ISTITUTO LOMBARDO
ACCADEMIA di SCIENZE e LETTERE



Rendiconti

CLASSE DI SCIENZE MATEMATICHE E NATURALI

Vol. 150
(2016)

Istituto Lombardo di Scienze e Lettere

MILANO
pISSN 1974-6989
eISSN 2384-986X

ISTITUTO LOMBARDO - ACCADEMIA di SCIENZE e LETTERE

RENDICONTI

CLASSE DI SCIENZE MATEMATICHE E NATURALI

Vol. 150
(2016)

Istituto Lombardo di Scienze e Lettere

MILANO

ISTITUTO LOMBARDO - ACCADEMIA DI SCIENZE E LETTERE

RENDICONTI

CLASSE DI SCIENZE MATEMATICHE E NATURALI

Responsabile scientifico
Carlo Domenico Pagani

Comitato di redazione
Silvio Beretta
Stefano Maiorana
Cinzia Bearzot
Carlo Domenico Pagani

Pubblicato con il contributo di
Politecnico di Milano
Università Cattolica del Sacro Cuore, Milano
Università Commerciale L. Bocconi, Milano
Università degli Studi dell'Insubria, Varese
Università degli Studi di Milano
Università degli Studi Milano-Bicocca
Università degli Studi di Pavia

pISSN 1974-6989
eISSN 2384-986X
Copyright 2020

PAGEPress Publications
via A. Cavagna Sangiuliani, 5
27100 Pavia, Italy
T. +39.0382.464340
F: +39.0382.34872

ENTANGLEMENT QUANTISTICO[1]

ALBERTO RIMINI (*)

Nota presentata dal m.e. Attilio Rigamonti
(Adunanza del 21 gennaio 2016)

SUNTO. – La memoria riguarda una presentazione essenzialmente di carattere didattico-divulgativo, e a partire da principi primi, dello stato di *entaglement* di due particelle. Dopo alcuni richiami di carattere generale sulla meccanica quantistica e sulle teorie fisiche, ci si riferisce dapprima al caso di particella singola definendone stato, relazioni di incertezza e funzione d'onda nello spazio degli stati. Si considerano successivamente il sistema di due particelle e i suoi possibili stati, passando alla condizione di *entanglement* a partire dai primi articoli di Einstein Podolsky Rosen e Schroedinger. Viene quindi dettagliatamente descritto il problema della misurazione quantistica e il suo ruolo nello stato di *entanglement*. Si descrivono infine la soluzione ortodossa e le conclusioni più significative.

***

ABSTRACT. – This extended note deals with a pedagogical description of the entangled state of two particles, starting from first principles. After some general remarks about quantum mechanics and physical theories, the single particle case is discussed by defining state, uncertainty relations and wave function in the state space. The system of two particles is then considered, with its possible states, starting from the original papers by Einstein Podolsky Rosen and by Schroedinger. The quantum measurement problem is then introduced, together with its role in the entanglement state. Finally the orthodox solution and the relevant conclusions are drawn.

(*) Professore Emerito, Dipartimento di Fisica, Università di Pavia, Italia.
E-mail: info@istitutolombardo.it

[1] Questa nota è basata sulla presentazione orale da parte di Alberto Rimini (1937-2017) nell'adunanza del 21 gennaio 2016. La sua stesura è stata curata da Oreste Nicrosini e Attilio Rigamonti, i quali hanno raccolto gli appunti originali, nella forma di trasparenze proiettate, e li hanno opportunamente rielaborati. In questo lavoro particolare cura è stata dedicata a mantenere il testo il più possibile fedele all'originale: allo scopo, gli interventi effettuati sono stati limitati ad adattare la presentazione orale alla forma di articolo scientifico di carattere divulgativo, e ad integrare la presentazione stessa con alcune (poche) osservazioni esplicative. In particolare, è stato aggiunto un breve paragrafo di conclusioni fedele, a quanto ricordano gli editori, al pensiero di Alberto.

La meccanica quantistica

È la teoria fisica nata, tra il 1925 il 1926, per descrivere i sistemi atomici. È tuttora usata, in una forma compatibile con la relatività speciale, in ambito atomico, molecolare, dei sistemi estesi e in ambito subatomico.

La meccanica quantistica è sperimentalmente "verificata", nel senso che mai è risultata falsificata, in una grande varietà di campi nei quali è ritenuta applicabile e in uno sterminato numero di esperimenti e dispositivi tecnici.

Ne esistono varie formulazioni. Formulazioni diverse possono essere del tutto equivalenti quanto alle previsioni, oppure differire in modo quantitativamente non rilevabile (almeno attualmente).

Nell'esposizione che segue si farà riferimento alla formulazione standard, quella introdotta dai padri fondatori, basata sulla cosiddetta interpretazione di Copenhagen, comunemente esposta nei libri di testo e insegnata nei corsi di studio universitari.

Lo stato nelle teorie fisiche

Il concetto di *stato di un sistema* riveste un'importanza fondamentale in fisica. Lo stato di un sistema è la rappresentazione matematica più esauriente, ammessa dalla teoria nel cui ambito è descritto il sistema, delle proprietà non permanenti del sistema stesso.

In quanto rappresentazione delle proprietà non permanenti del sistema, lo stato, in generale, dipende dal tempo. Il tipo di ente matematico che rappresenta lo stato dipende sia dal tipo di sistema fisico considerato che dalla teoria usata per descriverlo.

Il sistema fisico più semplice: una singola particella

Nella meccanica razionale classica è detto *punto materiale*. Una particella non è necessariamente un oggetto piccolissimo. È un sistema che viene trattato come privo di qualsiasi struttura, privo anche della pura e semplice estensione spaziale, quindi puntiforme.

Di norma, trattare un oggetto come una particella significa introdurne un modello (il più semplice possibile) accettando un certo livello

di approssimazione nella descrizione del sistema considerato.

Un esempio: la dinamica del sistema solare può essere studiata trattando i pianeti, i satelliti, gli asteroidi come particelle.

Lo stato di una particella

In meccanica classica – È costituito dalla posizione $x, y, z = \boldsymbol{x}$ e dalla velocità $v_x, v_y, v_z = \boldsymbol{v}$ della particella.

A un certo tempo, la posizione dice *dove è* la particella, la velocità dice *come si muove*. In generale entrambe dipendono dal tempo.

In meccanica quantistica – S'intende che la particella è ora microscopica. Lo stato è costituito dalla funzione d'onda $\psi(\boldsymbol{x})$.

La funzione d'onda è una sorta di *nuvoletta*. La particella è sì puntiforme, ma la sua posizione è *distribuita* (spalmata) come è indicato dalla nuvoletta. È più probabile trovarla dove la nebbia è più fitta, e ha meno probabilità di essere rinvenuta dove la nebbia è meno fitta. La funzione d'onda e la conseguente nuvoletta dipendono dal tempo.

La particella è puntiforme ma la sua posizione è una nuvoletta! Questo può apparire un'insensatezza. In realtà si può dare un significato a quest'affermazione. Per ora accontentiamoci: in qualche senso la funzione d'onda indica dove si trova la particella.

Come si muove la particella?

Il fatto è che la funzione d'onda $\psi(\boldsymbol{x})$ è sì una nuvoletta ma è anche molto di più. Tanto che genera una *funzione sorella* $\varphi(\boldsymbol{v})$. La sorella è pur essa una nuvoletta, la quale informa su come è distribuita la velocità della particella, cioè come si muove la nuvoletta che descrive la posizione. La sorella $\varphi(\boldsymbol{v})$, come la funzione d'onda $\psi(\boldsymbol{x})$, dipende dal tempo.

Confronti

Riassumendo:

Meccanica classica – La posizione $\boldsymbol{x}$ e la velocità $\boldsymbol{v}$ sono due terne di numeri che indicano *senza incertezza* dove si trova la particella e come si muove.

Meccanica quantistica – La funzione d'onda $\psi(\boldsymbol{x})$ e la funzione sorella $\varphi(\boldsymbol{v})$ sono entrambe *diffuse*.

Meccanica classica – Lo stato è costituito dalle due terne $\boldsymbol{x}$ e $\boldsymbol{v}$ e queste (a un dato tempo) sono indipendenti l'una dall'altra.

Meccanica quantistica – Lo stato è costituito dall'una $\psi(\boldsymbol{x})$ o dall'altra funzione $\varphi(\boldsymbol{v})$, le quali si implicano a vicenda.

Meccanica classica – L'evoluzione dello stato (la posizione e la velocità) è *deterministica* ed è dettata dall'*equazione di Newton*.

Meccanica quantistica – L'evoluzione dello stato (la funzione d'onda oppure sua sorella) è pure *deterministica* ed è dettata dall'*equazione di Schrödinger*.

La relazione d'incertezza

In meccanica quantistica la posizione e la velocità sono entrambe nuvolette. Si potrebbe tuttavia pensare di ridurre a nostro piacimento l'estensione delle nuvolette fino a soddisfare la nostra repulsione per l'incertezza.

Questo non è possibile, o meglio non è possibile per entrambe le nuvolette. La relazione tra $\psi(\boldsymbol{x})$ e $\varphi(\boldsymbol{v})$ è tale che se noi riduciamo l'estensione di una necessariamente aumenta l'estensione dell'altra: le incertezze non possono essere entrambe ridotte *ad libitum*.

Lo spazio degli stati

Una funzione $\psi(\boldsymbol{x})$ è una funzione d'onda, cioè costituisce uno stato della particella, solo se soddisfa a certi requisiti.

Allora appartiene allo spazio degli stati della particella.

Per il seguito interessa solo una proprietà fondamentale dello spazio degli stati. Se $\psi_a(\boldsymbol{x})$ e $\psi_b(\boldsymbol{x})$ sono due possibili stati della particella, anche la somma o sovrapposizione $\psi(\boldsymbol{x}) = \psi_a(\boldsymbol{x}) + \psi_b(\boldsymbol{x})$ è un possibile stato.

Una proprietà analoga vale per gli stati di ogni e qualsiasi sistema quantistico.

Il sistema di due particelle e il suo stato

Consideriamo ora un sistema appena un po' più complicato, il sistema costituito da due particelle.

Diamo loro un nome, chiamiamole 1 e 2.

In meccanica classica – L'estensione del concetto di stato al nuovo sistema è del tutto naturale. lo stato di *una* particella è costituito dalla posizione $\boldsymbol{x}$ e dalla velocità $\boldsymbol{v}$, lo stato di *due* particelle è costituito dalle posizioni e dalle velocità di entrambe, $\boldsymbol{x}_1$, $\boldsymbol{v}_1$, $\boldsymbol{x}_2$, $\boldsymbol{v}_2$.

In meccanica quantistica – Di nuovo l'estensione del concetto di stato è del tutto naturale. lo stato di *una* particella è costituito dalla funzione d'onda $\psi(\boldsymbol{x})$, lo stato di *due* particelle è costituito dalla funzione d'onda $\psi(\boldsymbol{x}_1, \boldsymbol{x}_2)$. Anche $\psi(\boldsymbol{x}_1, \boldsymbol{x}_2)$ è una nuvoletta, una nuvoletta nello spazio delle posizioni delle due particelle (a 6 dimensioni). Indica come è distribuita la coppia $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ delle loro posizioni.

Stati delle due particelle e stato del sistema di due particelle

In meccanica classica – Se gli stati di due particelle sono $\boldsymbol{x}_1$, $\boldsymbol{v}_1$ e $\boldsymbol{x}_2$, $\boldsymbol{v}_2$, lo stato del sistema delle due particelle è $\boldsymbol{x}_1$, $\boldsymbol{v}_1$, $\boldsymbol{x}_2$, $\boldsymbol{v}_2$, *e viceversa*.

In meccanica quantistica – Se gli stati di due particelle sono $\psi(\boldsymbol{x}_1)$ e $\psi(\boldsymbol{x}_2)$, lo stato del sistema delle due particelle è il prodotto $\psi(\boldsymbol{x}_1, \boldsymbol{x}_2) = \psi(\boldsymbol{x}_1)\,\psi(\boldsymbol{x}_2)$. Ma, in generale, a uno stato $\psi(\boldsymbol{x}_1, \boldsymbol{x}_2)$ di un sistema di due particelle *non corrispondono* stati definiti delle due particelle. Il motivo per cui in *meccanica quantistica* succede questo è semplice: in generale una funzione di $\boldsymbol{x}_1$ e $\boldsymbol{x}_2$ *non è* il prodotto di una funzione di $\boldsymbol{x}_1$ e una funzione di $\boldsymbol{x}_2$.

Ad esempio $\psi(\boldsymbol{x}_1, \boldsymbol{x}_2) = \psi_a(\boldsymbol{x}_1)\,\psi_b(\boldsymbol{x}_2) + \psi_c(\boldsymbol{x}_1)\,\psi_d(\boldsymbol{x}_2) \neq \psi_e(\boldsymbol{x}_1)\,\psi_f(\boldsymbol{x}_2)$ (a meno che non si verifichino condizioni particolari). Allora lo stato $\psi(\boldsymbol{x}_1, \boldsymbol{x}_2)$ è *entangled*, nel senso che mentre le due particelle *come un tutto* si trovano in uno stato definito $\psi(\boldsymbol{x}_1, \boldsymbol{x}_2)$, ciascuna di esse individualmente prese non si trova in uno stato definito; le due particelle sono in un certo senso "intrecciate".

Questa situazione non è eccezionale, priva di interesse pratico. Al contrario è piuttosto comune.

Stato del sistema delle due particelle e stati delle due particelle

Un esempio quantistico – Due particelle, siano 1 e 2, si muovono nello spazio; siano $\boldsymbol{x}_1$ e $\boldsymbol{x}_2$ le loro coordinate.

Come possiamo descrivere il sistema?

Il sistema *come un tutto* si muove nello spazio e come tale è in uno stato definito. Esso ha una sua *struttura interna* e come tale è in uno stato definito.

La funzione d'onda $\psi(\boldsymbol{x}_1, \boldsymbol{x}_2)$ è il prodotto di una funzione d'onda $\Phi(\boldsymbol{X})$ che descrive il movimento delle due particelle nello spazio, $\boldsymbol{X}$ essendo la coordinata del loro centro di massa, e di una funzione d'onda $(\boldsymbol{x})$ che descrive il suo stato interno, $\boldsymbol{x}$ essendo la coordinata relativa delle due particelle: $\psi(\boldsymbol{x}_1, \boldsymbol{x}_2) = \Phi(\boldsymbol{X})\ (\boldsymbol{x})$.

Tuttavia, in generale, $\Phi(\boldsymbol{X})\ (\boldsymbol{x}) \neq \psi_1(\boldsymbol{x}_1)\ \psi_2(\boldsymbol{x}_2)$, cosicché le due particelle *come un tutto* si trovano in uno stato ben definito, mentre ciascuna di esse no.

Un esempio classico – Trattiamo, con la meccanica classica, il solo sistema terra-luna. Anche in questo caso conviene trattare il sistema per mezzo della posizione del centro di massa del sistema terra-luna $\boldsymbol{X}$, della posizione relativa della luna rispetto alla terra $\boldsymbol{x}$ e delle corrispondenti velocità $\boldsymbol{V}$ e $\boldsymbol{v}$. Ma, a differenza che in meccanica quantistica, istante per istante i valori di $\boldsymbol{X}$, $\boldsymbol{V}$ e $\boldsymbol{x}$, $\boldsymbol{v}$ *determinano* valori corrispondenti di $\boldsymbol{x}_{\text{terra}}$, $\boldsymbol{v}_{\text{terra}}$ e $\boldsymbol{x}_{\text{luna}}$, $\boldsymbol{v}_{\text{luna}}$, cioè gli stati della terra e della luna separatamente presi.

Torniamo al caso delle due particelle microscopiche. È un caso di *entanglement*, le singole particelle non sono in stati definiti. La circostanza non è sconvolgente. Il sistema è un tutt'uno, che ha una sua identità come tale, il fatto che le due particelle perdano in qualche senso la loro identità ci appare quasi naturale.

Ma esistono due tipi di situazione in cui al contrario l'*entanglement* è difficile da accettare.

Entanglement problematici

Nel 1935, a dieci anni di distanza dall'introduzione della meccanica quantistica, uscirono due articoli che evidenziavano certi aspetti "paradossali" o almeno sorprendenti dell'*entanglement* quantistico.

Uno dei due articoli, di Albert Einstein, Boris Podolsky e Nathan Rosen, discuteva gli aspetti sorprendenti dell'*entanglement* tra oggetti lontani, aspetti ora noti come "paradosso" EPR.

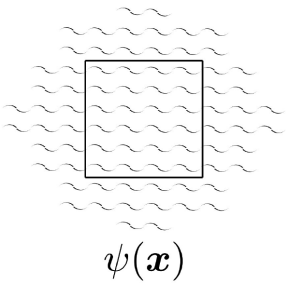
Nell'altro articolo, firmato da Erwin Schrödinger, si mostrava come dar luogo a uno stato *entangled* tra un oggetto microscopico e

uno macroscopico, un gatto. In questo secondo caso, il coinvolgimento di un essere vivente, il *gatto di Schrödinger*, esaltava l'inconcepibilità del risultato.

MISURAZIONE

A questo punto dobbiamo fare una digressione per descrivere un elemento imprescindibile dell'apparato formale della meccanica quantistica.

Secondo i principi della meccanica quantistica standard le *misurazioni* eseguite sul sistema considerato sono *processi di natura speciale* per i quali valgono *regole speciali*.

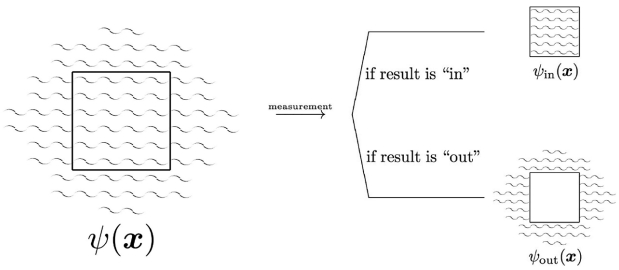


Un esempio semplice ma paradigmatico. Un rivelatore accerta se la particella studiata è dentro un certo volume o no.

Entrambi i risultati sono possibili:
"in" – la particella è dentro,
"out" – la particella non è dentro, cioè è fuori.

Solo se la nuvoletta è tutta dentro o tutta fuori, il risultato è certamente "in" o "out".

Ma che cosa succede allo stato in conseguenza della misurazione?



La funzione d'onda viene *ridotta* dalla misurazione. In corrispondenza della misurazione emerge il carattere *non deterministico*, probabilistico della meccanica quantistica, sia per quanto riguarda il valore della grandezza misurata, sia per quanto riguarda la corrispondente evoluzione, la riduzione.

Entanglement – Il paradosso EPR

Una molecola biatomica è un sistema di due particelle, i due atomi. La sua funzione ha la forma

$$\Phi(X)\ \ (x)$$

dove i simboli sono già stati illustrati in precedenza.

Lo stato interno della molecola non è unico, in generale ne esistono molti, consideriamone due, $_a(x)$ e $_b(x)$.

Per illustrare l' *entanglement* EPR ci servono due molecole, di nome 1 e 2.

Gli stati possibili della molecola 1 sono

$$\Phi_1(X_1)\ \ _a(x_1) \text{ e } \Phi_1(X_1)\ \ _b(x_1),$$

quelli della molecola 2 sono

$$\Phi_2(X_2)\ \ _a(x_2) \text{ e } \Phi_2(X_2)\ \ _b(x_2).$$

Considerando il sistema costituito dalle due molecole, una *qui* e l'altra *là*, stati possibili sono

$$(\Phi_{qui}(X_1)\ \ _a(x_1))\ (\Phi_{là}(X_2)\ \ _b(x_2)) = (\Phi_{qui}(X_1)\ \Phi_{là}(X_2))\ (\ _a(x_1)\ \ _b(x_2)),$$
$$(\Phi_{qui}(X_1)\ \ _b(x_1))\ (\Phi_{là}(X_2)\ \ _a(x_2)) = (\Phi_{qui}(X_1)\ \Phi_{là}(X_2))\ (\ _b(x_1)\ \ _a(x_2)).$$

Ma lo stato può anche essere

$$(\Phi_{qui}(X_1)\ \Phi_{là}(X_2))\ (\ _a(x_1)\ \ _b(x_2) +\ _b(x_1)\ \ _a(x_2)).$$

E qui appaiono elementi di disagio.

Ciò che preoccupava Einstein, Podolsky e Rosen è quanto segue. Misuriamo lo stato interno della molecola 1, quella che si trova *qui*. Sono possibili due risultati:

- la molecola 1 è nello stato interno $_a(x_1)$ oppure
- la molecola 1 è nello stato interno $_b(x_1)$.

Supponiamo che il risultato sia $_a(x_1)$. Che cosa succede? Ha luogo la riduzione, cioè

$$(\Phi_{qui}(X_1)\ \Phi_{là}(X_2))\ (\ _a(x_1)\ \ _b(x_2) +\ _b(x_1)\ \ _a(x_2)) \text{ à } \Phi_{qui}(X_1)\ \Phi_{là}(X_2)\ \ _a(x_1)\ \ _b(x_2)$$

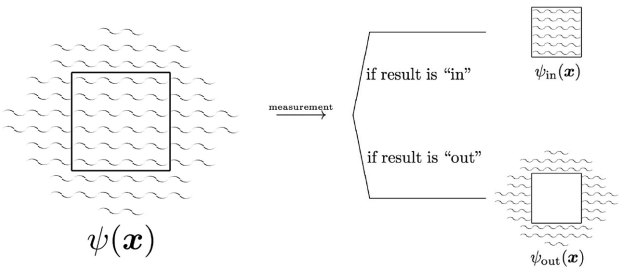
Abbiamo *agito a distanza* sullo stato interno della molecola 2. Secondo i tre di EPR questa azione a distanza, questa non-località, è inconcepibile, non può sussistere. Quindi lo stato del sistema non può

essere costituito dalla sola funzione d'onda, devono esistere delle *variabili addizionali* che *determinano* sia il risultato della misurazione sulla molecola 1, sia ciò che avviene alle due molecole.

La ricerca negli anni successivi (basata sulla disuguaglianza di Bell e sulla sua analisi sperimentale) ha mostrato che è effettivamente possibile sostituire alla meccanica quantistica standard una meccanica quantistica *deterministica* in cui lo stato è costituito dalla funzione d'onda più opportune variabili addizionali. Ma queste variabili addizionali hanno necessariamente comportamenti non-locali. In questo senso si può sostenere che anche Einstein sbagliava.

ENTANGLEMENT – IL *GATTO DI SCHRÖDINGER*

Torniamo alla descrizione della misurazione della posizione di una particella secondo la meccanica quantistica standard.



Si possono fare due considerazioni, tra di loro connesse. La prima è che la misurazione è descritta come una *situazione speciale*, nella quale si verifica una rottura dell'evoluzione ordinaria. La seconda è che l'apparato misuratore *non entra nella descrizione*, esso deve esistere e palesare il risultato, ma è solo lo stato della particella che viene descritto, $\psi(\boldsymbol{x})$ prima, $\psi_{\text{in}}(\boldsymbol{x})$ o $\psi_{\text{out}}(\boldsymbol{x})$ dopo.

Sorge spontanea l'idea di

- includere l'apparato nella descrizione,
- trattare il sistema "particella microscopica più apparato" secondo le regole ordinarie,
- riottenere in questo modo le regole speciali della formulazione standard.

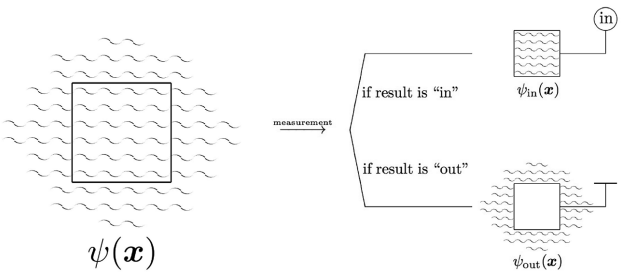
In realtà cosa è l'apparato? È un sistema molto complesso, che

certamente deve contenere un *indicatore macroscopico*, i cui diversi stati mostrino agli umani il risultato della misurazione.

Caratteristiche essenziali di un apparato misuratore sono

- una parte, di regola microscopica, interagisce con il sistema microscopico misurato,
- la modificazione della parte microscopica scatena un processo di amplificazione,
- la modificazione amplificata provoca un mutamento dello stato dell'indicatore.

La seguente figura illustra la situazione:



Se dunque si descrive il sistema particella più apparato secondo la meccanica quantistica, il suo stato è soggetto alla seguente evoluzione:

$$(\psi_{in}(x) + \psi_{out}(x))[a(\ldots)\ldots\psi^{giù}{}_{paletta}(\ldots)] \rightarrow$$
$$\psi_{in}(x)\ [b(\ldots)\ldots\psi^{su}{}_{paletta}(\ldots)] + \psi_{out}(x)\ [a(\ldots)\ldots\psi^{giù}{}_{paletta}(\ldots)]\ ,$$

$\psi^{giù}{}_{paletta}(\ldots)$ e $\psi^{su}{}_{paletta}(\ldots)$ rappresentano lo stato dell'indicatore (paletta su o giù rispettivamente), e a(…) e b(…) rappresentano gli stati di parti intermedie dell'apparato di misurazione.

Dunque il risultato che si ottiene applicando la sola evoluzione quantistica al sistema che includa l'apparato di misurazione *non corrisponde* a quanto prevede la teoria standard e appare *incomprensibile*, essendo la sovrapposizione di stati macroscopicamente diversi dell'indicatore.

Schrödinger rese particolarmente evidente il carattere paradossale dello stato entangled che viene a crearsi in seguito alla misurazione immaginando la seguente situazione. Si supponga che all'interno del volume in cui si accerta la presenza della particella, o meno, sia presente un gatto; si supponga inoltre che quando la particella viene trovata *in,* un dispositivo diffonda nel volume un gas venefico che uccide il gatto; allora l'evoluzione quantistica dello stato del sistema "particella +… + gatto" è data da

$$(\psi_{in}(\boldsymbol{x}) + \psi_{out}(\boldsymbol{x}))[a(\ldots)\ldots\psi^{vivo}{}_{gatto}(\ldots)] \rightarrow$$
$$\psi_{in}(\boldsymbol{x})\ [b(\ldots)\ldots\psi^{morto}{}_{gatto}(\ldots)] + \psi_{out}(\boldsymbol{x})\ [a(\ldots)\ldots\psi^{vivo}{}_{gatto}(\ldots)],$$

cioè uno stato in cui il povero gatto è in una sovrapposizione di *gatto vivo* e *gatto morto*.

La soluzione ortodossa

A questo punto l'interpretazione standard taglia la testa al toro. I "gradi di libertà" *su* o *giù* della paletta o *morto* o *vivo* del gatto per il loro carattere macroscopico sono *classici*. Come tali essi sono necessariamente sempre in uno stato definito.

Pertanto ha luogo la riduzione

$$\psi_{in}(\boldsymbol{x})\ b(\ldots)\ldots\psi^{morto}{}_{gatto}(\ldots) + \psi_{out}(\boldsymbol{x})\ a(\ldots)\ldots\psi^{vivo}{}_{gatto}(\ldots)$$

$$\rightarrow \quad \begin{array}{l} \psi_{in}(\boldsymbol{x})\ b(\ldots)\ldots\psi^{morto}{}_{gatto}(\ldots) \text{ se il risultato è } \textit{in}, \\ \psi_{out}(\boldsymbol{x})\ a(\ldots)\ldots\psi^{vivo}{}_{gatto}(\ldots) \text{ se il risultato è } \textit{out}. \end{array}$$

Conclusioni

L'*entanglement*, cioè l'esistenza di stati quantistici di sistemi composti tali per cui esistano proprietà globali ma non proprietà individuali dei componenti, è una delle caratteristiche più sorprendenti della meccanica quantistica. Lungi dall'essere una situazione particolare, di scarsa rilevanza pratica, al contrario esso è una situazione piuttosto comune. Mentre l'*entanglement* di particelle microscopiche su distanze microscopiche appare come in un certo senso naturale, quando si abbia a che fare o con *entanglement* di particelle microscopiche su grandi distanze o con *entanglement* di sistemi microscopici e macroscopici compaiono aspetti paradossali.

La prima delle due situazioni ha dato origine al cosiddetto paradosso EPR: secondo gli estensori di questa linea di pensiero la meccanica quantistica non sarebbe una teoria completa, nel senso che devono esistere variabili addizionali, dette variabili nascoste, che predeterminano il risultato delle misurazioni. L'intensa attività di ricerca su questo argomento (basata sulla disuguaglianza di Bell e sulle successive analisi sperimentali di Aspect e altri) ha evidenziato che un eventuale comple-

tamento della meccanica quantistica mediante una teoria a variabili nascoste non è possibile se queste ultime sono locali: in questo senso Einstein e collaboratori sbagliavano, una visione del mondo contemporaneamente realista, deterministica e locale non è sostenibile.

La seconda delle due situazioni mette in evidenza come la misurazione abbia un ruolo speciale: se la soluzione al problema sia l'interpretazione ortodossa della meccanica quantistica, con il collasso della funzione d'onda all'atto della misurazione, o altre possibili interpretazioni/estensioni della meccanica quantistica stessa, è argomento di vivace ricerca scientifica contemporanea.

Scienze naturali
Geologia

Istituto Lombardo (Rend. Scienze) 150, 15-30 (2016)

LO STATO ATTUALE DEI GHIACCIAI ITALIANI E LA LORO RECENTE EVOLUZIONE
RISULTATI DI UN NUOVO INVENTARIO

CLAUDIO SMIRAGLIA (*) e GUGLIELMINA ADELE DIOLAIUTI (**), *et al.*[#]

Nota presentata dal m.e. Giuseppe Orombelli
(Adunanza del 4 febbraio 2016)

SUNTO. – I ghiacciai montani rappresentano non solo un'importante risorsa idrica e paesaggistica, ma con la loro recente evoluzione sono divenuti presso l'opinione pubblica la testimonianza più evidente dei cambiamenti climatici in atto. Uno strumento fondamentale per descrivere e gestire l'ambiente glacializzato di alta montagna e per valutarne le trasformazioni è sicuramente la preparazione di un inventario o catasto dei ghiacciai esistenti che ne quantifichi caratteristiche ed evoluzione. Il nostro Paese ha una grande tradizione in questo settore. Fu infatti il primo a dotarsi alla metà del secolo scorso di un inventario dei ghiacciai, prodotto dal Comitato Glaciologico Italiano e del CNR, che evidenziò una superficie glaciale di circa 530 km^2. Un recente progetto, coordinato dall'Università Statale di Milano, con il supporto di enti privati e la collaborazione del Comitato EvK2CNR e del Comitato Glaciologico Italiano, ha portato alla preparazione del Nuovo Catasto dei Ghiacciai Italiani, un atlante nazionale prodotto sulla base dell'analisi di ortofoto a colori ad alta definizione acquisite nel periodo 2005-2011. Nel Nuovo Catasto dei Ghiacciai Italiani vengono descritti 903 ghiacciai estesi complessivamente su un'area di 370 km^2. La maggior parte della superficie glaciale è localizzata in Valle d'Aosta (36.15% del totale), in Lombardia e in Alto Adige. L'84% dei ghiacciai (per numero) è rappresentato da apparati di

(*) già Università Statale di Milano, Milano, Italia.
E-mail: claudio.smiraglia@unimi.it

(**) Università Statale di Milano, Dipartimento di Scienze e Politiche Ambientali, Milano, Italia. E-mail: guglielmina.diolaiuti@unimi.it

(#) La nota è il risultato di un lavoro di ricerca condotto in collaborazione con: Roberto Sergio Azzoni, Carlo D'Agata, Davide Fugazza, Antonella Senese, Davide Maragno.

dimensioni minori di 0.5 km^2 che insieme coprono il 21% della superficie totale. Ghiacciai più estesi di 1 km^2 sono solo il 9.4% del numero complessivo, ma coprono il 67.8% dell'area glaciale nazionale. Un confronto tra i dati riportati nel Nuovo Catasto dei Ghiacciai Italiani e quanto pubblicato nel Catasto CGI-CNR (1959-1962) indica una riduzione della superficie glaciale nazionale del 30%; se si considera il Catasto Glaciale Internazionale (World Glacier Inventory o WGI), pubblicato alla fine degli anni '80 e che descrive 1381 ghiacciai estesi su un'area di 609 km^2, ne risulta una perdita di 478 apparati ed una contrazione areale di 239 km^2 (-39%). Questa riduzione ha portato a importanti e rapide modifiche del paesaggio dell'alta montagna, in particolare alla frammentazione dei ghiacciai, all'incremento delle aree scoperte dal ghiaccio, alla formazione di numerosi laghi proglaciali, allo sviluppo della vegetazione pioniera.

***

ABSTRACT. – Mountain glaciers represent an important hydrological and touristic resource, and their recent evolution provides a dramatic evidence of climate change for the general public. Glacier inventories, quantifying glacier characteristics and evolution, are an important tool to describe and manage high mountain glacier environments and Italy has developed a long tradition in this sector. Our country was the first to provide itself with a glacier inventory, compiled by Comitato Glaciologico Italiano and CNR, showing a glacier surface of 530 km^2. A recent project, coordinated by Università Statale di Milano with the support of private bodies and the cooperation of Comitato EvK2CNR and Comitato Glaciologico Italiano, led to the development of the new Italian Glacier Inventory, a national atlas produced from the analysis of color orthophotos at high resolution acquired between 2005 and 2011. The New Italian Glacier Inventory lists 903 glaciers, covering an area of 370 km^2. The largest part of glacier area is located in Val d'Aosta (36.15% of the total), followed by Lombardia and South Tyrol. 84% of glaciers (considering the number of glaciers) have an area lower than 0.5 km^2 and jointly account for 21% of the total glacier surface. Glaciers larger than 1 Km^2 make up 9.4% of the total number, but cover 67.8% of the total glacier area. The comparison between data from the New Italian Glacier Inventory and the CGI-CNR inventory (1959-1962) shows a 30% reduction in glacier area in Italy; considering instead the World Glacier Inventory or WGI, published at the end of the '80s, which reported 1381 glaciers and an area of 609 km^2, glacier loss sums up to 478 glaciers and an area of 239 km^2 (-39%). This shrinkage has led to rapid and significant changes to high mountain landscapes, notably glacier fragmentation, an increase in deglaciated areas, the formation of proglacial lakes and the development of pioneer vegetation.

LO STUDIO DEI GHIACCIAI ALPINI

I ghiacciai non polari ("locali" o *mountain glacier*) costituiscono una ridottissima parte della criosfera terrestre. La quasi totalità di superficie e volume del ghiaccio terrestre è infatti concentrata nelle gigantesche calotte polari, in particolare in Antartide, che copre oltre

l'89% del volume globale [1, 2]. I ghiacciai situati sulle catene montuose raccolgono meno dell'1% del volume del ghiaccio terrestre con una superficie di circa 0,5 x 10^6 km^2. Sulle Alpi la superficie complessiva è evidentemente molto minore (poco più di 2000 km^2) [3], quindi praticamente insignificante a livello statistico rispetto alla criosfera totale. Il loro significato, e quindi anche il senso del loro studio, trascendono tuttavia le limitate dimensioni. I ghiacciai alpini nella loro evoluzione recente e in corso sono infatti entrati nella percezione comune come la più evidente ed avvertibile testimonianza dei cambiamenti climatici in atto. Le fotografie che mostrano impietosamente la riduzione e talora l'estinzione dei ghiacciai alpini con le conseguenti intense trasformazioni del paesaggio alpino, sono divenute negli ultimi anni il simbolo più diffuso nei mezzi di comunicazione degli effetti del *global change*. A questi aspetti si aggiungono tuttavia altre motivazioni importanti che spingono allo studio della criosfera. Fra queste va sottolineato che i ghiacciai alpini rappresentano anche una importante risorsa idrica ed energetica. Importante non certo a livello quantitativo (a fronte di un afflusso medio annuo per il bacino del Po di circa 77 miliardi di m^3, la riserva idrica totale dei ghiacciai padani è stimabile in circa 13 milioni di m^3 con un afflusso medio annuo di circa 360 milioni di m^3), ma piuttosto a livello strategico. Si tratta infatti di una risorsa modesta, ma essenziale durante le estati più siccitose. A ciò si aggiungono fattori di polarizzazione turistica e paesaggistica, che fanno dei ghiacciai alpini e delle aree contermini siti di elevata frequentazione turistica e sportiva. A questo proposito va ricordato che uno dei più recenti fattori di interesse applicativo di questo tipo di studio è rappresentato dall'incremento del rischio nella frequentazione delle aree di alta montagna glacializzata o recentemente deglaciata, sia a livello strettamente alpinistico sia soprattutto a livello di escursionismo o turismo in senso generale. Nata a fine Settecento nell'Europa illuministica ad opera dei primi naturalisti come De Saussure, la scienza dei ghiacciai (glaciologia) ha seguito in duecento anni la rapidissima evoluzione delle scienze, divenendo oggi una disciplina prevalentemente quantitativa, basata su metodi e tecniche delle scienze di base come la fisica e la chimica. Al contempo tuttavia l'esigenza di nuovi paradigmi che aprano ad una lettura e ad un'analisi integrata dell'ambiente, spingono oggi al recupero nello studio dei ghiacciai di aspetti anche storici, umanistici ed economici (*Figg. 1-2*).



Fig. 1 – Alta montagna e ghiacciai: l'impressione idilliaca e romantica. Il Ghiacciaio dei Forni in alta Valtellina (foto C. Smiraglia, Agosto 2017).



Fig. 2 – Un'impressione più realistica del settore inferiore del ghiacciaio: il "teschio" dei Forni (foto da drone UNIMI-Levissima-Agricola 2000, 2013).

I METODI PER LO STUDIO DEI GHIACCIAI

Sulla spinta di questi elementi di interesse e soprattutto delle rapidissime trasformazioni dell'oggetto di studio, i metodi di analisi del glacialismo hanno subito una vera e propria rivoluzione metodologica e strumentale nell'arco di pochi decenni [4]. Si è infatti dapprima passati dai rilievi sistematici delle variazioni delle fronti glaciali (misure di variazione di lunghezza) alle misure di variazione di spessore (bilanci di massa) che ormai contano serie più che secolari nel primo caso e pluridecennali nel secondo. Una consapevolezza essenziale delle prime commissioni nazionali per il monitoraggio dei ghiacciai (la prima fu fondata in Svizzera

nel 1893) fu infatti l'importanza della continuità temporale delle misure. Queste tipologie di misure proseguono infatti tutt'oggi, seppur con strumenti topografici molto più sofisticati (ai primitivi teodoliti e alle bussole e bindelle si sono infatti sostituiti, o meglio aggiunti, GPS e *laser scanner*) (*Fig. 3*). Soprattutto si è enormemente accentuata la potenza di calcolo, che ha favorito lo sviluppo della modellistica, il tutto accompagnato dallo sviluppo degli strumenti geofisici (dalle prime prospezioni sismiche per la misura degli spessori dei ghiacciai si è passati a quelle con georadar) e anche meteorologici (con stazioni meteorologiche automatiche collocate sulla superfice dei ghiacciai (*Fig. 4*). La maggiore rivoluzione strumentale è forse quella legata allo sviluppo del telerilevamento, basato sulla ripresa, elaborazione e interpretazione di immagini da terra e soprattutto da aereo e da satellite, nonché recentemente anche da droni [5].

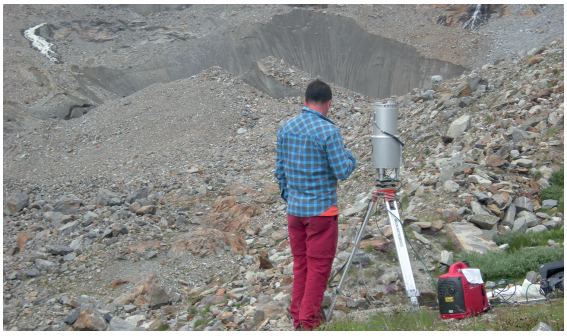


Fig. 3 – Rilievi con laser scanner presso la fronte del Ghiacciaio del Lys (Monte Rosa) (foto C. Smiraglia, 2017).



Fig. 4 – Stazione meteorologica automatica (AWS) collocata sulla superficie del Ghiacciaio dei Forni (foto C. Smiraglia, 2013).

Con queste strumentazioni è possibile ottenere da un lato informazioni quantitative sul glacialismo di aree vastissime (interi gruppi montuosi) e spesso non facilmente accessibili, dall'altro disporre di dati con altissimo dettaglio e altissima accuratezza (*Fig. 5*).

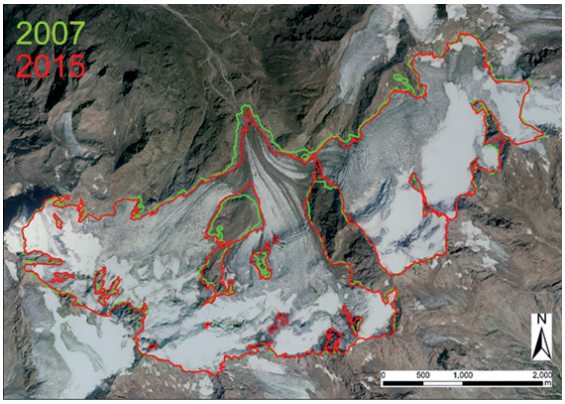


Fig. 5 – Variazioni del Ghiacciaio dei Forni fra il 2007 e il 2015 (elaborazione D. Maragno su base ortofoto 2007 Regione Lombardia: TerraItaly-Compagnia Generale Ripreseaeree S.p.A.).

In Italia gli studi glaciologici iniziarono ufficialmente verso la fine dell'Ottocento con la fondazione nel 1895 della Commissione Glaciologica nell'ambito del Club Alpino Italiano che raccolse i maggiori studiosi del settore e diede il via a rilievi sistematici alle fronti di numerosi ghiacciai del versante meridionale delle Alpi. Nel 1913 la commissione si trasformò nel Comitato Glaciologico Italiano (CGI), ente che tuttora coordina questo tipo di ricerche e promuove la diffusione dei dati raccolti, con la collaborazione di numerose strutture e organizzazioni locali [6].

Uno strumento tradizionale di conoscenza glaciologica, che si è giovato dell'evoluzione scientifica e metodologica sopra accennata, è l'inventario dei ghiacciai, cioè il censimento o catasto dei corpi glaciali di un'area più o meno ampia, che può andare da un singolo gruppo montuoso, ad una regione amministrativa o ad un'intera catena montuosa, fino ad arrivare all'intero pianeta. E' chiaro che gli inventari realizzati in tempi diversi nella medesima regione, solitamente a distanza

di decenni, danno la possibilità di analizzare l'evoluzione del glacialismo di quell'area; si possono quindi valutare le variazioni numeriche e dimensionali dei singoli apparati glaciali e quindi correlare queste variazioni con i parametri meteo-climatici.

L'Italia vanta una lunga ed importante tradizione nel campo degli inventari dei ghiacciai. Sempre nell'ambito del CGI venne infatti realizzato da Carlo Porro nel 1925 uno dei primissimi censimenti a livello mondiale, basato su cartografia e rilievi di terreno; nell'occasione vennero censiti 774 ghiacciai. Una realizzazione veramente all'avanguardia per il suo tempo, sempre ad opera del CGI con il CNR, fu il Catasto dei Ghiacciai Italiani pubblicato fra il 1959 e il 1962, anch'esso realizzato su cartografia e rilievi di terreno. In quell'occasione vennero censite 835 unità glaciali (comprendendo ghiacciai in senso stretto e glacionevati, cioè piccoli apparati privi di movimento) con una superficie totale di poco più di 500 km^2 [6, 7] (*Fig. 6*).

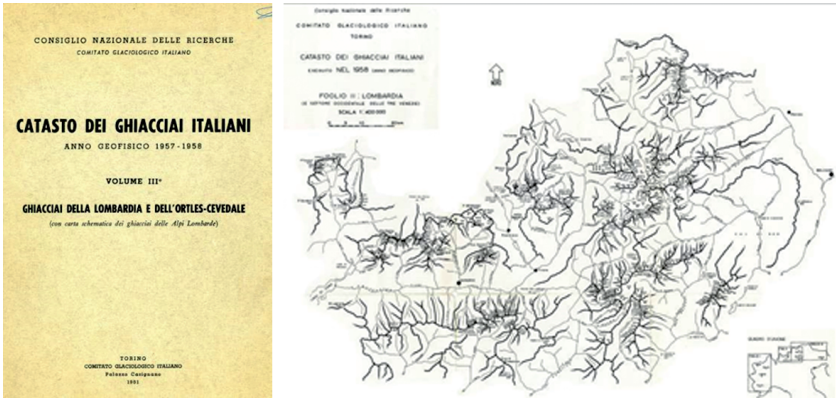


Fig. 6 – Il Catasto dei Ghiacciai Italiani del CGI-CNR. Volume delle Alpi Centrali (a sinistra copertina; a destra carta allegata con la distribuzione dei ghiacciai).

Successivamente con il coordinamento del *World Glacier Monitoring Service* (WGMS) fu realizzato il catasto internazionale dei ghiacciai (*World Glacier Inventory-WGI,* 1985 e successivi aggiornamenti anche in rete), cui per la parte italiana contribuì il CGI. La superficie totale dei ghiacciai italiani risultò di circa 600 km^2 con un notevole incremento numerico degli apparati. Dopo la compilazione

di numerosi inventari regionali e settoriali ad opera di vari enti e strutture locali, negli ultimi anni nell'ambito di diversi progetti sono stati predisposti degli elenchi aggiornati dei ghiacciai italiani. Fra questi il progetto internazionale GLIMS (*Global Land Ice Measurements from Space*) [1], il progetto CNR-CGI *Next Data* [8] e un progetto Università di Milano-Associazione EvK2CNR – Sanpellegrino Spa. Quest'ultimo si può considerare come un vero e proprio aggiornamento del catasto CGI degli anni '50-'60 ed è stato pubblicato nel 2015 (è liberamente consultabile e scaricabile dal sito *https://sites.unimi.it/glaciol/*) [9].

Lo stato attuale dei ghiacciai italiani: una sintesi dal nuovo inventario

Il nuovo catasto è stato realizzato nell'ambito del Gruppo di Glaciologia dell'Università Statale di Milano, si è avvalso della collaborazione scientifica del CGI, ha coinvolto numerose realtà regionali e provinciali dell'arco alpino, nonché i principali enti culturali esperti del settore, ed ha avuto il patrocinio del WGMS. L'obiettivo principale del lavoro, il cui interesse non è solo scientifico, ma anche culturale, divulgativo e applicato, è stato quello di fornire una quadro completo, aggiornato e quantitativo della risorsa paesaggistica rappresentata dai ghiacciai italiani (*Figg. 7-8*).

La metodologia di base per l'individuazione dei singoli ghiacciai e dei loro parametri geometrici è stata l'utilizzo, l'elaborazione e l'interpretazione in ambiente GIS di ortofoto a grande scala (1:5000) riprese fra il 2005 e il 2011, fornite cortesemente da tutte le Regioni alpine. L'accuratezza nominale delle ortofoto (pixel 0.5 m x 0.5 m) ha permesso di delimitare i perimetri dei ghiacciai con elevata precisione (errore areale medio +-2%); va tuttavia tenuto conto che in alcuni casi, in particolare a causa delle frequenti coperture detritiche che stanno sempre più caratterizzando la superficie dei ghiacciai, l'errore areale può essere superiore. I ghiacciai sono stati raggruppati in base alla distribuzione regionale amministrativa (unica eccezione il Trentino-Alto Adige, dove si è tenuto conto della suddivisione provinciale). Per ogni unità glaciale sono stati inseriti quindici parametri in parte descrittivi e qualitativi (ad esempio *Nome*, *Gruppo montuoso*, *Regione amministrativa*, *Bacino idrografico*, *Tipologia*), in parte

numerici e quantitativi (come *Area in km²*, *Area CGI*, *Area WGI*). In un apposito *Annex* sono stati inseriti altri parametri quantitativi, come *Lunghezza massima*, *Altezza minima*, *Altezza massima*, *Altezza media*, *Inclinazione media.* Vengono così resi disponibili i principali parametri numerici indicati dal WGMS, la cui elaborazione ha permesso non solo di evidenziare le caratteristiche attuali del glacialismo italiano, ma anche di individuare le tendenze dell'ultimo mezzo secolo [10, 11].

A livello di sintesi l'elaborazione dei dati del catasto indica che sulle montagne italiane sono situati 903 apparati glaciali (fra ghiacciai in senso stretto e glacionevati) con una superficie complessiva di 369 km², per la quasi totalità distribuiti sulle Alpi. Le Regioni con le maggiori superfici glaciali risultano essere la Valle d'Aosta (134 km²), la Lombardia (88 km²), l'Alto Adige (85 km²) e il Trentino (46 km²), mentre la Regione alpina meno glacializzata è il Friuli (0,2 km².) Sugli Appennini sono presenti solo due piccoli glacionevati, derivanti dalla frammentazione del preesistente Ghiacciaio del Calderone sul Gran Sasso (0,04 km²) [12]. Una delle più interessanti peculiarità del glacialismo italiano è la sua distribuzione su tutti i settori della catena alpina, dalle Alpi Marittime alle Alpi Giulie, caratteristica che lo differenzia da tutti gli altri Stati alpini. A questo si affianca la ridotta dimensione dei ghiacciai; prevalgono infatti gli apparati di piccole dimensioni (<0,5 km²), mentre quelli con area maggiore di 10 km² sono solamente tre: il Ghiacciaio dei Forni in Lombardia, quello dell'Adamello in Lombardia e Trentino, quello del Miage in Valle d'Aosta, che insieme coprono oltre il 10% dell'intero glacialismo italiano (va tuttavia ricordato che negli ultimissimi anni il Ghiacciaio dei Forni si è frammentato in almeno due settori, perdendo quindi la sua struttura unitaria). Va anche sottolineato che l'84% del numero dei ghiacciai italiani si colloca nella fascia areale <0,1-0,5 km², ma nel contempo ricopre solo il 21% della superficie totale. A questa dimensione prevalentemente molto ridotta, che li rende fra l'altro molto sensibili ai mutamenti climatici anche lievi, si aggiunge la tipologia poco differenziata. I ghiacciai "vallivi", caratterizzati da una lingua ben evidente che scende anche a quote relativamente basse, sono solo il 3%, mentre vi è l'assoluta prevalenza dei ghiacciai "montani", annidati sui versanti montuosi alle quote più elevate (57%), e dei "glacionevati", piccoli corpi glaciali caratteristici della fase terminale (o embrionale) del glacialismo (40%) [9].

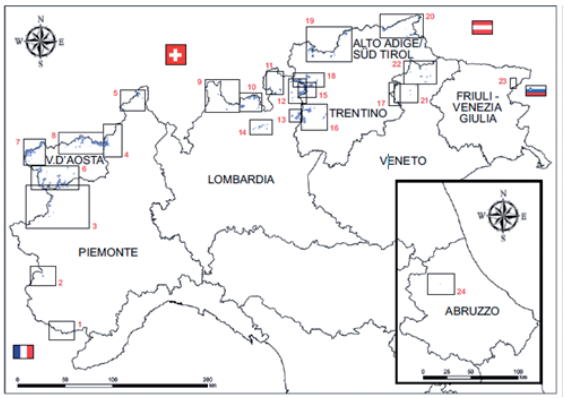


Fig. 7 – La distribuzione dei ghiacciai italiani. I riquadri e i numeri indicano i settori della suddivisione geografica utilizzata nel Nuovo Catasto [9] (elaborazione D. Maragno).



Fig. 8 – Il Nuovo Catasto dei Ghiacciai Italiani (consultabile e scaricabile dal sito https://sites.unimi.it/glaciol/).

Come sono variati i ghiacciai italiani

La disponibilità di inventari multitemporali permette, come si è accennato, il confronto fra situazioni diversificate e distribuite nel tempo. Va tuttavia sottolineato che questo confronto non è agevole a livello quantitativo e spesso offre solo indicazioni di massima. Le metodologie di raccolta dati, in particolare fra il catasto CGI e il nuovo catasto, sono infatti

molto diverse e i dati presentano accuratezze altrettanto diverse. Va ricordato infatti che per il catasto CGI i dati sono stati raccolti con misure di terreno e cartografia spesso non aggiornata, mentre per il nuovo catasto si sono utilizzate ortofoto ad alta definizione e anche immagini da satellite. La tendenza risultante appare tuttavia chiara e indiscutibile: su tutte le montagne italiane i ghiacciai dalla metà del secolo scorso sono in netta riduzione areale (complessivamente l'area è diminuita di 157 km^2 corrispondente a circa il 30% della superficie complessiva iniziale). Le perdite maggiori si sono avute nelle Regioni dove i ghiacciai coprivano superfici limitate, come il Friuli (-50%) e il Veneto (-48%); più contenute sono apparse le riduzioni dove i ghiacciai coprivano superfici più vaste, come la Valle d'Aosta (-26%) e la Lombardia (-24%). Se poi si mettono in relazione le superfici dei ghiacciai del catasto CGI e le rispettive variazioni areali, si osserva chiaramente come gli apparati con una superficie iniziale minore presentino la maggiore variabilità nelle variazioni (dallo 0 al 90%), mentre i ghiacciai maggiori presentano variazioni sempre meno intense (*Fig. 9*).

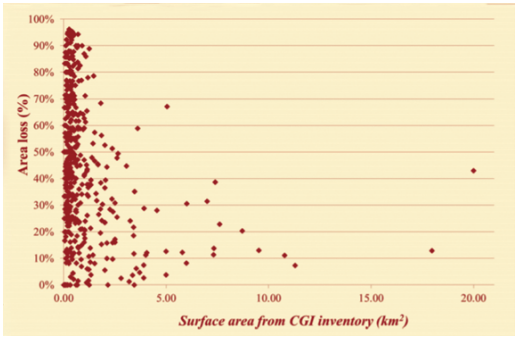


Fig. 9 – Relazione fra superficie dei ghiacciai nel catasto CGI e riduzione areale percentuale rispetto al nuovo catasto (da Smiraglia e Diolaiuti [9]).

Se il confronto viene effettuato con i dati del Catasto Internazionale (*World Glacier Inventory* o WGI), pubblicato alla fine degli anni '80 del XX secolo [13], il regresso appare ancora più intenso. Nel periodo 1965-1985 si ebbe infatti una limitata fase di espansione glaciale che secondo i dati WGI portò ad un lieve aumento nel numero e nella superficie dei ghiacciai italiani (e anche nel resto delle Alpi) rispetto alla metà del XX secolo. Nel WGI i ghiacciai elencati erano 1381 estesi su un'area di 609 km^2; i risultati del nuovo catasto segnalano quindi una perdita di 478 ghiacciai e una riduzione areale di 240 km^2 (-39%) [9]. La tendenza che

emerge dal confronto dei dati è sicuramente molto chiara, ma sono soprattutto le modifiche del paesaggio glaciale ad evidenziare l'intensità dei fenomeni avvenuti a partire dalla metà degli anni '50 del secolo scorso in alta montagna, come risulta ben illustrato anche dai semplici e ormai classici confronti fra foto riprese in periodi diversi (*Fig. 10*).

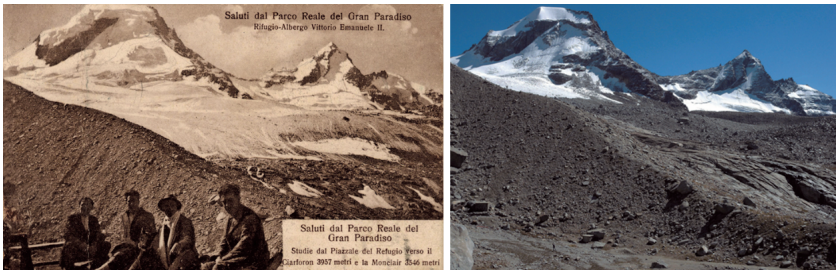


Fig. 10 – Il Ghiacciaio di Moncorvè nel gruppo del Gran Paradiso (Valle d'Aosta) in una cartolina viaggiata nel 1920 (Archivio Società Meteorologica Italiana) a sinistra; lo stesso ghiacciaio nel 2014 (foto Società Meteorologica Italiana) a destra.

Spicca soprattutto la frammentazione, rapida più di quanto prevedibile, dei ghiacciai preesistenti in più corpi minori, con l'apparentemente paradossale risultato di ampliare, almeno per il momento, il numero totale dei ghiacciai (da 835 a 903). Ben conosciuti, fra gli altri, sono i casi di ghiacciai, spesso vallivi, o anche di minori dimensioni, come la Lex Blanche e la Brenva sul Monte Bianco, il Lys sul Monte Rosa, il Ventina e il Fellaria Orientale sul Bernina-Disgrazia, i Forni e il Careser sull'Ortles-Cevedale, il Pisgana Occidentale in Adamello, il Malavalle nelle Breonie, e anche il minuscolo Calderone in Appennino, frammentatisi dall'inizio del Duemila in due o più tronconi (*Figg. 11-13*) [14, 15].

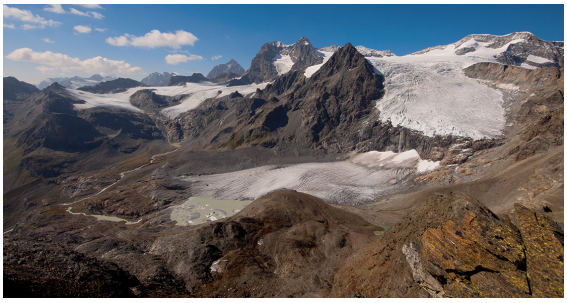


Fig. 11 – Il Ghiacciaio di Fellaria Orientale nel gruppo del Bernina con il settore inferiore ormai nettamente staccato dal corpo principale (foto R. Scotti, 2013).

Questo fenomeno è dovuto alla riduzione dello spessore delle colate glaciali, derivante a sua volta dal bilancio negativo che si è instaurato negli ultimi decenni fra alimentazione nevosa invernale e ablazione estiva (in massima parte fusione). Sulla superficie dei ghiacciai, nei settori più ripidi, si scoprono "finestre" rocciose che nell'arco di pochi anni aumentano la loro estensione fino a separare in più tronconi il ghiacciaio precedentemente unitario. A questo si uniscono la riduzione di lunghezza delle lingue glaciali, il collasso delle zone frontali dove si aprono crepacci, caverne e fratture, la formazione di laghi di contatto glaciale (*Fig. 12*), fenomeno ormai diffusissimo di classica retroazione, che con l'effetto meccanico e termico dell'acqua a contatto con il ghiaccio provoca un'intensa accelerazione del regresso. Altro fenomeno diffusissimo, con effetto contrario almeno nella fase iniziale, è l'incremento della copertura detritica superficiale, derivante dal più intenso crioclastismo sulle pareti rocciose che racchiudono il ghiacciaio; se lo spessore del detrito supera un livello critico (5-10 cm), la fusione del ghiacciaio sottostante viene rallentata.



Fig. 12 – Il lago di contatto glaciale alla fronte del Ghiacciaio Pisgana Occidentale nel gruppo dell'Adamello visto da monte (foto W. Belotti, 2016).

Gli scenari futuri, creati sulla base di modelli climatici e di modelli glaciali, non lasciano molto spazio all'ottimismo sulla sopravvivenza o la ripresa di questi elementi naturali che possono veramente essere considerati il simbolo dell'alta montagna. Le ultimissime calde estati del 2016 e del 2017, dopo la caldissima estate del 2003, hanno inferto un altro duro colpo alle masse glaciali alpine e soprattutto italiane. E' quindi molto pro-

babile che il regresso glaciale proseguirà e accelererà nei prossimi anni e decenni, fino a portare all'estinzione dei ghiacciai piccoli e medi. Per il Ghiacciaio dei Forni, ad esempio, uno dei più grandi ghiacciai italiani, da alcuni anni già frammentato in tre spezzoni, il volume complessivo potrebbe ridursi entro fine secolo al 5-0% di quello attuale [16]. La catena alpina, i cui grandiosi paesaggi sono già molto mutati nei periodi più recenti, si sta dunque avviando ad assumere un aspetto quasi totalmente privo di masse glaciali, molto diverso da quello che conoscevamo fino a mezzo secolo fa (*Fig. 13*).



Fig. 13 – I residui del Ghiacciaio del Calderone sul Gran Sasso in Appennino (foto R. Tonelli, 2005).

Ringraziamenti

Il Nuovo Catasto dei Ghiacciai Italiani è stato ideato e realizzato nel laboratorio Fotogeologia e Remote Sensing dell'Università degli Studi di Milano nell'ambito di un progetto di ricerca sostenuto da Sanpellegrino SPA e di una convenzione di collaborazione scientifica tra UNIMI, Associazione EvK2CNR e Sanpellegrino SPA. Il Comitato Glaciologico Italiano ed esperti delle amministrazioni di numerose regioni alpine italiane hanno partecipato alla fase di validazione dei dati raccolti e hanno fornito supporto scientifico e materiale iconografico (fotografie) pubblicato nel volume.

BIBLIOGRAFIA

[1] J.S. Kargel, A.B.G. Bush, J.G. Cogley, G.J. Leonard, B.H. Raup, C. Smiraglia, M. Pecci and R. Ranzi, *A world of changing glaciers: Summary and climatic context,* in J. S. Kargel et al. (eds.), Global Land Ice Measurements from Space, Springer Praxis Books, DOI: 10.1007/978-3-540-79818-7_33, _ Springer-Verlag, Berlin Heidelberg, 2014.

[2] M. Zemp (editor), *Global Glacier Changes: Facts and Figures,* UNEP-WORLD GLACIER MONITORING SERVICE, 2008.

[3] F. Paul, H. Frey and R. Le Bris, *A new glacier inventory for the European Alps from Landsat TM scenes of 2003: challenges and results,* Annals of Glaciology, **52** 59 (2011).

[4] C. Smiraglia, *Some considerations from recent footprints of the cryospheric sciences to the forthcoming steps,* in E. Wuillermoz, R. Toffolon and A. Cortinovis editors, Proceedings International Conference on Mountains and Climate Change "High Summit", Lecco, CNR, Roma, 2014.

[5] D. Fugazza, A. Senese, R. Azzoni, C. Smiraglia, M. Cernuschi, D. Severi, G. Diolaiuti, *High resolution mapping of glacier surface features. The UAV survey of the Forni Glacier (Stelvio National Park, Italy)*, Geografia Fisica e Dinamica Quaternaria, 38, 25-33 (2015).

[6] www.glaciologia.it

[7] Comitato Glaciologico Italiano – Consiglio Nazionale delle Ricerche, *Catasto dei Ghiacciai Italiani*, Torino, I-II-III-IV (1959-1962).

[8] M.C. Salvatore, T. Zanoner, C. Baroni, A. Carton, F.A. Banchieri, C. Viani, M. Giardino and L. Perotti, *The state of Italian glaciers: a snapshot of the 2006-2007 hydrological period*, Geografia Fisica e Dinamica Quaternaria, 38, 175-198 (2015).

[9] C. Smiraglia e G.A. Diolaiuti (a cura di), *Il Nuovo Catasto dei Ghiacciai Italiani*, Bergamo, Ev-K2-CNR, 2015, p. 400.

[10] C. Smiraglia,. R.S. Azzoni, C. D'Agata, D. Maragno, D. Fugazza and G. Diolaiuti, *The evolution of the Italian glaciers from the previous data base to the New Italian Inventory. Preliminary considerations and results,* Geografia Fisica e Dinamica Quaternaria, 38 (2015), 79-87.

[11] C. Smiraglia, R.S. Azzoni, C. D'Agata, D. Maragno, D. Fugazza and G. Diolaiuti, *The New Italian Glacier Inventory: a didactic tool for a better knowledge of the Alpine natural environment*, Journal of Research and Didactics in Geography, 1, 4 (2015).

[12] M. Pecci, C. D'Agata, C. Smiraglia, *Ghiacciaio del Calderone (Apennines, Italy): the mass balance of a shrinking mediterranean glacier.* Geografia Fisica e Dinamica Quaternaria, 31 (2008), 55-62.

[13] W. Haeberli, H. Bösch, K. Scherler, G. Østrem and C.C. Wallèn (editors), *World Glacier Inventory - Status 1988,* AHS-UNEP-UNESCO, 1989.

[14] C. Smiraglia e G. Diolaiuti, *L'evoluzione del paesaggio glaciale dell'alta montagna alpina,* in G. SCANU (a cura di), *Paesaggi, Ambienti, Culture, Economie*", Bologna, Patron, 2013, 417-426.

[15] L. Mercalli, D. Cat Berro, G. Mortara e C. Smiraglia, *Effetti dei Cambiamenti Climatici sui Ghiacciai*, In S. Castellari e V. Artale (a cura di), *I cambiamenti climatici in Italia: evidenze, vulnerabilità e impatti*, Bologna, Bonomia Press, 2009, pp. 221-239.

[16] R. Garavaglia, A. Marzorati, G. Confortola, D. Bocchiola, G. Cola, E. Manzata, A. Senese, C. Smiraglia e G. Diolaiuti, *Evoluzione del Ghiacciaio dei Forni. La possibile evoluzione del più grande ghiacciaio vallivo italiano attraverso approcci modellistici monodimensionali,* iNeve e Valanghe, 4 (2014), 60-67.

Istituto Lombardo (Rend. Scienze) 150, 31-54 (2016)

GIOVANNI BATTISTA PALLETTA (1748-1832) «DOTTO CHIRURGO E PRATICO ECCELLENTE» A CAPO DELLA CHIRURGIA DELL'*OSPEDALE MAGGIORE*

Nota del m.e. RENZO DIONIGI (*) e ANDREA TERRENI (**)

(Adunanza del 3 marzo 2016)

SUNTO. – Giovanni Battista Palletta, formatosi alla scuola di Bernardino Moscati, di Pietro Moscati e di Giovanni Alessandro Brambilla fu apprezzatissimo primario dell'Ospedale Maggiore di Milano. Molte furono le sue ricerche cliniche che lo resero noto anche a livello internazionale. Non vi è dubbio che la sua vastissima esperienza, la sua abilità, i successi ottenuti, le innovazioni introdotte, il suo tratto di uomo di poche parole giustificano il giudizio della maggior parte degli storici della chirurgia che ritengono la sua reputazione di molto superiore a quella dei suoi stessi maestri.

***

ABSTRACT. – Giovanni Battista Palletta, trained at the school of Bernardino Moscati, Pietro Moscati and Giovanni Alessandro Brambilla, was a highly appreciated surgeon of the Ospedale Maggiore in Milan. Many of his clinical researches made him known in the international surgical community. There is no doubt that his vast experience, his ability, the successes achieved, the innovations introduced, his trait as a man of few words, justify the opinion of most of the historians of surgery who believe his reputation to be much higher than that of his own teachers.

(*) Università degli Studi dell'Insubria, Varese, Italia.
E-mail: renzodionigi@hotmail.com
(**) Università Cattolica del Sacro Cuore, Milano, Italia.
E-mail: terreni11@gmail.com

1. Fino al primo Settecento l'attività chirurgica dell'*Ospedale Maggiore* di Milano era limitata a pochi interventi, da considerarsi impegnativi per l'epoca: qualche rara amputazione, la "legatura" incruenta delle ernie, la litotomia per calcolosi vescicale e l'intervento per la cataratta. Si trattava di una chirurgia empirica, esercitata dai chirurghi "norcini" o "preciani" (dai nomi dei borghi umbri di Norcia e Preci) che, da *celebres castratores* per antica tradizione, erano diventati esperti nell'applicare cinti erniari, a «trar la pietra di vescica» con una incisione in regione perineale e a «deponere la cataratta».

Questa chirurgia empirica, che in altri paesi era esercitata prevalentemente dai barbieri, si protrasse grosso modo sino alla metà del XVIII secolo, quando la medicina milanese – in particolare in seguito all'arrivo da Pisa di Bernardino Moscati (1705-1798),[1] chiamato nel 1735 dal Capitolo dell'Ospedale Maggiore per ricoprire il posto di incisore anatomico, e poi assurto alla direzione della chirurgia – entrò in una fase di risveglio dovuto prevalentemente ad alcune innovazioni tecniche e, soprattutto, alla grande attenzione che le riforme attuate nel periodo teresiano-giuseppino ebbero per l'insegnamento e l'esercizio della medicina e la riorganizzazione dell'assistenza medica.

Rientrato a Milano dopo un soggiorno a Parigi, ove divenne amico di André Levret, considerato ancor oggi la figura più rilevante dell'ostetricia francese del XVIII secolo, Bernardino Moscati nel 1760 organizzò l'insegnamento dell'ostetricia agli alunni della scuola ospedaliera. Successivamente, l'amministrazione dell'ospedale ritenne opportuno aprire anche un reparto ostetrico. In breve egli divenne il maestro di una vera e propria scuola chirurgica, che ebbe fra i suoi allievi, oltre al figlio Pietro, Giovanni Battista Palletta.

All'*Ospedale Maggiore* verso la fine del Settecento si venne a formare quindi una prestigiosa scuola chirurgica che ebbe la non comune opportunità di sfruttare, quale fondamentale palestra anatomica, la celeberrima *brugna*, termine dialettale che indicava il locale per il deposito dei cadaveri e le loro sezioni, la cui abbondanza, elogiata anche da medici stranieri di passaggio a Milano, era tale da rispondere con efficacia a tutte le esigenze poste dall'insegnamento e dalla ricerca.

In quegli stessi anni Giovanni Alessandro Brambilla, assurto a

[1] R. Dionigi, *Bernardino Moscati. Maestro di chirurgia e riformatore della sanità milanese nel Settecento*, Milano, Edra, 2017.

grande notorietà come chirurgo militare durante la Guerra dei Sette Anni e divenuto chirurgo personale del futuro imperatore Giuseppe II, sostenne la superiorità della chirurgia rispetto alla medicina, potendo la prima risolvere alcune situazioni per le quali la seconda era assolutamente impotente. Su sua proposta fu istituita a Vienna la scuola chirurgica militare denominata *Josephs Akademie*, in onore del sovrano Giuseppe II, con il fine di promuovere l'elevazione tecnica, culturale e sociale dei chirurghi. Un'altra delle imprese del Brambilla fu quella di mettere ordine nello strumentario chirurgico, con finalità pratiche.

Di questa strumentazione si servì certamente Giovanni Battista Palletta, formatosi alla scuola di questi maestri. Molte furono le sue ricerche cliniche originali che lo resero noto anche a livello internazionale. Non vi è dubbio che la sua vastissima esperienza chirurgica, la sua abilità, i successi ottenuti, le innovazioni introdotte, il suo tratto di uomo di poche parole giustificano il giudizio della maggior parte degli storici della chirurgia che ritengono la sua reputazione di molto superiore a quella dei suoi stessi maestri. Con Palletta, l'"arte del tagliare", per secoli affidata agli empirici, divenne medicina operativa. Basandosi sulle recenti scoperte della fisiologia, della struttura di tessuti ed organi e delle sperimentazioni animali, con coraggio e abilità egli fu tra i primi chirurghi a tracciare un nuovo percorso, quello della chirurgia moderna.

Ciononostante, di lui e della sua attività professionale e scientifica poco si sa e poco si è scritto, se non in maniera sommaria, e non si trova un'opera che riassuma compiutamente la sua attività di clinico, docente e ricercatore. Esistono solo alcuni profili biografici o elogi scritti da allievi, parenti o amici, basati su notizie e giudizi raccolti da persone che lo ammirarono in vita; persone quindi non in grado di valutare con la necessaria obiettività e di mettere in evidenza il ruolo rilevante che egli svolse nel periodo di transizione dalla chirurgia empirica alla chirurgia moderna.

2.

Giovanni Battista Palletta[2] nacque il 18 aprile 1748 nel borgo ossolano di Montecrestese, nella frazione di Cardone, da Giacomo e Domenica Leonardi. I genitori, socialmente di condizioni modeste, erano

[2] Il cognome Palletta, già durante la sua vita milanese, certamente dopo la morte e quasi sempre nelle citazioni dei testi in inglese, cominciò a essere scritto con una *l* sola: Paletta.

dediti ai lavori agricoli di alcuni loro fondi da cui traevano di che vivere. Sebbene dovessero affrontare gravi sacrifici, essi, evidentemente già consapevoli delle particolari doti del loro primogenito, decisero per i suoi primi studi di mandarlo, all'età di otto anni, nel 1756, a Briga nel Vallese, presso il collegio dei Gesuiti *Spiritus Sanctus*, luogo di studi, prevalentemente in lettere e filosofia. Esemplare fu il comportamento di Giovanni Battista in collegio, tanto che per le sue ottime qualità scolastiche e morali gli vennero attribuiti diversi premi durante il suo *cursus studiorum*.

Terminati gli studi nel Collegio di Briga, il giovane Palletta decise di recarsi a Milano, dedicandosi dapprima allo studio della giurisprudenza. Tuttavia, in breve tempo si rese conto di essere più portato verso studi meno speculativi, più operativi e, se possibile, più coinvolgenti dal punto di vista sociale, essendo grande il suo desiderio di rendersi utile al prossimo. Il 24 ottobre 1763, all'età di 15 anni, entrò così come studente di chirurgia nel *Collegio-Convitto degli Allievi Chirurghi*, nella *Magna Domus Hospitalis*, la *Ca' Granda*, ovvero l'*Ospedale Maggiore.* L'11 settembre 1769 fu nominato "Allievo fisso". Agli studenti del *Collegio-Convitto* veniva concesso vitto e alloggio gratuito, ma si dava loro l'incombenza di dispensare il vitto e le medicine agli ammalati, di fare le medicazioni, in definitiva di comportarsi come dei *frati ospitalieri*.

Alla conclusione del percorso formativo presso l'*Ospedale Maggiore* il Palletta si trasferì, nel 1770, a Padova, per iscriversi al corso di laurea in Filosofia e Medicina. È probabile che il suo interesse non solo per la chirurgia, ma anche per la filosofia fosse alla base della sua scelta a favore dell'Ateneo patavino. Questa sua curiosità ed attenzione sul senso dell'essere e dell'esistenza umana è peraltro comprovata da uno scritto inedito del 1766, *Philosophia rationalis et experimentalis*, di cui si conserva il manoscritto presso il Fondo Belloni alla biblioteca dell'*Istituto Lombardo.*[3]

Come risulta dall'Archivio storico dell'Università di Padova, il Palletta si iscrisse al *Collegio veneto artista* per il corso triennale 1770-

[3] [Giovanni Battista] Palletta, *Philosophia ratio[na]lis et experimentalis. Pars [pr]ima – Logica, Pars sec[unda] – Physica Generalis*, ms 21x14,5 cm (1766), 146 pp. non numerate, di cui dalla 92 alla 94 e dalla 138 alla 146 bianche. Il manoscritto, in lingua latina, è scritto a penna su carta. Copertina in cartone con il titolo manoscritto a penna. Il manoscritto contiene due foglietti, 11x16,5 cm il primo, 17x6,8 cm il secondo, manoscritti a penna su carta, in lingua latina, ciascuno da entrambe le parti, inseriti rispettivamente alle pagine 82-83 e 96-97). La data 1766 è indicata sulla copertina e a p. 92.

1773,[4] conseguendo il *Publicum Doctoratus Privilegium* il 18 giugno 1773.[5]

Alcuni anni dopo, il 17 marzo 1778 Giovanni Battista Palletta si presentò a Pavia a sostenere gli esami per ottenere il *Magisterium in Chirurgia major*.

Da quanto si evince dal verbale dell'esame,[6] il Palletta non dovette incontrare particolari difficoltà nel presentare la propria casistica operatoria e nel rispondere ai quesiti posti dai membri della commissione: è assai verosimile che questo giovane chirurgo originario di Domodossola avesse già maturato un'esperienza superiore a quella dei suoi stessi esaminatori. Il *Magisterium* conseguito fu per il Palletta una semplice formalità, tant'è che a soli trentadue anni, nel 1780, era già considerato un maestro della chirurgia.

3. Nell'aprile del 1804 gli amministratori dell'*Ospedale Civile di Milano* invitarono il «Capo Chirurgo e Professore Cittadino Paletta»[7] ad esprimere il suo parere sul nuovo *Piano disciplinare per l'Istruzione e promozione de' Chirurghi nell'Ospitale Maggiore di Milano*. Egli diede un notevole contributo alla stesura di questo regolamento, che rappresenta un documento di particolare interesse e assolutamente innovativo per lo studio della formazione professionale dei giovani chirurghi dell'epoca, dedicando particolare attenzione e cura anche ai minimi dettagli.

A questo proposito, è bene sottolineare che al principio dell'Ottocento, ovunque in Europa e naturalmente anche in Italia, la formazione chirurgica, e soprattutto la formazione nei suoi fondamentali aspetti pratici, era prerogativa quasi esclusiva dei grandi ospedali, le uniche strutture che potevano garantire una patologia vasta e variata, nonché la possibilità dell'esercitazione sui cadaveri, che, come appunto si può desumere dal "Piano disciplinare", era condizione imprescindibile nella preparazione e successiva carriera del chirurgo.

---

4 ARCHIVIO STORICO DELL'UNIVERSITÀ DI PADOVA (d'ora in poi ASUPD), 270, *Collegio veneto artista. Documenti per dottorati*, 1770-1773, f. 672.

5 Domodossola, Civico Museo di Palazzo Silva e Archivio Storico della Fondazione Gian Giacomo Galletti. Esposto in bacheca.

6 ARCHIVIO DI STATO DI PAVIA, *Università*, *Medicina*, cart. 357.

7 ARCHIVIO DELL'OSPEDALE MAGGIORE, *Piano disciplinare per l'istruzione e promozione dei Chirurghi dell'Ospedale*, n. 363 del 1804. È questo il primo documento ufficiale sin qui osservato in cui il cognome è scritto con una sola *l*: Paletta invece di Palletta.

L'*Ospedale Maggiore* rappresentava pertanto una delle strutture per quei tempi assolutamente all'avanguardia, tant'è che lo stesso Antonio Scarpa, insigne anatomista dell'Ateneo pavese, spesso si rivolgeva con fiducia al Palletta per sottoporgli i suoi progetti e le sue osservazioni e per chiedergli alcuni pezzi patologici necessari per il completamento di alcuni suoi studi.

Al principio del XIX secolo si può ragionevolmente ritenere che, almeno per quanto riguarda l'insegnamento della chirurgia, l'*Ospedale Maggiore* di Milano fosse ormai da anni una struttura efficiente, efficace e assai collaudata per l'insegnamento.

4. Negli scritti di storia della chirurgia, Antonio Scarpa e Giovanni Battista Palletta sono frequentemente citati insieme, non tanto per raffrontare i loro meriti e le loro glorie, quanto per sottolineare come – pur in ruoli diversi (lo Scarpa essendo più conosciuto come "l'anatomico pavese", il Palletta più come "il chirurgo milanese") ed appartenendo a istituzioni differenti (l'Università e l'Ospedale) – abbiano nello stesso periodo contribuito a risollevare il prestigio dell'arte chirurgica, da troppo tempo in declino, riuscendo con i loro studi e la loro attività clinica ad essere considerati "uomini sommi" della comunità chirurgica internazionale.[8]

Scarpa e Palletta erano pressoché coetanei, essendo il primo nato nel 1752 e il secondo nel 1748. Morirono anche nel medesimo anno: l'anatomico pavese il 30 ottobre 1832, preceduto di soli pochi mesi dal chirurgo milanese. Lo Scarpa onorò la Scuola pavese e la scienza; il Palletta, «dotto chirurgo e pratico eccellente» a capo della chirurgia dell'*Ospedale Maggiore*, è tuttora considerato il fondatore della scuola chirurgica milanese. I due eminenti medici furono legati da sinceri vincoli di reciproca amicizia, cordialità e stima, come si evince dalla lettura delle missive conservate del carteggio tra loro intercorso.[9] Infatti, traspare costantemente un intenso scambio e una significativa condivisione di conoscenze, di

[8] A questi è inoltre corretto unirvi anche Giovanni Battista Monteggia, tra i migliori discepoli del Palletta ai suoi corsi di anatomia e chirurgia, di cui subito si intravvidero le grandi qualità, il quale tuttavia nel 1815 premorì al maestro, ancor giovane e nel pieno della attività.

[9] Nelle numerose – ben 659 – e interessantissime lettere pubblicate nel suo epistolario (cfr: GUIDO SALA, *Antonio Scarpa. Epistolario (1772-1832)*, Pavia, Tip. Già

intuizioni, di osservazioni e riflessioni scientifiche. Nel corso degli anni Palletta ebbe occasione di inviare numerosi pezzi anatomici a Scarpa, dall'Ospedale Maggiore di Milano a Pavia: si trattava di materiale assolutamente fondamentale e prezioso per la conduzione degli studi di anatomia, che solo un grande ospedale come la *Ca' Granda* poteva garantire con la sua ampia offerta di casistiche patologiche.[10]

L'*Ospedale* di Milano, per la vastità, ricchezza e sovrabbondanza di mezzi, avrebbe potuto essere la struttura ideale per l'attivazione di una facoltà di Medicina, in grado di garantire una preparazione soprattutto "pratica", come a quei tempi avveniva presso le facoltà mediche di Parigi, Londra e Berlino.

5. Se è vero che Scarpa e Palletta, dopo un lungo periodo di sopore dell'accademia medico-chirurgica nazionale, erano riusciti, verso la fine del Settecento, a riportare la chirurgia italiana ad una notorietà su scala internazionale, va peraltro detto che tale rinascita fu per vari aspetti anticipata di almeno vent'anni dall'opera di un altro grande chirurgo italiano, Giovanni Alessandro Brambilla.[11] Nato nel 1728 a San Zenone Po, nel pavese, da famiglia appartenente alla classe media,

---

Cooperativa di Bortolo Bianchi in Pavia, 1938), lo Scarpa tratta problemi chirurgici, riferisce casi particolarmente interessanti, dà consigli a colleghi ed espone sue vedute personali. Le undici lettere di Scarpa a Palletta, raccolte nell'Epistolario e seguite da una lettera del Palletta allo Scarpa, erano già state trascritte e pubblicate da Umberto Calamida, ordinario di Clinica otorinolaringoiatrica a Milano, e cultore di storia della medicina, che si interessò appunto della pubblicazione di alcune lettere inedite dello Scarpa (UMBERTO CALAMIDA, *Antonio Scarpa. Lettere a G. Palletta e ad altri con un saggio bibliografico del suo epistolario. Per le nozze del Dott. Prof. Leopoldo Rossi con la Sig.na Maria Letizia dei Conti Sottocasa*, Milano, Tip. Conti, 1933).

[10] In maniera eloquente, Scarpa ebbe occasione di definire il Maggiore di Milano «il vaso più grande che abbia Pandora in Italia».

[11] Su Giovanni Alessandro Brambilla, cfr. CRISTIANO ANTONIO RIGONI, *Elogio del Cav. Gio. Alessandro Brambilla*, Pavia, Bizzoni, 1830; ARTURO CASARINI, *Giovanni Alessandro Brambilla (1728-1800)*, in *Profili di chirurghi militari. Dalle campagne napoleoniche alla grande guerra mondiale,* volume pubblicato in occasione dell'VIII Congresso Internazionale di Storia della Medicina di Roma, Roma, Istituto Poligrafico dello Stato, 1930, pp. 1-5; UGO BALDINI, *Brambilla Giovanni Alessandro*, in *Dizionario Biografico degli Italiani*, vol. 13, Roma, Istituto della Enciclopedia italiana, 1971, pp. 733-734; L. BONANDRINI, *Giovanni Alessandro Brambilla*, in «*…parlano un suon, che attenta Europa ascolta*», pp. 225-232.

Brambilla frequentò l'Università di Pavia effettuando pratica ospedaliera presso l'*Ospedale San Matteo*, mostrando sin dall'inizio uno spiccato interesse per la chirurgia.

Terminati gli studi, entrò come chirurgo nell'esercito asburgico. Nel 1757, postosi in evidenza con una dissertazione tenuta all'Università di Vienna, fu promosso chirurgo maggiore di guerra. Due felici operazioni, una sulla principessa Poniatowski, cognata del re di Polonia, e l'altra sul generale Botta Adorno, a cui riuscì a salvare il braccio sinistro dall'amputazione, posero in risalto le sue qualità chirurgiche presso la casa imperiale. Nel 1763 fu nominato chirurgo della guardia imperiale e, l'anno successivo, chirurgo dell'arciduca Pietro Leopoldo. Poco dopo, nel 1765, il fratello di Leopoldo, Giuseppe, associato in quel medesimo anno al trono sui domini familiari di casa d'Asburgo con la madre Maria Teresa, ottenne che Brambilla rimanesse al suo servizio quale chirurgo personale. Nel 1778 Brambilla fu nominato "Oberstabchirurg" o protochirurgo di Stato Maggiore e, l'anno successivo, sovraintendente unico del Servizio Sanitario Militare.

Soprattutto all'opera e all'intervento di Giovanni Alessandro Brambilla vanno attribuite le origini e l'ispirazione delle riforme giuseppine riguardanti gli ospedali, le carceri, gli asili d'infanzia e per vecchi, le case di maternità, gli istituti di ricerca clinica e dell'apparato sanitario dell'esercito. Tramite l'imperatore, egli intervenne anche nella riforma dell'Università di Pavia, Nel 1780, alla morte dell'imperatrice Maria Teresa, quando Giuseppe successe alla madre negli stati ereditari, il Brambilla, che era stato nominato capo chirurgo dello Stato Maggiore dell'esercito imperiale, inviò nelle scuole chirurgiche più avanzate i migliori fra i giovani medici, affinché, al ritorno, potessero trasmettere agli altri le loro esperienze. Questa iniziativa si concretizzerà poi nella fondazione della scuola medico-chirurgica militare (1781), con un Piano per l'istruzione compilato dal Brambilla stesso, insieme a regolamenti per l'annesso ospedale, che verranno pubblicati a Vienna nel 1784.

Nel 1785, l'orazione latina da lui pronunciata in occasione dell'inaugurazione della *Josephs Akademie*, scuola e accademia chirurgica ad alto livello scientifico, destò in tutta Europa vive discussioni fra i clinici, poiché vi si sostenne la superiorità della chirurgia rispetto alla medicina, potendo la prima risolvere casi di fronte ai quali la seconda era impotente. Per suggerimento del Brambilla, l'accademia venne presto equiparata alle facoltà universitarie di chirurgia e dotata di vasti laboratori e collezioni.

Intensi furono i contatti epistolari e le relazioni intercorse tra Brambilla, Scarpa e Palletta. In particolare, Brambilla aveva fatto la conoscenza di Scarpa durante una visita alla *Académie Royale de Chirurgie* e, sempre ricordandosi dell'ateneo in cui aveva mosso i suoi primi passi da chirurgo, non solo aveva favorito il trasferimento dello Scarpa da Modena a Pavia, ma si operò pure per assegnare in dono all'ateneo pavese un cospicuo numero di volumi della sua biblioteca. Nacque così, tra la Scuola chirurgica militare del Brambilla e l'ateneo pavese, rappresentato dall'anatomico, una collaborazione molto intensa.

Altrettanto significativi furono i rapporti intercorsi tra Brambilla e Palletta, il quale, nella sua biblioteca, conservava le opere più rilevanti del chirurgo castrense di Vienna, due delle quali con dedica dell'autore. Stando alla corrispondenza rinvenuta[12] non si riesce ad evincere se i due si siano mai incontrati – probabilmente no – anche se certa è la grande considerazione che il chirurgo della *Ca' Granda* aveva per il protochirurgo viennese. Lo ammirava a tal punto che gli chiese la disponibilità ad accettare la dedica di una delle sue opere maggiori, gli *Adversaria chirurgica*.

Come già aveva fatto con Scarpa, anche a Brambilla il Palletta inviò pezzi anatomici provenienti dal "vaso di Pandora" milanese.

Brambilla, resosi conto del valore e delle notevoli capacità di Palletta, gli espresse più volte l'intenzione di nominarlo socio corrispondente dell'*Accademia Giuseppina* in modo da poterlo successivamente proporlo quale membro effettivo dell'illustre consesso. Nel 1790, tuttavia, in seguito alla prematura morte dell'imperatore Giuseppe II, del quale non solo era stato chirurgo personale, ma soprattutto amico e consigliere, il ruolo del protochirurgo viennese nel contesto della chirurgia asburgica declinò rapidamente. Gli venne a mancare un importante sostegno e l'ambiente accademico lo abbandonò a tal punto che diede le dimissioni dall'accademia da lui stesso voluta e diretta. A Palletta rimase comunque la soddisfazione di essere nominato, su proposta di Brambilla, membro della *Cesarea Accademia Giuseppina* di Vienna.

6. Nell'elenco dei membri dell'*Istituto Lombardo di Scienze e Lettere*, Giovanni Battista Palletta compare come effettivo a partire dal

[12] Archivio dei Luoghi Pii Elemosinieri di Milano (d'ora in poi ALPEMi), *Autografi*, b. 10, fasc. "Brambilla".

1803, ovvero sin dal primo anno di attività dell'*Istituto Nazionale*, voluto da Napoleone Bonaparte allo scopo di riorganizzare anche in ambito italiano il mondo degli "eruditi" e dei "saggi" che da tempo era in attesa dell'avvio di un complessivo processo di riforma delle istituzioni della cultura nei suoi vari gradi, non solo in ambito scolastico e universitario, dunque, ma anche nella cosiddetta "alta cultura".

A tale scopo, il 6 novembre 1802 Napoleone procedette alla nomina dei primi trenta membri dell'*Istituto Nazionale*, i cui nominativi comparivano così elencati nel Decreto:

> «Scarpa, Oriani, Volta (professore a Pavia), Cagnoli, Fontana, Pino Ermenegildo, Moscati, Appiani (Pittore), Isimbardi, Dandolo, Saladini, Cassiani, Mondini (Anatomico a Bologna), Savioli, Canterzani, Monti Vincenzo, Brugnatelli, Longhi (Legislatore e bibliotecario), Spannocchi, Villa, Paradisi, Morcelli, Bettinelli, Rosa (già Professore a Modena), Fantoni Pio, Padre Soave, Bonati (di Ferrara), Bianconi Carlo (già segretario), Monga, Mari (di Mantova)».

Nelle settimane seguenti Napoleone e i suoi consiglieri fecero le loro ulteriori scelte nella lista dei sessanta nomi che erano stati presentati, e in data 18 aprile 1803 fu pubblicato a Milano il Decreto del 6 aprile che nominava i trentuno membri – vi è anche la nomina di Napoleone stesso – necessari per arrivare alla pienezza del corpo accademico dell'*Istituto*.

Dei nuovi prescelti i fisici e matematici erano ben undici: Giovanni Aldini, Michele Araldi, Giuseppe Avanzini, Vincenzo Brunacci, Paolo Delanges, Mariano Fontana, Giambattista Guglielmini, Paolo Pozzo, Paolo Ruffini, Simone Stratico e Giambattista Venturi, ai quali si affiancavano gli astronomi Angelo Giovanni Cesaris, Giuseppe Piazzi e Francesco Reggio, i naturalisti Carlo Amoretti e Luigi Castiglioni e il geologo Alberto Fortis. La medicina era rappresentata da Giuseppe Atti, Luigi Laghi, Giovanni Battista Palletta, Giuseppe Testa e Gaetano Uttini. Ai letterati, storici e giuristi risultarono attribuiti sette posti, occupati da Luigi Bossi, Benedetto Del Bene, Alberto De Simoni, Angelo Fumagalli, Luigi Lamberti, Gian Carlo Passeroni e Francesco Venini.

Nell'occasione della nomina, «il Ministro degli Affari Interni» inviò «al cittadino Paletta Membro dell'Istituto Nazionale» la seguente comunicazione:

«Il Presidente della Repubblica Italiana veduta la proposizione fatta in conformità della legge 21 agosto, ed il successivo decreto 22 ottobre 1802 dei Membri dell'Istituto Nazionale riuniti in Bologna per la scelta degli individui che ancora mancano a completarne lo stabilito numero di 60, si è compiaciuto di nominarvi Membro dell'Istituto Nazionale con suo decreto 7 aprile corrente, di cui si unisce un esemplare per la conveniente vostra intelligenza e direzione.

Incaricato di parteciparvi tale nomina io non posso che compiacermi del tributo d'onore che vi ha reso il primo Magistrato della Repubblica, e di quello che dal vostro nome ne ridonderà al primo fra gli stabilimenti della Istruzione Nazionale. Questo grande oggetto è troppo caro alle mie cure per non sentire colla massima soddisfazione l'incalcolabile vantaggio che ne verrà a questo Ministero dell'associazione d'uomini distinti per rare cognizioni e consecrati per istituto ad accrescere il deposito dell'umano sapere ed a propagarlo nelle più utili forme».

Passarono solo pochi giorni e il 7 maggio il medesimo ministro degli Interni si rivolse a Giovanni Battista Palletta in virtù della sua specifica competenza socio-sanitaria, invitandolo ad elaborare un progetto di *Regolamento generale nell'ambito di Sanità e di Polizia Medica*.

Sin dall'inizio della sua partecipazione alla vita dell'*Istituto*, Giovanni Battista Palletta fu molto attivo, non solo attraverso la presentazione di memorie di natura prevalentemente medico-chirurgica, ma anche nei lavori di vari gruppi di studio o commissioni che di volta in volta vennero costituiti con scopi differenti.

Nel 1804, quando alcuni membri proposero di riunire i materiali necessari alla formazione del primo volume degli *Atti* dell'*Istituto* – mentre nel medesimo tempo altri furono dell'avviso che sarebbe stato opportuno procrastinare la pubblicazione, in attesa di un maggior numero di "memorie" – Palletta si rese conto che la pubblicazione degli *Atti* avrebbe di fatto comportato molte difficoltà, tra cui certamente un laborioso intervento di revisione redazionale. A questo proposito, indirizzò questa breve lettera al segretario Araldi:

«Cittadino Collega.

Secondando l'invito fattomi con vostra circolare 29 maggio, ed in virtù del Regolamento Organico, ho il piacere di trasmettervi il mio penso, tremando nello stesso tempo pel giudizio che debb'essere proferito da un Consesso più rispettabile, qual è quello dell'Instituto. Non ignoro che abbisogna di molte e molte correzioni in materia di lingua

> specialmente, il che sarà per voi un lavoro di di più, che volentieri avrei evitato, se possedessi la bella letteratura a quel grado che voi l'avete.[13]

Il Palletta, sempre il più assiduo fra i medici e i chirurghi, rimase comunque del parere che gli atti dovessero essere pubblicati e ne sollecitò la rapida pubblicazione in una lettera all'Araldi del 30 novembre 1804, ove fece presente le necessità del «decoro» dell'*Istituto* e la convenienza degli autori. Sempre su questo argomento, suggerì in un'altra lettera i tempi di consegna delle memorie:

> «Illustre Collega.
> Per le solide ragioni da voi addotte nella circolare approvo interamente la scelta dei quattro quesiti, che è stata fatta nell'Adunanza Ordinaria, sembrandomi che per la soluzione dei due primi debbansi accordare due anni, ed un anno per quella degli ultimi».[14]

Grande risonanza ebbe, tra le altre cose, la relazione che Pietro Moscati e Giovanni Battista Palletta lessero il 12 giugno 1817 sui milanesi "Annali di medicina", di cui Annibale Omodei aveva intrapreso la pubblicazione dal 1814: una rivista che era stata accolta molto favorevolmente dalla commissione appositamente istituita, perché indicava alla comunità medica una serie di informazioni e aggiornamenti sui progressi della medicina nei paesi stranieri. Era quella anche un'occasione per auspicare che le autorità governative sostenessero, anche finanziariamente, le attività sanitarie e in particolare quelle dei medici condotti, favorendone l'aggiornamento scientifico-professionale e la formazione permanente:

> «Sarebbe desiderabile che quest'opera periodica […] fosse fra le mani de' nostri medici, e principalmente di quelli di campagna, che per la generale tenuità de' loro stipendi non sono in istato di tenersi al giorno delle mediche scoperte e pratiche osservazioni che, nell'esercizio dell'arte, loro riescirebbero utilissime: né essi per l'anzidetta perentoria ragione, né le Comuni potrebbero caricarsi di questo annuo dispendio. Quindi sarebbe della generosità del R.I. Governo o il degnarsi d'incoraggiare il benemerito autore con una giustificazione durabile almeno per un triennio, finché un più abbondante smercio lo abiliti alla continuazione; ovvero il rilevarne un numero di copie da farsi distribuire gratuita-

---

[13] Lettera di Palletta ad Araldi, recante la data Milano, 19 giugno 1804, conservata tra le carte del *Fondo Luigi Belloni* presso la biblioteca dell'*Istituto Lombardo di Scienze e lettere.*

[14] *Ibid.,* Milano, 21 aprile 1805, Lettera di Palletta ad Araldi.

> mente ai medici condotti, la qual cosa non è nuova ne' fasti del benefico austriaco Governo; e l'immortale augusta Maria Teresa ordinò simil cosa per opere reputate di molta utilità specialmente alla campagna dell'austriaca Lombardia».[15]

7. Verso la fine del Settecento, dopo due secoli di letargo, la chirurgia italiana fu riportata alla ribalta internazionale proprio con la produzione scientifica, in particolare, di Antonio Scarpa e di Giovanni Battista Palletta. Quest'ultimo fu certamente lo studioso più prolifico nei primissimi anni del XIX secolo, proprio nel medesimo tempo in cui furono creati l'*Istituto Italiano* e successivamente l'*Istituto Lombardo di Scienze e Lettere*. Egli risultò costantemente inesauribile ed insaziabile nel riferire di argomenti anche lontani dalla chirurgia,[16] quali la filosofia,[17] l'agronomia, la fisiologia, la pediatria. Gli editori veneti delle sue *Dissertazioni di Chirurgia* (1795) scrissero in fronte alla raccolta:

> Gli argomenti intorno ai quali si aggirano sono dei più importanti della patologia e della pratica chirurgica; la maniera con cui vi sono trattati corrisponde alla loro importanza, ed il chiarissimo autore si manifesta da per tutto quale è in fatti il più dotto ed esperto chirurgo che abbia ora l'Italia.[18]

---

15 F. DELLA PERUTA, *Cultura e organizzazione del sapere nella Lombardia dell'Ottocento. L'Istituto Lombardo di Scienze e Lettere dalla fondazione all'Unità d'Italia, in L'Istituto Lombardo Accademia di Scienze e Lettere (secoli XIX-XX)*, vol. I, Storia istituzionale, a cura di A. Robbiati Bianchi, Milano, 2007, p. 298.

16 Per un'esauriente rassegna critica delle più rilevanti memorie e relazioni presentate dal Palletta all'Istituto, cfr. PAOLO MAZZARELLO, *La medicina e le discipline affini nelle pubblicazioni dell'Istituto Lombardo*, in *L'Istituto Lombardo Accademia di Scienze e Lettere*, vol. II, *Storia della Classe di Scienze Matematiche e Naturali*, Milano, Libri Scheiwiller, 2008, pp. 565-670.

17 Della filosofia professata da Palletta dà una chiara descrizione lo storico Francesco Cusani: «Il profondo studio dell'anatomia lo aveva persuaso che l'armonica struttura del corpo umano appalesa, come disse Galeno, la sapienza d'una divinità. Le angosce e i disinganni dei morienti l'avevano persuaso dell'immortalità dell'anima e d'una vita futura, slontanandolo dall'abbiezione dei materialisti, i quali avendo invano cercata l'anima col coltello anatomico arguiscono che cessa d'esistere dopo che la morte l'ha separata dal corpo» (PASETTI, *Giovanni Battista Palletta*, p. 746). A questo proposito si cita il manoscritto conservato tra le carte del *Fondo Luigi Belloni* presso la biblioteca dell'*Istituto Lombardo di Scienze e lettere*: [G.B.] PALLETTA, *Philosophia ratio[na]lis et experimentalis. Pars [pr]ima – Logica; Pars sec[unda] – Physica Generalis* (1766).

18 Per una disamina ragionata delle opere più significative di Palletta ci si

Nelle sue opere, Palletta fa costantemente intendere della vastissima cultura e dell'approccio razionale che egli adotta nell'affrontare i più svariati argomenti, integrando le profonde cognizioni anatomiche con numerose osservazioni cliniche, che gli derivano da una grande esperienza nella pratica d'ospedale. Dalle sue parole si può dedurre l'accuratezza, la concisione, il rigore, la serietà, la modestia e l'onestà intellettuale di un grande studioso che in svariate occasioni ebbe occasione di dedicarsi a ricerche che oggi definiremmo traslazionali, ovvero con finalità applicative sul paziente, volte a migliorare concretamente la pratica clinica corrente.

In cinquant'anni di attività clinica e scientifica, il Palletta diede alle stampe più di trenta scritti di anatomia, di chirurgia e di ostetricia, molti dei quali presentati come memorie all'*Istituto nazionale italiano* e al *Lombardo*. Nei lunghi anni di attività all'*Ospedale Maggiore*, inoltre, aveva raccolto i manoscritti di una quantità immensa di osservazioni cliniche e di descrizioni di dissezioni anatomiche. Al termine della sua carriera, con l'intento di raccogliere in un'unica opera gli appunti e i risultati dei suoi studi, decise di pubblicare la sua opera più voluminosa, le *Exercitationes pathologicae*,[19] grande opera di chirurgia teorico-pratica, in due volumi *in quarto* con molte tavole.

I temi trattati dal Palletta prendono in considerazione gli argomenti che lo avevano interessato maggiormente. Tra questi: l'ortopedia (*De quibusdam coxae vitiis*, *De tubercolis ossivoris*, *De tibiæ diaphysi regenerata*, *De nonnullis vertebrarum fracturis*), la ginecologia (*De uteri polypis*, *De duobus rarioribus cervicis tumoribus*) e le affezioni congenite (*De morbis quibusdam congenitis*).

Con la sua vastissima esperienza il Palletta dimostra di essere effettivamente un chirurgo generale capace di affrontare ogni patologia di interesse chirurgico, ed infatti egli affronta altri importantissimi argomenti quali: le formazioni ascessuali in ogni distretto corporeo, la disfagia, l'angiologia, gli aneurismi, la patologia tiroidea, le affezioni urologiche, le ernie, l'oncologia, etc. Il materiale raccolto è particolamente

---

permette di rinviare il lettore a: R. Dionigi, *Giovanni Battista Palletta (1748-1832). Protochirurgo all'Ospedale Maggiore di Milano*, Milano, Cisalpino, 2016,, in particolare alle pp. 81-109, nonché alla *bibliografia essenziale* conclusiva del volume.

[19] G.B. PALLETTA, *Exercitationes pathologicae auctore Joanne Baptista Palletta equite a Corona Ferrea honorisque legione*, Mediolani, excudebat Societas Typographica e Classicis Scriptoribus Italicis nuncupata, 1820, 2 voll.

vasto e di grande utilità per gli studiosi e per i "pratici", costantemente corredato e arricchito da rassegne complete ed esaurienti in ordine alle singole patologie.

Si ricorda inoltre – sempre nell'ambito della sua vasta attività editoriale – che, nel 1780, Palletta aveva pubblicato il *Trattato delle malattie dei bambini*, traduzione dell'*Underrättelser om barns sjukdomar och deras botemedel* di Nils Rosen Von Rosenstein,[20] opera certamente importante ed influente, che era stata subito collocata dagli studiosi al vertice della letteratura pediatrica. La traduzione del Palletta, realizzata a sedici anni dalla prima pubblicazione del *Trattato* di Rosen Von Rosenstein, rispetto ai diversi trattati di pediatria di cui è ricco il secolo XVIII, appare di particolare interesse, non tanto per la parte dedicata alle malattie dei bambini, quanto per l'originalità, la completezza e la precisione delle note – talora non brevi – aggiunte dal traduttore, che si sofferma a presentare e descrivere alcuni aspetti della società del suo tempo.[21]

Inoltre Palletta si impegna con costanza a contraddire alcune delle secolari usanze che caratterizzano la pediatria del Settecento. Egli è indiscutibilmente un innovatore e, prescindendo dalle inevitabili inesattezze dovute all'incompleta conoscenza della fisiopatologia contemporanea, alcune sue frasi parrebbero scritte con linguaggio moderno ed alcuni precetti essere quelli adottati attualmente.

---

[20] Il medico svedese Nils Rosen Von Rosenstein (1706-1773) è considerato il fondatore della pediatria moderna. Nel suo *Trattato sulle malattie dei bambini* (1764), si occupa particolarmente delle cure infantili generali, alimentari ed igieniche, con numerosi riferimenti alla letteratura contemporanea. L'opera è costituita da una serie di articoli scritti per promuovere una maggiore attenzione sui bisogni dei bambini, problematica avvertita come particolarmente urgente in quanto i dati raccolti su base nazionale dagli uffici statistici registrano, in quegli anni, un alto tasso di mortalità infantile. A tutt'oggi quest'opera è considerata uno dei primi testi moderni di pediatria. Cfr. *Biographie médicale par ordre chronologique d'après Daniel Leclerc, Éloy, mise dans un nouvel ordre, revue et complétée par M.M. Bayle et Thillaye*, tome II, Paris, Adolphe Delahays, 1855, p. 335 e *Dictionary of medical biography*, vol. 4, edited by W.F. Bynum and Helen Bynum, Westport-London, Greenwood Press, 2007, p. 1083, p. 1083; NICOLA LATRONICO, *Storia della Pediatria*, Torino, Edizione Minerva Medica, 1977.

[21] Per un accurato commento critico dell'attività del Palletta nell'ambito della pediatria e della puericultura, cfr. RENATO BETTICA-GIOVANNINI, *Gian Battista Palletta, medico e puericultore del Settecento*, in "Castalia", n. 5-6 (1948), pp. 3-8.

8. Giovanni Battista Palletta fu popolare non solo per essere medico e chirurgo eccellente e accurato ricercatore, ma anche per l'originalità del suo carattere. Molti sono gli aneddoti che fanno riferimento sia alla sua fama sia alla sua personalità, dai tratti stravaganti.[22]

In particolare, anche Carlo Porta menziona il Palletta nell'ultimo dei *Dodes sonitt all'Abaa Giavan* (*Per fagh vedè, e toccà proppi con mann*) tra le glorie degli anatomici milanesi. La composizione del Porta si riferisce ad un articolo, apparso sul secondo numero della "Biblioteca Italiana" (febbraio 1816), in cui Pietro Giordani condannava l'iniziativa dell'editore Francesco Cherubini, scorgendone l'affermazione del particolarismo italiano, che egli voleva superare con l'uso della comune lingua nazionale, solo «istrumento a mantenere e diffondere la civiltà» e premessa al miglioramento delle classi inferiori della società. In risposta, Carlo Porta non perde naturalmente l'occasione di indirizzargli dodici sonetti satirici, chiamandolo *Abaa Giavan*:

Ch'el me sporgia on poo chì, sur don Giavan,
El fregacœucc de quella soa manina,
E ch'el tocca, ch'el studia e ch'el combina
Sti poch donzenn de nomm italian.
S'el leggiarà polit cont attenzion,
E se de gionta el vorrarà notà
Sti nomm sul mennabò di citazion,
Sur don Giavanin d'ôr, ch' el lassa fà
Che ghe mettaroo insemma un regalon
Ch'el vorrarà stantà a portali a cà.
Prest donch ch'el vegna za,
Leggemm: […]
Anatomega: Buzz, Gasper Asell,
Carchen, Bianch, Bium, Monteggia, Magistrell

[22] Per l'aneddotica ed un approfondimento della personalità cfr. GIUSEPPE DE FILIPPI, *Gio. Battista Palletta*, in "Biblioteca italiana o sia giornale di letteratura, scienze ed arti compilato da varj letterati", XVII, tomo LXVII (1832), pp. 276-279; FERRARIO, *Della vita*; PALLETTA, *Elogio*; PORTA, *Elogio*; PASETTI, *Giovanni Battista Palletta*; SCOTTI FOGLIENI, *Gian Battista Palletta*; CASTELLI, *G.B. Palletta*; VITTORIO PUTTI, *Giovanni Battista Palletta*, in *Biografie di chirurghi del XVI e XIX secolo*: *Magati, Palletta, Scarpa, Mathijsen, Fabbri, Rizzoli, Margary, Paci*, Bologna, L. Cappelli, 1941, pp. 19-22; BETTICA-GIOVANNINI, *Gian Battista Palletta*; L. BELLONI, *Per la storia della Medicina*, Bologna, Arnaldo Forni Editore, 1985, pp. VI-VIII.

Tron, Cuni, Mazzucchell,
E Riboli, e Paletta, e l'Ospedaa.[23]

In vita, Palletta ebbe grande considerazione sia da parte della cittadinanza che delle autorità locali. Numerosi segni dell'ammirazione e dell'apprezzamento tributatigli in vita emergono nella produzione poetica e letteraria di alcuni noti autori della sua epoca, tra i quali, ad esempio, Alessandro Manzoni.

In effetti, la reputazione del Palletta, Protochirurgo dell'*Ospedale Maggiore*, era tale che lo stesso Napoleone, nel 1797, in un periodo di timore circa lo stato della propria salute, lo avrebbe chiamato a consulto nella Villa Pusterla di Mombello, dove abitava con le sorelle.[24]

Secondo tutti i primi e più autorevoli biografi (De Filippi, Ferrario, Marco Palletta, Porta, Pasetti, Castelli, etc.), fu lo stesso Napoleone a nominare il Palletta membro della Legion d'Onore e cavaliere della Corona ferrea d'Italia. Se ciò è sicuramente vero per la Legion d'Onore, non può tuttavia esserlo per la nomina a cavaliere della Corona ferrea.[25]

---

[23] CARLO PORTA, *Dodes Sonitt all'Abaa Giavan*, in *Poesie milanesi di Carlo Porta e Tomaso Grossi con alcune inedite*, Milano, Casa Editrice Guigoni, 1891, pp. 285-286.

[24] CASTELLI, *G.B. Palletta*, p. 176. Anche secondo Volante il Palletta sarebbe stato medico consulente di Napoleone durante il soggiorno a Mombello nel 1797: F. VOLANTE, *L'opera anatomo-patologica di Gio. Batt. Palletta (nel centenario della sua morte). Comunicazione fatta nella seduta del 15 luglio 1932*, in "Giornale della R. Accademia di Medicina di Torino", 95 (1932), pp. 1-6. Sul fatto che Napoleone stimasse grandemente il Palletta non vi sono dubbi e diversi biografi sottolineano questo aspetto della vita del chirurgo. È lecito però avanzare qualche dubbio su quanto scrive il Castelli a proposito della chiamata a consulto di Napoleone nella Villa Pusterla di Mombello. Di questo episodio non vi è documentazione ed il Castelli è l'unico a citarlo. In effetti, il chirurgo di Bonaparte era Pietro Moscati, che per lui divenne quasi una persona di famiglia. Fu più volte invitato a cena, fu certamente il suo medico curante e il suo consigliere più fidato nelle faccende interne della vita milanese. Fu anche il medico di Giuseppina: "Voi siete un uomo onesto: vi affido mia moglie" disse a Moscati un giorno che si doveva allontanare da lei. E sempre a Moscati affidò altri ammalati fra i politici e i militari del suo *entourage* (ETTORE ROTA, *Milano napoleonica*, in *Storia di Milano*, vol. XIII, Milano, Fondazione Treccani degli Alfieri per la Storia di Milano, 1959, pp. 1-350, p. 80).

[25] L'onorificenza di Cavaliere della Corona ferrea è stata verosimilmente conferita da Francesco I e non da Napoleonecfr.: Giuseppe Ferrario, *Della vita del Professore Gio. Batt. Palletta*, Milano, coi tipi di Paolo Andrea Molina, 1833, p. 23.

Nel giugno del 1811, in compagnia del ministro del Culto Bovara, di salute malferma, Palletta si recò a Parigi per assistere al battesimo del figlio di Napoleone e, per la considerazione di cui godeva tra i dignitari di corte, fu invitato a trattenersi nella capitale francese per un periodo di ben sette mesi, avendo così l'opportunità di visitare i maggiori ospedali ed incontrare i clinici più illustri, suoi conoscenti e corrispondenti, che in massima parte già lo conoscevano di fama. Durante il suo soggiorno parigino, un personaggio eminente si trovò nella necessità di sottoporsi ad un rilevante intervento chirurgico, per il quale venne richiesto il suo autorevole parere. Il Palletta sconsigliò l'intervento, senza tuttavia essere ascoltato, e l'esito infelice, da lui previsto, dandogli ragione accrebbe la sua reputazione.[26]

La sua celebrità e l'affabile cordialità nei confronti dei pazienti gli procurarono una fiorentissima clientela in Milano, nella Lombardia ed in tutta l'Italia settentrionale. In seguito alla precoce scomparsa di Giovanni Battista Monteggia e all'abbandono della professione da parte di Antonio Scarpa, Palletta rimase a lungo il principale riferimento chirurgico del Paese.

Per la sua grandezza come scienziato e clinico chirurgo egli fu in contatto con gruppi di persone di ogni estrazione sociale, portando ovunque lo stesso abito e contegno, e mostrando repulsione verso ogni forma di ambizione e superbia intellettuale: egli esercitò la professione medica essenzialmente come una missione e un servizio.

Il Palletta non fu ideatore di nuovi strumenti chirurgici, di nuove "macchine", di farmaci od operazioni "nuove", ma spesso verificò, criticò ed eliminò dal suo repertorio metodi altrui che aveva avuto modo di dimostrare errati. Fu uno sperimentatore, le cui ricerche scientifiche possono essere considerate le prime dimostrazioni di quella disciplina che oggi definiamo "chirurgia sperimentale". L'anatomia e la chirurgia, invece, gli sono debitrici di scoperte, monografie, illustrazioni e di una vastissima serie di opinioni e di "pensieri nuovi", che qualificano il Palletta come "classico" e originale.

Tra i "pensieri nuovi" vi sono certamente l'uso dell'"aria fissa" (anidride carbonica) nel trattamento della diarrea, dello scorbuto, delle ulcere torpide e dei tumori; l'olio e le carni di lucertola nella scrofola; l'applicazione dell' "idrocontasterio"[27] nell'impaccamento fecale. Sua

26 PASETTI, *Giovanni Battista Palletta*.

27 «Termine di chirurgia. Chiamasi con questo nome una macchina che serve ad

fu la proposta di favorire la fistolizzazione dell'intestino necrotico nell'ernia strozzata e fu certamente il primo ad abbandonare la semplice legatura dei voluminosi polipi uterini, proposta precedentemente dal chirurgo francese Pierre-Joseph Desault, adottando invece la resezione e asportazione.

Una chirurgia, quindi, caratterizzata da un saggio eclettismo, dal desiderio di sapere e di conoscere come stimolo intellettuale, dalla necessità di accertare, mediante prove e sperimentazioni, l'esistenza, la qualità, la conformità, la riproducibilità di ipotesi, di terapie e di procedure chirurgiche, spesso proposte o adottate solo per consuetudine.

Determinato, deciso, essenziale nel gesto chirurgico, fu anche leale e sincero nell'ammettere pubblicamente i propri errori. Il 24 giugno 1812, ad esempio, lesse all'*Istituto di Scienze, Lettere ed Arti* una memoria che riguardava l'asportazione per errore di un utero durante un intervento che prevedeva solo l'asportazione di un sarcoma ulcerato della cervice uterina.

Nel 1826, affaticato e stanco, Giovanni Battista Palletta decise di lasciare l'*Ospedale Maggiore*, al quale pur era molto affezionato e dove aveva trascorso tanti anni di lavoro. Aveva creato una scuola che aveva reso l'*Ospedale* milanese un centro di istruzione di eccezionale livello. Collocato a riposo, continuò ad essere rispettato e onorato dalla sua città e, nel 1829, la Zecca di Milano coniò in suo onore una medaglia recante, nel *recto*, intorno alla sua immagine, la seguente iscrizione: «Giovanni Battista Palletta Medico Chirurgo Filosofo» e, nel *verso*: «Lume ed Onore della Chirurgia incomparabile per filantropia». Il celebre chirurgo morì la sera del 27 agosto 1832, all'età di 84 anni.

9. La biblioteca di Giovanni Battista Palletta costituisce uno dei principali fondi della biblioteca dell'*Ospedale Maggiore* di Milano, fondo che gli stessi curatori del catalogo,[28] pubblicato nel 1991, defini-

introdurre per l'ano negl'intestini, fin anche oltre la valvola del colon, una maggiore o minor copia di fluido, con mite, o, giusta l'iopo, con veemente velogità» (*Dizionario della lingua italiana*, vol. IV, Padova, Nella Tipografia della Minerva, 1828, s.v. *idrocontasterio*).

[28] *Catalogo del fondo Giovanni Battista Palletta*, a cura di Gigliola Marsala e Laura Panzeri, introduzione di Giorgio Cosmacini, Milano, Editrice Bibliografica, 1991.

scono il più omogeneo e interessante della raccolta ospedaliera. La biblioteca del "protochirurgo" milanese fu donata all'ospedale dal nipote Marco: un lascito la cui consistenza era, al momento della donazione, di 1855 volumi e 703 opuscoli. Per vicissitudini che hanno vessato le raccolte librarie dagli anni della Seconda guerra mondiale sino a oggi,[29] l'attuale consistenza del fondo è di 1056 opere appartenute al chirurgo e di 68 volumi del nipote Marco.

È particolarmente interessante scorrere con attenzione l'elenco delle opere appartenute al Palletta, il quale, attraverso i molti libri studiati e raccolti, palesa un elevato profilo culturale ed un aggiornamento scientifico professionale che giustificano appieno la fama da lui raggiunta a cavallo di due secoli in cui, anche grazie ai suoi contributi, la chirurgia si affacciò alla modernità.

Nel catalogo si apprezza la grande curiosità dell'uomo, la molteplicità delle sue conoscenze e competenze, che non si limitano alle opere strettamente mediche, nonostante comprenda naturalmente le opere più rilevanti dei maggiori medici e chirurghi del suo tempo.

Un ulteriore interesse di Palletta pare siano stati il termalismo, la balneoterapia e le acque minerali. Infatti, tra i libri da lui raccolti, vi sono pubblicazioni che riguardano l'effetto benefico e le proprietà terapeutiche delle acque di diverse stazioni termali.

10. Giovanni Battista Palletta con passione e serietà si dedicò in gioventù agli studi presso il *Kollegium Spiritus Sanctus*, il collegio dei Gesuiti a Briga, dove studiò per ben dieci anni. La formazione umanistica, prevalentemente in lettere e filosofia, ricevuta in questo famoso e reputato collegio pose le basi che gli avrebbero permesso di affermarsi

[29] I bombardamenti che infierirono su Milano non risparmiarono la biblioteca dell'Ospedale e gran parte dei suoi libri furono trasferiti nella nuova sede di Niguarda, in attesa di essere ricollocati in via Festa del Perdono, in quella che sarebbe dovuta diventare la biblioteca dell'ospedale. Tutte le opere sono schedate secondo le *Regole Italiane di Catalogazione per Autori* (RICA) ed ogni scheda, in ordine alfabetico, con data di nascita e di morte degli autori, è corredata da note, che spesso forniscono la storia dell'esemplare, ricostruita attraverso appunti manoscritti, dediche e note di possesso (MAURIZIO GIROLAMO, *Regione Lombardia. Servizio biblioteche e beni librari e documentari, Catalogo del fondo Giovanni Battista Palletta*, recensione in "Biblioteche oggi", n. 2 (1993), pp. 78-79).

come uno dei primi chirurghi eruditi, la cui preparazione culturale avrebbe contribuito a modificare l'immagine stereotipata del chirurgo rozzo e ignorante, che solitamente usciva dalla corporazione dei barbieri.

Anche l'apprendimento del latino e del tedesco molto contribuì al suo successo. Basti ricordare che la laurea in Filosofia e Medicina conseguita all'università di Padova fu il risultato, oltre che del suo notevole *cursus studiorum*, soprattutto dell'ammissione alla prestigiosa *natio germanica* dell'ateneo patavino, essendo stato considerato *helvetius* proprio grazie alla sua padronanza della lingua tedesca.

Ancor prima di laurearsi a Padova, aveva già deciso di intraprendere la carriera chirurgica. Infatti, lo studio della giurisprudenza, a cui si era in precedenza dedicato nei pochi mesi trascorsi a Milano, non lo aveva convinto affatto. Si era reso conto di essere portato verso studi meno speculativi, più operativi e, se possibile, più coinvolgenti dal punto di vista sociale, grande essendo il suo desiderio d'essere utile al prossimo. Così, a 15 anni, era entrato come studente di chirurgia nel *Collegio-Convitto degli Allievi Chirurghi*, alla *Magna Domus Hospitalis*, la *Ca' Granda*, ovvero l'*Ospedale Maggiore*.

Cionondimeno, Palletta continuò a riflettere sul mondo e sull'uomo, sul senso dell'essere e dell'esistenza umana ed anche indagò l'applicazione concreta dei principi desunti con la riflessione e il pensiero.

La storia della chirurgia è unanime nel ritenere che Antonio Scarpa, a Pavia, e Giovanni Battista Palletta, a Milano, abbiano contribuito a risollevare il prestigio dell'arte chirurgica, da troppo tempo in declino, riuscendo con i loro studi e la loro attività clinica ad essere considerati "uomini sommi" della comunità chirurgica internazionale.

Un raffronto tra questi due personaggi, uniti da «amicale affetto e stima», non è semplice, e induce comunque a porre in evidenza alcune differenze sostanziali. Scarpa diventò presto arbitro di ogni controversia all'interno dell'ateneo pavese: per vari anni egli fu il vero *dominus* delle carriere universitarie. Era un magistrale anatomico, ma la sua abilità chirurgica venne spesso messa in discussione. Molta della sua fama internazionale fu certamente dovuta alla collaborazione con Faustino Anderloni, eccellente intagliatore di rami di carattere scientifico e didattico, che, dal 1799 in avanti, curò tutte le opere pubblicate dallo Scarpa, incidendo splendide tavole, da solo o con la collaborazione di una sceltissima bottega, dove crebbero grandi interpreti dell'incisione come il fratello Piero Anderloni e Giovita

Garavaglia.[30] Le opere dello Scarpa si rivolgevano agli addetti: agli anatomici, ai chirurghi, alla comunità scientifica nell'accezione più restrittiva.

Palletta, invece, era uomo di non molte parole, anche se le sue erano frequentemente espressioni piene di arguzia. Era cauto nel pronunciarsi davanti al malato, pur dimostrando sicurezza nel giudizio. Non disponendo della collaborazione di grandi disegnatori e incisori, le sue descrizioni delle dissezioni anatomiche sono particolarmente accurate e precise. Leggendo le sue opere, ci si rende conto della sua profonda cultura storico-umanistica. Fa sempre precedere, alla descrizione di ogni malattia, un'accurata revisione storica, con citazioni delle fonti antiche più accreditate. Una conoscenza storica, la sua, intesa soprattutto come supporto indispensabile e punto di partenza per procedere in maniera corretta nella ricerca e nella sperimentazione.

Anche le malattie più rare e solo eccezionalmente descritte in passato avevano ai suoi occhi un'importanza non marginale. Al riguardo, scrive, con grande acume e lungimiranza:

> Non sono da trascurarsi per niun conto le storie delle malattie rare e sconosciute, perché, sebbene il più delle volte manchino dell'immediata utilità, sono però dei materiali che, messi da parte per poco, indi congiunti ad altri, possono servire di base ad un solido edificio.

Nei suoi scritti Palletta, ancor più che alla comunità scientifica, pare rivolgersi al "medico pratico", affinché possa prevenire e curare, evitando gli errori del passato che ostinatamente permangono anche nel presente. Egli si impone come un nuovo e moderno esponente di un ormai inevitabile "revisionismo chirurgico". Con decisione, e talora con espressioni non sempre moderate, critica terapie e tecniche chirurgiche che considera inadeguate o comunque superate, proponendo soluzioni più adatte ed efficaci, dettate dalla sua eccezionale esperienza maturata nell'*Ospedale Maggiore.*

Piace sottolineare il grande interesse che il Palletta nutriva per la patologia neonatale, sulla scia dell'opera svolta nei decenni precedenti dal suo maestro Pietro Moscati, e pure dal di lui padre Bernardino, i

30 SERGIO MOMESSO, *La collezione di Antonio Scarpa (1752-1832)*, Cittadella, Bertoncello Artigrafiche, 2007, p. 25.

quali, come è noto, ebbero un ruolo preminente nell'elaborazione di un progetto di riforma dell'assistenza agli esposti, individuando gli interventi necessari, secondo approcci diversi.[31]

Il Moscati prima e il Palletta successivamente si proponevano di favorire lo sviluppo dei piccoli infelici in giovani robusti da inserire nella società con un onesto lavoro.

Delle sue opere di maggior rilievo si è già detto: non vi è settore medico-chirurgico nel quale non ebbe da esprimere la sua opinione. Due furono i principi a cui spesso si riferiva sia nella sua attività professionale, sia nell'insegnamento: confidare nella natura e non trascurare le sempre presenti ed attive energie dell'organismo. Per Palletta, una vigile attesa era più efficace di un incauto ed immediato intervento ed alcune guarigioni si potevano ottenere spontaneamente evitando inutili e dannose aggressioni chirurgiche.

Un atteggiamento, questo, di semplice buon senso, che si ripresentava ad esempio nella vera e propria battaglia che per tanti anni fece contro la pratica del salasso.

La sua "modernità" emerge anche dal favore con cui sempre accolse la sperimentazione, sia clinica che di laboratorio. Con la sperimentazione clinica riuscì a smascherare le presunzioni di improvvisati "sapienti" ciarlatani (si pensi al caso della stravagante cura del cancro di tale dottor Flores che proponeva l'ingestione di lucertole).

Per quanto riguarda la sperimentazione animale, a lui si rivolsero Giuseppe Baronio e anche Alessandro Volta. Baronio praticò infatti alcuni esercizi di rinoplastica su cadavere nella sala anatomica dell'*Ospedale Maggiore* di Milano. Medico fisico, l'autore degli *Innesti* aveva necessità di coinvolgere anche maestri della chirurgia per soddisfare le finalità applicative delle proprie ricerche, cosa che fece fin dagli esperimenti sul montone, coinvolgendo Palletta nelle operazioni.[32] Alessandro Volta, dal canto suo, incuriosito dalle nuove esperienze del Galvani sull'elettricità animale, dopo un iniziale scetticismo, infine si «convertì», come egli stesso disse, cominciando «ad essere testimonio oculare e operatore stesso

---

31 Cfr.: A. Terreni, "Parere di Bernardino e Pietro Moscati intorno alla scelta d'un metodo più atto alla maggiore possibile conservazione de' bambini esposti nello Spedal Maggiore di Milano, scritto per superior comando nel giugno dell'anno 1771", Appendice documentaria I, in R. Dionigi, Bernardino Moscati, cit., pp. 127-146.

32 Giuseppe Baronio, *Degli innesti animali*, Milano, Dalla Stamperia e Fonderia del Genio, 1804, pp. 12, 32, 45, 46, 58, 64.

dei miracoli, e passato forse dall'incredulità al fanatismo». Nella fase iniziale di questi suoi esperimenti si avvalse prevalentemente delle rane, ma successivamente decise di ripeterli su animali a sangue caldo «aiutato la prima volta dall'eccellente chirurgo e Anatomico di Milano D.r Palletta, coll'assistenza pure del D.r Baronio».[33]

Palletta, contrariamente al suo maestro Moscati, che divenne il consigliere più fidato di Napoleone nelle faccende interne della vita milanese, mai volle interessarsi di politica, mai si schierò apertamente e, pur avendo messo da parte un consistente patrimonio grazie alla sua intensa attività chirurgica, fu sempre coscienzioso e alieno da ogni pratica di affari.

In conclusione, Giovanni Battista Palletta può considerarsi il primo chirurgo che in Italia, tra il Settecento e l'Ottocento, ossia agli albori della chirurgia moderna, si sia presentato come un vero innovatore, grazie alla sua opera di revisione di teorie e tecniche obsolete, diventando uno dei massimi esponenti di una nuova cultura medico-chirurgica liberata da secolari incrostazioni, sovvertendo in tal modo antiche "certezze" e credenze fisio-anatomo-patologiche e lasciando in eredità alle nuove generazioni di chirurghi, in molteplici settori, una innumerevole quantità di felici intuizioni, alcune delle quali possono considerarsi tuttora valide.

[33] ALESSANDRO VOLTA, *Collezione dell'opere del Cavaliere Conte Alessandro Volta, patrizio comasco*, tomo II, parte I, Firenze, nella Stamperia di Guglielmo Piatti, 1816, pp. 36-37.

ALCHIMIA, PRELUDIO ALLA CHIMICA?

Nota del m.e. GIANNANTONIO SACCHI LANDRIANI (*)

(Adunanza del 21 aprile 2016)

SUNTO. – Scopo di questa comunicazione è la ricerca, nella Biblioteca scientifica antiquaria dell'Istituto Lombardo, di documenti che preludano alla chimica. Come noto la chimica ha avuto un rigoroso assetto alla fine del XVIII secolo. Di fatto abbiamo trovato alcuni libri concernenti complicate e raffinate empiriche tecnologie particolarmente relative alla lavorazione di metalli e alla realizzazione di colori per la realizzazione di ceramiche. È necessario ricordare che durante la Rinascenza il clima culturale scientifico era alchimistico. La diffusione di informazioni riguardanti queste attività erano gelosamente custodite fra i membri delle corporazioni. Questa può essere considerata una ragione della povertà di documenti originali.

***

ABSTRACT. – This paper aims to a review, in the antiquarian scientific library of the Istituto Lombardo, able to find forerunner documents of chemistry. As it is well known chemistry found a rigorous settlement at the end of XVIII century. Actually we have found some books concerning complicated refined empirical technology particularly devoted to metal and making of colours for ceramic processing. It is necessary remember that during Renaissance the scientific cultural climate was alchemistical. So the spread of information concerning this kind of activities among members of guilds was nearly secret. This is the reason why the number of documents found is very poor.

CONSIDERAZIONI INTRODUTTIVE

Il proposito di questo intervento è dettato dal desiderio di avvicinarci alla conclusione di un itinerario iniziato da un paio d'anni. In precedenti comunicazioni avevamo presentato notizie relative a esplorazioni di cinquecentine scientifiche, appartenenti alla biblioteca del nostro

(*) Politecnico di Milano, Istituto Lombardo Accademia di Scienze e Lettere, Milano, Italia. E-mail: glansacchi.landriani@gmail.com

Istituto, in tema di matematica e fisica, di medicina e scienze naturali, di astronomia. Sembrava non irragionevole (al fine di ampliare l'esplorazione dei documenti scientifici del XVI secolo) rivolgere l'attenzione a quanto poteva preludere alla nascita della chimica, coscienti della circostanza che questa disciplina avrebbe avuto un assetto scientificamente definito a partite dal XVIII secolo.

Si tratta di un patrimonio librario (quello strettamente relativo a questo argomento) dotato di un numero esiguo di esemplari (e modesto contributo è pubblicato in cataloghi generali, quali ICCU e EDIT16).

Gli ingegnosi pensieri e le drammatiche speranze del XVI secolo erano, sotto la cappa dell'Astrologia, un insieme di tentativi di osservazioni di fenomeni complessi, di conclusioni fumose, ma formulate con acume, fantasia e ingegno. In quanto segue daremo rilievo ad alcuni autori che ci sembrano i più significativi, tra quelli della nostra biblioteca, quali anticipatori di concetti propri di innovatrici tecnologie.

La circostanza che nel XVI secolo la cultura non fosse organizzata, come l'attuale, in distinte discipline definite in ambiti precisi, rende complesso il discernere in un autore quanto potesse preludere alla non ancor definita Chimica.

Tutto era confluente in un ambito immaginifico dove la botanica, l'astronomia, la farmacologia, la medicina, confinavano con la teologia o con la sapienza aristotelica.

E' ammirevole comunque la profonda aspirazione degli antichi autori a prefigurare una conoscenza organizzata.

Significative notizie sono contenute nell'ambito di dottrine dalla forte tradizione come la botanica, la farmacia, l'anatomia.

Un argomento che merita una particolare attenzione è relativo alla preparazione dei colori (grandi vetrate di cattedrali, dipinti, affreschi).

Il Cinquecento è il secolo d'oro della maiolica italiana. Richiesta dai Signori, le stoviglie istoriate ormai di svariate forme, facevano bella mostra nei saloni ed erano spesso affiancate da stoffe pregiate e suppellettili in oro e argento, quali simboli di potenza economica.

La produzione ceramica in questo periodo fece un salto di qualità, arrivando a dialogare strettamente con la pittura contemporanea (spesso copiandone le composizioni grazie al circolare delle incisioni) con la nascita di botteghe di pittori specializzati in questo tipo di decorazioni.

Le tecniche utilizzate presupponevano conoscenze empiriche complesse relative al valore cromatico di una sostanza, tenendo conto delle dimensioni degli oggetti e delle particelle usate nei pigmenti. La

composizione dei colori era eseguita secondo ricette spesso tenute segrete dagli artefici, o nell'ambito delle specifiche corporazioni.

La composizione chimica dei materiali usati in quell'epoca è stata, e non completamente, analizzata soltanto a partire dal XIX secolo. Oggi lo è sulla base delle proprietà studiate per sistemi nanotecnologici.

Si ha notizia di interessanti contributi cinquecenteschi, che citano la problematica connessa alla produzione di colorazioni e pertanto con osservazioni che appartengono a prolegomeni della futura chimica, almeno dal punto di vista tecnologico.

Stupisce che Giorgio Vasari (1511-1574) ne LE VITE, opera edita nel 1550, ammette di non aver saputo (o forse voluto) svelare i segreti degli smalti, dei colori, delle arti applicate alla tecnologia del fuoco (vetro, mosaico, ceramica).

Grande notorietà ebbe all'epoca, e tutt'ora la conserva, l'opera del 1548 di Cipriano Piccolpasso (1524-1579) *[riprodotta per intero alla pagina line Google] LI TRE LIBRI DELL'ARTE DEL VASAIO [IN QUIETE-Introduzione di Giulio Vada MMII]* il cui contenuto è dedicato in gran parte alla smaltatura delle terrecotte con interessanti indicazioni, non certe ma attendibili, sugli elementi e relative applicazioni usate da Luca della Robbia nel XV secolo e dai suoi eredi nel XVI e XVII secolo.

Si tratta di un'opera che tende a svelare i segreti dei «vasari» senza tuttavia abbordare una elaborazione abbastanza sistematica da appoggiare ad un tecnica scientifica. Non è trascurabile che il Piccolpasso fosse coevo ed interlocutore di Donatello e di Brunelleschi. Inoltre è opportuno ricordare che la ceramica arrivò a dialogare con i pittori contemporanei e con le loro botteghe.

Il procedimento contemplava la preparazione e la foggiatura dell'impasto argilloso, la sua essicazione e cottura a biscotto, quindi la smaltatura. Materia prima è sabbia, silice non inquinata da ferro o titanio.

La silice fonde a più di 1700° C, temperatura non raggiungibile allora, ed era quindi necessario aggiungere fondenti per aumentarne la fusibilità e però assicurare la scorrevolezza degli smalti ossia alcali, sodio, potassio, e piombo. Era importante fondere il piombo con la silice (silicato di piombo) ossia vetro a bassa fusibilità dotato di alta brillantezza (che si deve al piombo). 3 parti di piombo+1 parte di sabbia.

Il Piccolpasso afferma di conoscere il segreto di spalmare l'ORPIMENTO (pigmento d'oro+solfuro d'arsenico in cristalli di color giallo oro) sui bordi del BISCOTTO (terracotta o ceramica di prima cottura destinata a ricevere il rivestimento vetroso).

Un adagio del tempo affermava:

L'ossido di stagno conferisce bianco. L'arsenico dà brillantezza.
Il piombo fa lustro, lo stagno fa bianco.

La stima della temperatura era effettuata, in quel tempo precedente all'invenzione dei termometri, mediante l'osservazione ad occhio nudo del colore assunto dal pezzo in lavorazione.

Le tecniche utilizzate presupponevano conoscenze empiriche complesse relative a colori, dimensioni, particelle usate nei pigmenti.

Benché non strettamente attinente al tema di prologo della chimica, ma piuttosto alla fisica (in particolare l'ottica) e con accenni di anatomia (oftalmia), comunque con caratteristiche scientifiche marcate e preveggenti per l'epoca, è l'opera, presente nella nostra biblioteca di Cartesio [RENATUS DES CARTES] (1569-1650) dal titolo de HOMINI FIGURIS. In buono stato di conservazione consta di 122 pagine + tavole fuori testo, h 20 cm.

L'opera del senese Vannuccio Biringuccio

Al senese Vannuccio BIRINGUCCIO (1480-1539), autore significativo nella storia della scienza, ma soprattutto della tecnologia, dobbiamo riconoscere notevole maestria nella presentazione di tecniche applicate alla produzione di oggetti metallici, come illustreremo nel seguito. Tecnologie che facevano appello all'uso della fusione dei metalli per creare armi, oppure campane, tanto che troviamo nel frontespizio della sua maggiore opera la sintesi del contenuto che segue.

P I R O T E C H N I A
L I D I E C I L I B R I D E L L A
P I R O T E C H N I A
Nelli quali fi tratta non solo la diuerfita
delle minere,ma ancho quanto fi ricer
ca alla prattica di effe:e di quan//
s'appartiene all'arte della
fufione ouer getto de me
talli,e d'ogni altra co//
fa a quefta fo//
migliante
Far campane, artiglieria, fuochi artificiali e diuerse cose utilissime.

Per ottenere pregevoli risultati ci si valse della sapienza ereditata da antichi studiosi esperti nell'individuare le miniere dalle quali trarre il grezzo per isolare soprattutto metalli. E con questi creare leghe quali il bronzo, l'ottone, l'acciaio.

Nella stessa utopia alchimistica, di trasformare qualunque vile sostanza in oro, v'è un germe delle trasformazioni tipiche dell'attuale disciplina chimica.

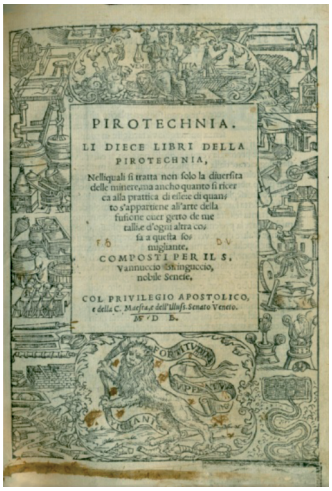

PIROTECHNIA.
LI DIECE LIBRI DELLA PIROTECHNIA,
Nelliquali si tratta non solo la diuersita delle minere ma anchoquanto si ricerca alla prattica di esse in quanto s'appartiene all'arte della fusione cioe getto de ogni sorte d'ogni altra cosa a questa sogetta
COMPOSTI PER IL S. Vannuccio Biringaccio, nobile Senese,
COL PRIVILEGIO APOSTOLICO, e della C. Maestà dell'Illustr. Senato Veneto.
M. D L.

Fig. 1 – L A PIROTECHNIA LI DIECI LIBRI DELLA PIROTECHNIA edizione 1550.

In *Fig. 1* si presenta L A PIROTECHNIA LI DIECI LIBRI DELLA PIROTECHNIA editi per il s. Vannuccio Biringuccio nobile senese (1480-1539) COL PRIVILEGIO APOSTOLICO, e della C.Maestà dell'Illuss. Senato Veneto nel MDL

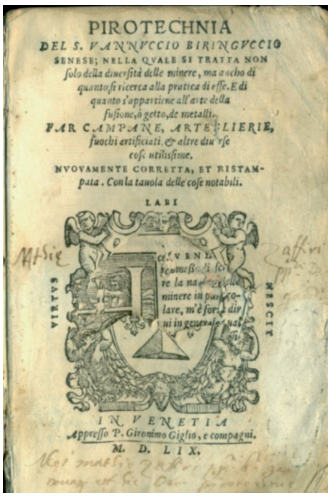
Si tratta probabilmente di una seconda edizione. Questa edizione del MDL non è catalogata ICCU.

La prima edizione citata [IT/ICCU/RAV/1930982] sembra essere del 1540 con un solo esemplare. Ma gode di ampia bibliografia del XX secolo.

In *Fig. 2* è presentata l'edizione datata MDLIX in VENETIA – *Appresso Geronimo Giglio e compagni.* Si tratta di un esemplare purtroppo fortemente danneggiato in frontespizio e con scritte a mano illeggibili. È catalogato [IT\ICCU\BVEE\014052].

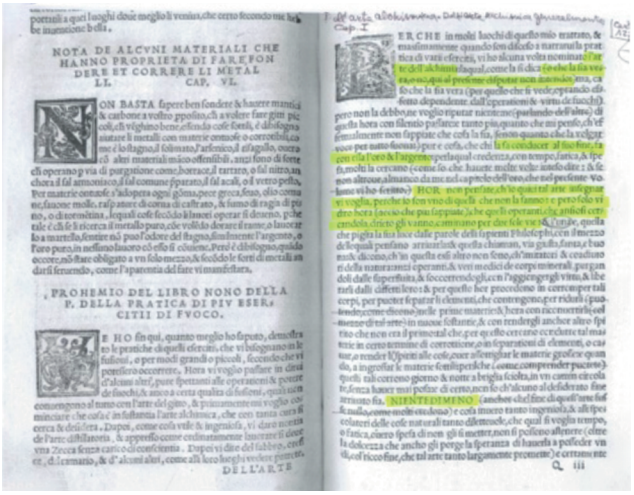
Si tratta di una pubblicazione avvenuta un decennio dopo la

morte dell'autore. Molto diffusa all'epoca e tuttora molto studiata dagli specialisti.



PIROTECHNIA
DEL S. VANNVCCIO BIRINGVCCIO
SENESE; NELLA QVALE SI TRATTA NON
solo della diuersità delle minere, ma a pieno di
quanto si ricerca alla pratica di esse. E di
quanto s'appartiene all'arte della
fusione, o getto, de metalli.
E PAR. CAMPANE, ARTE, E LIERIE,
fuochi artificiati, & altre diuerse
cose utilissime.
NVOVAMENTE CORRETTA, ET RISTAM-
pata. Con la tauola delle cose notabili.
LABI
IN VENETIA
Appresso P. Girolimo Giglio, e compagni.
M. D. LIX.

Fig. 2 – Successiva edizione datata MDLIX.



NOTA DE ALCVNI MATERIALI CHE
HANNO PROPRIETA DI FARE FON
DERE ET CORRERE LI METAL
LI. CAP. VI.

PROEMIO DEL LIBRO NONO DELLA
P. DELLA PRATICA DI PIV ESERC
CITII DI FVOCO.

DEL L'ARTE

Fig. 3 – La pagina 122 dedicata al PROEMIO DEL LIBRO NONO.

In *Fig. 3* è rappresentata la pagina 122, dell'edizione del MDL, dedicata al PROEMIO DEL LIBRO NONO con DELLA PRATICA DI PIÙ ESERCIZI DI FUOCO (cartella 123).

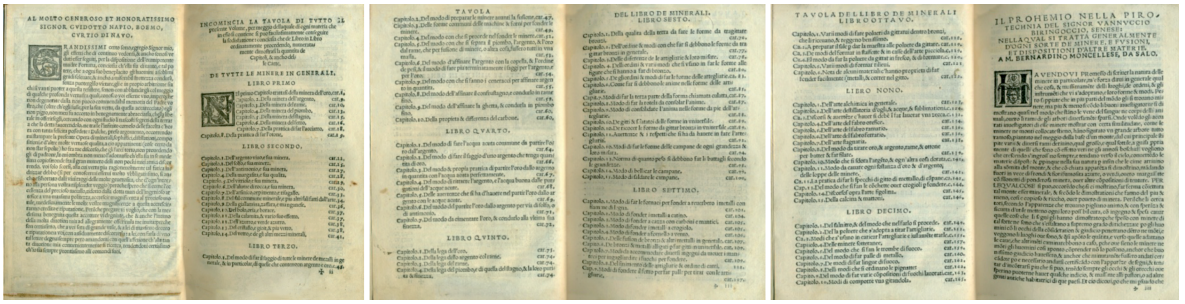


Fig. 4 – Completo indice analitico di PIROTECHNIA edizione del MDL.

L'insieme della *Fig. 4* è la presentazione completa dell'indice analitico di PIROTECHNIA edizione del MDL. I dieci libri sono descritti con precisione e trattano delle miniere d'argento, rame, piombo, ferro e della pratica di fare l'acciaio e l'ottone. Indicano inoltre come lavorare i minerali per isolare gli elementi con tecniche di fonderia.

Considerazioni di Biringuccio sull'arte alchimica

La pagina a destra (*Fig. 3*) è dedicata al tema DELL'ARTE ALCHIMICA IN GENERALE cap I (car 123).

Sembra opportuno tracciare una sintesi delle considerazioni in tema di Alchimia di Biringuccio, riconoscendogli una notevole capacità di preveggenza sugli sviluppi della conoscenza.

L'autore, nelle cartelle 123 e 124 della sua opera, si chiede cosa sia *"l'arte alchimica che oggi con tanta cura si cerca e desidera"*.

Premette la distinzione di due vie di praticare l'alchimia: la prima rispettabile ancorché immatura, la seconda cattiva ed esecrabile. Si pone in una posizione prossima alla prima pur dimostrando eccezionale capacità critica.

In particolare riconosce la notorietà della prima via pur essendo scettico sui risultati, consistenti essenzialmente nella trasformazione della materia con il miraggio di ottenere oro e argento. Ma è un estimatore di quanti, da lui definiti buoni alchimisti, cercano di capire in definitiva la natura delle materie, in particolare metalliche. Si professano medici dei corpi minerali e collaboratori della natura in quanto cercano di purgare i detti corpi dalle impurità, di liberarli dei difetti e se necessario di aggiungere virtù. La tecnica ricorrente è quella della frantuma-

zione, del ridurre in polvere per isolare le sostanze di base, e mescolandole di trovare nuove sostanze.

In generale costoro soni colti, rigorosi e in itinere cercano di trovare benefici per l'umanità. Esempio fondamentale è lo sviluppo della farmacopea. Il buon alchimista non deve essere povero, poiché la sua disciplina è costosa; non deve essere avaro, perché moralmente tenuto a far conoscere i propri risultati; deve essere un genuino amante della conoscenza.

Biringuccio ha parole severe nei confronti di coloro, definiti cattivi alchimisti, che si valgono dell'alchimia con mala intenzione. Operano con tecniche scorrette e disordinate (per esempio nella cattiva gestione dei fuochi) e vogliono spacciare l'ottone come oro, il vetro come cristallo, gli smalti come gemme. Si comportano in sostanza come avidi imbroglioni.


LIBRO NONO

HO ANCHO veduto un'altro fornello con una torre in mezzo quadro, più semplice, & con quattro nasse con li suoi regitti come l'antidetto. N'HO ancho veduto uno diuerso da tutti questi, con molti nasse, quello era un forno alto circa due braccia e mezzo, con una volta cotta, largo di diametro circa tre quarti, l'era quasi simile a una stufa secca, & haueua da piedi, due comunicate a far la volta, ne l'piu eccellente nella grossezza si erano murati a ordine molti urinaletti di terra attorno per dentro il fondi de quali passauano dentro al fornello, la bocca larga vn'oncia fuori del fornello, sopra ingual li commettena un capello di vetro col lambicco, di sopra nella fornitta era una campanetta, sopra le fornello li gli daua fuoco di fiamme per fin che fi scaldaua, mettena figli poi carbone perche lo mantenesse, ingual ardendo in primo alpetto mi parue bello, perfando poi maglio non fapeua fil douesse lauda re, ma per parermi piu presto cosa appartenente che utile, perche noto qual che fi faceua con tanti nasi, lambichi, e recipienti come è bisogna fare li poteua con due o tre campane in un stesso fornello, & oltre per ardere l'hauea piu fuoco il primo ordine chel feconde,el feconde chel terzane cosi successiue, però mi rifolsi a dire che non mi piacesse, e che per hauerli a far una cosa fimile, correi che hauea fatte una troppo per laquel fi faldasseno tanti fondi delli urinaletti da capo folo hauesse un puoco di fitto per dar occasione a fattine alo fi calore. MA HORA dire a forni da distillar assi vno inuestire in disegno le forme di dui altri fornetti, acciochè fentendoli nominare,& non li hauesti mai ordinati, foppiate qual che fi nomina,ze primo qual di reuerberoe trouato da alchimisti per ritenere qual da molti è detto,Matracre,credo per fottigliezza del luoco che fi dice effer nell'inferno,perche contiene molte humidita'alqu e buon a distillare, fublimare,e (bisognando fondere) a

DEL LAVORAR IN VNA ZECCA 131

DISCORSO ET AVERTENTIE SOPRA IL LAVORAR GIVSTAMENTE, E CON VANTAGGIO VNA ZECCA. CAP. III.

Fig. 5 – Forni per la distillazione con molti alambicchi.

Nella *Fig. 5* vengono illustrati accuratamente forni per la distillazione con gli scarichi dei fumi e soprattutto dotati di molti alambicchi atti a raccogliere all'esterno il distillato.

Si noti nella pagina di destra l'illustrazione di forni a riverbero. L'autore desidera mostrare in disegno le forme di due altri fornetti. Il primo, quello a riverbero, "inventato" da alchimisti e realizzato per contenere molte fiamme. Il secondo permette di distillare, sublimare, e fondere. Ingegnosa e delicata la tecnica del forno a riverbero perché il calore è fornito da un gas ad alta temperatura che lambisce il materiale

in trattamento. Una difficoltà rilevante è connessa con la circostanza che all'epoca le temperature facilmente ottenibili si aggiravano sui novecento gradi centigradi (con stima a vista) generati per lo più da polvere di carbone combusta.

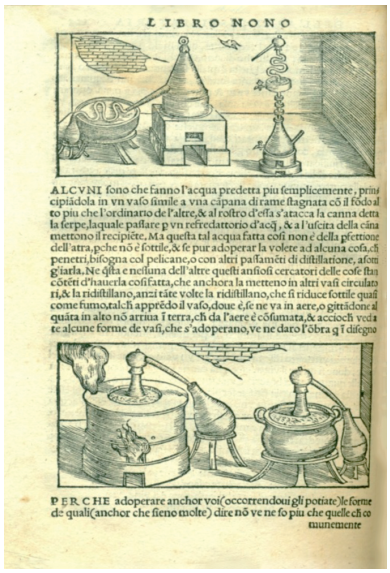
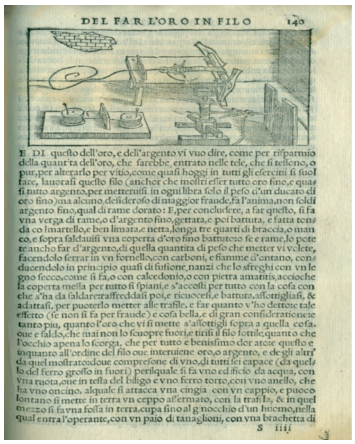


Fig. 6 – Tecniche di raffreddamento del distillato.

Nella *Fig. 6* si illustrano altri forni per la distillazione con particolare riguardo alle tecniche di raffreddamento del distillato prima di essere raccolto nei vasi. Si illustrano serpentine per prolungare il percorso del distillato, e tecniche di stabilizzazione della temperatura di distillazione con dispositivi ad acqua in Bagno-Maria.

In *Fig. 7* è illustrata una raffinata tecnica per "far l'oro in filo". Si illustra la possibilità di realizzare sottili cavi la cui anima sia d'argento o rame e soltanto la parte esterna in oro. Il fine è di ridurre il costo di realizzazione sí che risplendano come auree massicce (come ad esempio nel caso di uniformi militari).

Si realizza facilmente una verga sottile d'argento o di rame. La si lavora a martello e la si rifinisce a lima. Si introduce poi in un fornello con carbone e fiamma d'ontano, portandola sino a quasi fusione per poterla infine trafilare sí da renderla "TANTO FINE CHE QUASI L'OCCHIO NON COGLIE".


DEL FAR L'ORO IN FILO 140

E DI questo dell'oro, e dell'argento vi vuo dire, come per risparmio della quantità dell'oro, che farebbe, entrato nelle tele, che si fellono, o pur, per alterarlo per vitio, come quasi hoggi in tutti gli esercitii si fuol fare, lauorali questo filo (anchor che meglio si esser tutto oro fine, e quasi tanto argento, per metterui si in ogni libra solo il peso d'un ducato di oro fino) una alcuno, desiderolo di maggior fraude, fa l'anima, non solo di argento fino, qual di rame dorato: E per concludere, a far questo, si fa vna verga di rame, o d'argento fino, gettata, e poi battuta, e fatta rendi da co l martello, e ben limata, e netta, longa tre quarti di braccia, o manco, e sopra faldauisi vna coperta d'oro fino battuto: se e rame, lo pone tra anche far d'argento, di quella quantità di peso che metter vi volete, facendolo ferrar in vn forniello, con carboni, e fiamme d'ontano, consducendolo in principio quasi di fusione, nanzi che lo sfreghi con vn legno fecco, come si fa, o con calcedonio, o con pietra amatitis, accioche la coperta messa per tutto si spianti, e s'accosti per tutto con la cosa con che s'ha da saldare: raffredalli poi, e ricuocerli, e battuta, assottigliali, & adattali, per puoterlo metter alle trafile, e far quanto v'ho detto: tale effetto (se non si fa per fraude) e cosa bella, e di gran considerationi: se tanto piu, quanto l'oro che vi si mette s'assottigli sopra a quella cosa, oue e saldo, che mai non lo scuopre fuori: e tirili il filo sottile quanto che l'occhio apena lo scorga, che per tutto è benissimo der atteie questo e tuquanto all'ordine del filo que interuiene oro, o argento, e de gli altri da quel mostrato adoue comprestione di vno, di tutti fei capace (da quello del ferro grosso in fuori) perlquale si fa vno ed ficio da acqua, con vna ruota, one in tella del bulgo è vno ferro torto, con vno anello, che ha vno oncino, alquale si attacca vna cingia, con vn cappio, e pucco bottano si mette in terra vn ceppo affermato, con la trafila, & in quel mezzo si fa vna fossa in terra, cupa sino al ginocchio d'un huomo, nella qual entra l'operante, con vn paio di tanaglioni, con vna brachetta di

S iiij

Fig. 7 – Dispositivo per produrre fili d'oro a partire dal massello.


LIBRO DECIMO

giuato, fisso nella cognitione della desiderata cosa, dalche si semplicissimo intelletto nostro, accecato dal senso, da le medesimo s'inuilisce, tratto dalla vaghezza, o dall'immoderata speranza di posseder cio che cusa nella sensual & lasciua volonta, nutrita dalli molti, & vani giouanili pensieri che da esser (per efaltato) son detti, amore: ch'è in cielo, Volcano & Venere nacinato, che ancho è in terra, cui dalla gentilezza, & natural nobilta dell'animo, con la particolar elemion d'una molta bellezza accompagnatore piu oltre, ch'è molto corrigiato dall'inclinatione delli affetti, in conformita delle cause superienti: & va riconfondo con vn fascio di catene e lacci, che son l'humane grauissimi, & oscuri sguardi, de gratie carezze, virtu, modestia, honesta, & ogn'opra pellicciata di cotal cose solo si pasce e nutrisce dequal cose da lui degnate, le cenuere in viuo fuoco, che tanto piu brucia & consuma, quanto che delle dette cose piu si nue scorge dequal riscono poi in vn inchiostro del fuoco, che non solo nutrende interamente, ma e anche preparabile per: heva moltiplicando nelli petti humani, si come vna pestifera contagione, senza ripentar a sapienta, o fortezza, o a ricchezza, ouer a quella, ne forti anche (come si legge) alli deine mono alli terribili animali: e salda infino li cuori delli frigidi pesci, & acquatici conderlatche (defendendo alli infelici) hebbe ancho maggior forza verso il gran Platone vende confessar bile gra, quello esser vn fuoco dal ciel sceso, che scopra ogni virtute: ch fa influire del modo, che infino nelli arbori (come ne fa fede il nerdeggiante Lauro, che non patorisse senza il vscentro di questo caldo fuoco, & cosi il dottissimo) ma, non potendo così terminatamente compren der la sua forza, dico, che fia vn appetito, che si conserte in esso fuoco, dal mondo (amoroso) chiamato. Io adonque (parlando per speculatione) affermo, che così è fuoco è maggior ricetta di sua natura che non vi n'ho sapputo, fatene contenti (per fede) d'accettare il buon voler mio pero che essendo, come è partenne vn'ample, e capo pelago, entra sicurta di portar donque bisogno gettar il ferro, lasciandouine far giudicio, quando ui troueiate oue farainno le ardenti ferrace li tormenti ch'il martello li mouimenti, circ, le aberrationi, gelosie, timori, & infiniti altri fastidiosi effetti, c'hanno potentia d'accender le continue tormentanti et fiamme (qualche le ha prouate, e proua, render ne puo sicra tello mortanza. E qui fine, altro dire, mettero fine al lungo mio parlare.

IL FINE.

In Vinegia per Giouan Padoano, a istantia di Curtio di Nauò.
M D L.

CVRTIO FABIO

Fig. 8 – L'ultima pagina dell'opera di Vannuccio Biringuccio.

Nella pagina (*Fig. 8*) di chiusura la marca è di estrema raffinatezza grafica con l'indicazione del nome dell'editore CURTIO FABIO. Si noti che la dicitura al piede della pagina precedente recita: IL FINE e la dicitura in Vinegia, per Giouan Padoano, a istantia di Curtio Nauò. Segue una scritta a mano, probabilmente un EX LIBRIS. h cm 21, pag. 169.

Molto chiara la data di edizione MDL. Segue una firma difficilmente interpretabile.

Due opere di impostazione alchimica

Nelle due Figure seguenti (*Figg. 9-10*) si illustrano due opere non propriamente relative ad aspetti riguardabili come prolegomeni della chimica, ma piuttosto da ritenersi contenitori di aspetti molto elementari della chimica. Sono infatti destinati allo studio della farmacopea e costituiscono comunque due reperti interessanti dal punto di vista bibliografico orientato alla scienza.

In *Fig. 9* è presentata un'opera dal titolo ***Examen omnium Syruporum publicus vfus*** [usus] ***est***.

Si tratta di un volume contenente tre contributi di carattere terapeutico. I primi due sono dello stesso autore indicato in figura dedicati a medicamenti. Il primo, di pagine 131+index, è edito in VENEZIA MDXLV e il secondo, di pagine 205, è datato MDXLIX. In [ICCU\BVEE\0011476] si dichiara l'esistenza di 17 esemplari. Preziosa la marca dell'editore.


ANTONII
MVSAE BRASA-
uoli Ferrarienſis, Examen om-
nium Syruporum, quo
rum publicus
vſus eſt.
Omnia ab autore recognita.
GVM INDICE.
Venetijs Apud Cominum de Tridino
Montisferrati, Anno
I. R. ISTITUTO DI
SCIENZE LETTERE ARTI

Fig. 9 – Volume contenente due contributi di ANTONIO MUSA BRASAVOLA ferrarese (1500-1555) ed un contributo di ORIBASIO SARDANI.

Il terzo contributo è di ORIBASIO SARDIANI ed è legato nel volume precedente. Dedicato alla preparazione di medicinali. Edito in VENETIJS MDLVIII reca un ex libris scritto a mano che potrebbe leggersi Ex Zacharia' Caimi, h cm 14,7, pagine 216+note+index.

L'intera opera contiene tracce di attenzione a elementi di studio

analitico che potrebbero indurre ad un vago preludio di sistematica chimica. Si tratta di opera in buono stato di conservazione e di un certo pregio bibliografico.

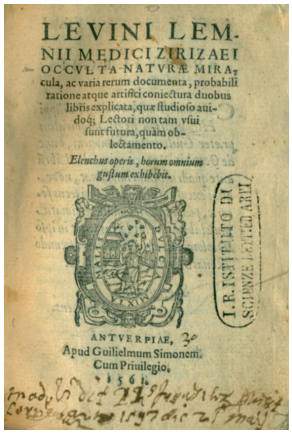

LEVINI LEM-
NII MEDICI ZIRIZAEI
OCCVLTA NATVRÆ MIRA-
cula, ac varia rerum documenta, probabili
ratione atque artifici coniectura duobus
libris explicata, quæ studioso auti-
dog; Lectori non tam visu
sunt futura, quàm ob-
lectamento.
Elenchus operis, horum omnium
iustum exhibebit.
ANTVERPIAE, 3o
Apud Guillelmum Simonem.
Cum Priuilegio.
1561.
I.R. ISTITUTO DI SCIENZE, LETTERE ED ARTI

Fig. 10 – Opera di LEVINO LEMNO (1505-1568).

In *Fig. 10* è presentata l'opera di LEVINO DEMNIO dal titolo ***Occulta Natura Miracula***, pag. 164, in folio h cm 14,5. Rilegatura originale, stato di conservazione mediocre, prima pagina interamente strappata, molte scritte a mano in inchiostro quasi illeggibili, timbro IL del periodo Lombardo-Veneto. In basso alla pagina di frontespizio una scrittura leggibile nella parte terminale con la dicitura *1597 die 21 may*, nella seconda pagina di copertina una lunga scritta di difficile interpretazione. Presenta un elenco numerato degli argomenti trattati, seguito da RERUM ET SENTENTIARUM HUIUS OPERIS INDEX, tratta di malattie fisiche e psichiche senza indicazioni terapeutiche e con un linguaggio ispirato ad interpretazioni misteriose e magiche.

Quest'opera ha avuto una grande fortuna editoriale come risulta dalla lista che segue:
ANTWERPIAE 1561 ed Guillelmum Simonem, in latino, non catalogata ICCU, VENETIA 1560 ed Ludovico Auanzi, in italiano, 14 esemplari in [ICCU\BVEE\004751]
1563 idem IT 16 esemplari in [ICCU\BVEE\004802]
1567 idem IT 12 esemplari in [ICCU\BVEE\004821]
1573 COLONIAE ed Johann Birkmann, in latino, 2 esemplari in IT[ICCU\BVEE\004825]

Un'opera postuma di Paracelso

Dotato di una personalità prepotente TEOPHRASTUS PARACELSO (Einsiendeln 1493- Salisburgo 1541) si è dedicato a vastissimi ambiti culturali. Si professa infatti medico, con qualche lampo di fisiologia, farmacologo nuovissimo perché fu tra i primi ad affermare che anche sostanze minerali, e non soltanto vegetali come sino ad allora si riteneva, potessero proficuamente essere utilizzate per creare medicamenti. Non disdegna di valersi s'una vaga psicologia che appartiene disciplinarmente alla magia, definisce lo zolfo come nuovo elemento. Un aspetto di rilevante interesse risiede nei commenti finali agli Aforismi di Ippocrate. Gli Aforismi non hanno specifico carattere riguardante la chimica, ma interessanti osservazioni riguardanti l'atteggiamento corretto di chi intende dedicarsi a studi scientifici.

Embrioni di chimica vanno cercati piuttosto nei ricettari e negli elenchi di elementi come vedremo illustrati.

Fu coevo, all'università di Ferrara, di Copernico. Fu corrispondente di Marsilio Ficino e Piero della Mirandola.

La concezione che Paracelso ebbe della scienza tradisce una notevole apologia.

Riteneva la scienza fosse "talmente alta" da non poter essere raggiunta con lo studio, come comunemente veniva sostenuto, ma d'essere accessibile solo per rivelazione alla quale si poteva a sua volta giungere con la perfetta conoscenza di Dio. Lo studio doveva quindi essere inteso come un mezzo soltanto necessario per raggiungere la scienza.

Alcuni autori hanno attribuito a Paracelso un concetto globale di Scienza e Religione.

A lui viene attribuita l'espressione:

> *Come si conosce la Scienza nella conoscenza di Dio, così si giunge meglio a Dio nella conoscenza della Scienza.*

Paracelso avversava Aristotele perché ne rifiutava i quattro fondamentali elementi (aria, terra, acqua, fuoco). Nel linguaggio di Paracelso sono presenti elementi squisitamente astrologici e aspirazioni alla futura scienza in particolare chimica. Singolare la pretesa di spiegare una fenomenologia cosmica con termini chimici: certi astri sulforeggiano, altri arsenicano, altri salano, altri mercuriano. Nel mare magnum cioè

nell'atmosfera in tal modo procurano le malattie. Sottolineare la necessità di fornire all'organismo i minerali *(che sono nel macrocosmo e deficienti nel microcosmo)*.

Il vero scopo della chimica non consiste nella preparazione dell'oro, bensì nella preparazione della medicina (Iatrochimica).

In tempi d'assenza di definiti metodi sperimentali, Paracelso predicava la necessità del lavoro di laboratorio e di osservazione, soprattutto in medicina.

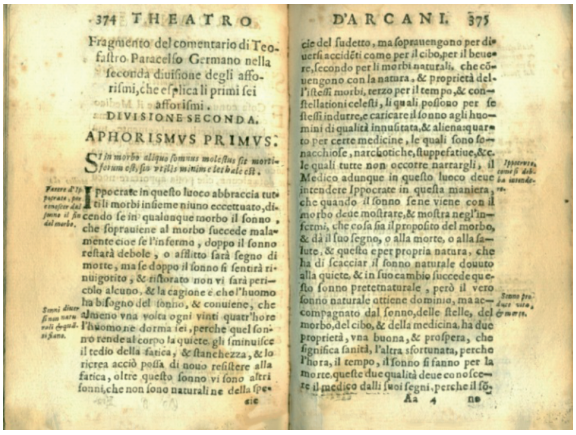
Col nome di chimico indicava chi preparava medicine, estratti e sali. Chi eseguiva ricerca chimica veniva definito *"philosophus per ignem"*.

Una interessante raccolta di **Afforismi** di Ippocrate commentati da Paracelso si trova nell'Opera **Theatro d'Arcani** di Ludovico Locatelli edita a Milano (*Fig. 11*) nel M:DC:XLIV, stato di conservazione mediocre, pag. 436, molte scritte a penna indecifrabili, contributo di Paracelso da pag. 374, Copertina originale , h. cm 17.2.

In *Fig. 12* si presenta la pagina 174 iniziale del testo dovuto a Paracelso. Segue nella *Fig. 13* la lista intitolata Esplicazione de nomi de Filosofi dedicata a nomenclatura di sostanze dovute alle nomenclature dei sapienti, ovvero filosofi, e seguita dalla dicitura volgare. Si rilevi la preoccupazione di divulgare la conoscenza "scientifica" atteggiamento proiettato sul futuro di un autore, come Paracelso decisamente Alchimista.


THEATRO
D'ARCANI
Del Medico
LODOVICO LOCATELLI
Da Bergamo;
NEL QVALE SI TRATTA
Dell'Arte CHIMICA, & suoi
ARCANI,
Con gli Afforismi d'Ippocrate Commentati da Paracelso, &
L'esposition d'alcune Cifre, & Caratteri oscuri de Filosofi,
CON DVE TAVOLE
Vna de Capitoli, & l'altra delle cose più notabili.
CON PRIVILEGIO DELL'ECCELmo SENATO
Di Milano per Anni dodeci futuri.
IN MILANO,
Per Gio. Pietro Ramellati, M. DC. XLIV.
Con licenza de' Superiori.

Fig. 11 – Frontespizio del Teatro d'Arcani.



374 THEATRO

Fragmento del comentario di Teofastro Paracelso Germano nella seconda diuisione degli affo-rismi, che esplica li primi sei affo-rismi.

DIVISIONE SECONDA.

APHORISMVS PRIMVS.

Si in morbo aliqua somnus molestus sit mortiferum est, sin vero minime tum bene est.

Ippocrate in questo luoco abbraccia tutti li morbi insieme niuno eccettuato, dicendo se in qualunque morbo il sonno, che soprauiene al morbo succede malamente cioe se l'infermo, doppo il sonno restarà debole, o afflitto sarà segno di morte, ma se doppo il sonno si sentirà rinuigorito, & ristorato non vi sarà pericolo alcuno. & la cagione è che l'huomo ha bisogno del sonno, & conuiene, che almeno vna volta ogni vinti quattr'hore l'huomo ne dorma sei, perche quel sonno rende al corpo la quiete, gli sminuisce il tedio della fatica, & stanchezza, & lo ricrea acciò possa di nouo resistere alla fatica, oltre questo sonno vi sono altri sonni, che non sono naturali ne della spe-

D'ARCANI. 375

cie del sudetto, ma soprauengono per diuersi accidēti come per il cibo, per il beuere, secondo per li morbi naturali, che cōuengono con la natura, & proprietà dell'istessi morbi, terzo per il tempo, & constellationi celesti, li quali possono per se stessi indurre, & caricare il sonno agli huomini di qualità innusitata, & aliena quarto per certe medicine, le quali sono sonnachiose, narcotiche, fluppefattiue, &c. le quali tutte non occorre narrargli, il Medico adunque in questo luoco deue intendere Ippocrate in questa maniera, che quando il sonno se ne viene con il morbo deue mostrare, & mostra negl'infetti, che cosa sia il proposito del morbo, & dà il suo segno, o alla morte, o alla salute, & questo e per propria natura, che ha di scacciar il sonno naturale douuto alla quiete, & in suo cambio succede questo sonno pretenaturale, però il vero sonno naturale ottiene dominio, ma accompagnato dal sonno, delle stelle, del morbo, del cibo, & della medicina, ha due proprietà, vna buona, & prospera, che significa sanità, l'altra sfortunata, perche l'hora, il tempo, il sonno si fanno per la maggior parte queste due qualità deue conoscere il medico dalli suoi segni, perche il so-

Aa 4 no

Fig. 12 – Inizio del testo di Paracelso.

Nei commenti di Paracelso appare almeno in due circostanze l'espressione CHIMICA.

Infatti, con riferimento al TEATRO D'ARCANI, si legge:

alla riga 20 di pagina 402, *Perché tù possi intendere li libri CHIMICI, che ti capiteranno nelle mani …*
alla riga 9 di pag 403, *… vocaboli ofcuri che , chi non è buon CHIMICO vero figlio dell'arte non è possibile intenderli …*



Esplicatione de nomi de Filosofi.

ATramentum,	vol dire Vitriolo
Alchaich,	argento viuo
Almifadir, & mixadir,	Sale Armeniaco
Arfenico	Orpimento
Aremaros	Cinabrio
Almarcach	Litargirio
Achitram	Oglio di Ginepro

Ce 2 Aga-

D'ARCANI. 403

Afor	argento viuo
Alfur	Zaffarano
Alacob	Sal armoniaco
Alabrot	Vn sale Confetto
Alhenet	Piombo
Alhufel	aquila
argento del popolo	sale amaro
algerie	calcina
altimar	Rame abbrucciato
alafor	alchali
alaphor	canfora
almifa	mosco
alfanet	sale armoniaco
acqua faluatica,	fatta di sale
allume liquido	cioe di amonio
allume fascolo	Palchali
allume crepo	tartaro di bon vino
allume di piuma	il ramino
allume alessandrino	Vetro fallo
allume di cabra	sale alchali
afphaltom	bitume giudaico
armenio	il bolo
amalgra	bolo ciprino
Baurac	ogni sorte di sale
Bolefis	il corallo
Bolo	Pafrea
Croco di marte	cioe del ferro
Croco	orcalo ottone
Calcanto	rame bruciato

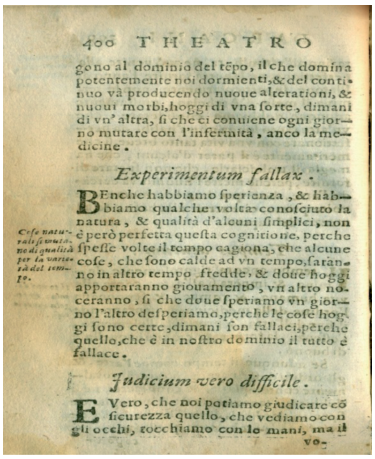
Cc 3 Cal-

406 THEATRO

Calchithor	Marchasita
Croccus hortolanus	Sal Alchali
Cadima	Tutia
Capricorno	Piombo
Calcadis	Sal Alchali
Dragmo	Vitriolo
Denoquor	Borace
Elchalei	Stagno
Epatis	Allue Epatico
Ezimar	Fior di rame
Electum	Oro
Elemo	Orpimento
Febo	Puro Vergine
Iupiter	Stagno
Latte di papauero	Opio
Latro	Argento viuo
Luna	Argento
Mercurio	Argento viuo
Mars	alpalto
Minieta	Ferro
Miſio	Alpalto
Maruch	Ol. Commune
Prafio Verde	Fior di rame
Sidea	Cerusa bianca
Solnere	In humare
Spodio	Cenere d'oro
Vchar	Medicina
Zaraca	Stagno

Fig. 13 – Lista dei nomi dati dai Filosofi e usati nel volgare.

Si tratta di citazioni esplicite della parola CHIMICA per intendere fenomeni di difficile interpretazione alla luce dell'alchimia.



400 THEATRO

gono al dominio del tẽpo, il che domina potentemente noi dormienti, & del continuo và producendo nuoue alterationi, & nuoui morbi, hoggi di vna sorte, dimani di vn' altra, si che ci conuiene ogni giorno mutare con l'infermità, anco la medicine.

Experimentum fallax.

BEnche habbiamo sperianza, & habbiamo qualche volta conosciuto la natura, & qualità d'alcuni simplici, non è però perfetta questa cognitione, perche spesse volte il tempo cagiona, che alcune cose, che sono calde ad vn tempo, saranno in altro tempo fredde, & doue hoggi apportaranno giouamento, vn altro nocceranno, si che doue speriamo vn giorno l'altro desperiamo, perche le cose hoggi sono certe, dimani son fallaci, perche quello, che è in nostro dominio il tutto è fallace.

Cose naturali si mutano di qualità per la scorsa del tempo.

Judicium vero difficile.

EVero, che noi potiamo giudicare cõ sicurezza quello, che vediamo con gli occhi, tocchiamo con le mani, ma il vo-

Fig. 14 – Chiusa dei commenti di Paracelso agli AFFORISMI.

Nella *Fig. 14* è rappresentata una pagina della chiusa dei commenti di Paracelso agli aforismi di Ippocrate con i paragrafi: ***Experimentum fallax*** e ***Judicium vero difficile***. Si presenta una sintesi dei due contributi la cui impostazione epistemologica sembra essere di sorprendente attualità. Il che risulta particolarmente interessante qualora si pensi che si debbono ad un autore la cui attività ha avuto luogo nella prima metà del XVI secolo.

Experimentum fallax

Le cose naturali mutano di qualità con lo scorrere del tempo. … le cose che oggi sono certe domani saranno fallaci, perché quanto è nel nostro dominio è tutto fallace.

Judicium vero difficile

E' vero che possiamo giudicare con sicurezza quello che vediamo con gli occhi e tocchiamo con le mani. Ma il voler giudicare quello che l'uomo immagina non è cosa facile, perché mentre giudichiamo inganniamo noi stessi con un'opinione, e formuliamo giudizi di quanto non abbiamo mai visto né conosciuto avendo investigato con argomenti sofistici e fallaci … [Noi] vediamo con gli occhi solamente il corpo, ma non possiamo vedere lo spirito nel quale è posto il fondamento.

OPERE (XVI-XVII sec.) CONSULTATE

BIRINGUCCIO Vannuccio, (1550) Pirotechnica, Venezia MDL.
BIRINGUCCIO Vannuccio (1559) Pirotechnica, Venezia MDLIX.
LOCATELLI Lodovico (1644) Teatro d'Arcani, Milano MDCXLIV.
DES CARTES Renatus (1662) De Homine Figuris, MDCLXII s.i. h 20.4 cm.
GESNERUM (TEGURINUM) Conradum (s.d.) Historia plantarum (a penna, pag. 281+contributo di Theofrasto) h 14,5 cm, .
LEVINI LEMNII (1561)…Occulta naturae miracula… Testo latino, Anrverpiae, h 24,3 cm.
MUSAE BRASA'UOLI Antonii (1545) SCIROPPI,… Venezia MDXLV, h 14,7 cm.
PARACELSO THEOPHRASTI, COMMENTI ad Aforismi di Ippocrate (in THEATRO D'ARCANI di Locatelli Lodovico). Milano MDCXLIV.
PICCOLPASSO Cipriano (1548) Li Tre Libri dell'arte Del Vasaio, Firenze MDXLVIII.

Istituto Lombardo (Rend. Scienze) 150, 73-116 (2016)

ORDER, DISORDER AND PHASE TRANSITIONS IN QUANTUM MANY BODY SYSTEMS

ALESSANDRO GIULIANI (*)

Nota presentata dal s.c. Antonio Giorgilli
(Adunanza del 5 maggio 2016)

Sunto. – In queste note offro una panoramica di diversi risultati della teoria dei sistemi quantistici a molti corpi, al confine tra la meccanica statistica quantistica matematica e la teoria della materia condensata. In particolare, discuto alcuni risultati recenti sull'universalità dei coefficienti di trasporto di sistemi di elettroni interagenti, e, nello specifico, sull'indipendenza della conduttività Hall dall'interazione elettrone-elettrone. In questo contesto, lo scambio di idee tra matematica e fisica teorica si è dimostrato estremamente fruttuoso e ha aiutato a chiarire il ruolo giocato dalle leggi di conservazione quantistiche (identità di Ward) e dalle proprietà di decadimento delle correlazioni corrente-corrente sull'indipendenza della conduttività dall'interazione tra elettroni.

* * *

Abstract. – In this paper, I give an overview of some selected results in quantum many body theory, lying at the interface between mathematical quantum statistical mechanics and condensed matter theory. In particular, I discuss some recent results on the universality of transport coefficients in lattice models of interacting electrons, with specific focus on the independence of the quantum Hall conductivity from the electron-electron interaction. In this context, the exchange of ideas between mathematical and theoretical physics proved particularly fruitful, and helped in clarifying the role played by quantum conservation laws (Ward Identities), together with the decay properties of the Euclidean current-current correlation functions, on the interaction-independence of the conductivity.

(*) Dipartimento di Matematica e Fisica, Università degli Studi Roma Tre, Roma, Italy.
E-mail: giuliani@mat.uniroma3.it

1. Introduction

The attempt to understand the macroscopic properties of quantum matter starting from first principles is a central challenge in theoretical physics, which dates back to the early days of quantum mechanics. There are several exotic phenomena that are still far from being fully understood, some of which are well known since more than a century: think, for example, to superconductivity and superfluidity, which have first been observed in 1911 [48] and 1937 [3, 39], respectively. Other important quantum phenomena have been discovered or observed only much more recently: think, *e.g.*, to the integer and fractional quantum Hall effect (observed in 1980 [40] and 1982 [53], respectively), to high-T_c superconductivity (discovered in 1986 [13]), or to Bose-Einstein condensation (observed in a gas of cold atoms in 1995 [5, 16, 24]).

In several cases, approximate or heuristic theories are available for understanding the microscopic mechanism behind these exotic phenomena: I refer here, *e.g.*, to the BCS theory of standard superconductivity [10, 11], to the Bogoliubov theory of Bose-Einstein condensation [15], or to the theory of the integer quantum Hall effect [41, 51, 34, 52]. Other phenomena remain less understood, such as high-T_c superconductivity [4] or the fractional quantum Hall effect [42], whose microscopic origin is still debated, and are among the central challenges of the current research in condensed matter theory.

Even in the cases that are better understood, mathematical proofs of the quantum phase transitions of interest are typically missing, which indicates an incomplete understanding of the subject: for instance, in standard superconductivity, there is no systematic way of controlling the corrections beyond mean field, which is the basic approximation that BCS theory is based on. In the context of Bose-Einstein condensates, a full proof of condensation for a homogenous gas in the thermodynamic limit is missing, which is a signal of our poor understanding of the phenomenon of continuous symmetry breaking in quantum many-body systems. In the context of the quantum Hall effect, a proof of the stability of the Hall plateaux in the presence of disorder and electron-electron interactions has still to come, which is a signal of our poor understanding of the interplay between disorder and many-body interactions.

Notwithstanding these limitations, the last years witnessed very important developments in the rigorous understanding of these elusive phenomena, at the interface between mathematical and theoretical physics. By combining the ideas developed in the last decades in the condensed matter com-

munity, including the use of Ward Identities in formal perturbation theory, the use of effective field theories, and the proposal of geometrical indices characterizing the 'quantum topological phases', with sophisticated mathematical tools, such as functional inequalities, localization bounds, constructive field theory and multiscale analysis, we acquired a better understanding of several remarkable phenomena, including, *e.g.*, Bose-Einstein condensation [44], quantum magnetism [23], and the universality of quantum transport coefficients [26, 28, 29, 36], just to mention a few.

In this paper I will review some of the latest developments in the universality theory of quantum transport coefficients, which allowed to clarify certain debated issues connected with the optical conductivity in graphene [28]. The key new technical tool that emerged from the combination of theoretical and mathematical physics ideas, is the implementation of Ward Identities within the constructive scheme ('multiscale fermionic cluster expansion') that is currently able to control the analyticity and decay of correlation functions for the ground state of several two-dimensional interacting electron systems. Note that a formal use of Ward Identities in the effective field theory description of quantum phenomena can easily lead to inconsistent results, particularly as far as the computation of transport coefficients is concerned, cf., *e.g.*, with [37].

In order to make the ideas behind these recent applications as transparent as possible, I will restrict my attention to the study of the universality properties of the Kubo conductivity, and, in particular, of its transverse component (Hall conducitivity) in weakly interacting lattice fermions characterized by a gapped ('massive') reference non-interacting Hamiltonian. For these systems, the construction of the ground state correlation functions and the proof of their analyticity properties is particularly simple and, strictly speaking, does not require a multiscale expansion at all. The extension of the proof to the gapless case, in particular in the case of graphene-like systems, requires the use of a multiscale analysis (constructive fermionic renormalization group), which goes beyond the purpose of this review.

The argument presented here is based on [29], which I will refer to for some technical aspects of the proof. However, compared to [29], the proof presented here has some important simplifications in the proof of analyticity and exponential decay of correlations, as well as in the use of the Schwinger-Dyson equation.

The plan is to first introduce the context, the model and the main results. Next, I will present the proof, first giving an overview of the structure

of the proof, and then explaining in some details the different steps.

2. The quantum Hall effect

Before presenting the main results, let me clarify the context under consideration: the quantum Hall effect is a peculiar electronic transport phenomenon, which is observed in thin conducting, or semi-conducting, materials. By 'thin', here, I mean that the material samples under consideration are two-dimensional, or quasi-two-dimensional. The quantum phenomenon of interest has a classical counterpart, which is important to keep in mind: the classical Hall effect is observed in thin conducting materials subject both to a longitudinal electric field E (along the direction of the current) and to a transverse magnetic field B, as shown in the picture.

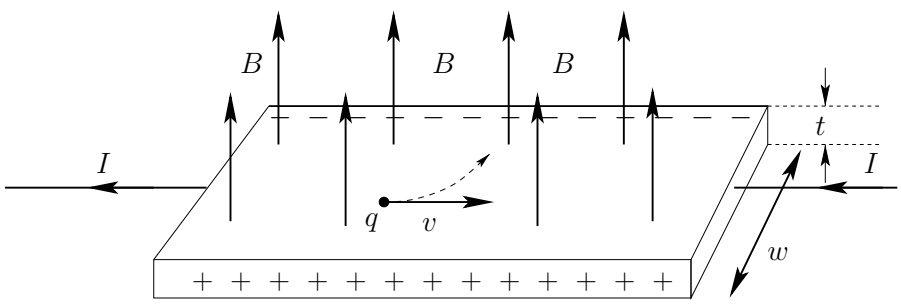


Fig. 1 – A sketch of the setting for the Hall effect: a steady longitudinal current I flows in a thin conducting material, of thickness t and width w, in the presence of a transverse magnetic field B. In the picture, the charge q is negative, and the deflection due to the Lorenz force is schematically represented.

Let t be the film thickness, and w the sample width. As the electrons move in the direction of the current, due to the presence of a transverse magnetic field, they are subject to the Lorentz force $F_L = qv \wedge B$ (here q is the electron charge and v its velocity), which tends to deflect them towards the edges of the sample. In this way, electrons start to accumulate at the edges, and the accumulation goes on until the transverse electric field produced by these electrons is strong enough to compensate the Lorentz force exactly. The equilibrium transverse voltage that is generated in this way is known as the

Hall voltage [33] and is equal to:

$$V_H = vBw.$$

By rewriting the electron velocity as $v = I/(n^{3d}qtw)$, where I is the intensity of the current, and n^{3d} the (three-dimensional) charge carriers density, one gets

$$V_H = \frac{IB}{n^{3d}qt}.$$

Note that V_H is inversely proportional to the film thickness, which explains why the Hall effect is easier to observe in thin samples than in thick ones. For two-dimensional samples, we get the analogous formula: $V_H = IB/(nq)$, where $n = n^{2d}$ is the two-dimensional charge carriers density. Letting $I = jw$, where j is the sheet current density, and defining the Hall field as $E_H = V_H/w$, one finds that the Hall conducitivity has the following expressions:

$$\sigma_H = \frac{j}{E_H} = \frac{nq}{B}.$$

Often, this formula for σ_H is equivalently rewritten in 'natural units' as

$$\sigma_H = \frac{q^2}{h}\nu,$$

where h is Planck's constant, and, if $\Phi_0 = h/q$ is the flux quantum, $\nu = n\Phi_0/B$ is called the 'filling factor'. Written in these terms, the Hall law is a linear relation between the transverse conductivity and the filling factor, with an explicit proportionality coefficient, equal to the 'conductivity quantum' q^2/h.

Experimentally [40] it turns out that at low temperatures and/or large magnetic fields, when the filling factor is of the order of a few units, the system displays a quantization effect: σ_H displays 'plateaus' on which σ_H is constant at a remarkable precision ($\sim 10^{-9}$). This quantization effect is well understood, at a mathematically rigorous level, in the case of non-interacting electron models, including, possibly, disorder, *i.e.*, a random potential [2, 9, 14, 52].

The key observation behind the mathematical theory of the integer quantum Hall effect is the interpretation of the transverse Kubo conductivity as a

geometric index (first Chern number of the Bloch bundle). Unfortunately, the representation breaks down in the presence of electron-electron interactions.

In this review I will give an introduction to the theory of the *interacting* quantum Hall effect, for a class of two-dimensional lattice electron models with weak interactions and no disorder. For more general models, in particular for larger values of the interaction strength, new quantization effects are expected to take place (fractional quantum Hall effect [50]).

3. The models

The models that I will consider are defined as follows. Consider a finite portion Λ of a two-dimensional Bravais lattice generated by two independent vectors $\ell_1, \ell_2 \in \mathbb{R}^2$. I will assume for definiteness that periodic conditions are imposed at the boundary, and that the system is periodic of period L in both directions ℓ_1 and ℓ_2, so that $|\Lambda| = L^2$. At each site x of the lattice I associate a finite number of fermionic creation/annihilation operators, $\psi^{\pm}_{x,\sigma}$, with 'color' σ taking values in a finite index set I. In typical examples, σ can represent a spin index, and/or the position index within the unit cell associated with x.

The 'configuration space' consists of vectors in a fermionic Fock space, labelled by the occupation numbers $\underline{n} = \{n_{x,\sigma}\}_{x\in\Lambda,\sigma\in I}$ of the fermions, with $n_{x,\sigma} = 0, 1$. The canonical anti-commutation rules for the creation/annihilation operators are $\{\psi^+_{x,\sigma}, \psi^+_{y,\sigma'}\} = \{\psi^-_{x,\sigma}, \psi^-_{y,\sigma'}\} = 0$, and

$$\{\psi^+_{x,\sigma}, \psi^-_{y,\sigma'}\} = \delta_{x,y}\delta_{\sigma,\sigma'}.$$

The vector in Fock space representing the state with occupation number $\underline{n}$ is obtained by acting on the fermionic vacuum $|0\rangle$ with $\prod_{x,\sigma}(\psi^+_{x,\sigma})^{n_{x,\sigma}}$, where the product of the different, mutually anti-commuting, operators, must be performed in some prescribed order, fixed once and for all.

The grand-canonical Hamiltonian of the system will be assumed to have the following form:

$$H = H_0 + UV - \mu N, \tag{1}$$

where

$$H_0 = \sum_{x,y\in\Lambda}\sum_{\sigma,\sigma'\in I} \psi^+_{x,\sigma} H^0_{\sigma,\sigma'}(x-y)\psi^-_{y,\sigma'},$$

$$V = \sum_{x,y\in\Lambda} \sum_{\sigma,\sigma'\in I} n_{x,\sigma} v_{\sigma,\sigma'}(x-y) n_{y,\sigma'}, \qquad \text{with} \qquad n_{x,\sigma} = \psi^+_{x,\sigma}\psi^-_{x,\sigma},$$

and

$$N = \sum_{x\in\Lambda} \sum_{\sigma\in I} n_{x,\sigma},$$

In these formulas, H_0 is the 'hopping Hamiltonian', and $H^0_{\sigma\sigma'}(x-y)$ are the 'hopping strengths', assumed to be of finite range and such that $H^0_{\sigma\sigma'}(x) = [H^0_{\sigma'\sigma}(-x)]^*$. The operator V has the interpretation of a density-density interaction, with the coefficients $v_{\sigma\sigma'}(x)$ that are assumed to be real, symmetric (*i.e.*, $v_{\sigma\sigma'}(x) = v_{\sigma'\sigma}(-x)$), and of finite range. Finally, μ is the chemical potential, to be fixed in a spectral gap of the unperturbed Hamiltonian, as discussed in the following.

The spectral properties of H_0 are most conveniently described in terms of the 'Bloch Hamiltonian' $\hat{H}^0(k)$, which is the Fourier transform of the $|I| \times |I|$ hopping matrix $H^0(x)$:

$$\hat{H}^0(k) = \sum_{x\in\Lambda} e^{-ikx} H^0(x),$$

where k is an element of the discretized Brillouin zone,

$$\mathcal{B}_L := \left\{ k \in \mathbb{R}^2 \,:\, k = \frac{n_1}{L} G_1 + \frac{n_2}{L} G_2,\ n_i \in [0, L) \cap \mathbb{Z} \right\},$$

and G_1, G_2 form a basis of the dual lattice Λ^*, *i.e.*, $G_i \cdot \ell_j = 2\pi\delta_{ij}$. I will also denote $\mathcal{B} \equiv \mathcal{B}_\infty$.

If, as assumed, $H^0(x)$ has finite range, then the (thermodynamic limit of the) Bloch Hamiltonian is a Hermitian matrix, depending *analytically* on the momentum k over the Brillouin zone $\mathcal{B}$, with periodic conditions at the boundary of $\mathcal{B}$. The Bloch Hamiltonian, $\sigma(\hat{H}^0(k)) = \{\epsilon_\sigma(k)\}_{\sigma\in I}$ has a real spectrum, and the functions $\epsilon_\sigma(k)$ are known as the 'energy bands'.

I assume the chemical potential to be in a spectral gap:

$$\delta_\mu = \inf_{k\in\mathcal{B}} \operatorname{dist}\big(\mu, \sigma(\hat{H}^0(k))\big) > 0. \tag{2}$$

An important fact is that, if $\delta_\mu > 0$, then the projector $P_-(k)$ over the filled bands, defined as

$$P_-(k) = \sum_{\alpha\,:\,\epsilon_\alpha(k)<\mu} P_\alpha(k),$$

with $P_\alpha(k)$ the projector over the α-th band, is *analytic* in k.

Observables O correspond to self-adjoint operators on the fermionic Fock space, and their expectation value at inverse temperature β is

$$\langle O\rangle_{\beta,L} = \frac{\mathrm{Tr}\, e^{-\beta H} O}{\mathrm{Tr}\, e^{-\beta H}}$$

I denote by $\langle\cdot\rangle$ the thermodynamic and zero temperature limit of $\langle\cdot\rangle_{\beta,L}$:

$$\langle O\rangle = \lim_{\beta\to\infty}\lim_{L\to\infty}\langle O\rangle_{\beta,L}.$$

The goal is to characterize the properties of the interacting, infinite-volume, ground state $\langle\cdot\rangle$, in terms of its 'correlation functions' $\langle O_1\cdots O_n\rangle$, with O_i being fermionic operators depending on a finite number of fermionic fields. I will be particularly interested in current-current correlations, in terms of which one can compute the *Kubo conductivity*, to be defined in the following.

Before that, let me discuss a couple of explicit examples of hopping Hamiltonian H_0, which are characterized by a non-trivial behavior of the Hall conductivity, as the free parameters of the model are properly varied.

3.1 Examples

The Hofstadter model with rational flux. The non-interacting version of this model describes tight-binding electrons, hopping between nearest neighbor sites of a square lattice, in the presence of an external, constant, transverse magnetic field. I let Λ be a square lattice, with $\ell_1 = (1,0)$ and $\ell_2 = (0,1)$, and

$$H_0 = -t\sum_{x\in\Lambda}\sum_{i=1,2}\psi^+_x\psi^-_{x+\ell_i}e^{iq\int_x^{x+\ell_i}A\cdot d\ell} + H.c.,$$

where $H.c.$ indicates the Hermitian conjugate, $t>0$ is the hopping strength, and A is a vector potential corresponding to a constant transverse magnetic field. Choosing, *e.g.*, $A = (0, Bx_1)$, the hopping Hamiltonian takes the form:

$$H_0 = -t\sum_{x\in\Lambda}\Big(\psi^+_x\psi^-_{x+\ell_1} + \psi^+_x\psi^-_{x+\ell_2}e^{iqBx_1}\Big) + H.c.$$

I assume qB to be a rational multiple of 2π, *i.e.*, $qB = 2\pi m/n$, with m, n relative prime integers, and the system size L to be proportional to n. Note that the hopping coefficients are invariant under translations by multiples of $(n, 1)$. Therefore, it is convenient to introduce a unit cell labelled by n colors, $\sigma \in \{1, \dots, n\}$, so that H_0 can be re-written, after a natural re-labeling of the field and site indices, as

$$H_0 = -t \sum_{x\in\Lambda} \Big[\sum_{\sigma=1}^{n-1} \psi^+_{x,\sigma}\psi^-_{x,\sigma+1} + \psi^+_{x,n}\psi^-_{x+\ell_1,1} + \sum_{\sigma=1}^{n} \psi^+_{x,\sigma}\psi^-_{x+\ell_2,\sigma} e^{i2\pi m\sigma/n} \Big].$$

The structure of the energy bands depends in a very peculiar (fractal) way on the value of m/n: more precisely, the value of the transverse Kubo conductivity (to be defined below), if plotted against the magnetic field $qB = 2\pi m/n$ and the chemical potential μ displays a fractal structure, which gives rise to the famous *Hofstadter butterfly* [38]. The band theory of this model is very rich and interesting, see, *e.g.*, [7] for details.

The Haldane model. The non-interacting version of this model describes tight binding electrons, hopping between nearest-neighbor and next-to-nearest-neighbor sites of an hexagonal lattice, in the presence of an external, *dipolar*, transverse magnetic field, as well as of a staggered chemical potential. I think the hexagonal lattice as the union of two translated copies of a triangular lattice Λ, with $\ell_{1,2} = (3/2, \mp\sqrt{3}/2)$. The sites of the two copies can be represented as being white (corresponding to a 'color label' $\sigma = 1$) and black (corresponding to a 'color label' $\sigma = 2$). I think Λ as coinciding with the sub-lattice of white sites, and the black sites as belonging to $\Lambda + (1, 0)$. The creation/annihilation operators associated with the white site located at $x \in \Lambda$ will be denoted by $\psi^\pm_{x,1}$, while the same operators associated with the black site located at $x + (1, 0)$ will be denoted by $\psi^\pm_{x,2}$.

The hopping Hamiltonian of the Haldane model is defined as follows:

$$\begin{aligned} H_0 \;=\; & -t_1 \sum_{x\in\Lambda} \Big[\psi^+_{x,1} \big(\psi^-_{x,2} + \psi^-_{x-\ell_1,2} + \psi^-_{x-\ell_2,2} \big) + H.c. \Big] \\ & -t_2 \sum_{x\in\Lambda} \sum_{j=1}^{3} \sum_{\alpha=\pm} \Big(e^{i\alpha\phi} \psi^+_{x,1}\psi^-_{x+\alpha\gamma_j,1} + e^{-i\alpha\phi} \psi^+_{x,2}\psi^-_{x+\alpha\gamma_j,2} \Big) \\ & +M \sum_{x\in\Lambda} \Big(\psi^+_{x,1}\psi^-_{x,1} - \psi^+_{x,2}\psi^-_{x,2} \Big), \end{aligned}$$

where: $\gamma_1 = \ell_1 - \ell_2$, $\gamma_2 = \ell_2$, $\gamma_3 = -\ell_1$; $t_1 > 0$ and $t_2 > 0$ are the nearest-neighbor and next-to-nearest-neighbor hopping strengths; ϕ is a parameter that measures the strength of the dipolar magnetic field; and M is the amplitude of the staggered chemical potential.

The nice feature of this model, which makes it a very interesting playground, is that the energy bands are very easy to calculate, and so is the Kubo conductivity. Moreover, as we shall see, the transverse conductivity has a non-trivial behavior, as the parameters M and ϕ are varied. In fact, the Bloch Hamiltonian takes the form:

$$\hat H^0(k) = \begin{pmatrix} -2t_2\alpha_1(k)\cos\phi + m(k) & -t_1\Omega^*(k) \\ -t_1\Omega(k) & -2t_2\alpha_1(k)\cos\phi - m(k) \end{pmatrix},$$

where $\alpha_1(k) = \sum_{j=1}^3 \cos k\cdot\gamma_j$, and $m(k) = M - 2t_2\alpha_2(k)\sin\phi$, with $\alpha_2(k) = \sum_{j=1}^3 \sin k\cdot\gamma_j$. Moreover, $\Omega(k) = 1 + e^{-ik\cdot\ell_1} + e^{-ik\cdot\ell_2}$. Note that $|\Omega(k)|$ vanishes if and only if $k = (\frac{2\pi}{3}, \pm\frac{2\pi}{3\sqrt{3}}) \equiv k_F^\pm$: the two points $k_F^\pm$ are called the *Fermi points*. The energy bands associated with the Bloch Hamiltonian are:

$$\epsilon_\pm(k) = -2t_2\alpha_1(k)\cos\phi \pm \sqrt{t_1^2|\Omega(k)|^2 + m^2(k)}.$$

I assume that $t_2/t_1 < 1/3$: in this way, using the fact that $\max_k |\Omega(k)| = |\Omega(0)| = 3$ and that $\alpha_1(k) = |\Omega(k)|^2/2 - 3/2$, one sees that the two bands can touch only if $m(k_F^\pm) = M \pm 3\sqrt{3}t_2\sin\phi = 0$, which is the equation for the *critical lines* of the Haldane model. The graph of the critical, massless, lines in the (ϕ, M) plane is shown in *Fig. 2*

For values of (ϕ, M) in the complement of the two critical lines, the two bands $\epsilon_\pm(k)$ are separated by a gap, so that μ can be chosen in such a way that the gap condition (2) is satisfied. Note that the complement of the critical lines (*i.e.*, the gapped region) is naturally partitioned in four disconnected portions, corresponding to the cases $m(k_F^+) > 0, m(k_F^-) > 0$, or $m(k_F^+) > 0 > m(k_F^-)$, or $m(k_F^+) < 0 < m(k_F^-)$, or $m(k_F^+) < 0, m(k_F^-) < 0$.

4. Current and conductivity

The total (d.c.) current of an electron system is defined as $J = q\hat v$, with $\hat v = i[H, X]$ the velocity of the system: here X is the position operator,

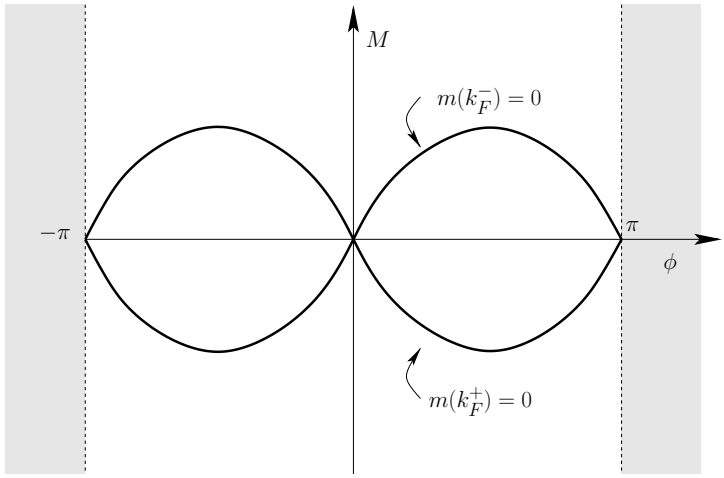


Fig. 2 – The critical lines $m(k_F^{\pm}) = 0$ of the Haldane model in the (ϕ, M) plane.

which takes the form

$$X = \sum_{x\in\Lambda} \sum_{\sigma\in I} n_{x,\sigma}(x + r_\sigma)$$

and r_σ represents the position of the site of color σ, relative to the position x of the cell it belongs to. Using the definitions of $\hat{v}$, H and X, one obtains:

$$J = -iq \sum_{x,y\in\Lambda} \sum_{\sigma,\sigma'\in I} \psi^+_{x,\sigma} H^0_{\sigma,\sigma'}(x - y)\psi^-_{y,\sigma'}(x - y + r_\sigma - r_{\sigma'}).$$

The **Kubo conductivity** is the linear response coefficient of the current to a small external electric field E, which is assumed to be uniform in space and adiabatically switched on in the far past. More precisely, I introduce the time-dependent Hamiltonian $H(t)$, which is obtained from H as follows:

$$H(t) = H - qe^{\eta t} \sum_{\substack{x\in\Lambda \\ \sigma\in I}} n_{x,\sigma} E \cdot (x + r_\sigma), \qquad t \in (-\infty, 0].$$

Here $\eta > 0$ is an adiabatic parameter, to be eventually sent to zero. I assume that, at $t = -\infty$, the density matrix of the system is the equilibrium one,

associated with the Hamiltonian H and the inverse temperature β:

$$\rho(t)\big|_{t=-\infty} = e^{-\beta H} \equiv \rho_{-\infty}.$$

For any $t \le 0$, the density matrix $\rho(t)$ is defined as the solution to the differential equation

$$\dot\rho(t) = -i[H(t), \rho(t)],$$

which gives, at first order in E,

$$\rho(t) = \rho_{-\infty} + iq \int_{-\infty}^{t} e^{\eta\tau} e^{-iH(t-\tau)} [E\cdot X, \rho_0] e^{iH(t-\tau)} d\tau + O(E^2).$$

Therefore, the mean value of the current per unit area at time $t = 0$, at first order in E, in the thermodynamic, zero temperature and adiabatic limits, is:

$$\lim_{\eta\to 0^+} \lim_{\beta\to\infty} \lim_{L\to\infty} \frac{iq}{|\ell_1 \wedge \ell_2|} \frac{1}{L^2} \int_{-\infty}^{0} e^{\eta\tau} \frac{\mathrm{Tr} J e^{iH\tau} [E\cdot X, e^{-\beta H}] e^{-iH\tau}}{\mathrm{Tr} e^{-\beta H}} d\tau.$$

Correspondingly, the linear response coefficient of the i-th component of the current J_i to the j-th component of the electric field E_j is

$$\sigma_{ij} = -\frac{q^2}{|\ell_1 \wedge \ell_2|} \int_{-\infty}^{0} e^{\eta\tau} \big\langle\!\!\big\langle \Big[e^{-iH\tau} [H, X_i] e^{iH\tau}, X_j \Big] \big\rangle\!\!\big\rangle_\infty d\tau,$$

where $\big\langle\!\!\big\langle \cdot \big\rangle\!\!\big\rangle_\infty = \lim_{\beta\to\infty} \lim_{L\to\infty} L^{-2} \langle\cdot\rangle_{\beta,L}$. The matrix of elements σ_{ij} is known as the Kubo conductivity matrix. It is customary to rewrite σ_{ij} in a different, equivalent, form. By using the Leibniz rule for the commutator, one finds

$$\begin{aligned} \big\langle\!\!\big\langle \Big[e^{-iH\tau} [H, X_i] e^{iH\tau}, X_j \Big] \big\rangle\!\!\big\rangle_\infty &= \big\langle\!\!\big\langle [e^{-iH\tau}, X_j]\, [H, X_i] e^{iH\tau} \big\rangle\!\!\big\rangle_\infty \\ &+ \big\langle\!\!\big\langle \Big[[H, X_i], X_j \Big] \big\rangle\!\!\big\rangle_\infty + \big\langle\!\!\big\langle e^{-iH\tau} [H, X_i]\, [e^{iH\tau}, X_j] \big\rangle\!\!\big\rangle_\infty. \end{aligned}$$

Moreover, $[e^{-iH\tau}, X_j]$ can be rewritten as

$$[e^{-iH\tau}, X_j] = -i \int_0^\tau e^{-iHt} [H, X_j] e^{iH(t-\tau)} dt.$$

By using these relations in the expression of σ_{ij}, and recalling the definition of J_i, one gets:

$$\begin{aligned}\sigma_{ij} &= \frac{1}{|\ell_1 \wedge \ell_2|} \lim_{\eta\to 0^+} \frac{1}{\eta} \times \\ &\times \Big\{ -i \int_{-\infty}^{0} e^{\eta t} \langle [J_i, J_j(t)] \rangle_\infty - q^2 \langle \Big[[H, X_i], X_j \Big] \rangle_\infty \Big\},\end{aligned} \tag{3}$$

where $J_j(t) = e^{iHt} J_j e^{-iHt}$. This is one of the standard expressions for the Kubo conductivity. The second term in braces is known as the diamagnetic, or Schwinger, term.

Note that σ_{ij} is a function of the parameters entering the Hamiltonian H of the system, in particular of the strength U of the electron-electron interaction. The main result reviewed in these notes is the following.

Theorem 1 [29]. *In the context above, let $\delta_\mu > 0$. There exists $U_0 > 0$ such that, if $|U| \leq U_0$, then*

$$\sigma_{ij} = \sigma_{ij}\Big|_{U=0},$$

that is, if the chemical potential is chosen so that the gap condition is verified, then the Kubo conductivity is independent of the strength of the interaction U.

Remarks.

1) Under the gap condition, the non-interacting Kubo conductivity can be re-written as q^2 times the first Chern number of the Bloch bundle, *i.e.*, of the vector bundle associated with the linear space $\mathrm{Ran} P_-(k)$, with $k \in \mathcal{B}$:

$$\sigma_{ij}\Big|_{U=0} = -iq^2 \int_{\mathcal{B}} \mathrm{Tr}\, P_-(k) [\partial_{k_i} P_-(k), \partial_{k_j} P_-(k)] \frac{dk}{(2\pi)^2}, \tag{4}$$

which is known to be proportional to an integer [9, 52]. More precisely, $\sigma_{11}\big|_{U=0} = \sigma_{22}\big|_{U=0} = 0$, while $\sigma_{12}\big|_{U=0} = -\sigma_{21}\big|_{U=0} \in \frac{q^2}{2\pi}\mathbb{Z}$. The off-diagonal Kubo coefficient $\sigma_{12}\big|_{U=0}$ is the non-interacting *Hall conductivity*, and the representation in terms of a Chern number shows that it is quantized. Theorem 1 shows that, under the gap condition and at weak enough coupling, the interacting Hall conductivity is also quantized, at the very same

value as the reference non-interacting system. A sketch of the proof of the formula (4) is given in Appendix A.

2) In order to prove that the Hall conductivity is non-trivial (*i.e.*, different from zero), one needs to perform an explicit computation that, of course, is model-dependent. The two examples mentioned above, *i.e.*, the Hofstadter and the Haldane models, are among the simplest examples of gapped systems displaying a non-trivial Hall conductivity. The computation of the value of the Hall conductivity for the **Hofstadter model** as a function of the magnetic flux is a very non-trivial and interesting exercise: the computation is reduced to the study of a Diophantine equation, which can be solved numerically. The resulting 'topological phase diagram' (*i.e.*, the plot of the value of the Hall conductivity as the magnetic flux and the chemical potential are varied) leads to the well known 'Hofstadter butterfly', whose construction goes beyond the purpose of this review [7, 38, 52]. The computation of the Hall conductivity for the **Haldane model** is much simpler: in fact, the evaluation of $\sigma_{12}\big|_{U=0}$ based on (4) is recommended as a very instructive exercise. The interested reader can consult, *e.g.*, [29, Appendix B]. The resulting topological phase diagram is shown in *Fig. 3*. It first appeared in [32].

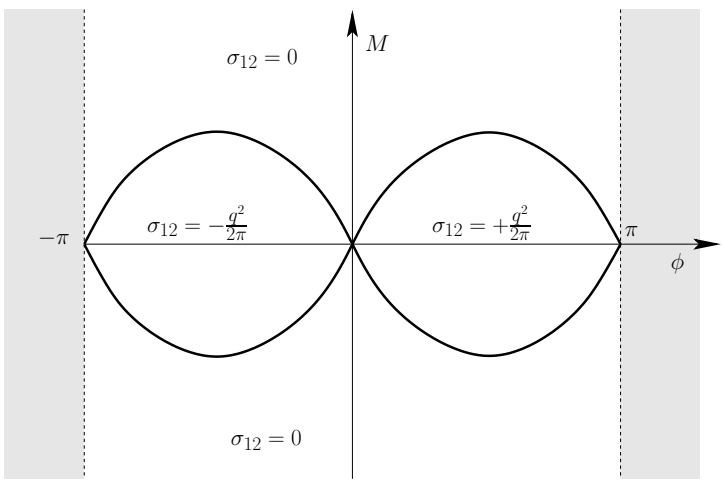


Fig. 3 – The 'topological phase diagram' of the Haldane model, where the different values of σ_{12} in the four regions $\{m(k_F^+) > 0, m(k_F^-) > 0\}$, $\{m(k_F^+) > 0 > m(k_F^-)\}$, $\{m(k_F^+) < 0 < m(k_F^-)\}$, and $\{m(k_F^+) < 0, m(k_F^-) < 0\}$ are indicated.

3) The range of interaction strengths for which Theorem 1 holds, $|U| \leq U_0$, is, in general, non uniform in the spectral gap δ_μ. However, in specific models, it is possible to obtain estimates uniform in the gap: *e.g.*, for the interacting Haldane model, it was proved in [26, 30] that the interacting system is gapped (*i.e.*, the Euclidean correlation functions decay exponentially) for *all* the values of (ϕ, M) outside a pair of *renormalized critical lines*, provided μ is chosen appropriately. The shape of the renormalized critical lines is qualitatively the same as that in *Fig. 3* and the values of σ_{12} inside or outside the curves are the same as in the non-interacting setting. The proof in [26, 30] is valid arbitrarily close to the renormalized critical lines, that is, uniformly in the gap (which vanishes as the critical lines are approached). The generalization of Theorem 1 to this setting requires an infrared multiscale analysis, reminiscent of the one developed in [27] for a model of graphene with short range interactions.

4) The methods reviewed here can be used to show that, under the gap condition, for U sufficiently small, the interacting Hamiltonian has a spectral gap above the unique ground state, uniformly in the volume L. For a recent independent proof of this fact, see [35]. Knowing this fact, one can show that the interacting Hall conductivity σ_{12} has an interpretation in terms of a geometrical index [8, 36], analogous to the representation in terms of the first Chern number of the Bloch bundle that is valid for the non-interacting one. Note, however, that such a representation involves a many-body (rather than a one-body) projector, that is the projector Π on the interacting, many-body, ground state. The control parameter one needs to average over is not the quasi-momentum k (which is not a priori well-defined in the many-body case), as in (4), but, rather, an angle θ controlling the (twisted) boundary conditions. The geometrical representation underlying the results in [8, 36] is not suitable for direct analytical computations: therefore, while the results in [8, 36] show that the Hall conductivity is quantized, they do not provide one with a good representation formula for computing the explicit value of the Hall conductivity in given models. On the contrary, the approach reviewed in this notes provides a very simple explicit formula for such a value.

5. Overview of the proof

Let me now start to discuss the main strategy to be followed for proving the main theorem. In order to compute the Kubo conductivity (3), the key point

is to control the current-current correlations, which enter the first term in the right side of (3). Actually, the direct control of the real-time interacting correlations $\langle\!\langle [J_i, J_j(t)] \rangle\!\rangle_\infty$ is very difficult. In order to go around the obstacle, one can take advantage of the powerful methods (applicable both in the massive and in the massless cases), based on ideas and techniques of *constructive field theory*, which can be used to control and compute the *Euclidean, imaginary-time*, current-current correlations. These methods will be briefly reviewed in the following. Before discussing them, let me explain how to use informations on the Euclidean correlations in order to infer informations on the Kubo conductivity.

5.1 Euclidean correlations

Let $O^{(1)}, \dots, O^{(n)}$ be observables, each of which is assumed to be a linear combination of normal-ordered, even, monomials in the fermionic creation/annihilation operators. I let

$$O^{(i)}_t \equiv O^{(i)}(-it) = e^{tH} O^{(i)} e^{-tH}$$

be the imaginary-time evolution of the i-th observable, with $0 \le t < \beta$, and

$$\langle T\, O^{(1)}_{t_1} \cdots O^{(n)}_{t_n} \rangle_{\beta,L} = \frac{\mathrm{Tr}\, e^{-\beta H} O^{(1)}_{t_1} \cdots O^{(n)}_{t_n}}{\mathrm{Tr}\, e^{-\beta H}}$$

where T is the (multilinear) time-ordering operator, acting on monomials in the fields $\psi^\pm_{(t,x)\sigma} = e^{tH} \psi^\pm_{x,\sigma} e^{-tH}$ as:

$$T\, \psi^{\epsilon_1}_{(t_1,x_1)\sigma_1} \cdots \psi^{\epsilon_n}_{(t_n,x_n)\sigma_n} = (-1)^\pi \psi^{\epsilon_{\pi_1}}_{(t_{\pi_1},x_{\pi_1})\sigma_{\pi_1}} \cdots \psi^{\epsilon_{\pi_n}}_{(t_{\pi_n},x_{\pi_n})\sigma_{\pi_n}}$$

where π is the permutation of $(1, \dots, n)$ such that $t_{\pi_1} > \cdots > t_{\pi_n}$ (in case of coinciding times, the variables at equal time will be assumed to be normal ordered, after the action of the ordering operator T). I also indicate by $\langle T\, O^{(1)}_{t_1}; \cdots; O^{(n)}_{t_n} \rangle_{\beta,L}$ the corresponding cumulants, or truncated correlations, *i.e.*, $\langle T\, O^{(1)}_{t_1}; O^{(2)}_{t_2} \rangle_{\beta,L} = \langle T\, O^{(1)}_{t_1} O^{(2)}_{t_2} \rangle_{\beta,L} - \langle O^{(1)} \rangle_{\beta,L} \langle O^{(2)} \rangle_{\beta,L}$, etc. Moreover, at finite β, I introduce the notion of Fourier transform with respect to the Euclidean time:

$$\hat O_\omega = \int_0^\beta dt\, e^{-i\omega t} O_t,$$

with $\omega \in \frac{2\pi}{\beta}\mathbb{Z}$ the Matsubara frequency. Note that the following momentum conservation rule holds:

$$\int_0^\beta dt_1 \cdots \int_0^\beta dt_n \langle T\, O^{(1)}_{t_1}; \cdots ; O^{(n)}_{t_n} \rangle_{\beta,L}\, e^{-i\omega_1 t_1 \cdots - i\omega_n t_n} = \\ = \delta_{\omega_1+\cdots+\omega_n,0} \langle T\, \hat{O}^{(1)}_{\omega_1}; \cdots ; \hat{O}^{(n)}_{\omega_n} \rangle_{\beta,L}.$$

A natural question arises: why are these definitions interesting and useful? The main reason is that the Euclidean correlation functions appear naturally in the perturbation theory for the equilibrium correlations, via the use of Trotter's product formula and the Duhamel expansion. In many cases, there are powerful methods (fermionic Renormalization Group) for controlling the interacting equilibrium correlation functions at weak enough coupling, and to derive detailed informations about their analyticity properties.

Moreover, typically, once one knows how to control the equilibrium correlations, one also knows how to control the interacting, Euclidean-time, correlations, including control of their analyticity properties in the complex time plane. Finally, if one can prove that the complex time plane is free of singularities of the Euclidean correlation, one can a posteriori reconstruct the (integral of the) real-time ones via a 'Wick rotation' in the complex time plane. In order to give an idea of what I am referring to, let me consider our main object of interest, *i.e.*, the Kubo conductivity eq.(3), and let me focus on the first term in the right side, $\frac{-i}{\eta}\int_{-\infty}^0 dt\, e^{\eta t} \big\langle [J_i, J_j(t)] \big\rangle_\infty$: this can be thought of as the sum of two terms: (i) $\frac{-i}{\eta}\int_{-\infty}^0 dt\, e^{\eta t} \big\langle J_i J_j(t) \big\rangle_\infty$, and (ii) $\frac{i}{\eta}\int_{-\infty}^0 dt\, e^{\eta t} \big\langle J_j(t) J_i \big\rangle_\infty$. Now, if the two terms are free of singularities in the second and third quadrants of the complex plane, respectively, then the two integrals can be rotated as shown in *Fig. 4*, provided the contributions from the integrals over the two quarter circles at infinity give vanishing contribution.

After having performed the rotation, the sum of the two terms can be rewritten as $-\frac{1}{\eta}\hat{K}_{ij}(-\eta)$, where

$$\hat{K}_{ij}(\omega) = \lim_{\beta\to\infty} \lim_{L\to\infty} \frac{1}{\beta L^2} \langle T\, \hat{J}_{i,\omega}; \hat{J}_{j,-\omega} \rangle_{\beta,L}. \tag{5}$$

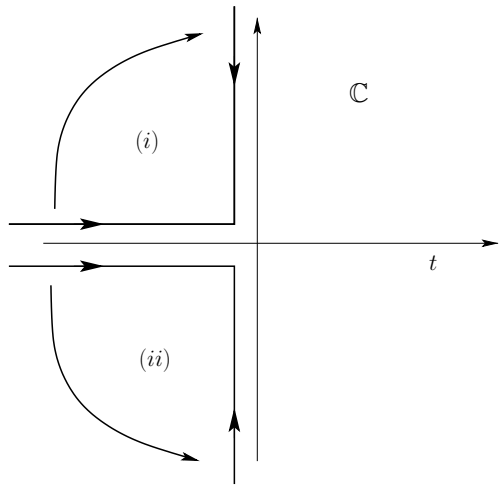


Fig. 4 – The 'complex rotations' of the integration paths that we perform on the two contributions (i) and (ii) to the Kubo conductivity.

Moreover, one can check that the diamagnetic contribution to the Kubo conductivity can be rewritten as $\hat K_{ij}(0)/\eta$. As a result, the Kubo conductivity is equal to the Euclidean one: $\sigma_{ij} = \bar\sigma_{ij}$, with

$$\bar\sigma_{ij} = \frac{1}{|\ell_1 \wedge \ell_2|} \lim_{\omega\to 0^-} \frac{\hat K_{ij}(\omega) - \hat K_{ij}(0)}{\omega}. \tag{6}$$

At this point, I am finally in the position of describing the general strategy of the proof of the main theorem:

1) One first discusses the existence of the thermodynamic and zero temperature limit of the Euclidean correlations, and proves their analyticity in U and in ω in suitable regions of the complex plane. Using the analyticity of the Euclidean correlations in the complex t-plane, as well as the existence of the real-time correlations in the $\beta, L \to \infty$ limit (which follows from the Lieb-Robinson bounds [43]), one proves that the 'Wick rotation strategy' sketched above is rigorously justified, thus showing that the Kubo conductivity is equal to its Euclidean counterpart.

2) In order to show that the interacting Euclidean Kubo conductivity equals the non-interacting one, one combines two classes of remarkable cancella-

tions/identities satisfied by the Euclidean correlation functions, known as Ward Identities (more precisely, the Ward Identities associated with the continuity equation for the current) and Schwinger-Dyson equations. *Note*: the idea of using these two classes of identities to prove the vanishing of the interaction corrections to the Euclidean conductivity is due to [22], whose strategy of proof is adapted here to the case of the Hall conductivity of interacting lattice fermions.

In the following sections, I discuss these two steps in some detail.

6. Equilibrium perturbation theory

Let me start, for simplicity, by discussing the perturbation theory for the partition function of the system at finite β, L:

$$\mathrm{Tr}\, e^{-\beta H} = \mathrm{Tr}\, e^{-\beta(H_0 - \mu N + UV)}.$$

Since the fermionic operators are all bounded at finite L, one can rewrite this trace by using the Duhamel's expansion (see, *e.g.*, [31, Section 3.2]), thus getting:

$$\begin{aligned} \mathrm{Tr} e^{-\beta H} &= \mathrm{Tr} e^{-\beta(H_0-\mu N)} \times \\ &\quad \Big[1 + \sum_{n\geq 1} (-U)^n \int_{\beta > t_1 > \cdots > t_n > 0} dt_1 \cdots dt_n \, \langle V_{t_1} \cdots V_{t_n} \rangle^0_{\beta,L}\Big] \\ &= \mathrm{Tr} e^{-\beta(H_0-\mu N)} \times \\ &\quad \Big[1 + \sum_{n\geq 1} \frac{(-U)^n}{n!} \int_0^\beta dt_1 \cdots \int_0^\beta dt_n \langle T\, V_{t_1} \cdots V_{t_n} \rangle^0_{\beta,L}\Big], \end{aligned} \tag{7}$$

where $V_t = e^{t(H_0-\mu N)} V e^{-t(H_0-\mu N)}$ and $\langle\cdot\rangle^0_{\beta,L}$ is the finite volume, finite temperature, Gibbs average with respect to the Hamiltonian $H_0 - \mu N$.

At finite volume, the series in U for $Z(U) := \mathrm{Tr} e^{-\beta H}$ is convergent (even more: the function $Z(U)$ is *entire* in U, for any finite L). The series for the logarithm of $Z(U)$ is a priori convergent in a very small neighborhood of the origin. In fact,

$$Z(U) \geq 1 - \sum_{n\geq 1} \frac{1}{n!} (\beta|U|)^n \|V\|^n_\infty = 2 - e^{\beta|U|\,\|V\|_\infty},$$

which does not vanish if $U \in B := \{U \in \mathbb{C} : \ |U| < \frac{\log 2}{\beta \|V\|_\infty}\}$. Therefore, the series of $\log Z(U)$ is convergent in any compact set K belonging to the open ball B. Note that typically $\|V\|_\infty$ is proportional to L^2, so that the radius of B goes to zero as $L \to \infty$. The goal of the following discussion is to show that the analyticity region of $\log Z(U)$ can be extended to a ball centered at the origin, with radius *independent* of L.

Before I proceed with the discussion, let me note that similar considerations are valid for the series in U for the Euclidean correlations: if $\beta > t_1 > t_2 > 0$, the Euclidean correlation $\langle O^{(1)}_{t_1} O^{(2)}_{t_2} \rangle_{\beta,L}$ is equal to the ratio between $\mathrm{Tr} e^{-\beta H} e^{t_1 H} O^{(1)} e^{-(t_1-t_2)H} O^{(2)} e^{-t_2 H}$ and $Z(U)$. The denominator is expanded as described above, and the numerator admits a similar expansion:

$$\langle O^{(1)}_{t_1} O^{(2)}_{t_2} \rangle_{\beta,L} = \langle O^{(1)}_{t_1} O^{(2)}_{t_2} \rangle^0_{\beta,L} + \sum_{n\geq 1} \frac{(-U)^n}{n!} \int_0^\beta ds_1 \cdots \int_0^\beta ds_n \langle T\, O^{(1)}_{t_1} O^{(2)}_{t_2} V_{s_1} \cdots V_{s_n} \rangle_{\beta,L}.$$

Therefore, $\langle T\, O^{(1)}_{t_1} O^{(2)}_{t_2} \rangle_{\beta,L}$ is the ratio of two entire functions, the denominator being different from zero in the analyticity domain of $\log Z(U)$.

Now, how do we evaluate $\log Z(U)$, and how do we prove bounds on its analyticity domain? The starting point is an explicit representation of the n-th Taylor coefficient of $Z(U)$: from (7) this coefficient is $(n!)^{-1}$ times the integral over $t_1 \cdots t_n$ of $\langle T\, V_{t_1} \cdots V_{t_n} \rangle^0_{\beta,L}$. Such an average can be computed via the fermionic Wick's rule, which is readily explained in terms of the following graphical representation. Associate with every

$$V_{t_i} = \sum_{x_i, y_i, \sigma_i, \sigma'_i} \psi^+_{(t_i,x_i)\sigma_i} \psi^-_{(t_i,x_i)\sigma_i} v_{\sigma_i\sigma'_i}(x_i - y_i) \psi^+_{(t_i,y_i)\sigma'_i} \psi^-_{(t_i,y_i)\sigma'_i}$$

the 'four-legged' vertex in *Fig. 5*.

Draw n such vertices and pair their 'external legs' in all possible ways , in such a way that the orientations of the paired lines are compatible two by two. The graph associated with any such pairing is called a 'Feynman diagram', and $\langle T\, V_{t_1} \cdots V_{t_n} \rangle^0_{\beta,L}$ can be expressed as the sum over all the possible Feynman diagrams of their values, which are defined as follows. For an example of a Feynman diagram of order $n = 3$, see *Fig. 6*.

First of all, every Feynman diagram $\mathcal{G}$ comes with a permutation $\pi = \pi(\mathcal{G})$, associated with the corresponding pairing of fermionic fields, which is

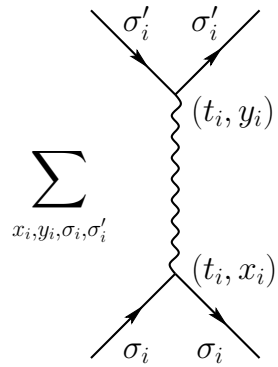


Fig. 5 – The graphical representation of V_{t_i} as a 'four-legged' vertex. The wiggly line represents the interaction $v_{\sigma_i\sigma_i'}(x_i - y_i)$, while the oriented solid lines (the 'external legs') represent the operators $\psi^{\pm}_{(t_i,x_i)\sigma_i}$ and $\psi^{\pm}_{(t_i,y_i)\sigma_i'}$, with the convention that the entering (resp. exiting) lines represent the annihilation (resp. creation) operators ψ^- (resp. ψ^+).

the one needed to move every creation operator $\psi^+_{(t',x')\sigma'}$ immediately to the right of the annihilation operator $\psi^-_{(t,x)\sigma}$, which it is connected with. Moreover, every contracted pair $(\psi^-_{(t,x)\sigma}, \psi^+_{(t',x')\sigma'})$ comes with the value

$$g^{\beta,L}_{\sigma\sigma'}(x-x', t-t') = \langle T\, \psi^-_{(t,x)\sigma}\psi^+_{(t',x')\sigma'}\rangle^0_{\beta,L} = \frac{1}{L^2}\sum_{k\in\mathcal{B}_L} e^{-ik(x-x')} \times$$

$$\times\Big\{e^{-(\hat{H}^0(k)-\mu)(t-t')}\Big[\frac{\mathbb{1}_{t>t'}}{1+e^{-\beta(\hat{H}^0-\mu)}} - \frac{\mathbb{1}_{t\le t'}}{1+e^{\beta(\hat{H}^0-\mu)}}\Big]\Big\}_{\sigma\sigma'},$$

which is called the 'propagator'. Note that its $\beta, L \to \infty$ limit is

$$g_{\sigma\sigma'}(x,t) = \int_{\mathcal{B}}\frac{dk}{|\mathcal{B}|}e^{-ikx}\{e^{-(\hat{H}^0(k)-\mu)t}[\mathbb{1}_{t>t'}P_+(k) - \mathbb{1}_{t\le t'}P_-(k)]\}_{\sigma\sigma'}.$$

In the following, I will use the shorthand notation g_ℓ for the propagator associated with the pair $\ell = \big(((t,x),\sigma),((t',x'),\sigma')\big)$. I also denote by $G(\underline{x}, \underline{y}, \underline{t}, \underline{\sigma}, \underline{\sigma}')$ the set of Feynman diagrams obtained from the contraction of the quartic monomials labeled by the time/space/color variables $\big\{\big(t_i,(x_i,y_i),(\sigma_i,\sigma_i')\big)\big\}_{i=1,\dots,n}$ (here I used the symbol $\underline{x}$ to denote the n-ple $(x_1,\dots,x_n)$, etc.)

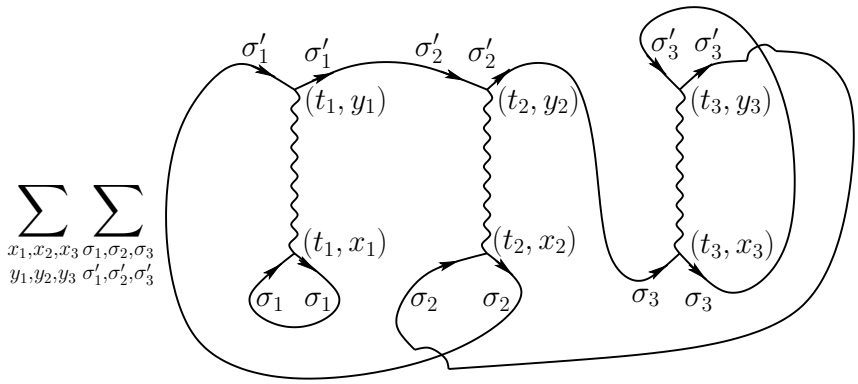


Fig. 6 – An example of a Feynman diagram of order $n = 3$. The solid lines are obtained by the pairing, or contraction, of two of the external legs of the vertices under consideration. Any pair of contracted legs must have compatible orientation, i.e., an exiting external leg can only be contracted with an entering one. Each solid line obtained by the contraction of two legs corresponds to a propagator, with indices readable from the labels of the two contracted legs.

Given these definitions, the value $\mathrm{Val}(\mathcal{G})$ of the Feynman diagram $\mathcal{G} \in G(\underline{x}, \underline{y}, \underline{t}, \underline{\sigma}, \underline{\sigma}')$ is

$$\mathrm{Val}(\mathcal{G}) = (-1)^{\pi} \prod_{i=1}^{n} v_{\sigma_i \sigma'_i}(x_i - y_i) \prod_{\ell \in \mathcal{G}} g_\ell \tag{8}$$

and

$$\langle T\, V_{t_1} \cdots V_{t_n} \rangle^0_{\beta,L} = \sum_{\substack{\underline{x},\underline{y} \\ \underline{\sigma},\underline{\sigma}'}} \sum_{\mathcal{G} \in G(\underline{x},\underline{y},\underline{t},\underline{\sigma},\underline{\sigma}')} \mathrm{Val}(\mathcal{G}),$$

so that

$$Z(U) = 1 + \sum_{n \geq 1} \frac{(-U)^n}{n!} \int_0^\beta dt_1 \cdots \int_0^\beta dt_n \sum_{\substack{\underline{x},\underline{y} \\ \underline{\sigma},\underline{\sigma}'}} \sum_{\mathcal{G} \in G(\underline{x},\underline{y},\underline{t},\underline{\sigma},\underline{\sigma}')} \mathrm{Val}(\mathcal{G}).$$

A remarkable combinatorial theorem, known as the linked cluster theorem [1, 47], shows that the logarithm of the partition function admits a very similar expansion, with the important difference that the Feynman diagrams contributing to $\log Z(U)$ are all and only the *connected* diagrams in

$G(\underline{x}, \underline{y}, \underline{t}, \underline{\sigma}, \underline{\sigma}')$. A similar discussion is valid for the correlation functions: in particular, truncated correlations are expressed as sums of values of connected Feynman diagrams, with vertices corresponding both to the quartic interactions introduced above and to extra vertices corresponding to the insertion of the operators one is computing the correlations of. In conclusion,

$$\log Z(U) = \sum_{n \geq 1} \frac{(-U)^n}{n!} \int_0^\beta dt_1 \cdots \int_0^\beta dt_n \sum_{\substack{\underline{x}, \underline{y} \\ \underline{\sigma}, \underline{\sigma}'}} \sum_{\mathcal{G} \in G_{conn}(\underline{x}, \underline{y}, \underline{t}, \underline{\sigma}, \underline{\sigma}')} \mathrm{Val}(\mathcal{G}),$$

where $G_{conn}(\underline{x}, \underline{y}, \underline{t}, \underline{\sigma}, \underline{\sigma}')$is the set of connected diagrams in $G(\underline{x}, \underline{y}, \underline{t}, \underline{\sigma}, \underline{\sigma}')$.

Now, how do we estimate the n-th order contribution to this series? The basic object one needs to evaluate is the integral over the time, space and color variables of a generic connected Feynman diagram with n quartic vertices, *i.e.*, $\int \sum \left|\mathrm{Val}(\mathcal{G})\right|$, where the integral is over the time variables, and the sum over the space and color indices. Recalling that $\left|\mathrm{Val}(\mathcal{G})\right| = \prod_i \left|v_{\sigma_i \sigma_i'}(x_i - y_i)\right| \prod_{\ell \in \mathcal{G}} |g_\ell|$, one can evaluate the desired integral by selecting a minimal connecting subset of the set of wiggly and solid lines, which the Feynman consists of (this minimal connecting subset is by construction a tree, to be called a 'spanning tree'). I choose the spanning tree in such a way that it contains all the wiggly lines of the diagram, and a suitably chosen set of $n-1$ solid lines. I bound the propagators on the solid lines outside the spanning tree by their L_∞ norm; once this is done, the integration over the space and time variables of the product of the interactions and propagators on the spanning tree produces the product of their L_1 norms, so that, recalling that $|I|$ is the cardinality of the set of allowed colors,

$$\int \sum \left|\mathrm{Val}(\mathcal{G})\right| \leq \beta L^2 \|v\|_1^n |I|^{2n} \|g\|_1^{n-1} \|g\|_\infty^{n+1},$$

where

$$\|g\|_\infty = \sup_{\sigma, \sigma'} \sup_{x,t} \left|g_{\sigma\sigma'}^{\beta, L}(x,t)\right|, \quad \|g\|_1 = \sup_{\sigma, \sigma'} \int dt \sum_x |g_{\sigma\sigma'}^{\beta, L}(x,t)|,$$

and similarly for $\|v\|_1$. Recall that v is of finite range, so its L_1 norm is finite. Moreover, under the gap condition, the propagator decays exponentially in space-time, uniformly in β, L: therefore, both its L_1 and L_∞ norms are finite,

uniformly in the temperature and volume. In conclusion, the n-th order term in the series of $\frac{1}{\beta L^2}\log Z(U)$ is bounded by

$$\frac{C^n}{n!}|U|^n\{\#\text{ Feynman diagrams of order } n\},$$

for a suitable constant C. Now, the number of Feynman diagrams of order n equals $(4n-1)!!$, which grows like $(\text{const.})^n(n!)^2$ as[1] $n\to\infty$ and, therefore, the bound that we just derived is not summable over n. In other words, *the series of Feynman diagrams is not absolutely convergent*.

Remarkably, it is possible to properly resum the series of connected Feynman diagrams, by carefully taking into account the $(-1)^\pi$ signs in (8), in such a way that the resummed series is convergent. The basic observation is that

$$\langle T\,\psi^-_{(t_1,x_1)\sigma_1}\psi^+_{(t'_1,y_1)\sigma'_1}\cdots\psi^-_{(t_n,x_n)\sigma_n}\psi^+_{(t'_n,y_n)\sigma'_n}\rangle^0_{\beta,L}=\det\big[g_{\sigma_i,\sigma'_j}(x_i-y_j,t_i-t'_j)\big]_{i,j=1,\dots,n}.$$

By expanding the determinant, one obtains $n!$ terms, each of which corresponds to one specific pairing of the fermionic fields (i.e, to one 'Feynman diagram'); each of these terms is bounded by $\|g\|^n_\infty$, so this expression can be bounded by $\|g\|^n_\infty n!$. However, one can do much better than this: if instead of expanding the determinant, one recalls that it is equal to the product of its eigenvalues, one sees that its absolute value typically scales like $(\text{const.})^n$, which is combinatorially better by a factor $1/n!$. This observation suggests that by grouping the sum over connected Feynman diagrams in determinant form (which should be possible, thanks to the $(-1)^\pi$ signs coming from the fermionic statistics), then one should be able to improve the bound derived above by a factor $(\text{const.})^n/n!$. The rough idea is to group together the connected Feynman diagrams sharing a common spanning tree, and then to resum the result of the pairings outside the spanning tree in the form of a determinant, to be bounded by $(\text{const.})^n$. Since the number of spanning trees

[1]By taking into account the connectedness condition, one could improve the bound above, by replacing $\{\#$ Feynman diagrams of order $n\}$ by $\{\#$ connected Feynman diagrams of order $n\}$. However, one can easily convince oneself that both quantities grow like $(\text{const.})^n(n!)^2$ for large n, so the connectedness condition does not qualitatively change things.

over n vertices is of the order $(\mathrm{const.})^n n!$ for large n, this procedure seems in fact to improve the previous bound by the desired combinatorial factor.

Of course, one needs to proceed with care, in order to avoid overcounting of the diagrams, when summing over the spanning trees. The key combinatorial formula was proposed long ago by Brydges, Battle and Federbush [12, 19, 20], later improved in collaboration with Kennedy [21], and reads as follows: let $\Psi(P_i)$ be a shorthand for the quartic monomial associated with the i-th quartic vertex, *i.e.*,

$$\Psi(P_i) = \psi^+_{(t_i,x_i)\sigma_i}\psi^-_{(t_i,x_i)\sigma_i} v_{\sigma_i\sigma'_i}(x_i - y_i)\psi^+_{(t_i,y_i)\sigma'_i}\psi^-_{(t_i,y_i)\sigma'_i} ,$$

with P_i the set of four labels (to be called 'field labels') associated with the four creation/annihilation operators, then

$$\begin{aligned}\langle \Psi(P_1); \cdots ; \Psi(P_n)\rangle^0_{\beta,L} \\ = \sum_T \alpha_T \prod_{\ell\in T} g_\ell \int dP_T(\underline{s}) \det[s_{i(f),i(f')}g_{(f,f')}],\end{aligned} \tag{9}$$

where the sum is over the collections T of $n-1$ propagators (graphically, of solid lines, each corresponding to the contraction of a pair of field labels $f_1, f_2 \in \cup_i P_i$) guaranteeing minimal connection among the n vertices (*i.e.*, the union of T and of the wiggly lines of the n vertices forms a spanning tree), α_T is a sign (irrelevant for the purpose of proving convergence of the series), and $dP_T(\underline{s})$ is a probability measure, with the variables $\underline{s} = (s_{ij})_{i,j=1,\dots n}$ playing the role of interpolation parameters, supported on a set of s_{ij}'s such that $s_{ij} = (u_i, u_j)$, for a family of vectors $u_i \in \mathbb{C}^n$ of unit norm. Finally, the matrix $[s_{i(f),i(f')}g_{(f,f')}]$ has elements labelled by field labels $f, f' \in \{\cup_i P_i\}\setminus P_T$, where P_T is the set of field labels corresponding to the propagators in T (*i.e.*, $[s_{i(f),i(f')}g_{(f,f')}]$ is labelled by fields outside the spanning tree); in writing $s_{i(f),i(f')}$, $i(f)$ indicates the vertex which the field labelled by f belongs to; moreover, $g_{(f,f')}$ is a shorthand for $g_{\sigma(f),\sigma(f')}(x(f) - x(f'), t(f) - t(f'))$.

In order to use effectively the 'BBFK interpolation formula', I must discuss how to bound the determinant. Life is easy if $g_{(f,f')}$ is in Gram form, *i.e.*, if there exist vectors of finite norm in a separable Hilbert space, such that $g_{(f,f')} = (a_f, b_{f'})$, in which case (Gram-Hadamard inequality, see, *e.g.*, [25, Appendix A.3])

$$\det[s_{i(f),i(f')}g_{(f,f')}] = \det(u_{i(f)} \otimes a_f, u_{i(f')} \otimes b_{f'}) \le \prod_f \|a_f\| \, \|b_f\|.$$

In our case the propagator is not in Gram form, but almost so (we give here the formulas in the $\beta, L \to \infty$ limit; similar, but slightly more cumbersome, expressions are valid in the general case):

$$g_{(f,f')} = \mathbb{1}_{t(f)>t(f')} A^+_{f,f'} - \mathbb{1}_{t(f)\le t(f')} A^-_{f,f'},$$

where

$$A^\pm_{f,f'} = \int_{\mathcal{B}} \frac{dk}{|\mathcal{B}|} e^{-ik(x(f)-x(f'))} \big[e^{-(\hat H^0(k)-\mu)(t(f)-t(f'))} P_\pm(k) \big]_{\sigma(f),\sigma(f')}.$$

Note that both $A^+_{f,f'}$ and $A^-_{f,f'}$ are in Gram form: in fact,

$$\begin{aligned} A^\pm_{f,f'} &= \iint\limits_{\mathcal{B}\times\mathbb{R}} \frac{dk d\omega}{|\mathcal{B}|\pi} \sum_{s\in I} \big[a^\pm_{x(f),t(f)\sigma(f)}(k,\omega,s) \big]^* a^\pm_{x(f'),t(f'),\sigma(f')}(k,\omega,s) \\ &\equiv \Big(a^\pm_{x(f),t(f),\sigma(f)}, a^\pm_{x(f'),t(f'),\sigma(f')} \Big), \end{aligned} \tag{10}$$

where

$$a^\pm_{x,t,\sigma}(k,\omega,s) = e^{ikx+i\omega t} \left[\sqrt{\pm(\hat H^0(k)-\mu)} \frac{P_\pm(k)}{-i\omega + \hat H^0(k) - \mu} \right]_{s,\sigma}.$$

From the definitions, letting $\|\cdot\|$ be the norm induced by the inner product in (10), one finds

$$\|a^\pm_{x,t,\sigma}\|^2 = \int \frac{dk}{|\mathcal{B}|} \Big[P_\pm(k) \Big]_{\sigma,\sigma},$$

which is ≤ 1, for both signs $\pm$. Summarizing, the determinant in the right side of (9) admits the following 'time-ordered' Gram representation:

$$\begin{aligned} \det[s_{i(f),i(f')} g_{(f,f')}] = \det \big[& \mathbb{1}_{t(f)>t(f')} \big(u_{i(f)} \otimes a^+_f, u_{i(f')} \otimes a^+_{f'} \big) \\ & - \mathbb{1}_{t(f)\le t(f')} \big(u_{i(f)} \otimes a^-_f, u_{i(f')} \otimes a^-_{f'} \big) \big], \end{aligned} \tag{11}$$

which is in fact enough for our purposes: as shown in [18, 49], the right side of (11) is bounded above by $(\|a^+\| + \|a^-\|)^{2d}$, where d is the dimension of the matrix (in our case, $d = n+1$). In conclusion, recalling that

$\|a^{\pm}_{x,t}\| \leq 1$, one finds that the integral of the determinant in the right side of (9) is bounded from above by $2^{2(n+1)}$. Recalling also that the number of collections T ('number of spanning trees') is bounded above by $(\mathrm{const.})^n n!$, one sees that this estimate is enough for our purposes, *i.e.*, to obtain uniform convergence of the specific free energy, for U small enough. In fact, in view of the combinatorial gain obtained above, the n-th order term in the series for $\log Z(U)$ is bounded above by $C^n|U|^n$, for a suitable constant $C > 0$, which is summable over n for $|U| < C^{-1}$.

A similar argument can be repeated to prove the convergence of the series in U for the Euclidean truncated correlation functions of local operators in the same domain of analyticity, as well as their exponential decay at large space-time distances, uniformly in β and L. In particular, there exists $C, c > 0$ such that

$$\langle J_{i,t}; J_j \rangle_\infty \leq C e^{-c\delta_\mu t}, \qquad \forall t \geq 0. \tag{12}$$

The details of the proof will not be belabored here and are left as an instructive exercise to the reader.

6.1 Wick rotation for the Kubo conductivity

Using the analyticity and exponential decay in space-time of the Euclidean correlations discussed above, one can prove that the Wick rotation for the Kubo conductivity, sketched in Section 5.1, can be made rigorous. The key steps in the proof are the following (for details, consult [29, Section 6]):

1) Let $J_i(z) = e^{izH} J_i e^{-izH}$, $z \in \mathbb{C}$, be the complex time evolution of the current. Note that $J_{i,t} = J_i(-it)$, $t \in \mathbb{R}$. At finite β, L, the correlation $L^{-2}\langle J_i(z); J_j \rangle_{\beta,L}$ is entire in z. Moreover, thanks to the results on the Euclidean correlations sketched above, these finite volume correlations converge as $\beta, L \to \infty$ along the negative imaginary axis, and they are uniformly bounded in the whole complex lower half-plane. Therefore, by Vitali's theorem, the limit as $\beta, L \to \infty$ of the complex time correlations, $\langle J_i(z); J_j \rangle_\infty$ exists and is analytic in the whole open lower half-plane. Moreover, by Cauchy-Schwarz inequality and (12),

$$|\langle J_i(z); J_j \rangle_\infty| \leq \langle J_i(\mathrm{Re}\, z); J_i \rangle_\infty^{1/2} \langle J_j(\mathrm{Re}\, z); J_j \rangle_\infty^{1/2} \leq C e^{\delta_\mu \mathrm{Re}\, z}.$$

2) For $t \in \mathbb{R}$, the limit

$$\langle [J_i, J_j(t)] \rangle_\infty = \lim_{\beta\to\infty} \lim_{L\to\infty} L^{-2} \langle [J_i, J_j(t)] \rangle_{\beta,L}$$

exists, uniformly in $t \in \mathbb{R}$. This follows from the Lieb-Robinson bounds on the quantum evolution of lattice systems [43], and in particular from one of its corollary, worked out in [17, 46], which implies the existence of the infinite volume, zero temperature, real time correlations for a large class of quantum lattice systems.

Given these two results, and using the fact that $\langle J_i \rangle_{\beta,L} = 0$, one rewrites, for $\omega < 0$,

$$\hat K_{ij}(\omega) = \lim_{\beta\to\infty} \lim_{L\to\infty} \frac{1}{L^2} \times \int_0^{\beta/2} \Big(e^{-i\omega t} \langle J_i(-it) J_j \rangle_{\beta,L} + e^{+i\omega t} \langle J_j(-it) J_i \rangle_{\beta,L} \Big) dt \tag{13}$$

Since the integrand, thought of as a function of $z = -it$, is analytic in the complex lower half-plane, and since it decays exponentially in $-\mathrm{Re}\, z$, one can rotate the integrals over the negative imaginary axis as shown in *Fig. 7*.

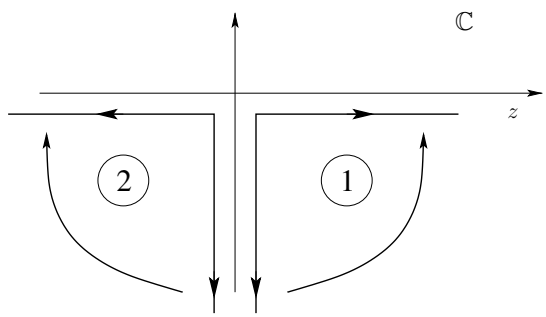


Fig. 7 – The complex rotations of the integration paths performed on the two contributions in the right side of (13)*. The rotation labelled (1) (resp. (2)) refers to the first (resp. second) term in the right side of the equation.*

After the rotation shown in the picture, one rewrites:

$$\hat K_{ij}(\omega) = \lim_{\epsilon\to 0^+} \lim_{\beta\to\infty} \lim_{L\to\infty} \frac{i}{L^2} \times \int_{-\beta/2}^{0} \Big(e^{-\omega(t+i\epsilon)} \langle J_i(-t-i\epsilon) J_j \rangle_{\beta,L} - e^{-\omega(t-i\epsilon)} \langle J_j(t-i\epsilon) J_i \rangle_{\beta,L} \Big) dt$$

The two terms under the integral sign represent complex time current- current correlations, with times having a small imaginary part: in other words, the integration paths shadow the real axis from below. In order to 'push' the paths on the real axis one uses the information that the real-time correlation exist, thanks to the Lieb-Robinson bounds. Using this fact, together with the a priori bounds on the complex time correlations discussed above, one finds that the limit as $\epsilon \to 0^+$ of the above integrals exist, and are equal to the time integral of the limiting (real-time) correlation functions, as desired. For details, see [29, Section 6].

In conclusion, as announced in Section 5.1, the first term in the right side of the Kubo conductivity (3) equals $-\frac{1}{\eta}\hat{K}_{ij}(-\eta)$. Moreover, as claimed in the same section, the diamagnetic term equals $\frac{1}{\eta}\hat{K}_{ij}(0)$ (for a proof of this fact, based on the use of Ward Identities, see (26) below), thus leading to the desired identity $\sigma_{ij} = \bar{\sigma}_{ij}$, with $\bar{\sigma}_{ij}$ as in (6).

7. Ward Identities, Schwinger-Dyson equations and cancellations in perturbation theory

From the theory of the Euclidean correlations discussion above, we inferred that $\sigma_{ij} = \bar{\sigma}_{ij}$, that $\bar{\sigma}_{ij}$ is analytic in U in a small ball centered at the origin, and that the truncated multi-point Euclidean current correlations decay exponentially to zero at large space-time separation. I now want to combine these ingredients in order to show that $\bar{\sigma}_{ij} = \bar{\sigma}_{ij}\big|_{U=0}$, that is, all the interaction corrections to the Euclidean Kubo conductivity vanishes, which implies Theorem 1.

As anticipated above, these remarkable cancellations can be proved via a combination of two classes of identities, namely the Ward Identities and the Schwinger-Dyson equations, which are briefly reviewed in the following.

7.1 Ward Identities

The Ward Identities that I want to discuss here are a consequence of the continuity equation for the lattice current. The starting point is the definition of the imaginary-time evolution of the fermionic density: $n^{\sigma}_{(t,x)} := e^{tH} n_{x,\sigma} e^{-tH}$.

By deriving it with respect to t one gets:

$$\frac{\partial}{\partial t} n^{\sigma}_{(t,x)} = \big[H, n^{\sigma}_{(t,x)}\big] = i \sum_{y,\sigma'} \Big(J^{\sigma\sigma'}_{xy} \Big)_t, \tag{14}$$

where $\Big(J^{\sigma\sigma'}_{xy} \Big)_t$ is the imaginary time evolution of the lattice current:

$$J^{\sigma\sigma'}_{xy} := i\big(\psi^+_{x,\sigma} H^0_{\sigma,\sigma'}(x-y)\psi^-_{y,\sigma'} - \psi^+_{y,\sigma'} H^0_{\sigma',\sigma}(y-x)\psi^-_{x,\sigma}\big).$$

Note that by letting σ vary in I in (14), one obtains $|I|$ independent local conservation laws. It is convenient to consider the following linear combination of these $|I|$ conservation laws: let $\tilde J_{0,(t,p)} := q \sum_{x,\sigma} e^{-ipx_\sigma} n^{\sigma}_{(t,x)}$, where $x_\sigma = x + r_\sigma$; by deriving with respect to t, and using the anti-symmetry of the lattice current under the exchange $(x,\sigma) \leftrightarrow (y,\sigma')$, one finds:

$$\frac{\partial}{\partial t} \tilde J_{0,(t,p)} = \frac{iq}{2} \sum_{\substack{x,y \\ \sigma,\sigma'}} (e^{-ipx_\sigma} - e^{-ipy_{\sigma'}}) \Big(J^{\sigma\sigma'}_{xy} \Big)_t \equiv -p \cdot \tilde J_{(t,p)},$$

where in the last equation I defined $\tilde J_{(t,p)}$ to be

$$\tilde J_{(t,p)} := \frac{q}{2} \sum_{\substack{x,y \\ \sigma,\sigma'}} e^{-ipx_\sigma} \eta\big(p(y_{\sigma'} - x_\sigma)\big)\, (y_{\sigma'} - x_\sigma) \Big(J^{\sigma\sigma'}_{xy} \Big)_t,$$

with $\eta(x) := (1 - e^{-ix})/(ix)$. After Fourier transform with respect to (w.r.t.) the imaginary time variable, and letting $\mathbf{p} := (\omega, p)$, this conservation law reads

$$i\omega \hat J_{0,\mathbf{p}} + p \cdot \hat J_{\mathbf{p}} = 0, \tag{15}$$

where $\hat J_{0,\mathbf{p}}$ is the Fourier transform w.r.t. t of $\tilde J_{0,(t,p)}$, and similarly for $\hat J_{\mathbf{p}}$, whose components will be denoted $\hat J_{i,\mathbf{p}}$, $i = 1, 2$.

Remark. The variable p is introduced here as an auxiliary variable: the equations I will eventually be interested in are obtained by deriving w.r.t. p the Ward Identities for the correlations of $\hat J_{\mathbf{p}}$ with other observables, and

then setting this momentum variable to zero. Note that $\tilde{J}_{(t,p)}\big|_{p=0} = J_t$, which clarifies the connection with the current observable introduced above.

The desired Ward Identities are obtained by plugging the conservation law (15) into the formula for the multi-point current correlations: let

$$\hat{K}_{\alpha_1,\dots,\alpha_n}(\mathbf{p}_1,\dots,\mathbf{p}_{n-1}) := \lim_{\beta\to\infty}\lim_{L\to\infty}\frac{1}{\beta L^2}\langle T\,\hat{J}_{\alpha_1,\mathbf{p}_1};\cdots;\hat{J}_{\alpha_1,\mathbf{p}_1}\rangle_{\beta,L}, \tag{16}$$

where: $\alpha_i \in \{0,1,2\}\cup I$ and, if $\alpha \in \{0,1,2\}$, $\hat{J}_{\alpha,\mathbf{p}}$ has already been defined above, while, if $\alpha = \sigma \in I$, then $\hat{J}_{\sigma,\mathbf{p}} \equiv \hat{n}^\sigma_{\mathbf{p}}$, that is the Fourier transform w.r.t. t of $\big(n^\sigma_x\big)_t$. Moreover, in the right side of (16), $\mathbf{p}_n := -\mathbf{p}_1\cdots-\mathbf{p}_{n-1}$. Note that $\hat{K}_{i,j}\big((\omega,0)\big)$ equals the Euclidean current-current correlation $\hat{K}_{ij}(\omega)$ defined in (5) above. By using (15) into the definition (16) one gets

$$\begin{aligned} i\omega_1\hat{K}_{0,\alpha_2,\cdots,\alpha_n}(\mathbf{p}_1,\quad\dots\quad,\mathbf{p}_{n-1}) + \sum_{i=1}^{2} p_{1,i}\hat{K}_{i,\alpha_2,\cdots,\alpha_n}(\mathbf{p}_1,\dots,\mathbf{p}_{n-1}) \\ = \sum_{j=2}^{n}\hat{S}_{\alpha_j;\hat{\alpha}_j}(\mathbf{p}_1,\dots,\mathbf{p}_{n-1}), \end{aligned} \tag{17}$$

where $\hat{\alpha}_j = (\alpha_2,\dots,\not{\alpha}_j,\dots,\alpha_n)$ and, if I let

$$\hat{\Delta}_{\alpha_j}(\mathbf{p}_1,\mathbf{p}_j) := \int_0^\beta dt\, e^{-i(\omega_1+\omega_j)t}\big[\tilde{J}_{0,(t,p_1)},\tilde{J}_{\alpha_j,(t,p_j)}\big], \tag{18}$$

then

$$\begin{aligned} \hat{S}_{\alpha_j;\hat{\alpha}_j}(\mathbf{p}_1,\dots,\mathbf{p}_{n-1}) := \lim_{\beta,L\to\infty}\frac{1}{\beta L^2}\times \\ \langle T\,\hat{\Delta}_{\alpha_j}(\mathbf{p}_1,\mathbf{p}_j);\hat{J}_{\alpha_2,\mathbf{p}_2};\cdots;\hat{J}_{\alpha_{j-1},\mathbf{p}_{j-1}};\not{J}_{\alpha_j,\mathbf{p}_j};\cdots;\hat{J}_{\alpha_n,\mathbf{p}_n}\rangle_{\beta,L}. \end{aligned} \tag{19}$$

In order to prove (17), one can start from the very definition of $i\omega_1\hat{K}_{0\alpha_2\cdots\alpha_n}$:

$$\begin{aligned} &i\omega_1\hat{K}_{0,\alpha_2,\cdots,\alpha_n}(\mathbf{p}_1,\dots,\mathbf{p}_{n-1}) \\ &= \lim_{\beta\to\infty}\lim_{L\to\infty}\frac{1}{\beta L^2}\int_0^\beta dt_1\, e^{-i\omega_1 t_1}\frac{\partial}{\partial t_1}\langle T\,\tilde{J}_{0,(t_1,p_1)};\hat{J}_{\alpha_2,\mathbf{p}_2};\cdots;\hat{J}_{\alpha_n,\mathbf{p}_n}\rangle_{\beta,L}. \end{aligned}$$

In the right side, the derivative w.r.t. t_1, when acting on

$$\langle T\, \tilde{J}_{0,(t_1,p_1)}; \hat{J}_{\alpha_2,\mathbf{p}_2}; \cdots ; \hat{J}_{\alpha_n,\mathbf{p}_n}\rangle_{\beta,L},$$

can either act on $\tilde{J}_{0,(t_1,p_1)}$, in which case we get

$$-p_1 \cdot \langle T\, \tilde{J}_{(t_1,p_1)}; \hat{J}_{\alpha_2,\mathbf{p}_2}; \cdots ; \hat{J}_{\alpha_n,\mathbf{p}_n}\rangle_{\beta,L},$$

or on the theta functions that implement the action of the time-orderding operator T, in which case we get

$$\sum_{j=2}^{n} \langle T\, \big[\tilde{J}_{0,(t_1,p_1)}, \tilde{J}_{\alpha_j,(t_1,p_j)}\big]; \hat{J}_{\alpha_2,\mathbf{p}_2}; \cdots ; \not{J}_{\alpha_j,\mathbf{p}_j}; \cdots ; \hat{J}_{\alpha_n,\mathbf{p}_n}\rangle_{\beta,L}.$$

Putting together the two contributions, we get (17).

Now, recalling the fact that the truncated Euclidean correlation functions of local operators decay exponentially in space and (Euclidean) time, one has that their Fourier transforms are analytic in $\mathbf{p}_i$ in a neighborhood of the origin. In particular, I can differentiate (17) w.r.t. p_1 and then set $p_1 = 0$. Let me do so in a few special cases, which will be useful in the following. Consider (17) in the case $n \geq 2$ and $\alpha_j \in I$, $\forall j \geq 2$, in which case it reduces to

$$i\omega_1 \hat{K}_{0,\underline{\sigma}}(\mathbf{p}_1, \dots, \mathbf{p}_{n-1}) + \sum_{j=1}^{2} p_{1,j} \hat{K}_{j,\underline{\sigma}}(\mathbf{p}_1, \dots, \mathbf{p}_{n-1}) = 0,$$

where $\underline{\sigma} = (\alpha_2, \dots, \alpha_n)$. The reason why the right side of the Ward Identities is zero is that $\tilde{J}_{0,(t_1,p_1)}$ commutes with $\tilde{J}_{\sigma_j,(t_1,p_j)}$, $\forall \sigma_j \in I$. By differentiating the previous identity w.r.t. p_1, and then setting $p_1 = 0$, one gets:

$$\begin{aligned} &\hat{K}_{j,\underline{\sigma}}\big((\omega,0), \mathbf{p}_2, \dots, \mathbf{p}_{n-1}\big) \\ &\qquad = -i\omega \frac{\partial}{\partial p_{1,j}} \hat{K}_{0,\underline{\sigma}}\big((\omega,0), \mathbf{p}_2, \dots, \mathbf{p}_{n-1}\big), \quad j = 1, 2. \end{aligned} \tag{20}$$

Similarly, if $n \geq 3$, $\alpha_2 = j' \in \{1, 2\}$, and $(\alpha_3, \dots, \alpha_n) \in I^{n-2}$, then the Ward Identity (17) reduces to

$$\begin{aligned} i\omega_1 \hat{K}_{0,j',\underline{\sigma}}(\mathbf{p}_1, \dots, \mathbf{p}_{n-1}) + \sum_{j=1}^{2} p_{1,j} \hat{K}_{j,j',\underline{\sigma}}\,(\mathbf{p}_1, \dots, \mathbf{p}_{n-1}) = \\ \hat{S}_{j';\underline{\sigma}}(\mathbf{p}_1, \dots, \mathbf{p}_{n-1}), \end{aligned} \tag{21}$$

while, if $\alpha_2 = 0$,

$$i\omega_1 \hat{K}_{0,0,\underline{\sigma}}(\mathbf{p}_1, \ldots, \mathbf{p}_{n-1}) + \sum_{j=1}^{2} p_{1,j} \hat{K}_{j,0,\underline{\sigma}}(\mathbf{p}_1, \ldots, \mathbf{p}_{n-1}) = 0,$$

or, by exchanging the roles of $\mathbf{p}_1$ and $\mathbf{p}_2$,

$$i\omega_2 \hat{K}_{0,0,\underline{\sigma}}(\mathbf{p}_1, \ldots, \mathbf{p}_{n-1}) + \sum_{j=1}^{2} p_{2,j} \hat{K}_{0,j,\underline{\sigma}}(\mathbf{p}_1, \ldots, \mathbf{p}_{n-1}) = 0. \quad (22)$$

Now, by differentiating (21) w.r.t. p_1 and then setting $p_1 = 0$, one gets

$$\begin{aligned} \hat{K}_{j,j',\underline{\sigma}}\big((\omega_1, 0), \mathbf{p}_2, \ldots, \mathbf{p}_{n-1}\big) \\ = -i\omega_1 \frac{\partial}{\partial p_{1,j}} \hat{K}_{0,j',\underline{\sigma}}\big((\omega_1, 0), \mathbf{p}_2, \ldots, \mathbf{p}_{n-1}\big) \\ + \frac{\partial}{\partial p_{1,j}} \hat{S}_{j';\underline{\sigma}}\big((\omega_1, 0), \mathbf{p}_2, \ldots, \mathbf{p}_{n-1}\big), \end{aligned} \quad (23)$$

Moreover, by differentiating (22) w.r.t. p_2 and then setting $p_2 = 0$, one gets

$$\begin{aligned} \hat{K}_{0,j',\underline{\sigma}}(\mathbf{p}_1, (\omega_2, 0), \mathbf{p}_3, \ldots, \mathbf{p}_{n-1}) \\ = -i\omega_2 \frac{\partial}{\partial p_{2,j'}} \hat{K}_{0,0,\underline{\sigma}}(\mathbf{p}_1, (\omega_2, 0), \mathbf{p}_3, \ldots, \mathbf{p}_{n-1}). \end{aligned} \quad (24)$$

By plugging (24) into (23), we obtain

$$\begin{aligned} \hat{K}_{j,j',\underline{\sigma}}\big((\omega_1, 0), (\omega_2, 0), \mathbf{p}_3, \ldots, \mathbf{p}_{n-1}\big) \\ = -\omega_1\omega_2 \frac{\partial^2}{\partial p_{1,j} \partial p_{2,j'}} \hat{K}_{0,0,\underline{\sigma}}\big((\omega_1, 0), (\omega_2, 0), \mathbf{p}_3, \ldots, \mathbf{p}_{n-1}\big) \\ + \frac{\partial}{\partial p_{1,j}} \hat{S}_{j';\underline{\sigma}}\big((\omega_1, 0), (\omega_2, 0), \mathbf{p}_3, \ldots, \mathbf{p}_{n-1}\big). \end{aligned} \quad (25)$$

In the following, I intend to combine the identities (20) and (25) with the Schwinger-Dyson equation. Before I do so, let me conclude this section by

proving that the diamagnetic term in the Kubo conductivity is equal to $\frac{1}{\eta}\hat{K}_{ij}(0)$: consider (17) with $n=2$ and $\alpha_2 = j \in \{1,2\}$:

$$i\omega\hat{K}_{0,j}(\mathbf{p})+p_1\hat{K}_{1,j}(\mathbf{p})+p_2\hat{K}_{2,j}(\mathbf{p}) = \hat{S}_j(\mathbf{p}), \qquad \text{where} \qquad \mathbf{p}=(\omega,p).$$

Deriving this equation w.r.t. p and then setting $\mathbf{p}=0$, one gets:

$$\hat{K}_{i,j}(\mathbf{0}) = \frac{\partial\hat{S}_j}{\partial p_i}(\mathbf{0}). \tag{26}$$

Recalling the definition of $\hat{S}_j(\mathbf{p})$, eq.(19), I can write

$$\hat{S}_j(\mathbf{p}) = q\big\langle\big[\sum_{x,\sigma} e^{-ip\cdot x_\sigma} n^\sigma_{(0,x)}, J_{j,-p}\big]\big\rangle_\infty$$

(here I denoted by $J_{j,-p}$ the j-th component of $\tilde{J}_{(t,-p)}\big|_{t=0}$), so that

$$\begin{aligned}\hat{K}_{i,j}(\mathbf{0}) &= -iq\big\langle\big[X_i, J_j\big]\big\rangle_\infty \\ &= q^2\big\langle\big[X_i,[H,X_j]\big]\big\rangle_\infty = -q^2\big\langle\big[[H,X_i],X_j\big]\big\rangle_\infty,\end{aligned}$$

as desired.

7.2 Schwinger-Dyson equation

The Schwinger-Dyson equation that I want to discuss is an interpolation identity relating the current-current correlation $\hat{K}_{i,j}$ of the interacting model, with interaction strength U, to higher point correlation functions (such as current-current-density-density correlations) at interaction strength U', where U' is an interpolation parameter between 0 and U. More general Schwinger-Dyson equations can be obtained along the same lines outlined below, for different multi-point correlations.

For the purpose of this and the following section, I add a superscript U to the interacting current-current correlation at interaction strength U (and similarly for the higher point correlations), in order to emphasize its dependence upon this parameter. By the fundamental theorem of calculus,

$$\hat{K}^U_{i,j}(\mathbf{p}) = \hat{K}^0_{i,j}(\mathbf{p}) + \int_0^U dU'\,\frac{d}{dU'}\hat{K}^{U'}_{i,j}(\mathbf{p}).$$

Recall that $\hat{K}^{U'}_{i,j}$ is the current-current correlation computed in the Gibbs state with Hamiltonian $H = H_0 - U'V - \mu N$, see (1). Therefore, by computing the derivative in U', one finds

$$\hat{K}^{U}_{i,j}(\mathbf{p}) = \hat{K}^{0}_{i,j}(\mathbf{p}) + \int_0^U dU'\, \hat{K}^{U'}_{i,j,V}(\mathbf{p}. - \mathbf{p}), \tag{27}$$

where

$$\hat{K}^{U'}_{i,j}(\mathbf{p}, -\mathbf{p}) = \lim_{\beta\to\infty} \lim_{L\to\infty} \frac{1}{\beta L^2} \langle T\, \hat{J}_{i,\mathbf{p}}; \hat{J}_{j,-\mathbf{p}}; \int_0^\beta V_t\, dt \rangle^{U'}_{\beta,L},$$

and, again, the superscript U' in the right side has been added to emphasize that the Gibbs measure is associated with the Hamiltonian $H = H_0 - U'V - \mu N$, with interaction strength U'.

Now, in the limit $\beta, L \to \infty$, $\int_0^\beta V_t\, dt$ formally tends to

$$\int_{\mathbb{R}} \frac{dq_0}{2\pi} \int_{\mathcal{B}} \frac{dq}{|\mathcal{B}|} \sum_{\sigma,\sigma'\in I} \hat{n}^{\sigma}_{\mathbf{q}} \hat{v}_{\sigma,\sigma'}(q) \hat{n}^{\sigma'}_{-\mathbf{q}},$$

where $\mathbf{q} = (q_0, q)$; at finite β and finite L, $\int_0^\beta V_t\, dt$ can be written in the same form, but the integrals over q_0 and q should be interpreted as suitable, discrete, Riemann sum approximations. By using this explicit expression for $\int_0^\beta V_t\, dt$, one rewrites:

$$\hat{K}^{U'}_{i,j,V}(\mathbf{p}) = \lim_{\beta\to\infty} \lim_{L\to\infty} \frac{1}{\beta L^2} \times \\ \times \int \frac{dq_0 dq}{2\pi|\mathcal{B}|} \sum_{\sigma,\sigma'\in I} \hat{v}_{\sigma,\sigma'}(q) \langle T\, \hat{J}_{i,\mathbf{p}}; \hat{J}_{j,-\mathbf{p}}; \hat{n}^{\sigma}_{\mathbf{q}} \hat{n}^{\sigma'}_{-\mathbf{q}} \rangle^{U'}_{\beta,L}.$$

where, again, the integrals over q_0 and q should be interpreted as appropriate Riemann sums. Note that there is no semicolon between $\hat{n}^{\sigma}_{\mathbf{q}}$ and $\hat{n}^{\sigma'}_{-\mathbf{q}}$, *i.e.*, the correlation function in the right side is not 'completely truncated'. Of course, if desired, one can rewrite it as a linear combination of (products of) completely truncated correlations. If I do so, and then perform the limit

$\beta, L \to \infty$, I find that (27) can be rewritten as:

$$\begin{aligned}\hat K^U_{i,j}(\mathbf{p}) &= \hat K^0_{i,j}(\mathbf{p}) + \int_0^U dU' \sum_{\sigma,\sigma'\in I} \times \\ &\times \Bigg[\int_{\mathbb{R}} \frac{dq_0}{2\pi} \int_{\mathcal{B}} \frac{dq}{|\mathcal{B}|} \hat v_{\sigma,\sigma'}(q) \hat K^{U'}_{i,j,\sigma,\sigma'}(\mathbf{p},-\mathbf{p},\mathbf{q}) \\ &+ 2\hat v_{\sigma,\sigma'}(0) \hat K^{U'}_{i,j,\sigma}(\mathbf{p},-\mathbf{p}) \hat K^{U'}_{\sigma'} \\ &+ 2\hat v_{\sigma,\sigma'}(-p) \hat K^{U'}_{i,\sigma}(\mathbf{p}) \hat K^{U'}_{j,\sigma'}(-\mathbf{p})\Big)\Bigg].\end{aligned} \tag{28}$$

7.3 Cancellations and proof of the main theorem

By combining the Ward Identity, or, better, the two corollaries of the Ward Identity, eqs.(20) and (25), with the Schwinger-Dyson equation, one can easily prove the desired cancellations, leading to the claim of our main theorem. Let me denote by $\bar\sigma^U_{ij}$ the Euclidean Kubo conductivity at interaction strength U:

$$\bar\sigma^U_{ij} = \frac{1}{|\ell_1 \wedge \ell_2|} \lim_{\omega\to 0} \frac{\partial \hat K^U_{i,j}(\omega,0)}{\partial\omega}. \tag{29}$$

Remember that $\hat K^U_{i,j}(\mathbf{p})$ is smooth (analytic!) in its argument. By using the Schwinger-Dyson equation (28) in the right side of (29), one gets:

$$\begin{aligned}\bar\sigma^U_{ij} &= \bar\sigma^0_{ij} + \frac{1}{|\ell_1 \wedge \ell_2|} \sum_{\sigma,\sigma'\in I} \int_0^U dU' \times \\ &\times \lim_{\omega\to 0} \Bigg\{ \int_{\mathbb{R}} \frac{dq_0}{2\pi} \int_{\mathcal{B}} \frac{dq}{|\mathcal{B}|} \hat v_{\sigma,\sigma'}(q) \frac{\partial}{\partial\omega} \hat K^{U'}_{i,j,\sigma,\sigma'}\big((\omega,0),-(\omega,0),\mathbf{q}\big) \\ &+ 2\hat v_{\sigma,\sigma'}(0) \frac{\partial}{\partial\omega} \hat K^{U'}_{i,j,\sigma}\big((\omega,0),-(\omega,0)\big) \hat K^{U'}_{\sigma'} \\ &+ 2\hat v_{\sigma,\sigma'}(0) \frac{\partial}{\partial\omega} \Big[\hat K^{U'}_{i,\sigma}\big((\omega,0)\big) \hat K^{U'}_{j,\sigma'}(-(\omega,0))\Big]\Bigg\}.\end{aligned} \tag{30}$$

I now focus separately on the three contributions displayed in the three lines of this equation. Let me start with the first term, *i.e.*, the one involving

$$\frac{\partial}{\partial\omega}\hat{K}^{U'}_{i,j,\sigma,\sigma'}\big((\omega,0),-(\omega,0),\mathbf{q}\big).$$

By using (25), one can rewrite

$$\begin{aligned}\frac{\partial}{\partial\omega}\hat{K}^{U'}_{i,j,\sigma,\sigma'}\big((\omega,0),-(\omega,0),\mathbf{q}\big) = \\ = -\frac{\partial}{\partial\omega}\Big[\omega^2\frac{\partial^2}{\partial p_{1,i}\partial p_{2,j}}\hat{K}^{U'}_{0,0,\sigma,\sigma'}\big((\omega,0),-(\omega,0),\mathbf{q}\big)\Big],\end{aligned} \tag{31}$$

where I used the fact that $\frac{\partial}{\partial p_{1,i}}\hat{S}_{j;\underline{\sigma}}\big((\omega,0),-(\omega,0),\mathbf{q}\big)$ is independent of ω, by the very definition of $\hat{S}_{j;\underline{\sigma}}$, cf. with (18) and (19) and recall that in this case $\omega_1 = \omega = -\omega_2$. By recalling the smoothness and boundedness properties of $\hat{K}^{U'}_{0,0,\sigma,\sigma'}$ and of its derivatives, we see that (31) implies that

$$\frac{\partial}{\partial\omega}\hat{K}^{U'}_{i,j,\sigma,\sigma'}\big((\omega,0),-(\omega,0),\mathbf{q}\big)$$

tends to zero as $\omega \to 0$. Correspondingly, the contribution to $\bar{\sigma}^U_{ij}$ from the second line of (30) vanishes (the only issue to be discussed is the exchange of the limit with the integral over q_0 – for details, see [29, Section 4]).

A similar discussion shows that the contributions to $\bar{\sigma}^U_{ij}$ from the third and fourth lines of (30) are zero, as well, and this completes the proof of Theorem 1.

8. Conclusions

I reviewed a recent proof of universality of the Kubo conductivity for a class of interacting lattice fermions in two dimensions. 'Universality' refers here to the independence of the transport coefficient from the strength of the electron-electron interaction: *i.e.*, the interacting conductivity is equal to the non-interacting one, provided the non-interacting Hamiltonian is chosen to have a spectral gap, and the strength of the interaction is small, compared to the gap size. The class of models which this result applies to includes an interacting version of the Haldane model. The nice feature of this model is that

it displays a transition from a 'topologically trivial' insulating phase (characterized by vanishing transverse Kubo conducitivity) to a non trivial one, as the parameters ϕ and M characterizing the hopping term are varied. The main theorem of these notes, Theorem 1, proves that the same transition takes place in the presence of interactions, thus providing an example of 'topological phase transition' in a model of interacting electrons.

Building upon the ideas and methods reviewed here, in particular combining them with an infrared multiscale analysis of the ground state of the interacting fermionic system, one can even characterize the nature of transition line from the topologically non-trivial phase to the trivial one, and it turns out that, in the case of the interacting Haldane model, the critical theory belongs to the same universality class as the non-interacting one. Nevertheless, the interactions have the remarkable effect of renormalizing ('dressing') the transition line: repulsive interactions tend to enlarge the topologically non-trivial region. For details, see [26, 30]. Constructive fermionic renormalization group methods have also been used to prove universality of the longitudinal ('optical') conductivity for interacting graphene [28] and, more in general, for the interacting Haldane model on the renormalized critical line [26, 30].

A lot remains to be done, but it is likely that extensions of the methods reviewed here, and further exchange of ideas between the mathematical physics and condensed matter communities, will allow us to attack and solve new problems that are currently beyond the state of the art, most notably the universality of quantum transport coefficients in interacting electron systems, in the presence of edges and disorder (in the case of systems with edges and no disorder, see [6, 45] for recent progress). I hope that the next decades will also witness advances in other challenging open problems in mathematical physics, such as the theory of the superconducting phase in interacting Fermi systems, of the condensed phase in interacting Bose systems, and of the ferromagnetic phase in ferromagnetic quantum spin systems.

A. The non-interacting Hall conductivity as the first Chern number of the Bloch bundle

In this appendix, I give a sketch of the proof of (4). I assume for simplicity that $r_\sigma = 0, \forall \sigma \in I$, and I set $q = -1$. In this case, after Fourier transform,

one gets

$$J = \sum_{\sigma,\sigma'} \int_{\mathcal{B}} \frac{dk}{|\mathcal{B}|} \hat{\psi}^+_{k,\sigma} \partial_k \hat{H}^0_{\sigma,\sigma'}(k) \hat{\psi}^-_{k,\sigma'},$$

where $\hat{\psi}^\pm_{k,\sigma} = \sum_x e^{\mp ik\cdot x} \psi^\pm_{x,\sigma}$. Using the Wick rule, one gets

$$\big\langle [J_i, J_j(t)] \big\rangle_\infty = \sum_{\sigma_1,\dots,\sigma_4} \int_{\mathcal{B}} \frac{dk}{|\mathcal{B}|} \partial_{k_i} \hat{H}^0_{\sigma_1,\sigma_2}(k) \times$$
$$\big\langle \hat{\psi}^-_{k,\sigma_2} \hat{\psi}^+_{k,\sigma_3}(t) \big\rangle_\infty \partial_{k_j} \hat{H}^0_{\sigma_3,\sigma_4}(k) \big\langle \hat{\psi}^+_{k,\sigma_1} \hat{\psi}^-_{k,\sigma_4}(t) \big\rangle_\infty - \Big(i \leftrightarrow j\Big),$$

where

$$\big\langle \hat{\psi}^-_{k,\sigma_2} \hat{\psi}^+_{k,\sigma_3}(t) \big\rangle_\infty = \big[P_+(k) e^{i(\hat{H}^0(k)-\mu)t}\big]_{\sigma_2,\sigma_3}, \tag{32}$$
$$\big\langle \hat{\psi}^+_{k,\sigma_1} \hat{\psi}^-_{k,\sigma_4}(t) \big\rangle_\infty = \big[P_-(k) e^{-i(\hat{H}^0(k)-\mu)t}\big]_{\sigma_4,\sigma_1}. \tag{33}$$

I plug these expressions in the previous formula, and I use the fact that

$$\partial_k \hat{H}^0(k) = \partial_k \sum_{\alpha=\pm} P_\alpha(k) \hat{H}^0(k) P_\alpha(k).$$

After a straightforward computation, one finds that

$$\big\langle [J_i, J_j(t)] \big\rangle_\infty = G''_{ij}(t),$$

with

$$G_{ij}(t) =$$
$$- \int_{\mathcal{B}} \frac{dk}{|\mathcal{B}|} \mathrm{Tr}\Big[\partial_{k_i} P_-(k) e^{i(\hat{H}^0(k)-\mu)t} \partial_{k_j} P_-(k)\, P_-(k) e^{-i(\hat{H}^0(k)-\mu)t}\Big]$$
$$+ \Big(i \leftrightarrow j\Big).$$

The contribution to the Kubo conductivity involving $\big\langle [J_i, J_j(t)] \big\rangle_\infty$ can then be written as $\frac{1}{|\ell_1 \wedge \ell_2|}$ times the limit as $\eta \to 0^+$ of

$$\frac{-i}{\eta} \int_{-\infty}^0 e^{\eta t} G''_{ij}(t) dt = \frac{-i}{\eta} G'_{ij}(0) + i G_{ij}(0) - i\eta \int_{-\infty}^0 e^{\eta t} G_{ij}(t) dt.$$

Now, the third term in the right side goes to zero as $\eta \to 0^+$. Moreover, a straightforward computation shows that the first term in the right side (which is singular as $\eta \to 0^+$) is equal to

$$\frac{-i}{\eta} G'_{ij}(0) = \frac{1}{\eta} \int_{\mathcal{B}} \frac{dk}{|\mathcal{B}|} \mathrm{Tr}\, \partial_{k_i} \hat{H}^0(k) \partial_{k_j} P_-(k),$$

which cancels exactly with the diamagnetic contribution to the Hall conductivity. In fact, after performing a Fourier transform, and using (33), one finds

$$\begin{aligned} -\frac{1}{\eta} \langle [[H, X_i], X_j] \rangle_\infty &= \frac{1}{\eta} \int_{\mathcal{B}} \frac{dk}{|\mathcal{B}|} \langle \hat{\psi}^+_k \partial_{k_i} \partial_{k_j} \hat{H}^0(k) \hat{\psi}^-_k \rangle_\infty \qquad (34) \\ &= \frac{1}{\eta} \int_{\mathcal{B}} \frac{dk}{|\mathcal{B}|} \mathrm{Tr}\, P_-(k) \partial_{k_i} \partial_{k_j} \hat{H}^0(k) \qquad (35) \\ &= -\frac{1}{\eta} \int_{\mathcal{B}} \frac{dk}{|\mathcal{B}|} \mathrm{Tr}\, \partial_{k_i} \hat{H}^0(k) \partial_{k_j} P_-(k), \qquad (36) \end{aligned}$$

that is $\frac{-i}{\eta} G'_{ij}(0) - \frac{1}{\eta} \langle [[H, X_i], X_j] \rangle_\infty = 0$.

In conclusion, plugging all these relations in (3) at $U = 0$, and recalling the fact that $|\ell_1 \wedge \ell_2|\, |\mathcal{B}| = (2\pi)^2$, one finds

$$\begin{aligned} \sigma_{ij}\Big|_{U=0} &= \frac{i}{|\ell_1 \wedge \ell_2|} G_{ij}(0) \\ &= -i \int_{\mathcal{B}} \frac{dk}{(2\pi)^2} \mathrm{Tr} \Big\{ P_-(k) \big[\partial_{k_i} P_-(k), \partial_{k_j} P_-(k) \big] \Big\}, \end{aligned}$$

as desired.

Acknowledgments. This review is based on a presentation given at the Istituto Lombardo, Accademia di Scienze e Lettere, in Milano (Italy) on May 5, 2016, as well as on the notes of a course given at the EMS-IAMP summer school in mathematical physics *Universality, Scaling Limits and Effective Theories*, held in Roma (Italy) on July 11-15, 2016. I gratefully thank the members of the Istituto Lombardo, Accademia di Scienze e Lettere, and in particular Dario Bambusi and Antonio Giorgilli, for their invitation to present my work on the occasion of an 'Adunanza' of the Istituto Lombardo. The research leading to these results has received funding from the European Research Council under the European Union's Horizon 2020 Programme, ERC Consolidator Grant UniCoSM (grant agreement n. 724939).

REFERENCES

[1] A. A. Abrikosov, L. P. Gorkov, I. E. Dzyaloshinski: *Methods of quantum field theory in statistical physics*, Prentice-Hall 1963.

[2] M. Aizenman, G. M. Graf: *Localization bounds for an electron gas*, J. Phys.A: Math.Gen. **31**, 6783 (1998).

[3] J. F. Allen, A. D. Misener: *Flow of liquid helium II*, Nature **141**, 75 (1938).

[4] P. W. Anderson: *The Theory of Superconductivity in the High-T_c Cuprates*, Princeton Series in Physics, Princeton Univ. Press 1997.

[5] M. H. Anderson, J. R. Ensher, M. R. Matthews, C. E. Wieman, E. A. Cornell: *Observation of Bose-Einstein condensation in a dilute atomic vapor*, Science **269**, n. 5221, 198-201 (1995).

[6] G. Antinucci, V. Mastropietro, M. Porta: *Universal edge transport in interacting Hall systems*, Comm. Math. Phys. **362**, 295-359 (2018).

[7] J. E. Avron: *Colored Hofstadter butterflies*, in: Multiscale methods in quantum mechanics, Trends Math. (Boston, MA: Birkhauser Boston, 2004), pp. 11-22

[8] J. E. Avron, R. Seiler: *Quantization of the Hall conductance for general, multiparticle Schrödinger Hamiltonians*, Phys. Rev. Lett. **54**, 259-262 (1985).

[9] J. E. Avron, R. Seiler, B. Simon: *Homotopy and quantization in condensed matter physics*, Phys. Rev. Lett. **51**, 51 (1983).

[10] J. Bardeen, L. N. Cooper, and J. R. Schrieffer: *Microscopic theory of superconductivity*, Phys. Rev. **106**, 162 (1957).

[11] J. Bardeen, L. N. Cooper, and J. R. Schrieffer: *Theory of superconductivity*, Phys. Rev. **108**, 1175 (1957).

[12] G. A. Battle, P. Federbush: *A note on cluster expansions, tree graph identities, extra 1/N! factors!*, Lett. Math. Phys. **8**, 55-57 (1984).

[13] J. G. Bednorz, A. Müller: *Possible high T_c superconducitivity in the Ba-La-Cu-O system*, Z. Phys. B Cond. Matt. **64**, 189-193 (1986).

[14] J. Bellissard, A. van Els, H. Schulz-Baldes: *The non-commutative geometry of the quantum Hall effect*, J. Math. Phys. **35**, 5373 (1994).

[15] N. N. Bogoliubov: *On the theory of superfluidity*, Izv. Akad. Nauk USSR **11**, 77 (1947). Eng. Trans. J. Phys. (USSR) **11**, 23 (1947).

[16] C. C. Bradley, C. A. Sackett, J. J. Tollett, R. G. Hulet: *Evidence of Bose-Einstein condensation in an atomic gas with attractive interactions*, Phys. Rev. Lett. **75**, 1687 (1995).

[17] J. B. Bru, W. A. de S.Pedra: *Lieb-Robinson Bounds for Multi-Commutators and Applications to Response Theory*, Springer Briefs in Mathematical Physics **13**, Springer, 2016.

[18] J. B. Bru, W. A. de S.Pedra: *Universal bounds for large determinants from non-commutative Hölder inequalities in fermionic constructive quantum field theory*, Math. Mod. Meth. Appl. Sc. **12** (27) (2017).

[19] D. C. Brydges: *A short course on cluster expansions*, Phénomènes critiques, systèmes aléatoires, théories de jauge, Part I, II (Les Houches, 1984), 129-183, North-Holland, Amsterdam, 1986.

[20] D. C. Brydges, P. Federbush: *A new form of the Mayer expansion in classical statistical mechanics*, J. Math. Phys. **19**, 2064-2067 (1978).

[21] D. Brydges, T. Kennedy: *Mayer expansions and the Hamilton-Jacobi equation*, J. Stat. Phys. **48**, 19-49 (1987).

[22] S. Coleman, B. Hill: *No more corrections to the topological mass term in QED3*, Phys. Lett. B. **159**, 184 (1985).

[23] M. Correggi, A. Giuliani, R. Seiringer: *Validity of spin wave theory for the quantum Heisenberg model*, Europhys. Lett. **108**, 20003 (2014); and *Validity of the spin-wave approximation for the free energy of the Heisenberg ferromagnet*, Comm. Math. Phys. **339**, 279-307 (2015).

[24] K. B. Davis, M.-O. Mewes, M. R. Andrews, N. J. van Druten, D. S. Durfee, D. M. Kurn, W. Ketterle: *Bose-Einstein condensation in a gas of sodium atoms*, Phys. Rev. Lett. **75**, 3969 (1995).

[25] G. Gentile, V. Mastropietro: *Renormalization group for one-dimensional fermions. A review on mathematical results*, Phys. Rep. **352** 273-437 (2001).

[26] A. Giuliani, I. Jauslin, V. Mastropietro, M. Porta: *Topological phase transitions and universality in the Haldane-Hubbard model*, Phys. Rev. B **94**, 205139 (2016).

[27] A. Giuliani, V. Mastropietro: *The two-dimensional Hubbard model on the honeycomb lattice*, Comm. Math. Phys. **293**, 301-346 (2010).

[28] A. Giuliani, V. Mastropietro, M. Porta: *Absence of interaction corrections in the optical conductivity of graphene*, Phys. Rev. B **83**, 195401 (2011); and *Universality of conductivity in interacting graphene*, Comm. Math. Phys. **311**, 317-355 (2012).

[29] A. Giuliani, V. Mastropietro, M. Porta: *Universality of the Hall Conductivity in Interacting Electron Systems*, Comm. Math. Phys. **111**, 1107-1161 (2017).

[30] A. Giuliani, V. Mastropietro, M. Porta: *Quantization of the interacting Hall conductivity in the critical regime*, J. Stat. Phys. (2019). https://doi.org/10.1007/s10955-019-02405-1.

[31] C. Golschmidt, D. Ueltschi, P. Windridge: *Quantum Heisenberg models and their probabilistic representations*, Contemp. Math. **552**, Entropy and the quantum, II, R. Sims and D. Ueltschi, Ed. 2011.

[32] F. D. M. Haldane: *Model for a quantum Hall effect without Landau levels: condensed-matter realization of the 'Parity Anomaly'*, Phys. Rev. Lett. **61**, 2015 (1988).

[33] E. Hall: *On a New Action of the Magnet on Electric Currents*, American Jour. Math. **2** (3), 287-292 (1879).

[34] B. I. Halperin: *Quantized Hall conductance, current-carrying edge states, and the existence of extended states in a two-dimensional disordered potential*, Phys. Rev. B **25**, 2185 (1982).

[35] M. B. Hastings: *The stability of free Fermi Hamiltonians*, J. Math. Phys. **60**, 042201 (2019).

[36] M. B. Hastings and S. Michalakis: *Quantization of Hall Conductance for Interacting Electrons on a Torus*, Comm. Math. Phys. **334**, 433-471 (2015).

[37] I. F. Herbut, V. Juricic, and O. Vafek: *Coulomb Interaction, Ripples, and the Minimal Conductivity of Graphene*, Phys. Rev. Lett. **100**, 046403 (2008); V. Juricic, O. Vafek, and I. F. Herbut: *Conductivity of interacting massless Dirac particles in graphene: Collisionless regime*, Phys. Rev. B **82**, 235402 (2010); I. F. Herbut, V. Juricic, and O. Vafek: *Comment on 'Minimal conductivity in graphene: Interaction corrections and ultraviolet anomaly' by Mishchenko E. G*, e-print arXiv:0809.0725.

[38] D. R. Hofstadter: *Energy levels and wave functions of Bloch electrons in rational and irrational magnetic fields*, Phys. Rev. B **14**, 2239-2249 (1976).

[39] P. Kapitza: *Viscosity of liquid helium below the λ-point*, Nature **141**, 74 (1938).

[40] K. von Klitzing, G. Dorda, M. Pepper: *New method for high-accuracy determination of the fine-structure constant based on quantized Hall resistance*, Phys. Rev. Lett **45**, 494 (1980).

[41] R. B. Laughlin: *Quantized Hall conductivity in two dimensions*, Phys. Rev. B **23**, 5632(R) (1981).

[42] R. B. Laughlin: *Anomalous Quantum Hall Effect: An Incompressible Quantum Fluid with Fractionally Charged Excitations*, Phys. Rev. Lett. **50**, 1395 (1983).

[43] E. H. Lieb, D. W. Robinson: *The finite group velocity of quantum spin systems*, Comm. Math. Phys. **28**, 251-257 (1972).

[44] E. H. Lieb, R. Seiringer, J.-P. Solovej, J. Yngvason: *The Mathematics of the Bose Gas and Its Condensation*, Oberwolfach Seminars **34**, Birkhauser, Basel (2005).

[45] V. Mastropietro, M. Porta: *Spin Hall insulators beyond the Helical Luttinger model*, Phys. Rev. B **96**, 245135 (2017).

[46] B. Nachtergaele, Y. Ogata, R. Sims: *Propagation of correlations in quantum lattice systems*, J. Stat. Phys. **124**, 1-13 (2006); B. Nachtergaele, R. Sims: *Lieb-Robinson bounds in quantum many-body physics*, Contemp. Math. **529**, 141-176 (2010).

[47] J. W. Negele, H. Orland: *Quantum many-particle systems*, Addison-Wesley Pub. Co., 1988.

[48] H. K. Onnes: *The resistance of pure mercury at helium temperatures*, Commun. Phys. Lab. Univ. Leiden **12**, 120 (1911).

[49] W. A. de S. Pedra, M. Salmhofer: *Determinant Bounds and the Matsubara UV Problem of Many-Fermion Systems*, Comm. Math. Phys. **282**, 797-818 (2008).

[50] H. L. Stormer, D. C. Tsui, A. C. Gossard: *The fractional quantum Hall effect*, Rev. Mod. Phys. **71**, S298 (1999).

[51] D. J. Thouless: *Localisation and the two-dimensional quantum Hall effect*, J. Phys. C **14**, 3475 (1981).

[52] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs: *Quantized Hall Conductance in a Two-Dimensional Periodic Potential*, Phys. Rev. Lett. **49**, 405 (1982).

[53] D. C. Tsui, H. L. Stormer, A. C. Gossard: *Two-dimensional magnetotransport in the extreme quantum limit*, Phys. Rev. Lett **48**, 1559 (1982).

Istituto Lombardo (Rend. Scienze) 150, 117-131 (2016)

"IN PROIEZIONE": SPAZIO E IMMAGINE IN *LAS MENINAS* DI DIEGO VELÁZQUEZ

LUIGI COCCHIARELLA (*)

Nota presentata dal m.e. G. Caglioti
(Adunanza del 25 maggio 2016)

SUNTO. – Lo studio qui presentato riguarda l'analisi spaziale del notissimo capolavoro pittorico *Las Meninas*, di Diego Velázquez. Estetica, forma e simbolismo sono in quest'opera indissolubilmente legati al pattern geometrico che ne struttura l'immagine in funzione della posizione del punto di vista designato, obbedendo severamente ai dettami della prospettiva lineare che si ripropone, gemma nella gemma, perfino nel riflesso speculare, a cui è sorprendentemente affidata la rivelazione di una delle principali chiavi di lettura della scena. La sofisticata costruzione geometrica e la sublime esecuzione grafica cooperano in quest'opera a sancire il ruolo *creativo* e non meramente *imitativo* della pittura, sorpresa nell'atto stesso di allestire la realtà pittorica, tutto questo al culmine della cosiddetta *Età d'Oro* dell'arte Spagnola.

***

ABSTRACT. – This study focuses on the spatial analysis of *Las Meninas*, the worldwide well-known pictorial art masterpiece by Diego Velázquez. Aesthetics, form, and symbolism, are here interlinked with a geometric pattern dependent on the exact location of the designated viewpoint, which severely obeys to the projective rules of linear perspective, where, as a gem in the gem, even the image of a mirror appears, to masterly highlight the final moral of the story, surprisingly explicated by the appeal to reflection. Sophisticated geometry feature and sublime graphic execution work together in this painting to establish the role of pictorial art as a *creative* besides *imitative* activity, representing painting itself in the very act of setting a pictorial reality, all this at the height of the so-called *Golden Age* of the Spanish art.

(*) Politecnico di Milano, Dipartimento di Architettura e Studi Urbani (DASTU), Italia. E-mail: luigi.cocchiarella@polimi.it

La scena

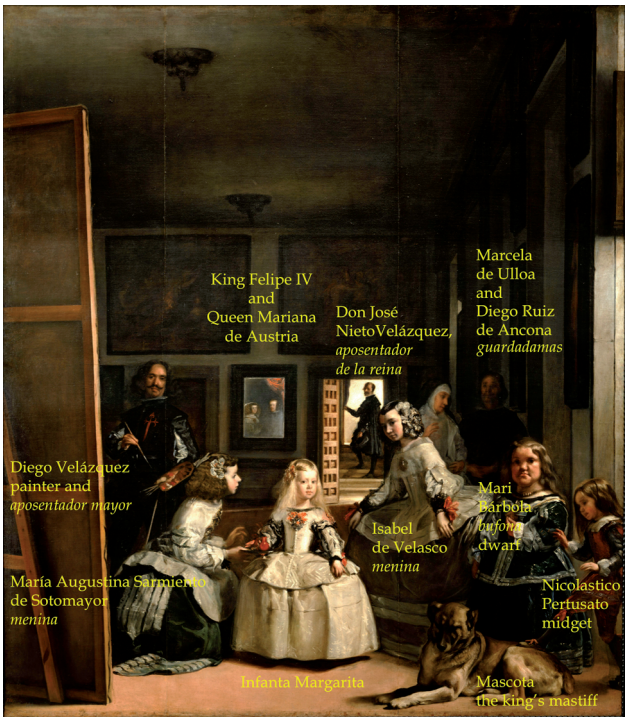


Fig. 1. Las Meninas*, di Diego Velázquez, 1656, olio su tela, cm 276x318, Madrid, Museo del Prado (fonte web): la scena e lo spazio (annotazioni dell'autore).*

Dipinta ad olio su tela da Diego Rodríguez Velázquez de Silva y Velázquez nel 1656, la monumentale opera pittorica oggi conosciuta come *Las Meninas*, mostra una scena all'apparenza imperscrutabile, ambientata in un pigro pomeriggio all'interno della Galería de Mediodía dell'Alcázar de Madrid, l'atelier del pittore situato all'interno del Palazzo Reale dell'epoca. Si tratta di una scena abbastanza densamente popolata, come mostra la *Fig. 1.* A partire dal fondo, troviamo il re Felipe IV e la regina Mariana de Austria, sposata in seconde nozze, le cui sembianze si riconoscono riflesse nel piccolo specchio collocato al centro della parete. Poco più a destra Don José Nieto Velázquez, *aposentador de la reina*, attendente personale della regina, appare in attesa sui gradini della rampa di accesso alla sala, pronto a preparale il passag-

gio. A circa metà sala, intenti in riservata conversazione, troviamo Marcela de Ulloa e Diego Ruiz de Ancona, *guardadamas*, responsabili delle damigelle d'onore, *las meninas* appunto, segnatamente Maria Augustina Sarmiento de Sotomayor e Isabel de Velasco, ritratte rispettivamente a sinistra e a destra, per chi osserva il dipinto, della Infanta Margarita, la giovanissima principessa, figlia dei reali di Spagna, la cui figura risplende luminosa al centro della scena. Più a destra scorgiamo Maria Barbola e Nicolastico Pertusato, una nana e uno gnomo, figure sovente di casa nelle corti reali, la cui particolare fisiognomica risulta all'attenzione del pittore anche in altre opere. In primissimo piano, quasi a presidiare la scena, sonnacchioso e rassicurante, ecco Mascota, lo splendido e fedele mastino del re. Un po' più indietro, sulla sinistra, subito oltre la grande tela di cui distinguiamo l'intelaiatura posteriore, la sagoma del pittore Diego Velázquez, *aposentador major*, maresciallo del castello, maestro delle cerimonie e attendente personale del re, si staglia imperiosa per altezza e fiera per postura sui presenti, nell'atto di apportare gli ultimi ritocchi all'opera in esecuzione sulla grande tela, e che non ci è dato vedere direttamente. Tutti i protagonisti sono rivolti verso chi osserva il dipinto, e non sembrano 'in posa' ma piuttosto intenti a perscrutare lo spazio 'al di qua' della cornice, ovvero il nostro stesso spazio, e infine noi stessi, con manifesta curiosità e un certo stupore. Osservando più in alto, ci si rende conto che la metà superiore del dipinto rappresenta uno spazio sostanzialmente vuoto in cui domina la superficie del lungo soffitto dell'alta sala, su cui si evidenziano ampie macchie di umidità, ironico paradosso per un luogo che sarà invece distrutto dal fuoco la notte di Natale del 1734. Sulle pareti sono riprodotti con precisione i dipinti realmente presenti nella sala al momento dell'esecuzione dell'opera. A destra, in alto e tra i vani delle grandi imposte finestrate, si riconoscono opere di Juan Batista del Mazo, genero di Velázquez, molte di esse essendo copie di opere di Rubens. Sempre in alto, sul fondo, campeggiano due imponenti tele, eseguite ancora da Mazo da originali di Rubens. Si tratta delle scene più significative: a sinistra *Minerva e Aracne*, perduta nell'incendio del 1734, a destra *Apollo e Marsia*, oggi al museo del Prado. Con straordinaria maestrìa la luce pervade l'intera scena con felpata morbidezza, accentuando i caratteri e stabilendone gerarchie per mezzo di sapientissimi bilanciamenti cromatici, che culminano nella fulgida immagine della principessa Margarita, a cui il regno di Spagna e gli stessi regnanti, a valle di dolorose vicende familiari, rivolgono ora affetto e speranze.

L'impianto prospettico

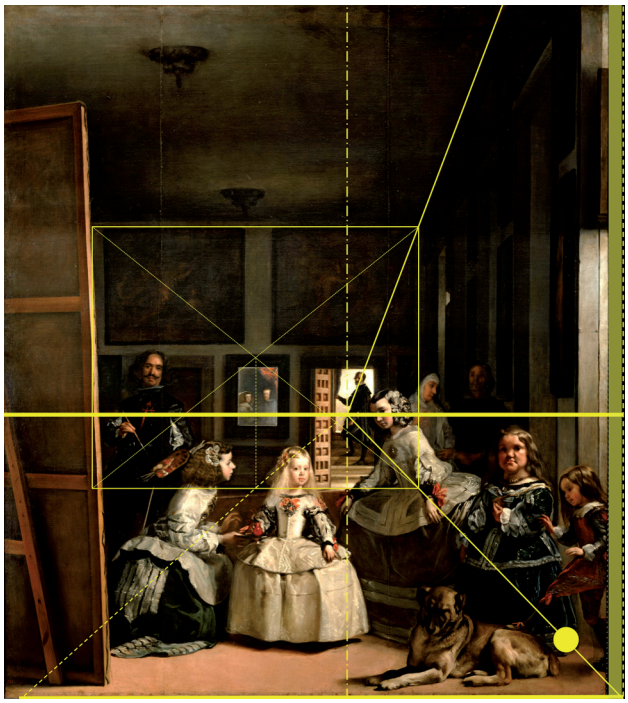


Fig. 2. Identificazione dell'impianto prospettico (diagramma dell'autore).

Tecnicamente l'impianto prospettico è di tipo classico, col punto principale di osservazione sensibilmente decentrato e il conseguente effetto di un potente dinamismo percettivo, come si può vedere nella *Fig. 2*.

Nell'identificazione del pattern proiettivo si è tenuto conto della rifilatura operata sul lato destro della tela a seguito dei danni patiti durante l'incendio dell'Alcázar de Madrid, il cui ingombro è graficamente evidenziato in *Fig. 2*. La disposizione dei personaggi all'interno della scena propone una quadripartizione del registro delle profondità, la cui suddivisione può considerarsi sostanzialmente scandita dai gruppi di soggetti già elencati nella precedente sezione. Vi è quindi un primissimo piano, fra il bordo della tela e la principessa con le damigelle d'onore. Più in là la porzione di spazio in cui stanno l'autore e i due *guardadamas*. Un terzo registro si estende fino alla parete di fondo. Insondabile, e tuttora non completamente chiaro nella effettiva configurazione dell'epoca, è infine

lo spazio accessibile al di là di quella parete, visibile attraverso lo squarcio dell'uscio e caratterizzato da una luminosità quasi teatrale per la suggestione creata dalla pesante tenda accostata allo stipite dal Nieto per favorire il passaggio della regina. Sappiamo che in tale regione al di là della soglia di fondo era collocato, lungo la parete meridionale, la stessa che appare a destra nel dipinto, l'ingresso principale dell'Alcázar de Madrid. In tale sistema di sequenze prospettiche Diego Velázquez occupa uno spazio autonomo, non risultando di fatto la sua figura allineata con alcuno dei gruppi ritratti. A ben vedere, inoltre, il perimetro della parete di fondo, evidenziato in *Fig. 2*, definisce un sottoinsieme scenico quasi autonomo, il cui perimetro include idealmente i soli membri principali della famiglia reale, significativamente sovrastati dalle effigi delle due grandi scene mitologiche, dedicate a Minerva e ad Apollo. Il dinamismo prospettico è altresì amplificato dal ritmo complessivo della composizione pittorica e dalle linee di forza che essa sembra idealmente evocare, alcune delle quali sono in via d'ipotesi evidenziate nella *Fig. 3*.

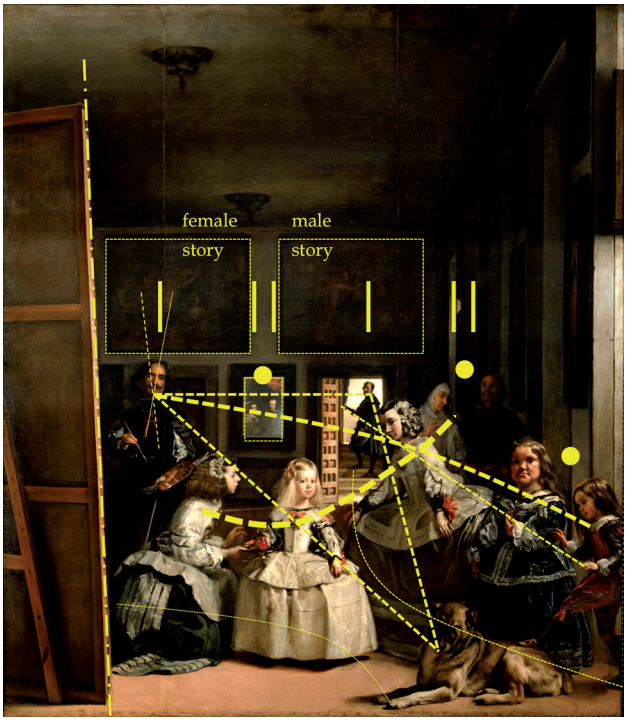


Fig. 3. L'impianto compositivo della scena (diagrammi dell'autore).

L'apparente perfezione della struttura prospettica ha dunque suggerito di verificare la corrispondenza fra l'immagine dipinta e l'effettivo spazio fisico della sala rappresentata, proposito di non facile perseguimento trattandosi di un ambiente ormai demolito, insieme agli altri spazi dell'Alcázar a seguito dell'incendio del 1734, per far posto all'attuale Palazzo Reale di Madrid. Il dipinto non offre inoltre, volutamente, i più tradizionali riferimenti grafici per la determinazione dei valori di scorcio, come ad esempio il disegno della pavimentazione, elementi un tempo utili, in primis agli artisti per la corretta esecuzione prospettica, e tuttavia non più necessari all'arte pittorica ormai pienamente matura del '600. Si è quindi dovuto prendere in considerazione il contributo di ulteriori materiali documentali, fra i quali di particolare ausilio è risultato il rilievo grafico del castello, eseguito nel 1626 da Juan Gomez de Mora in vista delle trasformazioni che avrebbero condotto alla configurazione degli spazi in essere al tempo in cui il dipinto fu eseguito. Pur risultando non sempre della massima precisione grafica i dettagli, le planimetrie hanno offerto un valido supporto alla comprensione delle proporzioni degli ambienti, nonché della generale organizzazione compositiva della fabbrica. Questi documenti grafici, unitamente alla più recente ed esaustiva ricostruzione storica delle vicende costruttive dell'Alcázar de Madrid pubblicata da José Manuel Barbeito nel 1992, hanno fornito le principali basi per la ricostruzione grafica della *Galería de Mediodía*, la sala rappresentata nel dipinto. A integrazione e completamento dell'indagine, ci si è infine riferiti a un *Inventario* dettagliato delle opere in essa presenti, risalente al 1686, che ha consentito di identificare con precisione i dipinti presenti nella sala, la loro collocazione e la loro dimensione, contribuendo a una più attendibile verifica delle ipotesi formulate. Sulla base dei dati metrici desunti, previa opportuna conversione dell'unità di misura originaria in *pie castellano* nel nostro sistema metrico, si è quindi proceduto all'estrazione di alcuni parametri angolari ricavati sulle due giaciture principali, orizzontale e verticale selezionando, in particolare, quelli evidenziati in *Fig. 4*, al fine di impostare correttamente il riferimento geometrico per la ricostruzione grafica dello spazio tramite prospettiva inversa. Tale procedimento si basa sostanzialmente sul preliminare tracciamento, sull'immagine data, di rette-traccia di piani noti e dei relativi circoli sottendenti angoli noti in quei piani, dalla cui interazione risulta possibile ricostruire la posizione esatta del centro di proiezione, ovvero la collocazione fisica dell'osservatore rispetto allo spazio rappresentato. La conoscenza di questo para-

metro fornisce infine la chiave di accesso alla procedura grafica per la ricostruzione della 'vera forma' dello spazio a partire dall'immagine prospettica.

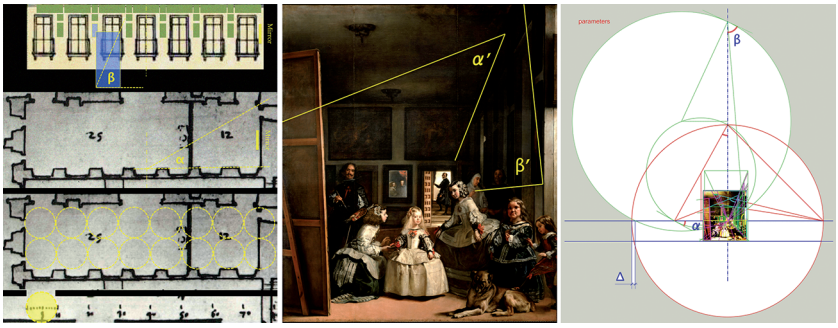


Fig. 4. Identificazione e valutazione dei principali indizi metrici utili per la ricostruzione prospettica (rielaborazioni grafiche dell'autore).

Ricostruzione dello spazio scenico

Impostato il riferimento prospettico sulla base dei dati metrici desunti come si è detto, si è proceduto alla ricostruzione grafica dello spazio in pianta e in alzato, mettendo in relazione i due registri proiettivi: quello prospettico, basato sulla proiezione conica, più fedele alla visione fisiologica ma con forti distorsioni metriche e angolari, e quello ortografico, basato sulla proiezione cilindrica, meno fedele alla visione fisiologica ma assai comodo per via della conservazione del parallelismo e di alcune proporzioni geometriche nell'immagine risultante. In particolare, mediante opportune costruzioni basate sul principio dell'omologia, la trasformazione che regola gli spazi proiettivi nel rispetto dei relativi invarianti fondativi, si è proceduto alla determinazione grafica delle principali viste ortogonali, riproiettandone in versione 'raddrizzata' e opportunamente ridimensionata, gli elementi presenti nelle corrispondenti sezioni del dipinto, ed estraendone, come in un tour virtuale, vedute inedite, riproposte da angolazioni non previste dall'autore, e tuttavia utili allo studio dell'opera. La *Fig. 5* mostra un dettaglio della ricostruzione grafica di due alzati e due viste orizzontali.

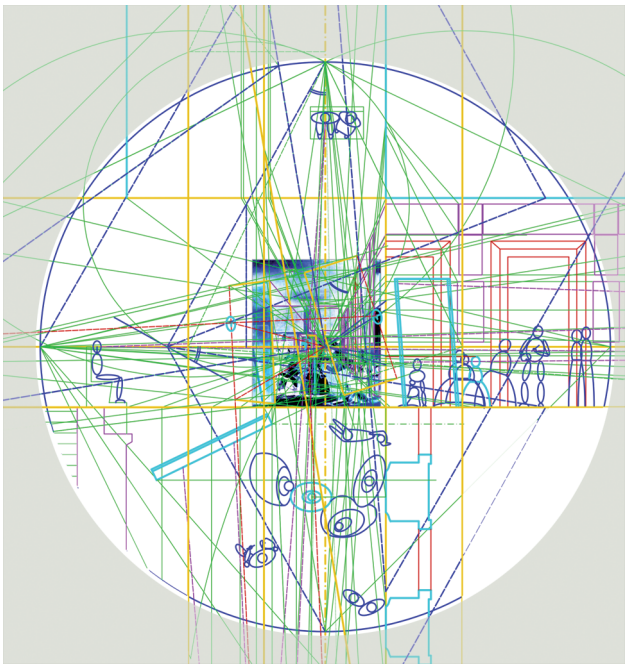


Fig. 5. Un dettaglio delle ricostruzioni grafiche in prospettiva inversa, eseguite mediante omologia di ribaltamento (grafico dell'autore).

Al termine delle operazioni di restituzione grafica le informazioni sono state ricomposte in un modello tridimensionale della scena, cui si riferiscono la proiezione planimetrica e l'alzato presentati in *Fig. 6*, per una più diretta valutazione geometrica della scena e delle traiettorie ottico-proiettive che ne strutturano l'impianto iconografico.

Particolare attenzione è stata rivolta allo studio dell'immagine riflessa dei regnanti di Spagna nel piccolo specchio collocato in fondo alla sala, che, come si anticipava, custodisce una delle più potenti chiavi di lettura dell'opera. Si scopre così che il dipinto è idealmente eseguito per rappresentare la scena dalla posizione di osservazione che il re stesso, Felipe IV, occupava nel momento in cui la scena si svolgeva. Una posizione decentrata dunque, trovandosi in quel momento egli e la regina Mariana nel semispazio *al di qua* del quadro, in prossimità dell'accesso opposto a quello visibile sul fondo con l'*aposentador* Nieto in attesa del passaggio dei reali di Spagna. Si comprendono allora gli

sguardi attenti dei presenti, e massimamente quello attentissimo, severo e concentrato di Velázquez, intento a ultimare il ritratto del suo sovrano sulla grande tela, probabilmente destinata, come *Las Meninas*, alla collezione privata della casa reale. Osservando il quadro dal corretto punto di traguardo risultiamo, allora, noi stessi ogni volta ricollocati in 'posizione regale', per il favore dell'arte pittorica capace di concedere anche a noi, per sempre, questo eccezionale privilegio. Rimane l'ultimo enigma, riguardante l'immagine riflessa nello specchio. Da dove proviene quell'immagine? Come le traiettorie ottiche ricostruite indicano, quell'immagine riflette l'immagine appena dipinta da Diego Velázquez sul lato della tela che non possiamo osservare direttamente, ma che essendo opportunamente rivolta verso lo specchio, attraverso questo si riflette, risultando visibile dal punto di vista designato. Ovvero al Re in persona, che poteva così seguire direttamente il lavoro. Il diagramma presentato in *Fig. 7* mostra alcuni particolari dello studio delle traiettorie di riflessione, eseguite in sovrapposizione grafica rispetto all'immagine originale, unitamente ad alcune costruzioni ausiliari.

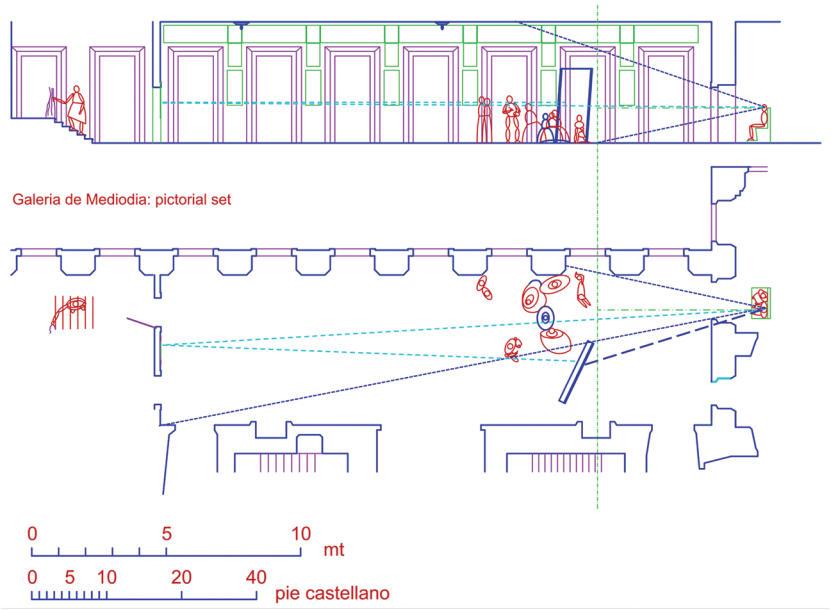


Fig. 6. Ricostruzione ortografica dello spazio scenico e valutazione dei parametri ottico-proiettivi (disegni dell'autore).

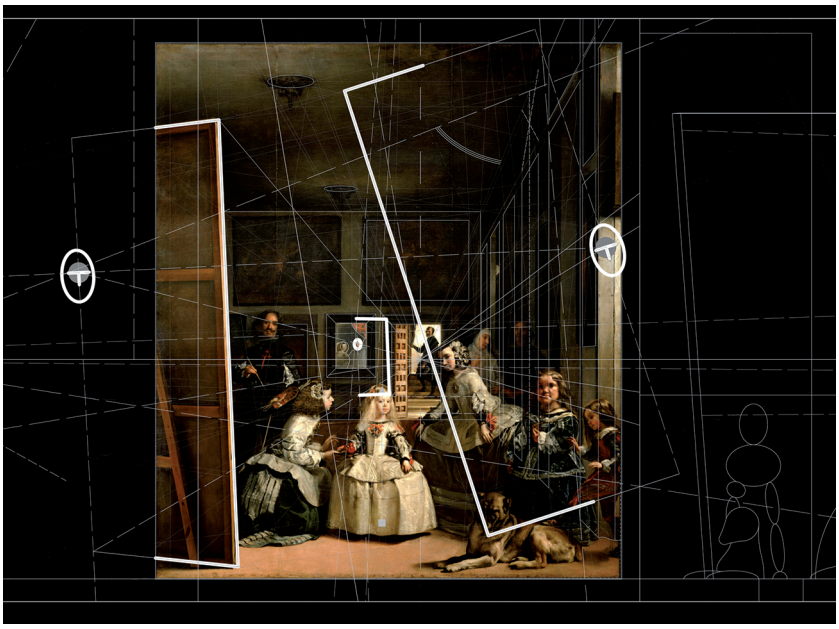


Fig. 7. L'enigma dello specchio attraverso l'analisi grafica delle relazioni omologiche che regolano le traiettorie di riflessione (elaborazione grafica dell'autore).

Corre l'obbligo di segnalare che le ipotesi qui proposte e le relative costruzioni su cui si basano, sono state eseguite non già su una copia del dipinto prodotta *in loco* e *ad hoc*, bensì su un'immagine selezionata fra quelle disponibili in rete, successivamente comparata con altre riproduzioni, a stampa e digitali. I risultati risentono inoltre della diversa natura grafica delle linee pittoriche fisicamente riconoscibili nel dipinto rispetto alle pressoché incorporee linee di ricalco tracciate in ambiente digitale. Fatte salve queste e altre specifiche approssimazioni, lo studio ha confermato l'esistenza di una precisa costruzione prospettica nella scena dipinta, circostanza del resto attesa, anche in considerazione dell'interesse che l'artista nutriva per la geometria e l'ottica, come dimostrano i 154 testi di matematica, geometria, ottica, astronomia, includendo i principali trattati del tempo, unitamente a molti strumenti grafici e a una nutrita serie di specchi, ritrovati alla sua morte nella sua biblioteca. Naturalmente non si esclude la possibilità che grazie alla sua raffinatissima maestrìa Velázquez abbia effettivamente tracciato 'di getto' sulla tela le linee caratterizzanti la

complessa struttura prospettica del quadro, in ogni caso la coerenza metrico-proiettiva fin qui illustrata permane, verificabile 'con riga e compasso', a testimonianza del profondo legame fra rigore scientifico e sensibilità artistica caratterizzante quest'opera, in linea col sentimento dominante nella cultura europea, non solo artistica, del tempo.

Conclusioni

Lo stupefacente risultato estetico della composizione si basa dunque e parimenti sulla straordinaria maestria pittorica di Diego Velázquez e sull'intelligente impiego dello statuto prospettico, il cui impianto geometrico sostiene organizzazione ed espressione grafica al pari, si potrebbe dire, della grammatica nella poetica o della matematica nella formalizzazione scientifica. La sapientissima articolazione a tutto campo delle componenti tecniche e visuali del linguaggio pittorico consente al Maestro di superare il registro denotativo, e di costruire su tela un vero e proprio teorema celebrativo dell'autorevolezza della pittura e del ruolo del pittore nell'esplicitare, manifestandolo per via iconica, il nesso fra i registri umano e divino, di cui la condizione regale è insostituibile mediatrice e garante. Lo dichiara l'immagine di re Felipe IV e della regina Mariana, riflessa sullo sfondo nello specchio situato proprio al di sotto delle grandi tele narranti leggende di umani irrimediabilmente sopraffatti nel tracotante atto di sfidare gli dei. Con gran precisione, *Minerva e Aracne*, scena di genere femminile, sovrasta il quarto sinistro, riferendosi evidentemente al profilo riflesso della regina, *Apollo e Marsia*, scena di genere maschile, sovrasta il quarto destro, corrispondendo dunque al profilo riflesso del re. Ma perché lo specchio? Al di là della ricorrenza di simili spiazzanti virtuosismi ottico-prospettici nella produzione pittorica di Velázquez, l'ispirazione sembra rimontare anche al mito e alla tradizione letteraria del tempo. Affidabile quanto inclemente, risorsa decisiva nel compimento della vicenda di Perseo e Medusa, lo specchio appare infatti anche in altri capolavori pittorici fra il '400 e il '600, basti pensare al *Ritratto dei coniugi Arnolfini* di Jan van Eyck (siamo ancora nel 1434), o alla *Venere Rokeby* dello stesso Velázquez, tuttavia in questi e in altri esempi l'immagine riflessa è rappresentata insieme alla figura che le corrisponde. In *Las Meninas* invece, l'immagine riflessa non ha ulteriori referenti visibili nel dipinto, se non, come si è ipotizzato, il ritratto dei reali di Spagna in esecuzione sulla grande tela, che tuttavia non possiamo vedere direttamente.

Il ricorso alla rappresentazione dei sovrani mediante la sola immagine riflessa, al di là della ritrosìa del re a comparire ritratto più nitidamente e più in grande in ragione della sua non più giovane età, risiederebbe altresì nella metaforica adesione a un genere letterario allora in auge, noto in spagna come *espejo de principe*, letteralmente lo specchio del principe, riguardante una serie di opere e libelli dedicati alla educazione del buon sovrano, sull'onda di una tradizione classica che rimonta a Isocrate, in particolare all'opera *A Nicocle*, risalente al IV secolo A.C. e ripresa in seguito da vari autori europei, specie nel '500, fra i quali l'italiano Nicolò Machiavelli e gli spagnoli Baldasar Castiglione e Balthasar Gracian. Il riferimento più diretto per Velázquez sarebbe tuttavia il volume *Idea de un principe politico Christiano* pubblicato nel 1640 da Don Diego de Saavedra Fajardo, ministro e ambasciatore di Spagna, dunque persona vicinissima al Re, nonché Cavaliere dell'*Orden de Santiago*, l'antico e potente ordine monastico-militare fondato nel XII secolo (*Fig. 8*).



Fig. 8. Il frontespizio del testo Idea de un principe politico Christiano, pubblicato nel 1640 da Don Diego de Saavedra Fajardo (fonte: web).

È chiaro a questo punto il principale intento celebrativo dell'opera, la cui composizione gravita intorno allo splendido ritratto della Infanta Margarita, con i più fulgidi auspici per il futuro del regno di Spagna. Regno al momento retto dai migliori regnanti, come pare sottolineare la scelta di rappresentarli per pura immagine riflessa, dunque in forma pressoché idealizzata, non contaminata dalla realtà e altresì sottoposta alla sacra egida di Minerva patrona di saggezza e di Apollo patrono d'arte. Quasi a voler esprimere l'essenza stessa del regnare in quanto supremo modello e precetto per l'adorata Infanta. Ma un ulteriore anelito pervade l'opera. L'adamantino realismo delle forme, il sotteso intreccio di metafore e simbolismi, l'affascinante articolazione prospettica impreziosita dal gioco della riflessione speculare, la sacralità dell'azione suggerita dalla diffusa presenza del mito, l'intensità espressiva della scena, colta nell'apice dell'estremo stupore da parte degli astanti un attimo prima che la tensione scenica si dissolva, ebbene questo straordinario spettacolo è ideato e allestito dall'artista, eseguito per sua mano e concepito per sua intelligenza. Esso esiste in quanto è dipinto, sancendo pure in pittura il primato della *creazione* sull'*imitazione*. Si tratta di un punto cruciale, legato anche alla personale vicenda del Maestro. Al tempo del dipinto già *Pintor de camera* e *Ayudo de camera* del re, *Aposentador de palacio* e *Superintendente de obras particulares*, Velázquez era ormai in attesa da più di cinque anni di essere insignito del titolo di Cavaliere dell'*Orden de Santiago*, a completamento della sua già straordinaria carriera presso la corte di Spagna. La richiesta di attribuzione del titolo onorifico, tenacemente perorata fin dal 1650 e fortemente respinta per quasi un decennio dall'Ordine per via della mancata origine nobiliare del richiedente, sarà finalmente accolta, dopo due intercessioni papali e l'intervento personale del re, soltanto nel 1659, pochi mesi prima della morte dell'artista. Leggenda vuole che la croce dell'Ordine, di colore rosso, che nel quadro vediamo segnata sul corpetto di Velázquez, sia stata dipinta per mano di Felipe IV medesimo dopo la morte dell'amato artista e fedele ufficiale, non essendo ancora questi, come già si diceva, insignito del titolo onorifico di cavaliere dell'*Orden de Santiago* nel 1656, l'anno in cui l'opera, che oggi conosciamo col titolo *Las Meninas*, fu realizzata. Potremmo quindi concludere che il finissimo ricorso a molteplici livelli semantici e simbolici, magistralmente articolati nella non banale adesione a un registro fedelmente realistico, la perfetta integrazione fra articolazione scenica, composizione spaziale e intrecci ottico-prospettici nella straordinaria subli-

mazione pittorica, inducano a pensare che *Las Meninas* intendesse travalicare i limiti stessi del capolavoro artistico, proponendosi come colossale progetto di redenzione dell'arte pittorica. Significative le parole di Jonathan Brown a proposito di questo imponente capolavoro: "A painting as rich in ambiguity as it is in sublety, *Las Meninas* has long been recognized as a masterpiece of Western art, a pictorial tour de force rarely equalled and never surpassed. But when we attempt to explain its greatness, we soon realize how it seems to evade the grasp both of intuitive and rational understanding".

Ringraziamenti

Ringrazio il prof. Giuseppe Caglioti, professore emerito (fisica della materia), Politenico di Milano, per il gradito invito a presentare pubblicamente questo studio, e l'Istituto Lombardo Accademica di Scienze e Lettere per avermene offerto l'opportunità. Sono altresì grato al prof. Carlo Pagani per le indicazioni redazionali e il supporto editoriale.

BIBLIOGRAFIA

J.M. Barbeito, *El Alcázar de Madrid*, Colegio Oficial de Arquitectos de Madrid, Madrid, 1992.

J.M. Barbeito, Velázquez y la decoración excultórica del Alcázar, in: *Velázquez: esculturas para el Alcázar*, L. Alba Carcelén, 113-131, Real Academia de Bellas Artes de San Ferdinando, Madrid, 2007.

J. Brown, *Images and Ideas in seventeenth-century Spanish painting*, Princeton niversity Press, Princeton, N.J., 1978.

J. Brown, *Velázquez*, *Painter and Courtier*, Yale University Press, London, 1986.

K. Clark, *Looking at pictures*, Holt Rinehart and Winston, New York, 1960.

L. Cocchiarella, When Image sets Reality: Perspectival Alchemy in Velázquez's Las Meninas, *KoG. Scientific – Professional Journal of Croatian Society for Geometry and Graphics*, 2015: Vol.19: pp. 65-83.

M. Kemp, *The Science of Art. Optical themes in western art from Brunelleschi to Seurat*, Yale University Press, New Haven and London, 1990.

M. Miller Kahr, Velázquez and Las Meninas, *The Art Bulletin*, 1975: Vol.57(2): pp. 225-246.

J.R. Saerle, "Las Meninas" and the Paradoxes of Pictorial Representation, *Critical Inquiry*, 1980: Vol. 6(3): pp. 477-488.

A. Sgrosso, *Rigore scientifico e sensibilità artistica tra Rinascimento e Barocco*, Vol. 2 della serie *La geometria nell'immagine: storia dei metodi di rappresentazione*, UTET, Torino, 2001.

J. Snyder, "Las Meninas" and the Mirror of the Prince, *Critical Inquiry*, 1985: Vol. 11(4): 539-572.

A.S. Velázquez, Genius Loci. Anaparástasis del Alcázar vejo de Madrid a partir de la reconstrucción tridimensional de Las Meninas, *Cuadernos de los Amigos de los Museos de Osuna*, 2013: Vol. 15: pp. 138-147.

Source of the image of Velázquez's *Las Meninas*: wikipedia.org

Web source: en.m.wikipedia.org

Istituto Lombardo (Rend. Scienze) 150, 133-134 (2016)

DALLA SCOPERTA DEL DNA AD OGGI: LA MEDICINA DENTRO AL GENOMA[1]

Nota del s.c. ELENA CATTANEO (*) e del m.e. ORSETTA ZUFFARDI (**)

(Adunanza del 23 giugno 2016)

In passato, l'approccio al paziente si basava sulla storia clinica e sulla risposta ai trattamenti farmacologici. Oggi invece la definizione delle basi molecolari della malattia è preliminare all'identificazione di nuove strategie terapeutiche basate sulla specifica alterazione di quel gene o sul ruolo del gene in quel certo tessuto o organo. Se la genetica è stata in passato il fanalino di coda della medicina, paradossalmente oggi è la medicina a dipendere dalla genetica, appunto la medicina dentro al genoma.

Di fatto con le nuove modalità di sequenziamento del DNA che permettono di sequenziare interi genomi a bassi costi è diventato usuale cercare l'alterazione del genoma causativa della malattia, almeno per disordini a base strettamente genetica o per disordini cosiddetti multifattoriali dove la genetica svolge comunque un ruolo primario. Nonostante la disponibilità di laboratori dedicati al sequenziamento del DNA, l'interpretazione delle varianti rimane non banale. Infatti ciascun individuo sano contiene nel suo DNA centinaia di varianti il cui ruolo nella salute o nella malattia non è tuttora chiaro. Per superare questa difficoltà, il DNA del soggetto indice (cioè quello malato) viene sequenziato con quello dei genitori o di altri fratelli sia sani che malati. Le varianti identificate nel soggetto malato ma non nei genitori sani o le

(*) Università degli Studi di Milano, Italia. E-mail: elena.cattaneo@unimi.it
(**) Università degli Studi di Pavia, Italia. E-mail: orsetta.zuffardi@unipv.it

1 Sintesi dell'intervento.

varianti che sono presenti su ambedue gli alleli nel soggetto malato e su un singolo allele nei genitori sani diventano varianti candidate ad essere causative della malattia. L'ipotesi di gene "candidato" viene avvalorata attraverso la consultazione di database dedicati che indicano la frequenza della variante nella popolazione sana (tanto più è rara, tanto più è probabilmente causativa), le conseguenze sulla proteina codificata dal gene con la variante, e altri dati utili a definire se la variante abbia o meno un ruolo nella malattia in esame. Le varianti che interessano regioni genomiche che non codificano per proteine sono comunque di più difficile interpretazione.

Allo stato attuale, la conoscenza dell'alterazione genomica causativa sta aprendo le porte ai trattamenti per una serie di malattie, fra le quali le malattie tumorali. Una copia non alterata del gene può essere introdotta in una persona, di solito per mezzo di un vettore virale o il gene alterato può essere silenziato prevenendone la trascrizione. Addirittura, esperimenti su animali fanno sperare che i disturbi genetici potranno essere riparati nell'utero, in modo che il neonato nasca libero da malattia.

Inoltre, la terapia genica potrà permettere di trattare anche alcune infezioni virali, utilizzando il DNA virale per indurre l'organismo a creare gli anticorpi monoclonali per combattere l'agente patogeno.

La terapia genica rimane comunque un percorso medico costoso che richiederà nuovi modelli etici affinché questa procedura non sia aggio di poche classi sociali o di pochi paesi con sistemi sanitari equi.

Istituto Lombardo (Rend. Scienze) 150, 135-146 (2016)

NUMERICAL SOLUTION OF BLOOD DYNAMICS IN ATHEROSCLEROTIC CAROTIDS

CHRISTIAN VERGARA (*)

Nota presentata dal m.e. Carlo Domenico Pagani
(Adunanza del 23 giugno 2016)

SUNTO. – In questo lavoro verrà discussa la modellazione matematico-numerica della fluido-dinamica del sangue in carotidi aterosclerotiche. Verranno richiamati i concetti di base per la ricostruzione delle geometrie computazionali, la modellazione matematica del problema di interazione-fluido struttura e una possibile strategia per la sua risoluzione numerica. Verranno inoltre presentati alcuni risultati numerici ottenuti per un caso sia prima che dopo la rimozione chirurgica della placca.

***

ABSTRACT. – In this work we present the mathematical and numerical modeling of blood fluid-dynamics in atherosclerotic carotids. We review how to obtain the computational domain, the mathematical formulation of the fluid-structure interaction problem, and a possible strategy for its numerical discretization. We also report some numerical results for a case both before and after the surgical removal of the plaque.

1. INTRODUCTION

Atherosclerosis is the leading cause of death in Western countries and consists in the formation of atherosclerotic plaques in arteries of medium or large size, with possible obstruction of the vessel and formation of blood clots [18]. Under normal conditions, this process is a protective response of the endothelium and muscle cells of the artery to stress, and consists in the formation of fibrous lesions. When this

(*) LABS - Dipartimento di Chimica, Materiali e Ingegneria Chimica "Giulio Natta" - Politecnico di Milano, Italia. E-mail: christian.vergara@polimi.it

response becomes excessive as a result of prolonged and high fluid viscous stresses, a matrix of connective tissue grows up and lipids begin to accumulate. This accumulation can lead to the formation of a plaque, that obstructs the vessel (stenosis), reducing its lumen, and could degenerate leading to a complete obstruction of the vessel and to a stroke.

The plaque may be calcified, fibrous or lipid, depending on its composition, and mostly develops at the bifurcation sites, such as the carotid one. The carotids are two arteries located at the level of the neck and are crucial to maintain the human life, since they supply the blood to the brain. In the case of a lipid plaque in the carotids, the thin fibrous cap may crumble inducing embolization of fragments into the brain tissue.

The mechanisms of formation and progression of the atherosclerotic plaque are related to several factors, such as vessel morphology and vascular cell biology, interacting with each other. In particular, the blood fluid-dynamics seems to play a crucial role in the plaque development. Indeed, it has been extensively shown in the last thirty years that when vascular endothelial cells are subjected to stress, they change their polygonal shape and become elongated in the direction of the blood flow. A number of molecular and biological changes are mediated by shear stress exposure of endothelial cells, including changes in gene expression, production of vaso-active compound, adhesion molecules and changes in cell cytoskeleton. As a result of this biological response to flow, the intact endothelium is able to initiate a cascade of events leading to the thickening of the arterial wall and to plaque formation in regions of low and oscillating wall shear stress (WSS), whereas areas of vessel wall characterized by uniformly directed and laminar WSS are spared from development of vascular disease [11]. All these phenomena, induced by the blood fluid-dynamics, are highly correlated to the formation and progression of the plaque, and should be considered in the study of these clinical conditions, not only in general terms but at patient-specific level.

The mathematical and numerical modeling is nowadays a valuable tool for the study and the understanding of many physical phenomena. In particular, the use of numerical approximations of the equations describing these phenomena is a synergistic instrument (in some cases even alternative) to experimental studies. There are many areas in which the numerical modeling significantly supports the understanding of physical phenomena and the industrial design, such as aerodynamics, shipbuilding, geo-science, structural mechanics, electronics

and, most recently, the biomedical environment. In the context of patient-specific numerical simulations for the study of the cardiovascular system, the research has been active for at least twenty years, see, *e.g.*, [15]. In addition, the opportunities given by modern computers and high performance computing allow to achieve meaningful patient-specific results in a reasonable time. This has been also possible thanks to the great advances in recent years of the techniques to obtain biomedical data and imaging and the related techniques to reconstruct geometries and process data, which are a necessary starting point in view of the numerical studies.

In this work, we present the numerical solution of blood flow in carotids. To this aim we consider the fluid-structure interaction problem due to the compliant nature of the carotid vessel wall and its numerical discretization. In particular, we address the case of stenotic carotids, so that we need to introduce a suitable model for the plaque. We present some numerical results obtained in realistic geometries and with patient-specific boundary conditions, both before and after the plaque removal.

2. Generation of the computational domains

The equations that describe blood flow in a compliant vessel (see Cap. 3 for their formalization) should be numerically solved in a realistic computational domain. This is mandatory to obtain significant results from a clinical perspective. To do this, the following steps should be performed: i) *reconstruction of the lumen geometry*, ii) *deflation of the domain*, iii) *generation of the vessel wall domain*, and iv) *generation of the computational meshes*.

Reconstruction of the lumen geometry. The reconstruction of the lumen geometry provides a first guess of the domain where the fluid problem will be solved. To do this, a standard procedure is based on the segmentation of medical images based on *Magnetic Resonance Imaging* (MRI) or *Computed Tomography* (CT) technology. Since these images are often affected by noise, an imaging enhancement is usually performed at the beginning of the procedure [19, 16, 6]. Then, image segmentation starts. This consists in the construction of the shape of the vascular district, by detecting those points which (presumably) belong to the boundary of the vessel lumen (*i.e.* to the interface with the vessel

wall). Typically, a *thresholding* technique is used, where a threshold k is selcted to identify the points whose intensity in grey-scale is larger than k. This is motivated by the assumption that the vessel lumen is characterized by intensity values larger than the background [1]. More sophisticated class of segmentation methods are *front propagation methods*, where the propagation of a suitable wavefront is tracked, based on noticing that the speed of the wave is small in regions where the image intensity changes rapidly, so that the wavefront is supposed to slow down when approaching the boundary [22]. See [17] for a review of other methods. In *Fig. 1*, an example of final reconstruction of an ascending aorta is provided.

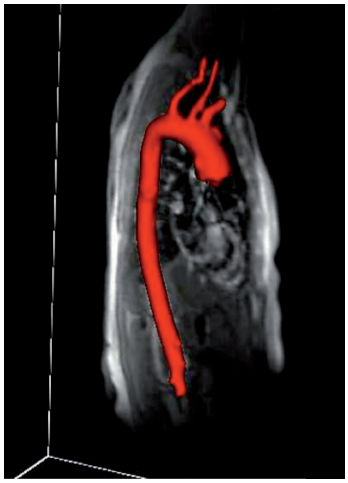


Fig. 1 – Example of segmentation of the lumen boundary of an ascending aorta. Courtesy of E. Faggiano.

In this work, we consider MRI images of the carotid of a patient who underwent to carotid endarterectomy (TEA) (*i.e.* the surgical plaque removal) both before (*i.e.* in the stenotic case) and after (*i.e.* without plaque) TEA. These data have been provided by the Ospedale Maggiore Policlinico di Milano. The segmentation of such geometries has been performed with the open source software VMTK (www.vmtk.org).

Deflation of the computational domain. In this work, we consider a non-linear elastic material (see Cap. 3), so that we could not ignore the non-null diastolic blood pressure which characterizes a radiological image. This means that the geometry we have reconstructed, corre-

sponds to an internal pressure of about 70*mmHg*. To account for this, here we use the strategy introduced in [12, 3]. In particular, the idea is to recover the *zero-stress geometry* (that is the one would have without the blood inside the lumen) by suitably deflating the diastolic one. This is obtained by introducing an inverse problem that has been solved by means of fixed point iterations (see [12] for more details). In *Fig. 2* the deflation of a carotid domain is shown.

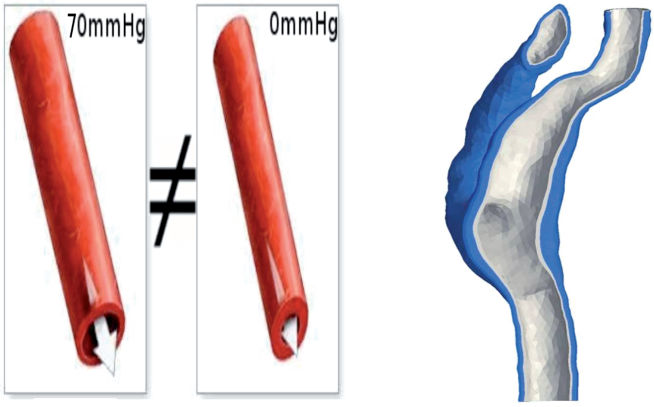


Fig. 2 – Deflation of a carotid.
In blue the original segmentation, in grey the geometry after the deflation.

Generation of the structure domain. The vessel wall is not detectable from standard radiological images. Instead, the lumen where blood flows and the plaque could be identified. Once the lumen geometry has been reconstructed (see previous point), the region occupied by the plaque is reconstructed as well (see *Fig. 3*, left). Thus, the vessel wall domain was obtained by extrusion, from the surface delimiting the lumen/plaque reconstruction, with a total wall thickness equal to 20% of the local vessel radius, intended as the sum of the lumen and plaque transversal distance from the center [4], see *Fig. 3*, left.

Generation of the computational meshes. Once the three geometries characterizing the different part of a carotid (the lumen where blood flow, the plaque, and the vessel wall) have been reconstructed, the internal volume of these regions were discretized using tetrahedral elements.

In *Fig. 3*, middle and right, we report the computational meshes of the lumen vessel after deflation considered in this work, both before and after TEA.

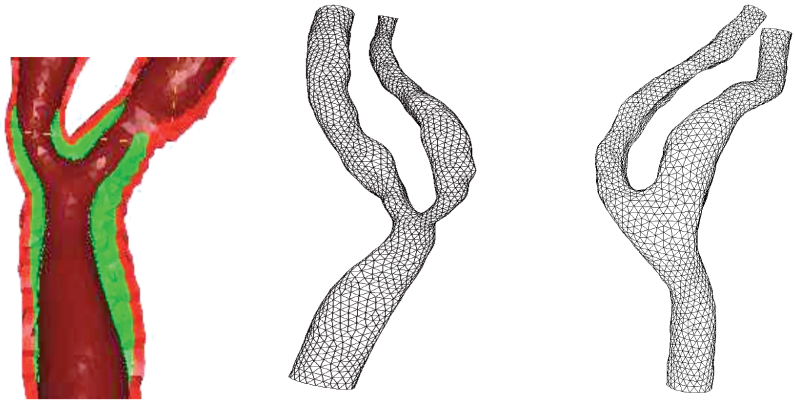


Fig. 3 – Left. Reconstruction of the lumen/plaque/vessel wall geometry. In green, the plaque; in red, the vessel wall obtained by extrusion. Middle and right. Computational meshes for the fluid problem before the plaque removal (stenotic case) and after the plaque removal, respectively.

3. Mathematical and numerical modeling

3.1 Mathematical modeling

As highlighted in the Introduction, blood flows in compliant vessels. From a mathematical point of view, this can be described by a *fluid-structure interaction* (FSI) problem.

Well accepted hypotheses for blood are incompressibility and homogeneity. Moreover, its rheology could be considered Newtonian for medium and large vessels such as carotids [15]. The deformation of the vessel wall is mathematically described by the elastodynamics equation. It is often assumed a non-linear finite elastic law which relates the stress tensor to the strain, and a nearly incompressible behavior [10].

The fluid problem is usually written in an Eulerian framework, whereas the vessel wall problem in a Lagrangian configuration. Given a function g defined in the current configuration Ω^t, we denote by $\hat{g} = g \circ \mathcal{L}$ the function in the reference configuration Ω, $\mathcal{L}$ being the Lagrangian map. Moreover, $\boldsymbol{F} = \nabla \boldsymbol{x}$ is the deformation tensor and $J = \det(\boldsymbol{F})$ represents the change of volume between the reference and the current configuration. The stress tensor in the reference configuration *(First Piola-Kirchhoff* stress tensor) is mapped into the current con-

figuration (Cauchy stress tensor), as follows: $\widehat{\boldsymbol{T}}_s = J\boldsymbol{T}_s\boldsymbol{F}^{-T}$. The common boundary between the two regions is the fluid-solid (FS) interface Σ^t, which coincides with the physical lumen boundary and the internal vessel wall/plaque boundary, see *Fig. 4*, left.

Thus, the FSI problem at each time reads as follows:

$$\rho_f \left(\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{v} \right) - \nabla \cdot \boldsymbol{T}_f(\boldsymbol{v}, p) = \boldsymbol{0} \qquad \text{in } \Omega_f^t(\boldsymbol{d}_f), \tag{1a}$$

$$\nabla \cdot \boldsymbol{v} = 0 \qquad \text{in } \Omega_f^t(\boldsymbol{d}_f), \tag{1b}$$

$$\boldsymbol{v} = \frac{\partial \boldsymbol{d}}{\partial t} \qquad \text{at } \Sigma^t(\boldsymbol{d}_f), \tag{1c}$$

$$\boldsymbol{T}_s(\boldsymbol{d})\,\boldsymbol{n} = \boldsymbol{T}_f(\boldsymbol{v}, p)\,\boldsymbol{n} \qquad \text{at } \Sigma^t(\boldsymbol{d}_f), \tag{1d}$$

$$\rho_s \frac{\partial^2 \widehat{\boldsymbol{d}}}{\partial t^2} - \nabla \cdot \widehat{\boldsymbol{T}}_s(\widehat{\boldsymbol{d}}) = \boldsymbol{0} \qquad \text{in } \Omega_s, \tag{1e}$$

$$\boldsymbol{d}_f = \boldsymbol{d} \qquad \text{at } \Sigma^t, \tag{1f}$$

where v is the blood velocity, p the blood pressure, $\boldsymbol{T}_f = 1/2\mu\,(\nabla\boldsymbol{u} + \nabla\boldsymbol{u}^T) - pI$ the Cauchy fluid stress tensor, μ the blood viscosity, $\boldsymbol{d}$ the vessel wall displacement, and $\boldsymbol{d}_f$ the fluid domain displacement; ρ_f and ρ_s denote the fluid and vessel wall densities. $\widehat{\boldsymbol{T}}_s(\widehat{\boldsymbol{\eta}_s})$ is the first Piola-Kirchhoff tensor for a nearly incompressible exponential material, that is

$$\widehat{\boldsymbol{T}}_s(\boldsymbol{F}) = GJ^{-2/3}\left(\boldsymbol{F} - \frac{1}{3}\mathrm{tr}(\boldsymbol{F}^T\boldsymbol{F})\boldsymbol{F}^{-\mathrm{T}}\right) e^{\gamma(J^{-\frac{2}{3}}\mathrm{tr}(\boldsymbol{F}^T\boldsymbol{F})-3)} + \frac{\kappa}{2}\left(J - 1 + \frac{1}{J}\ln(J)\right) J\boldsymbol{F}^{-\mathrm{T}}, \tag{2}$$

where κ is the bulk modulus, G the shear modulus, and γ regulates the exponential behaviour. For small deformations this material behaves as a linear structure. In order to accout for the plaque, we will consider different values of κ and G in the region of the vessel wall and in that of the plaque, accounting for the increased stiffness in the latter, see *Fig. 4*, right.

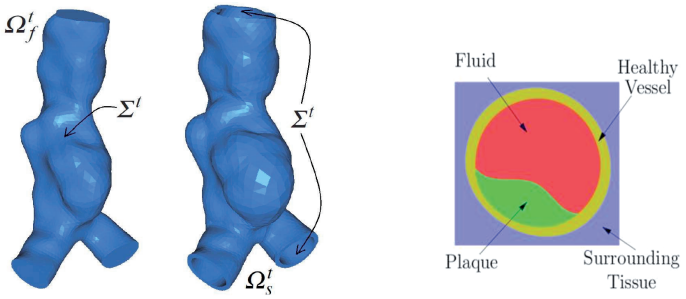


Fig. 4 – Left. Fluid and structure domains.
Right. schematic representation of the different computational domains depicted through a section.

In the coupled problem above, together with the Navier-Stokes equations (1a)-(1b) and the elastodynamics equation (1e), three interface coupling conditions are identified: the *kinematic* condition (1c) stating a no-slip assumption between fluid and structure particles; the *dynamic* condition (1d) expressing the action-reaction principle (III Newton law); and the *geometric* condition (1f) enforcing the perfect adherence between fluid and structure domains.

Problem (1) should be completed with suitable initial condition for $\boldsymbol{v}$, $\boldsymbol{d}$and $\dot{\boldsymbol{d}}$, and boundary conditions. For the structure, here we consider homogeneous Dirichlet and Neumann conditions at the artificial sections in the normal and tangential directions, respectively, and a Robin condition with coefficient α_e to account for the presence of the surrounding tissue at the external surface [17]. For the fluid problem, we prescribe a flow rate at the inlet and an absorbing resistance condition at the outlet [14].

3.2 Numerical approximation

For the numerical solution of the FSI problem (1), segregated algorithms are here considered, where the solution of the fluid and structure problems are solved separately in an iterative way.

After time discretization (BDF1 for the fluid and mid-point scheme for the structure), the fluid problem is linearized by means of a semi-implicit approach, where the convective term and the fluid domain are extrapolated from previous time steps. This choice introduces a (mild) CFL-like restriction on the time step Δt to preserve absolute stability. We indicate with * extrapolated quantities. Given $\sigma_f \neq \sigma_s$, a general segregated algorithm reads as follows:

For $n \geq 1$, $k \geq 1$, at time step t^n/iteration k, until convergence or $k = k_{max}$,

1. solve the Oseen problem with a Robin condition at the FS interface:

$$\frac{\rho_f}{\Delta t}\boldsymbol{v}^n_{(k)} + \rho_f((\boldsymbol{v}^* - \boldsymbol{v}^*_f)\cdot\nabla)\boldsymbol{v}^n_{(k)} - \nabla\cdot\boldsymbol{T}_f(\boldsymbol{v}^n_{(k)}, p^n_{(k)}) = \boldsymbol{g}^n_f \qquad \text{in } \Omega^*_f, \qquad (3a)$$

$$\nabla\cdot\boldsymbol{v}^n_{(k)} = 0 \qquad \text{in } \Omega^*_f, \qquad (3b)$$

$$\sigma_f\boldsymbol{v}^n_{(k)} + \boldsymbol{T}_f\left(\boldsymbol{v}^n_{(k)}, p^n_{(k)}\right)\boldsymbol{n}^* = \sigma_f\left(\frac{1}{\Delta t}\boldsymbol{d}^n_{(k-1)} + \boldsymbol{g}^n_{fs}\right) + \boldsymbol{T}_s\left(\boldsymbol{d}^n_{(k-1)}\right)\boldsymbol{n}^* \qquad \text{on } \Sigma^*; \qquad (3c)$$

2. then, solve the non-linear vessel wall problem with another Robin condition at the FS interface:

$$\frac{\rho_s}{\Delta t^2}\widehat{\boldsymbol{d}}^n_{(k)} - \nabla\cdot\widehat{\boldsymbol{T}}_s\left(\widehat{\boldsymbol{d}}^n_{(k)}\right) = \widehat{\boldsymbol{g}}^n_s \qquad \text{in } \Omega_s, \qquad (4a)$$

$$\frac{\sigma_s}{\Delta t}\widehat{\boldsymbol{d}}^n_{(k)} + \widehat{\boldsymbol{T}}_s\left(\widehat{\boldsymbol{d}}^n_{(k)}\right)\widehat{\boldsymbol{n}} = \sigma_s\widehat{\boldsymbol{v}}^n_{(k)} + \widehat{\boldsymbol{T}}_f\left(\widehat{\boldsymbol{v}}^n_{(k)}, \widehat{p}^n_{(k)}\right)\widehat{\boldsymbol{n}} - \sigma_s\widehat{\boldsymbol{g}}^n_{fs} \qquad \text{on } \Sigma, \qquad (4b)$$

where $v_f = \boldsymbol{d}_f$ and the terms $\boldsymbol{g}$ account for the quantities at previous time steps. For the choice of the interface parameters σ_f and σ_s, several choices have been proposed to obtain fast convergence, see *e.g.* [2, 7, 8, 21].

At each iteration the fluid problem is solved in a (known) moving domain Ω^*_f. This corresponds to an explicit treatment of the geometric interface condition (1f). This choice produces stable results in hemodynamics, due to the restrained wall displacements [5]. For its solution, we consider the *Arbitrary Lagrangian-Eulerian* configuration, *i.e.* we write the fluid equations in a framework moving with the fluid domain [9]. For the reconstruction of the latter, an harmonic extension of the geometric interface condition is adopted. On the other side, the vessel wall problem is linearized by means of the Newton method.

4. Numerical results

We present here some numerical results obtained in the two geometries described above, namely the carotid of a patient before and after the removal of the atherosclerotic plaque. These numerical results were obtained using the Finite Elements library *LifeV* (www.lifev.org), *P*1 – *Bubble*/*P*1 Finite Elements for the fluid problem and P1 Finite Elements for the vessel wall/plaque problem; the backward Euler scheme and the midpoint method have been used for the time discretization of the fluid and vessel wall/plaque problems, respectively. The implicit Robin-Robin partitioned scheme has been used for the solution of the FSI problem. For the choice of the interface parameters σ_f and σ_s, we used the values proposed in [2, 13], which guarantee excellent convergence properties.

We also set $\rho_f = 1.0$, g/cm^3, $\rho_s = 1.1$ g/cm^3, $\mu = 0.035$ cm^2/s For the healthy vessel wall, we set $\kappa = 10^7$ $dyne/cm^2$, $G = 1.034 \cdot 10^6$ $dyne/cm^2$ (corresponding for small displacements to $E = 3 \cdot 10^6$ $dyne/cm^2$ and $\nu = 0.45$), $\gamma = 1$, whereas for the plaque $\kappa = 6.67 \cdot 10^7$ $dyne/cm^2$, $G = 6.897 \cdot 10^6$ $dyne/cm^2$ (corresponding for small displacements to $E = 2 \cdot 10^7$ $dyne/cm^2$ and $\nu = 0.45$), $\gamma = 1$. The time discretization parameter was $\Delta t = 0.002$ s.

In *Fig. 5*, left, we show the flow rate prescribed at the inlet and obtained by means of Echo-Doppler measurements at Ospedale Maggiore Policlinico di Milano. This datum has been prescribed by means of a Dirichlet condition obtained by selecting a parabolic profile for the velocity. At the outlet, a resistance condition to avoid spurious reflections has been prescribed, see [14]. In *Fig. 5*, middle, the computed pressures at the level of the bifurcation has been reported both before and after the plaque removal. From this comparison, we observe the reduced pressure obtained after the removal of the plaque, in accordance with the clinical evidences. In *Fig. 5*, right, the computed blood velocity streamlines together with the vessel wall/plaque displacements (in arrows) at the systolic peak have been reported.

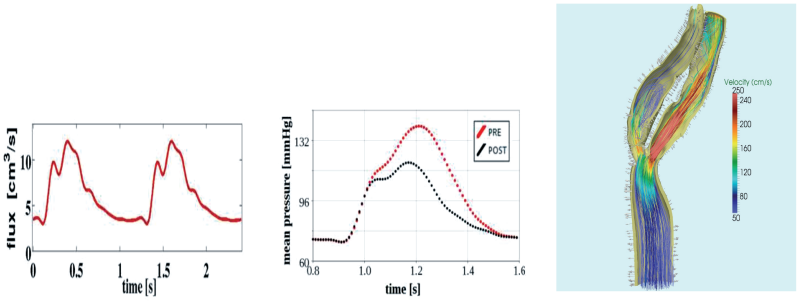


Fig. 5 – Left. flow rate prescribed at the inlet obtained by means of Echo-Doppler measurements; Middle. Computed pressure at the level of the bifurcation; Right. Computed blood velocity stream-lines and vessel wall/plaque displacements at the systole.

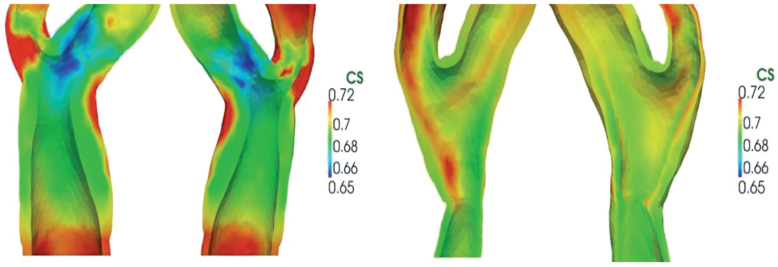


Fig. 6 – Cyclic strain before (left) and after (right) the plaque removal.

In *Fig. 6*, we report the *cyclic strain* in the vessel wall/plaque domain both before and after the plaque removal. This quantity is

defined for each point as $(V_{syst} - V_{diast})/ V_{syst}$, where V_{syst} and V_{diast} are the systolic and diastolic *Von Mises stress* [20]. This quantity seems to be correlated with high values of the permeability of the wall which should favour the plaque development. We observe the decrease of the cyclic strain after the plaque removal.

ACKNOWLEDGEMENTS

The author would like to thank M. Domanin, E. Faggiano, R.M. Lancellotti, and M. Pozzoli.

REFERENCES

[1] L. Antiga, J. Peiro, and D.A. Steinman. From image data to computational domains. In *Cardiovascular mathematics, edited by L. Formaggia, A. Quarteroni, A. Veneziani, Chapter 4*, pages 123-175. Springer, 2009.

[2] S. Badia, F. Nobile, and C. Vergara. Fluid-structure partitioned procedures based on Robin transmission conditions. *Journal of Computational Physics*, 227:7027-7051, 2008.

[3] J. Bols, J. Degroote, B. Trachet, B. Verhegghe, P. Segers, and J. Vierendeels. A computational method to assess the in vivo stresses and unloaded configuration of patient-specific blood vessels. *Journal of Computational and Applied Mathematics*, 246:10-17, 2013.

[4] E. Faggiano, L. Formaggia, and L. Antiga. An open-source tool for patient-specific fluid-structure vessel mesh generation. In *V International Symposium on Modelling of Physiological Flows (MPF)*, pages 25-26, 2013.

[5] M.A. Fernandez, J.F. Gerbeau, and C. Grandmont. A projection semi-implicit scheme for the coupling of an elastic structure with an incompressible fluid. *International Journal for Numerical Methods in Engineering*, 69(4):794-821, 2007.

[6] A.F. Frangi, W.J. Niessen, R.M. Hoogeveen, T. Van Walsum, and M.A. Viergever. Model-based quantitation of 3-d magnetic resonance angiographic images. *IEEE Transactions on Medical Imaging*, 18(10):946-956, 1999.

[7] L. Gerardo Giorda, F. Nobile, and C. Vergara. Analysis and optimization of robin-robin partitioned procedures in fluid-structure interaction problems. *SIAM Journal on Numerical Analysis*, 48(6):2091-2116, 2010.

[8] G. Gigante and C. Vergara. Analysis and optimization of the generalized schwarz method for elliptic problems with application to fluid-structure interaction. *Numer. Math.*, 131(2):369-404, 2015.

[9] C.W. Hirt, A.A. Amsden, and J.L. Cook. An arbitrary lagrangian eulerian computing method for all flow speeds. *Journal of Computational Physics*, 69:277-324, 1974.

[10] G.A. Holzapfel and R.W. Ogden. Constitutive modelling of arteries. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 466(2118):1551-1596, 2010.

[11] D.N. Ku, D.P. Giddens, C.K. Zarins, and S. Glagov. Pulsatile flow and atherosclerosis in the human carotid bifurcation. positive correlation between plaque location and low oscillating shear stress. *Arteriosclerosis*, 5(3):293-302, 1985.

[12] R.M. Lancellotti. *Numerical Computations of Deflated Vascular Geometries fo Fluid-Structure Interaction in Haemodynamics*. PhD thesis, Università degli Studi di Napoli Federico II, July 2012.

[13] F. Nobile, M. Pozzoli, and C. Vergara. Inexact accurate partitioned algorithms for fluid-structure interaction problems with finite elasticity in haemodynamics. *Journal of Computational Physics*, 273:598-617, 2014.

[14] F. Nobile and C. Vergara. An effective fluid-structure interaction formulation for vascular dynamics by generalized Robin conditions. *SIAM J Sc Comp*, 30(2):731-763, 2008.

[15] K. Perktold, E. Thurner, and T. Kenner. Flow and stress characteristics in rigid walled and compliant carotid artery bifurcation models. *Medical and Biological Engineering and Computing*, 32(1):19-26, 1994.

[16] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629-639, 1990.

[17] A. Quarteroni, A. Manzoni, and C. Vergara. The cardiovascular system: Mathematical modelling, numerical algorithms and clinical applications. *Acta Numerica*, 26:365-590, 2017.

[18] S.G. Timsit, R.L. Sacco, J.P. Mohr, M.A. Foulkes, T.K. Tatemichi, P.A. Wolf, T.R. Price, and D.B. Hier. Early clinical differentiation of cerebral infarction from severe atherosclerotic stenosis and cardioembolism. *Stroke*, 23:486-491, 1992.

[19] M. Unser. Splines: a perfect fit for signal and image processing. *IEEE Transactions on Signal Processing Magazine*, 16(6):22-38, 1999.

[20] H.F. Younis, M.R. Kaazempur-Mofrad, R.C. Chan, A.G. Isasi, D.P. Hinton, A.H. Chau, L.A. Kim, and R.D. Kamm. Hemodynamics and wall mechanics in human carotid bifurcation and its consequences for atherogenesis: investigation of inter-individual variation. *Biomechanics and Modeling in Mechanobiology*, 3(1):17-32, 2004.

[21] Y. Yu, H. Baek, and G.E. Karniadakis. Generalized fictitious methods for fluid-structure interactions: Analysis and simulations. *Journal of Computational Physics*, 245:317-346, 2013.

[22] F. Zhu and J. Tian. Modified fast marching and level set method for medical image segmentation. *J Xray Sci Technol*, 11(4):193-204, 2003.

Istituto Lombardo (Rend. Scienze) 150, 147-165 (2016)

FROM DANTE'S UNIVERSE TO CONTEMPORARY COSMOLOGY

Nota del s.c. MARCO BERSANELLI (*)

(Adunanza del 6 ottobre 2016)

SUNTO. – Una lettura della Divina Commedia con gli occhi di uno scienziato moderno rivela che Dante ha riservato grande attenzione alla descrizione di una vasta gamma di fenomeni naturali, in particolare quelli legati all'astronomia, all'ottica e alla geometria. Un esempio notevole riguarda la struttura generale del cosmo così come emerge dagli ultimi canti del Paradiso nei quali, come fu notato per la prima volta dal matematico elvetico Andreas Speiser nel 1925, viene suggerita una struttura non-Euclidea dello spazio. In tal modo Dante concepisce un universo in cui le sfere astronomiche e quelle angeliche sono disposte simmetricamente e mosse dal Punto divino, il vero centro dell'universo, posto agli antipodi della Terra. Tale struttura presenta una sorprendente somiglianza con il modello cosmologico di Einstein del 1917, caratterizzato da curvatura positiva (3-sfera) e reso statico dalla famosa introduzione di un valore *ad-hoc* della costante cosmologica. Tale modello, così come altre soluzioni basate su valori marcatamente non-nulli della costante di curvatura, risultano oggi esclusi dalle osservazioni astrofisiche. In questo contributo si descrive l'intuizione geometrica di Dante e se ne discute la plausibilità nel contesto della sua poetica e del suo orizzonte culturale. Si descrive inoltre la coincidenza topologica della "3-sfera dantesca" con la curvatura dello spazio-tempo nel contesto del modello cosmologico standard, oggi fortemente confermato da osservazioni di alta precisione quali quelle ottenute dal satellite Planck dell'ESA.

***

ABSTRACT. – A reading of the *Divina Commedia* with the eyes of a modern scientist reveals that Dante devoted great attention to the description of a wide variety of natural phenomena, particularly those involving astronomy, optics and geometry. A remarkable case is the structure of the cosmos emerging from the *Paradiso*, which foreshadows a non-Euclidean geometrical structure with remarkable similarities to Einstein's 1917 static cosmological solution. Such model, however, as well as other solutions with positive spatial curvature, are ruled out by current astrophysical observations. Here I dis-

(*) Dipartimento di Fisica, Università degli Studi di Milano, Italia.
E-mail: marco.bersanelli@mi.infn.it

cuss Dante's geometric intuition and show its close analogy with the shape of the observable cosmic space-time in the standard CDM expanding model, fully supported by present-day cosmological data.

DANTE'S VISION OF THE NATURAL WORLD

A modern scientist running into Dante's *Divina Commedia* will discover several passages where the description of a wide variety of natural phenomena related to astronomy, optics, geometry, physics, mathematics, geology, ethology, are treated with remarkable attention.[1] For example, light and its diverse manifestations provide a central element in the scenography of the *Commedia*. Dante frequently uses light as a metaphor of divine realities and of moral virtues, but in so doing he also occasionally describes its physical properties exhibiting the skill of a careful naturalistic observer. In *Purg.* XV (16-21), for example, when Dante encounters the angel of mercy, appearing to him as a glare of reflected divine light, he goes as far as giving a rigorous description of the law of reflection expressed in verses:

Come quando da l'acqua o da lo specchio
salta lo raggio a l'opposita parte,
salendo su per lo modo parecchio
a quel che scende, e tanto si diparte
dal cader de la pietra in igual tratta,
sì come mostra esperïenza e arte;

This passage shows that Dante's understanding of optics ("*perspectiva*") was not based just on commonsense, but it was rooted in some level of methodical study, likely based on the works of Alhazen and Witelo (Nardi 1990, Gilson 1999). It also shows that he felt no discomfort with including detailed accounts of natural phenomena in his poetry, rather, he used them to introduce additional richness in his narration. In *Purg.* XXV, 91-93, we find a beautiful description of the rainbow that goes as follows:

---

[1] The significance of the scientific dimension in Dante's writing has been the subject of a vast literature, see *e.g.* Capasso (1967), Gizzi (1974), Di Scipio & Scaglione (1983), D'Amore (1994), Stabile (1997), Gilson (1999), Ledda (2001), Peterson (2002), Ricci (2005), Buti & Bertagni (2008).

E come l'aere, quand'è ben pïorno,
per l'altrui raggio che 'n sé si reflette,
di diversi color diventa addorno.

Here, in the same *terzina*, Dante gives a poetic rendering of the rainbow (*l'aere… di diversi color diventa addorno*), but he also provides a synthetic account of the physical mechanism responsible for its appearance, *i.e.*, the reflection of solar light (*l'altrui raggio*) in the droplets of moist in the atmosphere (*l'aere… ben pïorno*).[2]

Another notable example is the discussion of the nature of the dark spots on the Moon in *Paradiso II*. Here we find a combination of empirical observation, logical reasoning, and even the suggestion of an experimental approach. Indeed, in making her final case to rule out the hypothesis that the lunar dark spots are due to sunlight reflected from deeper regions in the Moon surface, Beatrice proposes the following optics experiment (*Par.* II, 97-105):

Tre specchi prenderai; e i due rimovi
da te d'un modo, e l'altro, più rimosso,
tr'ambo li primi li occhi tuoi ritrovi.
Rivolto ad essi, fa che dopo il dosso
ti stea un lume che i tre specchi accenda
e torni a te da tutti ripercosso.

Three mirrors should be placed at different distances, as to represent a model of the irregular surface of the Moon, supposed to be everywhere reflective, with the central mirror further away than the two lateral ones. The light from a candle, which mimics the Sun rays, is reflected back by the three mirrors into the eyes of the observer. The conclusion is that:

Ben che nel quanto tanto non si stenda
la vista più lontana, lì vedrai
come convien ch'igualmente risplenda.

---

[2] Several medieval scholars studied in detail the physics of the rainbow, including Robert Grosseteste and Albert the Great. In 1310 a contemporary Dante's, the German Dominican Theodoric of Freiberg, in his work *De iride et radialibus impressionibus* correctly described the shape and colours of the rainbow as a combination of reflection and refraction of sunlight in the water droplets in the atmosphere (Grant 1974).

Although the solid angles – as we would say in the language of modern physics – of the three images of the flame are different, the one coming from the mirror placed further away being smaller (*ben che nel quanto tanto non si stenda)*, their surface brightness is the same (they shine with the same intensity: *igualmente risplenda*). Since the light from the Sun covers completely the exposed surface of the Moon, this would result in equal luminosity in every part of the Moon disc, which falsifies[3] the hypothesis. This is a remarkably accurate, non-trivial physical observation. The fact that the surface brightness (not the flux density) is independent of distance[4] is a straightforward concept for a modern physicist, but surely not obvious for a medieval poet. In fact, a number of non-physicist commentators of the *Commedia* have seriously misunderstood Dante's "three-mirror experiment", labelling his conclusions as obviously wrong or absurd, or just insisting on the purely allegorical nature of the situation being described.[5]

Dante's deep interest in the physical world, as well as his audacious imagination, were rooted in a cultural view where everything was believed to have a role and a meaning. In the medieval mentality, each creature was thought to be created for a purpose and was in relation with the whole. From the rainbow to the whole universe, every natural reality was considered a sign of the Creator, and for this reason worthy of attention and consideration. As Beatrice states in Par I (103-108):

[…] Le cose tutte quante
hanno ordine tra loro, e questo è forma
che l'universo a Dio fa simigliante.

---

[3] Interestingly, here Beatrice uses precisely the word "falsified" (*e s'elli avvien ch'io l'altro cassi, / falsificato fia lo tuo parere. Par*. II, 83-84), a term that is retained in modern philosophy of science.

[4] This is true assuming that absorption from the medium (air in this case) is negligible, a fully justified condition in this case.

[5] *E.g.*, Parodi (1922), Corsano (1975), Boyde (1993). Interestingly, in discussing this topic Patrick Boyde reproduces a XV Century miniature representing Dante's three-mirror experiment in which the flame images look exactly the same. This testifies that, not surprisingly, not only modern commentators but also ancient ones have been deceived by the optical setup described by Dante.

In Dante's poetry there is a distinction, but not a dichotomy, between natural phenomena and their metaphysical dimension. Physics and aesthetics live together. Clearly Dante's mindset was quite different from the one, far more specialized and fragmented, typical of our modern era. Now days, no scientist would think of including in an academic paper any aesthetic appreciation of the system under study, and probably no poet would dear giving a rigorous account of a physical law in her/his artwork.

The fascinating interplay between physical and metaphysical, between geometry and virtue, reaches an apex in Dante's narration in the last *Canti* of *Paradiso*, where an unprecedented cosmic picture is suggested.

Dante's universe

The iconography of Dante's cosmos, as normally depicted in textbooks, appears rather cumbersome. The central Earth is surrounded by the spheres of the seven planets, by the *Stellatum* with the fixed stars, by the *Primum Mobile*, and ultimately by the Empyrean; then, separated from the astronomical universe, we normally see an additional structure composed by a set of nine concentric circles; and then the *Candida Rosa* is often drawn in some intermediate region. It is hard to believe that Dante, with his deep appreciation for symmetry and unity, at the apex of his cosmic journey presents us with such disconnected image of the universe. Indeed, this awkward cosmic map is the consequence of an apparent internal contradiction present within the text of the *Divina Commedia*.

In Dante's times the traditional cosmography was a simplified Aristotelian-Ptolemaic universe, a system that was enthusiastically adopted by medieval scholars after the rediscovery of Greek authors. Importantly, however, the medieval Christian culture added a tenth sky beyond the *Primum Mobile*, the Empyrean, inhabited by the angels, the blessed ones, and by God himself. Often the Empyrean was further divided into nine concentric circles representing the nine "angelic choirs", each corresponding to an astronomical sphere. The remarkable XIV century fresco by Piero di Puccio, in Piazza dei Miracoli, Pisa, shows the nine Empyrean heavens surrounding the *Primum Mobile* and the nested astronomical orbs around the central Earth.

In several places Dante's description seems to follow closely such traditional world picture. For example, when Dante and Beatrice approach the *Primum Mobile,* the Empyrean is depicted as the outmost sphere full of light and love surrounding the entire universe (*Par.* XXVII, 112-114):

> *Luce e amor d'un cerchio lui comprende,*
> *sì come questo li altri; e quel precinto*
> *Colui che 'l cinge solamente intende*

Here, *Colui che l'cigne* is God himself, imagined as the ultimate outer rim of the Empyrean. This is coherent with *Par.* XIV, 30, where God is described as *non circunscritto e tutto circunscrive*.[6]

However, in *Par.* XXVIII, 13-39, when Dante and Beatrice look out into the Empyrean from the *Primum Mobile*, they experience a surprising view. They see a bright point of pure light – so bright that Dante is nearly dazzled – surrounded by nine concentric fiery circles, widely extending over the whole visible space (*Par.* XXVIII, 31-34). The speed and the luminosity of the nine circles increase in proportion to their proximity to the center. The bright central point represents God himself, as a foreshadow of what will become Dante's ultimate divine vision at the end of the poem. As Beatrice explains, the nine circles are the angelic orders, symmetric to the astronomical spheres, forming a mirror image of the physical universe[7] (*Fig. 1*). However, contrary to the physical spheres, their rotation speed increases towards the centre. Puzzled by this velocity asymmetry Dante asks Beatrice, who clarifies that the correspondence between the Empyrean and astronomical spheres is to be found in their degree of *virtute*, not in their physical size (*Par.* XXVIII, 46-78).

In summary, in some passages Dante's Empyrean is the traditional all-encompassing tenth sphere around the *Primum Mobile*, God being its ultimate limit; in others, it is pictured as a separated structure, literally as «another universe» (*l'altro universo*, XXVIII, 71), a reverse image of the astronomical world (Lewis 2006), with God as a vivid

---

6 See also: *Purg.* XI, 2; *Par.* II, 112-114; *Conv.* IV, IX 3.

7 The number 9 completes the recursive structure encountered in the nine circles of Hell and in the nine levels of Purgatory.

infinitesimal Point at its centre. How is this possible? How can God be represented as the outmost sphere embracing the cosmos and, at the same time, as a single point of light? In *Par.* XXX, speaking of the divine Point, Dante exacerbates the paradox by stating in one single verse that they are both simultaneously (*Par.* XXX, 10-12):

> *...il trïunfo che lude*
> *sempre dintorno al punto che mi vinse,*
> *parendo inchiuso da quel ch'elli 'nchiude.*

There is a further puzzle. The place on the *Primum Mobile* from where Dante and Beatrice contemplate the Empyrean scene is a generic, unimportant location. This is explicitly remarked by Dante, as he states that the *Primum Mobile* is so uniform that he can't tell where on the sphere Beatrice lead him (*Par.* XXVII, 100-102). But if the Empyrean is "another universe", somehow adjacent to the physical one, Dante and Beatrice must find themselves at the gate to the Empyrean, a very special viewpoint where they can contemplate the angelic universe that they will soon enter. How can such heavenly door be indicated as an insignificant place?

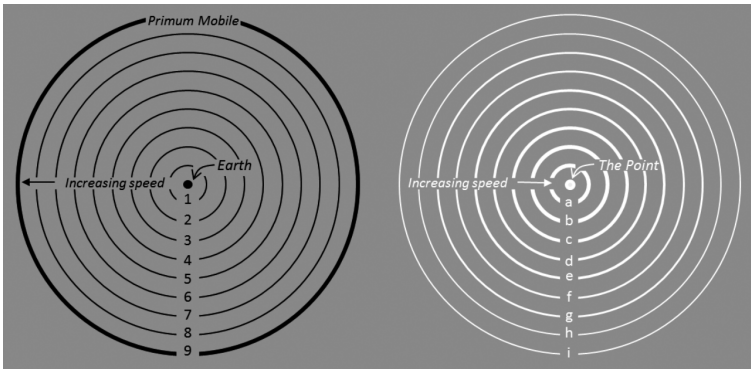


Fig. 1 – Schematic of Dante's cosmos. Left: the arrangement of the astronomical spheres around the Earth. From 1 to 9 the circles represent the spheres of the Moon, Mercury, Venus, the Sun, Mars, Jupiter, Saturn, the fixed stars, and the Primum Mobile. Right: the Empyrean circles with the nine angelic orders. From *a* to *i* the circles represent: Seraphim, Cherubim, Thrones, Dominions, Virtues, Powers, Principalities, Archangels, Angels (*Par.* XXVIII, 97-126). Their luminosity and rotation speed increase towards the centre. The correspondence between the physical and angelic spheres is according to their virtue, which is in inverse order with their size, *i.e.*: *a* → 9, *b* → 8, … *i* → 1.

When Dante and Beatrice leave the Primum Mobile to move beyond, the vision of the Point and of the angelic circles fades away (*Par.* XXX, 1-15). New and deeper levels of vision of the Empyrean, and of God himself, will become accessible to Dante as he will be more and more deeply immersed in the heavenly realm (*Par.* XXX, 1-15). This is coherent with what Dante experienced throughout his journey in the *Paradiso*: as he ascends, his sight increasingly gains the capability to sustain and grasp the ineffable realities he encounters (*e.g.*, *Par.* XXX, 76-81, XXXIII, 109-114). Indeed, the geometrical structure described in Canto XXVIII is just a limited vision of the Empyrean, the partial image of it that can be glimpsed from within the *Primum Mobile* – which after all is still part of the *mondo sensibile*. While such geometric structure does not capture the whole nature of the Empyrean, its peculiar arrangement and the apparent inconsistency it implies require some kind of explanation.

A SPHERICAL SPACE

Is there a way to reconcile Dante's cosmic contradictions? An implicit, apparently obvious, assumption in all of the above discussion is that the underlying space where the scheme of *Fig. 1* is drawn is a flat sheet of paper: an Euclidean two-dimensional space. However, if we draw the same scheme on a 2-d *spherical* space, like the surface of a globe (see *Fig. 2*), all the apparent contradictions are readily resolved. This is best seen by using our modern geometric language. A spherical surface of radius R centred at the origin is described by the familiar equation $x_0^2 + x_1^2 + x_2^2 = R^2$. The sections $x_2 = k$, with $|k| \leq R$, are circles parallel to the (x_0, x_1) plane with radius $(R^2 - k^2)^{1/2}$. An observer moving in any direction from the point $(0, 0, R)$ (the Earth in Dante's scheme, north pole in *Fig. 2*), will encounter increasingly larger circles (the astronomical orbs), each *containing* all the previous (and smaller) ones. Moving beyond the equator ($x_2 = 0$), the observer will now cross increasingly smaller circles, *i.e.*, the angelic choirs. Now, however, as the observer moves on, the new sections crossed will *include* larger sections. Eventually the crossed circles converge around a single point at $(0, 0, -R)$ («The Point» in *Fig. 1*).

Now the physical universe and the Empyrean are symmetric to

each other, like the two hemispheres of Fig. 2. The Empyrean surrounds everywhere the *Primum Mobile* and, at the same time, it is ultimately converging on a single point, the antipodal of the Earth. The velocities of all circles (astronomical and angelic) are now naturally ordered in proportion to their proximity to the divine Point (a proximity that defines their *virtute*), and their sizes are reversed in the two hemispheres. Also, it is now clear why the vision of the angelic circles does not depend on the specific position of the observer on the *Primum Mobile*: the scheme is completely symmetric around the equator.

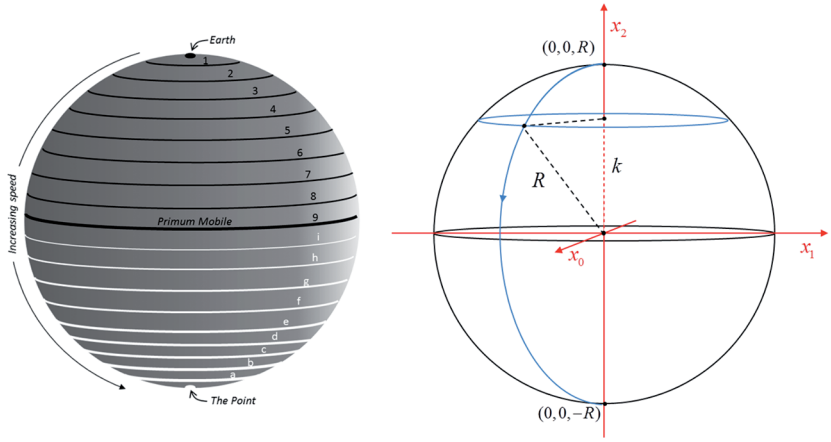


Fig. 2 – Dante's spherical universe. *Left*: The two arrangements of *Fig. 1* (the physical and angelic circles of Dante's cosmos), are here drawn on a 2-d spherical space. The numbers and letters correspond to those of *Fig. 1*. Note that the rotation velocities in this case increase regularly from "Earth" to "The Point". *Right*: coordinate system for the sphere (see text).

Of course *Figs. 1-2* represent two-dimensional maps of the true 3-d space, since each circle represents a sphere. In the case of *Fig. 2*, the extra dimension implies that the underlying space is a 3-sphere, or hypersphere, *i.e.* the analogue of a 2-d spherical surface in one more dimension. The analytical expression for a 3-sphere centred in the origin and with radius R is simply:

$$\sum_{j=0}^{3} x_j^2 = x_0^2 + x_1^2 + x_2^2 + x_3^2 = R^2 \qquad [1]$$

We cannot visualize a 3-sphere, but we can visualize its sections with constant values of $x_3 = k$ with $|k| \leq R$. Similarly to previous case, these sections are a set of concentric *spheres* (not circles) of radius $(R^2 - k^2)^{1/2}$. Their radius, as seen by an observer moving outwards in a generic direction, increases up to $x_3 = 0$ and then diminishes until the spheres converge around a point at $x_3 = - R$. In Dante's image, that antipodal point includes all the spheres of the universe, astronomical and angelic. In this view Dante's enigmatic description of the divine Point makes perfect sense: *parendo inchiuso da quel ch'elli 'nchiude* (XXX, 10-12). God is the true centre of the cosmos: *Da quel punto depende il cielo e tutta la natura* (*Par*. XXXIII, 41-42). The Earth is the anti-centre of the universe, *l'infima lacuna de l'universo* (*Par*. XXXIII, 22-23), with Lucifer buried at its centre, *da tutti li pesi del mondo costretto* (*Inf*. XXIX, 57). All the apparent contradictions are resolved, the poetry can be better understood and more deeply enjoyed.

The first modern-time suggestion of this reading of Dante's cosmic geometry, as far as I know, was made in 1925 by Swiss mathematician Andreas Speiser (Speiser 1925). Interestingly, Speiser was working on non-Euclidean geometry, an expertise that enabled him to readily recognize the nature of Dante's intuition. A number of authors have further discussed this interpretation (Callahan 1976, Peterson 1979, Osserman 1995, Patapievici 2004, Egginton 1999; Bersanelli 2012, Rovelli 2014). Of course Dante did not have the mathematical language we have today, and his notion of a curved space must have been purely qualitative, far from the rigorous description given above. In fact, it is straightforward for us to describe a non-Euclidean space, while it surely required a stunning imagination for Dante to think something like that without the aid of mathematics. Is it realistic that he could have conceived such counterintuitive, yet coherent, geometrical structure?

A few elements should be considered. First, at Dante's time the geometrical imagination was less rigidly anchored to the Newtonian paradigm than it is today. As a matter of fact, Dante was more familiar with spherical astronomy than with plane Euclidean geometry (Egginton 1999). Furthermore, at that time the notion of a spherical Earth, while widely accepted, was sufficiently paradoxical that conceiving the same situation in one more dimension may have been not such a great leap of imagination as it is for us today. Interestingly, Dante's

mentor Brunetto Latini, in the first book of his encyclopedic work *Li Tresor*, to explain the spherical shape of the Earth does not describe it as a sphere seen from outside. Instead, he tells the story of two horsemen travelling in opposite directions, who end up finding themselves in front of one another at the opposite side of the globe (Rovelli 2014). This must have been an astonishing thought at the time. Brunetto sees spatial curvature from within space, not from outside. Indeed, this is the most straightforward way for us to grasp curvature in 3-d space as well. Just as, on the 2-d spherical Earth, a horseman going straight in any direction will find himself at the antipodes, in a 3-spherical space an astronaut flying straight on a spaceship will reach the antipodal point of the universe. And if the horseman and the astronaut will keep going long enough, they will both come back to their starting points.

THE 21ST CENTURY VIEW

Since Dante's times, we have learnt a great deal about the nature of our physical universe. Today we routinely observe large numbers of galaxies – each containing ~10^{11} stars – distributed rather uniformly in deep space. We have also realised that our universe is expanding and we can measure accurately the time to the beginning of cosmic expansion: about 13.8 billion years (Planck Collaboration 2016). In spite of light's great speed, $c \sim 10^8$ m/s, cosmic distances are so great that light takes a long time to cross them. The photons from the furthest galaxies travelled up to 12-13 billion years before reaching us, and therefore we see those galaxies as belonging to a relatively young universe, only 1 or 2 billion years old. However, because of expansion, that young universe was also much smaller than it is today. We may suspect that those spherical layers of the universe that we see far away, which surround us at very large distances, are actually *smaller* than the huge space they contain. This is indeed the case. A paradox that immediately reminds us of Dante's spherical geometry: smaller spheres that *include* larger ones.

To look at the picture in more detail we need to use again mathematical language. In relativistic cosmology we describe space-time in an expanding isotropic and homogeneous universe through the Friedman-Lemaître-Robertson-Walker (FLRW) metric:

$$ds^2 = -c^2dt^2 + R(t)^2\left[d\chi^2 + S_k^2(\chi)d\psi^2\right] \qquad [2]$$

where dt is the infinitesimal time interval, $d\chi$ is the comoving coordinate separation,[8] and $d\psi^2 = d\theta^2 + \sin^2\theta d\phi^2$ is the angular distance on the sphere. The function $R(t)$ is the expansion term, measured in units of length. The function $S_k(\chi)$ takes the values $\sin\chi$, χ, $\sinh\chi$ for closed ($k = +1$), flat ($k = 0$), and open ($k = -1$) models, respectively. The proper distance $D(t)$ between the origin and a given point in space-time is defined on a surface of constant time, $dt = 0$. If we then consider radial distances only, so that $d\psi = 0$, equation [2] reduces to $ds = R(t)d\chi$, and by integration we have $D(t) = R(t)\chi$. Differentiating this relation yields:

$$\dot{D} = \dot{R}\chi + R\dot{\chi} = v_{\text{rec}} + v_{\text{pec}} \qquad [3]$$

where over-dot represents a time derivative. The velocity in terms of proper distance is thus the combination of two terms: a recession term, v_{rec}, representing the effect cosmic expansion; and a peculiar velocity, v_{pec}, representing the motion relative to the comoving coordinates system. The first term can be rewritten as the familiar Hubble law:

$$v_{\text{rec}}(t) = \dot{R}(t)\chi = H(t)D(t) \qquad [4]$$

where $H = \dot{R} / R$ is the Hubble parameter, and its value at present time, $H(t_0) = H_0$, is defined as the Hubble constant.

The second term is useful to understand the motion of a photon in the expanding space. Light rays move along null geodesics in the FLRW metric, *i.e.* along paths corresponding to $ds = 0$. Considering again radial propagation ($d\psi = 0$), from [2] we obtain $cdt = R(t)d\chi$. This can be rewritten as:

$$\dot{\chi}_{\text{light}} = \frac{c}{R(t)}, \qquad [5]$$

which shows that the velocity of a photon through comoving coordinates is not constant, but depends on the expansion factor $R(t)$. Therefore, to calculate the comoving distance of a galaxy that emitted light at time t_{em}, we need to integrate the changing comoving speed of

[8] Comoving coordinates are defined such that they remain fixed for object that follow the expansion flow, *i.e.*, with zero peculiar motion. (In this work we follow the notation of Peacock 1999).

light (eq. [5]) from that time to the present. This yields the equation for our "past light cone" in the expanding universe:

$$\chi_{\text{light cone}}(t_{\text{em}}) = c\int_{t_{\text{em}}}^{t_0} \frac{dt'}{R(t')} \qquad [6]$$

The past light cone represents the set of all the events in the universe that are visible to us today. Light rays travel following the light cone and bring to us images of different galaxies at different stages of their evolution. Integration of equation [6] shows that, for any realistic cosmological model,[9] our past light cone has the shape schematically shown in *Fig. 3*.

Assuming negligible peculiar motion, our world line (*i.e.*, the path of our location in space-time) is the central vertical line of *Fig. 3*. World lines of distant galaxies, such as G1, G2 and G3, diverge from our world line as a consequence of cosmic expansion.[10] We see those galaxies at the time when their world lines intersect our past light cone. The galaxies G1, G2 and G3 are at increasing distances into our past light cone, therefore we observe them to recede from us at increasing rates.

For galaxies relatively nearby (say, up to G1), our light cone is nearly linear (*Fig. 3*). In this region the proper distance is simply proportional to the time it took for light to reach us, as it would be in a static universe. But as we move further away – *e.g.*, between G1 and G2 – the shape of our light cone is curved towards the origin. This means that, as we look back into the past, the proper distance of the objects we see grows more and more slowly. At G2, the proper distance reaches a maximum. This is the so-called "maximum emission distance", *i.e.*, the region of the universe where the galaxies we observe today were furthest away from us when they emitted the light we see.[11] Beyond that

---

[9] Here we assume the standard CDM cosmological model, *i.e.*, a flat expanding space and with matter and vacuum energy densities contributing, respectively, for ~30% and ~70% of the total energy density.

[10] In *Fig. 3* the world lines are straight lines. A more rigorous calculation would show a slight bending with positive curvature, due to time variation of the Hubble term. However, our simplification has no impact on our discussion.

[11] In the standard model this happens at a cosmic time of ~5 billion years, corresponding to a redshift of ~1.5.

point, something remarkable happens: as we look at earlier epochs we see galaxies that were closer and closer to us when they emitted the light we see today. In other words, we see those galaxies as belonging to spheres, centred on our observing point, smaller than the spheres they encompass. Note that beyond G2 the slope of the light cone becomes negative, *i.e.*, photons emitted at those early times were *receding* from us. This is because at that time the expansion rate exceeded the speed of light,[12] $(v_{rec} - c) > 0$. This is in no contradiction with special relativity: the expansion rate can be superluminal since it is a velocity that describes the dynamics of space itself, not the movement of an object through space.

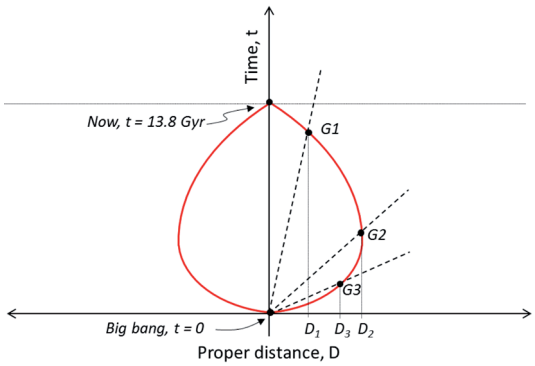


Fig. 3 – The shape of our past light cone (red line). Time (vertical axis) spans from the big bang (t = 0) to the present epoch (13.8 billion years, or Gyr). The slope of the light cone at any time is the velocity of light relative to us at that time: this is not constant, but it is equal to $v_{rec} - c$. A rotation around the vertical axis of the red curve is a 2-d surface topologically equivalent to a sphere. In analogy with *Fig. 2*, any constant-time section in *Fig. 3* (circles parallel to the vertical axis, not shown), represents a sphere. We see the galaxies G1, G2 and G3 as belonging to spheres centred on us and nested into one another. As we look beyond G2, we observe smaller spheres encompassing larger ones (G3 is beyond G2, but $D_3 < D_2$). In scale of this drawing, the last scattering surface is much smaller than the black point indicated as "big bang" at $t = 0$.

We may ask: how far can we actually see into our past light cone? How deep can we see towards the initial point in space-time, the "big

[12] The limit beyond which, at any given time, the recession of galaxies exceeds the speed of light is known as the "Hubble sphere".

bang"? The deepest view we have of the early universe is the cosmic microwave background (CMB), a sea of photons released when the universe was about 3.8×10^5 years old, only 0.003% of its present age. The CMB photons traveled rather undisturbed in their 13.8 billion year journey, so they bring to us a remarkably faithful view of the so-called "last scattering surface" (LSS), the sphere of space-time surrounding us where the CMB photons last interacted with matter. The LSS surrounds us as an ultimate cosmic photosphere seen from the inside. Beyond that surface, the universe was opaque to light. Today we have a detailed image of the LSS thanks to accurate full-sky observations of the CMB, such as those recently obtained by the ESA Planck satellite shown in *Fig. 4* (Planck Collaboration 2016; see also Bersanelli 2013).

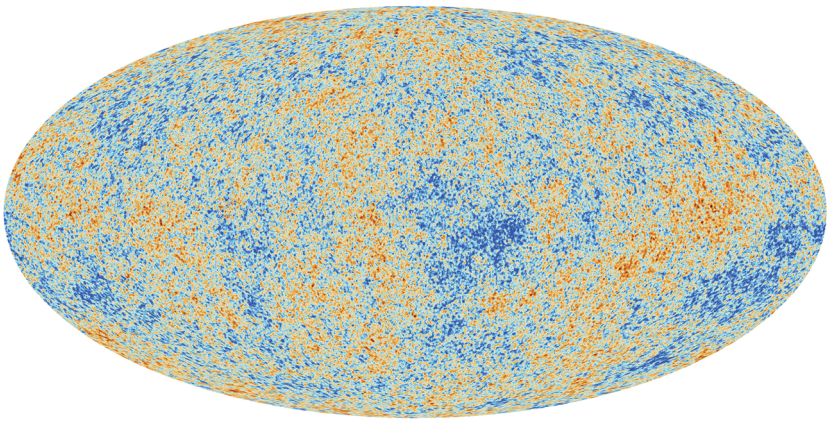


Fig. 4 – Full-sky image of the last scattering surface from the latest data from the Planck satellite (Planck Collaboration 2015). The intensity of the CMB (color scale in brightness temperature, ±300μK) traces small fluctuations in the density and velocity of the primordial hot plasma, whose statistics encode a wealth of information on the structure, composition, evolution and geometry of the universe (credit: ESA and Planck Collaboration).

How large is the last scattering surface? The answer requires care. Calculated at the present cosmic epoch, the material that last interacted with the CMB is now at a distance of ~42 billion light-yeas (about 3 times what we would expect in a static universe). However, this is *not* what we see: we observe the LSS at the time it released the photons, *i.e.*, when its radius was much smaller than today. Calculation shows that the radius of the LSS as we actually see it, is only 38 million light-years,

much smaller than the space it contains. Paradoxically, the ultimate visible sphere in the sky surrounding us in all directions, is dramatically smaller than the space occupied by all the galaxies we observe in between (see *Fig. 5*).

The shape of our past light cone is symmetric around the time axis, and it points in any direction towards the initial big bang. This means that, when we look up in the sky in any direction, we are actually looking towards a single point in space-time, the origin of the universe.

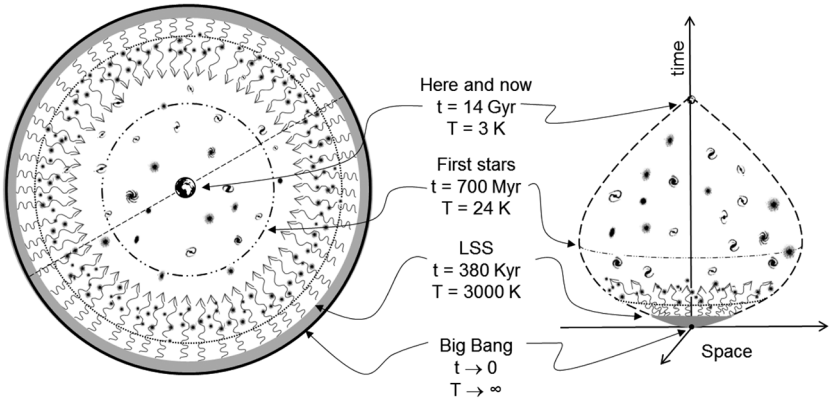


Fig. 5 – Schematic showing some of the key phases in the evolution of the universe: the big bang, the last scattering surface, the formation of the first stars, and the present time. Left: events are shown in concentric spheres, as they appear to our observation; right: the same events are depicted on our past light cone.

Conclusions

This picture has eloquent similarities with Dante's cosmos. Indeed, the topology of our past light cone is identical to that of the hypersphere imagined by Dante in his *Paradiso*. Clearly, we should note an important difference: in the modern vision we deal with a curvature of space-time, while in Dante's case the curvature was spatial only.

Most previous discussions of the subject (*e.g.*, Speiser 1925, Callahan 1976, Peterson 1979, Lachieze-Rey & Luminet 1995, Buonanno & Quercellini 2009, Rovelli 2014) highlighted the similarity

between Dante's universe and a positively curved Riemann space (*i.e.*, a model with k = +1 in eq. [2]), such as the famous static model by Einstein (1917) with a fine-tuned cosmological constant. Such closed models, however, are highly disfavoured by present-day observations. Here, following Osserman (1995), we have discussed Dante's geometrical vision in connection with the shape of the observable space-time in a generic expanding cosmology, the simplest case of which being a flat expanding universe. In recent years, precise cosmological observations strongly indicate a spatially flat universe with matter density $\Omega_M = \Omega_B + \Omega_{CDM} \approx 0.3$ (including baryonic matter, Ω_B, and cold dark matter, Ω_{CDM}) and vacuum energy density $\Omega_\Lambda \approx 0.7$, the so-called standard ΛCDM cosmological model. In particular, the recent full-sky observations of the CMB by the Planck satellite provide measurements of the basic parameters of the CDM model with unprecedented accuracy, at ~1% level or less (Planck Collaboration 2015). Surprisingly, therefore, it turns out that the visionary intuition of the medieval poet Dante Alighieri has a strong analogy with the best general description of cosmic space-time as revealed by 21st century science.

While remarkable, these analogies should not be misunderstood. Dante had no physical or quantitative argument to support his ideas, and his cosmic vision was based entirely on imagination. It would be a mistake to treat his cosmological insight as a scientific theory in a modern sense. Nonetheless, it is noteworthy that his open-minded thinking, his relentless search for a beautiful, meaningful and coherent description of the world lead him to foreshadow concepts that modern science would confirm and develop on scientific basis seven centuries later.

Perhaps the situation is not dissimilar from what happened with some of the great ancient Greek thinkers. Consider, for example, Democritus' intuition of the atomic structure of matter; or Anaxagoras' insight that heavenly bodies are made out of earthly materials: in these cases too, modern science has shown the basic correctness of these ideas, in spite of the fact that their ancient authors could not support them with empirical evidence. Similarly the hypersphere architecture of Dante's universe, originated by his genius rooted in the medieval culture, represents a fascinating original thought in cosmology and should be regarded as a wonderful hidden jewel in the history of scientific ideas.

BIBLIOGRAPHY

Bersanelli, M., 2012. *Light in the Beginning*, in «Light from light. Scientists and theologians in dialogue», Ed. by Gerald O'Collins, S.J. and Mary Ann Meyers, Cambridge: Wm. B. Eerdmans Publishing Co., pp. 80-101.

Bersanelli, M., 2013. *Missione spaziale Planck: Verso l'alba del tempo*, Istituto Lombardo, Rend. Scienze, 147, 183-209.

Boyde, P., 1993. *L'esegesi di Dante e la Scienza*, in «Dante e la Scienza», Ed. by Patrick Boyde & Vittorio Russo, Longo Editore, pp. 14-16.

Buonanno, R. and Quercellini, C., The equations of medieval cosmology, New Astronomy 14 (2009), pp. 347-348.

Buti, G., and Bertagni, R., (2008). Commento astronomico della Divina Commedia, Edizioni Remo Sandron, Firenze (2008).

Callahan, J.J., 1976. *The curvature of space in a finite universe*, Scientific American, vol. 235, pag. 90-100.

Capasso, I., 1967. *L'astronomia nella Divina Commedia*, Domus Galileana.

Corsano, A., 1975. *Dante e la scienza del suo tempo*, Bollettino di Storia della Filosofia, Università degli Studi di Lecce, vol. 3, pp. 69-80.

D'Amore, B., 1994. *La matematica nella Divina Commedia*, Alma Mater Studiorum, Università di Bologna, anno VII, n. 1.

Davis, T.M., and Lineweaver, C.H, 2004. *Expanding Confusion: Common Misconceptions of Cosmological Horizons and the Superluminal Expansion of the Universe*; Publications of the Astronomical Society of Australia; Vol. 21, Issue 1, pp. 97-109.

Di Scipio, G.C., and Scaglione, A., 1983. The Divine Comedy and the Encyclopedia of Arts and Sciences, Acta of the International Dante Symposium, 13-16 Nov. 1983, Hunter College, New York.

Egginton, W., 1999. *On Dante, Hyperspheres, and the Curvature of the Medieval Cosmos*, Journal of the History of Ideas, Volume 60, n. 2, pp. 195-216.

Einstein, A. 1917. Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie. Sitz. König. Preuss. Akad. 142-152. Or 'Cosmological considerations in the general theory of relativity' in (Lorentz et al. 1952): 175-188. CPAE 6 (Doc. 43).

Gilson, S., 1999. *Light Reflection, Mirror Metaphors, and Optical Framing in Dante's Comedy*: *Precedents and Transformations,* Neophilologus, n. 83, pp. 241-52.

Gizzi, C., 1974. *L'astronomia nel poema sacro*, Lofredo.

Grant, E., 1974. *A Source Book in Medieval Science.* Harvard University Press.

Lachieze-Rey, M., Luminet, J.P., Cosmic Topology, Phys.Rept. 254 (1995) 135-214.

Ledda, G., 2001. *Poesia, scienza e critica dantesca: due libri recenti sull'ottica e l'astronomia in Dante*, L'Alighieri A. 42, n. 18, pp. 99-113.

Lewis, C.S., 2006. *The discarded image*, Cambridge University Press, p.116.

Nardi, B., 1990. *Il Canto XV del Purgatorio*, Lecturae ed altri studi danteschi.

Osserman, R., 1995. *A Mathematical Exploration of the Cosmos*, Anchor Books, 1995

Parodi, E.G., 1922. *Il Canto II del Paradiso, letto da Ernesto Giacomo Parodi nella sala di Dante in Orsanmichele*, Sansoni.

Patapievici, H.-R., 2004. *Gli occhi di Beatrice*, Mondadori.

Peacock, J.A., 1999. *Cosmological Physics*, Cambridge University Press.
Peterson, M.A., 1979. *Dante and the 3-sphere,* American Journal of Physics, 12, 1031.
Peterson, M.A., 2002. *Galileo's discovery of the scaling laws*, American Journal of Physics, 70, 575.
Planck Collaboration, 2016. *Planck 2015 results. I. Overview of products and results*; Astronomy and Astrophysics, 594, A1.
Ricci, L., 2005. *Dante's insight into Galilean invariance*, Nature vol. 434, p. 717.
Rovelli, C., 2014. *La realtà non è come ci appare*, Cortina, pp. 83-95.
Speiser, A., 1925. *Das Raumende und das Jenseits*, in «Klassiche Stücke der Mathematik», Orell Füssli.
Stabile, G., 1997. *Bartolomeo da Parma e l'astronomia di Dante*, Seventh Centenary of the Teaching of Astronomy in Bologna, 1297-1997, Proceedings of the meeting held in Bologna at the Accademia delle Scienze, pp. 99-122.

Istituto Lombardo (Rend. Scienze) 150, 167-181 (2016)

SANT'AGOSTINO E LE SCIENZE

CLAUDIO CITRINI (*)

Nota presentata dal m.e. Carlo Domenico Pagani
(Adunanza del 6 ottobre 2016)

SUNTO. – Si riportano numerosi esempi delle conoscenze scientifiche di Sant'Agostino, che dimostrano la sua costante ricerca e adesione alle verità della ragione, in un rapporto complesso ma mai conflittuale con quelle della fede. Con ragionamenti a volte nitidi, altre volte contorti o erronei egli cerca di superare ogni contrasto con l'interpretazione, letterale o allegorica, delle Scritture.

***

ABSTRACT. – Several examples of Saint Augustin's learning in sciences are reported. They show his constant research of the truth both in human rational knowledge and in faith, which must never be in disagreement. He always tries to overcome, with sometimes straightforward and sometimes sophistic arguments, any inconsistency of the scientific knowledge with the literal or allegoric interpretation of the Holy Scriptures.

1. FEDE E RAGIONE

Non si può comprendere il rapporto di Sant'Agostino con le scienze se non si tiene in conto il suo percorso culturale e spirituale. La sua formazione comprendeva tutte le arti che poi furono dette liberali, e, per quanto portato per la retorica, in cui acquisì un tale prestigio da meritarsi la nomina a maestro di retorica nella città allora capitale imperiale, non dimenticò mai la sua preparazione scientifica. Quando, venuto a Milano per assumere la carica, incontrò il vescovo Ambrogio, la sua vita ebbe una svolta, culminata con la conversione e il battesimo, tutta

(*) Istituto Lombardo Accademia di Scienze e Lettere, Milano, Italia.
E-mail: claudio.citrini@polimi.it

la sua vita. La sua produzione apologetica, tuttavia, non fu mai in contrasto con la verità scientifica appresa, ma si sforzò sempre di conciliarla con la verità di fede.

Nella sua catechesi del mercoledì 21 novembre 2012 papa Benedetto XVI descrisse magistralmente il ruolo di Agostino nel percorso di razionalizzazione della fede: "Oggi in questa catechesi vorrei soffermarmi sulla ragionevolezza della fede in Dio. La tradizione cattolica sin dall'inizio ha rigettato il cosiddetto fideismo, che è la volontà di credere contro la ragione. *Credo quia absurdum* (credo perché è assurdo) non è formula che interpreti la fede cattolica. Dio, infatti, non è assurdo, semmai è mistero. Il mistero, a sua volta, non è irrazionale, ma sovrabbondanza di senso, di significato, di verità. …

Sant'Agostino, prima della sua conversione, cerca con tanta inquietudine la verità, attraverso tutte le filosofie disponibili, trovandole tutte insoddisfacenti. La sua faticosa ricerca razionale è per lui una significativa pedagogia per l'incontro con la Verità di Cristo. Quando dice: «comprendi per credere e credi per comprendere» (Discorso 43, 9), è come se raccontasse la propria esperienza di vita. Intelletto e fede, dinanzi alla divina Rivelazione non sono estranei o antagonisti, ma sono ambedue condizioni per comprenderne il senso, per recepirne il messaggio autentico, accostandosi alla soglia del mistero. Sant'Agostino, insieme a tanti altri autori cristiani, è testimone di una fede che si esercita con la ragione, che pensa e invita a pensare. Su questa scia, Sant'Anselmo dirà nel suo Proslogion che la fede cattolica è *fides quaerens intellectum*, dove il cercare l'intelligenza è atto interiore al credere. Sarà soprattutto San Tommaso d'Aquino – forte di questa tradizione – a confrontarsi con la ragione dei filosofi, …".

Vediamo dunque alcuni esempi.

1.1 Razionalità del creato

Il mondo, per Agostino, è fatto con tale perfezione che non può assolutamente essere dovuta al caso. Nel De Ordine (I, 1. 2) egli scrive: "O le membra di animali anche piccolissimi sono strutturate dal caso in dimensioni tanto proporzionate ed esatte; ovvero si ammette che deriva da un principio razionale che non può esser prodotto dal caso; o infine noi oseremmo, per pregiudizi di vana filosofia, non attribuire all'occulta legge del divino potere l'ordine che ammiriamo in ogni essere nella successione di tutti i fenomeni naturali e indipendentemente dalla

razionale produttività dell'uomo. … Ma a questo proposito supponiamo che un tale abbia la vista tanto limitata che in un pavimento a mosaico il suo sguardo possa percepire soltanto le dimensioni di un quadratino per volta. Egli rimprovererebbe all'artista l'imperizia nell'opera d'ordinamento e composizione nella convinzione che le diverse pietruzze sono state maldisposte. Invece è proprio lui che non può cogliere e rappresentarsi in una visione d'insieme i pezzettini armonizzati in una riproduzione d'unitaria bellezza. … La medesima condizione si verifica per le persone incolte. Incapaci di comprendere e riflettere sull'universale e armonico ordinamento delle cose, se qualche aspetto, che per la loro immaginazione è grande, li urta, pensano che nell'universo esiste una grande irrazionalità".

Dunque i dettagli, visti separatamente o non compresi, non permettono di ammirare il meraviglioso quadro della natura, che solo la legge divina ha potuto concepire e attuare.

1.2 Razionalità dell'uomo

Tutto dunque è razionale nel creato, e anche gli animali agiscono secondo un istinto dettato dalla ragione. Naturalmente, l'uomo è a un livello più elevato in quanto dotato di consapevolezza.

Ancora nel De Ordine (II, 19. 49) leggiamo: "Da molti elementi sparsi disordinatamente e poi radunati secondo una struttura unitaria io costruisco una casa. Io valgo di più perché la faccio ed essa è fatta. … Ma non per lo stesso motivo valgo di più della rondine e dell'ape poiché la prima tanto ingegnosamente costruisce i nidi e la seconda i favi. Valgo di più perché sono un animale dotato di ragione. Ma se la ragione consiste nelle misure razionalmente disposte, forse che la costruzione degli uccelli non è misurata proporzionatamente e convenientemente? Anzi ha esattezza matematica. Quindi valgo di più non perché eseguo opere matematicamente esatte ma perché conosco l'esattezza matematica. Ma allora? Gli uccelli possono realizzare cose matematicamente esatte pur non avendone scienza? Lo possono certamente. Da dove l'hanno appreso? Da quello stesso principio per cui anche noi adattiamo la lingua ai denti e al palato perché ne escano parole. Non pensiamo tuttavia nel parlare con quale movimento della bocca lo facciamo".

L'istinto, dunque, o comunque il comportamento appreso per mimesi e applicato senza controllo specifico, per Agostino è ragione inconsapevole.

1.3 Scienza e scienziati

Purtroppo l'uomo talvolta si inorgoglisce della potenza del suo pensiero, e dimentica che essa non è altro che il riflesso della sapienza divina. Ecco un bel passo dalle Confessioni (Conf. V, 3. 4): "Investigando questi misteri con l'intelligenza e l'ingegno da te ricevuti, [gli astronomi] fecero molte scoperte, predissero con anticipo di molti anni le eclissi della luce del sole e della luna, con il giorno, l'ora e la misura in cui sarebbero avvenuti, senza errare nei calcoli. I fenomeni si verificarono puntualmente secondo le loro predizioni, ed essi misero per iscritto le leggi scoperte, tuttora consultate e usate per predire l'anno, il mese dell'anno, il giorno del mese, l'ora del giorno e la misura in cui la luce del sole e della luna scomparirà; e il fenomeno avverrà puntualmente secondo le predizioni. Il popolo ne è ammirato, gli ignari stupiti, gli esperti imbaldanziti ed esaltati. Ma se, lontani ed eclissati dalla tua luce per la loro empia superbia, prevedono con tanto anticipo l'offuscamento futuro del sole, non vedono però il loro, presente, poiché non ricercano con spirito religioso l'origine del proprio ingegno, con cui eseguono queste ricerche; o, se si scoprono tue creature, non si donano a te con slancio affinché tu conservi le tue creature".

Qui entra in gioco l'atteggiamento di Agostino. La scienza lo affascina, ma nulla conta senza la fede. È un mezzo per giungere a intuire che la razionalità del mondo è manifestazione di Dio, ma nulla conta nella storia della salvezza dell'uomo. Più avanti, a proposito degli errori scientifici dei Manichei, afferma (Conf. V, 5. 9): "Ascoltando qua o là un mio fratello cristiano, che in materia è inesperto e ha idee sbagliate, io considero le sue opinioni pazientemente né vedo come gli nuoccia l'ignorare accidentalmente la posizione e la condotta di enti corporei creati da te, allorché su di te, Signore, creatore di tutto, non ha opinioni sconvenienti. Gli nuocerebbe invece il pensare che questa scienza faccia parte proprio dell'insegnamento religioso e l'affermare con sfrontata ostinazione quanto ignora".

Comunque la scienza ha valore di verità assoluta: "La scienza infatti non è costituita da una qualsiasi rappresentazione, ma da una rappresentazione tale che chi l'ha formulata non commetta errore e, anche se sottoposto a qualsiasi critica, non possa dubitare. Ne consegue la indiscussa verità dell'affermazione di alcuni filosofi che in nessuno essa si può ritrovare se non nel filosofante. Egli non deve soltanto aver

conoscenza della dottrina che sostiene, ma trovarle anche un fondamento incrollabile" (Contra Academicos, I, 7. 19).

1.4 Scienza e scrittura

La scienza dunque ha le sue verità, ma anche la Scrittura è fonte di verità, e di livello superiore. Non ci può e non ci deve essere un contrasto tra le due. Ma se la scrittura è presa alla lettera, come facevano i Manichei, può anche affermare cose che alla scienza appaiono erronee. Questo era uno dei motivi per cui il giovane Agostino non si era accostato alla Scrittura come avrebbe voluto la madre Monica, profondamente cristiana.

Giunto a Milano, Agostino aveva appreso da Sant'Ambrogio a fare una lettura metaforica della Scrittura, per cavarne le verità più profonde che si celano in essa; tuttavia egli cercò sempre di salvare anche il significato letterale, pur affermando esplicitamente che esso non ha e non vuole avere un contenuto scientifico.

Per esempio, a proposito della forma del cielo vi sono passi delle Scritture in apparente contrasto: Nei salmi si legge "Qui extendit coelum sicut pellem" (Ps 103, 2), ma Isaia dice che "il cielo sta sospeso come una volta" (Is 40, 22 sec. LXX). La prima concezione, così ingenua, è ripresa in un affresco del 1246 che si trova nell'oratorio di San Silvestro adiacente alla chiesa romana dei Santi Quattro Coronati. Accanto alle storie di Costantino e di papa Silvestro I si può ammirare un Giudizio universale, nel quale un angelo arrotola il cielo stellato proprio come fosse un tappeto!

Il primo commento di Agostino è che il problema non è rilevante per la felicità e salvezza eterna: "Che importa infatti se il cielo racchiude da ogni parte, come una sfera, la terra mantenuta in equilibrio al centro del mondo oppure la copra come un disco solo dalla parte superiore? Ma qui è in gioco la credibilità della Scrittura per il motivo più volte da me ricordato" (De Gen. II, 9, 20).

Pertanto Agostino così argomenta: "Ecco perché è necessario dire in breve che i nostri agiografi conoscevano quanto è conforme alla verità per ciò che riguarda la figura del cielo, ma lo Spirito di Dio, che parlava per mezzo di essi, non ha voluto insegnare agli uomini queste cognizioni per nulla utili alla salvezza dell'anima".

Questo argomento, come è ben noto, comprese le due precedenti citazioni, sarà ripreso da Galileo, nella celebre Lettera a Cristina di

Lorena: "E se l'istesso Spirito Santo a bello studio ha pretermesso d'insegnarci simili proposizioni, come nulla attenenti alla sua intenzione, ciò è alla nostra salute, come si potrà adesso affermare, che il tener di esse questa parte, e non quella, sia tanto necessario che l'una sia *de Fide*, e l'altra erronea? … Io qui direi che quello che intesi da persona ecclesiastica costituita in eminentissimo grado [il cardinal Baronio], ciò è l'intenzione delle Spirito Santo essere d'insegnarci come si vadia al cielo, e non come vadia il cielo".

Tuttavia, quando la scienza si pronuncia con certezza, bisogna rivedere l'interpretazione della Scrittura, in modo da evitare ogni conflitto: "Ma se per caso quelli [che attribuiscono al cielo la figura d'una sfera] potessero provare la loro opinione con argomenti di cui non si dovrebbe dubitare, bisognerebbe dimostrare che l'immagine della pelle usata dai nostri Libri sacri non è in contraddizione con quelle argomentazioni razionali" (De Gen. II, 9. 21). O anche: "Qualora ciò si avverasse, non era quello il senso della sacra Scrittura, ma un'opinione dell'umana ignoranza" (De Gen. I, 19, 38).

Non si può non rilevare la consonanza con il pensiero del cardinale San Roberto Bellarmino, a proposito di un tema analogo, quello del sistema eliocentrico, ovviamente contrario alle teorie geocentriche adottate nella Scrittura: "Dico che quando ci fusse vera dimostrazione che il sole stia nel centro del mondo e la terra nel 3° cielo, allhora bisogneria andar con molta consideratione in esplicare le Scritture che paiono contrarie, e più tosto dire che non le intendiamo, che dire che sia falso quel che si dimostra" (*Lettera a Paolo Antonio Foscarini*, 12 aprile 1615, in [1] vol. XII, pp. 171-172).

Entrambi i contendenti, Galileo e Bellarmino, si rifacevano dunque al pensiero di Agostino! Peraltro, Bellarmino così prosegue: "Ma io non crederò che ci sia tal dimostratione, fin che non mi sia mostrata: né è l'istesso dimostrare che supposto ch'il sole stia al centro e la terra nel cielo, si salvino le apparenze, e dimostrare che in verità il sole stia nel centro e la terra nel cielo; perché la prima dimostratione credo ci possa essere, ma della seconda ho grandissimo dubbio, et in caso di dubbio non si dee lasciare la Scrittura Santa, esposta dai Santi Padri". Tale posizione viene lodata come scientificamente corretta da Feyerabend [2], e certamente sarebbe piaciuta ad Agostino.

Per tornare al problema della forma del cielo, è chiaro che Agostino

vede benissimo la contraddizione: "In realtà che c'è di tanto diverso e contrario quanto la superficie piana e distesa d'una pelle e la curva d'una volta convessa?" (De Gen. II, 9. 21). Tuttavia egli cerca comunque di cavarsi d'impaccio anche dal punto di vista letterale: "Se infatti una volta può chiamarsi correttamente non solo curva ma anche piatta, certamente anche una pelle può essere stesa non solo come una superficie piana ma anche in forma d'una tasca rotonda. In realtà sono una pelle non solo un otre ma anche una vescica" (9. 22). La spiegazione appare decisamente "stiracchiata", perché, anche se non poteva certo conoscere il *theorema egregium* che Gauss dimostrò nel 1827, Agostino aveva già ammesso che una volta e un piano sono ben diversi!

2. Il tempo

Uno dei temi più affascinanti della scienza e della filosofia è da sempre la natura del tempo. Diverse concezioni si sono affrontate nel pensiero dei popoli: dal tempo illimitato ab eterno al tempo circolare e ricorrente. Agostino sposa quella del tempo limitato. Oltre che nel De Genesi, vi sono bellissime pagine nelle Confessioni che trattano del tempo. Non si può non citare la famosa dichiarazione, che esprime la difficoltà di affrontare questo concetto: "Quid est ergo tempus? si nemo ex me quaerat, scio; si quaerenti explicare velim, nescio / Cos'è dunque il tempo? Se nessuno m'interroga, lo so; se volessi spiegarlo a chi m'interroga, non lo so" (Conf. XI, 14. 17).

Poco oltre (Conf. XI, 20. 26) Agostino descrive l'aspetto psicologico del tempo: "Tempora sunt tria: praesens de praeteritis, praesens de praesentibus, praesens de futuris. Sunt enim haec in anima tria quaedam et alibi ea non video. Praesens de praeteritis memoria; praesens de praesentibus contuitus; praesens de futuris expectatio / I tempi sono tre: presente del passato, presente del presente, presente del futuro. Questi tre tempi sono nella mia anima e non li vedo altrove. Il presente del passato è la memoria; il presente del presente, la percezione; il presente del futuro l'attesa".

Il rapporto tra tempo e movimento era già stato investigato da numerosi filosofi.

Platone (Timeo, 37d) parla del tempo come di "un'immagine mobile dell'eternità … che procede secondo il numero, e che noi abbiamo chiamato tempo"; e spiega poi che l'*era* e il *sarà* sono forme di

tempo, ma che all'eternità si addice solo l'*è*. Agostino ribadisce (Conf. XI, 11. 13) che "nell'eternità nulla passa, ma tutto è presente, a differenza del tempo, mai tutto presente".

2.1 La creazione del tempo

Ma inframmezzate con queste considerazioni psicologiche vi sono altri ragionamenti più scientifici.

Il tempo per Agostino è stato creato assieme al mondo. Prima non esisteva, anzi, non c'è un prima, perché l'inizio è stato appunto l'atto della creazione. "Ecco come rispondo a chi chiede: «Cosa faceva Dio prima di fare il cielo e la terra?». Non rispondo come quel tale, che, dicono, rispose, eludendo con una facezia l'insidiosità della domanda: «Preparava la geenna per chi scruta i misteri profondi». Altro è capire, altro è schernire. Io non risponderò così. Preferirei rispondere: «Non so ciò che non so», anziché in modo d'attirare il ridicolo su chi ha posto una domanda profonda, e la lode a chi diede una risposta falsa. Invece dico che tu, Dio nostro, sei il creatore di ogni cosa creata; e se col nome di cielo e terra s'intende ogni cosa creata, arditamente dico: «Dio, prima di fare il cielo e la terra, non faceva alcunché». Infatti, se faceva qualcosa, che altro faceva, se non una creatura?" (Conf. XI, 12. 14).

2.2 Tempo e movimento

Nel De Genesi (V 5. 12) Agostino afferma: "Così, dunque, il decorso del tempo iniziò con il movimento – mutamento delle creature; invano quindi si ricerca il tempo prima della creazione, come se fosse possibile trovare il tempo prima del tempo. Se infatti non ci fosse alcun movimento delle creature, spirituali o corporali, mediante il quale al passato succede il futuro attraverso il presente, non vi sarebbe affatto il tempo. La creatura poi non potrebbe muoversi – mutarsi, se non esistesse. Il tempo dunque è iniziato con la creazione anziché la creazione col tempo; l'uno e l'altra poi provengono da Dio, poiché da lui, grazie a lui e in lui sono tutte le cose".

Aristotele afferma che "il tempo è misura di un movimento secondo il prima e il poi" (Phys. IV, 10-11). Al contrario, Plotino dichiara che "quello che viene misurato dalla rotazione celeste, ossia quello che viene rivelato, sarà propriamente il tempo; ma questo non è creato da

quella rotazione, ma semplicemente è messo in luce. Per tale motivo [alcuni filosofi], anziché sostenere che il movimento è la misura del tempo e aggiungere quale cosa sia effettivamente misurata dal movimento, tacendo il fatto che ciò avviene solo per accidente e per un certo tipo di tempo, furono portati a credere che il tempo fosse la misura del movimento e a invertire l'ordine dei fatti" (Enn. III, 7, 12-13).

2.3 Movimento degli astri

Se il tempo è legato al movimento, il più significativo movimento che lo esprime è quello degli astri, che danno anche lo spunto per le prime unità di misura del tempo stesso: anno, mese e giorno, che Agostino, nel De Genesi (II, 14-15) descrive con puntigliosa esattezza, spiegando gli anni bisestili, le fasi della luna e la sua forma sferica, che evidentemente qualcuno, ancora ai suoi tempi, credeva fosse piatta o di forma mutevole.

Si noti che il giorno della Creazione non può che essere allegorico, perché "quando da noi è notte, la presenza della luce illumina le altre parti del mondo che il sole percorre prima di tornare dalla parte ove tramonta a quella ove sorge; per questo motivo nello spazio di tutte le ventiquattro ore c'è sempre, lungo il percorso circolare del sole, una parte [della terra] ov'è giorno e un'altra ov'è notte" (De Gen. I, 10, 21); e certo Dio non stava in un particolare punto della Terra quando creò il mondo!

Tuttavia per Agostino (Conf. XI, 23, 29-30) il tempo non è tale movimento: "Ho udito dire da una persona istruita che il tempo è, di per sé, il moto del sole, della luna e degli astri; e non assentii. Perché il tempo non sarebbe piuttosto il moto di tutti i corpi? Qualora si arrestassero gli astri del cielo, e si muovesse la ruota del vasaio, non esisterebbe più il tempo per misurarne i giri e poter dire che hanno durate uguali, oppure, se si svolgono ora più lenti, ora più veloci, che gli uni sono più lunghi, gli altri meno?".

E prosegue, di nuovo con un esempio tratto dalla Bibbia, cui ho accennato prima citando il card. Bellarmino: "Dunque non mi si dica che il tempo è il movimento dei corpi celesti. Quando il sole si fermò all'appello di un uomo per dargli modo di concludere una battaglia vittoriosa, il sole era fermo, ma il tempo procedeva, tant'è vero che la battaglia fu condotta e finita nello spazio di tempo ad essa sufficiente".

Quest'ultima osservazione naturalmente si riferisce al famoso epi-

sodio narrato nel libro di Giosuè 10, 12-14: "Quel giorno, quando il Signore diede a Israele la vittoria sugli Amorrei, Giosuè pregò il Signore e gridò alla presenza di tutti gli Israeliti: 'Sole, fermati su Gabaon! e tu, luna, sulla valle di Aialon! Il sole si fermò, la luna restò immobile, un popolo si vendicò dei suoi nemici'. Questo avvenimento è descritto nel 'Libro del Giusto'; per quasi un giorno intero il sole restò in alto in cielo, senza avviarsi al tramonto. Un giorno come quello non c'è mai stato né prima né dopo di allora, quando il Signore ubbidì a un essere umano e combatté al fianco d'Israele".

È di nuovo interessante notare come lo stesso esempio sia preso da Galileo, nella Lettera a don Benedetto Castelli, a sostegno della teoria eliocentrica. Molto più semplice, egli argomenta con un lungo ma convincente ragionamento, è pensare che sia stata la Terra ad aver interrotto la sua rotazione e non il Sole, e con esso tutta la volta celeste e il primo mobile, responsabile del movimento diurno degli astri.

3. L'astronomia e l'astrologia

Agostino, così esperto di astronomia, rifiutava con assoluta determinazione l'astrologia. Da giovane invero aveva consultato gli astrologi, come ammette nelle Confessioni (IV, 3): "non rinunziavo ancora a consultare quegli imbroglioni che sono detti *mathematicos*", ma più tardi si era ricreduto e aveva "rifiutato le false divinazioni e le empie fantasticherie degli astrologi (*mathematicorum*)" (Conf., VII, 6).

Questo rifiuto era dovuto, naturalmente, alle Scritture, che in vari passi condannano tali prassi, ma era già implicito nella discussione sui gemelli.

Secondo Agostino "il concepimento dei gemelli, che si attua con una sola unione, come attestano i medici, la cui scienza è molto più sicura ed evidente, avviene in un tempo così rapido da non oltrepassare due secondi". E siccome gli astrologi dei secondi non si curano ("Minutas autem minutarum iam in constellationibus … non inveniunt", 83 questioni diverse, n. 45 – Contro gli astrologi), due gemelli dovrebbero avere lo stesso destino, il che non è. Sono argomentazioni che avrebbe tranquillamente potuto riprendere da Cicerone (De Divinatione, II, 87-99), ma volendosi basare sull'evidenza fattuale diretta riporta un fatto di cronaca, raccontato da un certo Firmino (Conf. VII, 6, 8-10), per cui due ragazzi, nati nello stesso momento, avevano

avuto vite differenti. Del resto, annota, anche Esaù e Giacobbe erano gemelli!

Concludendo queste argomentazioni, Agostino osserva acutamente: "Si dice che [gli astrologi] hanno predetto molte cose vere, ma questo dipende dal fatto che gli uomini dimenticano le loro falsità ed errori. Unicamente preoccupati di quanto si accordava alle loro aspettative, dimenticano ciò che non corrispondeva e ricordano solo gli avvenimenti che capitano accidentalmente, non per arte divinatoria, del tutto inesistente, ma per qualche fortuita coincidenza". Si tratta di un fenomeno psicologico che possiamo verificare ancora ai giorni nostri, e non solo riguardo all'astrologia (che incredibilmente sopravvive ancora) ma anche a riguardo di molti altri discorsi antiscientifici che si diffondono con la complicità dei mezzi di comunicazione e che nessuno è in grado di arginare.

4. Le scienze naturali e la fisica

Vediamo ora alcuni esempi di fenomeni fisici descritti da Agostino. Anche se sovente le descrizioni sono corredate da spiegazioni arzigogolate e assolutamente prive di valore scientifico, esse dimostrano uno spirito di osservazione degno di nota e una vasta conoscenza delle cognizioni dell'epoca. Vorrei anche notare che Agostino, da maestro di retorica qual era, puntigliosamente esamina tutte le spiegazioni alternative e tutti i significati che si possono dare alle parole, perché ovviamente il contesto apologetico non consente la precisione tipica del linguaggio scientifico.

4.1 Le acque sopra il cielo e il ciclo delle acque

Un primo esempio si trova nel commento al secondo giorno della creazione: "Dio inoltre disse: Vi sia un firmamento in mezzo alle acque e tenga separate le acque dalle acque. E così avvenne. E Dio fece il firmamento e separò le acque ch'erano al di sopra del firmamento. E Dio chiamò 'cielo' il firmamento. E Dio vide ch'essa è cosa buona. E venne sera e poi mattina: secondo giorno".

Agostino (De Gen. II, 1.2), dopo aver notato che "per loro natura, queste acque non possono trovarsi al di sopra del cielo stellato poiché, a causa del loro peso, sono costrette a scorrere sulla terra oppure,

evaporando, ad esser trasportate nell'atmosfera vicina alla terra" espone alcune esperienze di fisica. Dapprima nota (2.5) che "un recipiente immerso [nell'acqua] con la bocca rivolta verso il basso non può riempirsi d'acqua: questa è una prova sufficiente che l'aria tende per sua natura verso l'alto. Il recipiente sembra bensì vuoto ma si dimostra esser pieno d'aria quando lo s'immerge nell'acqua con la bocca rivolta verso il basso; in realtà poiché [l'aria] non trova una via d'uscita attraverso la parte rivolta verso l'alto e, data la sua natura, non può passare sotto l'acqua penetrandovi a forza, la respinge dal recipiente ch'è pieno e non ve la lascia entrare". Poi descrive il gorgoglio dell'aria quando fuoriesce perché l'acqua entra nel vaso inclinato.

Peraltro si lascia anche andare a considerazioni del tutto prive di fondamento quando per esempio attribuisce alle acque che stanno sopra il cielo la capacità di raffreddare Saturno, che, secondo lui, siccome nel suo moto diurno si muove più rapidamente degli altri dovrebbe piuttosto essere caldo. (De Gen. II, 5, 9).

Nel libro incompiuto del De Genesi, in cui si riprendono molti dei ragionamenti degli altri libri, è descritto con precisione il ciclo delle acque (14. 47): "D'altra parte la Scrittura, parlando delle sorgenti e dei fiumi, non dice neppure come sono stati creati. Ora, coloro che indagano e discutono questi fenomeni, affermano che, per via del movimento superficiale dell'aria, dal mare si solleva invisibilmente il vapore d'acqua dolce evidentemente per via di queste esalazioni che noi non possiamo osservare in alcun modo: queste esalazioni si condensano in nubi, e in tal modo la terra bagnata dalle piogge emette a goccia a goccia l'acqua in caverne assai occulte e ve la trasuda nella misura in cui, dopo essersi raccolta e penetrata per diversi meati, scaturisce in sorgenti sia piccole sia capaci di formare dei fiumi.

Gli scienziati sostengono che di questo fenomeno è prova il fatto che l'evaporazione dell'acqua marina fatta bollire e raccolta in una serpentina presenta, a chi l'assaggia, acqua dolce. A tutti inoltre appare evidente che le sorgenti diminuiscono di portata poiché risentono l'effetto della scarsezza delle piogge".

4.2 Il cervello

Nel settimo libro del De Genesi Agostino tratta della creazione dell'anima, e del suo rapporto col corpo. In particolare descrive la

struttura del cervello, che secondo la medicina dell'epoca era costituito da tre ventricoli: "il primo vicino al volto, dal quale si dipartono tutti i nervi sensori; il secondo è quello posteriore situato presso la base del cervello, che regola tutti i movimenti; il terzo è sito tra gli altri due, ove gli scrittori dimostrano che ha sede la memoria" (De Gen. VII, 18. 24).

E prosegue con una dimostrazione molto moderna di tale partizione: "L'esistenza di siffatti ventricoli, al dire di quegli scrittori, è dimostrata da segni sicuri, in casi in cui quelle rispettive zone del cervello sono state affette da una malattia o da un difetto patologico. Anche quando sono menomate le funzioni della sensazione o il movimento delle membra o il ricordo dei movimenti del corpo, [i medici] indicano assai chiaramente la funzione di ciascuno dei ventricoli e, applicando a questi la cura [opportuna], hanno appurato a quale delle zone cerebrali ha giovato la cura apprestata". La medicina moderna ovviamente ha strumenti molto fini per isolare le zone attive del cervello in ciascuna delle operazioni che compie, ma concettualmente il metodo è corretto e funzionale.

4.3 Qualche altro errore

Alcune delle teorie scientifiche esposte da Agostino, oltre al geocentrismo sono (per noi moderni!) palesemente erronee.

Per esempio, egli credeva alla generazione spontanea di animali, seguendo sia la Bibbia che Aristotele; ma per confutare tale teoria si dovette attendere il XVII secolo, con le Esperienze intorno alla generazione degl'insetti di Francesco Redi del 1668 seguito, nel secolo successivo, dal Saggio di osservazioni Microscopiche sul Sistema della Generazione de' Signori di Needham e Buffon di Lazzaro Spallanzani (1765) e infine, ancora cent'anni dopo, da Louis Pasteur, che nel 1864 chiuse la questione con una dimostrazione premiata dall'Accademia delle Scienze di Parigi.

Anche nella teoria della visione era in errore, pensando che il raggio si propagasse dagli occhi alla sorgente luminosa e non viceversa. Questo errore forse nasce da una interpretazione fisica del linguaggio puramente geometrico dell'ottica di Euclide, che ovviamente si mette dal punto di vista dell'osservatore per descrivere l'angolo o il cono sotto cui si vede l'oggetto; ma Euclide non aveva alcuna intenzione di fare una teoria fisica della visione.

5. La matematica

Non intendo spendere troppe parole sulla matematica di Sant'Agostino, avendone già trattato diffusamente in una relazione svolta nel convegno Matematica e Letteratura 2 (Fisciano, SA, 6-8 aprile 2016) e riportata negli atti di quel convegno [3].

Non posso però non ricordare che non solo Agostino aveva una conoscenza non banale della aritmetica dei suoi tempi (numeri primi, perfetti, triangolari ecc.), ma che al solito ne faceva strumento di interpretazione allegorica della Scrittura.

Per esempio, dopo aver dimostrato che 6 è un numero perfetto, e avendo anche chiosato, senza uno scopo specifico, ma per il puro piacere di esprimere delle proprietà matematiche, che pure 28 lo è e che i numeri perfetti si diradano sempre più (di fatto, all'epoca se ne conoscevano solo quattro: 6, 28, 496 e 8128), egli spiega così i tempi della creazione: "Noi quindi non possiamo dire che il sei è un numero perfetto per il fatto che Dio ha compiuto tutte le sue opere in sei giorni, ma possiamo dire che Dio ha compiuto le sue opere in sei giorni per il fatto che il sei è un numero perfetto. Questo numero perciò sarebbe perfetto anche se queste opere non ci fossero state; se invece esso non fosse perfetto, Dio non avrebbe compiuto le sue opere attenendosi a questo numero" (De Gen. IV, 7.14). Naturalmente, i sei giorni sono puramente allegorici: "Ecco perché, quando pensiamo alla creazione primordiale degli esseri, cioè alle opere dalle quali Dio si riposò il settimo giorno, non dobbiamo immaginare quei giorni come i nostri giorni solari né l'operazione di Dio come se fosse l'attività con cui ora compirebbe qualcosa nel tempo, ma dobbiamo pensare piuttosto il modo con cui operò ciò da cui cominciò il tempo, il modo cioè con cui fece tutte le cose simultaneamente dando loro anche un ordine risultante non da intervalli temporali ma dalla connessione delle cause; in tal modo gli esseri creati simultaneamente furono anche portati a compimento alla perfezione mediante la ripetizione del "giorno" [della creazione] fatto presente per sei volte" (De Gen., V, 5. 12).

Così i 153 grossi pesci della pesca miracolosa sono per lui simbolo della Chiesa dei santi: infatti "se al numero dieci, proprio della legge, aggiungiamo il numero sette, proprio dello Spirito Santo, abbiamo diciassette. Se si scompone questo numero in tutti i numeri che lo formano, e si sommano tutti questi numeri, si ha come risultato centocin-

quantatré". In termini matematici, il numero triangolare di 17 è appunto 1+2+3+…+17=153 (in Io. Ev. 122, 8).

Ma il punto in cui Agostino si distacca dalla tradizione matematica antica e in qualche modo precorre la modernità è nella concezione della totalità dei numeri non più solo come potenziale, ma come una infinità attuale, compresente nella mente di Dio, e quindi per ciò stesso concepibile anche dall'uomo.

"Riguardo poi all'altra loro teoria che neanche con la scienza di Dio può essere rappresentato l'infinito, rimane loro che osino affermare, immergendosi nell'abisso profondo della irreligiosità, che Dio non conosce tutti i numeri.

È assolutamente certo che il numero è infinito, perché qualunque sia il numero che si prende come limite, non dico che è possibile aumentarlo di un'unità, ma per quanto sia grande e comprensivo di una indefinita quantità numerica, in base all'idea stessa del numero, non solo si può raddoppiare, ma anche moltiplicare per se stesso. Infatti qualsiasi numero è così determinato dalle sue proprietà che non v'è numero eguale ad un altro. Sono dunque disuguali per quantità e qualità, ognuno è finito, e tutti insieme sono infiniti.

Dio dunque non conoscerebbe a causa dell'infinità la totalità dei numeri e la sua scienza arriverebbe fino a una certa quantità numerica e ignorerebbe il resto? Non lo potrebbe dire neanche il più insensato".

Molti altri esempi potrebbero essere riportati a testimonianza della accuratezza della descrizione dei fenomeni fisici e matematici da parte di Sant'Agostino, oltre che della eleganza del suo eloquio, ma spero che questi bastino a illustrare questo aspetto della sua personalità di uomo capace di coniugare così felicemente la fede e la ragione.

NOTA BIBLIOGRAFICA

I testi di Agostino sono tratti, per comodità, dal sito ufficiale [4], che riporta l'intero corpus agostiniano sia in latino che in italiano, oltre che in altre lingue moderne. Per questo non vengono riportati i testi cartacei consultati, anche se ricchi di pregevole materiale commentario.

1. Galilei G., *Opere*, a cura di A. Favaro, Giunti-Barbera, Firenze 1968.
2. Feyerabend P., Galileo e la tirannia della verità, *L'Astronomia* n. 71, 1987, 28-36.
3. Citrini C., Sant'Agostino e la matematica, in *Parole, formule, emozioni*, a cura di Maroscia P., Toffalori C., Tortoriello F.S. Vincenzi G., 145-181, UTET, 2018.
4. www.augustinus.it

Scienze naturali
Biologia

Istituto Lombardo (Rend. Scienze) 150, 183-190 (2016)

GLI ESOSOMI NELLA COMUNICAZIONE CELLULA-CELLULA

STEFANIA RAIMONDO (*)

Nota presentata dal m.e. Gianpiero Sironi
(Adunanza del 20 ottobre 2016)

SUNTO. – La comunicazione cellula-cellula è un processo essenziale per il coordinamento e la corretta organizzazione dei sistemi multicellulari. Le cellule comunicano attraverso svariati meccanismi tra cui il rilascio di fattori di crescita, chemochine, piccole molecole bioattive, il contatto cellula-cellula. Negli ultimi anni è stato descritto un nuovo e sofisticato meccanismo di comunicazione cellulare basato su vescicole extracellulari. Le vescicole extracellulari sono vescicole specializzate, rilasciate nello spazio extracellulare dalla maggior parte dei tipi cellulari, in condizioni fisiologiche e patologiche. Tra i diversi sottotipi di vescicole extracellulari gli esosomi hanno recentemente ricevuto la maggior parte dell'attenzione vista la loro capacità di fungere da messaggeri nella comunicazione intercellulare.

***

ABSTRACT. – Cell to cell communication is essential for the coordination and proper organization of different cell types in multicellular systems. Cells exchange information through a multitude of mechanisms such as secreted growth factors and chemokines, small molecules (peptides, ions, bioactive lipids and nucleotides), cell-cell contact and the secretion of extracellular matrix components. Over the last few years a new and sophisticated mechanism of cell-cell communication based on extracellular vesicles has been described. Extracellular vesicles are specialized vesicles released in the extracellular space by most of cell types, under physiological and pathological conditions. Among different extracellular vesicles subtypes, exosomes (30-100 nm) have recently received most of the attention do to their ability to be messenger in intercellular communication.

(*) Dipartimento di Biomedicina, Neuroscienze e Diagnostica avanzata, Sezione di Biologia e Genetica, Università degli Studi di Palermo, Palermo, Italia.
E-mail: stefania.raimondo@unipa.it

1. La comunicazione cellula-cellula

E' ormai ampiamente descritto che nessuna cellula vive isolata nel nostro organismo e che esiste una complessa rete di comunicazione tra le cellule che garantisce l'omeostasi di tessuti e organi. Le cellule comunicano e interagiscono tra loro tramite il fenomeno della segnalazione cellulare. In alcuni casi la comunicazione cellula-cellula è diretta, mediante giunzioni serrate o comunicanti, in altri casi le cellule devono poter comunicare a distanza: una cellula segnalatrice produce una molecola segnale, riconosciuta da una cellula bersaglio, per mezzo di una proteina recettore, che a sua volta produce segnali intracellulari. L'intero processo che traduce l'informazione portata dal messaggero extracellulare in cambiamenti intracellulari è chiamato trasduzione del segnale (*Fig. 1*). I segnali prodotti possono avere efficacia a breve o ampio raggio, in base a questo i segnali extracellulari si dividono in endocrini, paracrini e autocrini. Nella segnalazione endocrina le cellule rilasciano nel sangue sostanze dette ormoni, ovvero segnali che agiscono su cellule bersaglio lontane. Nella segnalazione paracrina, la cellula bersaglio si trova in prossimità della cellula segnale e la sostanza che funge da segnale agisce sul gruppo di cellule bersaglio adiacenti. Infine nella segnalazione autocrina le cellule agiscono contemporaneamente da segnale e da bersaglio.

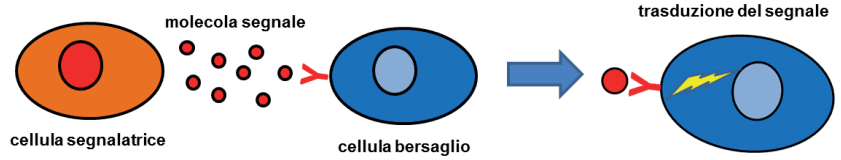


Fig. 1 – Trasduzione del segnale. Una cellula segnalatrice produce una molecola segnale che viene riconosciuta dalla cellula bersaglio per mezzo di un recettore. Il legame della molecola segnale al recettore attiva una via biochimica determinando una risposta cellulare.

2. Le vescicole extra-cellulari

Nell'ultimo decennio è stato descritto un nuovo meccanismo di comunicazione cellula- cellula, basato sul rilascio di un gruppo eterogeneo di vescicole lipoproteiche che fungono da veicoli di informazioni intercellulari [1] . Questi vettori, rilasciati da cellule animali e vegetali [2], e generalmente definite vescicole extracellulari, possono essere suddivisi in tre

classi principali: esosomi, microvescicole e corpi apoptotici. Tra queste, gli esosomi hanno recentemente ricevuto maggiore attenzione. Mentre gli esosomi sono vescicole di origine endocitica, omogenee in dimensione e morfologia (40-100 nm), i corpi apototici sono rilasciati dalle cellule in fase di morte cellulare programmata, e infine le microvescicole derivano direttamente dalla membrana plasmatica e mostrano una dimensione variabile (100-1000 nm). Le vescicole extracellulari vengono inoltre denominate in base sulla base della loro origine, per esempio i prostasomi sono le vescicole secrete dalle cellule epiteliali della prostata e secrete nel liquido seminale [3], gli oncosomi quelle rilasciate dalle cellule tumorali [4].

3. Gli esosomi

Il termine "esosomi" è stato introdotto per la prima volta nel 1983, per identificare, negli eritroblasti in via di differenziamento, le vescicole di membrane in grado di inglobare al loro interno il recettore della trasferrina (Tfr) [5]. Più recentemente la ricerca si è focalizzata su queste vescicole sia perché si sono rivelate dei veri e propri vettori di scambio di informazioni biologiche tra le cellule, sia per il loro ruolo in numerose condizioni fisiologiche e patologiche [6]. Gli esosomi sono una classe di nanoparticelle bioattive, del diametro compreso tra i 30 e i 100 nm e derivate dal compartimento endosomiale. Gli esosomi vengono rilasciati in vitro da diversi tipi di cellule, tra cui le cellule del sistema immunitario, le cellule staminali mesenchimali, le cellule epiteliali, le cellule endoteliali e cellule tumorali di quasi tutti gli istotipi. Inoltre gli esosomi sono stati purificati da una varietà di fluidi corporei come plasma, urina, saliva, liquido cerebrospinale, amniotico e fluido sinoviale. Gli esosomi hanno una specifica morfologia e, come indicato da analisi biochimiche e proteomiche, hanno una membrana arricchita in particolari lipidi (colesterolo, sfingomielina e ceramide) e una composizione proteica unica che li caratterizza e ne rende facile l'identificazione. Anche se gli esosomi provenienti da cellule diverse hanno una differente composizione proteica, alcune proteine sono comuni a tutti i tipi di esosomi; tra queste vi sono Alix, alcune tetraspanine e proteine chaperon. Ulteriori studi hanno inoltre evidenziato la presenza di acidi nucleici all'interno degli esosomi, quali DNA [7], mRNA [8], microRNA [9] e long non-coding RNA [10] (*Fig. 2*). Numerosi studi hanno dimostrato che la presenza di questi acidi nucleici all'interno degli esosomi non è casuale ma è finemente regolato da processi selettivi.

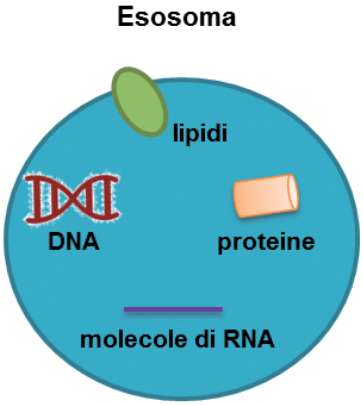


Fig. 2 – Composizione degli esosomi. Gli esosomi sono costituiti da un doppio strato lipidico, contengono proteine, metaboliti e molecole di RNA (mRNA, microRNA, long non-coding RNA).

3.1 Gli esosomi in condizioni fisiologiche e patologiche

Gli esosomi sono ad oggi considerati parte integrante del microambiente intercellulare e possono agire come regolatori della comunicazione cellula-cellula [6]. Gli esosomi possono attivare meccanismi di trasduzione del segnale mediante interazioni fisiche ligando/recettore o mediante fusione con le cellule bersaglio e trasferimento dei loro contenuti. Possono anche essere endocitate dalle cellule bersaglio oppure rilasciare il loro contenuto nello spazio extracellulare. Essi possono agire come modulatori del sistema immunitario con effetti immunosoppressivi o promuovendone l'attivazione; sono inoltre implicati nel riciclo e smaltimento di proteine di membrana e nello scambio di materiale genetico. Una volta rilasciati dalla cellula che li produce, gli esosomi possono influenzare il comportamento della cellula bersaglio sia localmente sia a distanza attraverso i fluidi biologici. Gli esosomi sono stati purificati dal latte materno e diversi studi sottolineano come questi siano in grado di regolare l'immunità materno-fetale. Sono inoltre stati ritrovati nelle urine dove giocano un ruolo nel regolare la funzionalità renale. Un numero crescente di evidenze sperimentali sottolinea un ruolo degli esosomi nell'insorgenza e progressione di svariate patologie, quali cancro, malat-

tie neurodegenerative, malattie infiammatorie, reumatologiche, cardiache e muscolari.

3.2 Gli esosomi come biomarcatori

Gli esosomi a causa della loro origine cellulare, del loro ruolo sia in condizioni fisiologiche e patologiche e dei recenti sviluppi tecnologici che consentono la loro rapida purificazione e caratterizzazione, hanno cambiato in modo significativo molte aree delle scienza clinica. I primi approcci per l'applicazione degli esosomi in diagnostica clinica risalgono a diversi anni fa. Infatti, grazie al loro contenuto proteico e di acidi nucleici, che riflette la natura e lo stato delle cellule che li rilasciano, gli esosomi sono considerati una ricca fonte di informazioni [11]. Le molecole veicolate dagli esosomi sono utili pertanto ai fini diagnostici e per il monitoraggio di trattamenti terapeutici; inoltre, il contenuto, che è ben protetto da un doppio strato lipidico che conferisce un elevato grado di stabilità, comprende alcuni biomarcatori che sarebbero altrimenti scarsamente rilevabili. Molti progressi sono stati fatti in campo oncologico, dove micro-RNA contenuti negli esosomi sono stati associati alla progressione della patologia. Ad esempio, Rabinowits e colleghi hanno dimostrato la differenza significativa in livelli di esosomi e di miRNA tra pazienti affetti da tumore al polmone e pazienti sani, suggerendo che i microRNA veicolati dagli esosomi potrebbero essere utili per lo sviluppo di test di screening per questa neoplasia [12]. Così come in ambito oncologico, gli esosomi possono rappresentare importanti biomarcatori in numerose patologie quali quelle reumatologiche [13], neurodegenerative [14] e cardiache [15]. E' ormai noto che citochine, fattori di crescita, molecole di adesione e proteine della matrice extracellulare sono secrete dalle cellule tumorali e mediano la comunicazione cellula-cellula all'interno del microambiente. Recentemente gli esosomi sono stati considerati come veicoli di queste molecole e, per questa ragione, sono considerati nuovi attori nel crosstalk tra le cellule del microambiente tumorale [16].

3.3 Gli esosomi: un nuovo sistema di drug delivery

Il drug delivery è lo sviluppo di sistemi alternativi di indirizzamento dei farmaci nell'organismo, con l'obiettivo si circoscriverne l'effetto biologico su una cellula bersaglio, migliorandone l'efficacia e ridu-

cendone la tossicità. Ad oggi tante sono le informazioni riguardanti gli esosomi come sistemi di trasporto di citochine, fattori di crescita, molecole di adesione e acidi nucleici. Gli esosomi potrebbero anche essere utilizzati nel campo della terapia genica in quanto possono fungere da veicolo di composti terapeutici rappresentati da miRNA, siRNA e agenti chemioterapici [17] (*Fig. 3*). Di più recente interesse è l'utilizzo di esosomi come "strumenti terapeutici", in quanto recenti studi che analizzano i profili proteomici e trascrittomici di queste vescicole, hanno dimostrato che esprimono molecole sulla loro superficie, alcune delle quali si ritrovano in tutti gli esosomi (LAMP2B e CD63) mentre altre sono tessuto-specifiche e questo può essere sfruttato per indirizzare in modo specifico esosomi caricati con farmaci selettivi per determinate cellule [18]. Il drug delivery che utilizza esosomi può fornire vantaggi unici rispetto ad altri sistemi, tra cui una limitata immunogenicità ed una maggiore stabilità nel sangue. Le membrane degli esosomi possono inoltre essere modificati per mostrare epitopi che possano legarsi a specifici ligandi presenti su determinati tipi cellulari e migliorarne il delivery specifico verso tale tipo cellulare.

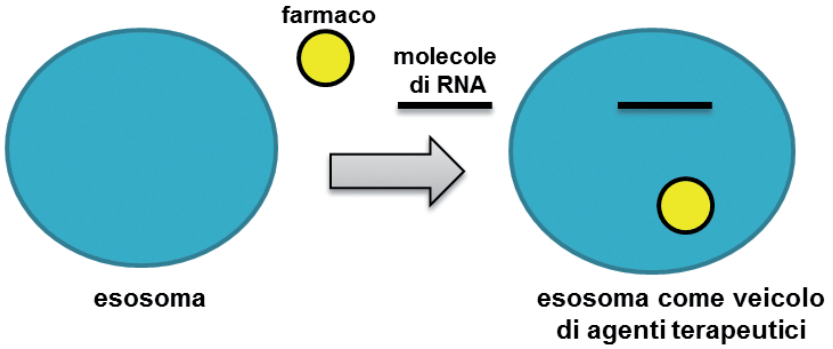


Fig. 3 – Gli esosomi come veicoli di agenti terapeutici. Gli esosomi possono essere impiegati per veicolare farmaci o molecole di RNA terapeutiche su cellule bersaglio.

Ringraziamenti

Ringrazio il Prof. Riccardo Alessandro per il supporto scientifico. Ringrazio inoltre la Fondazione Italiana per la ricerca sul Cancro (FIRC) per il finanziamento ricevuto.

BIBLIOGRAFIA

1. Raposo G and Stoorvogel W. Extracellular vesicles: exosomes, microvesicles, and friends. J Cell Biol. 2013; 200(4):373-383.
2. Raimondo S, Naselli F, Fontana S, Monteleone F, Lo Dico A, Saieva L, Zito G, Flugy A, Manno M, Di Bella MA, De Leo G and Alessandro R. Citrus limon-derived nanovesicles inhibit cancer cell proliferation and suppress CML xenograft growth by inducing TRAIL-mediated cell death. Oncotarget. 2015; 6(23):19514-19527.
3. Aalberts M, Stout TA and Stoorvogel W. Prostasomes: extracellular vesicles from the prostate. Reproduction. 2014; 147(1):R1-14.
4. Meehan B, Rak J and Di Vizio D. Oncosomes - large and small: what are they, where they came from? J Extracell Vesicles. 2016; 5:33109.
5. Harding C, Heuser J and Stahl P. Receptor-mediated endocytosis of transferrin and recycling of the transferrin receptor in rat reticulocytes. J Cell Biol. 1983; 97(2):329-339.
6. Corrado C, Raimondo S, Chiesi A, Ciccia F, De Leo G and Alessandro R. Exosomes as intercellular signaling organelles involved in health and disease: basic science and clinical applications. Int J Mol Sci. 2013; 14(3):5338-5366.
7. Aucamp J, Bronkhorst AJ, Badenhorst CP and Pretorius PJ. A historical and evolutionary perspective on the biological significance of circulating DNA and extracellular vesicles. Cell Mol Life Sci. 2016; 73(23):4355-4381.
8. Valadi H, Ekstrom K, Bossios A, Sjostrand M, Lee JJ and Lotvall JO. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. Nat Cell Biol. 2007; 9(6):654-659.
9. Thind A and Wilson C. Exosomal miRNAs as cancer biomarkers and therapeutic targets. J Extracell Vesicles. 2016; 5:31292.
10. Hewson C and Morris KV. Form and Function of Exosome-Associated Long Non-coding RNAs in Cancer. Curr Top Microbiol Immunol. 2016; 394: 41-56.
11. Gonzalez E and Falcon-Perez JM. Cell-derived extracellular vesicles as a platform to identify low-invasive disease biomarkers. Expert Rev Mol Diagn. 2015; 15(7):907-923.
12. Rabinowits G, Gercel-Taylor C, Day JM, Taylor DD and Kloecker GH. Exosomal microRNA: a diagnostic marker for lung cancer. Clin Lung Cancer. 2009; 10(1):42-46.
13. Withrow J, Murphy C, Liu Y, Hunter M, Fulzele S and Hamrick MW. Extracellular vesicles in the pathogenesis of rheumatoid arthritis and osteoarthritis. Arthritis Res Ther. 2016; 18(1):286.
14. Howitt J and Hill AF. Exosomes in the Pathology of Neurodegenerative Diseases. J Biol Chem. 2016; 291(52):26589-26597.
15. Kishore R, Garikipati VN and Gumpert A. Tiny Shuttles for Information Transfer: Exosomes in Cardiac Health and Disease. J Cardiovasc Transl Res. 2016; 9(3):169-175.
16. Raimondo S, Corrado C, Raimondi L, De Leo G and Alessandro R. Role of

Extracellular Vesicles in Hematological Malignancies. Biomed Res Int. 2015; 2015:821613.
17. Kim OY, Lee J and Gho YS. Extracellular vesicle mimetics: Novel alternatives to extracellular vesicle-based theranostics, drug delivery, and vaccines. Semin Cell Dev Biol. 2016.
18. Ha D, Yang N and Nadithe V. Exosomes as therapeutic drug carriers and delivery vehicles across biological membranes: current perspectives and future challenges. Acta Pharm Sin B. 2016; 6(4):287-296.

Scienze naturali
Biologia

Istituto Lombardo (Rend. Scienze) 150, 191-202 (2016)

IDENTIFICATION AND FUNCTIONAL CHARACTERIZATION OF GERMLINE GENETIC VARIANTS PREDISPOSING TO DIFFERENTIATED THYROID CANCER RISK

GISELLA FIGLIOLI (*), *et al.*[#]

Nota presentata dal s.c. Andrea Mattevi
(Adunanza del 3 novembre 2016)

SUNTO. – Il carcinoma tiroideo differenziato (DTC) è il tumore endocrino più frequente ed in Italia si registra una elevata incidenza. Il rischio d'insorgenza di questa malattia è significativamente maggiore nei parenti di primo grado degli individui affetti suggerendo che anche le varianti genetiche germinali possano contribuire alla sua insorgenza. Precedenti studi di associazione sull'intero genoma (GWASs) hanno permesso l'identificazione di polimorfismi a singolo nucleotide (SNPs) associati con il rischio di DTC localizzati nelle regioni 2q35, 9q22.33 e 14q13.3. Al fine di individuare nuovi *loci* di suscettibilità al DTC, è stato eseguito un GWAS sulla popolazione italiana. In seguito, gli SNPs più interessanti sono stati analizzati in una popolazione italiana più ampia e in altre casistiche europee. I risultati di questo studio hanno confermato il ruolo del *loci* 2q35 e 9q22.33 nella suscettibilità al DTC. Inoltre, i polimorfismi nelle regioni 14q24.3 e 20q11.22-q12 sono risultati associati ad un incremento del rischio di DTC nell'analisi combinata di tutte le popolazioni analizzate, e gli SNPs localizzati in 3q25.32, 5q14, 7q21, 9q34.3, 11p15, 13q12.12 e 20p11 sono risultati associati alla malattia soltanto negli Italiani. Secondo i dati del progetto ENCODE, molti di questi polimorfismi si trovano all'interno di regioni di regolazione della trascrizione e lo studio delle eQTL ha mostrato che cinque degli SNPs identificati in questo studio sono associati alla regolazione dei loro geni più vicini in diversi tessuti, compreso il tessuto tiroideo. In conclusione, tramite questo studio sono stati identi-

(*) IFOM, the FIRC Institute of Molecular Oncology, Milan, Italy.
E-mail: gisella.figlioli@ifom.eu

(#) La nota è il risultato di un lavoro di ricerca condotto in collaborazione con: Aleksandra Köhler, Bowang Chen, Rossella Elisei, Cristina Romei, Kari Hemminki, Stefano Landi, Asta Försti, Federica Gemignani.

ficati nuovi alleli di suscettibilità al DTC ed è stato proposto un loro possibile ruolo funzionale.

***

ABSTRACT. – Differentiated thyroid cancer (DTC) is the most common endocrine tumor, showing a high incidence in Italy. A significant higher risk of this cancer is described in first-degree relatives of DTC patients compared to the general population suggesting that germline genetic variants may contribute to its development. Previous genome-wide association studies (GWASs) on DTC have identified robust associations with single nucleotide polymorphisms (SNPs) at chromosomes 2q35, 9q22.33 and 14q13.3. In order to identify additional DTC susceptibility *loci*, a novel GWAS on the Italian population was conducted. The GWAS was followed by validation studies, where the most interesting SNPs were replicated in a larger Italian population and other European cohorts. Previously observed association for 2q35 and 9q22.33 was confirmed. Moreover, a strong relationship of DTC risk was found with SNPs on 14q24.3 and 20q11.22-q12 across all populations and SNPs on 3q25.32, 5q14, 7q21, 9q34.3, 11p15, 13q12.12 and 20p11 only among Italians. According to ENCODE Project data, many of these SNPs are located in transcription regulatory regions and eQTL analyses showed that five of the associated SNPs may affect the expression levels of their closest genes in different human tissues, including thyroid. In conclusion, novel DTC risk alleles were identified and new insights into their possible functional role were discovered.

1. INTRODUCTION

Thyroid cancer (TC) comprises approximately 1% of all human malignancies and it is the most common endocrine malignancy, representing up to 80% of all cancers originating from endocrine organs [1]. Worldwide, thyroid cancer incidence varies in different geographic regions and it is overall higher in more economically developed countries. In Europe particularly elevated age-standardised rates (ASRs) were observed in Lithuania (ASR=15.5/100,000), Italy (ASR=13.5/100,000) and Austria (ASR=12.4/100,000) (http://eco.iarc.fr/EUCAN/ and http://globocan.iarc.fr).

Approximately 90% of diagnosed TCs arise from the thyrocytes, the thyroid hormone–producing cells, and are classified as differentiated thyroid cancers (DTCs). Among them, the most frequent subtype is papillary (PTC, 75%), followed by follicular (FTC, 10%), Hurthle cells (5%), and poorly differentiated carcinomas (1%–6%). Only a small proportion of TCs, called medullary thyroid cancer (MTC), originates from thyroid parafollicular cells [2].

There are few known TC risk factors except for exposure to ionizing radiation, female gender and a previous benign thyroid disease

[3]. Interestingly, TC is characterized by having one of the highest heritability among all cancer sites, as demonstrated by a significantly higher risk in first-degree relatives of patients with DTC than in the general population and suggesting that also germline genetic variants (*e.g.* single nucleotide polymorphisms, SNPs) may contribute to the risk of the disease [4].

In order to identify SNPs associated with DTC genetic predisposition case-control association studies were conducted [5]. To date, candidate gene/pathway association studies remain the most prevalent type of investigation with more than 100 studies published and more than 300 SNPs examined. Of them, SNPs within genes involved in DNA repair pathways, cell-cycle control, xenobiotic metabolism and in the MAPK pathway were frequently investigated. While some of these variants could represent true associations with DTC, many failed to be replicated among additional populations and could be false-positive [6].

During the past decade genome-wide association studies (GWASs) emerged worldwide as an unbiased approach, independent of *a priori* knowledge on the disease, to reveal SNPs and genomic regions associated with human cancer risk. GWASs rely on the phenomenon of linkage disequilibrium (LD) wherein SNPs are not inherited individually but instead are in LD blocks, with many nearby SNPs being highly correlated. This allows obtaining information on ~50,000 variants by analyzing only one SNP (formally named "tag-SNP") and therefore it may be used to study the entire genome by evaluating ~500,000-700,000 tag-SNPs [7]. GWASs are commonly performed using a multistage approach. In the first stage (*e.g.* the discovery phase), all tag-SNPs are tested on a small subset of cases and control. During the second stage (*e.g.* the validation phase), significant tag-SNPs, are analyzed in a population similar to that used in the previous phase with the purpose of ruling out false positive associations. Moreover, validations in different populations represent also an important task to check if the observed effects are specific for the discovery population or if they may be extended to different geographic groups [8, 9]. Once the association of a SNP is confirmed, the next challenge lies in discovering its functional role and in the identification of its target gene. While few GWAS-identified SNPs are predicted to disrupt protein-coding regions, the great majority of them, approximately 88%, lie in intergenic or intronic regions and could be located within transcriptional regulatory elements (such as promoters, enhancers, transcription factor binding sites [TFBSs] and DNaseI hypersensitive sites [DHSs]).

All these functional elements of the human genome were annotated by the ENCODE (ENCylopedia Of DNA Elements) project and they could be mined to uncover the possible functional role of GWAS-identified SNPs [10, 11]. Moreover, it was clearly reported that inherited genetic variants within several genomic regions, called expression quantitative-trait *loci* (eQTLs), are associated with expression of many transcripts. Thus, studying the associations between SNPs and gene expression levels may represent a useful way to connect risk variants to their putative genes or transcripts. eQTLs can be located either near (1 Mb) the gene they regulate (*cis*-eQTL) or at a significant distance away from it (*trans*-eQTL) [12, 13]. At this regard, the Genotype-Tissue Expression (GTEx) program has recently provided expression levels of human genes in transformed fibroblasts and in several tissues from healthy individuals allowing the characterization of the eQTLs and the interpretation of GWAS-identified SNPs [14].

In 2009 Gudmusson and collaborators published the first GWAS on DTC. The strongest signals were found for the SNP rs965513 on 9q22.33, 57 kb upstream *FOXE1* with (OR=1.75, 95% CI 1.59-1.94) and rs944289 on 14q13.3, near *NKX2-1* (OR=1.37, 95% CI 1.24- 1.52) [15]. The importance of *FOXE1* in the disease genetic predisposition was confirmed in a second GWAS performed on radiation-related PTCs. In particular, the association for rs965513 was confirmed (OR=1.65, 95% CI 1.43-1.91) and a statistically significant association was found for rs1867277 (OR=1.48, 95% CI 1.27-1.71) [16]. Additional DTC risk variants were found through a GWAS on circulating TSH levels. In this study the strongest association was observed for rs116909374 on 14q13.3 with an OR of 2.09 (95% CI 1.68-2.60). Similarly, the variants rs966423 on 2q35 (within *DIRC3*) and rs2439302 on 8p12 (within *NRG1*) were significantly associated with DTC with OR of 1.34 (95% 1.22-1.47) and 1.36 (95% CI 1.23-1.50), respectively [17].

In order to search for additional variants predisposing to DTC a novel GWAS on the high-incidence Italian population was performed thanks to a successful collaboration among the Department of Endocrinology and Metabolism of the University Hospital Cisanello of Pisa (Italy), the Department of Biology of the University of Pisa (Italy) and the Division of Molecular Genetic Epidemiology of the German Cancer Research Center (DKFZ, Heidelberg, Germany). The GWAS was followed by the replication of SNPs that showed suggestive evidence of association in other European cohorts and by the assessment

of the cumulative effect of the identified SNPs. The functional role of the best-associated SNPs was also investigated by using ENCODE project experimental data and by eQTL analyses.

2. Risk variants associated with differentiated thyroid cancer risk

The local Ethical Committee approved this study and all participants gave their written informed consent to participate according to Helsinki declaration.

The discovery phase of the GWAS was performed on Italian histologically confirmed DTC patients and healthy controls. A total of 701 cases were recruited at the Department of Endocrinology and Metabolism of the University Hospital Cisanello in Pisa. The control group comprised 499 healthy individuals without any thyroid disease and cancer history and included workers who underwent a routine check-up at the same hospital of Pisa and blood donors from the Meyer Hospital in Florence. After the application of strict quality controls 572,042 SNPs were analyzed for association with DTC in 690 cases and 497 controls. The results of this phase confirmed the role of the SNPs identified in previous GWASs on DTC predisposition. In particular, a robust evidence for a relationship between *FOXE1* and the risk of the disease with multiple SNPs annotating in the region having *p-value*$<5.0\times10^{-8}$. Because the association between *FOXE1* and DTC was already well-established, SNPs in this region were excluded from the following phases of analysis (*Fig. 1*) [18].

Validation studies were conducted on four European cohorts. The Italian cohort (Italian1+Italian2) included 1,539 patients and 1,719 controls collected in the same hospitals of the samples of the discovery phase. The Polish group comprised 468 patients with DTC and 470 healthy controls from the Department of Nuclear Medicine and Endocrine Oncology, Maria Sklodowska-Curie Memorial Cancer Center and Institute of Oncology in Gliwice. The Spanish cohort consisted of 446 patients, recruited at the Department of Genetics and Microbiology of Autonomous University of Barcelona, and 420 healthy individuals. The United Kingdom series consisted of 509 cases, ascertained through the Institute of Cancer Research/Royal Marsden Hospital National Health Service Trust, and 1,118 controls recruited through the National Study of

Colorectal Cancer. A total of 109 SNPs, selected within regions showing stronger association signals (*e.g.* lower *p-value*) with DTC risk, were validated in the larger Italian population. Of them, 24 SNPs positively replicate the GWAS associations and were also analyzed in Polish, UK and Spanish cohorts. The joint analysis of all populations (2,985 cases and 3,727 controls) revealed a genome-wide significant association with DTC for rs6759952 near *DIRC3* on 2q35 (OR=1.25, 95% CI 1.16-1.34, *p-value*=6.4×10^{-10}), rs10136427 near *BATF* on 14q24.3 (OR=1.30, 95% CI 1.17-1.44, *p-value*=9.30×10^{-7}) and rs7267944 near *DHX35* on 20q11.22-q12 (OR=1.32, 95% 1.20-1.46, *p-value*=1.34×10^{-8}). The role of the SNPs was also assessed only in the Italian population, totaling 2,260 cases and 2,218 controls. The most significant associations were observed for rs7617304 upstream *RARRES1* on 3q25.32 (OR=1.25, 95% CI 1.12-1.39, *p-value*=4.6×10^{-5}), rs13184587 within *ARSB* on 5q14.1 (OR=1.28, 95% CI 1.15-1.43, *p-value*=8.54×10^{-6}), rs10238549 (OR=1.27, 95% CI 1.15-1.40, *p-value*=4.1×10^{-6}) and rs7800391 downstream *IMMP2L* on 7q21 (OR=1.25, 95% CI 1.14-1.38, *p-value*=5.7×10^{-6}), rs10781500 downstream *SNAPC4* on 9q34.3 (OR=1.23, 95% CI 1.12-1.36, *p-value*=3.5×10^{-5}), rs7935113 within *GALNTL4* on 11p15.3 (OR=1.36, 95% CI 1.20-1.53, *p-value*=7.41×10^{-7}), rs1220597 within *SPATA13* on 13q12.12 (OR=1.26, 95% CI 1.14-1.38, *p-value*=3.25×10^{-6}) and rs1203952 upstream *FOXA2* on 20p11 (OR=1.29, 95% CI 1.16-1.44, *p-value*=4.42×10^{-6}) (*Tab. 1*) [18-20]. The cumulative effect of the 11 independent susceptibility SNPs in 10 genes identified in the Italian population was also evaluated. A dose-dependent increase in risk of DTC was observed with an increasing number of risk alleles (OR=1.30, 95% CI 1.26-1.35, *p-trend*=3.13×10^{-47}). In particular, individuals carrying ≥14 risk alleles had 7.68 times higher risk of getting DTC as compared to those with ≤7 risk alleles (*Fig. 2*) [19].

3. Bioinformatic and eQTL analyses of the risk variants

Computational approaches were employed to functionally annotate the associated SNPs. Briefly, ENCODE Project data were quarried using the HaploReg v2 tool and eQTL analyses were performed taking advantage of data free available on GTEx Portal [14, 21]. Among the three variants that were found associated with the risk of DTC in the combined analysis of all the European populations, rs6759952 (*DIRC3*) and rs10136427 (*BATF*) were found to map in regions of weak and

strong enhancers. Both SNPs alter the binding sites for regulatory protein: rs6759952-T risk allele introduces the consensus sequence for Pou1f1, Pou2f2, TLX1 and YY1 proteins and rs10136427-C risk allele removes a FOXD3 site. Little support for the functional role of rs7267944 (*DHX35*) was found.

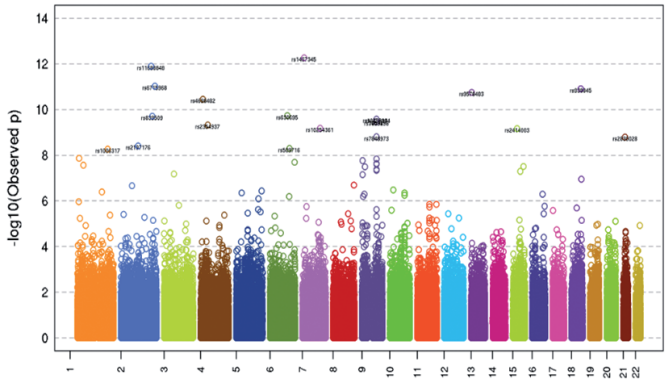


Fig. 1 – Manhattan plot of simple ^{2} allelic test of association *p-values* from the GWAS. The plot shows the –log10*p-values* for each tagSNP against chromosomal location. Values for each chromosome are reported in different colors for visual effect.

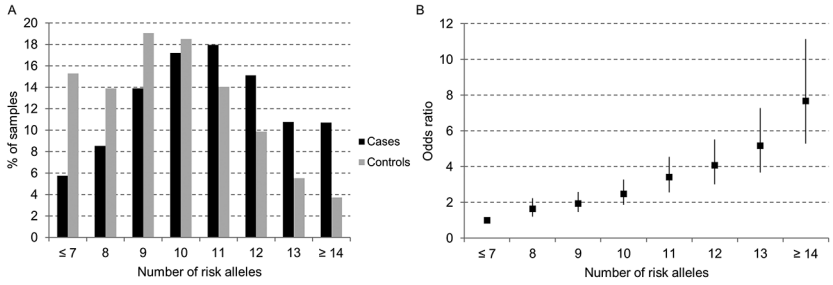


Fig. 2 – Cumulative risk assessment. (A) Sample distribution according to the number of risk alleles in eleven SNPs associated in the Italian DTC cases (black columns) and controls (grey columns). (B) Plot of the increasing ORs for DTC with increasing number of risk alleles. The category ≤7 was chosen as reference (OR=1.0); vertical bars correspond to 95% confidence intervals.

A possible regulatory consequence was identified for all the SNPs associated with the disease only in the Italian population, except for rs1220597 (*SPATA13*). The variants rs7617304 (*RARRES1*), rs10238549 and rs7800391 (*IMMP2L*) and rs7935113 (*GALNTL4*) are

Tab. 1 – Risk of differentiated thyroid cancer associated with single nucleotide polymorphisms identified in the present study. Significant results at a genome-wide level are highlighted in bold.

SNP	*Locus*	Closest Gene	Risk allele	Population	Number of cases/controls	Risk allele frequency (cases/controls	Allelic OR (95% CI)(*)	*p-value*(*)
rs6759952	2q35	*DIRC3*	T	GWAS	649/431	0.47/0.38	1.44 (1.21-1.72)	4.7×10^{-5}
				Italian1	1160/927	0.49/0.42	1.32 (1.17-1.49)	9.8×10^{-6}
				Italian2	322/710	0.44/0.40	1.16 (0.96-1.41)	0.12
				Polish	413/420	0.46/0.44	1.07 (0.88-1.30)	0.48
				Spanish	397/393	0.51/0.39	1.58 (1.29-1.93)	8.2×10^{-6}
				UK	497/1090	0.46/0.43	1.14 (0.98-1.32)	0.09
				Italian cohorts	2131/2068	-	1.30 (1.18-1.43)	7.3×10^{-8}
				Joint analysis	**3438/3971**	**-**	**1.25 (1.16-1.34)** P_{het}=0.03, I^2=58.5%	**6.4×10^{-10}**
rs7617304	3q25.32	*RARRES1*	A	GWAS	649/430	0.31/0.25	1.37 (1.13-1.67)	1.4×10^{-3}
				Italian1	1170/933	0.28/0.25	1.16 (1.01-1.34)	0.03
				Italian2	320/705	0.31/0.27	1.22 (1.00-1.50)	0.05
				Polish	437/431	0.19/0.21	0.92 (0.73-1.17)	0.50
				Spanish	388/394	0.26/0.25	1.09 (0.87-1.37)	0.47
				Italian cohorts	**2139/2068**	**-**	**1.25 (1.12-1.39)**	**4.6×10^{-5}**
				Joint analysis	2964/2893	-	1.17 (1.07-1.27) P_{het}=0.12, I^2=46.0%	7.0×10^{-4}
rs1318458 7	5q14.1	*ARSB*	G	GWAS	646/414	0.77/0.69	1.51 (1.24-1.83)	3.60×10^{-5}
				Italian1+2	1429/1541	0.76/0.72	1.19 (1.06-1.33)	4.15×10^{-3}
				Polish	442/434	0.76/0.74	1.10 (0.88-1.36)	0.41
				Spanish	382/403	0.72/0.74	0.91 (0.73-1.14)	0.42
				Italian cohorts	**2075/1955**	**-**	**1.28 (1.15-1.43)**	**8.54×10^{-6}**
				Joint analysis	2899/2792	-	1.17 (1.07-1.27) $P_{het}=6.4\times10^{-3}$, I^2=75.6%	7.16×10^{-4}
rs1023854 9	7q21	*IMMP2L*	C	GWAS	649/431	0.72/0.63	1.48 (1.23-1.77)	3.2×10^{-5}
				Italian1	1170/933	0.68/0.64	1.21 (1.06-1.37)	4.7×10^{-3}
				Italian2	323/695	0.71/0.68	1.16 (0.94-1.42)	0.16
				Polish	451/438	0.76/0.73	1.17 (0.94-1.45)	0.16
				Spanish	390/391	0.65/0.72	0.73 (0.59-0.91)	4.4×10^{-3}
				UK	503/1101	0.72/0.71	1.02 (0.86-1.20)	0.85
				Italian cohorts	**2142/2059**	**-**	**1.27 (1.15-1.40)**	**4.1×10^{-6}**
				Joint analysis	3486/3989	-	1.13 (1.05-1.22) $P_{het}=2.0\times10^{-4}$, I^2=79.5%	1.1×10^{-3}
rs7800391	7q21	*IMMP2L*	T	GWAS	649/430	0.43/0.34	1.45 (1.21-1.73)	5.2×10^{-5}
				Italian1	1165/934	0.40/0.39	1.22 (1.07-1.38)	2.3×10^{-3}
				Italian2	320/679	0.41/0.40	1.06 (0.87-1.28)	0.55
				Polish	444/422	0.42/0.40	1.06 (0.88-1.29)	0.53
				Spanish	407/394	0.40/0.44	0.86 (0.71-1.05)	0.14
				Italian cohorts	**2134/2043**	**-**	**1.25 (1.14-1.38)**	**5.7×10^{-6}**
				Joint analysis	2985/2859	-	1.14 (1.05-1.23) $P_{het}=5.0\times10^{-3}$, I^2=73.1%	1.3×10^{-3}
rs1078150 0	9q34.3	*SNAPC4*	C	GWAS	649/431	0.69/0.60	1.51 (1.23-1.86)	7.9×10^{-5}
				Italian1	1155/916	0.67/0.63	1.18 (1.04-1.34)	0.01
				Italian2	322/706	0.66/0.62	1.20 (0.98-1.46)	0.07
				Polish	404/424	0.58/0.59	0.99 (0.81-1.21)	0.91
				Spanish	386/392	0.63/0.61	1.09 (0.89-1.34)	0.42
				Italian cohorts	**2126/2053**	**-**	**1.23 (1.12-1.36)**	**3.5×10^{-5}**
				Joint analysis	2916/2869	-	1.17 (1.08-1.27) P_{het}=0.18, I^2=36.6%	1.1×10^{-4}
rs7935113	11p15.3	*GALNTL 4*	C	GWAS	647/431	0.23/0.16	1.50 (1.20-1.88)	3.26×10^{-4}
				Italian1+2	1454/1572	0.20/0.16	1.28 (1.12-1.46)	2.20×10^{-4}
				Polish	452/443	0.13/0.13	0.99 (0.75-1.30)	0.93
				Spanish	352/407	0.18/0.18	1.02 (0.78-1.32)	0.90
				Italian cohorts	**2101/2003**	**-**	**1.36 (1.20-1.53)**	**7.41×10^{-7}**
				Joint analysis	2905/2853	-	1.24 (1.12-1.38) P_{het}=0.05 , I^2=61.96	2.71×10^{-5}
rs1220597	13q12.12	*SPATA13*	C	GWAS	646/414	0.48/0.39	1.42 (1.20-1.70)	7.11×10^{-5}
				Italian1+2	1458/1575	0.47/0.43	1.20 (1.08-1.33)	5.02×10^{-4}
				Polish	435/449	0.44/0.43	1.01 (0.84-1.22)	0.92
				Spanish	332/407	0.47/0.47	1.00 (0.82-1.23)	0.99
				Italian cohorts	**2104/1989**	**-**	**1.26 (1.14-1.38)**	**3.25×10^{-6}**
				Joint analysis	2871/2845	-	1.16 (1.07-1.25)	2.64×10^{-4}

							P_{het}=0.02, I^2=70.1%	
rs10136427	14q24.3	*BATF*	C	GWAS	646/414	0.88/0.81	1.62 (1.28-2.06)	5.73×10^{-5}
				Italian1+2	1515/1598	0.87/0.84	1.26 (1.09-1.45)	1.32×10^{-3}
				Polish	458/444	0.79/0.76	1.19 (0.95-1.48)	0.12
				Spanish	380/406	0.84/0.83	1.05 (0.81-1.37)	0.72
				Italian cohorts	2158/2012	-	1.40 (1.23-1.60)	4.35×10^{-7}
				Joint analysis	**2999/2862**	**-**	**1.30 (1.17-1.44)**	**9.30×10^{-7}**
							P_{het}=0.12, I^2=49.1%	
rs1203952	20p11	*FOXA2*	G	GWAS	647/429	0.28/0.21	1.49 (1.21-1.83)	1.37×10^{-4}
				Italian1+2	1453/1514	0.27/0.23	1.25 (1.11-1.41)	2.10×10^{-4}
				Polish	453/437	0.21/0.23	0.88 (0.70-1.10)	0.25
				Spanish	351/404	0.23/0.21	1.10 (0.86-1.40)	0.44
				Italian cohorts	**2100/1943**	**-**	**1.29 (1.16-1.44)**	**4.42×10^{-6}**
				Joint analysis	2904/2784	-	1.20 (1.09-1.31)	1.20×10^{-4}
							P_{het}=6.1×10^{-3} , I^2=75.80	
rs7267944	20q11.22-q12	*DHX35*	C	GWAS	646/414	0.26/0.18	1.54 (1.24-1.90)	6.60×10^{-5}
				Italian1+2	1477/1601	0.23/0.19	1.29 (1.14-1.45)	5.83×10^{-5}
				Polish	454/442	0.17/0.16	1.10 (0.85-1.41)	0.45
				Spanish	374/406	0.21/0.18	1.23 (0.96-1.57)	0.11
				Italian cohorts	2123/2015	-	1.39 (1.24-1.56)	2.13×10^{-8}
				Joint analysis	**2951/2863**	**-**	**1.32 (1.20-1.46)**	**1.34×10^{-8}**
							P_{het}=0.18, I^2=38.6%	

Tab. 2 – eQTL analysis based on Genotype-Tissue Expression (GTEx) program data. Significant results on thyroid tissues are highlighted in bold. Up=up-regulation, Down=down-regulation.

SNP	SNP *locus*	Risk allele	Gene	Gene *locus*	*p-value*	Risk allele effect on gene expression	Tissue
rs6759952	2q35	T	*DIRC3*	2q35	2.2×10^{-7}	Up	Testis
rs7617304	3q25.32	A	*GFM1*	3q25.32	2.5×10^{-5}	Down	Pancreas
			LXN	3q25.32	7.0×10^{-5}	Up	Cells - Transformed fibroblasts
			MLF1	3q25.32	8.7×10^{-6}	Up	Heart - Left Ventricle
			RP11-379F4.4	3q25.32	1.3×10^{-6}	Down	Adipose - Visceral (Omentum)
			RP11-379F4.4	3q25.32	2.4×10^{-5}	Down	Adrenal Gland
			RP11-379F4.4	3q25.32	3.5×10^{-6}	Down	Artery - Aorta
			RP11-379F4.4	3q25.32	4.9×10^{-5}	Down	Artery - Tibial
			RP11-379F4.4	3q25.32	6.2×10^{-5}	Down	Colon - Sigmoid
			RP11-379F4.4	3q25.32	7.0×10^{-6}	Down	Esophagus - Gastroesophageal Junction
			RP11-379F4.4	3q25.32	2.6×10^{-9}	Down	Lung
			RP11-379F4.4	3q25.32	4.7×10^{-6}	Down	Nerve - Tibial
			RP11-379F4.4	3q25.32	4.7×10^{-10}	Down	Skin - Not Sun Exposed (Suprapubic)
			RP11-379F4.4	3q25.32	9.5×10^{-10}	Down	Skin - Sun Exposed (Lower leg)
			RP11-379F4.4	3q25.32	2.6×10^{-12}	Down	Testis
			RP11-	**3q25.32**	**9.3×10^{-11}**	Down	**Thyroid**
			RP11-379F4.7	3q25.32	5.2×10^{-6}	Up	Nerve - Tibial
			RP11-379F4.8	3q25.32	3.1×10^{-7}	Down	Testis
rs7800391	7q21	T	*AC003088.1*	7q31.1	2.2×10^{-20}	Down	Cells - Transformed fibroblasts
rs10781500	9q34.3	C	*CARD9*	9q34.3	8.4×10^{-7}	Up	Cells - Transformed fibroblasts
			CARD9	9q34.3	1.0×10^{-18}	Down	Whole Blood
			GPSM1	9q34.3	3.0×10^{-5}	Up	Muscle - Skeletal
			GPSM1	**9q34.3**	**2.0×10^{-5}**	**Up**	**Thyroid**
			INPP5E	9q34.3	1.9×10^{-7}	Up	Brain - Caudate (basal ganglia)
			INPP5E	9q34.3	1.7×10^{-7}	Up	Brain - Cerebellum
			INPP5E	9q34.3	2.1×10^{-5}	Up	Esophagus - Mucosa
			INPP5E	9q34.3	2.3×10^{-5}	Up	Lung
			INPP5E	9q34.3	3.4×10^{-6}	Up	Skin - Sun Exposed (Lower leg)
			INPP5E	9q34.3	1.6×10^{-9}	Up	Testis
			INPP5E	**9q34.3**	**5.3×10^{-28}**	**Up**	**Thyroid**
			PMPCA	9q34.3	7.8×10^{-6}	Down	Liver
			PMPCA	**9q34.3**	**4.1×10^{-8}**	**Up**	**Thyroid**
			SDCCAG3	9q34.3	2.1×10^{-5}	Up	Artery - Tibial
			SDCCAG3	9q34.3	1.1×10^{-5}	Up	Cells - Transformed fibroblasts
			SDCCAG3	**9q34.3**	**1.8×10^{-5}**	**Up**	**Thyroid**
			SDCCAG3	9q34.3	2.0×10^{-10}	Up	Whole Blood
			SEC16A	9q34.3	1.4×10^{-5}	Up	Pancreas
			SEC16A	9q34.3	2.2×10^{-11}	Up	Whole Blood
			SNAPC4	9q34.3	3.9×10^{-6}	Down	Whole Blood
rs1203952	20p11	G	***FOXA2***	**20p11**	**1.6×10^{-18}**	**Down**	**Thyroid**
			LINC00261	**20p11**	**3.6×10^{-11}**	**Down**	**Thyroid**
			RP4-	**20p11**	**5.0×10^{-17}**	**Down**	**Thyroid**

5p15.33, 6q14.1, 10q26.12, 10q24.33 and 15q22.33 emerged as novel DTC risk factors [22, 23]. However, current evidence from available GWASs explains only a small proportion of the disease heritability. Several explanations for the missing heritability were proposed. These included rare variants in novel pathways that are undetectable through traditional GWAS study design, structural variants, such as CNVs, that are poorly captured by existing technologies, insufficient power to detect gene-gene interactions and environmental factors. Thus, further studies on large sample sets and based on novel experimental approaches, as array-based fine-mapping, next generation sequencing and gene- environment association studies, are warranted to identify the predisposing factors that could explain a greater percentage of DTC heritability.

5. Acknowledgements

The authors are grateful to the study participants and the study investigators and coordinators for the recruitment of samples. This study was supported by the Associazione Italiana Ricerca sul Cancro (AIRC, 2008) and by the Istituto Toscano Tumori (ITT, 2010).

REFERENCES

[1] Sipos JA, Mazzaferri EL, Thyroid cancer epidemiology and prognostic variables, *Clin Oncol (R Coll Radiol)*, 2010:22:395-404.

[2] Grande E, Diez JJ, Zafon C, Capdevila J, Thyroid cancer: molecular aspects and new therapeutic strategies, *J Thyroid Res*, 2012:2012:847108.

[3] Dal Maso L, Bosetti C, La Vecchia C, Franceschi S, Risk factors for thyroid cancer: an epidemiological review focused on nutritional factors, *Cancer Causes Control*, 2009:20:75-86.

[4] Fallah M, Pukkala E, Tryggvadottir L, Olsen JH, Tretli S, Sundquist K, Hemminki K, Risk of thyroid cancer in first-degree relatives of patients with non-medullary thyroid cancer by histology type and age at diagnosis: a joint study from five Nordic countries, *J Med Genet*, 2013:50:373-382.

[5] Landa I, Robledo M, Association studies in thyroid cancer susceptibility: are we on the right track?, *J Mol Endocrinol*, 2011:47:R43-58.

[6] Figlioli G, Elisei R, Romei C, Melaiu O, Cipollini M, Bambi F, Chen B, Kohler A, Cristaudo A, Hemminki K, Gemignani F, Forsti A, Landi S, A Comprehensive Meta-analysis of Case-Control Association Studies to Evaluate

Polymorphisms Associated with the Risk of Differentiated Thyroid Carcinoma, *Cancer Epidemiol Biomarkers Prev*, 2016:25:700-713.

[7] Kruglyak L, The road to genome-wide association studies, *Nat Rev Genet*, 2008:9:314-318.

[8] Hirschhorn JN, Daly MJ, Genome-wide association studies for common diseases and complex traits, *Nat Rev Genet*, 2005:6:95-108.

[9] Hosking FJ, Dobbins SE, Houlston RS, Genome-wide association studies for detecting cancer susceptibility, *Br Med Bull*, 2011:97:27-46.

[10] Cooper GM, Shendure J, Needles in stacks of needles: finding disease-causal variants in a wealth of genomic data, *Nat Rev Genet*, 2011:12:628-640.

[11] Consortium EP, The ENCODE (ENCyclopedia Of DNA Elements) Project, *Science*, 2004:306:636-640.

[12] Rockman MV, Kruglyak L, Genetics of global gene expression, *Nat Rev Genet*, 2006:7:862-872.

[13] Stranger BE, Forrest MS, Clark AG, Minichiello MJ, Deutsch S, Lyle R, Hunt S, Kahl B, Antonarakis SE, Tavare S, Deloukas P, Dermitzakis ET, Genome-wide associations of gene expression variation in humans, *PLoS Genet*, 2005:1:e78.

[14] Consortium GT, Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans, *Science*, 2015:348:648-660.

[15] Gudmundsson J, Sulem P, Gudbjartsson DF, Jonasson JG, Sigurdsson A, Bergthorsson JT, He H, Blondal T, Geller F, Jakobsdottir M, Magnusdottir DN, Matthiasdottir S, Stacey SN, Skarphedinsson OB, Helgadottir H, Li W, Nagy R, Aguillo E, Faure E, Prats E, Saez B, Martinez M, Eyjolfsson GI, Bjornsdottir US, Holm H, Kristjansson K, Frigge ML, Kristvinsson H, Gulcher JR, Jonsson T, Rafnar T, Hjartarsson H, Mayordomo JI, de la Chapelle A, Hrafnkelsson J, Thorsteinsdottir U, Kong A, Stefansson K, Common variants on 9q22.33 and 14q13.3 predispose to thyroid cancer in European populations, *Nat Genet*, 2009:41:460-464.

[16] Takahashi M, Saenko VA, Rogounovitch TI, Kawaguchi T, Drozd VM, Takigawa-Imamura H, Akulevich NM, Ratanajaraya C, Mitsutake N, Takamura N, Danilova LI, Lushchik ML, Demidchik YE, Heath S, Yamada R, Lathrop M, Matsuda F, Yamashita S, The FOXE1 locus is a major genetic determinant for radiation-related thyroid carcinoma in Chernobyl, *Hum Mol Genet*, 2010:19:2516-2523.

[17] Gudmundsson J, Sulem P, Gudbjartsson DF, Jonasson JG, Masson G, He H, Jonasdottir A, Sigurdsson A, Stacey SN, Johannsdottir H, Helgadottir HT, Li W, Nagy R, Ringel MD, Kloos RT, de Visser MC, Plantinga TS, den Heijer M, Aguillo E, Panadero A, Prats E, Garcia-Castano A, De Juan A, Rivera F, Walters GB, Bjarnason H, Tryggvadottir L, Eyjolfsson GI, Bjornsdottir US, Holm H, Olafsson I, Kristjansson K, Kristvinsson H, Magnusson OT, Thorleifsson G, Gulcher JR, Kong A, Kiemeney LA, Jonsson T, Hjartarson H, Mayordomo JI, Netea-Maier RT, de la Chapelle A, Hrafnkelsson J, Thorsteinsdottir U, Rafnar T, Stefansson K, Discovery of common variants associated with low TSH levels and thyroid cancer risk, *Nat Genet*, 2012:44:319-322.

[18] Kohler A, Chen B, Gemignani F, Elisei R, Romei C, Figlioli G, Cipollini M,

Cristaudo A, Bambi F, Hoffmann P, Herms S, Kalemba M, Kula D, Harris S, Broderick P, Houlston R, Pastor S, Marcos R, Velazquez A, Jarzab B, Hemminki K, Landi S, Forsti A, Genome-wide association study on differentiated thyroid cancer, *J Clin Endocrinol Metab*, 2013:98:E1674-1681.

[19] Figlioli G, Chen B, Elisei R, Romei C, Campo C, Cipollini M, Cristaudo A, Bambi F, Paolicchi E, Hoffmann P, Herms S, Kalemba M, Kula D, Pastor S, Marcos R, Velazquez A, Jarzab B, Landi S, Hemminki K, Gemignani F, Forsti A, Novel genetic variants in differentiated thyroid cancer and assessment of the cumulative risk, *Sci Rep*, 2015:5:8922.

[20] Figlioli G, Kohler A, Chen B, Elisei R, Romei C, Cipollini M, Cristaudo A, Bambi F, Paolicchi E, Hoffmann P, Herms S, Kalemba M, Kula D, Pastor S, Marcos R, Velazquez A, Jarzab B, Landi S, Hemminki K, Forsti A, Gemignani F, Novel genome-wide association study-based candidate loci for differentiated thyroid cancer risk, *J Clin Endocrinol Metab*, 2014:99:E2084-2092.

[21] Ward LD, Kellis M, HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants, *Nucleic Acids Res*, 2012:40:D930-934.

[22] Gudmundsson J, Thorleifsson G, Sigurdsson JK, Stefansdottir L, Jonasson JG, Gudjonsson SA, Gudbjartsson DF, Masson G, Johannsdottir H, Halldorsson GH, Stacey SN, Helgason H, Sulem P, Senter L, He H, Liyanarachchi S, Ringel MD, Aguillo E, Panadero A, Prats E, Garcia-Castano A, De Juan A, Rivera F, Xu L, Kiemeney LA, Eyjolfsson GI, Sigurdardottir O, Olafsson I, Kristvinsson H, Netea-Maier RT, Jonsson T, Mayordomo JI, Plantinga TS, Hjartarson H, Hrafnkelsson J, Sturgis EM, Thorsteinsdottir U, Rafnar T, de la Chapelle A, Stefansson K, A genome-wide association study yields five novel thyroid cancer risk loci, *Nat Commun*, 2017:8:14517.

[23] Mancikova V, Cruz R, Inglada-Perez L, Fernandez-Rozadilla C, Landa I, Cameselle-Teijeiro J, Celeiro C, Pastor S, Velazquez A, Marcos R, Andia V, Alvarez-Escola C, Meoro A, Schiavi F, Opocher G, Quintela I, Ansede-Bermejo J, Ruiz-Ponte C, Santisteban P, Robledo M, Carracedo A, Thyroid cancer GWAS identifies 10q26.12 and 6q14.1 as novel susceptibility loci and reveals genetic heterogeneity among populations, *Int J Cancer*, 2015:137:1870-1878.

Istituto Lombardo (Rend. Scienze) 150, 203-211 (2016)

COMPUTER QUANTISTICI: STATUS E PROSPETTIVE

LORENZO MACCONE (*)

Nota presentata dal m.e. Mauro D'Ariano
(Adunanza del 24 novembere 2016.)

SUNTO. – Il quantum computer è un sistema che è in grado di processare informazione quantistica. In questa presentazione spiegherò brevemente di cosa si tratta e quali sono i principali risultati della ricerca attuale in questo ambito.

***

ABSTRACT. – A quantum computer is a system that is able to process quantum information. In this presentation I'll briefly explain what is it and what are the main results of the current research in the field of quantum computation.

Strutturerò la mia presentazione come un dialogo tra una persona curiosa che pone domande e il sottoscritto che tenta di rispondere nel modo più chiaro possibile. Cercherò di comunicare i concetti di base piuttosto che entrare nel dettaglio tecnico dei risultati che presenterò.

– Cos'è un computer quantistico?

È un computer che usa fenomeni quantistici quali l'entanglement e la complementarità (che approfondirò di seguito) per svolgere calcoli.

– E cosa ci si guadagna?

Non si perde niente, nel senso che un computer quantistico può svolgere le stesse operazioni che può svolgere un computer convenzionale (che d'ora in poi chiamerò "computer classico"). Inoltre, si guada-

(*) Dipartimento di Fisica "A. Volta" & INFN Sezione di Pavia, Italia.
E-mail: maccone@unipv.it

gna in velocità per alcune specifiche computazioni (ma non per tutte). Quindi è sicuramente conveniente usare un computer quantistico, soprattutto se si vogliono svolgere questi tipi di computazioni.

– Come funziona?

Iniziamo a vedere come funziona un computer classico. Esso è un sistema che prende informazione, la elabora e restituisce un risultato. L'informazione in ingresso viene tradotta (tramite appropriate codifiche o alfabeti) in numeri ed è tipicamente codificata in bit. Infatti, invece di usare numeri in base 10 (quelli che usano gli umani), i computer usano numeri in base 2, cioè bit. Un bit può avere due soli valori, zero oppure uno. Si usano i bit nei computer elettronici perchè è semplice implementare un bit in elettronica: si suppone che si ha "uno" quando passa corrente e "zero" quando non passa corrente. Ad esempio il numero cinque si scrive come 101 in binario, che corrispone a un segnale elettronico dove ho corrente inizialmente, poi la corrente cessa, e alla fine la corrente passa nuovamente. L'elaborazione dei bit in ingresso viene svolta da circuiti elettronici che svolgono operazioni matematiche. Alla fine il computer restituisce il risultato.

I computer moderni sono talmente potenti che questi processi sono mascherati: invece di fornire numeri, possiamo fornire informazione tramite una tastiera, una telecamera, un microfono, internet, etc. Tutti questi segnali vengono tradotti in numeri e codificati in bit. Il computer elabora (matematicamente) questi numeri e restituisce il risultato che viene a sua volta tradotto come figure, filmati o lettere sullo schermo, suoni dall'altoparlante, etc.

Un computer quantistico funziona allo stesso modo: prende informazione, la elabora, e restituisce un risultato. La differenza con il computer classico è nel fatto che l'informazione è informazione quantistica, che non è facilmente traducibile in termini di numeri. Invece di usare bit, il computer quantistico usa quantum bit, chiamati "qubit". Sono sistemi *quantistici* a due livelli. L'elaborazione di informazione avviene tramite circuiti quantistici che sono in grado di elaborare l'informazione quantistica. Il risultato finale è, solitamente, qualcosa che un umano può utilizzare, e quindi è informazione classica, codificata in bit (ed, eventualmente, tradotta in immagini, filmati, lettere, suoni, etc.)

Quindi la filosofia dietro al computer quantistico è sostanzialmente la stessa che c'è dietro ad un computer convenzionale. Solo che

il computer classico elabora numeri, cioè bit, mentre il computer quantistico elabora informazione quantistica, cioè qubit.

– Cos'è un qubit?

È la generalizzazione quantistica del bit. Il bit può assumere due valori "zero" e "uno". Un qubit è un sistema quantistico a due livelli, a cui possiamo assegnare le etichette "zero" e "uno". Ad esempio, lo spin di un elettrone è un sistema quantistico a due livelli: lo spin può assumere valore $-1/2$ e $+1/2$ (a seconda se è orientato antiparallelamente o parallelamente al campo magnetico che si usa per definire la direzione di quantizzazione). Possiamo chiamare "zero" un elettrone con spin $-1/2$ e "uno" un elettrone con spin $+1/2$. Chiaramente, essendo questo un sistema quantistico, non è limitato ai valori "zero" e "uno", ma può accedere a fenomeni quantistici come l'entanglement e la complementarità.

– Complementarità?! Di cosa si tratta?

Ogni sistema quantistico può avere delle proprietà (complementari) che non si possono conoscere contemporaneamente con precisione. Per "proprietà" intendo qualunque cosa si possa misurare. Ad esempio, l'elettrone può avere spin con una proprietà di essere parallelo ("uno") o antiparallelo ("zero") al campo magnetico. Però può anche avere proprietà complementari.

Se "sovrappongo" più valori di una proprietà, ottengo un valore di una proprietà complementare. Questo è il principio di sovrapposizione della meccanica quantistica. Siccome non è intuitivamente chiaro a nessuno cosa vuol dire "sovrapporre" quantisticamente, preferisco parlare di complementarità, che è un concetto che si può comprendere.

Un modo per comprendere la complementarità è con una metafora. L'arancione si ottiene mescolando ("sovrapponendo") il giallo e il rosso. Quindi un'arancia non è nè gialla nè rossa. Si può dire che un'arancia sia contemporaneamente gialla e rossa. Quindi possiamo concludere che la proprietà di essere arancione ("arancionità") è complementare alla prorietà di essere giallo o rosso.

Purtroppo questa metafora è incompleta, perchè questi sono tutti valori di un'unica proprietà (il colore), mentre la complementarità quantistica riguarda proprietà diverse: sovrapponendo valori di una proprietà (ad esempio, sovrapponendo il giallo e il rosso) si dovrebbe ottenere il valore (arancione) di una proprietà diversa (qui la metafora fallisce perchè arancione è ancora un colore), e complementare alla prima.

La complementarità è alla base di una delle conseguenze più famose della teoria quantistica, la relazione di indeterminazione di Heisenberg. Questa dice che, meglio conosco una proprietà (ad esempio la posizione), meno conosco una proprietà complementare (ad esempio il momento). Matematicamente, questa relazione è espressa con $\Delta x \Delta p \geq \hbar$, dove Δx e Δp sono gli errori con cui determino la posizione e il momento rispettivamente, e dove $\hbar$ è una costante fisica. Questa disuguaglianza dice che non posso avere errori troppo piccoli per entrambi posizione e momento, perchè il prodotto degli errori deve essere maggiore di $\hbar$.

La nostra mente "classica" fatica a comprendere le implicazioni della complementarità e della indeterminazione: cosa vuole dire che la posizione della mia auto non è ben definita?! Dov'è la mia auto?! Sarebbe ben difficile guidare in autostrada se non potessi conoscere contemporaneamente sia la posizione che la velocità della mia auto! Nella vita di tutti i giorni non vediamo gli effetti dell'indeterminazione perchè la costante $\hbar$ ha un valore molto piccolo rispetto alle precisioni che possiamo raggiungere nella vita di tutti i giorni. La posizione della mia auto inizierebbe ad essere indefinita per qualche centimetro se io potessi determinare la sua velocità con una precisione di 32 cifre decimali. Una tale mostruosa precisione è irraggiungibile con la tecnologia odierna: le misure più precise che riusciamo a compiere attualmente non riescono a determinare nemmeno 20 cifre decimali. Questo è il motivo per cui riusciamo a guidare la nostra auto: possiamo conoscere sia la sua velocità che la sua posizione perchè non conosciamo nessuna delle due con precisione "quantistica". Il fatto che ciascuna delle due è determinata in modo "grossolano" (che è più che sufficiente per gli scopi ordinari), ci permette di definire entrambe.

– Perchè la complementarità è utile in un computer quantistico? Sembrerebbe più che altro una limitazione a quello che possiamo conoscere e definire!

Già, ma un quantum bit, un qubit, può accedere anche alle proprietà complementari a "zero" e "uno". Invece un bit di un computer classico è limitato a questa proprietà. Cioè un bit può solo avere valori {0, 1}. Invece un qubit può avere valori {0, 1}, ma anche {+, −}, {+i, −i}, etc., che sono tutte proprietà complementari. Ad esempio, la proprietà "+" si ottiene "sommando" (nel senso di sovrapposizione quantistica) le proprietà "zero" e "uno", e la proprietà "−" si ottiene "sot-

traendo" (di nuovo nel senso di sovrapposizione quantistica) le stesse proprietà.

Il fatto che queste siano proprietà complementari significa che posso accedere ad una sola di queste proprietà per volta. Ma se riesco ad essere sufficientemente furbo, posso usarle tutte nella mia computazione.

Un esempio di questo è l'algoritmo di Deutsch-Jozsa. Supponiamo di avere una scatola che accetta un bit come input e mi restituisce un valore a se il bit era zero o b se il bit era uno. Cioè la scatola calcola la funzione $f(0) = a$ e $f(1) = b$. Voglio scoprire se a è uguale a b oppure se sono diversi, *usando la scatola una volta sola*. (La scatola rappresenta una subroutine di una computazione.) Chiaramente questo è impossibile con una strategia classica: devo prima calcolare a e poi calcolare b per poterli confrontare, quindi devo usare la scatola due volte. Invece, con una strategia quantistica (che non descriverò qui), è possibile verificare se $a = b$ usando la scatola un'unica volta! Il trucco principale consiste nell'inserire nella scatola un qubit "+", che è complementare a "zero" o "uno". In un certo senso sto inserendo nella scatola "zero" e "uno" contemporaneamente e quindi riesco a scoprire se $a = b$ con un uso solo della scatola.

[Ho presentato una versione semplificata dell'algoritmo di Deutsch-Jozsa. La versione più generale è in grado di determinare se una funzione di n bit è costante oppure bilanciata, con una diminuzione esponenziale in n di calcoli della funzione rispetto a qualunque algoritmo classico deterministico. Se invece uno fa il confronto con gli algoritmi classici probabilistici, allora la diminuzione è solo di un fattore due come l'esempio semplificato che ho indicato sopra.]

– Questo esempio è interessante, ma poco utile in pratica. Ci sono esempi un po' più utili?

In realtà è molto difficile trovare algoritmi quantistici che hanno vantaggi rispetto ai corrispondenti algoritmi classici. Ma sono noti alcuni esempi importanti. I due più famosi sono i seguenti:

1. Fattorizzazione. Dato un numero c, voglio trovare i fattori a e b tali che $a \times b = c$. Ad esempio, se $c = 15$, devo trovare $a = 5$ e $b = 3$. Per numeri piccoli come 15, questo è semplice. Per numeri grandi (cioè numeri con centinaia di cifre) questo è estremamente complicato. È complicato a tal punto che tutti i principali protocolli crittografici in uso oggi su internet si basano sull'assunzione che fattorizzare numeri grandi sia talmente complicato che non sia possibile risolvere questo

problema in un tempo ragionevole. Questi protocolli sono quelli che usate ogni volta che fate un acquisto con la carta di credito oppure accedete al sito web della vostra banca: usate protocolli crittografici perchè non volete che nessuno possa scoprire il vostro numero di carta o possa accedere al vostro conto online.

L'assunzione che fattorizzare numeri grandi sia troppo complicato per un computer è falsa se usate un quantum computer! Questa fu la scoperta, ad opera di Peter Shor, che lanciò la quantum computation. Peter Shor trovò un algoritmo quantistico che può fattorizzare numeri di n cifre con un'efficienza esponenziale in n, cioè in modo molto più semplice che gli algoritmi classici.

In previsione dello sviluppo di quantum computers che siano in grado di eseguire l'algoritmo di Shor per numeri molto grandi, sono stati sviluppati nuovi protocolli crittografici "post-quantum" che sono robusti anche se attaccati da un quantum computer. Per ora stiamo ancora usando i vecchi protocolli perchè non esistono ancora quantum computer sufficientemente potenti da eseguire l'algoritmo di Shor per numeri sufficientemente grandi da creare un rischio di sicurezza per gli attuali protocolli crittografici.

2. Grover search: ricerca in un database non strutturato. Supponiamo di voler scoprire a chi è intestato un numero di telefono cercando il nome sull'elenco telefonico. Siccome l'elenco telefonico è ordinato per ordine alfabetico, dovrò cercare il numero in ogni pagina finchè non trovo il nome associato. (Al contrario di cercare il numero associato ad un nome che è molto semplice.) L'algoritmo di Grover permette ad un quantum computer di effettuare questa ricerca in modo più efficiente di un algoritmo classico.

Questo algoritmo ha anche un interesse concettuale, perchè ci dice che un computer quantistico è in grado di velocizzare tutti gli algoritmi NP, che sono gli algoritmi dove è difficile trovare una soluzione ad un problema, ma è facile verificare se ho una soluzione. L'algoritmo di Grover ci dice che un algoritmo quantistico può velocizzare gli algoritmi NP che non hanno struttura (come l'esempio dell'elenco telefonico) e dove devo ridurmi a testare tutte le possibili soluzioni una per una.

– Allora i quantum computer sono utili! Ma esistono?

Non ancora! O meglio, abbiamo quantum computer molto rudimentali che sono in grado di fare semplici operazioni su poche decine di qubit al massimo. Siamo ancora ad uno stadio della ricerca dove non

è neanche chiaro quale sia la tecnologia più promettente per implementare un quantum computer. Si stanno esplorando diverse possibili implementazioni: ad esempio, sistemi a stato solido (simili a quelli usati nei computer elettronici classici), ma anche molte altre possibilità, quali la risonanza magnetica nucleare, le trappole magneto ottiche, l'ottica quantistica, etc.

Se avete 10 milioni di dollari da buttare, potete acquistare un "quantum computer" dalla D-Wave, una ditta canadese che mette in commercio un computer basato sulla tecnologia del quantum simulated annealing. È stato acquistato da NASA, Google, Lockeed (che sono organizzazioni che effettivamente hanno 10 milioni di dollari da buttare). Purtroppo ad oggi non ha ancora dimostrato di essere un "vero" quantum computer. Cioè non ha dimostrato di saper risolvere dei problemi in modo più efficiente degli analoghi algoritmi classici. Ha solo saputo dimostrare di fare simulated annealing più velocemente del simulated annealing classico, ma questo non vuole necessariamente dire che può risolvere problemi in modo più veloce di *altri* algoritmi classici. Attualmente c'è un forte dibattito molto divertente tra i costruttori della D-Wave e i principali scienziati che lavorano in quantum computation.

– In Europa?

L'Europa ha riconosciuto che la tecnologia quantistica, di cui il quantum computer è un aspetto, sarà fondamentale per lo sviluppo tecnologico del futuro. Infatti, sta partendo un'iniziativa di finanziamento molto rilevante nei confronti della tecnologia quantistica, la "quantum technologies flagship" che finanzierà un miliardo di euro nei prossimi anni per sviluppare queste tecnologie, incluso il quantum computer.

In Italia ci sono molti gruppi di ricerca che si occupano di questi temi. Oltre al nostro gruppo di Pavia, anche qui a Milano ci sono dei gruppi importanti (ad esempio il gruppo di Matteo Paris all'università statale) e molti altri in tutto il resto d'Italia.

– Quando avremo un quantum computer, allora?

Non lo so. È molto difficile prevedere gli sviluppi tecnologici senza rendersi completamente ridicoli. Negli anni 40 il direttore dell'IBM aveva previsto "I think there is a world market for maybe five computers." Se uno conta solo gli i-phone, ne sono stati prodotti, ad oggi, più di un miliardo.

– Tu di cosa ti sei occupato in questo ambito?

La mia ricerca non è direttamente collegata alla quantum computation, ma nel corso degli anni mi sono occupato anche di questo. In particolare menziono tre risultati a cui ho contribuito:

1. Quantum random access memory: la memoria quantistica ad accesso casuale. La RAM è parte integrante dell'architettura dei computer elettronici. Consiste in una serie di celle di memoria a cui si può accedere in modo indipendente e in maniera casuale. Io e i miei collaboratori avevamo sviluppato e brevettato una architettura per la versione quantistica della RAM.

2. Quantum private queries. È un protocollo crittografico per interrogare un database (ad esempio Google) senza che il proprietario del database possa scoprire quale domanda è stata posta. È anche uscito un articolo su 'Le Scienze' su questo protocollo, che abbiamo provato a vendere a Google. Non si sono dimostrati interessati all'acquisto perchè loro vogliono sapere cosa noi chiediamo a Google: il loro business model lo richiede.

3. Blind quantum computation. È un altro algoritmo crittografico che permette ad Alice di eseguire una computazione sul quantum computer di Bob senza che lui possa sapere che tipo di computazione lei sta eseguendo, nè il risultato ottenuto. Questo sarebbe impossibile usando un computer convenzionale.

– Perchè un fisico dovrebbe interessarsi al quantum computer? Sembra più una sfida tecnologica che concettuale.

No! Analizzare la fisica dal punto di vista dell'elaborazione dell'informazione quantistica si è rivelata essere una rivoluzione concettuale di grossa portata nella fisica moderna. Ad esempio, il prof. D'Ariano nel nostro gruppo di Pavia sta portando avanti un programma molto interessante di riformalizzazione di tutta la fisica a partire da principi fisici operativi e informazionali. Inoltre, le ultime idee sulla gravità quantistica esplorano quali tipi di operazioni quantistiche si possono effettuare con un buco nero in termini di elaborazione di informazione.

Quindi, ragionare in termini di elaborazione di informazione porta a nuovi risultati perfino nella fisica fondamentale, quanto di più lontano ci sia dalle applicazioni tecnologiche!

– Interessante, mi puoi sintetizzare cosa hai detto?

Ho parlato del quantum computer. Ho illustrato come esso sia

una generalizzazione dei computer convenzionali, dove i bit vengono sostituiti da qubit. Ho fatto vedere che in questo modo si ottiene un guadagno grazie agli effetti quantistici, quali la complementarità, di cui ho presentato una metafora e di cui ho illustrato la connessione con l'indeterminazione di Heisenberg. Ad esempio la complementarità è importante nell'algoritmo di Deutsch-Jozsa che sfrutta il fatto che posso calcolare una funzione con un input che è "contemporaneamente" sia "zero" che "uno". Infine ho concluso con un brevissimo excursus sullo stato attuale della ricerca in questo campo. Il messaggio che vorrei che il lettore conservasse di questa presentazione è che la tecnologia quantistica non è solo il futuro della tecnologia, ma riguarda anche una rivoluzione concettuale: un modo diverso di analizzare la natura, analizzando i sistemi fisici in base alla loro capacità di elaborare informazione.

Scienze naturali
Biologia

Istituto Lombardo (Rend. Scienze) 150, 213-219 (2016)

STUDY ON THE GENETIC PREDISPOSITION TO BREAST CANCER AND THE DNA DAMAGE RESPONSE AT TELOMERES AND IN CELLULAR SENESCENCE

FRANCESCA ROSSIELLO (*), *et al.*[#]

Nota presentata dal s.c. Andrea Mattevi
(Adunanza del 24 novembre 2016)

SUNTO. – Le mutazioni germinali nella sequenza codificante di alcuni geni, in particolare BRCA1 e BRCA2, conferiscono un elevato rischio di insorgenza del tumore alla mammella. Tuttavia circa il 70% dei casi non sono associabili a nessuna mutazione conosciuta. Mediante uno screening per sequenziamento a partire da sangue periferico di pazienti e controlli sani, abbiamo identificato alcune mutazioni germinali non descritte in precedenza nella regione non codificante 3' UTR del gene BRCA1. Queste mutazioni possono causare un'alterazione dell'espressione genica, e possono quindi essere utili per predire, con una maggiore accuratezza, la predisposizione al tumore familiare alla mammella. Successivamente, per comprendere i meccanismi di invecchiamento cellulare in cellule non proliferanti abbiamo studiato il segnale di DNA damage response (DDR) persistente nelle cellule senescenti, sia in colture cellulari che in tessuti animali, e abbiamo scoperto che questi segnali localizzano principalmente ai telomeri, indipendentemente dalla loro lunghezza. L'accumulo di danno al DNA telomerico, dovuto all'inibizione dei meccanismi di riparo, se da una parte causa l'invecchiamento

(*) IFOM-The FIRC Institute of Molecular Oncology, Milan, Italia.
E-mail: francesca.rossiello@ifom.eu

(#) La nota è il risultato di un lavoro di ricerca condotto in collaborazione con: B.L. Brewster, J.D. French, S.L. Edwards, M. Wong, A. Wronski, P. Whiley, N. Waddell, X. Chen, B. Bove, KCONFAB, J.L. Hopper, E.M. John, I. Andrulis, M. Daly, S. Volorio, L. Bernard, B. Peissel, S. Manoukian, M. Barile, S. Pizzamiglio, P. Verderio, A.B. Spurdle, P. Radice, A.K. Godwin, M.C. Southey, M.A. Brown, P. Peterlongo, M. Fumagalli, M. Clerici, S. Barozzi, D. Cittaro, J.M. Kaplunov, G. Bucci, M. Dobreva, V. Matti, C.M. Beausejour, U. Herbig, M.P. Longhese, C. Mondello, F. d'Adda di Fagagna.

cellulare e dell'individuo, è un importante meccanismo anti tumorale, in quanto previene l'iperproliferazione delle cellule nelle fasi iniziali di trasformazione neoplastica.

***

ABSTRACT. – Germ line mutations in the coding sequence of some genes, mainly BRCA1 and BRCA2, confer a high risk of developing breast cancer. However, about 70% of tumors cases are not associated with any known mutation. By screening by sequencing peripheral blood of patients and healthy controls, we identified some previously unknown germline mutations in the 3 'UTR non-coding region of the BRCA1 gene. These mutations can modify gene expression, and can therefore be useful to predict, with greater accuracy, familiar breast cancer predisposition. Subsequently, to understand the mechanisms of aging in non-proliferating cells we studied the persistent DNA damage response (DDR) signal in senescent cells, both in cell cultures and in animal tissues, and we found that these signals mainly localize to telomeres, independently from their length. The accumulation of damage at the telomeric DNA, due to the inhibition of the repair mechanisms, causes cellular and individual aging, but it is also an important anti-tumor mechanism, as it prevents the uncontrolled cell growth during the early stages of neoplastic transformation.

During my master thesis internship, I participated to a project focused on understanding better the genetic predisposition to familiar breast cancer [Ch. 1].

Germline mutations with higher penetration, which confer a high risk of developing breast cancer, are located in the BRCA1 and BRCA2 coding region. Other mutations in other genes that determine a moderate risk of cancer are also known. However, about 70% of breast cancers are not associated with any known mutations.

We performed a screening by sequencing to identify germline mutations in the non-coding region 3' UTR of the BRCA1 gene, starting from peripheral blood of female breast cancer patients who had no mutations in the coding region of BRCA1 and 2 genes. Healthy donors were used for the control group. As a result of this screening, we have identified new variants not previously described, which were detected only in patients. Mutations in the 3' UTR regions can cause gene expression alteration by forming new miRNA interaction sites. By in silico studies, we have shown that some of these variants produce new sites for some miRNAs. BRCA1 acts as an oncosuppressor, so a reduction in its expression can cause a greater cancer incidence. To experimentally validate the hypothesis that these novel identified variants could determine an altered expression of the BRCA1 gene, we performed luciferase expression experiments, in which luciferase gene was cloned upstream of the 3' UTR

region, containing the mutation under investigation. We have therefore verified that some of these variants modify the expression of luciferase, if co-expressed with the related miRNA.

These results indicate that a more in-depth study of BRCA1 gene mutations in the 3'UTR region may be useful in predicting with greater accuracy the predisposition to familial breast cancer.

During my PhD internship, I moved to the field of DNA damage and telomeres [Ch. 2].

Replicative senescence, caused by progressive telomere shortening, has been deeply studied in the past fifty years. However it was not yet clear which molecular events regulated the aging of non-proliferating cells.

Cells have several DNA repair mechanisms that ensure the maintenance of the integrity of genetic information. In this study, we found that when telomeric DNA is damaged, it cannot be fixed because the telomeric protein TRF2 inhibits cellular repair mechanisms. This phenomenon causes a persistent DNA damage response (DDR) signal, which activates the senescence program, regardless of the length of the telomeres. To confirm these *in vitro* results, we observed an accumulation of DDR markers in aged baboons, in tissues made of non-proliferating cells, such as liver and hippocampus. This discovery has an important implication in the oncology field. In fact, a study in collaboration with some members of our group (Suram *et al.*, Embo Journal, 2012) showed that hyper-proliferation associated with the expression of an oncogene in the early stages of tumorigenesis causes replicative stress, leading to accumulation of telomeric damage, persistent DDR and activation and cellular senescence. Only when tumor cells accumulate further mutations, which cause the DDR to be deactivated at telomeres, tumor progression may take place.

These data suggest that the inhibition of repair mechanism at telomeres, despite causing cellular and organismal aging, is an essential anti-tumor mechanism.

Finally I was involved in another project, aimed to investigate more deeply the mechanisms regulating cellular senescence [Ch. 3].

Cellular senescence is a permanent cell cycle arrest characterized by persistent DDR activation. However, some reports suggested that DDR activation is a feature only of early cellular senescence that is then lost with time. This challenged the hypothesis of a causal relationship between DDR activation and cellular senescence.

We examined whether the DDR of telomeres in the senescent

cells was really persistent or whether it was activated at first and then turned off.

To better understand the differences in the experimental conditions that led to opposite conclusions, we analysed replicative senescence and ionizing radiation-induced senescence in different cell lines. We observed how the loss of DDR markers, which leaded to a temporary activation of DDR, correlated with the inability of the cell line to survive in culture for weeks or months. The death of the most damaged cells therefore caused a bias in the interpretation of the results. By analysing instead a cell line that was very resistant to the stress caused by *in vitro* culture, we have observed how DDR markers are constantly activated for weeks or months – the entire duration of the experiments. In support of these observations, human skin fibroblasts from ultra-centenarians donors, and cultured up to three years *in vitro*, still showed the activation of the DDR markers, indicating that also *in vivo* activated senescence is constantly supported by persistent DDR.

We concluded that different types of senescence (replicative, ionizing radiation-induced and *in vivo* activated) are associated with a constant DDR activation that acts as a trigger for the maintenance of cellular senescence program.

1. Identification of fifteen novel germline variants in the BRCA1 3'UTR reveals a variant in a breast cancer case that introduces a functional miR-103 target site[1]

Introduction

Germline mutations with higher penetration, which confer a high risk of developing breast cancer, are located in the BRCA1 and BRCA2 coding region. Other mutations in other genes that determine a moderate risk of cancer are also known. However, about 70% of breast cancers are not associated with any known mutations.

---

[1] Brewster BL*, Rossiello F*, French JD, Edwards SL, Wong M, Wronski A, Whiley P, Waddell N, Chen X, Bove B; kConFab, Hopper JL, John EM, Andrulis I, Daly M, Volorio S, Bernard L, Peissel B, Manoukian S, Barile M, Pizzamiglio S, Verderio P, Spurdle AB, Radice P, Godwin AK, Southey MC, Brown MA, Peterlongo P. Human Mutation, 2012. *Equal contribution.

Results

We performed a screening by sequencing to identify germline mutations in the non-coding region 3' UTR of the BRCA1 gene, starting from peripheral blood of female breast cancer patients who had no mutations in the coding region of BRCA1 and 2 genes. Healthy donors were used for the control group. As a result of this screening, we have identified new variants not previously described, which were detected only in patients. Mutations in the 3' UTR regions can cause gene expression alteration by forming new miRNA interaction sites. By in silico studies, we have shown that some of these variants produce new sites for some miRNAs. BRCA1 acts as an oncosuppressor, so a reduction in its expression can cause a greater cancer incidence. To experimentally validate the hypothesis that these novel identified variants could determine an altered expression of the BRCA1 gene, we performed luciferase expression experiments, in which luciferase gene was cloned upstream of the 3' UTR region, containing the mutation under investigation. We have therefore verified that some of these variants modify the expression of luciferase, if co-expressed with the related miRNA.

Conclusions

These results indicate that a more in-depth study of BRCA1 gene mutations in the 3'UTR region may be useful in predicting with greater accuracy the predisposition to familial breast cancer.

2. Telomeric DNA damage is irreparable and causes persistent DNA-damage-response activation[2]

Introduction

Replicative senescence, caused by progressive telomere shortening, has been deeply studied in the past fifty years. However it was not

[2] Fumagalli M*, Rossiello F*, Clerici M, Barozzi S, Cittaro D, Kaplunov JM, Bucci G, Dobreva M, Matti V, Beausejour CM, Herbig U, Longhese MP, d'Adda di Fagagna F. Nature Cell Biology, 2012. *Equal contribution.

yet clear which molecular events regulated the aging of non-proliferating cells.

Results

Cells have several DNA repair mechanisms that ensure the maintenance of the integrity of genetic information. In this study, we found that when telomeric DNA is damaged, it cannot be fixed because the telomeric protein TRF2 inhibits cellular repair mechanisms. This phenomenon causes a persistent DNA damage response (DDR) signal, which activates the senescence program, regardless of the length of the telomeres. To confirm these *in vitro* results, we observed an accumulation of DDR markers in aged baboons, in tissues made of non-proliferating cells, such as liver and hippocampus. This discovery has an important implication in the oncology field. In fact, a study in collaboration with some members of our group (Suram *et al.*, Embo Journal, 2012) showed that hyper-proliferation associated with the expression of an oncogene in the early stages of tumorigenesis causes replicative stress, leading to accumulation of telomeric damage, persistent DDR and activation and cellular senescence. Only when tumor cells accumulate further mutations, which cause the DDR to be deactivated at telomeres, tumor progression may take place.

Conclusions

These data suggest that the inhibition of repair mechanism at telomeres, despite causing cellular and organismal aging, is an essential anti-tumor mechanism.

3. Stable cellular senescence is associated with persistent DDR activation[3]

Introduction

Cellular senescence is a permanent cell cycle arrest characterized by persistent DDR activation. However, some reports suggested that

[3] Fumagalli M, Rossiello F, Mondello C, d'Adda di Fagagna F. PLoS One, 2014.

DDR activation is a feature only of early cellular senescence that is then lost with time. This challenged the hypothesis of a causal relationship between DDR activation and cellular senescence.

Results

We examined whether the DDR of telomeres in the senescent cells was really persistent or whether it was activated at first and then turned off.

To better understand the differences in the experimental conditions that led to opposite conclusions, we analysed replicative senescence and ionizing radiation-induced senescence in different cell lines. We observed how the loss of DDR markers, which leaded to a temporary activation of DDR, correlated with the inability of the cell line to survive in culture for weeks or months. The death of the most damaged cells therefore caused a bias in the interpretation of the results. By analysing instead a cell line that was very resistant to the stress caused by *in vitro* culture, we have observed how DDR markers are constantly activated for weeks or months – the entire duration of the experiments. In support of these observations, human skin fibroblasts from ultracentenarians donors, and cultured up to three years *in vitro*, still showed the activation of the DDR markers, indicating that also *in vivo* activated senescence is constantly supported by persistent DDR.

Conclusions

We concluded that different types of senescence (replicative, ionizing radiation-induced and *in vivo* activated) are associated with a constant DDR activation that acts as a trigger for the maintenance of cellular senescence program.

Istituto Lombardo (Rend. Scienze) 150, 221-237 (2016)

SIMMETRIA PARITÀ-TEMPO IN OTTICA: VERSO UNA NUOVA INGEGNERIA DELLA LUCE

STEFANO LONGHI (*)

Nota presentata dal m.e. Orazio Svelto
(Adunanza del 22 dicembre 2016)

SUNTO. – L'introduzione in ottica dei concetti di simmetria parità-tempo, ben noti nelle teorie di meccanica quantistica e di teoria quantistica dei campi, sta rivoluzionando la nostra capacità di ideare e realizzare materiali ottici sintetici e dispositivi fotonici per il controllo della luce su micro- e nano-scala, con nuove funzionalità inimmaginabili con materiali ordinari e con importanti applicazioni nell' ambito delle tecnologie laser, della fotonica integrata e della sensoristica. In questa Nota verranno illustrati in maniera semplificata i principi teorici dell'ottica in materiali con simmetria parità-tempo e presentate le principali applicazioni alle tecnologie fotoniche.

***

ABSTRACT. – The introduction of the concept of parity-time symmetry in optics, inspired by recent theoretical developments in quantum mechanics and quantum-field theories, is revolutionizing our ability to design and manufacturing synthetic optical materials and photonic devices for molding the flow of light at the micro- and nano-scale, with novel functionalities impossible with ordinary materials and with important applications in the fields of laser technologies, integrated photonics and sensor technology. In this note I will illustrate the basic theoretical grounds of optical materials with parity-time symmetry and I will present the main recent applications to photonic technologies.

1. INTRODUZIONE E MOTIVAZIONI

Il concetto di simmetria è fondamentale in molti ambiti della scienza, ed ha attratto grande interesse fin dall'antichità da parte di

(*) Dipartimento di Fisica, Politecnico di Milano, Milano, Italia.
E-mail: stefano.longhi@polimi.it

diversi pensatori, filosofi, artisti e scienziati [1]. Nella fisica moderna, i concetti di simmetria e di rottura spontanea di simmetria giocano un ruolo fondamentale e sono basilari per una comprensione moderna delle leggi fisiche. Basti pensare che le più grandi scoperte della fisica teorica del secolo XX, a partire dalla teoria della relatività ristretta fino alle teorie della superfluidità e della superconduttività nella fisica della materia ed alla teoria unificata delle interazioni elementari [2], possono essere interpretate in un certo senso come scoperte di invarianze, di simmetrie, e di rotture spontanee di simmetrie.

Alcune importanti simmetrie in Fisica sono legate a combinazioni di operatori quali la parità P, che possiamo identificare in maniera semplice con la simmetria spaziale di inversione, la coniugazione di carica C, che scambia ogni particella con la sua antiparticella (che ha carica opposta), e la trasformazione di inversione temporale T, che inverte la direzione del tempo (sostanzialmente come un film visto all'indietro). Attualmente si ritiene che le leggi fisiche debbano essere invarianti per simmetria TCP, e cioè che per ottenere invarianza delle leggi fisiche occorre scambiare, nelle immagini speculari degli eventi fisici, le particelle con le loro antiparticelle e guardare gli eventi nel loro fluire all'indietro.

Le simmetrie di parità P e di tempo T giocano un ruolo essenziale anche nell'ambito della meccanica quantistica non relativistica, dove la simmetria di coniugazione di carica non interviene. La teoria della meccanica quantistica ha ormai un secolo di vita e, in forza delle numerosissime verifiche sperimentali di tale teoria e della sua capacità di descrivere le proprietà della materia così come la osserviamo su scala microscopica (atomica), è comunemente accettata dalla scienza moderna. Uno degli assiomi centrali della meccanica quantistica è quello secondo cui ogni sistema fisico è descritto da un operatore autoaggiunto (o Hermitiano) H, il cosiddetto operatore Hamiltoniano del sistema, che ne compendia le sue proprietà intime e che consente di prevederne la sua evoluzione nel tempo. La proprietà menzionata, che in matematica implica l'uguaglianza dell'operatore Hamiltoniano con il suo aggiunto , resta da un punto di vista fisico piuttosto oscura, ma è considerata necessaria ed ineludibile per garantire alcuni elementari vincoli che il sistema deve soddisfare, come ad esempio il fatto che la particella non scompaia (conservazione della norma) e che le sue possibili energie siano quantità misurabili e quindi a valori reali.

Recentemente l'assioma che un sistema quantistico possa essere

descritto da un operatore non Hermitiano è stata dibattuta in ambito teorico da alcuni autori, soprattutto in alcuni lavori di Carl Bender [3-5]. In tali lavori, la condizione di Hermiticità viene sostituita dalla condizione (più debole) di simmetria parità-tempo (*PT*) per l'operatore Hamiltoniano, e cioè che il commutatore si annulli. Poichè l'operatore di inversione temporale *T* non è un operatore lineare, la proprietà non implica necessariamente che gli operatori Hamiltoniano e *PT* abbiano in comune uno stesso set di autofunzioni e autovalori. Si dice cha la simmetria *PT* non è rotta se le energie sono reali esattamente come per un operatore Hermitiano; solo quando c'è rottura della simmetria parità-tempo le energie diventano 'complesse'. La teoria avanzata da Bender è stata a lungo dibattuta e resta controversa, soprattutto perché in molti casi essa rappresenta una formulazione diversa ma equivalente alla formulazione ortodossa della meccanica quantistica mentre in altri casi pone difficoltà concettuali inerenti la violazione del principio 'no-signaling' di relatività (si veda, per esempio, [6]). Difatti, assumendo la simmetria *PT* (anziché l'Hermiticità dell'operatore Hamiltoniano) come fondamentale in una teoria quantistica, ci sarebbero radicali e nuove conseguenze, come la possibilità per un sistema quantistico di evolvere più rapidamente di quello previsto normalmente. In particolare, si potrebbe immaginare un esperimento in cui l'informazione viaggi più veloce della luce [6], il che è considerato impossibile dalla comunità scientifica in virtù delle limitazioni imposte dalla teoria della relatività ristretta di Einstein.

Come spesso accade in Fisica, fenomeni e sistemi molto diversi fra loro possono essere descritti da relazioni ed equazioni simili, e quindi possono mostrare dinamiche e fenomenologie simili. Analogie fra la meccanica quantistica (ondulatoria) e l'ottica sono ben note fin dai primordi delle teorie quantistiche della materia. Ad esempio, il principio di Fermat dell'ottica classica ed il principio della minima azione della meccanica classica esprimono sostanzialmente il medesimo principio per raggi luminosi e particelle classiche, e tale analogia fu usata per formulare una equazione ondulatoria per particelle, che ha dato origine alla cosiddetta meccanica ondulatoria (la forma primordiale della meccanica quantistica non relativistica). Un compendio recente di analogie fra ottica e meccanica quantistica, con applicazioni all'ottica ed alla fotonica integrata, può trovarsi nel lavoro [7]. La meccanica quantistica *PT* simmetrica trova una corrispondenza semplice ed elegante in ottica quando si considera la propagazione della luce in

strutture dielettriche in cui l'indice di rifrazione del mezzo è immaginario, e più precisamente in cui si ha amplificazione e dissipazione di luce in maniera spazialmente bilanciata. Mentre teorie *PT* simmetriche in meccanica quantistica e teoria dei campi hanno mostrato serie criticità [6] e sembrano destinate a non avere sviluppi di particolare interesse [8], l'introduzione in ottica dei medesimi concetti di simmetria parità-tempo sta rivoluzionando, in soli pochi anni, la nostra capacità di ideare e realizzare materiali ottici sintetici e dispositivi fotonici. Materiali ottici con distribuzioni bilanciate di guadagno e perdite possono essere progettati in modo da controllare la propagazione della luce sulla microscala in maniera nuova rispetto ai metodi convenzionali, quali le guide e le fibre ottiche ordinarie, e sono di interesse in applicazioni alle tecnologie laser, alla fotonica integrata, alla sensoristica, alle tecniche di imaging. Nuovi sviluppi ed applicazioni dei concetti di simmetria parità-tempo all'ottica ed alle tecnologie laser si possono trovare in alcuni recenti lavori di letteratura, cui si rimanda il lettore per approfondimenti più specialistici [9-17]. In questa Nota verranno discusse brevemente ed in maniera semplificata alcune importanti applicazioni dei concetti di simmetria parità-tempo in ottica: la possibilità di realizzare un laser-antilaser, cioè un dispositivo capace simultaneamente di emettere luce coerente (come in un convenzionale oscillatore laser) e di assorbire completamente radiazione coerente incidente su di esso; la possibilità di costruire un metamateriale che mostra una invisibilità unidirezionale, che risulta cioè invisibile (come se esso non ci fosse) se osservato in una direzione dello spazio, ma non nella direzione opposta; infine la possibilità di generare luce laser strutturata, che trasporta momento angolare orbitale ('optical vortices'), in micro-laser ad anelli che operano ad un punto eccezionale

2. Simmetria parità-tempo in ottica

L'idea di applicare a strutture ottiche dielettriche proprietà di simmetria partà-tempo è piuttosto recente e risale ad alcuni lavori di pochi anni fa [18-21]. In ottica, una struttura con simmetria parità-tempo è sostanzialmente costituita da un materiale dielettrico in cui l'indice di rifrazione n è complesso. Per ottenere un sistema ottico con simmetria *PT*, bisogna guardare all'indice di rifrazione da un punto di vista più generale e matematicamente più sottile di quanto non avvenga

solitamente. Nel caso più generale, l'indice di rifrazione è infatti espresso da una quantità complessa, la cui parte reale è correlata al rallentamento della luce nel materiale e all'angolo con cui essa viene piegata, mentre la parte immaginaria determina l'amplificazione (guadagno) o l'assorbimento (perdita) di luce all'interno di un materiale. Un mezzo dielettrico con simmetria *PT* ha sostanzialmente un profilo di indice di rifrazione reale che è una funzione pari per inversione, mentre la parte immaginaria dell'indice è una funzione dispari per inversione. In altri termini, il dielettrico può assorbire ed amplificare radiazione ottica in regioni diverse dello spazio ma in maniera bilanciata. In termini matematici, possiamo scrivere

$$n(\boldsymbol{r}) = n^*(-\boldsymbol{r}) \tag{1}$$

dove $\boldsymbol{r}$ è l'insieme delle coordinate spaziali. La propagazione della luce in tali strutture è descrivile mediante un operatore Hamiltoniano H con simmetria PT , così come avviene nel problema quantistico. Ad esempio, nel caso più semplice in cui il mezzo dielettrico ha una indice di rifrazione $n(\boldsymbol{r})$ invariate in un a direzione dello spazio z e che si discosta poco da un valore di riferimento (di substrato o bulk) n_s, la propagazione dell'onda luminosa nella direzione z è descritta da una equazione parassiale simile all'equazione di Schoedinger per una particella quantistica non relativistica [7]

$$\mathrm{i}\hbar\frac{\partial\psi}{\partial z} = H\,\psi \tag{2}$$

dove $\hbar=\lambda/(2\pi)$ è lunghezza d'onda normalizzata della luce, e l'operatore Hamiltoniano H è legato alla distribuzione dell'indice di rifrazione complesso del mezzo secondo la relazione

$$H = -\frac{\hbar^2}{2n_s}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right) + n_s - n(r) \tag{3}$$

Una distribuzione di indice di rifrazione che soddisfa la relazione (1) rende l'operatore H PT simmetrico. Quando la simmetria PT non è rotta, la luce si propaga nella struttura dielettrica senza che via sia, globalmente, assorbimento o amplificazione del segnale luminoso, sebbene fotoni vengano localmente creati e distrutti nel dielettrico. Quando la simmetria PT è rotta, vi è invece una amplificazione del segnale luminoso nella struttura, una situazione che generalmente viene detta instabilità ottica e che è alla base di fenomeni importanti quali l'emissione di luce laser per emissione stimolata dal rumore quantistico o di radiazione coerente in oscillatori ottici parametrici.

La rottura di simmetria *PT* è accompagnata dalla comparsa dei cosiddetti punti eccezionali o singolarità spettrali [22-25], che sono delle criticità matematiche dell'operatore *H* nelle quali esso perde alcune proprietà importanti (come la completezza del suo spettro). Sono proprio tali criticità che modificano in maniera unica le proprietà di trasporto della luce e che rendono possibili alcuni fenomeni quali l'invisibilità unidirezionale di un mezzo materiale o la possibilità di realizzare un laser-antilaser, come discuterò nel seguito. Un'importante applicazione dei punti eccezionali nella tecnologia laser è quella, ad esempio, di selezionare in strutture integrate micrometriche l'oscillazione monomodale [15, 16]. Due possibili schemi di strutture con simmetria parità-tempo per la selezione modale è mostrata nella *Fig. 1* [15, 16]. La prima consiste in una coppia di microanelli accoppiati per tunneling evanescente di luce, uno con guadagno e l'alto con perdite in modo da rispettare la condizione (1), mentre nella seconda struttura vi è un solo microanello con super-imposta una modulazione periodica delle perdite lungo l'anello. Nel primo caso, l'accoppiamento tra gli anelli con guadagno/perdite bilanciati induce un innalzamento del contrasto di estinzione dei modi longitudinali adiacenti dell'anello, rispetto al modo centrale che cade al centro della riga di guadagno del mezzo attivo, determinando così l'oscillazione su una singola risonanza longitudinale della cavità anche quando il laser emette molto al di sopra della sua soglia di oscillazione. Nel secondo caso la modulazione di perdite lungo l'arco dell'anello favorisce l'oscillazione di un singolo modo alla lunghezza d'onda che si sincronizza con il periodo della modulazione.

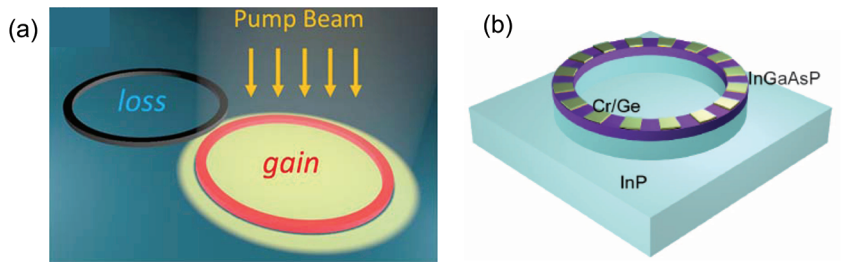


Fig. 1 – Schema di due strutture con simmetria parità-tempo in ottica integrata per la realizzazione di micro laser operanti in singolo modo.
(a) Microanelli accoppiati in modo evanescente, uno con perdite e l'altro con guadagno.
(b) Singolo micronaello con modulazione periodica delle perdite ottenuta mediante layer metallici di Cr/Ge in semiconduttori InP/InGaAsP.

3. Laser antilaser con simmetria parità-tempo

Una delle conseguenze più interessanti della simmetria *PT* applicata all'ottica è la possibilità di realizzare un dispositivo ottico attivo capace di emettere, come in un normale oscillatore laser, radiazione coerente che fuoriesce dal dispositivo ma in grado, simultaneamente, di assorbire in maniera perfetta una radiazione coerente che indice su di esso [10]. L'esistenza di un tale dispositivo sembrerebbe violare il senso comune secondo cui un mezzo materiale che amplifica luce e che quindi emette radiazione coerente quando posto in un cavità risonante (un laser) possa, simultaneamente, anche poter assorbire perfettamente luce coerente che incide su di esso (assorbitore coerente perfetto o antilaser). Tuttavia, l'esistenza di un tale dispositivo è una semplice e necessaria conseguenza della simmetria *PT* del mezzo ottico. Consideriamo, per semplicità, la propagazione della luce in una struttura ottica unidimensionale, ed indichiamo con x la coordinata spaziale di propagazione e con $\varepsilon(x) = \varepsilon_0 n^2(x)$ la constante dielettrica del mezzo spazialmente variabile in una regione di lunghezza L. La parte reale di ε consente di localizzare la luce nella regione spaziale limitata di lunghezza L (come ad esempio avviene nel caso di due specchi altamente riflettenti posti agli estremi della regione), mentre la parte immaginaria di ε descrive i processi di assorbimento e/o emissione stimolata di fotoni che avviene nel mezzo materiale. Un'onda di luce di frequenza (generalmente complessa) ω è descritta da un campo elettrico $E=E(x,t)$ che soddisfa l'equazione di Helmholtz

$$\frac{d^2E}{dx^2} + \frac{\omega^2}{c^2}\epsilon(x)E = 0 \tag{4}$$

dove c è la velocità della luce nel vuoto. L'equazione di Helmholtz, che è una conseguenza delle equazioni di Maxwell, è una equazione un po' più generale della equazione (2) introdotta nella precedente sezione, ma anche per essa possiamo introdurre degli operatori di inversione spaziale (P) e temporale (T). Supponiamo ora che il dispositivo sia una laser; allora la frequenza ω è reale e $E(x)$ descrive un'onda che fuoriesce dal mezzo, come indicato nel pannello di sinistra della *Fig. 2*. Se ora applichiamo alla equazione di Helmholtz l'operazione T di inversione temporale, e cioè consideriamo il complesso coniugato della medesima equazione, le onde uscenti dal mezzo sono ora entranti e vengono totalmente assorbite da esso. La inversione temporale genera un nuovo mezzo materiale con costante dielettrica $\varepsilon^*(x)$, che rappresenta un assorbitore perfetto coerente (il flusso di energia rispetto al dispositivo laser è ora invertito): infatti la coniugazione trasforma un mezzo amplificatore in un mezzo assorbitore; si veda il pannello di destra della *Fig. 2*.

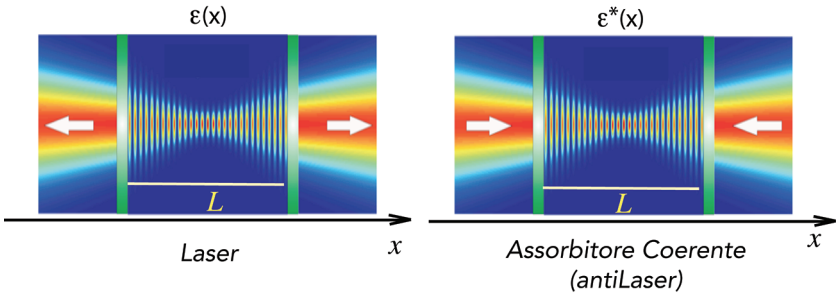


Fig. 2 – Principio di un dispositivo laser/antilaser basato sulle proprietà di simmetria parità-tempo. Il dispositivo a sinistra è un oscillatore laser che emette onde luminose uscenti dal mezzo (di lunghezza L). ε(x) è la costante dielettrica del mezzo. Se la costante dielettrica soddisfa la condizione di simmetria parità-tempo definita dalla eq. (5), il medesimo mezzo può perfettamente assorbire onde coerenti incidenti su di esso, ottenute mediante inversione temporale e spaziale delle onde emesse nel caso del laser (come in un film avvolto all'indietro e visto specularmente).

Un'analisi più completa, basata su un approccio semiclassico dell'interazione luce-materia, è presentata nella referenza [26]. Se ora il mezzo ottico è invariante per simmetria *PT*, e cioè se

$$\varepsilon(-x) = \varepsilon^*(x) \tag{5}$$

è chiaro che il medesimo dispositivo che funziona come un laser deve anche funzionare come un assorbitore coerente perfetto. La condizione (5) sulla costante dielettrica implica che il mezzo dielettrico debba avere una distribuzione spazialmente bilanciata di perdite e guadagno (parte immaginaria della costante dielettrica). Possibili realizzazioni di sistemi laser/antilaser sono mostrate nella *Fig. 3* e sono basate su strutture cosiddette a feedback distribuito, solitamente impiegate nella tecnologia dei laser a semiconduttore. Un struttura a feedback distribuito è sostanzialmente costituita da due mezzi semiconduttori trasparenti separati da una superficie che presenta una corrugazione periodica (modulazione) di periodo molto piccolo, dell'ordine della lunghezza d'onda di oscillazione. Tale corrugazione induce un accoppiamento di potenza tra le due onde contro-propaganti nel materiale (per un fenomeno noto come scattering o riflessione di Bragg) e consente così di intrappolare la luce al suo interno, agendo come uno specchio distribuito nello spazio. Per realizzare un dispositivo che possa simultaneamente operare come laser e come assorbitore coerente perfetto, metà del

materiale assorbe luce, cioè distrugge fotoni con probabilità P_1, mentre l'altra metà del materiale ne crea per emissione stimolata con una probabilità P_2 esattamente uguale a P_1 [si veda il pannello (a) della *Fig. 3*]. Quando il valore di $P_1 = P_2 = P$ è aumentato fino a raggiungere un valore critico, il sistema emette spontaneamente luce che esce nelle due direzioni opposte dello spazio (in prevalenza dalla parte dove il materiale amplifica luce), si è cioè al punto di rottura della simmetria *PT* per la presenza di una singolarità spettrale o punto eccezionale. Ma se simultaneamente due onde luminose di opportuna ampiezza e fase relative incidono da parte opposte sulla struttura, i fotoni di queste onde vengono totalmente distrutti. Una struttura fotonica simile che può realizzare la medesima funzionalità è mostrata nel pannello (b) della *Fig. 3*, ed è costituita da un microdisco (o microanello) dielettrico nel quale vi è una corrugazione simultanea dell'indice di rifrazione e delle perdite, sfasate l'una dall'altra di un quarto di periodo. Va osservato che la possibilità di poter simultaneamente distruggere completamente luce di alcune onde incidenti, e creare fotoni per le rispettive onde speculari uscenti dal mezzo è certamente un fenomeno non intuitivo ma non è paradossale, ed è appunto una conseguenza della simmetria parità-tempo e spiegabile in ultima analisi mediante processi di interferenza delle onde nelle diverse regioni di perdite e guadagno nella struttura. La dimostrazione sperimentale del primo dispositivo laser-antilaser, teoricamente previsto nel lavoro [10], è stata recentemente riportata da un gruppo di ricerca americano usando tecnologia laser a semiconduttore e pubblicata sulla prestigiosa rivista Nature Photonics [28].

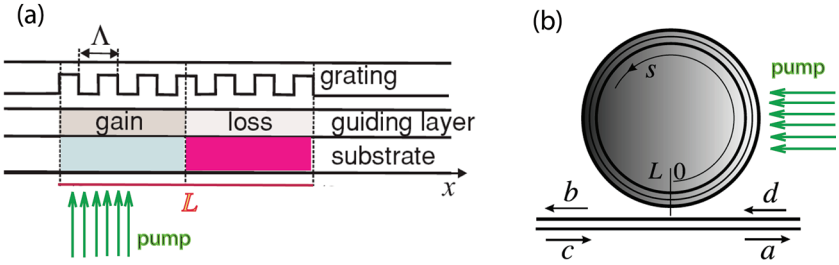


Fig. 3 – Esempi di strutture a feedback distribuito che realizzano un sistema laser/antilaser basato sul concetto di simmetria parità-tempo.
(a) Struttura a cavità lineare con reticolo di indice di rifrazione e con due regioni bilanciate di guadagno (zona verde) e perdite (zona rossa).
(b) Struttura ad anello (microanello accoppiato ad una guida d'onda di bus). L'anello presenta questa volta un reticolo di indice e di perdite.

4. Invisibilità unidirezionale in metamateriali con simmetria parità-tempo

Un'altra interessante applicazione, ispirata dai concetti di parità-tempo, è la possibilità di realizzare metamateriali 'attivi', cioè che richiedono inversione di popolazione, che appaiono invisibili in funzione della direzione di provenienza della luce [29-31]. Una struttura che può diventare totalmente invisibile quando la luce la attraversa da un lato, mentre può rimanere visibile quando viene illuminata dall'altro, si basa sul fenomeno fisico della rottura di simmetria nel fenomeno dello scattering di Bragg che avviene quando la luce attraversa un mezzo periodico, come un cristallo. Nei cristalli naturali lo scattering di Bragg avviene in egual misura per la luce che si propaga in direzioni opposte. Ma se il cristallo è sintetizzato artificialmente in modo da descrivere un cristallo 'complesso' con simmetria parità-tempo del suo indice di rifrazione, al punto della rottura della simmetria *PT* lo scattering di Bragg è completamente assente quando la luce incide in una direzione del cristallo, ma non in quella opposta [21, 24]; si veda la *Fig. 4*.

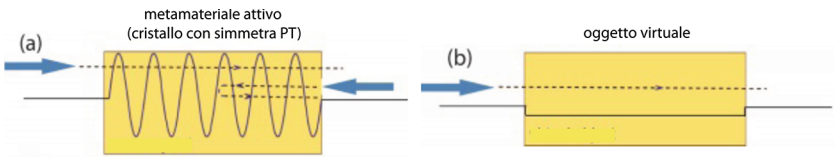


Fig. 4 – Principio dell'invisibilità unidirezionale in un metamateriale attivo con simmetria parità-tempo. Il metamateriale, mostrato schematicamente nel pannello (a) della figura, è costituito da un cristallo sintetico (meta-cristallo) che mostra una modulazione periodica dell'indice di rifrazione complesso. Quando le parti reale ed immaginarie dell'indice di rifrazione hanno la stessa ampiezze e sono sfasate di un quarto di periodo, il meta-cristallo è al punto della rottura della simmetria PT e risulta invisibile per un'onda che si propaga in una direzione (da sinistra a destra nella figura), ma non in quella opposta. Per tale onda il cristallo appare invisibile, e precisamente ad un osservatore appare come un mezzo omogeneo che non riflette la luce [pannello (b) della figura].

Tale fenomeno è stato sperimentalmente dimostrato recentemente da due gruppi sperimentali usando come sistemi attivi guide dielettriche e fibre ottiche attive [12, 13]; si veda la *Fig. 5*. Metamateriali capaci di rendere invisibili degli oggetti o, all'opposto, di ottenere una risoluzione delle immagini molto superiore al normale, sono già stati concepiti e dimostrati in ottica, ma nelle soluzioni precedentemente

adottate l'alterazione attenta e specifica della distribuzione spaziale dell'indice di rifrazione all'interno del metamateriale era di tipo passivo e solitamente richiedeva strutture anisotrope e magnetiche [32].

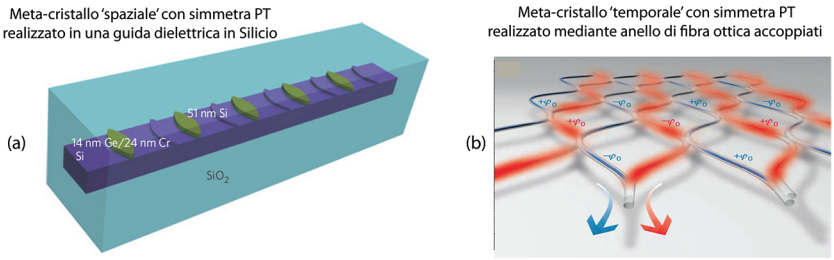


Fig. 5 – (a) Implementazione di un meta-cristallo con simmetria PT basato sulla propagazione spaziale di luce in una guida dielettrica di Silicio in SOI con reticolo di indice e di perdite sfasati di un quarto di periodo. (b) Implementazione di un meta-cristallo con simmetria PT basato sulla propagazione di impulsi ottici luminosi in anelli accoppiati di fibra ottica. Gli anelli sono disposti in una geometria simmetrica con alternanza di amplificazione e di perdite in anelli contigui.

Nelle strutture ottiche a simmetria parità-tempo, invece, il metamateriale è di tipo 'attivo', cioè richiede che l'indice di rifrazione sia complesso e vi sia una distribuzione bilanciata di perdite e guadagno. Le proprietà di invisibilità dei matemateriali attivi possono essere superiori di quelli dei soli metamateriali passivi [33], e questo apre quindi possibilità nuove di applicazioni all'imaging ottico. Inoltre, in tali mezzi è possibile introdurre forme parziali di invisibilità, ad esempio è possibile sintetizzare materiali sintetici che mostrano una invisibilità unidirezionale per metà dei colori della luce (ad esempio per lunghezze d'onda lunghe fino al rosso), ma non per l'altra metà (dal rosso alle lunghezze d'onda più corte) [34]. Una applicazione interessante della invisibilità unidirezionale, basata alla rottura di simmetria *PT* ad un punto eccezionale, è quella della generazione di luce strutturata, che trasporta momento angolare orbitale, direttamente da un microlaser ad anello [35]. Tale tipo di luce, che trova importanti applicazioni ad esempio nella codifica e trasmissione di informazione ad altissimo bit rate, è costituita da fasci di luce vorticosi (vedi la *Fig. 6*), in cui cioè la parte centrale del fascio di luce è buio (zero di campo) mentre la fase del campo varia di un multiplo intero di 2π (carica topologica del vortice) quando si percorre un cammino chiuso attorno al vortice. Luce vorticosa generata su scala micrometrica direttamente da un microla-

ser ad anello richiede di poter forzare l'oscillazione laser nell'anello in una sola direzione (oraria o antioraria), per evitare che la vorticità del fascio emesso dall'anello vari stocasticamente seguendo le fluttuazioni aleatorie della oscillazione oraria-antioraria del laser innescata da processi non-lineari e di competizione modale nel microanello. Tale unidirezionalità richiede solitamente l'utilizzo di un isolatore ottico (diodo ottico) basato su effetti magneto-ottici, che sono tuttavia impraticabili su scale spaziali piccole (dell'ordine di qualche micrometro). L'uso di un reticolo combinato di indice e di perdite lungo il microanello, ingegnerizzato in modo da operare al punto della rottura della simmetria *PT*, consente di stabilizzare l'oscillazione unidirezionale del laser, senza l'utilizzo di un isolatore ottico di tipo magnetico (rotatorie di Faraday).

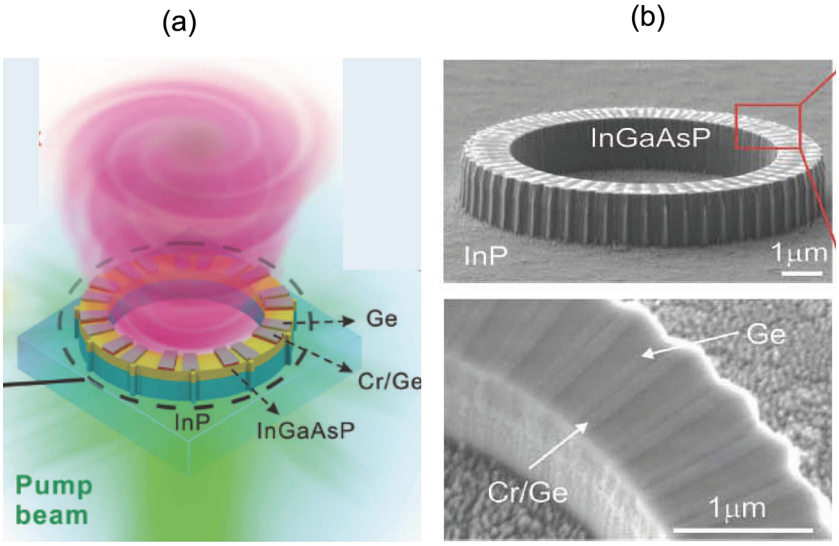


Fig. 6 – (a) Schema della generazione di luce strutturata (luce vorticosa che trasporta momento angolare orbitale) da un laser a microanello in cui l'oscillazione unidirezionale (oraria o antioraria) è ottenuta mediante l'uso di un cristallo PT simmetrico al punto di rottura della simmetria (punto eccezionale). (b) Microalaser in tecnologia semiconduttore InP/InGaAsP che ha recentemente consentito la dimostrazione sperimentale della generazione di luce strutturata direttamente da un microanello del diametro di soli 9 μm.

Come osservazione conclusiva, si deve menzionare che il problema dell'invisibilità in mezzi attivi, qui illustrato con riferimento ai concetti di simmetria parità-tempo, può essere anche esteso a famiglie di

sistemi ottici più generali che non godono di tale simmetria. Per esempio, utilizzando i concetti di supersimmetria noti nella meccanica quantistica, è possibile sintetizzare dei difetti invisibili in un cristallo, difetti che cioè non alterano le proprietà di propagazione dell'onda nel cristallo [36]. Questo risultato è piuttosto interessante e desueto, perché difetti reticolari solitamente producono scattering dell'onda. Recentemente, sono state anche scoperte ed investigate alcune proprietà interessanti, fra cui la possibilità di ottenere invisibilità unidirezionale o bidirezionale, in mezzi ottici nei quali le distribuzioni spaziali delle parti reali ed immaginarie dell'indice di rifrazione soddisfano certe relazioni speciali, le cosiddette relazioni di Kramers-Kronig [37-40] note anche in matematica come trasformazioni di Hilbert. Queste relazioni, che sono ben note in ottica ed esprimono il principio di causalità nella risposata di ogni mezzo materiale esistente in natura, garantiscono l'assenza di onde riflesse quando il mezzo è illuminato da un lato, indipendentemente dall'angolo di incidenza della luce, ma solitamente non dal lato opposto.

5. Conclusioni

In questa breve Nota si è discusso in maniera introduttiva e semplificata dei principi teorici e delle applicazioni dei concetti di simmetria parità-tempo in ottica, con particolare riguardo alla sintesi di materiali artificiali (metamateriali) che promettono interessanti sviluppi ed applicazioni nelle tecnologie laser, nella fotonica integrata, nel sensing e nelle tecniche di imaging ottico. L'ottica dei sistemi con simmetria parità-tempo è un settore di ricerca piuttosto recente ma molto promettente, che ha tratto molte idee e beneficiato di metodi di analisi sviluppati in meccanica quantistica. Oltre alla simmetria parità-tempo, altre proprietà di simmetria della meccanica quantistica e della teoria dei campi possono trovare interessanti applicazioni all'ottica. Menzioniamo brevemente, a titolo d'esempio, la simmetria di inversione temporale T, che trova applicazioni nella realizzazione di assorbitori coerenti anche policromatici [41], le proprietà di supersimmetria che possono essere ingegnosamente utilizzate per progettare circuiti fotonici integrati [42-44], ed altre proprietà di simmetria (come quella di chiralità) che garantiscono ad esempio l'esistenza di modi di propagazione di *edge* della luce insensibili al disordine della struttura

a causa di una protezione di natura topologica (si tratta cioè di strutture fotoniche con proprietà simili ai cosiddetti isolanti topologici studiati nella Fisica dello stato solido). Con le teorie di supersimmetria è possibile costruire, per esempio, intersezioni trasparenti di guide dielettriche, cioè incroci di guide le quali non interferiscono le une con le altre ed i fotoni si propagano indipendentemente in esse (si veda la *Fig. 7*). Visto l'elevato interesse che tali tematiche hanno suscitato nella comunità scientifica internazionale, è prevedibile nei prossimi anni un rapido sviluppo e una maturazione tecnologica di queste ricerche che consentiranno di progettare e realizzare circuiti fotonici innovativi e materiali sintetici di nuova generazione con importanti applicazioni alla fotonica, alle tecnologie ottiche, alla sensoristica ed alle tecnologie dei laser.

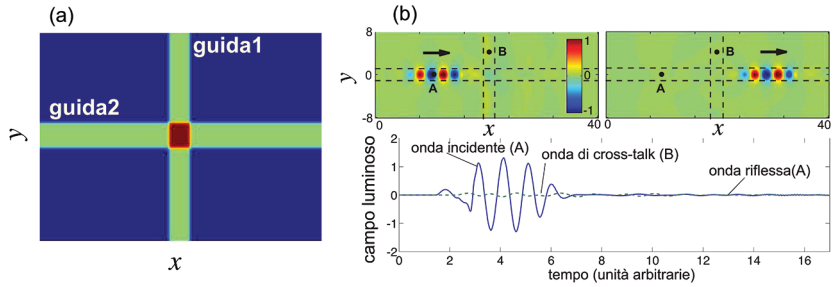


Fig. 7 – Intersezione trasparente di due guide ottiche basato sull'uso della supersimmetria della meccanica quantistica. (a) Schema di due guide di luce che si intersecano. L'intersezione provoca generalmente un disturbo per la luce che si propaga nelle due guide. Tuttavia, se l'indice di rifrazione delle guide soddisfa a certe proprietà di supersimmetria, le onde luminose in ciascuna delle due guide attraversano l'intersezione senza subire riflessioni o senza disturbare il canale dell'altra guida. In sostanza l'intersezione appare trasparente. La trasparenza indotta da supersimmetria è mostrata nel pannello (b), in cui sono indicati l'andamento nel tempo di un segnale luminoso misurato nei punti A e B delle due guide. La trasparenza dell'intersezione si evince dal fatto che in B non si osserva luce (assenza del cosiddetto segnale di cross-talk) ed in A il segnale ritardato nel tempo (eco) corrispondente a riflessione di luce alla discontinuità praticamente assente.

Ringraziamenti

L'autore ringrazia la Pia Fondazione Edoardo Kramer e l'Istituto Lombardo Accademia di Science e Lettere per l'assegnazione del premio Edoardo Kramer 2015.

BIBLIOGRAFIA

[1] E. Castellani, "*Simmetria e Natura*" (Laterza, Bari-Roma, 2001).

[2] "*Simmetria e realtà"* in: Le Scienze, Quaderni n. 118. A cura di E. Castellani, Febbraio 2001.

[3] C.M. Bender, S. Boettcher, "*Real Spectra in Non-Hermitian Hamiltonians Having PT Symmetry*", Physical Review Letters (1998), vol. 80 pp. 5243-5246 - DOI: 10.1103/PhysRevLett.80.5243.

[4] C.M. Bender, D.C. Brody, and H.F. Jones, "*Complex Extension of Quantum Mechanics*", Physical Review Letters (2002), vol. 89, pp. 270401 - DOI: 10.1103/PhysRevLett.89.270401.

[5] C. Bender, "*Making Sense of Non-Hermitian Hamiltonians*", Reports Progress Physics (2007), vol. 70, pp. 947-1018 - DOI:10.1088/0034-4885/70/6/R03.

[6] Y.-C. Lee, M.-H. Hsieh, S.T. Flammia, and R.-K. Lee, "*Local PT symmetry Violates the No- Signaling Principle"*, Physical Review Letters (2014), vol. 112, pp. 130404 – DOI: 10.1103/PhysRevLett.112.130404.

[7] S. Longhi, "*Quantum-optical analogies using photonic structures"*, Laser & Photonics Review (2009), vol. 3, pp. 243-261 - DOI: 10.1002/lpor.200810055.

[8] M. Schirber, "*Synopsis: Reflecting on an Alternative Quantum Theory*", Physics (https://physics.aps.org/synopsis-for/10.1103/PhysRevLett.112.130404).

[9] C.E. Ruter, K.G. Makris, R. El-Ganainy, D.N. Christodoulides, and M. Segev, *"Observation of parity-time symmetry in optics"*, Nature Physics (2010), vol.6, pp. 192-195 - DOI: 10.1038/nphys1515.

[10] S. Longhi, "*PT -symmetric laser-absorber*", Physical Review A (2010), vol. 82, pp. 031801- DOI: 10.1103/PhysRevA.82.031801.

[11] L. Feng, M. Ayache, J. Huang, Y.L. Xu, M.H. Lu, Y.F. Chen, Y. Fainman, and A. Scherer, "*Nonreciprocal light propagation in a silicon photonic circuit*," Science (2011), vol. 333, pp.729-733 - DOI: 10.1126/science.1206038.

[12] L. Feng, Y.-L. Xu, W. S. Fegadolli, M.-H. Lu, J.E.B. Oliveira, V. R. Almeida, Y.-F. Chen, and A. Scherer, *"Experimental demonstration of a unidirectional reflectionless parity-time metamaterial at optical frequencies"*, Nature Materials (2012), vol.12, 108-113 - DOI: 10.1038/nmat3495.

[13] A. Regensburger, C. Bersch, M.-A. Miri, G. Onishchukov, D.N. Christodoulides, and U. Peschel, *"Parity-time synthetic optical lattices",* Nature (2012), vol. 488, pp.167-171 - DOI: 10.1038/nature11298.

[14] L. Feng, Z.J. Wong, R.M. Ma, Y. Wang, and X. Zhang, *"Single-mode laser by parity-time symmetry breaking",* Science (2014) vol. 346, pp. 972-975 - DOI: 10.1126/science.1258479.

[15] H. Hodaei, M.A. Miri, M. Heinrich, D.N. Christodoulides, and M. Khajavikhan, *"Parity-time symmetric microring lasers"*, Science (2014), vol. 346, 975-978 - DOI: 10.1126/science.1258480.

[16] R. Fleury, D. Sounas, and A. Alù, *"Negative Refraction and Planar Focusing Based on Parity-Time Symmetric Metasurfaces"*, Physical Review Letters (2014), vol. 113, pp. 023903 - DOI: 10.1103/PhysRevLett.113.023903.

[17] R. Fleury, D. L. Sounas, and A. Alù, *"An Invisible Acoustic Sensor Based on*

Parity-Time Symmetry", Nature Communications (2015), vol. 6, pp. 5905 - DOI: 10.1038/ncomms6905.

[18] A. Ruschhaupt, F. Delgado, and J. G.Muga, *"Physical realization of PT -symmetric potential scattering in a planar slab waveguide"*, Journal of Physics A (2005), vol. 38, pp. L171-L176 - DOI: 10.1088/0305-4470/38/9/L03.

[19] R. El-Ganainy, K.G. Makris, D.N. Christodoulides, and Z.H. Musslimani, *"Theory of coupled optical PT -symmetric structures"*, Optics Letters (2007), vol. 32, 2632-2634 –DOI: 10.1364/OL.32.002632.

[20] K.G.Makris, R. El-Ganainy, D.N. Christodoulides, and Z.H. Musslimani, *"Beam dynamics in PT symmetric optical lattices"*, Physical Review Letters (2008), vol.100, pp.103904 - DOI: 10.1103/PhysRevLett.100.103904.

[21] S. Longhi, *"Bloch oscillations in complex crystals with PT symmetry"*, Physical Review Letters (2009), vol. 103, pp. 123601 - DOI: 10.1103/PhysRevLett.103.123601.

[22] M.V. Berry, *"Physics of non-Hermitian degeneracies",* Czech. Journal of Physics (2004), vol. 54, 1039-1047 - DOI: 10.1023/B:CJOP.0000044002.05657.04.

[23] W.D. Heiss, *"The physics of exceptional points"*, Journal of Physiscs A (2012), vol. 45, pp. 444016 - DOI: 10.1088/1751-8113/45/44/444016.

[24] S. Longhi, *"Spectral singularities and Bragg scattering in complex crystals"*, Physical Review A (2010), vol. 81 pp. 022102 - DOI: 10.1103/PhysRevA.81.022102.

[25] S. Longhi, *"Optical Realization of Non-Hermitian Quantum Mechanics"*, Physical Review Letters (2010), vol. 105, pp. 013903 - DOI: 10.1103/PhysRevLett.105.013903.

[26] S. Longhi, *"Coherent perfect absorption in a homogeneously broadened two-level medium"*, Physical Review A (2011), pp. 055804 - DOI: 10.1103/PhysRevA.83.055804.

[27] S. Longhi, L. Feng, *"PT-symmetric microring laser- absorber",* Optics Letters (2014), vol. 39 pp. 5026-5029 - DOI: 10.1364/OL.39.005026.

[28] Z.J. Wong, Y.-L. Xu, J. Kim, K. O'Brien, Y. Wang, L. Feng, and X. Zhang, *"Lasing and anti-lasing in a single cavity"*, Nature Photonics (2016), vol. 10, pp. 796-801 - DOI: 10.1038/nphoton.2016.216.

[29] Z. Lin, H. Ramezani, T. Eichelkraut, T. Kottos, H. Cao, and D.N. Christodoulides, *"Unidirectional Invisibility Induced by PT-Symmetric Periodic Structures"*, Physical Review Letters (2011), vol. 106, pp. 213901 - DOI: 10.1103/PhysRevLett.106.213901.

[30] S. Longhi, *"Invisibility in PT -symmetric complex crystals"*, Journal of Physics A (2011), vol. 44, pp. 485302 - DOI: 10.1088/1751-8113/44/48/485302.

[31] S. Longhi, *"A unidirectionally invisible PT -symmetric complex crystal with arbitrary thickness"*, Journal of Physics A (2014), vol. 47, pp. 485302 - DOI: 10.1088/1751-8113/47/48/485302.

[32] I. Fausto, *"Fisica dell'invisibilità. Metamateriali e tecnologie del futuro"* (Arance Editrice, Milano, 2013).

[33] S. Longhi, *"Invisibility in non-Hermitian tight-biniding lattices"*, Physical Review A (2010), vol. 82, pp. 032111 - DOI: 10.1103/PhysRevA.82.032111.

[34] S. Longhi, *"Half-spectral unidirectional invisibility in non- Hermitian periodic optical structures"*, Optics Letters (2015), vol. 40, pp. 5694-5697 - DOI: 10.1364/OL.40.005694.

[35] P. Miao, Z. Zhang, J. Sun, W. Walasik, S. Longhi, N.M. Litchinitser, and L. Feng, *"Orbital angular momentum microlaser"*, Science (2016), vol. 353, pp. 464-467 - DOI: 10.1126/science.aaf8533.

[36] S. Longhi and G. Della Valle, *"Invisible defects in complex crystals"*, Annals of Physics (2013), vol. 334, pp. 35-46 - DOI: 10.1016/j.aop.2013.03.011.

[37] S.A.R. Horsley, M. Artoni, G.C. La Rocca, *"Spatial Kramers-Kronig relations and the reflection of waves"*, Nature Photonics (2015), vol. 9, pp. 436-439 - DOI: 10.1038/nphoton.2015.106.

[38] S. Longhi, *"Wave reflection in dielectric media obeying spatial Kramers-Kronig relations"*, EuroPhysics Letters (2015), vol. 112, pp. 64001 - DOI: 10.1209/0295-5075/112/64001.

[39] S.A.R. Horsley and S. Longhi,*"Spatiotemporal deformations of reflectionless potentials"*, Physical Review A (2017), vol. 96, pp. 023841 - DOI: 10.1103/PhysRevA.96.023841.

[40] S. Longhi, *"Bidirectional invisibility in Kramers–Kronig optical media"*, Optics Letters (2016), vol. 41, pp. 3727-3730 - DOI: 10.1364/OL.41.003727.

[41] S. Longhi, *"Time-Reversed Optical Parametric Oscillation"*, Physical Review Letters (2011), vol. 107, pp. 033901 - DOI: 10.1103/PhysRevLett.107.033901.

[42] M.-A. Miri, M. Heinrich, and D.N. Christodoulides, *"SUSY-inspired one-dimensional transformation optics"*, Optica (2014), vol. 1, pp. 89-95 - DOI: 10.1364/OPTICA.1.000089.

[43] S. Longhi, *"Supersymmetric transparent optical intersections"*, Optics Letters (2015), vol. 40, pp. 463-466 - DOI: 10.1364/OL.40.000463.

[44] S. Longhi, *"Supersymmetric Bragg gratings"*, Journal of Optics (2015), vol.17, pp. 045803 - DOI: 10.1088/2040-8978/17/4/045803.

INDICE DEL VOLUME

RIMINI A., *Entanglement* quantistico 3

SMIRAGLIA C., DIOLAIUTI G.A., *et al.*, Lo stato attuale dei ghiacciai italiani e la loro recente evoluzione. Risultati di un nuovo inventario 15

DIONIGI R., TERRENI A., Giovanni Battista Palletta (1748-1832) «Dotto chirurgo e pratico eccellente» a capo della chirurgia dell'*Ospedale Maggiore* 31

SACCHI LANDRIANI G., Alchimia, preludio alla chimica? 55

GIULIANI A., Order, disorder and phase transitions in quantum many body systems 73

COCCHIARELLA L., "In proiezione": spazio e immagine in *Las Meninas* di Diego Velázquez 117

CATTANEO E., ZUFFARDI O., Dalla scoperta del DNA ad oggi: la medicina dentro al genoma 133

VERGARA C., Numerical solution of blood dynamics in atherosclerotic carotids 135

BERSANELLI M., From Dante's universe to contemporary cosmology 147

CITRINI C., Sant'Agostino e le scienze 167

RAIMONDO S., Gli esosomi nella comunicazione cellula-cellula 183

FIGLIOLI G., *et al.*, Identification and functional characterization of germline genetic variants predisposing to differentiated thyroid cancer risk 191

MACCONE L., Computer quantistici: status e prospettive 203

ROSSIELLO F., *et al.*, Study on the genetic predisposition to breast cancer and the DNA damage response at telomeres and in cellular senescence 213

LONGHI S., Simmetria parità-tempo in ottica: verso una nuova ingegneria della luce 221

INDICE PER MATERIA

ARCHITETTURA

COCCHIARELLA L., "In proiezione": spazio e immagine in *Las Meninas* di Diego Velázquez 117

CHIMICA

SACCHI LANDRIANI G., Alchimia, preludio alla chimica? 55

FISICA

BERSANELLI M., From Dante's universe to contemporary cosmology 147

LONGHI S., Simmetria parità-tempo in ottica: verso una nuova ingegneria della luce 221

MACCONE L., Computer quantistici: status e prospettive 203

RIMINI A., *Entanglement* quantistico 3

MATEMATICA

CITRINI C., Sant'Agostino e le scienze 167

GIULIANI A., Order, disorder and phase transitions in quantum many body systems 73

VERGARA C., Numerical solution of blood dynamics in atherosclerotic carotids 135

MEDICINA

CATTANEO E., ZUFFARDI O., Dalla scoperta del DNA ad oggi: la medicina dentro al genoma 133

DIONIGI R., TERRENI A., Giovanni Battista Palletta (1748-1832) «Dotto chirurgo e pratico eccellente» a capo della chirurgia dell'*Ospedale Maggiore* 31

Scienze naturali

Biologia

FIGLIOLI G., *et al.*, Identification and functional characterization of germline genetic variants predisposing to differentiated thyroid cancer risk 191

RAIMONDO S., Gli esosomi nella comunicazione cellula-cellula 183

ROSSIELLO F., *et al.*, Study on the genetic predisposition to breast cancer and the DNA damage response at telomeres and in cellular senescence 213

Geologia

SMIRAGLIA C., DIOLAIUTI G.A., *et al.*, Lo stato attuale dei ghiacciai italiani e la loro recente evoluzione. Risultati di un nuovo inventario 15

Finito di stampare nel mese di marzo 2020

Decreto Pres. Tribunale 10 gennaio 1961, n. 5504 registro cancelleria
Direttore responsabile: Adele Robbiati Bianchi