in O(m)]$. Quindi, questa definisce una misura invariante sulla varietà di Stiefel $V_{m,n}$ e il volume di $V_{m,n}$ è:

$$\text{Vol}(V_{m,n}) = \int_{V_{m,n}} (H_1^T dH_1)$$

Si può dimostrare che:

$$\int_{V_{m,n}} (H_1^T dH_1) = \frac{2^m \pi^{mn/2}}{\Gamma_m(\frac{n}{2})}$$

dal quale si ricava che:

$$\text{Vol}[O(m)] = \frac{2^m \pi^{m^2/2}}{\Gamma_m(\frac{m}{2})}$$

Le misure appena introdotte possono essere normalizzate in modo da diventare misure di probabilità. Nel caso generale di $V_{m,n}$, denotata con (dH) la forma differenziale:

$$(dH) \equiv \frac{1}{\text{Vol}[V_{mn}]} (H_1^T dH_1) = \frac{\Gamma_m(\frac{n}{2})}{2^m \pi^{mn/2}} \bigwedge_{i=1}^m \bigwedge_{j=i+1}^n \mathbf{h}_j^T d\mathbf{h}_i$$

allora $\int_{V_{m,n}} dH = 1$. Quindi la misura $\mu^*(\mathcal{D}) \equiv \int_{\mathcal{D}} (dH)$ è una misura di probabilità su $V_{m,n}$.

Introduciamo ora un concetto importante per il seguito della trattazione: quello di *prodotto di Kronecker* o *prodotto diretto*.

Definizione 2.4.2. Siano $A = (a_{ij})$ una matrice $p \times q$ e $B = (b_{ij})$ una matrice $r \times s$. Il prodotto di Kronecker di A e B , denotato con $A \otimes B$, è la matrice $pr \times qs$

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1q}B \\ a_{21}B & a_{22}B & \cdots & a_{2q}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1}B & a_{p2}B & \cdots & a_{pq}B \end{bmatrix}$$

Un caso particolare del prodotto di Kronecker si ha quando $A = \mathbb{I}_p$, ovvero A coincide con la matrice identità $p \times p$: se B è una generica matrice $r \times s$, allora $\mathbb{I}_p \otimes B$ è una matrice $pr \times ps$ a blocchi diagonali, ovvero:

$$\mathbb{I}_p \otimes B = \begin{bmatrix} B & 0 & \cdots & 0 \\ 0 & B & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B \end{bmatrix}$$

2.5 La funzione ipergeometrica generalizzata

Prima di introdurre l'ensemble di Wishart, dobbiamo prima fornire le nozioni di *polinomio zonale* e *funzione ipergeometrica generalizzata*.

Definiamo innanzitutto la *partizione di un numero intero k* : sia k un intero positivo, una partizione κ di k è denotata con $\kappa = (k_1, k_2, \dots)$, in modo tale che $\sum_i k_i = k$, con la convenzione che $k_1 \geq k_2 \geq \dots$ e $i k_i \in \mathbb{N}$. La partizione di un intero non è unica e tutte le partizioni possono essere ordinate lessicograficamente, ovvero, date $\kappa = (k_1, k_2, \dots)$ e $\lambda = (l_1, l_2, \dots)$ partizioni dello stesso intero k , scriveremo che $\kappa > \lambda$ se $\kappa_i > \lambda_i$ per il primo indice i per il quale le due partizioni sono diverse.

Siano κ e λ partizioni dello stesso k e siano $y_1, \dots, y_m$ delle variabili: se $\kappa > \lambda$ allora diremo che il monomio $y_1^{k_1} \dots y_m^{k_m}$ è di *peso maggiore* rispetto al monomio $y_1^{l_1} \dots y_m^{l_m}$.

Definizione 2.5.1. Siano Y una matrice simmetrica $m \times m$ con autovalori $y_1, \dots, y_m$ e $\kappa = (k_1, \dots, k_m)$ una partizione di k in non più di m interi. Il polinomio zonale di Y corrispondente a κ , denotato con $C_\kappa(Y)$, è un polinomio di grado k simmetrico e omogeneo negli autovalori $y_1, \dots, y_m$ tale che:

- (i) Il termine di peso maggiore nella $C_\kappa(Y)$ è $y_1^{k_1} \dots y_m^{k_m}$, tale che:

$$C_\kappa(Y) = d_\kappa y_1^{k_1} \dots y_m^{k_m} + \text{termini di ordine minore}$$

con d_κ costante;

- (ii) $C_\kappa(Y)$ è un'autofunzione dell'operatore differenziale ∇_Y data da:

$$\nabla_Y = \sum_{i=1}^m y_i^2 \frac{\partial^2}{\partial y_i^2} + \sum_{i=1}^m \sum_{\substack{j=1 \\ j \neq i}}^m \frac{y_i^2}{y_i - y_j} \frac{\partial}{\partial y_i}$$

ovvero:

$$\nabla_Y C_\kappa(Y) = \alpha C_\kappa(Y)$$

con α costante indipendente dagli autovalori di Y .

- (iii) Se consideriamo tutte le partizioni κ di k i polinomi zionali hanno coefficienti uguali a 1 nell'espansione di $(\text{Tr } Y)^k$, ovvero:

$$(\text{Tr } Y)^k = (y_1 + \dots + y_m)^k = \sum_{\kappa} C_\kappa(Y)$$

I polinomi zionali sono fondamentali nella generalizzazione delle *funzioni ipergeometriche* al caso di argomento matriciale.

Definizione 2.5.2. Siano X una matrice $m \times m$ complessa simmetrica e $a_1, \dots, a_p, b_1, \dots, b_q \in \mathbb{C}$, definiamo le funzioni ipergeometriche di argomento matriciale come:

$${}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; X) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a_1)_{\kappa} \dots (a_p)_{\kappa}}{(b_1)_{\kappa} \dots (b_q)_{\kappa}} \frac{C_{\kappa}(X)}{k!} \quad (2.4)$$

dove la sommatoria più interna viene effettuata su tutte le partizioni κ di k e il coefficiente ipergeometrico $(a)_{\kappa}$ è dato da:

$$(a)_{\kappa} = \prod_{i=1}^m (a - (i-1)/2)_{k_i} \quad (2.5)$$

dove $(a)_k = a(a+1) \dots (a+k-1)$, $(a)_0 = 1$.

I termini della (2.4) devono soddisfare le seguenti condizioni:

- (i) nessuno dei b_j può essere uguale a zero o essere un intero o semi-intero minore o uguale a $(m-1)/2$, altrimenti il denominatore si annullerebbe;
- (ii) la serie:
 - converge per ogni X quando $p \leq q$;
 - converge per $\|X\| < 1$ quando $p = q + 1$, dove $\|X\|$ è il massimo dei valori assoluti degli autovalori di X ;
 - diverge per ogni $X \neq 0$ se $p > q + 1$.

Tra i casi particolari della (2.4), abbiamo:

(i)

$${}_0F_0(X) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{C_{\kappa}(X)}{k!} = \sum_{k=0}^{\infty} \frac{(\text{Tr } X)^k}{k!} = \exp \{ \text{Tr } X \}$$

(ii)

$${}_1F_0(a; X) = \sum_{k=0}^{\infty} \sum_{\kappa} (a)_{\kappa} \frac{C_{\kappa}(X)}{k!} = \det(\mathbb{I}_m - X)^{-a}$$

quando $\|X\| < 1$.

Possiamo definire una funzione ipergeometrica avente due matrici come argomento.

Definizione 2.5.3. Le funzioni ipergeometriche con matrici $m \times m$ simmetriche X e Y come argomenti sono date da:

$${}_pF_q^{(m)}(a_1, \dots, a_p; b_1, \dots, b_q; X, Y) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(a_1)_{\kappa} \dots (a_p)_{\kappa}}{(b_1)_{\kappa} \dots (b_q)_{\kappa}} \frac{C_{\kappa}(X) C_{\kappa}(Y)}{k! C_{\kappa}(\mathbb{I}_m)} \quad (2.6)$$

Questa definizione è simmetrica rispetto allo scambio delle matrici:

$${}_pF_q^{(m)}(a_1, \dots, a_p; b_1, \dots, b_q; X, Y) = {}_pF_q^{(m)}(a_1, \dots, a_p; b_1, \dots, b_q; Y, X).$$

Inoltre, quando uno dei due argomenti coincide con la matrice identità, la (2.6) si riduce alla (2.4).

Forniamo ora alcuni risultati riguardanti il legame tra le funzioni ipergeometriche a una matrice e quelle a due.

Teorema 2.5.1. *Se X è una matrice definita positiva $m \times m$ e Y è una matrice simmetrica, allora:*

$$\int_{O(m)} {}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; XHYH^T)(dH) = {}_pF_q^{(m)}(a_1, \dots, a_p; b_1, \dots, b_q; X, Y) \quad (2.7)$$

dove (dH) è la misura invariante normalizzata su $O(m)$

Teorema 2.5.2. *Siano Z una matrice complessa simmetrica $m \times m$ con $\Re(Z) > 0$ e Y una matrice simmetrica $m \times m$, allora valgono:*

$$\begin{aligned} \int_{X>0} \exp\{\text{Tr}(-XZ)\} (\det X)^{a-(m+1)/2} {}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; X)(dX) \\ = \Gamma_m(a)(\det Z)^{-a} {}_{p+1}F_q(a_1, \dots, a_p; b_1, \dots, b_q; Z^{-1}) \end{aligned} \quad (2.8)$$

e

$$\begin{aligned} \int_{X>0} \exp\{\text{Tr}(-XZ)\} (\det X)^{a-(m+1)/2} {}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; X, Y)(dX) \\ = \Gamma_m(a)(\det Z)^{-a} {}_{p+1}F_q(a_1, \dots, a_p; b_1, \dots, b_q; Z^{-1}, Y) \end{aligned} \quad (2.9)$$

per $p < q$, $\Re(a) > (m-1)/2$; mentre per $p = q$, $\Re(a) > (m-1)/2$, $\|Z^{-1}\| < 1$ ($\|Y\| \leq 1$).

2.6 Matrici di Wishart

Prima di passare all'argomento principale del capitolo, abbiamo bisogno di introdurre un cospicuo numero di notazioni.

Diciamo che una matrice random $r \times s$ Y è *distribuita normalmente*, ovvero $Y \in N(M, C \otimes D)$, dove M è una matrice $r \times s$ e C e D sono matrici, rispettivamente, $r \times r$ e $s \times s$ entrambe definite positive, quando intendiamo che $\mathbb{E}(Y) = M$ e che $C \otimes D$ è la matrice di covarianza del vettore $\mathbf{y} = \text{vec}(Y')$, dove $\text{vec}(Y^T) = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_s)^T$ è un vettore $(rs \times 1)$. Quindi, l'affermazione " $Y \in N(M, C \otimes D)$ " equivalente all'affermazione che " $\mathbf{y}$ è $N_{rs}(\mathbf{m}, C \otimes D)$ ", con $\mathbf{m} = \text{vec}(M^T)$.

Forniamo il seguente risultato per la funzione densità congiunta degli elementi di Y .

Teorema 2.6.1. *Se la matrice $r \times s$ Y è $N(M, C \otimes D)$, dove C e D sono le matrici di cui sopra, allora la funzione densità di Y è:*

$$(2\pi)^{-rs/2} (\det C)^{-s/2} (\det D)^{-r/2} \exp \left\{ \text{Tr} \left[-\frac{1}{2} C^{-1} (Y - M) D^{-1} (Y - M)^T \right] \right\}$$

Ora, siano $\mathbf{X}_1, \dots, \mathbf{X}_N$ vettori random indipendenti $N_m(\mu, \Sigma)$ e Σ definita positiva. Sia X una matrice $N \times m$:

$$X = \begin{bmatrix} \mathbf{X}_1^T \\ \mathbf{X}_2^T \\ \vdots \\ \mathbf{X}_N^T \end{bmatrix}$$

allora:

$$\mathbb{E}(X) = \begin{bmatrix} \mu^T \\ \mu^T \\ \vdots \\ \mu^T \end{bmatrix} = \mathbb{I} \mu^T \quad \mathbb{I} = (1, \dots, 1)^T \in \mathbb{R}^N$$

e $\text{Cov}[\text{vec}(X^T)] = \mathbb{I}_N \otimes \Sigma$, così che X è $N(\mathbb{I} \mu^T, \mathbb{I}_N \otimes \Sigma)$.

Teorema 2.6.2. Se la matrice $N \times m$ X è $N(\mathbb{I} \mu^T, \mathbb{I}_N \otimes \Sigma)$, allora il vettore media campionaria $\bar{\mathbf{X}}$ e la matrice di covarianza S , definiti come:

$$\bar{\mathbf{X}} = \frac{1}{N} \sum_{i=1}^N \mathbf{X}_i = \frac{1}{N} X^T \mathbb{I} \quad S = \frac{1}{N-1} A$$

con:

$$A = \sum_{i=1}^N (\mathbf{X}_i - \bar{\mathbf{X}}) (\bar{\mathbf{X}} - \mathbf{X}_i)^T = (X - \mathbb{I} \bar{\mathbf{X}}^T)^T (X - \mathbb{I} \bar{\mathbf{X}}^T)$$

sono distribuiti indipendentemente; $\bar{\mathbf{X}}$ è $N_m(\mu, (1/N)\Sigma)$ e A ha la stessa distribuzione di $Z^T Z$, dove Z è una matrice $(N-1) \times m$ $N(0, \mathbb{I}_{N-1} \otimes \Sigma)$

Definizione 2.6.1. Se $A = Z^T Z$, dove Z è una matrice $(N-1) \times m$ $N(0, \mathbb{I}_{N-1})$, allora si dice che A ha la **distribuzione di Wishart** con $N-1$ gradi di libertà e matrice di covarianza Σ . Scriveremo quindi che A è $W_m(n, \Sigma)$, dove m è la dimensione della matrice.

Teorema 2.6.3. Se A è $W_m(n, m)$ con $n \geq m$ allora la funzione densità di A è:

$$f(A) = \frac{1}{2^{mn/2} \Gamma_m(n/2) (\det \Sigma)} \exp \left\{ \text{Tr} \left(-\frac{1}{2} \Sigma^{-1} A \right) (\det A)^{(n-m-1)/2} \right\}$$

Dimostrazione. La densità di Z è:

$$(2\pi)^{-mn/2} (\det \Sigma)^{-n/2} \exp \left\{ \text{Tr} \left[-\frac{1}{2} \Sigma^{-1} Z^T Z \right] \right\} (dZ) \quad (2.10)$$

dove l'elemento di volume $(dZ) \equiv \bigwedge_{i=1}^n \bigwedge_{j=1}^m dz_{ij}$ è stato incluso per facilitare il calcolo degli Jacobiani quando facciamo le trasformazioni su Z . Siccome $n \geq m$, Z ha rango m con probabilità 1. Poniamo $Z = H_1 T$, dove H_1 è una matrice $n \times m$ con $H_1^T H_1 = \mathbb{I}_m$ e T è una matrice triangolare superiore $m \times m$. Allora $A = Z^T Z = T^T T$ e l'elemento di volume (dZ) diventa:

$$(dZ) = 2^{-m} (\det A)^{(n-m-1)/2} (dA) (H_1^T dH_1)$$

, così che la densità congiunta di A e H_1 è:

$$(2\pi)^{-mn/2} (\det \Sigma)^{-n/2} \exp \left\{ \text{Tr} \left(-\frac{1}{2} \Sigma^{-1} A \right) \right\} \times 2^{-m} (\det A)^{(n-m-1)/2} (dA) (H_1^T dH_1) \quad (2.11)$$

La densità marginale di A data da (2.10) segue da (2.11) integrando rispetto a H_1 sulla varietà di Stiefel $V_{m,n}$:

$$\int_{V_{m,n}} (H_1^T dH_1) = \frac{2^m \pi^{mn/2}}{\Gamma_m(n/2)}$$

□

Corollario 2.6.3.1. Se $\mathbf{X}_1, \dots, \mathbf{X}_N$ sono vettori random indipendenti $N_m(\mu, \Sigma)$ e $N > m$, la funzione densità della matrice di covarianza campionaria

$$S = \frac{1}{N-1} \sum_{i=1}^N (\mathbf{X}_i - \bar{\mathbf{X}}) (\mathbf{X}_i - \bar{\mathbf{X}})$$

è

$$f(S) = \frac{1}{\Gamma_m(n/2) (\det \Sigma)^{n/2}} (n/2)^{mn/2} \exp \left\{ \text{Tr} \left(-\frac{1}{2} n \Sigma^{-1} S \right) \right\} (\det S)^{(n-m-1)/2} \quad (2.12)$$

Nel caso univariato $m = 1$, se poniamo:

$$\Sigma \equiv \sigma^2 \quad S \equiv s^2 = \frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})^2$$

la funzione densità di s^2 è, dalla (2.12):

$$\left(\frac{1}{2\sigma^2} \right)^{n/2} \frac{1}{\Gamma(n/2)} \exp \left(-\frac{ns^2/\sigma^2}{2} \right) (s^2)^{n/2-1}$$

Ponendo $v = ns^2/\sigma^2$, otteniamo la funzione densità di v :

$$\frac{1}{2^{n/2} \Gamma(n/2)} e^{-v/2} v^{n/2-1}$$

ovvero la funzione densità χ_n^2 .

La distribuzione di Wishart ha degli interessanti collegamenti con la *statistica T^2 di Hotelling*, definita nel modo seguente: siano $\mathbf{X}_1, \dots, \mathbf{X}_N$ dei vettori random $N(\mu, \Sigma)$ che hanno vettore media campionaria $\bar{\mathbf{X}}$ e matrice di covarianza campionaria S , allora:

$$T^2 = N \bar{\mathbf{X}}^T S^{-1} \bar{\mathbf{X}}$$

2.6.1 Autovalori delle matrici di Wishart

Teorema 2.6.4. Se A è una matrice $m \times m$ definita positiva con densità

$$f(A) = \frac{1}{2^{mn/2} \Gamma_m(n/2) (\det \Sigma)} \exp \left\{ \text{Tr} \left(-\frac{1}{2} \Sigma^{-1} A \right) (\det A)^{(n-m-1)/2} \right\}$$

allora la densità congiunta degli autovalori $\lambda_1 > \lambda_2 > \dots > \lambda_m > 0$ di A è:

$$f(\lambda_1, \dots, \lambda_m) = \frac{\pi^{m^2/2}}{\Gamma_m(m/2)} \prod_{i < j}^m (\lambda_i - \lambda_j) \int_{O(m)} f(H \Lambda H^T) (dH)$$

dove (dH) è la misura invariante di Haar normalizzata su $O(m)$ e $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_m)$.

Dimostrazione. Possiamo effettuare una trasformazione ortogonale da A ai suoi autovalori e autovettori, ovvero:

$$A = H \Lambda H^T$$

con $H \in O(m)$. La i -esima colonna di H è un autovettore normalizzato di A associato all'autovalore λ_i . Ricaviamo ora la Jacobiano di questa trasformazione. Notiamo innanzitutto che:

$$dA = dH \Lambda H^T + H d\Lambda H^T + H \Lambda dH^T$$

da cui otteniamo:

$$\begin{aligned} H^T dA H &= H^T dH \Lambda + d\Lambda + \Lambda dH^T H \\ &= H^T dH \Lambda - \Lambda H^T dH + d\Lambda \end{aligned} \tag{2.13}$$

siccome $H^T dH = -dH^T H$. Il prodotto esterno degli elementi distinti nella matrice simmetrica a sinistra dell'uguale nella (2.13) è:

$$(\det H)^{m+1} (dA) = (dA)$$

Il prodotto esterno degli elementi diagonali nella parte destra della (2.13) è:

$$\bigwedge_{i=1}^m d\lambda_i$$

e per $i < j$ l'elemento $i - j$ -esimo nella parte destra della (2.13) è $\mathbf{h}'_j d\mathbf{h}_i (\lambda_i - \lambda_j)$. Di conseguenza, il prodotto esterno degli elementi distinti della matrice simmetrica nella parte destra della (2.13) è:

$$\bigwedge_{i < j}^m \mathbf{h}'_j d\mathbf{h}_i \prod_{i < j} (\lambda_i - \lambda_j) \bigwedge_{j=1}^m d\lambda_i$$

Uguagliando i prodotti esterni, otteniamo:

$$\begin{aligned}
(dA) &= \bigwedge_{i < j}^m \mathbf{h}'_j d\mathbf{h}_i \prod_{i < j} (\lambda_i - \lambda_j) \bigwedge_{j=1}^m d\lambda_i \\
&= (H^T dH) \prod_{i < j} (\lambda_i - \lambda_j) \bigwedge_{j=1}^m d\lambda_i \\
&= \frac{2^m \pi^{m^2/2}}{\Gamma_m(m/2)} (dH) \prod_{i < j} (\lambda_i - \lambda_j) \bigwedge_{i=1}^m d\lambda_i
\end{aligned}$$

Sostituendo $A = H\Lambda H^T$ e (dA) in $f(A)dA$, integrando rispetto a (dH) su $O(m)$ e dividendo il risultato per 2^m , otteniamo la densità di $\lambda_1, \dots, \lambda_m$ come:

$$f(\lambda_1, \dots, \lambda_m) = \frac{\pi^{m^2/2}}{\Gamma_m(m/2)} \prod_{i < j} (\lambda_i - \lambda_j) \int_{O(m)} f(H\Lambda H^T) (dH)$$

□

La densità degli autovalori di una matrice di Wishart è diretta conseguenza di questo teorema:

Corollario 2.6.4.1. *Se A è $W_m(n, \Sigma)$ con $n > m - 1$, la funzione densità congiunta degli autovalori $\lambda_1, \dots, \lambda_n$ di A è:*

$$\begin{aligned}
f(\lambda_1, \dots, \lambda_m) &= \frac{\pi^{m^2/2} 2^{-mn/2} (\det \Sigma)^{-n/2}}{\Gamma_m(m/2) \Gamma(n/2)} \prod_{i=1}^m \lambda_i^{(n-m-1)/2} \prod_{i < j}^m (\lambda_i - \lambda_j) \\
&\quad \times \int_{O(m)} \exp \left\{ \text{Tr} \left(-\frac{1}{2} \Sigma^{-1} H \Lambda H' \right) \right\} (dH)
\end{aligned}$$

Nel caso generale, l'integrale nel corollario precedente può essere valutato tramite i polinomi zionali e le funzione ipergeometrica generalizzata.

Teorema 2.6.5. *Siano S una matrice di covarianza campionaria e sia $A = nS$ tale che questa sia $W_m(n, \Sigma)$, con $n > m - 1$, allora la densità congiunta degli autovalori $\lambda_1, \dots, \lambda_m$ di S può essere espressa nella forma:*

$$f(\lambda_1, \dots, \lambda_m) = \left(\frac{n}{2}\right)^{mn/2} \frac{\pi^{m^2/2}}{\Gamma_m(n/2) \Gamma_m(m/2)} \prod_{i=1}^m \lambda_i^{(n-m-1)/2} \prod_{i < j}^m (\lambda_i - \lambda_j) {}_0F_0^{(m)} \left(-\frac{1}{2} n \Lambda, \Sigma^{-1} \right) \quad (2.14)$$

dove $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_m)$, $\lambda_1 > \dots > \lambda_m > 0$ e

$${}_0F_0^{(m)} \left(-\frac{1}{2} n \Lambda, \Sigma^{-1} \right) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{C_{\kappa} \left(-\frac{1}{2} n \Lambda, \Sigma^{-1} \right) C_{\kappa}(\Sigma^{-1})}{k! C_{\kappa}(\mathbb{I}_m)} \quad (2.15)$$

Dimostrazione. Sappiamo dal teorema (2.6.4.1) che la densità congiunta degli autovalori è:

$$f(\lambda_1, \dots, \lambda_m) = \left(\frac{n}{2}\right)^{mn/2} \frac{\pi^{m^2/2} 2^{-mn/2} (\det \Sigma)^{-n/2}}{\Gamma_m(m/2) \Gamma(n/2)} \prod_{i=1}^m \lambda_i^{(n-m-1)/2} \prod_{i<j}^m (\lambda_i - \lambda_j) \times \int_{O(m)} \exp \left\{ \text{Tr} \left(-\frac{1}{2} \Sigma^{-1} H \Lambda H' \right) \right\} (dH) \quad (2.16)$$

La tesi si ottiene grazie a (2.5.1) e al fatto che:

$$\exp \left\{ \text{Tr} \left(-\frac{1}{2} n \Sigma^{-1} H \Lambda H^T \right) \right\} = {}_0F_0 \left(-\frac{1}{2} n \Sigma^{-1} H \Lambda H^T \right) \quad (2.17)$$

□

Nel caso in cui $\Sigma = \lambda \mathbb{I}_m$, la densità congiunta degli autovalori si semplifica notevolmente.

Corollario 2.6.5.1. *Se A è $W_m(n, \lambda \mathbb{I}_m)$, con $n > m - 1$, la densità congiunta degli autovalori $\lambda_1, \dots, \lambda_m$ di A è:*

$$f(\lambda_1, \dots, \lambda_m) = \left(\frac{n}{2\lambda}\right) \frac{\pi^{m^2/2}}{(2\lambda)^{mn/2} \Gamma_m(m/2) \Gamma_m(n/2)} \exp \left(-\frac{1}{2\lambda} \sum_{i=1}^m \lambda_i \right) \prod_{i=1}^m \lambda_i^{(n-m-1)/2} \prod_{i<j}^m (\lambda_i - \lambda_j)$$

Dimostrazione. Ponendo $\Sigma = \lambda \mathbb{I}_m$ nel teorema (2.6.4) e osservando che:

$$\int_{O(m)} \exp \left\{ \text{Tr} \left(-\frac{1}{2\lambda} H \Lambda H^T \right) \right\} (dH) = \exp \left\{ \text{Tr} \left(-\frac{1}{2\lambda} \Lambda \right) \right\} \int_{O(m)} (dH) = \exp \left(-\frac{1}{2\lambda} \sum_{i=1}^m \lambda_i \right)$$

dalle quali si ottiene la tesi. □

2.6.2 Legame con i polinomi di Laguerre

I *polinomi di Laguerre classici* costituiscono una delle classi più importanti tra i cosiddetti *polinomi ortogonali classici*, assieme a quelli di Hermite, Legendre e Chebyshev (vedi [32]) e vennero ricavati dal francese Edmond Nicolas Laguerre come soluzione dell'equazione differenziale del secondo ordine:

$$xy'' + (1-x)y' + ny = 0 \quad n \in \mathbb{N}_0$$

Un polinomio di Laguerre L_k^γ è dato:

$$L_k^\gamma(x) = (\gamma + 1)_k \sum_{s=0}^k \binom{k}{s} \frac{(-x)^s}{(\gamma + 1)_s}$$

per $\gamma > -1$. Scegliamo L_k^γ normalizzato in modo tale che il coefficiente di x^k sia $(-1)^k$. L_k^γ è un polinomio di grado k in x e gli $\{L_k^\gamma\}_k$ sono ortogonali per $x > 0$ rispetto alla funzione peso $e^{-x}x^\gamma$, ovvero:

$$\int_0^\infty e^{-x}x^\gamma L_k^\gamma(x)L_j^\gamma(x)dx = \delta_{jk}k!\Gamma(\gamma + 1 + k)$$

La funzione generatrice per L_k^γ è:

$$(1 - z)^{-\gamma-1} \exp\left(\frac{xz}{z-1}\right) = \sum_{k=0}^{\infty} \frac{L_k^\gamma(x)z^k}{k!} \quad (2.18)$$

dove $|z| < 1$.

Possiamo ora introdurre la nozione di *polinomio di Laguerre generalizzato*.

Definizione 2.6.2. Il polinomio di Laguerre generalizzato L_κ^γ corrispondente a una partizione κ di una matrice $m \times m$ simmetrica X è:

$$L_\kappa^\gamma(X) = (\gamma + p)_\kappa C_\kappa(\mathbb{I}_m) \sum_{s=0}^k \sum_{\sigma} \binom{\kappa}{\sigma} \frac{C_\sigma(-X)}{(\gamma + p)_\sigma C_\sigma(\mathbb{I}_m)}$$

con $\gamma > -1$, dove la sommatoria più interna viene effettuata su tutte le partizioni σ dell'intero s e $p = (m + 1)/2$.

Il polinomio L_κ^γ è un polinomio simmetrico di grado k rispetto agli autovalori di X . Osserviamo che:

$$L_\kappa^\gamma(0) = (\gamma + p)_\kappa C_\kappa(\mathbb{I}_m)$$

Anche i polinomi di Laguerre generalizzati soddisfano una relazione di ortogonalità, però rispetto alla densità di Wishart. Il risultato seguente è il motivo per il quale l'ensemble di Wishart è detto anche **ensemble di Wishart-Laguerre**.

Teorema 2.6.6. $L_\kappa^\gamma(X)$ e $L_\sigma^\gamma(X)$ sono ortogonali su $X > 0$ rispetto alla funzione densità

$$W(X) = \exp\{\text{Tr}(-X)\} (\det X)^\gamma$$

per $\kappa \neq \sigma$. In particolare,

$$\begin{aligned} \int_{X>0} \exp\{\text{Tr}(-X)\} (\det X)^\gamma L_\kappa^\gamma(X)L_\sigma^\gamma(X)(dX) \\ = \delta_{\kappa\sigma}k!C_\kappa(\mathbb{I})\Gamma_m(\gamma + p)(\gamma + p)_\kappa \end{aligned} \quad (2.19)$$

dove $p = (m + 1)/2$.

Dimostrazione. Dalla funzione generatrice (2.18), abbiamo:

$$\det(\mathbb{I} - Z)^{-\gamma-p} \int_{O(m)} \exp \text{Tr}(-XHZ(\mathbb{I} - Z)^{-1}H^T)(dH) = \sum_{k=0}^{\infty} \sum_{\kappa} \frac{L_\kappa^\gamma(X)C_\kappa(Z)}{C_\kappa(\mathbb{I}_m)k!}$$

Moltiplicando entrambe le parti per $\exp \{\text{Tr}(-X)\} (\det X)^\gamma C_\sigma(X)$, dove σ è una partizione di un qualsiasi intero s , e integriamo su $X > 0$. Il lato sinistro di (2.19) diventa:

$$\begin{aligned} & \det(\mathbb{I} - Z)^{-\gamma-p} \int_{O(m)} \exp \left\{ \text{Tr} \left[-X (\mathbb{I} + HZ(\mathbb{I} - Z)^{-1} H^T) \right] \right\} \\ & \quad \times (\det X)^\gamma C_\sigma(X) (dX) (dH) \\ & = \det(\mathbb{I} - Z)^{-\gamma-p} \Gamma_m(\gamma + p) (\gamma + p)_\sigma \int_{O(m)} \\ & \quad \times \det(\mathbb{I} + HZ(\mathbb{I} - Z)^{-1} H^T)^{-\gamma-p} (dH) C_\sigma(\mathbb{I} - Z) \end{aligned}$$

Per calcolare quest'ultima parte ci serviamo del seguente risultato:

Lemma 2.6.7. *Siano Z una matrice complessa simmetrica $m \times m$ con $\Re(Z) > 0$ e Y una matrice simmetrica $m \times m$. Allora:*

$$\begin{aligned} & \int_{X>0} \exp \{ \text{Tr}(-XZ) \} (\det X)^{a-(m+1)/2} C_\kappa(XY) (dX) = \\ & \quad (a)_\kappa \Gamma_m(a) (\det Z)^{-a} C_\kappa(YZ^{-1}) \end{aligned}$$

per $\Re(a) > (m-1)/2$.

Quindi, tornando alla dimostrazione:

$$\begin{aligned} & \det(\mathbb{I} - Z)^{-\gamma-p} \int_{O(m)} \exp \left\{ \text{Tr} \left[-X (\mathbb{I} + HZ(\mathbb{I} - Z)^{-1} H^T) \right] \right\} \\ & \quad \times (\det X)^\gamma C_\sigma(X) (dX) (dH) \\ & = \Gamma_m(\gamma + p) (\gamma + p)_\sigma C_\sigma(\mathbb{I} - Z) \\ & = \Gamma_m(\gamma + p) (\gamma + p)_\sigma (-1)^s C_\sigma(Z) + \text{termini di grado minore} \end{aligned} \tag{2.20}$$

Il lato destro di (2.19) diventa:

$$\sum_{k=0}^{\infty} \sum_{\kappa} \frac{C_\kappa(Z)}{C_\kappa(\mathbb{I}) k!} \int_{X>0} \exp \{ \text{Tr}(-X) \} (\det X)^\gamma C_\sigma(X) L_\kappa^\gamma(X) (dX) \tag{2.21}$$

Confrontando i coefficienti di $C_\kappa(Z)$ di entrambi i lati otteniamo che

$$\int_{X>0} \exp \{ \text{Tr}(-X) \} (\det X)^\gamma C_\sigma(X) L_\kappa^\gamma(X) (dX) = 0$$

per $k \geq s$, ovvero $L_\kappa^\gamma(X)$ è ortogonale a tutti i polinomi di Laguerre di grado minore. Siccome dalla definizione di polinomio di Laguerre generalizzato:

$$L_\sigma^\gamma(X) = (-1)^s C_\sigma(X) + \text{termini di grado minore}$$

segue che $L_\kappa^\gamma(X)$ è ortogonale a tutti i polinomi di Laguerre $L_\sigma^\gamma(X)$ dello stesso grado per $\kappa \neq \sigma$. Ponendo $\kappa = \sigma$ e confrontando i coefficienti di $C_\sigma(Z)$ nelle (2.20) e (2.21), si ha:

$$\int_{X>0} \exp \{ \text{Tr}(-X) \} (\det X)^\gamma C_\sigma(-X) L_\sigma^\gamma(X) (dX) = s! C_\sigma(\mathbb{I}) \Gamma_m(\gamma + p) (\gamma + p)_\sigma$$

dalla quale si ottiene che:

$$\int_{X>0} \exp \{ \text{Tr}(-X) \} (\det X)^\gamma [L_\sigma(X)]^2 (dX) = s! C_\sigma(\mathbb{I}) \Gamma_m(\gamma + p) (\gamma + p)_\sigma$$

in quanto

$$L_\sigma^\gamma(X) = C_\sigma(-X) + \text{termini di grado minore che integrati danno zero}$$

□

2.7 Teorema di Marchenko-Pastur

Nel loro lavoro del 1972 [24], i fisici matematici sovietici Vladimir Alexandrovich Marchenko e Leonid Andreevich Pastur ricavarono l'omonima distribuzione come distribuzione limite per $N \rightarrow \infty$ della densità congiunta degli autovalori di due classi speciali di matrici hermitiane e di una di matrici unitarie.

2.7.1 Nozioni preliminari

Prima di fornire la dimostrazione del teorema di Marchenko - Pastur, dobbiamo introdurre qualche nozione e risultato preliminari.

Definiamo una *misura complessa* come una misura che ha le proprietà canoniche della misura di Lebesgue, che assume, però, valori in $\mathbb{C}$.

Definizione 2.7.1. Siano μ una misura di Borel su X spazio di misura e E un sottoinsieme di Borel di X . La misura μ è detta **esternamente regolare** su E se:

$$\mu(E) = \inf \{ \mu(U) : U \supset E, U \text{ aperto} \} \quad (2.22)$$

e **internamente regolare** se:

$$\mu(E) = \sup \{ \mu(U) : U \subset E, U \text{ compatto} \}. \quad (2.23)$$

Una **misura di Radon** su X è una misura di Borel che è finita su tutti gli insiemi compatti, esternamente regolare su tutti gli insiemi di Borel e internamente regolare su tutti gli insiemi aperti.

Una **misura di Radon con segno** è una misura di Borel le cui variazioni positive e negative sono misure di Radon.

Una **misura di Radon complessa** è una misura di Borel complessa avente come parti reale e immaginaria misure di Radon con segno.

Lo spazio delle misure di Radon complesse ha la struttura di spazio vettoriale avente norma: $\|\mu\| = |\mu|(X)$ e viene denotato con $M(X)$.

Definizione 2.7.2. Definiamo il seguente spazio funzionale: sia X uno spazio topologico, definiamo

$$C_0(X) = \{f \text{ continua su } X : \forall \epsilon > 0 \{x : |f(x)| \geq \epsilon\} \text{ è compatto}\} \quad (2.24)$$

Si può dimostrare [16] che, se X è uno spazio topologico di Hausdorff localmente compatto, $M(X) = C_0(X)^*$ e che $M(X)$ può essere dotato di una *topologia debole* per la quale, date una successione $\{\mu_n\}$ e μ in $M(X)$:

$$\mu \rightharpoonup \mu \iff \int f d\mu_n \rightarrow \int f d\mu \quad \forall f \in C_0(X). \quad (2.25)$$

Tale topologia debole viene detta *topologia vaga* su $M(X)$.

Introduciamo ora l'ingrediente saliente della dimostrazione del teorema di Marchenko-Pastur: la *trasformata di Stieltjes*, dal nome del matematico olandese Thomas Joannes Stieltjes, su $\mathbb{R}$

$$s_\mu(z) = \int_{\mathbb{R}} \frac{1}{\mu - z} d\mu(x), \quad (2.26)$$

dove $z \in \mathbb{C}$ e non appartenente al supporto di μ e μ è una misura di probabilità su $\mathbb{R}$. La (2.26) è ben definita nei semipiani superiore ed inferiore di $\mathbb{C}$. Si può dimostrare che:

Lemma 2.7.1. Teorema di continuità di Stieltjes [33]. Sia $\{\mu_n\}$ una successione di misure di probabilità casuali su $\mathbb{R}$ e sia μ una misura di probabilità deterministica, allora:

- (i) $\{\mu_n\}$ converge quasi certamente a μ nella topologia vaga se e solo se $\{s_{\mu_n}(z)\}$ converge quasi certamente a $s_\mu(z)$ per ogni z nel semipiano superiore;
- (ii) $\{\mu_n\}$ converge in probabilità a μ nella topologia vaga se e solo se $\{s_{\mu_n}(z)\}$ converge in probabilità a $s_\mu(z)$ per ogni z nel semipiano superiore;
- (iii) $\{\mu_n\}$ converge in aspettazione a μ nella topologia vaga se e solo se $\{s_{\mu_n}(z)\}$ converge in aspettazione a $s_\mu(z)$ per ogni z nel semipiano superiore.

Riportiamo un risultato del quale ci serviremo nella dimostrazione del teorema di Marchenko-Pastur.

Lemma 2.7.2. Formula di Sherman-Morrison[30]. Siano $A \in \mathbb{R}^{n \times n}$ una matrice quadrata invertibile e $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$. Allora, supponendo che $A + \mathbf{v}\mathbf{w}^T$ sia non singolare, con $\mathbf{v}\mathbf{w}^T$ il prodotto esterno tra i due vettori, vale:

$$(A + \mathbf{v}\mathbf{w}^T)^{-1} = A^{-1} - \frac{A^{-1}\mathbf{v}\mathbf{w}^T A^{-1}}{1 + \mathbf{w}^T A^{-1}\mathbf{v}} \quad (2.27)$$

2.7.2 Dimostrazione del teorema

Definizione 2.7.3. Sia A una matrice reale simmetrica $p \times p$, definiamo la sua **densità spettrale empirica** come:

$$\mu = \frac{1}{p} \sum_{i=1}^p \delta_{\lambda_i}$$

dove δ_λ è la massa di Dirac $\delta_\lambda = \delta(x - \lambda)$ con $\lambda \in \mathbb{R}$ e $\lambda_1 \leq \dots \leq \lambda_p$ sono gli autovalori di A .

Teorema 2.7.3. Sia X_{pn} una matrice random $p \times n$ tale che le sue colonne $\{\mathbf{x}_{pk}\}_{k=1}^p$ sono copie indipendenti e distribuite identicamente di un vettore random $\mathbf{x}_p \in \mathbb{R}^p$. Se vale:

$$\frac{\mathbf{x}_p^T A_p \mathbf{x}_p - \text{Tr}(A_p)}{p} \rightarrow 0 \quad p \rightarrow +\infty$$

per tutte le successioni $\{A_p\}$ di matrici $p \times p$ complesse aventi norme spettrali $\|A_p\| = \{\sqrt{\lambda_{\max}(AA^\dagger)}\}$ uniformemente limitate, allora, per ogni $p = p(n)$ con $p/n \rightarrow c > 0$ per $n \rightarrow +\infty$, vale:

$$P(\mu_{pn} \rightarrow \mu_c \text{ debolmente}, n \rightarrow +\infty) = 1$$

dove μ_c è la **legge di Marchenko-Pastur** per gli autovalori:

$$\mu_c(x) = (1 - 1/c)^+ \delta_0 + \frac{\sqrt{(b-x)(x-a)}}{2\pi cx} \chi_{[a,b]} dx$$

con $x^+ = \max\{x, 0\}$, per $x \in \mathbb{R}$, $a = (1 - \sqrt{c})^2$ e $b = (1 + \sqrt{c})^2$ e μ_{pn} è la distribuzione spettrale empirica di $n^{-1} X_{pn} X_{pn}^T$.

Dimostrazione. Per il teorema (2.7.1), dobbiamo dimostrare solo che $s_n(z) \rightarrow s(z)$ quasi certamente per ogni $z \in \mathbb{C}$ con $\Im m(z) > 0$, dove $s_n = s_n(z)$ e $s = s(z)$ sono le trasformate di Stieltjes di μ_{pn} e μ_c definite da:

$$s_n(z) = \int_{\mathbb{R}} \frac{\mu_{pn}(d\lambda)}{\lambda - z} \quad e \quad s(z) = \int_{\mathbb{R}} \frac{\mu_c(d\lambda)}{\lambda - z}.$$

Dalla definizione di μ_{pn} , $s_n(z) = \text{Tr}\{n^{-1} X_{pn} X_{pn}^T - z \mathbb{I}_p\}^{-1}/p$.

Fissiamo $z \in \mathbb{C}$ con $\nu = \Im m(z) > 0$. Ci serviamo del seguente lemma per trovare che: $s_n(z) \rightarrow \mathbb{E} s_n(z)$ quasi certamente:

Lemma 2.7.4. [1] Siano A una matrice random $n \times p$ con righe indipendenti e $s : \mathbb{C}^+ \rightarrow \mathbb{C}$ della misura spettrale. Allora per ogni $\alpha = x + iy \in \mathbb{C}_+$ e ogni $\epsilon > 0$:

$$P(|s_n(\alpha) - \mathbb{E} s_n(\alpha)| \geq \epsilon) \leq C \exp(-c\epsilon^2 ny^2)$$

Per finire la dimostrazione, dobbiamo dimostrare che $\mathbb{E} s_n(z) \rightarrow s(z)$. Utilizziamo il seguente lemma:

Lemma 2.7.5. Sia C una matrice reale simmetrica definita semipositiva e $\mathbf{x} \in \mathbb{R}^p$. Se $\nu = \Im m(z) > 0$, allora:

- (i) $\|(C - z\mathbb{I}_p)^{-1}\| \leq 1/\nu$;
- (ii) $|\text{Tr}(C + \mathbf{x}\mathbf{x}^T - z\mathbb{I}_p)^{-1} - \text{Tr}(C - z\mathbb{I}_p)^{-1}| \leq 1/\nu$;
- (iii) $|\mathbf{x}^T(C + \mathbf{x}\mathbf{x}^T - z\mathbb{I}_p)^{-1}\mathbf{x}| \leq 1 + |z|/\nu$
- (iv) $\Im(z + z\text{Tr}(C - z\mathbb{I}_p)^{-1}) \geq \nu$ e $\Im(\text{Tr}(C - z\mathbb{I}_p)^{-1}) > 0$.
- (v) $\Im(z + z\mathbf{x}^T(C - z\mathbb{I}_p)^{-1}\mathbf{x}) \geq \nu$

Consideriamo $\mathbf{x}_p = \mathbf{x}_{p,n+1}$ indipendente da X_{pn} e distribuito come le colonne $\{\mathbf{x}_{pk}\}_{k=1}^n$. Definiamo:

$$A_n = X_{pn}X_{pn}^T = \sum_{k=1}^n \mathbf{x}_{pk}\mathbf{x}_{pk}^T \quad e \quad B_n = A_n + \mathbf{x}_p\mathbf{x}_p^T = \sum_{k=1}^{n+1} \mathbf{x}_{pk}\mathbf{x}_{pk}^T$$

La matrice $B_n - zn\mathbb{I}_p$ è non degenera e:

$$p = \text{Tr}((B - zn\mathbb{I}_p)(B - zn\mathbb{I}_p)^{-1}) = \sum_{k=1}^{n+1} \mathbf{x}_{pk}^T(B - zn\mathbb{I}_p)^{-1}\mathbf{x}_{pk} - zn\text{Tr}(B - zn\mathbb{I}_p)^{-1}$$

Considerando i valori di aspettazione e utilizzando l'intercambiabilità di $\{\mathbf{x}_{pk}\}_{k=1}^{n+1}$

$$p = (n+1)\mathbb{E}\mathbf{x}_p^T(B_n - zn\mathbb{I}_p)^{-1}\mathbf{x}_p - zn\mathbb{E}\text{Tr}(B_n - zn\mathbb{I}_p)^{-1} \quad (2.28)$$

Definiamo $S_n(z) = \text{Tr}(B_n - zn\mathbb{I}_p)^{-1}$ e osserviamo che $S_n(z) = (p/n)s_n(z)$. Dalla (ii) e dalla (iii):

$$\mathbb{E}\text{Tr}(B_n - zn\mathbb{I}_p)^{-1} = \mathbb{E}S_n(z) + O(1/n) \quad e \quad \mathbb{E}\mathbf{x}_p^T(B_n - zn\mathbb{I}_p)^{-1}\mathbf{x}_p = O(1).$$

Per la (2.27):

$$\mathbf{x}_p^T(B_n - zn\mathbb{I}_p)^{-1}\mathbf{x}_p = \mathbf{x}_p^T(A_n + \mathbf{x}_p\mathbf{x}_p^T - zn\mathbb{I}_p)^{-1}\mathbf{x}_p = \frac{\mathbf{x}_p^T(A_n - zn\mathbb{I}_p)^{-1}\mathbf{x}_p}{1 + \mathbf{x}_p^T(A_n - zn\mathbb{I}_p)^{-1}\mathbf{x}_p}$$

Per la (i) e l'indipendenza di $\mathbf{x}_p$ e A_n , otteniamo $\mathbf{x}_p^T(A_n - zn\mathbb{I}_p)^{-1}\mathbf{x}_p - S_n(z) \rightarrow 0$. Abbiamo anche:

$$S_n(z) - \mathbb{E}S_n(z) = (p/n)(s_n(z) - \mathbb{E}s_n(z)) \xrightarrow{p} 0$$

Quindi, dalla (iv), dalla (v) e dalla disuguaglianza $|1 + w| \geq \Im(z + zw)/|z|$, $w \in \mathbb{C}$:

$$\left| \frac{\mathbf{x}_p^T(A_n - zn\mathbb{I}_p)^{-1}\mathbf{x}_p}{1 + \mathbf{x}_p^T(A_n - zn\mathbb{I}_p)^{-1}\mathbf{x}_p} - \frac{\mathbb{E}S_n(z)}{1 + \mathbb{E}S_n(z)} \right| \leq \frac{|z|^2}{\nu^2} |\mathbf{x}_p^T(A_n - zn\mathbb{I}_p)^{-1}\mathbf{x}_p| \xrightarrow{p} 0.$$

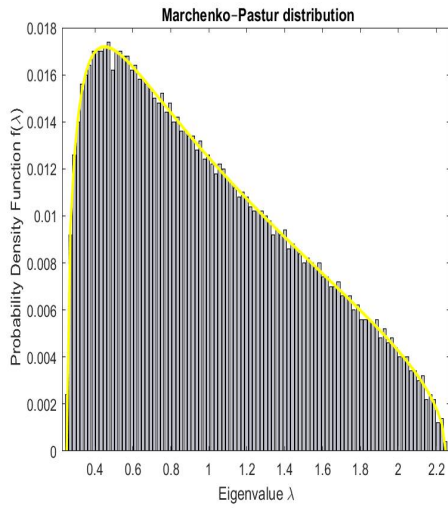
Per il teorema della convergenza dominata e per la (iii):

$$\mathbb{E}\mathbf{x}_p^T(B_n - zn\mathbb{I}_p)^{-1}\mathbf{x}_p = \frac{\mathbb{E}S_n(z)}{1 + \mathbb{E}S_n(z)} + o(1)$$

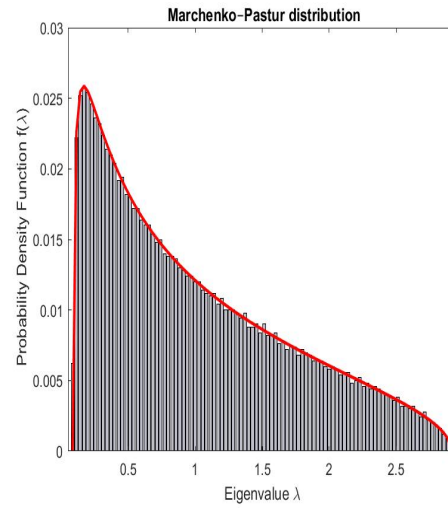
Allora la (2.28) si riduce a:

$$\frac{\mathbb{E}S_n(z)}{1 + \mathbb{E}S_n(z)} - z\mathbb{E}S_n(z) = c + o(1)$$

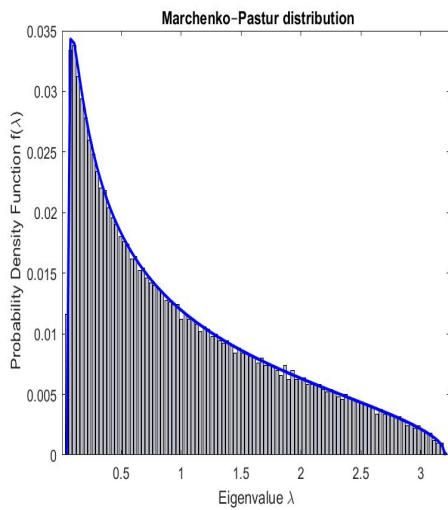
Dalla (i) e dalla (iv), $\mathcal{S} = (\mathbb{E}S_n(z))_{n=1}^\infty$ è una successione limitata con $\Im(\mathbb{E}S_n(z)) > 0$: di conseguenza, il punto che limita S ha parte immaginaria non negativa. Inoltre, si può provare che l'equazione quadratica limite: $S/(1+S) - zS = c$ ha una soluzione unica $S = S(z)$. Quindi: $\mathbb{E}S_n(z) = (p/n)\mathbb{E}s_n(z) \rightarrow S(z)$. Si può inoltre dimostrare [3] che $S(z) = cs(z)$ è l'unica soluzione, dove $S(z)$ è la trasformata di Stieltjes della legge di Marchenko-Pastur. Abbiamo quindi dimostrato che $s_n(z) \rightarrow s(z)$ quasi certamente. $\square$



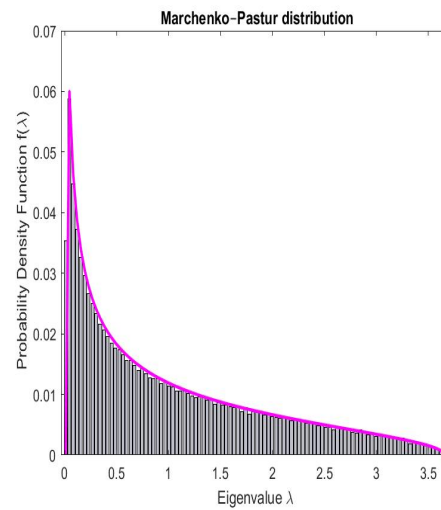
(a) $c=1/4$



(b) $c=1/2$



(c) $c=5/8$



(d) $c=5/6$

Distribuzioni di Marchenko-Pastur di matrici random aventi numero di righe $N=5000$. ©Youssef Khmou, Applied Mathematics ,30 January,2015

Capitolo 3

Teoria di Landauer

In questo capitolo viene esposto l'*approccio di Landauer* al problema del trasporto elettrico nei conduttori, il cui risultato fondamentale è la formula di quantizzazione della conduttanza o *formula di Landauer*, ricavata da Rolf Landauer nel 1957 [22].

L'oggetto di studio di questa sezione saranno principalmente sistemi in nanoscala o mesoscopici. I primi, come si può intuire dal prefisso *nano*—, sono sistemi aventi almeno una dimensione dell'ordine del nanometro, mentre i secondi presentano una struttura di grandezza maggiore dei precedenti, ma tali che non possano essere considerati come oggetti macroscopici.

Prima di dare una rassegna dei risultati fondamentali della teoria di Landauer, definiamo l'ambiente fisico e matematico nel quale si sviluppa tale teoria.

3.1 Cenni di elettromagnetismo e meccanica quantistica

Lo studio del trasporto di elettroni coincide di fatto con lo studio delle correnti nei mezzi conduttori. Il metodo più semplice per generare una corrente consiste nel prendere due conduttori (o elettrodi), caricarne uno negativamente e l'altro positivamente, creare una differenza di potenziale tra gli elettrodi e infine collegarli tra loro tramite un cavo conduttore. Ci concentriamo su quest'ultimo.

Come ogni manuale di elettromagnetismo classico insegna, possiamo concentrarci su una sezione S del conduttore e studiare le varie grandezze fisiche associate al flusso degli elettroni che attraversa S . Consideriamo l'elemento di superficie infinitesimo dS e il versore normale a dS $\hat{\mathbf{i}}$. Se consideriamo ogni elettrone come un oggetto classico, esso sarà caratterizzato da una velocità $\mathbf{v}_{tot}(\mathbf{r}, t) = \mathbf{v}_t(\mathbf{r}, t) + \mathbf{v}_d(\mathbf{r}, t)$, dove la prima componente è data dal contributo termico ed è orientata casualmente, mentre la seconda, detta *velocità di deriva*, è orientata come il flusso di corrente. Facendo la media su tutte le particelle considerate otteniamo che: $\mathbf{v}(\mathbf{r}, t) \equiv \langle \mathbf{v}_t(\mathbf{r}, t) \rangle + \langle \mathbf{v}_d(\mathbf{r}, t) \rangle = \langle \mathbf{v}_d(\mathbf{r}, t) \rangle$.

Definita $n(\mathbf{r}, t)$ la *densità numerica* di portatori di carica, possiamo calcolare la quantità di carica dQ che attraversa dS nell'intervallo di tempo dt :

$$dQ = en\mathbf{v} \cdot \hat{\mathbf{i}}dSdt$$

con e carica dell'elettrone. Si ricava immediatamente l'equazione per la corrente che fluisce attraverso dS :

$$dI = \frac{dQ}{dt} = en\mathbf{v} \cdot \hat{\mathbf{n}} dS \equiv \mathbf{j} \cdot d\mathbf{S}$$

dove abbiamo definito la *densità di corrente* $\mathbf{j}(\mathbf{r}, t) = en(\mathbf{r}, t)\mathbf{v}(\mathbf{r}, t)$ e considerato $d\mathbf{S} = dS\hat{\mathbf{n}}$. La corrente totale media attraverso la superficie S è data da:

$$I_S(t) = \int_S d\mathbf{S} \cdot \mathbf{j}(\mathbf{r}, t)$$

Anche la presenza di cariche e correnti localizzate contribuisce alla corrente totale nel conduttore. In particolare le cariche legate danno origine alla *polarizzazione*, mentre le correnti localizzate alla *magnetizzazione*. Chiamiamo q_i le cariche legate aventi posizione $\mathbf{r}_i$ e velocità $\mathbf{v}_i$. La polarizzazione è quindi data da:

$$\mathbf{P}(\mathbf{r}) = \frac{1}{V} \sum_i q_i \mathbf{r}_i$$

mentre la magnetizzazione:

$$\mathbf{M}(\mathbf{r}) = \frac{1}{2cV} \sum_i q_i (\mathbf{r}_i \times \mathbf{v}_i)$$

dove c della velocità della luce e V è il volume nel quale si trovano le cariche e le correnti localizzate.

La densità di corrente totale è quindi data da:

$$\mathbf{j}_{rot}(\mathbf{r}, t) = en(\mathbf{r}, t)\mathbf{v}(\mathbf{r}, t) + \frac{\partial \mathbf{P}}{\partial t} + c\nabla \times \mathbf{M}$$

e la densità totale:

$$en_{tot} = en - \nabla \cdot \mathbf{P}$$

Oltre a queste equazioni, è valida anche la *equazione di continuità*:

$$e \frac{\partial n(\mathbf{r}, t)}{\partial t} = -\nabla \cdot \mathbf{j}(\mathbf{r}, t)$$

Sperimentalmente si riscontra che la corrente decade con il tempo e che fluttua continuamente attorno a un valor medio. Se consideriamo un intervallo di tempo T grande abbastanza nel quale si ha un numero sufficientemente grande di oscillazioni della corrente, ma non così grande da avere una riduzione della corrente non trascurabile, possiamo definire la corrente media:

$$\langle I_S \rangle = \frac{1}{T} \int_{-T/2}^{T/2} dt I_S(t)$$

Nel limite di elettrodi infiniti e per $T \rightarrow +\infty$, otteniamo la *condizione stazionaria*:

$$\frac{d\langle I_S \rangle}{dt} = 0$$

Il limite di elettrodi infiniti ci porta ad introdurre il concetto di *riserva*.

Definizione 3.1.1. Una *riserva* è un sistema ideale che può fornire e ricevere un ammontare arbitrario di portatori di carica e di energia senza cambiare il suo stato interno.

Ci occupiamo ora dei concetti fondamentali di meccanica quantistica che verranno adoperati in seguito.

Sia $|\Psi_0\rangle \in \mathcal{H}$, dove $\mathcal{H}$ è lo spazio di Hilbert degli stati che caratterizzano un determinato sistema quantistico (nel nostro caso un elettrone), lo stato iniziale nel quale si trova il sistema. La sua evoluzione temporale $|\Psi(t)\rangle$ è data dalla soluzione dell'equazione di Schrödinger dipendente dal tempo:

$$i\hbar \frac{d|\Psi(t)\rangle}{dt} = \hat{H}|\Psi(t)\rangle$$

con condizione al contorno $|\Psi_0\rangle = |\Psi(0)\rangle$ e $\hat{H}$ Hamiltoniana dell'elettrone. Una soluzione dell'equazione di Schrödinger è data da:

$$|\Psi(t)\rangle = \hat{U}(t, t_0)|\Psi_0\rangle$$

con l'operatore unitario di evoluzione temporale dato da:

$$\hat{U}(t, t_0) = \mathcal{T} \left\{ \exp \left[-\frac{i}{\hbar} \int_{t_0}^t dt' \hat{H}(t') \right] \right\}$$

con $\mathcal{T}$ prodotto cronologico definito da:

$$\mathcal{T}[O_1(t_1), O_2(t_2)] \equiv \begin{cases} O_1(t_1)O_2(t_2), & \text{se } t_1 > t_2 \\ O_2(t_2)O_1(t_1), & \text{se } t_1 < t_2 \end{cases}$$

dove $O_1, O_2 : \mathcal{H} \rightarrow \mathbb{C}$ sono due operatori lineari. Dall'equazione di Schrödinger per lo stato $|\Psi(t)\rangle$ e dalla sua soluzione, otteniamo l'equazione del moto per l'operatore di evoluzione temporale:

$$i\hbar \frac{\partial \hat{U}(t, t_0)}{\partial t} = \hat{H}(t)\hat{U}(t, t_0)$$

con condizione al contorno $\hat{U}(t_0, t_0) = \mathbb{I}$. Nel caso in cui $\hat{H}$ non sia dipendente dal tempo, l'evoluzione temporale è semplicemente:

$$\hat{U}(t, t_0) \equiv \hat{U}(t - t_0) = e^{-i\hat{H}(t-t_0)/\hbar}$$

Passiamo ora a ridefinire le grandezze fisiche definite in precedenza come opportuni operatori su spazi di Hilbert.

Consideriamo un sistema di elettroni interagenti con un campo elettromagnetico classico caratterizzato da un potenziale vettore $\mathbf{A}(\mathbf{r}, t)$. L'operatore densità di corrente è dato da:

$$\hat{\mathbf{j}}(\mathbf{r}, t) = \frac{1}{2} \sum_i \{ \delta(\mathbf{r} - \hat{\mathbf{r}}_i), \hat{\mathbf{v}}_i \}$$

dove la somma viene effettuata su tutte le particelle aventi posizione $\hat{\mathbf{r}}_i$ e velocità

$$\hat{\mathbf{v}}_i = \frac{\hat{\mathbf{p}}_i - e\mathbf{A}(\hat{\mathbf{r}}_i, t)/c}{m}$$

dove l'operatore $\{\hat{A}, \hat{B}\}$ è l'anticommutatore canonico, δ è la distribuzione delta di Dirac e m è la massa dell'elettrone.

Se definiamo ora l'operatore densità numerica come $\hat{n}(\mathbf{r}) = \sum_i \delta(\mathbf{r} - \hat{\mathbf{r}}_i)$, otteniamo:

$$\hat{\mathbf{j}}(\mathbf{r}, t) = \hat{\mathbf{j}}_p(\mathbf{r}) - \frac{e}{mc} \hat{n}(\mathbf{r}) \mathbf{A}(\hat{\mathbf{r}}_i, t)$$

con l'operatore *densità di corrente paramagnetica* $\hat{\mathbf{j}}_p(\mathbf{r}) = \frac{1}{2m} \sum_i \{\delta(\mathbf{r} - \hat{\mathbf{r}}_i), \hat{\mathbf{p}}_i\}$ e l'operatore *densità di corrente diamagnetica* dato da $-\frac{e}{mc} \hat{n}(\mathbf{r}) \mathbf{A}(\hat{\mathbf{r}}_i, t)$.

In analogia con il caso classico, possiamo definire l'operatore di corrente passante per una data superficie S come:

$$\hat{I} = \int_S d\mathbf{S} \cdot \hat{\mathbf{j}}(\mathbf{r}, t) \quad (3.1)$$

L'operatore di corrente può presentare uno spettro di autosoluzioni sia discreto che continuo.

Nel caso discreto, gli autostati $|\Psi_{I_0}\rangle$ appartengono allo spazio di Hilbert del sistema e soddisfano la relazione di ortonormalità:

$$\langle \Psi_{I_0}, \Psi_{I'_0} \rangle = \delta_{I_0, I'_0}$$

Nel caso continuo, agli autostati della corrente $|\Psi_I\rangle$ corrispondono delle autofunzioni $\Psi \notin L^2$ e quindi non appartenenti allo spazio di Hilbert associato al sistema. Sono quindi normalizzabili in senso improprio:

$$\langle \Psi_I | \Psi_{I'} \rangle = \delta(I - I')$$

dove I e I' sono gli autovalori associati ai due autostati. Tuttavia, possiamo superare questa difficoltà considerando gli stati associati alle differenze di corrente, ovvero:

$$|\Psi_{\delta I}\rangle = \frac{1}{\sqrt{\delta I}} \int_I^{I+\delta I} dI |\Psi_I\rangle$$

che hanno norma al quadrato finita e quindi appartengono allo spazio di Hilbert del sistema. Questi stati soddisfano la seguente condizione di ortonormalità:

$$\lim_{\substack{\delta I \rightarrow 0 \\ \delta I' \rightarrow 0}} \langle \Psi_{\delta I} | \Psi_{\delta I'} \rangle = \delta_{I, I'}$$

Vediamo ora come si configura matematicamente il processo di misura di corrente nei casi di spettro discreto e continuo.

Nel caso di spettro discreto, la probabilità che una misura di $\hat{I}$ fornisca un autovalore I_0 è data da:

$$P(\hat{I} \rightarrow I_0 | t) = \langle \Psi(t) | \hat{P}_{I_0} | \Psi(t) \rangle$$

dove $\hat{P}_{I_0} = |\Psi_{I_0}\rangle\langle\Psi_{I_0}|$ è il proiettore sul sottospazio generato da $|\Psi_{I_0}\rangle$. Nell'istante successivo alla misura al tempo t il sistema si ritrova in $|\Psi_{I_0}\rangle$, ovvero: $|\Psi(t+dt)\rangle = |\Psi_{I_0}\rangle$. Nel caso continuo, alla misura di corrente viene associata una probabilità che questa fornisca un autostato in un intervallo $(I, I + \delta I)$:

$$P(\hat{I} \rightarrow I + \delta I | t) = \langle\Psi(t)|\hat{P}_I|\Psi(t)\rangle$$

dove $\hat{P}_I = \int_I^{I+\delta I} dI |\Psi_I\rangle\langle\Psi_I|$ è il proiettore sul sottospazio degli autovalori di corrente in $(I, I + \delta I)$. A seguito della misura, il sistema si trova in uno stato:

$$|\Psi(t+dt)\rangle = \frac{1}{\sqrt{\int_I^{I+\delta I} dI |c_I(t)|^2}} \int_I^{I+\delta I} dI c_I(t) |\Psi_I\rangle$$

se il sistema si trovava prima della misura in una sovrapposizione $\int dI c_I(t) |\Psi_I\rangle$ di tutti gli autostati dell'operatore corrente, dove i $c_I(t)$ sono i coefficienti associati a ciascun autostato di corrente.

L'operatore densità

Introduciamo ora un concetto matematico fondamentale che emerge dalla natura intrinsecamente probabilistica della meccanica quantistica: l'*operatore densità*. Infatti, per quello che è stato esposto in precedenza, se non conosciamo lo stato nel quale si trova al tempo iniziale il sistema, possiamo solo calcolare la probabilità che il sistema si trovi in un dato stato $|\Psi_i(t)\rangle$ a un dato tempo t . Per trattare al meglio questo problema, possiamo considerare un ensemble di copie identiche del sistema, tutte con le stesse condizioni iniziali, e attribuir loro un vettore di stato $|\Psi_i(t)\rangle$. A ogni vettore di stato è associata una probabilità p_i e queste sono tali che: $\sum_i p_i = 1$.

Definiamo quindi l'*operatore densità* $\hat{\rho}$ come:

$$\hat{\rho}(t) = \sum_i p_i |\Psi_i(t)\rangle\langle\Psi_i(t)|$$

Da questa definizione, le probabilità p_i sono:

$$p_i = \langle\Psi_i|\hat{\rho}|\Psi_i\rangle$$

L'insieme $\{|\Psi_i\rangle, p_i\}$ di tutti i possibili microstati $|\Psi_i\rangle$ e delle probabilità ad essi associati p_i è detto *macrostato* o *stato misto*. Nel caso particolare in cui le probabilità abbiano la forma $\{1, 0, 0, \dots, 0\}$, si parla di *stato puro* e l'operatore densità si riduce a: $\hat{\rho}(t) = |\Psi(t)\rangle\langle\Psi(t)|$.

Nella rappresentazione di Schrödinger, dato un operatore $\hat{A}$, il suo valore di aspettazione è:

$$\langle\hat{A}\rangle_t = \sum_i p_i \langle\Psi_i(t)|\hat{A}|\Psi_i(t)\rangle = \text{Tr}\{\hat{\rho}(t), \hat{A}\}$$

mentre in rappresentazione di Heisenberg:

$$\langle\hat{A}\rangle_t = \langle\hat{A}_H(t)\rangle \equiv \text{Tr}\{\hat{\rho}(t=0)\hat{A}(t)\}$$

L'operatore densità gode delle seguenti proprietà:

- $\text{Tr}\{\hat{\rho}(t)\} = \sum_i p_i = 1$;
- $\hat{\rho} = \hat{\rho}^\dagger$;
- $\langle \Psi | \hat{\rho} | \Psi \rangle \geq 0$ per ogni $|\Psi\rangle$ nello spazio di Hilbert.

Per quanto riguarda la dinamica dell'operatore densità, essa può essere ricavata dall'equazione di Schrödinger dipendente dal tempo:

$$\begin{aligned} i\hbar \frac{d\hat{\rho}(t)}{dt} &= i\hbar \sum_i p_i \frac{d|\Psi_i(t)\rangle\langle\Psi_i(t)|}{dt} = \sum_i p_i [\hat{H}(t)|\Psi_i(t)\rangle\langle\Psi_i(t)| - |\Psi_i(t)\rangle\langle\Psi_i(t)|\hat{H}(t)] = \\ &= \hat{H}(t) \left(\sum_i p_i |\Psi_i(t)\rangle\langle\Psi_i(t)| \right) - \left(\sum_i p_i |\Psi_i(t)\rangle\langle\Psi_i(t)| \right) \hat{H}(t) = \hat{H}(t)\hat{\rho}(t) - \hat{\rho}(t)\hat{H}(t) \end{aligned}$$

ottenendo quindi l'equazione di Liouville-von Neumann:

$$i\hbar \frac{d\hat{\rho}(t)}{dt} = [\hat{H}(t), \hat{\rho}(t)]$$

Consideriamo ora un sistema in equilibrio con una riserva alla temperatura θ con la quale può scambiare particelle: la probabilità p_i che il sistema possa essere misurato nello stato $|\Psi_i\rangle_N$ di N particelle e che abbia energia E_i^N è data dalla *distribuzione gran canonica* all'equilibrio

$$p_i^{eq} = \frac{e^{-\beta(E_i^N - \bar{\mu}N)}}{\sum_M \sum_j e^{-\beta(E_j^M - \bar{\mu}M)}}$$

dove $\bar{\mu}$ è lo scambio di energia causato dall'aggiunta di una ulteriore particella, ovvero il *potenziale chimico* e $\beta = 1/k_B\theta$. Utilizzando gli autostati energia $|\Psi_i\rangle_N$ e la definizione dell'operatore densità, possiamo riscrivere quest'ultimo come

$$\hat{\rho}_G^{eq} = \frac{e^{-\beta(\hat{H} - \bar{\mu}\hat{N})}}{\text{Tr}\{e^{-\beta(\hat{H} - \bar{\mu}\hat{N})}\}} = \frac{e^{-\beta(\hat{H} - \bar{\mu}\hat{N})}}{\mathcal{Z}_G}$$

con la funzione di partizione gran canonica $\mathcal{Z}_G = \text{Tr}\{e^{-\beta(\hat{H} - \bar{\mu}\hat{N})}\}$ e $\hat{N}$ è un operatore che conta il numero di particelle in un dato stato a molti corpi:

$$\hat{N}|\Psi_i\rangle_N = \left(\sum_{M,j} M |\Psi_j\rangle_M \langle\Psi_j| \right) |\Psi_i\rangle_N = N |\Psi_i\rangle_N$$

Se il numero di particelle nel sistema è fissato, allora l'operatore densità all'equilibrio è quello canonico:

$$\hat{\rho}_C^{eq} = \frac{e^{-\beta\hat{H}}}{\text{Tr}\{e^{-\beta\hat{H}}\}} = \frac{e^{-\beta\hat{H}}}{\mathcal{Z}_C}$$

con la funzione di partizione canonica $\mathcal{Z}_C = \text{Tr}\{e^{-\beta\hat{H}}\}$. Le probabilità canoniche sono

$$p_i^{eq} = \frac{e^{-\beta E_i}}{\sum_j e^{-\beta E_j}}$$

Quando le N particelle sono non interagenti o debolmente interagenti e indistinguibili, allora l'energia totale del sistema E è la somma sulle energie di singola particella $\mathcal{E}_j$, $E = \sum_j N_j \mathcal{E}_j$, dove N_j è il numero medio delle particelle che occupano il livello di singola particella con energia $\mathcal{E}_j$, con $\sum_j N_j = N$. Possiamo quindi trovare il *numero di occupazione* medio dei livelli di singola particella:

$$N_j = \frac{1}{e^{(\mathcal{E}_j - \bar{\mu})/k_B\theta} + 1}$$

per gli elettroni (*distribuzione di Fermi-Dirac*), mentre per i fononi (*distribuzione di Bose-Einstein*):

$$N_j = \frac{1}{e^{\mathcal{E}_j/k_B\theta} - 1}$$

Nel limite di temperatura nulla, la distribuzione di Fermi-Dirac tende alla funzione a gradino di Heaviside centrata sul potenziale chimico: $f(\epsilon) \rightarrow \Theta(\bar{\mu} - \epsilon)$ per $\epsilon \rightarrow 0$. In questo limite, gli elettroni formano quello che viene definito un *gas di Fermi degeneri* e il potenziale chimico coincide con l' *energia di Fermi* E_F , ovvero l'energia del livello più energetico occupato.

Ora che abbiamo introdotto i concetti principali concernenti il problema del trasporto di elettroni, dobbiamo comprendere da quale prospettiva affrontarlo. Due degli approcci storicamente rilevanti al problema sono i seguenti:

1. la corrente elettrica è conseguenza di un campo elettrico esterno;
2. la corrente è determinata dalle condizioni al contorno alle superfici del conduttore analizzato. Il flusso di cariche incidente sugli estremi del conduttore causa un accumulo di portatori su tali estremi, generando un campo elettrico non omogeneo. In sintesi, il campo elettrico è conseguenza del flusso di corrente.

Le principali teorie che si basano sul primo approccio sono il *modello classico di Drude*, il *formalismo di Kubo* e l'*equazione di Boltzmann*; per il secondo, abbiamo la *teoria di Landauer*, concezione che approfondiremo nella sezione seguente.

3.2 L'approccio di Landauer

L'approccio di Landauer, come accennato in precedenza, si basa sull'assunto che il flusso di corrente è determinato dalle condizioni al contorno e quindi il campo elettrico dipende dalle condizioni al contorno.

Il sistema che andremo a studiare sarà quello di una pila che viene collegata agli elettrodi in modo da avere una corrente costante in media e quindi ottenere una condizione di

stazionarietà.

Il nostro scopo è quello di descrivere la dinamica del flusso di corrente, dinamica che è legata alla forma dell'Hamiltoniana $\hat{H}_{tot}$ del sistema. $\hat{H}_{tot}$ consta di tre termini:

- (i) l'Hamiltoniana degli elettroni e degli ioni del conduttore $\hat{H}_S$;
- (ii) l'Hamiltoniana degli elettroni e degli ioni della pila $\hat{H}_p$;
- (iii) l'Hamiltoniana di interazione tra pila e conduttore $\hat{H}_{int}$.

Quindi potremo scrivere: $\hat{H}_{tot} = \hat{H}_S + \hat{H}_p + \hat{H}_{int}$.

Lo studio della dinamica associata a questa Hamiltoniana viene semplificato attraverso varie approssimazioni, che ci apprestiamo ad enunciare.

3.2.1 Approssimazioni

Innanzitutto, descriviamo il nostro sistema: esso è costituito da due parti di conduttore collegate tra loro tramite una giunzione alle nanoscale ed entrambe sono aperte a due riserve aventi potenziale elettrochimico diverso tra loro. Il caso di un sistema aperto è particolarmente difficile da studiare, di conseguenza, utilizzando, come vedremo tra poco, la condizione di stazionarietà, possiamo ricondurre lo studio del sistema aperto a quello di un sistema chiuso infinito. Vediamo quindi i vari passaggi che ci portano a scrivere le equazioni del moto del sistema in esame.

1. Stato stazionario

Supponiamo che la soluzione dell'equazione per l'operatore densità sia un'unica soluzione stazionaria denotata con $\hat{\rho}_S^{ss}$. Stiamo assumendo quindi che la condizione di stazionarietà sia valida ad ogni istante:

$$\text{Tr}\{\hat{\rho}_S^{ss}\} = \langle \hat{I} \rangle = \text{costante}$$

2. Sistema quantistico chiuso

L'approssimazione precedente ci porta a considerare le riserve come fonti di emissione continua nel tempo di elettroni sotto forma di pacchetti d'onda che si avvicinano o allontanano dalla giunzione. Questa rappresentazione delle riserve ci permette di imporre la seguente condizione al contorno: gli elettroni vengono emessi a un tempo $t \rightarrow -\infty$ e lontano dalla giunzione sotto forma di pacchetti d'onda. Questi pacchetti d'onda si muovono da queste regioni di spazio, dette *guide*, verso la giunzione, diffondono sul potenziale della giunzione e si allontanano sotto forma di pacchetti d'onda che propagano nelle guide. Quelle appena citate sono dette *condizioni al contorno di scattering*.

In sintesi, assumeremo che il sistema sia *chiuso e infinito*, aggiungendo l'ulteriore condizione che si trovi in uno *stato puro*.

3. Approssimazione di campo medio

Assumiamo che l'Hamiltoniana $\hat{H}_S$ possa essere separata in due termini:

$$\hat{H}_S = \hat{H}^{mf} + \hat{V}$$

dove $\hat{H}^{mf}$ descrive l'interazione di elettroni, che consideriamo indipendenti, con gli ioni che non fanno parte della giunzione e $\hat{V}$ è l'energia di interazione tra gli elettroni e quella tra gli elettroni e gli ioni nella nanogiunzione. In questa approssimazione di campo medio di fatto consideriamo solo l'Hamiltoniana di particella singola per un elettrone. Ad esempio, possiamo considerare come Hamiltoniana del sistema quella di Hartree:

$$H_S = H_{Hartree}(\mathbf{r}) = -\frac{\hbar^2}{2m}\nabla^2 + V_H(\mathbf{r}) + V_{ext}(\mathbf{r})$$

con $V_H(\mathbf{r}) = e^2 \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}$ potenziale di Hartree e $V_{ext}(\mathbf{r})$ è un potenziale esterno statico. In generale, quindi, assumeremo che l'Hamiltoniana di particella singola possa essere scritta come

$$H_S = -\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})$$

4. Canali indipendenti e loro riempimento

Definiamo innanzitutto cosa si intende per canale.

Definizione 3.2.1. Un **canale** è un insieme di numeri quantici $\{E, \alpha\}$ che descrive una soluzione di scattering

Nel seguito della trattazione assumeremo che i canali siano *indipendenti*: questo si traduce nel fatto che gli elementi non diagonali della matrice densità possono essere considerati trascurabili, se non nulli.

Assumiamo inoltre che gli elettroni siano emessi dalla riserva sinistra con una distribuzione di equilibrio locale data da

$$f_L(E) = \frac{1}{e^{(E-\mu_L)/k_B\theta} + 1}$$

e in maniera analoga dalla riserva destra: $f_R(E) = \frac{1}{e^{(E-\mu_R)/k_B\theta} + 1}$, dove μ_L e μ_R sono i potenziali elettrochimici delle due riserve.

Date le distribuzioni locali, se chiamiamo $|\Psi^L\rangle$ uno stato di particella singola emessa dalla riserva sinistra e $|\Psi^R\rangle$ uno stato di particella singola emessa da quella destra, possiamo scrivere l'operatore densità stazionario di particella singola come somma di due operatori statistici, uno che descrive la dinamica degli elettroni diretti verso destra e l'altro quella degli elettroni che si muovono verso sinistra:

$$\hat{\rho}_S^{ss} = \sum_L |\Psi^L\rangle f_L \langle \Psi^L| + \sum_R |\Psi^R\rangle f_R \langle \Psi^R|$$

3.2.2 Condizioni di scattering al contorno

Consideriamo ora la dinamica degli elettroni: assumiamo che questi siano particelle libere nella direzione x , ma siano confinati in una buca di potenziale nelle altre due direzioni.

L'Hamiltoniana $\hat{H}_S$ soddisfa le condizioni al contorno asintotiche:

$$\begin{cases} \lim_{x \rightarrow -\infty} H_S = -\frac{\hbar^2}{2m} \nabla^2 + V_L(\mathbf{r}_\perp) \equiv H_L \\ \lim_{x \rightarrow +\infty} H_S = -\frac{\hbar^2}{2m} \nabla^2 + V_R(\mathbf{r}_\perp) \equiv H_R \end{cases}$$

con $V_L(\mathbf{r}_\perp)$ e $V_R(\mathbf{r}_\perp)$ potenziali generici che confinano gli elettroni nel piano trasverso $y - z$ degli elettrodi, rispettivamente, sinistro e destro.

Gli autostati delle Hamiltoniane H_L e H_R possono essere ricavati dall'equazione di Schrödinger associata: ad esempio, per H_L

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + V_L(\mathbf{r}_\perp) \right] \psi_{\alpha k}(\mathbf{r}) = E_\alpha(k) \psi_{\alpha k}(\mathbf{r})$$

che può essere separata in un'equazione longitudinale lungo la direzione x :

$$-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} e^{ikx} = \frac{\hbar^2 k^2}{2m} e^{ikx}$$

e una nel piano $y - z$:

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + V_L(\mathbf{r}_\perp) \right] u_\alpha(\mathbf{r}_\perp) = \epsilon_\alpha u_\alpha(\mathbf{r}_\perp)$$

La soluzione complessiva è data da:

$$\psi_{\alpha k}(\mathbf{r}) = \sqrt{\frac{1}{L_x}} u_\alpha(\mathbf{r}_\perp) e^{ikx} \quad -\infty < k < +\infty \quad (3.2)$$

con energie:

$$E_\alpha(k) = \epsilon_\alpha + \frac{\hbar^2 k^2}{2m} \equiv \epsilon_\alpha + \frac{mv_\alpha^2}{2}(k)$$

dove L_x è una lunghezza di normalizzazione e $v_\alpha(k) = \hbar k/m$ è la velocità dell'elettrone lungo l'asse x .

Per quanto riguarda le energie associate alla dinamica nel piano $y - z$, esse dipendono dalla forma del potenziale. Ad esempio, nel caso di una buca di potenziale quantistica rettangolare infinita avente dimensioni L_y e L_z , avremo:

$$\epsilon_{\alpha_y, \alpha_z} = \frac{\pi^2 \hbar^2}{2m} \left[\frac{\alpha_y^2}{L_y^2} + \frac{\alpha_z^2}{L_z^2} \right] \quad \alpha_y, \alpha_z \in \mathbb{N}$$

Gli stati dati dalla (3.2) soddisfano la condizione di ortonormalità:

$$\langle \psi_{\alpha k} | \psi_{\alpha' k'} \rangle = \frac{1}{L_x} \int dx e^{i(k-k')x} \langle u_\alpha | u_{\alpha'} \rangle = \frac{2\pi}{L_x} \delta(k' - k) \delta_{\alpha\alpha'}$$

Vediamo quindi che, dal punto di vista fisico, le soluzioni dell'equazione di Schrödinger per il piano $y - z$ descrivono modi trasversi aventi spettro di energia discreto ϵ_α , mentre quelle per l'asse x sono onde piane, aventi quindi spettro continuo.

Data un'energia E , il numero di canali a quell'energia $N_c(E)$ è dato dal numero di modi aventi energia ϵ_α minore di E , ovvero: $N_c(E) = \sum_\alpha \Theta(E - \epsilon_\alpha)$. Il numero di canali nelle guide di sezione S può essere stimato calcolando quanti canali sono ammessi alla lunghezza d'onda degli elettroni aventi energia E in un tratto lungo $\approx \sqrt{S}$ in tutte le direzioni. Se consideriamo come energia di riferimento quella di Fermi E_F , alla quale è associato il vettore d'onda k_F , il numero di canali è dell'ordine di $N_c(E_F) \approx k_F^2 S$. Per un sistema d -dimensionale, la sezione S ha $d-1$ dimensioni, così che il numero di canali è dell'ordine di: $N_c(E_F) \approx k_F^{d-1} S$.

Oltre agli stati (3.2), esistono soluzioni dell'equazione di Schrödinger della forma:

$$\psi_{\alpha k}(\mathbf{r}) = \sqrt{\frac{1}{L_x}} u_\alpha e^{-\tilde{k}x} \quad \tilde{k} > 0$$

con energie:

$$E_\alpha(k) = \epsilon_\alpha - \frac{\hbar^2 \tilde{k}^2}{2m}$$

Questi stati sono detti *evanescenti*, in quanto la loro probabilità decade esponenzialmente nello spazio.

3.2.3 Probabilità di trasmissione e riflessione

Determiniamo ora la soluzione generale H_S .

Per una data energia E dobbiamo così risolvere la seguente equazione di Schrödinger:

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) \right] \Psi_{\alpha k}(\mathbf{r}) = E \Psi_{\alpha k}(\mathbf{r})$$

Consideriamo un elettrone con energia E_i che per $x \rightarrow -\infty$ si trovava in un autostato iniziale $\psi_{ik_i}(\mathbf{r})$ dell'Hamiltoniana H_L alla stessa energia. L'elettrone ha quindi momento positivo $\hbar k_i$ tale che

$$E_i(k_i) = \epsilon_i + \frac{\hbar^2 k_i^2}{2m}$$

Ci aspettiamo che all'estremità dell'elettrodo destro si abbia una combinazione lineare degli autostati di H_R , ovvero:

$$\Psi_{ik_i}^+(\mathbf{r}) \rightarrow \sum_{f=1}^{N_c^R} \mathcal{T}_{if} \psi_{fk_f}(\mathbf{r}) \quad x \rightarrow +\infty \quad (3.3)$$

All'estremità della guida sinistra, oltre alle onde entranti, abbiamo le onde associate agli elettroni che vengono diffusi dalla giunzione. Di conseguenza, la soluzione sarà del tipo:

$$\Psi_{ik_i}^+(\mathbf{r}) \rightarrow \psi_{ik_i}(\mathbf{r}) + \sum_{f=1}^{N_c^L} \mathcal{R}_{if} \psi_{fk_f}(\mathbf{r}) \quad x \rightarrow -\infty \quad (3.4)$$

combinazione lineare dell'onda entrante e tutte le possibili onde riflesse con momenti $\hbar k_f$ orientate nel verso negativo dell'asse x .

Calcoliamo ora la corrente attraverso una superficie S perpendicolare al flusso di elettroni trasportata dall'onda $\Psi_{ik_i}^+(\mathbf{r})$ utilizzando la definizione (3.1) per un singolo elettrone. Il valore di aspettazione dell'operatore densità di corrente paramagnetica nello stato $|\Psi_{ik_i}^+\rangle$ è dato da:

$$\begin{aligned} \mathbf{j}(\mathbf{r}) &= \langle \Psi_{ik_i}^+ | \hat{\mathbf{j}}(\mathbf{r}) | \Psi_{ik_i}^+ \rangle = \frac{\hbar}{2im} \left[[\Psi_{ik_i}^+(\mathbf{r})]^* \frac{\partial \Psi_{ik_i}^+(\mathbf{r})}{\partial x} - \Psi_{ik_i}^+(\mathbf{r}) \frac{\partial [\Psi_{ik_i}^+(\mathbf{r})]^*}{\partial x} \right] = \\ &= \frac{\hbar}{m} \Im \left\{ [\Psi_{ik_i}^+(\mathbf{r})]^* \frac{\partial \Psi_{ik_i}^+(\mathbf{r})}{\partial x} \right\} \end{aligned}$$

Integriamo ora sul piano perpendicolare all'asse x per ricavare la corrente media $I(E_i)$ trasportata dallo stato con energia E_i

$$I(E_i) = e \langle \Psi_{ik_i}^+ | \hat{I} | \Psi_{ik_i}^+ \rangle = \frac{e\hbar}{2im} \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} dz \left[[\Psi_{ik_i}^+(\mathbf{r})]^* \frac{\partial \Psi_{ik_i}^+(\mathbf{r})}{\partial x} - \Psi_{ik_i}^+(\mathbf{r}) \frac{\partial [\Psi_{ik_i}^+(\mathbf{r})]^*}{\partial x} \right] \quad (3.5)$$

Osserviamo che il sistema in considerazione si trova in uno stato ideale e stazionario: di conseguenza, il valore della corrente è indipendente dalla posizione in cui si trova la sezione S . Possiamo quindi sostituire la (3.4) nella (3.5) per ottenere la corrente all'estremità della guida sinistra:

$$\begin{aligned} I_L(E_i) &= I_i(E_i) + \sum_{f=1}^{N_c^L} |\mathcal{R}_{if}|^2 I_f(E_i) = I_i(E_i) - \sum_{f=1}^{N_c^L} |\mathcal{R}_{if}|^2 |I_f(E_i)| \\ &\equiv I_i(E_i) \left(1 - \sum_{f=1}^{N_c^L} R_{if}(E_i) \right) \end{aligned} \quad (3.6)$$

dove abbiamo definito la seguente quantità dalla (3.2):

$$\begin{aligned} I_i(E_i) &= \frac{e\hbar}{2im} \int_{-\infty}^{+\infty} dy \int_{-\infty}^{+\infty} dz \left[[\Psi_{ik_i}^+(\mathbf{r})]^* \frac{\partial \Psi_{ik_i}^+(\mathbf{r})}{\partial x} - \Psi_{ik_i}^+(\mathbf{r}) \frac{\partial [\Psi_{ik_i}^+(\mathbf{r})]^*}{\partial x} \right] = \\ &= \frac{\hbar k_i}{m L_x} \end{aligned} \quad (3.7)$$

In maniera analoga, abbiamo per la corrente riflessa verso la guida sinistra I_f

$$I_f(E_i) = \frac{\hbar k_f}{m L_x} = \frac{v_f(k_f)}{L_x} \quad (3.8)$$

La quantità

$$R_{if}(E_i) \equiv |\mathcal{R}_{if}|^2 \frac{|I_f(E_i)|}{|I_i(E_i)|}$$

è la *probabilità di riflessione* che un'onda incidente sulla giunzione con momento $\hbar k_i$ venga diffusa verso la guida sinistra in uno avente momento $\hbar k_f$, assumendo che l'energia sia conservata nel processo.

Possiamo effettuare un calcolo analogo per la (3.3), considerando l'estremità della guida destra: sostituendo la (3.3) in (3.5), ricaviamo

$$I_R(E_i) = \sum_{f=1}^{N_c^R} |\mathcal{T}_{if}|^2 |I_f(E_i)| \equiv I_i(E_i) \sum_{f=1}^{N_c^R} T_{if}(E_i) \quad (3.9)$$

dove

$$T_{if}(E_i) \equiv |\mathcal{T}_{if}|^2 \frac{|I_f(E_i)|}{|I_i(E_i)|}$$

è la *probabilità di trasmissione* che un'onda con momento iniziale $\hbar k_i$ venga trasmessa attraverso la giunzione verso la guida destra in uno stato finale con momento $\hbar k_f$ alla stessa energia.

Trovandosi il sistema in uno stato stazionario ideale, come accennato in precedenza, le due correnti $I_L(E_i)$ e $I_R(E_i)$ devono essere identiche. Di conseguenza, dalle (3.6) e (3.9), possiamo ricavare la relazione tra le probabilità di riflessione e di trasmissione per un'onda incidente nella guida sinistra:

$$\sum_{f=1}^{N_c^R} T_{if}(E_i) + \sum_{f=1}^{N_c^L} R_{if}(E_i) = 1 \quad (3.10)$$

Il procedimento effettuato può essere ripetuto per un'onda incidente proveniente dalla guida destra: in questo caso, l'onda viene trasmessa nella guida sinistra e riflessa verso quella destra. Vale quindi:

$$\sum_{f=1}^{N_c^L} T_{if}(E_i) + \sum_{f=1}^{N_c^R} R_{if}(E_i) = 1 \quad (3.11)$$

Le due equazioni precedenti sono due equazioni di conservazione per il flusso di elettroni.

Infine, a causa dell'invarianza rispetto all'inversione temporale, le relazioni appena ricavate devono essere valide se invertiamo le velocità degli stati iniziale e finale, cambiando nel contempo la direzionalità del processo di scattering: ad esempio, un processo nel quale un'onda viene diffusa da destra con momento $\hbar k_i$ in uno stato destro con momento $\hbar k_f$ è equivalente a uno nel quale un'onda dalla destra con momento $-\hbar k_f$ viene diffusa in uno stato a sinistra con momento $-\hbar k_i$. Le relazioni tra probabilità di trasmissione e di riflessione per questi processi inversi possono essere ottenute dalle (3.10) e (3.11) scambiando i e f e L e R , ricavando quindi:

$$\begin{aligned} \sum_{i=1}^{N_c^L} T_{fi}(E_f) + \sum_{i=1}^{N_c^R} R_{fi}(E_f) &= 1 \\ \sum_{i=1}^{N_c^R} T_{fi}(E_f) + \sum_{i=1}^{N_c^L} R_{fi}(E_f) &= 1 \end{aligned}$$

3.2.4 Corrente totale

A causa dell'approssimazione di canali indipendenti, la corrente totale nel sistema è la somma di tutte le correnti trasportate da tutti i canali per tutte le energie. Supponiamo inoltre che le due classi di canali, quelli che si muovono da destra a sinistra e quelli che si muovono da sinistra a destra, siano equamente popolate. Siccome ciascun canale è un sistema 1-dimensionale, la densità degli stati per spin per un momento $\hbar k_i$ è:

$$g_i(E_i)dE_i = \frac{L_x}{2\pi} dk_i \iff g_i(E_i) = \frac{L_x}{2\pi} \frac{dk_i}{dE_i} = \frac{L_x}{2\pi \hbar v_\alpha(k_i)} \quad (3.12)$$

Tenendo conto che siamo in uno stato stazionario ideale, la corrente è costante in qualsiasi regione dello spazio: scegliamo quindi un punto qualsiasi all'estremità della guida destra. Dalle (3.6) e (3.9) e dalle relazioni equivalenti per l'onda incidente dalla guida destra, otteniamo la corrente totale integrando su tutte le energie:

$$I = 2e \int dE \left\{ \sum_{i=1}^{N_c^L} \sum_{f=1}^{N_c^R} g_i(E_i) I_i(E_i) T_{if}(E_i) - \sum_{i=1}^{N_c^R} g_i(E_i) I_i(E_i) \left[1 - \sum_{f=1}^{N_c^R} R_{if}(E_i) \right] \right\}$$

dove il fattore 2 è dovuto alla degenerazione di spin,

$$= 2e \int dE \left\{ \sum_{i=1}^{N_c^L} \sum_{f=1}^{N_c^R} g_i(E_i) I_i(E_i) T_{if}(E_i) - \sum_{i=1}^{N_c^R} \sum_{f=1}^{N_c^L} g_i(E_i) I_i(E_i) T_{if}(E_i) \right\}$$

dove abbiamo utilizzato la (3.11),

$$= e2 \frac{1}{2\pi \hbar} \int dE \{ \mathcal{T}_{LR}(E) - \mathcal{T}_{RL}(E) \} = 0$$

dove abbiamo adoperato la (3.8) e la (3.7) per la corrente per canale e la (3.12) per la densità di stati per spin per mostrare che il prodotto è la costante $1/2\pi \hbar$. Nell'ultima uguaglianza abbiamo definito il *coefficiente di trasmissione totale* a una data energia:

$$\mathcal{T}_{RL}(E) = \sum_{i=1}^{N_c^R} \sum_{f=1}^{N_c^L} T_{if}(E) = \sum_{i=1}^{N_c^R} \tilde{T}_i(E)$$

e

$$\mathcal{T}_{LR}(E) = \sum_{i=1}^{N_c^L} \sum_{f=1}^{N_c^R} T_{if}(E) = \sum_{i=1}^{N_c^L} T_i(E)$$

con $\tilde{T}_i(E) = \sum_{f=1}^{N_c^L} T_{if}(E)$ probabilità di trasmissione totale che un'onda incidente da destra venga diffusa in uno qualsiasi dei canali di sinistra e $T_i(E) = \sum_{f=1}^{N_c^R} T_{if}(E)$ probabilità di trasmissione totale che un'onda incidente da sinistra venga diffusa in uno qualsiasi

canali di destra. Infine, siccome il flusso di elettroni deve essere conservato, il coefficiente di trasmissione totale da sinistra a destra deve essere uguale al coefficiente di trasmissione da destra a sinistra:

$$\mathcal{T}_{RL}(E) = \mathcal{T}_{LR}(E) \equiv \mathcal{T}(E)$$

Se assumiamo invece che gli stati non siano equamente popolati, ma che gli stati che viaggiano verso destra siano popolati fino al potenziale elettrochimico μ_L , quelli che viaggiano verso sinistra siano popolati fino a μ_R e che la differenza tra i due sia $\mu_L - \mu_R = eV$, allora la corrente totale non sarà nulla. In particolare, sapendo che $I = e\text{Tr}\{\rho_S^{ss}\hat{I}\}$, con calcoli analoghi a quelli del caso precedente, otteniamo:

$$I = e\text{Tr}\{\rho_S^{ss}\hat{I}\} = \frac{e}{\pi\hbar} \int dE [f_L(E)\mathcal{T}_{LR}(E) - f_R(E)\mathcal{T}_{LR}(E)]$$

Imponendo la conservazione del flusso otteniamo:

$$I = \frac{e}{\pi\hbar} \int dE [f_L(E) - f_R(E)] \mathcal{T}(E) \quad (3.13)$$

3.2.5 Limite di bias nullo

Nel limite $\mu_L - \mu_R \rightarrow 0$ possiamo approssimare la (3.13) facendo lo sviluppo di Taylor della distribuzione locale sinistra:

$$f_L(E) = f_R(E) - \left. \frac{\partial f_R(E)}{\partial E} \right|_{\mu_R} (\mu_L - \mu_R) + \mathcal{O}[(\mu_L - \mu_R)^2]$$

e otteniamo dalla (3.13):

$$I = \frac{2e}{h} (\mu_L - \mu_R) \int dE \left(- \left. \frac{\partial f_R(E)}{\partial E} \right|_{\mu_R} \right) \mathcal{T}(E) \quad (3.14)$$

Se poniamo la temperatura θ a zero, la distribuzione di Fermi-Dirac locale $f_R(E)$ diventa la Θ di Heaviside e la sua derivata rispetto all'energia diventa una δ di Dirac centrata sul potenziale elettrochimico μ_R . Di conseguenza, la (3.14) diventa:

$$I = \frac{2e}{h} (\mu_L - \mu_R) \mathcal{T}(E = \mu_R) = \frac{2e^2}{h} \mathcal{T}(E = \mu_R) V$$

In alternativa, avremmo potuto sviluppare la distribuzione locale destra di Fermi-Dirac:

$$f_R(E) = f_L(E) - \left. \frac{\partial f_L(E)}{\partial E} \right|_{\mu_L} (\mu_R - \mu_L) + \mathcal{O}[(\mu_L - \mu_R)^2]$$

e sostituiamo questa nella (3.13):

$$I = \frac{2e^2}{h} \mathcal{T}(E = \mu_L) V$$

Notiamo che però la corrente varia in base a quale delle due distribuzioni viene considerata nel calcolo della corrente. Tuttavia, entrambe le equazioni sono valide in quanto ci troviamo nel limite $\mu_L - \mu_R \rightarrow 0$ e possiamo quindi assumere che $\mu_L \approx E_F + \epsilon$ e $\mu_R \approx E_F - \epsilon$ o, equivalentemente, $(\mu_L + \mu_R)/2 \approx E_F$. Il coefficiente di trasmissione valutato al potenziale elettrochimico sinistro o destro è quindi:

$$(E_F \pm \epsilon) = \mathcal{T}(E_F) \pm \left. \frac{\partial \mathcal{T}(E)}{\partial E_F} \right|_{E_F} \epsilon + (\epsilon^2)$$

Se il coefficiente di trasmissione varia lentamente con l'energia al livello di Fermi, possiamo trascurare le sue derivate rispetto all'energia e possiamo effettuare un'ulteriore approssimazione: $\mathcal{T}(E_F) \approx \mathcal{T}(E = \mu_R) \approx \mathcal{T}(E = \mu_L)$. Con tale approssimazione, troviamo infine:

$$I = \frac{2e^2}{h} \mathcal{T}(E_F) V$$

per $\mu_L - \mu_R \rightarrow 0$ e $\theta \rightarrow 0$.

3.2.6 Conduttanza di sonda doppia

Possiamo finalmente calcolare la conduttanza per il sistema considerato.

Assumiamo di misurare idealmente le differenze di potenziale elettrochimico alla posizione delle riserve: la differenza di potenziale è quindi $V \equiv (\mu_L - \mu_R)/e$. Se definiamo la conduttanza come nel caso classico, ovvero:

$$G = \frac{I}{V}$$

otteniamo:

$$G = \frac{I}{V} = \frac{2e^2}{h} \mathcal{T}(E_F) = \frac{2e^2}{h} \sum_{i=1}^{N_c^L} \sum_{f=1}^{N_c^R} T_{if}(E_F) = \frac{2e^2}{h} \sum_{i=1}^{N_c^R} \sum_{f=1}^{N_c^L} T_{if}(E_F)$$

per $\mu_L - \mu_R \rightarrow 0$ e $\theta \rightarrow 0$. Se il coefficiente di trasmissione è l'unità, la conduttanza assume il valore quantizzato:

$$G_0 = \frac{2e^2}{h}$$

Capitolo 4

Fluttuazioni della conduttanza

In questo capitolo finale applicheremo le matrici di Wishart allo studio della conduttanza di mezzi mesoscopici disordinati.

Le *fluttuazioni universali della conduttanza* nella teoria del trasporto di elettroni vennero predette da Altshuler [2] e Stone e Lee [23] nel 1985 e rivelate sperimentalmente da Webb e Washburn[34].

4.1 Matrici di scattering

Consideriamo nuovamente il modello di doppia sonda descritto nel capitolo precedente: immaginiamo che il sistema disordinato venga posto tra due guide, ovvero due conduttori perfetti, semi-infinite di lunghezza finita. Assumiamo le medesime condizione al contorno del capitolo precedente per quanto riguarda le superfici trasverse: la parte trasversa della funzione d'onda ha quindi spettro di energia quantizzato. Gli stati di scattering all'energia di Fermi soddisfano la relazione $k_F^2 = k_n^2 + k^2$, dove k_F è il momento di Fermi, k è il momento longitudinale e k_n è il momento trasverso. I vari k_n definiscono N canali; siccome ogni canale può trasportare due onde che propagano in direzioni opposte, la funzione d'onda su ciascun lato della regione disordinata è data da un vettore a $2N$ componenti: le prime N componenti sono le ampiezze delle onde che propagano verso destra e le N rimanenti le ampiezze delle onde che viaggiano verso sinistra.

La *matrice di scattering* S lega i flussi uscente ed entrante:

$$S \begin{pmatrix} \mathbf{I} \\ \mathbf{I}' \end{pmatrix} = \begin{pmatrix} \mathbf{O} \\ \mathbf{O}' \end{pmatrix} \quad (4.1)$$

dove $\mathbf{I}$, $\mathbf{O}$, $\mathbf{I}'$ e $\mathbf{O}'$ sono vettori di dimensione N che descrivono le ampiezze d'onda, rispettivamente, sul lato sinistro e su quello destro.

In una geometria quasi 1-D, ovvero nel caso di un conduttore che ha una lunghezza $L \gg W$, con W sezione trasversa, la matrice S è una matrice $2N \times 2N$:

$$S = \begin{pmatrix} r & t \\ t' & r' \end{pmatrix} \quad (4.2)$$

dove t è la matrice $N \times N$ di trasmissione e r è quella di riflessione $N \times N$. La formula di Landauer del capitolo precedente può essere riscritta in funzione della matrice t in questo modo:

$$G = \frac{2e^2}{h} \text{Tr}\{tt^\dagger\}$$

La conservazione del flusso di corrente implica che:

$$|\mathbf{I}|^2 + |\mathbf{I}'|^2 = |\mathbf{O}|^2 + |\mathbf{O}'|^2 \quad (4.3)$$

che è equivalente all'unitarietà della matrice S .

Definiamo ora la *matrice di trasferimento* M di dimensione $2N \times 2N$, che lega le ampiezze sul lato sinistro a quelle del lato destro:

$$M \begin{pmatrix} \mathbf{I} \\ \mathbf{O} \end{pmatrix} = \begin{pmatrix} \mathbf{O}' \\ \mathbf{I}' \end{pmatrix} \quad (4.4)$$

Analogamente alla matrice S , possiamo scrivere la M in termini di quattro blocchi $N \times N$:

$$M = \begin{pmatrix} m_1 & m_2 \\ m_3 & m_4 \end{pmatrix}$$

e dalle (4.1), (4.2) e (4.4), si ricavano le relazioni:

$$\begin{aligned} m_1 &= (t^\dagger)^{-1} & m_2 &= r'(t')^{-1} & m_3 &= -(t')^{-1}r & m_4 &= (t')^{-1} \\ t &= (m_1^\dagger)^{-1} & r &= -m_4^{-1}m_3 \end{aligned} \quad (4.5)$$

Possiamo quindi imporre il vincolo di conservazione del flusso per M dalla (4.3):

$$|\mathbf{I}|^2 - |\mathbf{O}|^2 = |\mathbf{O}'|^2 - |\mathbf{I}'|^2$$

il che si traduce, dalla (4.4), nel fatto che M preserva la norma iperbolica del vettore al quale viene applicata. Quindi, $M \in U(N, N)$, ovvero M è una matrice *pseudo-unitaria*. Possiamo esprimere la conservazione del flusso su M come:

$$M^\dagger \Sigma_z M = \Sigma_z \quad (4.6)$$

dove:

$$\Sigma_z = \begin{pmatrix} \mathbb{I} & \mathbf{0} \\ \mathbf{0} & -\mathbb{I} \end{pmatrix}$$

con $\Sigma_i, i = x, y, z$ generalizzazione $2N \times 2N$ delle matrici di Pauli.

Una matrice che soddisfa la condizione di conservazione del flusso può essere parametrizzata come:

$$M = \begin{pmatrix} u^{(1)} & 0 \\ 0 & u^{(3)} \end{pmatrix} \begin{pmatrix} (\mathbb{I} + \lambda)^{1/2} & \lambda^{1/2} \\ \lambda^{1/2} & (\mathbb{I} + \lambda)^{1/2} \end{pmatrix} \begin{pmatrix} u^{(2)} & 0 \\ 0 & u^{(4)} \end{pmatrix} \equiv UTV \quad (4.7)$$

dove $u^{(i)}$ sono matrici unitarie $N \times N$ arbitrarie e λ è una matrice reale diagonale con autovalori non-negativi $\{\lambda_1, \dots, \lambda_n\}$. La (4.7) è detta *scomposizione polare* di M perché le matrici unitarie fungono da "coordinate angolari", mentre i $\{\lambda_i\}_i$, detti anche *autoparametri* di M , da quelle "radiali". La (4.7) soddisfa la (4.6).

4.1.1 Relazione tra matrice di trasferimento e conduttanza

Possiamo ora esprimere le matrici di trasmissione in termini di questi parametri utilizzando le (4.1) e (4.5). Si ricava che per la matrici di trasmissione valgono:

$$\begin{aligned} t' &= m_4^{-1} = -(u^{(4)})^\dagger (\mathbb{I} + \lambda)^{-1/2} u^{(3)\dagger} \\ t &= (m_1^\dagger)^{-1} = u^{(1)} (\mathbb{I} + \lambda)^{-1/2} u^{(2)} \end{aligned}$$

Dalle ultime due equazioni otteniamo:

$$t^\dagger t = u^{(2)} (\mathbb{I} + \lambda)^{-1} u^{(2)\dagger} \quad t' t'^\dagger = u^{(4)} (\mathbb{I} + \lambda)^{-1} u^{(4)\dagger} \quad (4.8)$$

Da queste ultime relazioni possiamo osservare che le due matrici hanno gli stessi autovalori e sono legate da una trasformazione unitaria.

Consideriamo ora la matrice $Q = M^\dagger M$: utilizzando la scomposizione polare di M , si ha:

$$Q = \begin{pmatrix} u^{(2)} (\mathbb{I} + 2\lambda) u^{(2)\dagger} & 2u^{(2)} \sqrt{\lambda(\mathbb{I} + \lambda)} u^{(4)\dagger} \\ 2u^{(4)} \sqrt{\lambda(\mathbb{I} + \lambda)} u^{(2)\dagger} & u^{(4)} (\mathbb{I} + 2\lambda) u^{(4)\dagger} \end{pmatrix} \quad (4.9)$$

Q è Hermitiana positiva e presenta una simmetria $U(N, N)$ espressa dalla (4.6). Da questa simmetria segue che se $\mathbf{v}_n$ è un autovettore di Q con autovalore q_n , allora lo è pure $\Sigma_z \mathbf{v}_n$ con autovalore q_n^{-1} . Questo implica che gli autovalori di Q sono numeri reali positivi legati in questo modo:

$$q_n \equiv e^{v_n} \quad q_n^{-1} \equiv e^{-v_n}$$

Possiamo valutare esplicitamente Q^{-1} utilizzando la scomposizione polare:

$$Q^{-1} = \Sigma_z Q \Sigma_z = \begin{pmatrix} u^{(2)} (\mathbb{I} + 2\lambda) u^{(2)\dagger} & -2u^{(2)} \sqrt{\lambda(\mathbb{I} + \lambda)} u^{(4)\dagger} \\ -2u^{(4)} \sqrt{\lambda(\mathbb{I} + \lambda)} u^{(2)\dagger} & u^{(4)} (\mathbb{I} + 2\lambda) u^{(4)\dagger} \end{pmatrix}$$

Quindi, definita la matrice X come $X = \frac{1}{4}(Q + Q^{-1} - 2)$, quest'ultima diventa:

$$X = \begin{pmatrix} u^{(2)} \lambda u^{(2)\dagger} & \mathbf{0} \\ \mathbf{0} & u^{(4)} \lambda u^{(4)\dagger} \end{pmatrix} \quad (4.10)$$

Usando le equazioni (4.8), (4.9) e (4.10) otteniamo la seguente identità:

$$[\mathbb{I} + X]^{-1} = \begin{pmatrix} t^\dagger t & \mathbf{0} \\ \mathbf{0} & t' t'^\dagger \end{pmatrix}$$

Osserviamo che X è una matrice $2N \times 2N$ diagonale a blocchi con autovalori positivi $\{\lambda_i\}_i$ aventi doppia degenerazione.

Infine, possiamo fornire la formula di Landauer in funzione della matrice X :

$$G = \frac{2e^2}{h} \text{Tr}[\mathbb{I} + X]^{-1} = \frac{2e^2}{h} \sum_n \frac{1}{1 + \lambda_n}$$

4.2 Varianza della conduttanza

Chiamiamo T_n gli autovalori associati alla matrice t . L'unitarietà della matrice s implica che $0 \leq T_n \leq 1$ per ogni n .

Studieremo le proprietà A del trasporto della forma:

$$A = \sum_{n=1}^N f(T_n)$$

Le quantità come A vengono dette *statistiche lineari* sugli autovalori di trasmissione. Per facilitare la trattazione, consideriamo gli autovalori di $\mathbb{I} + X$:

$$T_n \equiv \frac{1}{1 + \lambda_n}$$

con $0 \leq \lambda_n < +\infty$ per ogni n , in modo da lavorare con una statistica lineare della forma:

$$A = \sum_{n=1}^N f(\lambda_n) \quad (4.11)$$

La distribuzione dei λ_n è data da:

$$P(\{\lambda_n\}) = \mathcal{Z}^{-1} \exp[-\beta \mathcal{H}(\{\lambda_n\})] \quad (4.12)$$

$$\mathcal{H}(\{\lambda_n\}) = - \sum_{i < j} \ln |\lambda_i - \lambda_j| + \sum_i V(\lambda_i)$$

dove $\mathcal{Z}$ è tale che P sia normalizzata a 1:

$$\mathcal{Z} = \int_0^\infty d\lambda_1 \cdots \int_0^\infty d\lambda_N \exp[-\beta \mathcal{H}(\{\lambda_n\})]$$

Il parametro β è l'indice di Dyson e varia in base alle proprietà di simmetria dell'insieme delle matrici di scattering:

- se il sistema è simmetrico rispetto a inversione temporale e lo scattering è indipendente dallo spin, $\beta = 1$ (caso ortogonale);
- se la simmetria rispetto all'inversione naturale viene rotta (per esempio, a causa di un campo magnetico), $\beta = 2$ (caso unitario);
- se inoltre lo scattering è fortemente dipendente dallo spin, $\beta = 4$ (caso simplettico).

La $\mathcal{H}$ costituisce l'Hamiltoniana per gli autovalori e contiene due termini:

- un termine repulsivo tra gli autovalori;
- un potenziale confinante V .

La *densità media* degli autovalori di trasmissione $\langle \rho(\lambda) \rangle$ è definita come la media sull'ensemble della densità microscopica $\rho(\lambda) = \sum_{n=1}^N \delta(\lambda - \lambda_n)$:

$$\langle \rho(\lambda) \rangle = \frac{\int d\lambda_1 \cdots \int d\lambda_N \rho(\lambda) \exp(-\beta \mathcal{H})}{\int d\lambda_1 \cdots \int d\lambda_N \exp(-\beta \mathcal{H})} \quad (4.13)$$

Introduciamo la *funzione di correlazione a due punti* come:

$$K_2(\lambda, \lambda') = \langle \rho(\lambda) \rangle \langle \rho(\lambda') \rangle - \langle \rho(\lambda) \rho(\lambda') \rangle$$

Dimostriamo ora che la funzione di correlazione è legata alla derivata funzionale di ρ rispetto a V . Siccome $\delta \mathcal{H} / \delta V(\lambda) = \rho(\lambda)$, derivando la (4.13), si ottiene:

$$\begin{aligned} \frac{\delta \langle \rho(\lambda) \rangle}{\delta V(\lambda')} &= -\beta \frac{\int d\lambda_1 \cdots \int d\lambda_N \rho(\lambda) \rho(\lambda') \exp(-\beta \mathcal{H})}{\int d\lambda_1 \cdots \int d\lambda_N \exp(-\beta \mathcal{H})} \\ &+ \beta \left(\frac{\int d\lambda_1 \cdots \int d\lambda_N \rho(\lambda) \exp(-\beta \mathcal{H})}{\int d\lambda_1 \cdots \int d\lambda_N \exp(-\beta \mathcal{H})} \right) \left(\frac{\int d\lambda_1 \cdots \int d\lambda_N \rho(\lambda') \exp(-\beta \mathcal{H})}{\int d\lambda_1 \cdots \int d\lambda_N \exp(-\beta \mathcal{H})} \right) = \\ &= -\beta \langle \rho(\lambda) \rho(\lambda') \rangle + \beta \langle \rho(\lambda) \rangle \langle \rho(\lambda') \rangle \end{aligned}$$

Otteniamo quindi:

$$K_2(\lambda, \lambda') = \frac{1}{\beta} \frac{\delta \langle \rho(\lambda) \rangle}{\delta V(\lambda')}$$

La statistica lineare A può essere riscritta in termini di ρ :

$$A = \int_0^\infty d\lambda a(\lambda) \rho(\lambda)$$

La media sull'ensemble è data da:

$$\langle A \rangle = \int_0^\infty d\lambda a(\lambda) \langle \rho(\lambda) \rangle$$

in modo tale che la varianza $\text{Var} A = \langle A^2 \rangle - \langle A \rangle^2$ diventa:

$$\text{Var} A = - \int_0^\infty d\lambda \int_0^\infty d\lambda' a(\lambda) a(\lambda') K_2(\lambda, \lambda') = - \frac{1}{\beta} \int_0^\infty d\lambda \int_0^\infty d\lambda' a(\lambda) a(\lambda') \frac{\delta \langle \rho(\lambda) \rangle}{\delta V(\lambda')} \quad (4.14)$$

4.2.1 Equazione integrale per la densità degli autovalori

Per valutare la derivata funzionale di $\langle \rho \rangle$, dobbiamo sapere la dipendenza di quest'ultima dal potenziale V . Un problema analogo è stato risolto da Dyson [14] nell'ambito dei livelli energetici in fisica nucleare, con la differenza che in quel caso non era presente il vincolo della positività degli autovalori. In quel caso, $\langle \rho \rangle$ era soluzione della seguente equazione integrale:

$$\int_{\mathbb{R}} d\lambda' \langle \rho(\lambda') \rangle \ln|\lambda - \lambda'| + \frac{\beta - 2}{2\beta} \ln \langle \rho(\lambda) \rangle = V(\lambda) + \text{costante} \quad (4.15)$$

dove la costante viene ricavata dalla condizione di normalizzazione: $\int d\lambda \langle \rho(\lambda) \rangle = N$. Il secondo termine a sinistra nella (4.15) è di ordine $N^{-1} \ln N$ relativamente al primo e può essere quindi trascurato. Per calcolare la funzione di correlazione a due punti è quindi sufficiente risolvere l'equazione approssimata:

$$\int_{\mathbb{R}} d\lambda' \langle \rho(\lambda') \rangle \ln|\lambda - \lambda'| = V(\lambda) + \text{costante} \quad (4.16)$$

Ribadiamo però che gli autovalori con i quali stiamo lavorando sono positivi: dobbiamo quindi cercare di introdurre in qualche modo questo vincolo nella (4.16).

Come primo passo operiamo il seguente cambio di variabili: $x_n \equiv \ln \lambda_n$. Siccome $\lambda_n \in (0, +\infty)$ allora $x_n \in (-\infty, +\infty)$. In questo modo, la distribuzione di probabilità e l'Hamiltoniana diventano:

$$\begin{aligned} \tilde{P}(\{x_n\}) &= \mathcal{Z} \exp \left[-\beta \tilde{\mathcal{H}}(\{x_n\}) \right] \\ \mathcal{H}(\{x_n\}) &= - \sum_{i < j} \ln |e^{x_i} - e^{x_j}| + \sum_i V(e^{x_i}) - \beta^{-1} \sum_i x_i \end{aligned} \quad (4.17)$$

Il fattore $\exp(\sum_i x_i) = \prod_i \lambda_i$ è lo Jacobiano della trasformazione da λ a x . Utilizzando la seguente identità:

$$\ln |e^x - e^{x'}| - \frac{1}{2}(x + x') = \ln \left| 2 \sinh \left(\frac{x - x'}{2} \right) \right| \quad (4.18)$$

$\tilde{\mathcal{H}}(\{x_n\})$ può essere riscritta in termini di una somma di un'interazione invariante sotto traslazioni con un potenziale a un corpo $\tilde{V}$:

$$\tilde{\mathcal{H}}(\{x_n\}) = - \sum_{i < j} \ln \left| 2 \sinh \left(\frac{x_i - x_j}{2} \right) \right| + \sum_i \tilde{V}(x_i) \quad (4.19)$$

$$\tilde{V}(x) = V(e^x) - \left[\frac{1}{2}(N - 1) + \beta^{-1} \right] x$$

Notiamo che $\tilde{P}(\{x_n\})$ soddisfa per ogni $i = 1, \dots, N$ l'equazione differenziale:

$$\frac{\partial \tilde{P}}{\partial x_i} + \beta \tilde{P} \frac{\partial \tilde{H}}{\partial x_i} = 0$$

Moltiplichiamo quest'ultima per $\delta(x - x_i)$, integriamo su $x_1, \dots, x_N$ e sommiamo sulle i :

$$\frac{d}{dx} \langle \tilde{\rho}(x) \rangle + \beta \int_{-\infty}^{\infty} dx_1 \cdots \int_{-\infty}^{\infty} dx_N \tilde{P}(\{x_n\}) \sum_{i=1}^N \delta(x - x_i) \frac{\partial \tilde{H}}{\partial x_i} = 0 \quad (4.20)$$

dove $\langle \tilde{\rho}(x) \rangle$ è la densità media degli x_n :

$$\langle \tilde{\rho}(x) \rangle = \left\langle \sum_{i=1}^N \delta(x - x_i) \right\rangle \equiv \int_{-\infty}^{\infty} dx_1 \cdots \int_{-\infty}^{\infty} dx_N \tilde{P}(\{x_n\}) \sum_{i=1}^N \delta(x - x_i)$$

Sostituendo la (4.19) in (4.20) si ottiene:

$$\frac{d}{dx} \langle \tilde{\rho}(x) \rangle + \beta \langle \tilde{\rho}(x) \rangle \frac{d}{dx} \tilde{V}(x) - \frac{1}{2} \beta P \int_{-\infty}^{\infty} dx' \langle \tilde{\rho}_2(x, x') \rangle \coth \left(\frac{x - x'}{2} \right) = 0 \quad (4.21)$$

dove $P \int$ indica il valore principale dell'integrale e $\langle \tilde{\rho}_2(x, x') \rangle$ è la *densità di coppia* definita come:

$$\langle \tilde{\rho}_2(x, x') \rangle = \left\langle \sum_{i \neq j} \delta(x - x_i) \delta(x' - x_j) \right\rangle$$

La densità di coppia è simmetrica nei suoi argomenti e soddisfa la condizione di normalizzazione:

$$\int_{-\infty}^{\infty} dx' \langle \tilde{\rho}_2(x, x') \rangle = (N - 1) \langle \tilde{\rho}(x, x') \rangle$$

Possiamo scomporre la densità di coppia in una parte correlata e in una non correlata:

$$\langle \tilde{\rho}_2(x, x') \rangle = \langle \tilde{\rho}(x) \rangle \langle \tilde{\rho}(x') \rangle [1 - y(x, x')] \quad (4.22)$$

con $y(x, x') = y(x', x)$ tale che: $\int_{-\infty}^{\infty} dx' y(x, x') \langle \tilde{\rho}(x') \rangle = 1$. Sostituendo la (4.21) nella (4.22), otteniamo:

$$\frac{d}{dx} \langle \tilde{\rho}(x) \rangle + \beta \langle \tilde{\rho}(x) \rangle \frac{d}{dx} [\tilde{V}(x) + U(x)] + \frac{\beta}{2} \langle \tilde{\rho}(x) \rangle I(x) = 0 \quad (4.23)$$

dove:

$$I(x) = P \int_{-\infty}^{\infty} dx' \langle \tilde{\rho}(x') \rangle y(x, x') \coth \left(\frac{x - x'}{2} \right)$$

$$U(x) = - \int_{-\infty}^{\infty} dx' \langle \tilde{\rho}(x') \rangle \ln \left| 2 \sinh \left(\frac{x - x'}{2} \right) \right|$$

La (4.23) è ancora un'equazione esatta. Dobbiamo quindi considerare ulteriori approssimazioni.

Esprimiamo $y(x, x')$ in termini della differenza $s = x' - x$ e somma $t = \frac{x+x'}{2}$:

$$y(x, x') = Y[(x + x')/2, x' - x] \equiv Y(t, s)$$

$Y(t, s)$ è pari nella componente s . La condizione di normalizzazione diventa quindi:

$$\int_{-\infty}^{\infty} ds Y(x + s/2, s) \langle \tilde{\rho}(x + s) \rangle = 1 \quad (4.24)$$

L'integrale $I(x)$ si trasforma in:

$$I(x) = -P \int_{-\infty}^{\infty} ds Y(x + s/2, s) \langle \tilde{\rho}(x + s) \rangle \coth(s/2) \quad (4.25)$$

Se sostituiamo nella (4.25) i seguenti sviluppi di Taylor:

$$Y(x + s/2, s) = Y(x, s) + \frac{s}{2} \frac{\partial}{\partial x} Y(x, s) + \dots \quad (4.26)$$

$$\begin{aligned}\langle \tilde{\rho}(x+s) \rangle &= \langle \tilde{\rho}(x) \rangle + s \frac{d}{dx} \langle \tilde{\rho}(x) \rangle + \dots \\ \coth(s/2) &= 2/s + s/6 + \dots\end{aligned}\quad (4.27)$$

otteniamo un'espansione di $I(x)$ nei momenti di ordine maggiore $Y_p(x) = \int_{-\infty}^{\infty} ds Y(x, s) s^p$ di $Y(x, s)$ rispetto a s .

A causa della simmetria $Y(t, s) = Y(t, -s)$, tutti i momenti di ordine dispari si annullano e solo i momenti di ordine pari forniscono un contributo non nullo. Seguendo la trattazione di [14], trascuriamo i momenti di ordine maggiore o uguale al secondo, commettendo così un errore dell'ordine di N^2 .

Siccome $Y_{-1}(x)$ e $Y_1(x)$ sono nulli, solo $Y_0(x)$ contribuisce a $I(x)$ al secondo ordine. Sostituendo le (4.26)-(4.27) nella (4.25) otteniamo:

$$I(x) = -\langle \tilde{\rho}(x) \rangle \frac{d}{dx} Y_0(x) - 2Y_0(x) \frac{d}{dx} \langle \tilde{\rho}(x) \rangle$$

Analogamente, sostituendo le (4.26)-(4.27) nella (4.24):

$$\langle \tilde{\rho}(x) \rangle Y_0(x) = 1$$

Mettendo assieme le ultime due equazioni, si ricava:

$$I(x) = -\langle \tilde{\rho}(x) \rangle^{-1} \frac{d}{dx} \langle \tilde{\rho}(x) \rangle$$

Quindi la (4.23) diventa:

$$(1 - \beta/2) \frac{d}{dx} \langle \tilde{\rho}(x) \rangle + \beta \langle \tilde{\rho}(x) \rangle \frac{d}{dx} [\tilde{V}(x) + U(x)] = 0$$

o, equivalentemente:

$$\int_{-\infty}^{\infty} dx' \langle \tilde{\rho}(x') \rangle \ln \left| 2 \sinh \left(\frac{x-x'}{2} \right) \right| + \frac{\beta-2}{2\beta} \ln \langle \tilde{\rho}(x) \rangle = V(e^x) - \frac{N}{2}x + \frac{\beta-2}{2\beta}x + \text{costante} \quad (4.28)$$

Ritorniamo alle variabili $\lambda = e^x$. Le densità sono legate da $\langle \tilde{\rho}(x) \rangle dx = \langle \rho(\lambda) \rangle d\lambda$, dalla quale otteniamo:

$$\langle \tilde{\rho}(x) \rangle = \lambda \langle \rho(\lambda) \rangle$$

Utilizzando l'identità (4.18) e la normalizzazione $\int \langle \rho \rangle d\lambda = N$, otteniamo dalla (4.28):

$$\int_0^{\infty} d\lambda' \langle \rho(\lambda') \rangle \ln |\lambda - \lambda'| + \frac{\beta-2}{2\beta} \ln \langle \rho(\lambda) \rangle = V(\lambda) + \text{costante} \quad (4.29)$$

che coincide con la (4.15), con la differenza che la (4.29) incorpora anche il vincolo di positività degli autovalori.

Una volta ricavata la (4.29), la dobbiamo studiare per ottenere la funzione di correlazione a due punti $K_2(\lambda, \lambda')$ nel limite $N \rightarrow +\infty$. La derivata funzionale $\delta \langle \rho \rangle / \delta V$ è uguale al kernel di:

$$\int_0^{\infty} d\lambda' \psi(\lambda') \ln |\lambda - \lambda'| = \phi(\lambda) + \text{costante} \quad (4.30)$$

dove la costante è determinata in modo che ψ abbia media nulla, cioè

$$\int_0^\infty d\lambda \psi(\lambda) = 0 \quad (4.31)$$

siccome la variazione in $\langle \rho \rangle$ deve avvenire a N costante. La soluzione integrale

$$\psi(\lambda) = \int_0^\infty d\lambda' \beta K_2(\lambda, \lambda') \phi(\lambda') \quad (4.32)$$

della (4.30) determina direttamente la funzione di correlazione a due punti.

Soluzione dell'equazione integrale

L'equazione integrale (4.30) può essere risolta analiticamente tramite una tecnica basata sulla trasformata di Mellin.

Definiamo:

$$\begin{aligned} x &= \ln \lambda \\ \tilde{\psi}(x) &= e^x \psi(e^x) \\ \tilde{\phi}(x) &= \phi(e^x) \\ \tilde{K}_2(x, x') &= e^{x+x'} K_2(e^x, e^{x'}) \end{aligned} \quad (4.33)$$

Nella nuova variabile x le equazioni (4.30) e (4.31) diventano:

$$\int_{-\infty}^\infty dx' \tilde{\psi}(x') C(x, x') = \tilde{\psi}(x) + \text{costante} \quad (4.34)$$

$$\int_{-\infty}^\infty dx \tilde{\psi}(x) = 0 \quad (4.35)$$

con kernel $C(x, x') = \ln|e^x - e^{x'}|$. L'inverso di C è uguale $\beta \tilde{K}_2$, ovvero:

$$\tilde{\psi}(x) = \int_{-\infty}^\infty dx' \beta \tilde{K}_2(x, x') \tilde{\phi}(x') \quad (4.36)$$

in accordo con la (4.32). Notiamo che la (4.34) è invariante sotto la trasformazione:

$$C(x, x') \rightarrow C(x, x') + f(x) + g(x')$$

in quanto l'integrale $\int dx' \tilde{\psi}(x') f(x)$ si annulla per la (4.35), mentre l'integrale $\int dx' \tilde{\psi}(x') g(x')$ contribuisce per una costante indipendente da x .

Scegliamo $f(x) = -x/2$ e $g(x') = -x'/2$. Il kernel integrale diventa invariante sotto traslazioni:

$$\ln|e^x - e^{x'}| - (x + x')/2 = \ln \left| 2 \sinh \left(\frac{x - x'}{2} \right) \right|$$

In questo modo abbiamo ridotto la (4.30) a un integrale di convoluzione:

$$\int_{-\infty}^\infty dx' \tilde{\psi}(x') \ln \left| 2 \sinh \left(\frac{x - x'}{2} \right) \right| = \tilde{\psi}(x) + \text{costante} \quad (4.37)$$

che può essere risolto tramite una trasformazione di Fourier. Definiamo la trasformata di Fourier di una funzione arbitraria $f(x)$ come:

$$f(k) = \int_{-\infty}^{\infty} dx e^{ikx} f(x) \quad (4.38)$$

La trasformata di Fourier del kernel fornisce:

$$\int_{-\infty}^{\infty} dx e^{ikx} \ln \left| 2 \sinh \frac{x}{2} \right| = -\frac{\pi}{k} \coth(\pi k)$$

Di conseguenza, la (4.37) ha la soluzione nello spazio k :

$$\tilde{\psi}(k) = -\frac{k}{\pi} \tanh(\pi k) \tilde{\phi}(k) \quad (4.39)$$

Concludiamo dalle (4.36) e (4.37) che la funzione di correlazione a due punti nella variabile x è invariante sotto traslazioni, $\tilde{K}_2(x, x') = \tilde{K}_2(x - x')$, con trasformata di Fourier:

$$\tilde{K}_2(k) \equiv \int_{-\infty}^{\infty} dx e^{ikx} \tilde{K}_2(x) = -\frac{k}{\pi\beta} \tanh(\pi k) \quad (4.40)$$

L'inversione della trasformata di Fourier fornisce:

$$\tilde{K}_2(k) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-ikx} \tilde{K}_2(k) = -\frac{1}{\pi^2\beta} \frac{d^2}{dx^2} \ln \left| \tanh \frac{x}{4} \right| \quad (4.41)$$

Usando la relazione (4.33) tra K_2 e $\tilde{K}_2$, troviamo dalla (4.41) l'espressione:

$$K_2(\lambda, \lambda') = \frac{1}{\pi^2\beta} \frac{\partial}{\partial \lambda} \frac{\partial}{\partial \lambda'} \ln \left| \frac{\sqrt{\lambda} - \sqrt{\lambda'}}{\sqrt{\lambda} + \sqrt{\lambda'}} \right| \quad (4.42)$$

Il kernel $K_2(\lambda, \lambda')$ ha una singolarità integrabile per λ e λ' . La parte non singolare è ottenuta derivando la (4.42), con il risultato:

$$K_2(\lambda, \lambda') = \frac{1}{2\pi^2\beta} (\lambda\lambda')^{-1/2} (\lambda + \lambda') (\lambda - \lambda')^{-2}$$

se $\lambda \neq \lambda'$.

4.2.2 Formula per la varianza di una statistica lineare

Possiamo ora valutare la varianza della statistica lineare $A = \sum_n a(\lambda_n)$.

Definiamo $\tilde{a}(x) = a(e^x)$. Usando la definizione (4.33), l'equazione (4.14) assume la forma:

$$\text{Var } A = - \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dx' \tilde{a}(x) \tilde{a}(x') \tilde{K}_2(x, x') \quad (4.43)$$

Definiamo la trasformata di Fourier di $\tilde{a}(x)$ secondo la (4.38):

$$\tilde{a}(k) = \int_{-\infty}^{\infty} dx e^{ikx} \tilde{a}(x) = \int_0^{\infty} d\lambda \lambda^{ik-1} a(\lambda)$$

La trasformata di Fourier rispetto alla x è una *trasformata di Mellin* rispetto a λ . L'equazione (4.43) diventa nello spazio k :

$$\text{Var}A = -\frac{1}{2\pi} \int_{-\infty}^{\infty} dk |\tilde{a}(k)|^2 \tilde{K}_2(k)$$

Sostituendo la (4.40) otteniamo la formula:

$$\text{Var}A = \frac{1}{\beta} \frac{1}{2\pi^2} \int_{-\infty}^{\infty} dk |\tilde{a}(k)|^2 k \tanh(\pi k) \quad (4.44)$$

per la varianza di una statistica lineare sugli autovalori di trasmissione.

4.2.3 Formula della varianza

La conduttanza G è legata agli autovalori di trasmissione dalla formula di Landauer:

$$G/G_0 = \sum_{n=1}^N T_n = \sum_{n=1}^N \frac{1}{1 + \lambda_n}$$

dove $G/G_0 = 2e^2/h$ è il quanto di conduttanza. La conduttanza è una statistica lineare della forma (4.11), con $a(\lambda) = (1 + \lambda)^{-1}$. La trasformata di Mellin di $a(\lambda)$, cioè la trasformata di Fourier di $(1 + e^{-x})$ è:

$$\tilde{a}(k) = \int_{-\infty}^{\infty} dx e^{ikx} \frac{1}{1 + e^x} = -\frac{i\pi}{\sinh(\pi k)}$$

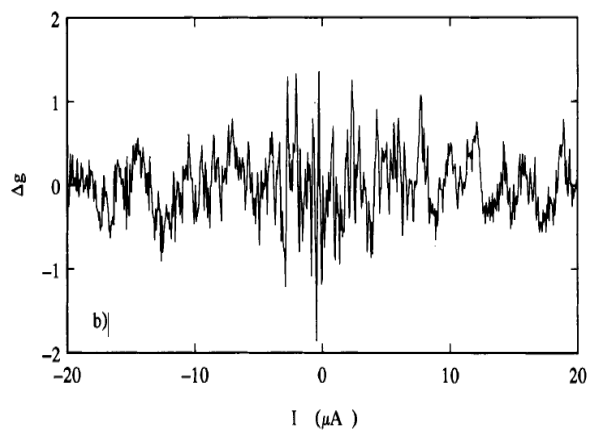
Sostituendo nella (4.44) fornisce la formula della varianza:

$$\text{Var}(G/G_0) = \beta^{-1} \int_{-\infty}^{\infty} dk \frac{k}{\sinh(2\pi k)} = \frac{1}{8} \beta^{-1}$$

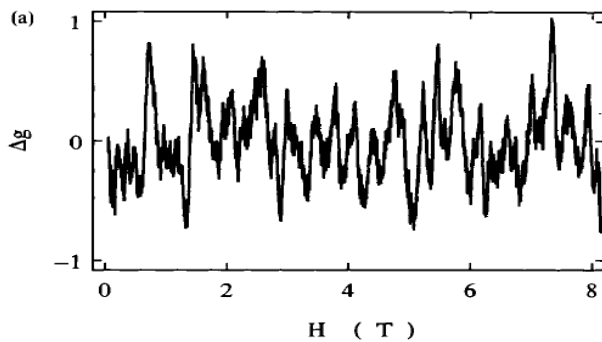
4.3 Teoria di Dorokhov-Mello-Pereyra-Kumar

Il caso generale *multicanale di N conduttori* è stato approfondito negli articoli [12] e [26]: in particolare nel primo lavoro viene ricavata l'equazione di diffusione per la distribuzione congiunta degli autovalori per un sistema non invariante sotto inversione temporale ($\beta = 2$), mentre nel secondo per un sistema invariante sotto inversione temporale, ma non dipendente dallo spin ($\beta = 1$). Entrambi sono quindi casi particolari dell' *equazione di Dorokhov-Mello-Pereyra-Kumar*: per un conduttore quasi 1-D di lunghezza L vale

$$\frac{\partial P}{\partial s} = \frac{2}{\gamma} \frac{\partial}{\partial \lambda_i} \lambda_i (1 + \lambda_i) J \frac{\partial}{\partial \lambda_i} J^{-1} P \equiv DP \quad (4.45)$$



(a)



(b)

(a) Fluttuazione della conduttanza di un cavo di antimonio di lunghezza $0,6\mu m$ come funzione della corrente nella spira ($\beta = 1$); (b) fluttuazione nella magnetoresistenza di un cavo d'oro in funzione del campo magnetico applicato ($\beta = 2$). Entrambe sono misurate tramite la tecnica della quadrupla sonda ©Elsevier

con $\gamma = \beta N + 2 - \beta$, $\beta \in \{1, 2, 4\}$, $s = L/l$, dove l è il libero cammino medio degli elettroni, J è la lo Jacobiano della matrice dello spazio degli autovalori :

$$J(\{\lambda_n\}) = \prod_{i < j} |\lambda_j - \lambda_i|^\beta \quad (4.46)$$

e D viene detto *operatore di Dorokhov-Mello-Pereyra-Kumar*.

Riportiamo ora la derivazione della (4.45) nel caso $\beta = 1$ fornita in [26].

4.3.1 Equazione di diffusione

Un insieme di conduttori di lunghezza macroscopica L viene descritto da un ensemble di matrici di trasferimento M , alle quali viene associata una probabilità differenziale $dP_L(M)$ dipendente parametricamente da L :

$$dP_L(M) = p_L(M) d\mu(M) = J(\lambda) \prod_{n=1}^N d\lambda_n d\mu(u) d\mu(v) \quad (4.47)$$

dove $p_L(R)$ è la densità di probabilità e $d\mu(M)$ è la misura invariante di Haar del gruppo preso in considerazione, con $J(\lambda) = \prod_{i < j} |\lambda_i - \lambda_j|$ e $d\mu$ è la misura invariante del gruppo unitario $U(N)$.

Dobbiamo imporre una condizione sulla densità di probabilità: supponiamo di collegare due fili di lunghezze L' e L'' e matrici di trasferimento M' e M'' , allora abbiamo $L = L' + L''$ e $M = M' M''$. Se poniamo $p_L(M')$ e $p_L(M'')$ le due densità di probabilità e assumiamo M' e M'' statisticamente indipendenti, la densità di probabilità risultante è data dalla convoluzione:

$$p_{L'+L''}(M) = \int p_{L''}(M M'^{-1}) p_{L'}(M') d\mu(M') \quad (4.48)$$

Utilizzando la notazione convenzionale della convoluzione, possiamo scrivere la (4.48) come

$$p_{L'+L''} = p_{L''} * p_{L'} \quad (4.49)$$

Si dice che la funzione $p_L(R)$ è *riproducibile sotto convoluzione*: l'effetto della convoluzione è la somma dei valori dei parametri L' e L'' , mantenendo la struttura della funzione inalterata. Ad esempio la *Gaussiana*:

$$p_L(x) = (2\pi DL)^{-1/2} e^{-(x^2/2DL)}$$

è riproducibile sotto convoluzione in $\mathbb{R}$.

Consideriamo la (4.48): poniamo $L'' = L$ e $L' = \delta L$, con δL piccolo, ma abbastanza grande per essere considerato macroscopico. Allora vale:

$$p_{L+\delta L}(M) = \int p_L(M M'^{-1}) p_{\delta L}(M') d\mu(M') \quad (4.50)$$

Possiamo osservare che, se conosciamo $p_{\delta L}$, possiamo costruire p_L tramite convoluzioni successive.

Definiamo ora il *coefficiente diffusivo* come:

$$D = \frac{1}{4N} \lim_{\delta L \rightarrow 0} \frac{\langle \text{Tr}(R'^{\dagger} R' - \mathbb{I}) \rangle_{\delta L}}{\delta L} \equiv \frac{1}{\alpha l} \quad (4.51)$$

ove l è una lunghezza e α un numero adimensionale.

Scegliamo ora come ansatz per la $p_{\delta L}(M')$ la distribuzione statistica che massimizza l'entropia di Shannon, definita come:

$$\mathcal{S}[p_{\delta L}] = - \int p_{\delta L}(M') \ln p_{\delta L}(M') d\mu(M')$$

con il vincolo (4.51). Il risultato è:

$$p_{\delta L}(M') = \exp \left[\tilde{\alpha} - \frac{\tilde{\beta}}{\delta L} \text{Tr}(M'^{\dagger} M') \right]$$

dove $\tilde{\alpha}$ e $\tilde{\beta}$ sono moltiplicatori di Lagrange.

Forniamo ora il risultato fondamentale della teoria di Mello-Pereyra-Kumar: *l'equazione di diffusione per la densità di probabilità*.

Innanzitutto, introduciamo le matrici Hermitiane:

$$H = MM^{\dagger} = \mathbb{I} + K$$

$$\tilde{H} = M^{\dagger}M = \mathbb{I} + \tilde{K}$$

Se consideriamo la decomposizione polare (4.7), per $\lambda \rightarrow 0$, $M \rightarrow UV$ e H e H' tendono ad assumere la forma della matrice identità. Il blocco H^{11} è dato da una matrice $N \times N$:

$$H^{11} = \mathbb{I} + 2u\lambda u^{\dagger} \equiv h$$

La distribuzione di probabilità $p_L(\lambda)$ è una funzione solo di λ e quindi solo funzione degli autovalori di h . Nella (4.50), $p_{\delta L}(M')$ dipenderà quindi dagli autovalori di h' e $p_L(MM'^{-1})$ da quelli di h'' dove:

$$h'' = (M''M''^{\dagger})^{11} = h + (M\Sigma_z \tilde{K}' \Sigma_z M^{\dagger})^{11} \quad (4.52)$$

Gli autovalori di h'' sono $1 + 2\lambda_i''$, quelli di h sono $1 + 2\lambda_i$, il secondo termine nella (4.52) può essere trattato come una perturbazione, in quanto $p_{\delta L}(M')$ nella (4.50) concentra λ' vicino all'origine. Possiamo scrivere $\lambda'' = \lambda + \delta\lambda$ ed esprimere la (4.50) come:

$$p_{L+\delta L}(\lambda) = \langle p_L(\lambda + \delta\lambda) \rangle_{\delta L} \quad (4.53)$$

dove viene effettuata una media su λ' , valutato sulla probabilità $dP_{\delta L}(M')$.

Per comodità, viene introdotta la densità di probabilità $w_L(\lambda)$ per la variabile λ , che dall'equazione per la misura invariante (4.47), è:

$$w_L(\lambda) = p_L(\lambda) J(\lambda) \quad (4.54)$$

$$\alpha \frac{\partial w_L(\lambda)}{\partial L} = \frac{2}{N+1} \sum_{i=1}^N \frac{\partial}{\partial \lambda_i} \left[\lambda_i (1 + \lambda_i) J(\lambda) \frac{\partial}{\partial \lambda_i} \frac{w_L(\lambda)}{J(\lambda)} \right] \quad (4.55)$$

con condizione iniziale:

$$w_0(\lambda) = \delta(\lambda)$$

Conseguenze notevoli dell'equazione di diffusione

Se consideriamo la (4.55), moltiplichiamo entrambi i lati dell'equazione per una funzione arbitraria $F(\lambda)$ e integriamo su λ otteniamo, per il valore atteso:

$$\langle F \rangle_L = \int F(\lambda) w_L(\lambda) d^N \lambda$$

la seguente equazione di evoluzione:

$$\begin{aligned} \alpha l \frac{\partial \langle F \rangle_L}{\partial L} = & \frac{2}{N+1} \left\langle \sum_{i=1}^N \left[\lambda_i (1 + \lambda_i) \frac{\partial^2 F}{\partial \lambda_i^2} + (1 + 2\lambda_i) \frac{\partial F}{\partial \lambda_i} \right] \right. \\ & \left. + \frac{1}{2} \sum_{i \neq j} \frac{1}{\lambda_i - \lambda_j} \left[\lambda_i (1 + \lambda_i) \frac{\partial F}{\partial \lambda_i} + \lambda_j (1 + \lambda_j) \frac{\partial F}{\partial \lambda_j} \right] \right\rangle_L \end{aligned}$$

Dalla (4.3.1) si ottiene un'equazione di evoluzione per $\langle \mathcal{T}^k \rangle$, con $\mathcal{T} = \sum_i (1 + \lambda_i)^{-1}$ fattore di trasmissione:

$$\begin{aligned} \alpha l \frac{\partial \langle \mathcal{T}^k \rangle_L}{\partial L} = & \frac{2}{N+1} \left\langle k(k-1) \mathcal{T}^{k-2} (\mathcal{T}_{(2)} - \mathcal{T}_{(3)}) \right. \\ & \left. - k \mathcal{T}^{k-1} \mathcal{T}_{(2)} - \frac{k}{2} \mathcal{T}^{k-1} (\mathcal{T}^2 - \mathcal{T}_{(2)}) \right\rangle_L \end{aligned}$$

dove:

$$\mathcal{T}_{(m)} = \sum_i \frac{1}{(1 + \lambda_i)^m}$$

Riscalando le varie quantità come:

$$\tau_m = \mathcal{T}_{(m)} / N$$

e considerando il limite $N \rightarrow \infty$, otteniamo, per i momenti $\tau = \tau_1$, la relazione di ricorsione:

$$\alpha l \frac{\partial \langle \tau^k \rangle_L}{\partial L} = -k \langle \tau^{k+1} \rangle_L$$

la soluzione della quale è:

$$\langle \tau^k \rangle_L = \frac{(\alpha l)^k}{(\alpha l + L)^k}$$

Se $N \gg 1$, il termine principale in $\langle \mathcal{T} \rangle_L$ è così:

$$\langle \mathcal{T} \rangle_L \approx \frac{N \alpha l}{\alpha l + L} \quad (4.56)$$

valida per $N \gg 1$ e $L/(\alpha l)$ arbitrario, ma $\ll N$. Così, per $1 \ll L/(\alpha l) \ll N$, la (4.56) indica che:

$$\frac{\langle \mathcal{T} \rangle_L}{N} \approx \frac{\alpha l}{L} \ll 1$$

Sappiamo che nel limite $N \rightarrow +\infty$, la conduttanza diventa:

$$\langle G \rangle_L \approx 2 \langle \mathcal{T} \rangle_L$$

Utilizzando la (4.56), otteniamo:

$$\langle G \rangle_L \approx \frac{2\alpha N l}{L} \quad l \ll L \ll N l$$

che coincide con la legge di Ohm in regime diffusivo, dove l è il *cammino libero medio* $N(k_F W)^2$, con W dimensione trasversa del conduttore.

4.3.2 Soluzione generale dell'equazione DMPK

La soluzione generale per l'equazione DMPK può essere ricavata utilizzando come punto di partenza i cosiddetti *modelli di Calogero-Sutherland* [8][31], i quali descrivono un sistema di N particelle interagenti poste su una retta, identificate dalle loro coordinate $\{x_i\}$, $i = 1, \dots, N$, con potenziale della forma $1/\sinh^2(x_i - x_j)$. Questi modelli sono particolarmente notevoli in quanto presentano N sono integrabili esattamente e, come vedremo, l'Hamiltoniana può essere mappata nella parte radiale di un operatore di Laplace-Beltrami definito su uno spazio simmetrico apposito. Queste proprietà sono dovute al fatto che il reticolo delle radici di un' *algebra di Lie* rimane invariato per potenziali costruiti tramite un qualche reticolo di radici associato canonicamente a un'algebra di Lie semplice.

Chiamiamo $\mathcal{V}$ lo spazio N -dimensionale definito dalle coordinate $\{x_i\}$ e $x = (x_1, \dots, x_N)$ un vettore in $\mathcal{V}$. Siano $R = \{\alpha\}$ un sistema di radici in $\mathcal{V}$ e R_+ il sottosistema delle radici positive di R . L'Hamiltoniana di Calogero-Sutherland assume la forma generale:

$$\mathcal{H} = -\frac{1}{2} \sum_{i=1}^N \frac{\partial^2}{\partial^2 x_i} + \sum_{\alpha \in R_+} \frac{g_\alpha^2}{\sinh^2(x_\alpha)} \quad (4.57)$$

dove gli accoppiamenti g_α^2 sono gli stessi per le radici equivalenti, ovvero per quelle radici che sono connesse tramite trasformazioni del *gruppo di Coxeter* W del sistema di radici.

I modelli CS rilevanti per la risoluzione del nostro problema sono quelli costruiti mediante il reticolo C_N . In questo caso il sistema di radici è $R = \{\pm 2e_i, \pm e_i \pm e_j, i \neq j\}$, con $\{e_1, \dots, e_N\}$ base canonica in $\mathbb{R}^n$ e il gruppo di Coxeter W coincide con il prodotto del gruppo delle permutazioni e delle trasformazioni che cambiano il segno del vettore $\{e_i\}$. L'Hamiltoniana corrispondente è:

$$\mathcal{H} = -\frac{1}{2} \sum_{i=1}^N \frac{\partial^2}{\partial^2 x_i} + \sum_{\alpha \in R_+} \frac{g_2^2}{\sinh^2(2x_i)} + \sum_{i < j} \left(\frac{g_1^2}{\sinh^2(x_i - x_j)} + \frac{g_1^2}{\sinh^2(x_i + x_j)} \right) \quad (4.58)$$

Utilizzando le identità valide per le funzioni iperboliche, la (4.58) può essere riscritta come:

$$\mathcal{H} = -\frac{1}{2} \sum_{i=1}^N \frac{\partial^2}{\partial^2 x_i} + \sum_i \frac{g_2^2}{\sinh^2(2x_i)} + c + 2g_1^2 \sum_{i < j} \frac{\sinh^2(2x_i) + \sinh^2(2x_j)}{(\cosh(2x_i) - \cosh(2x_j))^2} \quad (4.59)$$

con c costante.

La (4.58) può essere mappata nella parte radiale dell'operatore di Laplace-Beltrami B di uno spazio simmetrico apposito (vedi [28]):

$$H = \xi(x) \left[\frac{1}{2}(B + \rho^2) \right] \xi(x)^{-1} \quad (4.60)$$

con

$$B = [\xi(x)]^{-2} \sum_{k=1}^n \frac{\partial}{\partial x_k} [\xi(x)]^2 \frac{\partial}{\partial x_k} \quad (4.61)$$

e ρ costante. In questo modo possiamo legare l'operatore di DMPK a B (vedi [6]):

$$D = \frac{1}{2\gamma} [\xi(x)]^2 B [\xi(x)]^{-2} \quad (4.62)$$

In particolare, questa identificazione può avvenire solo se:

$$g_\alpha^2 = \frac{m_\alpha(m_\alpha - 2)}{8} |\alpha|^2 \quad (4.63)$$

dove $|\alpha|$ è la lunghezza della radice α e m_α la sua molteplicità. Nel caso preso in considerazione, $m_\alpha = \beta$ per le radici associate agli autovettori del tipo $\{\pm e_i, \pm e_j\}$ e $m_\alpha = 1$ per la radice avente lunghezza massima.

Se $\Phi_k(x)$, $x = \{x_1, \dots, x_N\}$, $k = \{k_1, \dots, k_N\}$ è un'autofunzione di B con autovalori k^2 , allora $\xi(x)^2 \Phi_k(x)$ dell'operatore DMPK con autovalore $k^2/(2\gamma)$. Le autofunzioni dell'operatore B sono i polinomi zonalì introdotti nel Capitolo 2; in particolare, utilizziamo le seguenti proprietà:

- tramite i polinomi zonalì possiamo generalizzare la trasformata di Fourier agli spazi simmetrici:

$$f(x) = \int \bar{f}(k) \Phi_k(x) \frac{dk}{|c(k)|^2} \quad (4.64)$$

e in nei tre casi di nostro interesse:

$$|c(k)|^2 = |\Delta(k)|^2 \prod_j \left| \frac{\Gamma(ik_j/2)}{\Gamma((1+ik_j)/2)} \right| \quad (4.65)$$

con:

$$|\Delta(k)|^2 = \prod_{m < j} \left| \frac{\Gamma\left(i \frac{k_m - k_j}{2}\right) \Gamma\left(i \frac{k_m + k_j}{2}\right)}{\Gamma\left(\frac{\beta}{2} + i \frac{k_m - k_j}{2}\right) \Gamma\left(\frac{\beta}{2} + i \frac{k_m + k_j}{2}\right)} \right|^2 \quad (4.66)$$

- per grandi valori di x , $\Phi_k(x)$ ha il seguente andamento asintotico:

$$\Phi_k(x) \sim \frac{1}{\xi(x)} \left(\sum_{r \in W} c(rk) e^{(rk, x)} \right) \quad (4.67)$$

dove rk è il vettore ottenuto applicando $r \in W$ su k .

Le equazioni (4.62) e (4.64) ci permettono di scrivere l'evoluzione di $P(\{x_n\}, s)$ dalle condizioni iniziali date (descritte dalla funzione $\bar{f}_0(k)$):

$$P(\{x_n\}, s) = |\xi(x)|^2 \int \bar{f}_0(k) e^{-\frac{k^2}{2\gamma}} \Phi_k(x) \frac{dk}{|c(k)|^2} \quad (4.68)$$

Inserendo l'espressione esplicita di $|c(k)|$ e utilizzando l'identità:

$$\left| \frac{\Gamma(\frac{1}{2} + i\frac{k}{2})}{\Gamma(i\frac{k}{2})} \right| = \frac{k}{2} \tanh\left(\frac{\pi k}{2}\right) \quad (4.69)$$

otteniamo la seguente espressione generale per $P(\{x_n\}, s)$ con condizioni iniziali balistiche, che coincidono con $\bar{f}_0(k) = \text{costante}$:

$$P(\{x_n\}, s) = |\xi(x)|^2 \int dk e^{-\frac{k^2}{2\gamma}} \frac{\Phi_k(x)}{|\Delta(k)|^2} \prod_j k_j \tanh\left(\frac{\pi k_j}{2}\right) \quad (4.70)$$

Nel caso $\beta = 2$, la forma esplicita di $\Phi_k(x)$ è stata ricavata in [5]:

$$\Phi_k(x) = \frac{\det[Q_m^j]}{\prod_{i < j} [(k_i^2 - k_j^2)(\sinh^2 x_i - \sinh^2 x_j)]} \quad (4.71)$$

dove gli elementi di matrice di Q sono:

$$Q_m^j = F\left(\frac{1}{2}(1 + ik_m), \frac{1}{2}(1 - ik_m), 1; -\sinh^2 x_j\right) \quad (4.72)$$

e $F(a, b, c; z)$ è la funzione ipergeometrica. Se poniamo la (4.72) nella (4.70) e utilizziamo l'identità:

$$P_\nu = F(-\nu, \nu + 1, 1; (1 - z)/2) \quad (4.73)$$

si ottiene:

$$P(\{x_n\}, s) = C(s) \prod_{i < j} (\sinh^2 x_j - \sinh^2 x_i) \prod_i (\sinh 2x_i) \times \det \left[\int_0^\infty dk e^{-\frac{k^2 s}{4n}} \tanh\left(\frac{\pi k}{2}\right) k^{2m-1} P_{\frac{1}{2}(ik-1)}(\cosh 2x_n) \right] \quad (4.74)$$

con $P_\nu(z)$ le funzioni di Legendre del primo tipo e $C(s)$ è una funzione che dipende solo da s e che può essere fissata imponendo una condizione di normalizzazione sulla densità di probabilità $P(\{x_n\}, s)$. La (4.74) si semplifica notevolmente quando consideriamo conduttori molto corti (*regime metallico*) o molto lunghi (*regime isolante*); mentre per il caso intermedio tra i due regimi, detto di *cross-over*, non è possibile effettuare alcuna approssimazione che semplifichi la (4.74).

Consideriamo quindi i due casi limite.

Regime isolante

Nel regime isolante vale $k \ll 1$, quindi possiamo utilizzare la seguente approssimazione:

$$\frac{\Gamma(\frac{\beta}{2} + iy)}{\Gamma(iy)} \sim_{y \rightarrow 0} iy. \quad (4.75)$$

Quindi, riscrivendo il prodotto $\prod_{i < j} (k_i^2 - k_j^2)$ e la somma sugli esponenziali in (4.67) come determinanti, si ottiene:

$$\begin{aligned} P(\{x_n\}, s) &= \prod_{i < j} |\sinh^2 x_j - \sinh^2 x_i|^{\beta/2} [(x_j^2 - x_i^2)] \\ &\times \prod_i [\exp(-x_i^2 \gamma / (2s)) x_i (\sinh 2x_i)^{1/2}] \end{aligned} \quad (4.76)$$

Ordinando gli x_n dal minore al maggiore e utilizzando il fatto che in questo regime $1 \ll x_1 \ll x_2 \ll \dots \ll x_N$, possiamo riscrivere la (4.76) come:

$$P(\{x_n\}, s) = \prod_{i=1}^N \exp [-(\gamma / (2s))(x_i - \bar{x}_i)^2] \quad (4.77)$$

con $\bar{x}_n = \frac{s}{\gamma}(1 + \beta(n - 1))$.

Regime metallico

In questo caso, $k \gg 1$ ed è quindi necessario adoperare la seguente espansione asintotica:

$$\frac{\Gamma(\frac{\beta}{2} + iy)}{\Gamma(iy)} \sim_{y \rightarrow \infty} |y|^{\frac{\beta}{2}} e^{\frac{i\pi\beta}{4}} \quad (4.78)$$

A differenza del caso precedente, l'integrazione rispetto a k è più difficoltosa e nel risultato finale vengono considerate solo le potenze di grado maggiore di $(x\sqrt{\frac{\gamma}{2s}})$; si ottiene quindi:

$$\begin{aligned} P(\{x_n\}, s) &= \prod_{i < j} |\sinh^2 x_j - \sinh^2 x_i|^{\beta/2} |x_j^2 - x_i^2|^{\frac{\beta}{2}} \\ &\times \prod_i [\exp(-x_i^2 \gamma / (2s)) (x_i \sinh 2x_i)^{1/2}] \end{aligned} \quad (4.79)$$

Bibliografia

- [1] R. Adamczak et al. On the marchenko-pastur and circular laws for some classes of random matrices with dependent entries. *Electronic Journal of Probability*, 16:1065–1095, 2011.
- [2] B. Altshuler. Fluctuations in the extrinsic conductivity of disordered conductors. *JETP lett*, 41(12):648–651, 1985.
- [3] Z. Bai and W. Zhou. Large sample covariance matrices without independence structures in columns. *Statistica Sinica*, pages 425–442, 2008.
- [4] C. Beenakker. Random-matrix theory of mesoscopic fluctuations in conductors and superconductors. *Physical Review B*, 47(23):15763, 1993.
- [5] C. Beenakker and B. Rajaei. Exact solution for the distribution of transmission eigenvalues in a disordered wire and comparison with random-matrix theory. *Physical Review B*, 49(11):7499, 1994.
- [6] C. Beenakker and B. Rajaei. Nonlogarithmic repulsion of transmission eigenvalues in a disordered wire. *Physical review letters*, 71(22):3689, 1993.
- [7] J. Bun, J.-P. Bouchaud, and M. Potters. Cleaning large correlation matrices: tools from random matrix theory. *Physics Reports*, 666:1–109, 2017.
- [8] F. Calogero. Ground state of a one-dimensional n-body system. *Journal of Mathematical Physics*, 10(12):2197–2200, 1969.
- [9] M. Caselle. Distribution of transmission eigenvalues in disordered wires. *Physical review letters*, 74(14):2776, 1995.
- [10] M. Di Ventra. Electrical transport in nanoscale systems. *etns*, 2008.
- [11] P. Diaconis and P. J. Forrester. Hurwitz and the origins of random matrix theory in mathematics. *Random Matrices: Theory and Applications*, 6(01):1730001, 2017.
- [12] O. Dorokhov. Localization and transmission coefficient for two coupled metal chains with disorder. *Solid state communications*, 44(6):915–919, 1982.
- [13] F. J. Dyson. The threefold way. algebraic structure of symmetry groups and ensembles in quantum mechanics. *Journal of Mathematical Physics*, 3(6):1199–1215, 1962.

- [14] F. J. Dyson. A class of matrix ensembles. *Journal of Mathematical Physics*, 13(1):90–97, 1972.
- [15] R. A. Fisher. The sampling distribution of some statistics obtained from non-linear equations. *Annals of Eugenics*, 9(3):238–249, 1939.
- [16] G. B. Folland. *Real analysis: modern techniques and their applications*, volume 40. John Wiley & Sons, 1999.
- [17] P. Forrester, N. Snaith, and J. Verbaarschot. Developments in random matrix theory. *Journal of Physics A: Mathematical and General*, 36(12):R1, 2003.
- [18] M. Girshick. On the sampling theory of roots of determinantal equations. *The Annals of Mathematical Statistics*, 10(3):203–224, 1939.
- [19] A. Haar. Über unendliche kommutative gruppen. *Mathematische Zeitschrift*, 33(1):129–159, 1931.
- [20] P. Hsu. On the distribution of roots of certain determinantal equations. *Annals of Eugenics*, 9(3):250–258, 1939.
- [21] A. Hurwitz. Ueber die entwicklungscoefficienten der lemniscatischen functionen. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 1897:273–276, 1897.
- [22] R. Landauer. Spatial variation of currents and fields due to localized scatterers in metallic conduction. *IBM Journal of research and development*, 1(3):223–231, 1957.
- [23] P. A. Lee and A. D. Stone. Universal conductance fluctuations in metals. *Physical review letters*, 55(15):1622, 1985.
- [24] V. Marchenko and L. Pastur. Eigenvalue distribution of some class of random matrices. *Matem. Sbornik*, 72:507, 1972.
- [25] M. L. Mehta. *Random matrices*. Elsevier, 2004.
- [26] P. Mello, P. Pereyra, and N. Kumar. Macroscopic approach to multichannel disordered conductors. *Annals of Physics*, 181(2):290–317, 1988.
- [27] R. J. Muirhead. *Aspects of multivariate statistical theory*, volume 197. John Wiley & Sons, 2009.
- [28] M. Olshanetsky and A. Perelomov. Quantum integrable systems related to lie algebras. *Physics Reports*, 94(6):313–404, 1983.
- [29] S. N. Roy. P-statistics or some generalisations in analysis of variance appropriate to multivariate problems. *Sankhya: The Indian Journal of Statistics*, pages 381–396, 1939.

- [30] J. Sherman and W. J. Morrison. Adjustment of an inverse matrix corresponding to a change in one element of a given matrix. *The Annals of Mathematical Statistics*, 21(1):124–127, 1950.
- [31] B. Sutherland. Quantum many-body problem in one dimension: Ground state. *Journal of Mathematical Physics*, 12(2):246–250, 1971.
- [32] G. Szeg. *Orthogonal polynomials*, volume 23. American Mathematical Soc., 1939.
- [33] T. Tao. *Topics in random matrix theory*, volume 132. American Mathematical Soc., 2012.
- [34] R. A. Webb, S. Washburn, C. Umbach, and R. Laibowitz. Observation of the Aharonov-Bohm oscillations in normal-metal rings. *Physical Review Letters*, 54(25):2696, 1985.
- [35] E. P. Wigner. On the statistical distribution of the widths and spacings of nuclear resonance levels. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 47, pages 790–798. Cambridge University Press, 1951.
- [36] E. P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. In *The Collected Works of Eugene Paul Wigner*, pages 524–540. Springer, 1993.
- [37] J. Wishart. The generalised product moment distribution in samples from a normal multivariate population. *Biometrika*, pages 32–52, 1928.
- [38] P. Yaskov. A short proof of the Marchenko–Pastur theorem. *Comptes Rendus Mathématique*, 354(3):319–322, 2016.