

Numero III
Anno 2014


DIPARTIMENTO DI MATEMATICA E FISICA
"Ennio De Giorgi"

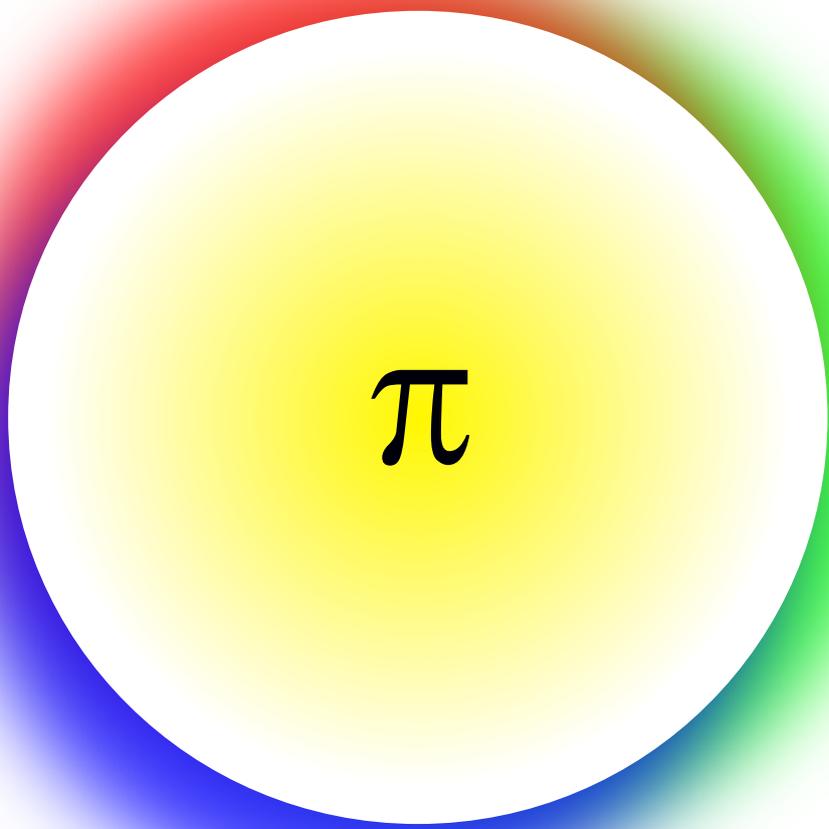

UNIVERSITÀ
DEL SALENTO

Viaggio nella Scienza

Ithaca



π



Arte e Scienza

Ithaca: Viaggio nella Scienza

Una pubblicazione del Dipartimento di Matematica e Fisica “Ennio De Giorgi” dell’Università del Salento.

Registrazione presso il Tribunale di Lecce n. 6 del 30 Aprile 2013.
e-ISSN: 2282-8079

Direttore Responsabile
Luigi Spedicato

Ideatore
Giampaolo Co’

Comitato di Redazione
Rocco Chirivì,
Maria Luisa De Giorgi,
Luigi Martina,
Giuseppe Maruccio,
Marco Mazzeo,
Francesco Paparella,
Carlo Sempi.

Segreteria di Redazione
Daniela Dell’Anna.

© 2013-2014 Dipartimento di Matematica e Fisica “*Ennio de Giorgi*”.

© 2014 per i singoli articoli dei rispettivi autori.

Il materiale di questa pubblicazione può essere riprodotto nei limiti stabiliti dalla licenza
“*Creative Commons Attribuzione – Condividi allo stesso modo 3.0 Italia*” (CC BY-SA 3.0 IT).

Per il testo della licenza: <http://creativecommons.org/licenses/by-sa/3.0/it/deed.it>

Ithaca: Viaggio nella Scienza

è disponibile sul sito:

<http://ithaca.unisalento.it/>

Scriveteci all’indirizzo:

ithaca@unisalento.it

- 3** In questo numero
- 5** Arte e scienza. Dal futurismo all'arte moltiplicata
Luca Zaffarano
- 15** Un matematico passeggia per l'Alhambra
Rafael Pérez Gómez
- 25** La fisica della musica
Giampaolo Co'
- 33** Una geometrizzazione dello spazio degli accordi
Mattia G. Bergomi, Riccardo D. Jadanza, Alessandro Portaluri
- 47** Bosoni, Fermioni e ...
Luigi Martina
- 61** Fermioni di Majorana in sistemi a due dimensioni?
Pasquale Scarlino
- 67** *La lezione mancata*
Breve storia dei numeri primi
Alessandro Zaccagnini

In questo numero

Abbiamo pensato di cambiare la scadenza semestrale di Ithaca, passando da Giugno-Dicembre ad Aprile-Ottobre, ecco perché questo numero appare soltanto dopo quattro mesi dalla pubblicazione di quello precedente.

Il tema principale di questo numero riguarda **Arte e Scienza**. Il primo articolo è di Luca Zaffarano ed affronta il rapporto tra l'arte futurista nel suo sviluppo storico e la scienza e la tecnica. Segue un articolo di Rafael Pérez Gómez che descrive come nelle tassellature delle piastrelle decorative di quel magnifico monumento che è l'Alhambra di Granada siano identificabili tutti i possibili gruppi cristallografici piani. Seguono due articoli che parlano di musica. L'articolo di Giampaolo Co', propedeutico a quello di Bergamo, Jadanza e Portaluri, tratta delle basi fisiche della musica e della relazione tra le frequenze delle note che costituiscono la scala musicale utilizzata nella musica occidentale. Nel secondo articolo sulla musica, vengono analizzati, in termini geometrici, i rapporti di frequenze di certe strutture armoniche musicali, ottenendo delle sorprese.

Accanto agli articoli legati al tema principale, presentiamo altri tre contributi. Gli articoli di Martina e quello di Scarlino sono legati tra loro. Martina discute la dipendenza dalla geometria di un osservabile delle particelle elementari tipicamente quantistico, lo spin, e indica come sistemi quantistici in dimensioni inferiori a tre possano avere spin né intero né semi-intero. L'articolo di Scarlino parla di una possibile scoperta di particelle di questo tipo che, in aggiunta, presentano anche la caratteristica di essere identiche alle loro antiparticelle. Infine, per la serie "la lezione man-

cata", l'articolo di Zaccagnini offre una approfondita panoramica dei più importanti risultati sulla distribuzione dei numeri primi.

Arte e Scienza sarà il tema che caratterizzerà anche il prossimo numero.

Buona lettura,
il Comitato di Redazione.

Arte e scienza. Dal futurismo all'arte moltiplicata

Luca Zaffarano munart.org

Il futurismo italiano è oggetto di una costante e crescente attenzione critica. Attraverso un breve *excursus* di alcune delle correnti artistiche più sperimentali che hanno accompagnato la trasformazione economica e culturale dell'Italia del XX secolo, si mette in evidenza come il vero big-bang dei grandi cambiamenti avvenuti nei linguaggi dell'arte sia rintracciabile nelle posizioni teoriche del movimento futurista.

Questo articolo è stato scritto mentre l'ufficio stampa del Guggenheim diffondeva le prime note sulla grande mostra newyorkese dedicata al movimento futurista in programma a partire dal 21 febbraio 2014 [1]. Per una volta l'Italia viene chiamata ad assumere il ruolo di paese esportatore, in particolare di quello che risulta essere il movimento artistico italiano più importante di tutto il nostro novecento ed anche quello internazionalmente più noto, il cui valore culturale ed economico (per marchio, identità, riconoscibilità) non è mai stato né stimato né sfruttato adeguatamente. La mostra in preparazione racconta un pezzo importante della nostra storia artistica con oltre 350 opere, un numero appena

sufficiente per descrivere le complesse vicende di questa avanguardia. Il nostro paese, anche in presenza di un passato da tutti riconosciuto come anticipatore di molte correnti artistiche che hanno dato voce alla modernità del XX secolo (dadaismo, surrealismo, costruttivismo, neoplasticismo), appare sostanzialmente ancora incapace di valorizzare il suo passato più noto e consolidato. Filippo Tommaso Marinetti e il futurismo sono stati "sdoganati" solo verso la fine degli anni '80 dall'importante storico d'arte (svedese) Pontus Hulten, già direttore del Museo Pompidou di Parigi, che nella rinnovata sede veneziana di Palazzo Grassi ha curato la mostra "*Futurismo & Futurismi*", riunendo in una sola esposizione tutti i principali capolavori sparsi nel mondo [2]. Analogamente l'impianto della mostra americana, che ha richiesto ben sei anni di lavoro e che tocca un ampio intervallo, dalla nascita del futurismo con la pubblicazione del primo manifesto a Parigi nel 1909 sul quotidiano *Le Figaro*, fino alla morte nel 1944 del suo fondatore e *deus ex machina* Filippo Tommaso Marinetti, è ampio e variegato e cerca di focalizzare l'attenzione sulla complessità multidisciplinare del movimento, senza alcuna distinzione artificiosa tra un futurismo delle origini (primo e secondo decennio del secolo) ed un "secondo" futurismo più tardivo (anni '30 e '40).

Arte, tecnologia, scienza

Quale legame può esistere tra un evento celebrativo dedicato ad una rivisitazione critico-storica della *"Ricostruzione Futurista dell'Universo"* e la tecnologia o la scienza? In generale esiste un ambito comune, un territorio di incontro tra l'arte visiva e le discipline scientifiche? A Marinetti e al movimento futurista va il merito di aver compreso che nella società moderna la cultura è in gran parte tecnologica ed estetica. Agli occhi dei protagonisti appare innegabile che il secolo sta per diventare visivo, con una predominanza della funzione dell'immagine all'interno dei mezzi di comunicazione di massa, ed è altresì evidente che la tecnica inizia a trasformare in modo pervasivo ogni aspetto della vita quotidiana. L'obiettivo principale del futurismo può essere riassunto nel tentativo irruento e fortemente provocatorio di aprire le porte della modernità all'Italia. Dal grande *big-bang* futurista si è poi sviluppato, attraverso un lungo processo evolutivo e sperimentale che è durato diversi decenni, un insieme eterogeneo di correnti e di autori che potremmo raggruppare genericamente sotto la definizione di "arte esatta" e che hanno tratto spunti, metodi e riflessioni dalle discipline scientifiche o dai risultati della ricerca scientifica. D'altro canto non ci può essere evoluzione di una cultura specifica, inclusi i linguaggi dell'arte, senza sperimentazione.

Metodi e processi

Si può parlare anche in arte di metodo scientifico, ripetibile, falsificabile? Vedremo alcuni esempi di come questo approccio sia stato sviluppato e affrontato, sia attraverso l'uso di un metodo fortemente sperimentale con la finalità di offrire alla comunicazione visiva soluzioni e codici innovativi (l'artista si trasforma in ricercatore), sia attraverso il formarsi di un rapporto proficuo con l'industria, attraverso le nuove tecnologie e l'impiego di materiali innovativi (l'artista diventa in un certo senso produttore). Sullo sfondo di queste grandi trasformazioni, che toccano il ruolo stesso dell'artista all'interno di una società industriale avanzata, vi è poi il tentativo estetico-filosofico di dar rappresentazione visiva alla instabilità della forma di un mondo, complesso e

fragile allo stesso tempo, in radicale e veloce trasformazione, per il quale vale la massima, per dirla con le parole di Umberto Boccioni, di una realtà che non riposa mai.

Contaminazioni

Le discipline artistiche nel corso del '900 hanno subito l'influsso delle grandi scoperte scientifiche e tecnologiche come le matematiche non euclidee, la teoria della relatività, la conquista dello spazio, le scoperte chimiche e dei nuovi materiali (si pensi alla plastica ed alle ricerche condotte dal premio Nobel Giulio Natta), la nascita di nuove discipline come la cibernetica e la teoria dell'informazione, la matematica dei frattali, la teoria dei sistemi complessi, la fisica subatomica e molto altro ancora. Anche le invenzioni tecnologiche hanno profondamente influito sul pensiero artistico: l'aereo, la fotografia, il cinema, la radio, la televisione, la fotocopiatrice, il computer. Spesso le teorie scientifiche o le invenzioni tecnologiche hanno prodotto un vero cambio di paradigma nei linguaggi dell'arte.

Sperimentazioni

L'arte del '900 è stata, da un punto di vista critico, suddivisa in due filoni primari: l'avanguardia e la tradizione. Alla luce delle grandi trasformazioni avvenute è però forse più ragionevole distinguere tra arte sperimentale, che fa ricerca di materiali, di tecniche, di processi, di metodi, di idee, ed arte che non possiede al contrario alcun elemento di progettualità sperimentale.

La storia, i protagonisti

Proviamo ora a raccontare alcune storie esemplari offrendo al lettore, attraverso una breve descrizione delle motivazioni dei protagonisti, una serie di esempi di arte prevalentemente sperimentale, certamente meritevoli di ulteriori e più dettagliati approfondimenti.

Il grande Big-Bang futurista (1909-1944)

Il fondatore del movimento futurista Filippo Tommaso Marinetti è una figura complessa e

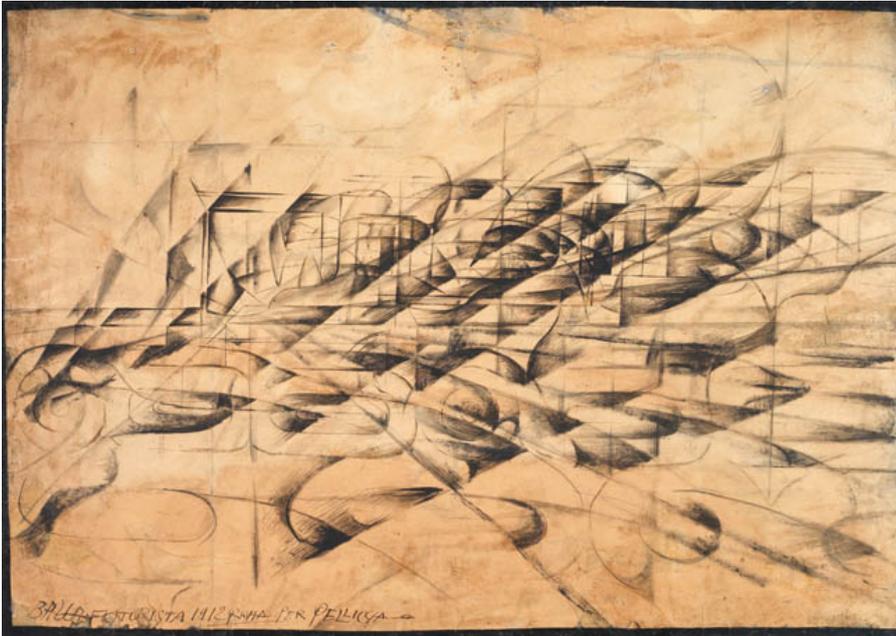


Figura 1: Giacomo Balla, *Disgregazione × velocità. Penetrazioni dinamiche di un'automobile*. (1912). Per gentile concessione di Sonia e Massimo Cirulli Archive / Christie's.

sfaccettata di artista, poeta, pensatore, editore, agitatore, mecenate. Una figura con idee ed abilità decisamente in anticipo rispetto ai suoi tempi, come, ad esempio, la capacità di usare i mass-media riuscendo a trasformare l'informazione in realtà (Marinetti pubblica il manifesto che dà origine al futurismo prima ancora della creazione di alcun quadro da parte di qualunque futurista); è in grado di generare scandalo facendo parlare di sé e del movimento con slogan che si fissano indelebili nella mente ("Uccidiamo il Chiaro di Luna!", "La guerra, sola igiene del mondo"); è organizzatore tenace di spettacoli che diventano azioni dimostrative e performance, che poi si trasformano in scazzottate e notizia per i giornali; è in contatto con tutti i massimi esponenti della cultura europea e mondiale che in varie occasioni ambiscono ad incontrarlo (da Ezra Pound a Picasso, da Majakovskij a Stravinsky, da Apollinaire a Borges). Il futurismo ha una *pars destruens* - certamente la più nota - il cui obiettivo è rovesciare la cultura *passatista* incapace di dare forma e visione alle grandi trasformazioni che si preannunciano. La stessa adorazione per la violenza ha per lo più lo scopo di scandalizzare, di fare *tabula rasa* del passato. La *pars costruens* invece si concretizza in un insieme straripante di invenzioni che si sviluppano in ogni ambito artistico: la poesia, la pittura, la scultura, l'architettura, le arti applicate, la musica, il cinema, la fotografia, la cucina, il volo, l'editoria, il teatro e molto altro ancora. Un'immagine stereotipata tende a ridurre il futu-

rismo a semplice apologia del "movimento", del "dinamismo" e della "velocità", proprio perché i fondamenti della sua poetica rispecchiano la meccanizzazione del mondo operata dalla rivoluzione industriale. Una diversa interpretazione della poetica futurista è invece centrata sulla instabilità e sulla tensione del divenire di forme in movimento. I pittori futuristi cercano di dimostrare che la vita è una continua evoluzione, un continuo divenire, e che "non esiste forma, in quanto la forma attiene a ciò che è immobile, mentre la realtà è movimento, ecco perché reale è soltanto il cambiamento continuo di forma" (Boccioni, *Pittura scultura futuriste*, 1912). Simultaneità, frammentazione del reale, dinamismo sono oggi concetti che rientrano nel patrimonio della nostra cultura visiva. Giacomo Balla, recentemente protagonista a New York nelle aste serali di Christie's e Sotheby's con un record milionario per il dipinto "Automobili in corsa" del 1912, ha dedicato uno studio pignolo, metodico e razionale, proprio come quello di uno scienziato naturalista, alle macchine transanti in via Veneto a Roma. Le automobili nei suoi studi, studi che per certi versi (come nell'opera riprodotta in figura 1) sono riconducibili ai disegni leonardeschi, diventano sempre più astratte fino a scomparire definitivamente per lasciare il posto alle linee di fuga del movimento di penetrazione nello spazio del mezzo. Il metodo è quello dell'indagine e della sperimentazione, finalizzata a dar forma pittorica (o scultorea, in altri casi) ad un concetto, la

mutevolezza della forma del mondo, offrendo al linguaggio stesso soluzioni estetiche che diventano poi, con il tempo, patrimonio consolidato della nostra cultura visiva. Se il cubismo ci spinge a considerare la visione di una scena da molti punti di vista contemporaneamente, suggerendoci che non esiste una realtà, ma una molteplicità di realtà che dipendono solo dai punti di osservazione, il futurismo - in modo più avanzato - ci fa osservare che la realtà è in costante mutamento ed è descrivibile solo come processo dinamico nello spazio-tempo, come sistema complesso costituito da una frammentazione di immagini che rappresentano la granularità stessa di questa complessità. Balla e Depero nel 1915 firmano il famoso manifesto "Ricostruzione futurista dell'Universo", che fornisce anche il nome alla mostra del Guggenheim, e si spingono ancora più avanti. Immaginando complessi plastici da mettere in moto che girano su un perno, su più perni, complessi plastici che si scompongono, a volumi, a strati, per trasformazioni successive, gettano le basi di un'arte che abbandona definitivamente la forma tradizionale del quadro per smaterializzarsi nello spazio. *"Daremo scheletro e carne all'invisibile, all'impalpabile, all'imponderabile, all'impercettibile"*, scrivono i firmatari del manifesto. Purtroppo un difetto noto del futurismo sta proprio nella contraddizione tra una elevata capacità teorica e una incapacità, altrettanto elevata, di dar forma visiva alle proprie teorizzazioni. Questo è il motivo principale per il quale molte delle idee futuriste trovano applicazione concreta solo nei decenni successivi, spesso ad esperienza futurista conclusa, e proprio grazie a quegli artisti che nel futurismo si sono formati e sono maturati. È un po' come se il futurismo fosse stato parzialmente più teorico che applicativo. Ma come si può ben immaginare prima dell'azione viene il tempo della teoria e della progettazione.

Arte Totale (anni '30)

Durante i primi decenni del secolo XX si affermano le grandi trasformazioni urbane, tecnologiche e industriali iniziate già verso la fine del secolo precedente (aereo, automobile, cinema, elettricità, radio). Nello stesso periodo nasce anche la matematica applicata moderna, che poi si svilup-

perà pienamente con l'arrivo della seconda guerra mondiale, quando si occuperà di codici segreti, calcoli per la bomba atomica, radar, calcolatori, teoria dell'informazione. Il futurismo risponde a queste sollecitazioni con un'arte totale, attraverso scomposizioni tipografiche di parole in libertà (Marinetti), scomposizioni pittoriche di velocità e dinamismo (Balla), o di forma (Boccioni), con foto dinamiche e cinema sperimentale (Bragaglia), con arte meccanica e complessi plastici (Balla, Depero, Prampolini), con visioni sintetiche di aeropittura, con l'architettura in vetro, cemento e ferro delle grandi stazioni (Sant'Elia), con la sintesi rumoristica (Russolo) e molto altro ancora. Il futurismo cerca di descrivere la nuova realtà con un linguaggio che gli si addice, con un approccio sperimentale sviluppato in ogni ambito formale ed estetico (totale, appunto), raggiungendo ogni anfratto della vita quotidiana, occupandosi, tra le tante cose, di progettazione di mobili, di tessuti, di comunicazione pubblicitaria, di grafica editoriale, di giochi per bambini. Questo approccio conduce a delle produzioni limitate di proto-design, ma soprattutto modifica in profondità il ruolo dell'artista moderno, non più alle prese con pennelli, paesaggi e nature morte. Ora l'artista diventa progettista e costruttore, ha metodi di lavoro sperimentali (di materiali, di procedure), procede a stretto contatto con l'industria, entra nelle nuove professioni legate alla comunicazione di massa (riviste, libri, cartellonistica, comunicazione aziendale). Milano è il centro di questa grande trasformazione paradigmatica. Gli spazi espositivi privati svolgono un ruolo importante, come la galleria-libreria del Milione che diffonde le idee razionaliste e neoplastiche che si sviluppano in Europa, importando i libri della Bauhaus, la rivista "Art Concret", riportando informazioni attraverso i "Bollettini". Al Milione espongono in Italia per la prima volta, e con tutte le difficoltà del caso, gli astrattisti ("degenerati" secondo la definizione nazista) Kandinsky, Vordemberge-Gildewart, Albers. Nascono i primi studi professionali di grafica come lo Studio R+M dei giovani futuristi Ricas e Munari, lo Studio Boggeri, le prime riviste come "Campo Grafico" (estetica e tecnica grafica), "Casabella" (architettura), "Domus" (architettura), le grandi occasioni di confronto: Biennali (Venezia), Triennali (Milano) e Quadriennali (Roma). Si gettano i semi



Figura 2: Filippo Tommaso Marinetti, *Il Poema del Vestito di Latte*, Snia Viscosa, (1937). Grafica di Bruno Munari. Collezione Nicoletta Gradella, Brescia.

per un cambio di mentalità nel ruolo dell'artista all'interno di una società industriale avanzata: attivo, tecnologico e sperimentale. La rivoluzione futurista di un'arte totale si realizza in pieno, soverchiando la presunta e fallimentare rivoluzione politica cercata da Marinetti e trasfigurata dal fascismo, che del futurismo è fortemente debitore. La lezione da trarre è quella di una superiorità dell'estetica e della cultura, nei processi di trasformazione della realtà e dell'uomo, sulla politica.

Il Movimento Arte Concreta (1948)

In Italia, alla fine della seconda guerra mondiale, il dibattito tra gli artisti provoca una scissione, determinata anche dalla presa di posizione ufficiale del segretario del Partito Comunista Italiano Palmiro Togliatti contro le opere non allineate allo stile del realismo socialista, tra due gruppi contrapposti di "realisti" ed "astrattisti". Ma la vera novità storica è rappresentata dal diffondersi delle idee del movimento di arte concreta

sviluppate in Europa da autori come Theo Van Doesburg, Jean Arp, Piet Mondrian. Questa corrente artistica si afferma in Italia con la prima mostra del M.A.C. (Movimento Arte Concreta, fondato a Milano da Bruno Munari, Gillo Dorfles, Gianni Monnet e Atanasio Soldati), che si tiene alla libreria (nonchè galleria) Salto di Milano nel dicembre del 1948 con la presentazione (non casuale) di una cartella di 12 serigrafie dei membri del gruppo. L'intento dei concretisti è quello di svincolare l'arte da ogni rappresentazione che non sia la visualizzazione estetica di pensieri astratti, fornendo al linguaggio visivo artistico quella autonomia speculativa che lo rende indipendente da ogni esigenza rappresentativa o neorealista. In questo senso il fatto che la mostra fondante del gruppo milanese sia centrata sulla presentazione di una cartella serigrafica la dice lunga sul fatto che il gruppo dia importanza al contenuto della comunicazione visiva, non tanto ai mezzi utilizzati per realizzarlo (un dipinto unico fatto a mano ha per loro lo stesso valore comunicativo di una stampa seriale). Il M.A.C. conta sull'appoggio di alcuni padri storici del futurismo italiano, come Giacomo Balla ed Enrico Prampolini, mentre tra i fondatori troviamo quella figura di cerniera, tra la cultura futurista e le rinnovate esigenze sperimentali, rappresentata dal poliedrico Bruno Munari. Dunque vi è un cambio di paradigma e di metodo, l'arte mostra una dipendenza sempre più stretta dalle tecnologie industriali (polimeri, neon, acciaio inox, ecc.) e attraverso una "sintesi delle arti", cioè una stretta collaborazione tra architetti, artisti, grafici, si orienta sempre di più verso il design. E' proprio Munari, per il quale l'arte è essenzialmente sperimentazione creativa, a scrivere sulla rivista AZ nel numero di aprile-maggio del 1950 un articolo dal titolo "Arte e Industria" il cui incipit è il seguente: "La vera funzione sociale dell'arte dovrebbe essere quella di migliorare non solo l'animo umano ma anche l'ambiente dove l'uomo vive. Il pubblico passa per strade grigie e monotone, viaggia su brutti veicoli, i suoi occhi si posano continuamente su una maggioranza di pubblicità volgare dai colori stonati. Perché non portiamo un poco della nostra sensibilità artistica all'industria e ai suoi prodotti? Credete che un apparecchio telefonico non possa essere anche bello oltre che pratico?" Come non ricondurre ai nostri tempi quest'ultima affermazione correlandola al-



Figura 3: Bruno Munari. *Composizione con materiali vari in vetrino per proiezione multifocale*. (1952). Collezione Fondazione Jacqueline Vodoz e Bruno Danese, Milano, Fotografie di Roberto Marossi. (A sinistra: variante dell'immagine proiettata; al centro: vetrino; a destra: vetrino visto di lato.)

la storia vincente dell'iPhone di Steve Jobs? Il merito del M.A.C. consiste proprio nell'aver gettato le fondamenta per l'affermazione dello stile italiano, attraverso la sperimentazione dei materiali, attraverso una cultura capace di incrociare la domanda industriale di apertura verso le forme innovative del design. Oltre alle mostre di arte concreta più importanti (di Milano, Parigi, Zurigo, Roma) vanno ricordate le mostre alla Saletta dell'Elicottero della Galleria dell'Annunciata a Milano nel 1952 dal titolo *"Materie plastiche in forme concrete"*, *"Studi per forme concrete nell'industria motociclistica"*, e a Torino nel 1954 al Salone dell'automobile *"Colore per le carrozzerie d'auto"*. Sono piccole mostre che danno il senso di un'atmosfera positiva e collaborativa, di raccordo tra produzione ed estetica. Sempre in questi anni, tra la fine degli anni '40 e i primi anni '50, si realizza una trasformazione cruciale nell'arte italiana rappresentata da alcuni eventi che segnano il passaggio, ampiamente teorizzato dal movimento futurista, dalla terza alla quarta dimensione, per mezzo di ambienti in cui lo spettatore partecipa, fisicamente e percettivamente, in modi più coinvolgenti del passato, all'opera d'arte. Inizia Bruno Munari con le *"Macchine inutili"* costruite a partire dal 1930 e poi con l'opera *"Concavo-convesso"* del 1946. Le prime sono delle opere d'arte composte da elementi leggeri appesi (bacchette, forme geometriche in cartoncino colorato) che ruotano in libertà e dove la programmazione costruttiva è bilanciata dalla aleatorietà dei movimenti degli elementi, spinti anche dal minimo soffio d'aria. Il secondo invece è un ambiente in semi-oscurità costruito con una rete metallica incurvata che ricorda certi oggetti non orientabili della topologia matematica,

come il noto nastro di Möebius. L'opera illuminata con luci puntiformi determina, con la rifrazione delle ombre sui muri, con il generarsi di immagini in movimento, dei veri e propri film astratti. Sempre a Milano Lucio Fontana realizza nel 1949 alla Galleria del Naviglio a Milano un *"Ambiente spaziale nero"* ottenuto sospendendo in uno spazio totalmente nero delle forme dipinte con colori fluorescenti illuminate da una luce di Wood. Nel 1950 Bruno Munari realizza delle micro-composizioni pittoriche da proiettare in grandi dimensioni sostenendo: *"Il vivere moderno ci ha dato la musica in dischi, ora ci dà la pittura proiettata"* (figura 3). Lucio Fontana realizza poi un *"Ambiente al neon"* nel 1951, ovvero dei segni grafici aerei luminosi per il salone di ingresso della Triennale di Milano. Infine Bruno Munari sfruttando un nuovo materiale prodotto dall'azienda americana Polaroid costruisce nel 1953 le *"Proiezioni polarizzate"*, ovvero delle micro composizioni a colori variabili provocati dalla rotazione del filtro Polaroid posto davanti al proiettore. Le pitture proiettate vengono presentate per la prima volta nell'autunno del 1953 allo Studio B24 di Milano e poi nel 1955 al MoMa di New York. E' come se le intuizioni futuriste avessero subito una improvvisa accelerazione e si fossero trasformate di colpo, nel giro di pochi anni, e dopo una lenta maturazione, in ambienti immersivi e fortemente tecnologici.

La Civiltà delle Macchine (1955)

Ricordiamo ora brevemente il ruolo del poeta ingegnere Leonardo Sinisgalli, chiamato giovanissimo da Fermi a partecipare alle ricerche sulla fissione dell'atomo all'interno del gruppo di



Figura 4: Bruno Munari, *Macchina inutile* (1956-68). Collezione Nicoletta Gradella, Brescia. Fotografia di Pierangelo Parimbelli.

ragazzi di via Panisperna (Roma), quale figura di grandissimo rilievo nazionale per gli impulsi che egli fornisce al connubio tra le arti visive, tecnica e scienza, nello sforzo di trovare un terreno comune tra cultura umanistica, tecnologia e industria [3]. Sullo sfondo di una Milano in cui artisti, tecnici, industriali, architetti vanno a braccetto e sono orientati verso la progettazione, Sinisgalli, dopo le esperienze in Pirelli e Olivetti, diventa direttore, dal 1953 al 1958, della rivista di Finmeccanica "La Civiltà delle Macchine". Sul periodico aziendale, di elegante fattura, Sinisgalli tratta, con taglio divulgativo, temi di fisica, matematica, cibernetica, architettura, filosofia, poesia, arti visive. Il grado di provocazione aperta all'integrazione delle varie discipline è un caso anomalo ed isolato nella cultura italiana del periodo. Così come è un caso anomalo l'ideazione, in collaborazione con il futurista Prampolini, di una mostra alla Galleria Nazionale di Arte Moderna di Roma nel 1955, in cui si confrontano, da un punto di vista estetico, espressivo, macchine (congegni) industriali e opere d'arte, con il titolo: "Le arti plastiche e la civiltà meccanica". Le macchine sono fornite da aziende come Finmeccanica, Alfa Romeo, Delta, Ansaldo, Pirelli, Cantieri dell'Adriatico. Tra gli artisti partecipanti vi sono: Arp, Bloc, Corpora, Munari, Perilli, Prampolini, Radice, Reggiani, Severini, Veronesi e molti altri. In questa strana esposizione convivono le invenzioni degli ingegneri e le creazioni interessate degli artisti, tra cui un esemplare di "macchina inutile" di Bruno Munari, un'opera d'arte cinetica che è forse la concretizzazione visiva più riuscita

di una poetica che sta esattamente all'incrocio tra arte, scienza e tecnica, arricchita da componenti come il paradosso, la casualità, la leggerezza, la spazialità.

L'arte programmata (1962)

Olivetti è la prima grande azienda italiana che in qualità di committente sponsorizza una mostra d'arte, sulla base della convinzione che la produzione industriale è il centro dell'azione dell'uomo e che l'arte deve raccontare il valore dell'innovazione. Nel maggio del 1962 nel negozio Olivetti di Milano, in galleria Vittorio Emanuele, a due passi dalla Scala, si apre la fortunata mostra di "Arte Programmata" che vede animatore principale Bruno Munari con la collaborazione di Umberto Eco e la presenza dei gruppi artistici Gruppo T di Milano, Gruppo N di Padova (i cui partecipanti firmano le opere in modo collettivo, dotandosi di una metodologia di lavoro tipica della ricerca scientifica) assieme ad alcune figure isolate come Enzo Mari e, successivamente, Getulio Alviani [4]. La mostra, che porta i sottotitoli di "arte cinetica", "opere moltiplicate" e "opera aperta", verrà poi spostata a Venezia nel negozio Olivetti allestito dall'architetto Carlo Scarpa, per transitare successivamente a Roma, Dusseldorf, Londra e in vari musei americani. L'esposizione è composta da opere in cui le componenti vengono messe in movimento da motori elettrici. La programmazione, basata su un equilibrio instabile tra regole ed elementi di casualità dovuti al tipo di materiale utilizzato ed alle forze fisiche

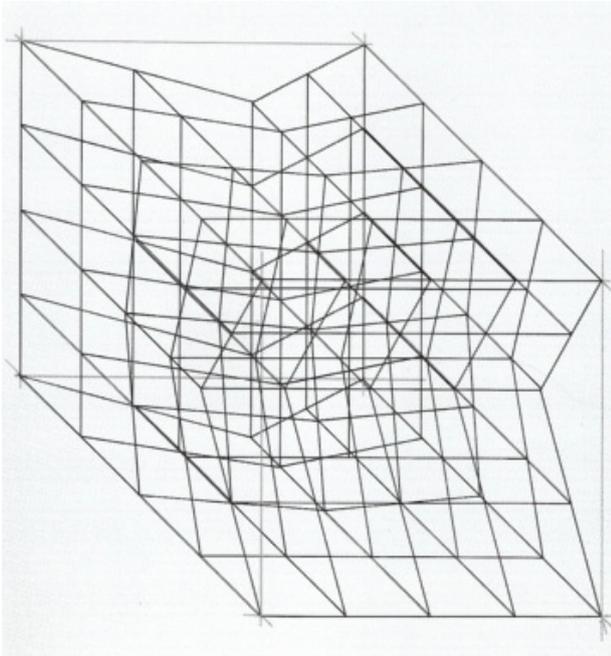


Figura 5: Gianni Colombo. *Progetto di Spazio elastico. Graz (1967)*. Per gentile concessione dell'Archivio Gianni Colombo, Milano.

in gioco (l'attrito, la gravità, la forza elastica, la forza magnetica, la dinamica dei fluidi), provoca infinite permutazioni al punto che l'opera d'arte si caratterizza proprio per una forma indefinita. La denominazione di "arte programmata" è accreditata al curatore-artista Bruno Munari ed ha ancora dei legami con i principi futuristi del movimento e del dinamismo. Gli artisti stabiliscono delle regole e studiano le possibili varianti. Gianni Colombo, esponente di punta del gruppo milanese T (la lettera indica la dimensione "Tempo" che entra in gioco in ogni composizione mostrando la variazione temporale di una forma), presenta una "strutturazione fluida" costruita con un lungo nastro di materiale plastico trasparente, racchiuso dentro una scatola di cristallo, che mosso da un motore genera per accumulazioni casuali immagini sempre diverse. Davide Boriani inserisce in un disco con scomparti irregolari della polvere di ferro che viene spostata da calamite poste sul retro del disco azionate da motori elettrici. Munari crea "nove sfere in colonna" che contengono dei segni bianchi. La prima sfera in basso è posta su una puleggia azionata da un motore che muove casualmente e per attrito tutte le sfere incolonnate dentro una struttura in cristallo essenziale e minimale. Il gruppo N di Padova presenta una serie di opere

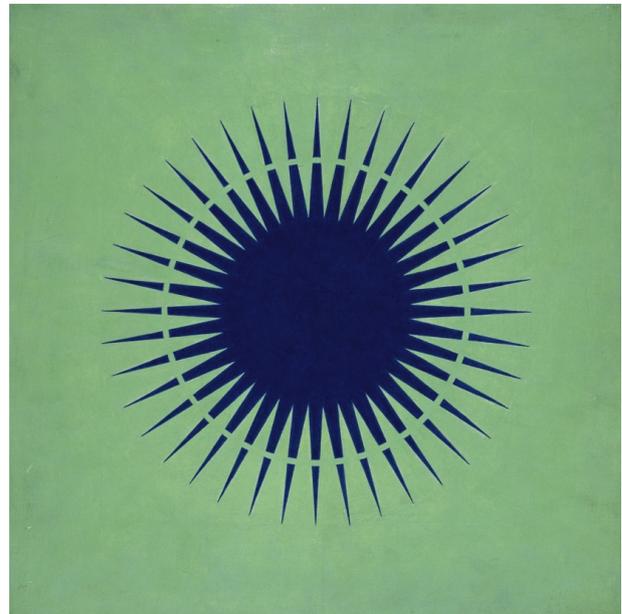


Figura 6: Mario Ballocco. *Struttura stellare blu con cerchio indotto. (1969)* Collezione privata.

tra cui una "visione dinamica" in cui si ha un movimento apparente di colore determinato non da un motore, ma dal movimento dello spettatore stesso. Questa esposizione fornisce un grande impulso alle ricerche cinetiche e percettive che indagano ogni aspetto della nostra visione e che mettono al centro dell'opera l'interazione con lo spettatore attraverso una forma "aperta", instabile, filosoficamente ed esteticamente interprete di una realtà in continua e rapida evoluzione. Da un punto di vista compositivo gli artisti traggono spunti dalla fisica e dalle matematiche, le opere nascono da una progettazione e la loro realizzazione richiede competenze tecniche non più artigianali ma industriali.

L'arte come ricerca (anni '60)

Milano, quale centro industriale e del terziario avanzato più importante in Italia, diventa il laboratorio di grandi innovazioni. Gli artisti interpretano la modernità con l'essenzialità, attraverso la deformazione delle superfici pittoriche: monocrome o vibratili, ritmiche e minimali, estroflesse, stratificate, fresate, sagomate, tagliate, forate, retroilluminate. Ricordiamo i nomi di alcuni tra i giovani protagonisti che in quegli anni hanno come maestri e *tutor* la coppia, quasi complementare, composta dallo "spazialista" Lucio Fontana e "il perfettissimo" Bruno Munari: Piero Man-

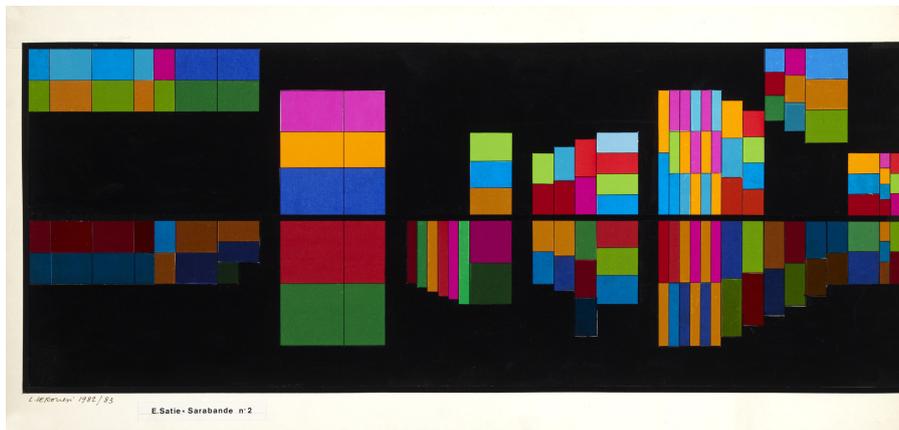


Figura 7: Luigi Veronesi. Visualizzazione cromatica delle misure dalla 1 alla 14 della Seconda Sarabanda per pianoforte di Erik Satie (1982-83). Collezione Archivio Luigi Veronesi, Milano.

zoni, Agostino Bonalumi, Getulio Alviani, Paolo Scheggi, Enrico Castellani, Enzo Mari e molti altri. Gianni Colombo sviluppa un'idea ambientale di arte modificabile, realizzata attraverso strutture geometriche delimitate da corde elastiche fosforescenti tese in ambienti bui ("spazi elastici", ambienti visuo-cine-estetici programmati, figura 5) [5]. L'artista Mario Ballocco, che inaugura all'Accademia di Brera l'insegnamento della "cromatologia", indaga in modo scientifico tutti gli aspetti della percezione, inclusi gli "inganni" della visione, al punto da concepire un quadro come concretizzazione e verifica percettiva di un dato risultato sperimentale (figura 6) [6]. Luigi Veronesi prosegue le proprie ricerche nell'ambito della fotografia e del cinema sperimentale e si occupa di indagare le visualizzazioni cromatiche della musica (figura 7) [7], mentre Munari usa la macchina fotocopiatrice Rank Xerox per sperimentare con la luce [8].

L'arte come progetto (anni '60-'70)

Nel dicembre del 1959 l'artista Daniel Spoerri presenta a Parigi i multipli delle Edizioni MAT (Multiplication d'Art Transformable), con opere cinetiche di Albers, Duchamp, Malina, Munari, Soto, Tinguely, Vasarely ed altri [9]. Nel febbraio del 1960 la mostra arriva a Milano, allo spazio Danese (lo showroom di una delle più innovative aziende milanesi di *industrial design*) e costituisce la prima esposizione in Italia dedicata al tema dell'opera d'arte prodotta in serie (multipli) a partire da un progetto, non come riproduzione seriale di un pezzo unico originale. Spoerri sottolinea quanto sia importante che i multipli abbiano delle caratteristiche cinetiche, di movimento: "Anche se i proprietari di un'opera cinetica

prodotta in un'edizione di cento copie hanno pezzi che rispondono alle stesse qualità specifiche, in realtà posseggono un'opera unica perché, in sé stessa, cambia continuamente". L'arte entra in produzione in tiratura limitata o anche illimitata, in base alle esigenze della comunicazione. E' un tentativo di demitizzare l'artista-star, di avvicinare il mondo dell'arte e della creazione estetica al singolo individuo e alla sua vita quotidiana. "Solo chi confonde il prezzo col valore non dà interesse a queste produzioni" (Munari). L'idea di una ricerca continua si scontra anche con i meccanismi fondanti del mercato dell'arte, centrato sulla costruzione del valore attorno ad opere facilmente riconoscibili (ostentate come *status symbol*). Poi arrivano, verso la fine degli anni '70 il terrorismo ed i cosiddetti "anni di piombo". Il design, dopo un periodo di forte ideologizzazione della realtà, diventa assurdamente iper-decorativo ed anti-funzionale, l'arte mette in scena materiali poveri e una certa banalità di idee, la pittura ritorna al quadro e diventa rassicurante (per il mercato) e citazionista. Insomma si attua un vero passo all'indietro rispetto alle istanze sperimentali che abbiamo brevemente citato [10].

L'arte ha un futuro?

Bruno Munari una volta scrisse: "quando tutto è arte, niente è più arte". Una constatazione del fatto che quando regna la confusione più totale è difficile poi mettere le cose al loro posto, in ordine. Fuori dalla metafora potremmo dire che questo paradosso è oggi ancora attuale. L'innumerabile quantità, statisticamente parlando, di artisti operanti nei principali centri dell'arte contemporanea ci spinge a supporre che molti di que-

sti, verosimilmente in una percentuale assai alta, non troveranno posto sui libri di storia dell'arte, nonostante la loro pervasiva presenza mediatica. Eppure, in conflitto con il ragionevole assunto statistico appena accennato, il numero degli artisti professionisti è in costante crescita. Ma il lavoro di tutti questi artisti è davvero utile? E a cosa serve realmente oggi l'arte, al di là di ogni ragionevole considerazione sulla funzione di bene di lusso, di rifugio o di investimento ad alta volatilità? L'arte ha senso se questa serve a stimolare la fantasia e la creatività, se aiuta gli individui a formarsi e a realizzarsi secondo la propria natura, se spiega se stessa e se svela i suoi processi, affinché possa essere replicata e compresa all'interno di un processo di apprendimento che si realizza nel momento stesso del fare, seguendo le regole del gioco. Un'arte "esatta", razionale, in grado di alimentare, evolvere ed espandere in modo continuativo le capacità espressive di ogni linguaggio e cultura visiva. Crediamo si possa interpretare l'arte come uno strumento per raggiungere obiettivi che vanno al di là dei confini stessi dell'opera d'arte, che dunque non può elevarsi a scopo finale di ogni agire artistico. Appunto: l'arte è il mezzo, non il fine.



- [1] VIVIEN GREEN (A CURA DI): *Italian Futurism, 1909-1944: Reconstructing the Universe*, Guggenheim, New York (2014).
- [2] PONTUS HULTEN (A CURA DI): *Futurismo & futurismi*, Bompiani, Milano (1986).
- [3] SEBASTIANO MARTELLI, FRANCO VITELLI (A CURA DI): *Il gusto della chiocciola. Studi su Leonardo Sinigalli*, Edisud Salerno, Salerno (2012).

- [4] MARCO MENEGUZZO, ENRICO MORTEO, ALBERTO SAIBENE: *Programmare l'arte. Olivetti e le neoavanguardie cinetiche*, Johan & Levi Editore, Milano (2012).
- [5] CAROLYN CHRISTOV-BAKARGIEV, MARCELLA BECCARIA: *Gianni Colombo*, Skira, Milano (2010).
- [6] PAOLO BOLPAGNI (A CURA DI): *Mario Ballocco*, Silvana Editoriale, Milano (2009).
- [7] PAOLO BOLPAGNI, ANDREINA DI BRINO, CHIARA SAVETTIERI (A CURA DI): *Ritmi visivi. Luigi Veronesi nell'astrattismo europeo*, Edizioni Fondazione Raggianti, Lucca (2011).
- [8] BRUNO MUNARI: *L'arte come mestiere*, Laterza, Bari (1966).
- [9] GUIDO BALLO: *La mano e la macchina. Dalla serialità artigianale ai multipli*, Sperling & Kupfer, Milano (1976).
- [10] GIAMPIERO BOSONI: *Il Modo Italiano. Design e avanguardia nel XX secolo*, Skira, Milano (2007).



Luca Zaffarano: Laureato in informatica, consulente IT in ambito finanziario nell'area del market risk management, collezionista di arte italiana, è autore di un progetto documentale dedicato all'arte di Bruno Munari consultabile all'indirizzo <http://www.munart.org>. Nel 2012 ha ideato ed è stato co-curatore della mostra personale "Bruno Munari. My Futurist Past" alla Estorick Collection of Modern Italian Art di Londra. Ha pubblicato scritti in: Sonia e Massimo Cirulli (a cura di), *Lo Stile Italiano*, Silvana Ed. 2011; M. Hájek, L. Zaffarano, *Bruno Munari. My Futurist Past*, Silvana Ed. 2012. Attualmente sta lavorando, in collaborazione con alcuni studiosi delle Università di Cambridge (UK) e UCL London, ad un libro dedicato alla poetica dell'artista Bruno Munari.

Un matematico passeggia per l'Alhambra

Rafael Pérez Gómez

Departamento de Matemática Aplicada, Universidad de Granada, Granada, Spagna

La búsqueda y recreación de la belleza del mundo ha sido una tarea incansable de todas las culturas aparecidas en nuestro mundo que se plasma en la Historia del Arte. Para lograrlo, han sido necesarios múltiples análisis, hechos en contextos muy diversos y variados, que han dado lugar a modelos teóricos que han permitido reflejarla en casos singulares. La belleza de la Alhambra de Granada, fruto de la cultura nazarí, nace de la Geometría con la que han sido creadas sus formas mostrando, de forma completamente abstracta, simbólica, la Unidad (de Allah) entre la multiplicidad que hoy interpretamos mediante la teoría matemática de Grupos Cristallográficos Planos.

La Alhambra de Granada fue declarada Patrimonio de la Humanidad en el año 1984. La justificación, breve, que figura en el correspondiente expediente de catalogación dice que: *El bien incluye logros artísticos únicos. Es un testimonio excepcional de la España musulmana del siglo XIV. Ofrece un ejemplo valioso de las residencias árabes del medioevo.*

Quien conoce la Alhambra sabe bien que es única, excepcional y paradigmática. Hay pocos

La ricerca e la ri-produzione della bellezza del mondo è stato uno scopo incessante di tutte le culture apparse nel nostro mondo che si plasma attorno alla Storia dell'Arte. Per ottenere questi risultati sono state necessarie analisi multiple, fatte in contesti molto differenti e vari, che hanno prodotto modelli teorici che hanno permesso di riflettersi in casi singolari. La bellezza dell'Alhambra di Granada, frutto della cultura nazarí, nasce dalla Geometria con la quale sono state create le sue forme mostrando in forma completamente astratta, simbolica, l'Unità (di Allah) nella molteplicità che oggi interpretiamo mediante la teoria matematica dei Gruppi Cristallografici Piani.

L'Alhambra di Granada fu dichiarata Patrimonio dell'umanità nell'anno 1984. La motivazione, breve, che appare nella pratica di catalogazione dice che: *Il bene incorpora apici artistici unici. È una testimonianza eccezionale della Spagna mussulmana del secolo XIV. Offre un valido esempio delle residenze arabe del medioevo.*

Chi conosce l'Alhambra sa bene che è unica, eccezionale e paradigmatica. Ci sono pochi luo-

lugares en los que pueda apreciarse el color en el aire, el murmullo de los pensamientos de un pueblo culto o los continuos quiebros a la interpretación de los sentidos. Desde este punto de vista se entiende fácilmente su carácter universal como obra de arte: utiliza todos los lenguajes posibles para que cualquiera pueda comunicarse con sus creadores a través del monumento. Es bien conocido que las paredes de la Alhambra hablan, que forman las páginas de un gran libro que trata, fundamentalmente, sobre Historia, Religión, Sociología, Ciencia y Tecnología y que está escrito haciendo uso de la Poesía y la Geometría como herramientas de la Arquitectura que, a su vez, utiliza el Amor como el mejor de los materiales de construcción de todas las épocas. *Cómo un edificio hecho con cuatro palitroques está aún en pie?* El Dr. García Gómez, uno de los mejores arabistas que hemos tenido, hace esta pregunta en su libro "Poemas en los muros y fuentes de la Alhambra"; a su vez, D. Emilio respondía: ha sido gracias al amor, porque la Alhambra ha sido siempre vivida, querida y, por tanto, conservada.

Desde hace muy poco tiempo, la Alhambra ha cobrado un inusitado interés en el mundo de las Matemáticas. Se sabe desde hace tiempo que la belleza de sus trazados obedece a la aplicación de una estética basada en el uso de proporciones tanto pitagóricas como áureas. El mejor ejemplo de esta última lo tenemos en la fachada del palacio de Comares que está diseñada milimétricamente haciendo uso de rectángulos áureos y sus correspondientes recíprocos internos.

Sin embargo, aún siendo todo esto de gran interés, el estudio de su simetría uni y bidimensional es del máximo interés. Desde que H. Weyl escribiera su libro sobre Simetría, se abre una historia para mí apasionante. Dedicué una temporada de mi vida a investigar qué Matemáticas servían para crear ese gran caleidoscopio de color llamado Alhambra. Como se verá más adelante, tuve la fortuna de dar con ciertas claves utilizadas en sus mosaicos, claves muy sencillas que se ocultan magistralmente y que hacen buena la afirmación de Ortega de que el arte es prestidigitación sublime y genial transformismo. Además, demostré que fueron agotadas empíricamente todas las combinaciones básicas posibles para componerlos. Todo esto hace que la Alhambra tenga hoy ese especial interés para los matemáticos al que

ghi nei quali si può apprezzare il colore nell'aria, il mormorio dei pensieri di un popolo colto o le continue finte all'interpretazione dei sentimenti. Da questo punto di vista si percepisce facilmente il suo carattere universale come opera d'arte: usa tutti i linguaggi possibili perché chiunque possa comunicare con i suoi creatori attraverso il monumento. È noto che le pareti dell'Alhambra parlano, che formano le pagine di un grande libro che tratta, fondamentalemente, di storia, religione, sociologia, scienza e tecnologia, e che è scritto facendo uso della poesia e della geometria come strumenti dell'Architettura che, a sua volta, usa l'Amore come il miglior materiale di costruzione di ogni epoca. *Com'è possibile che un edificio fatto con quattro paletti sia ancora in piedi?* Il Dr. García Gómez, uno dei migliori arabisti che abbiamo avuto, fa questa domanda nel suo libro "Poemas en los muros y fuentes de la Alhambra"; a sua volta, Don Emilio (García Gómez) rispondeva: è stato grazie all'amore, perché l'Alhambra è sempre stata vissuta, amata e, quindi, conservata.

Da molto poco tempo, l'Alhambra ha guadagnato un inusitato interesse nel mondo della Matematica. Si sa da tempo che la bellezza dei suoi tratti obbedisce all'applicazione di un'estetica basata sull'uso delle proporzioni tanto pitagoriche che auree. Il miglior esempio di queste ultime l'abbiamo nella facciata del palazzo de Comares che è disegnata millimetricamente facendo uso di rettangoli aurei e dei suoi corrispondenti reciproci interni.

Ad ogni modo, pur essendo tutto questo di grande interesse, lo studio della sua simmetria uni- e bi-dimensionale è del massimo interesse. Fin da quando H. Weyl scrisse il suo libro sulla Simmetria, si apre una storia che ritengo appassionante. Ho dedicato una parte della mia vita a studiare quali matematiche servissero per creare questo grande caleidoscopio di colori chiamato Alhambra. Come vedremo più in là, ho avuto la fortuna di trovare certe chiavi utilizzate nei suoi mosaici, chiavi molto semplici che si nascondono magistralmente e che confermano l'affermazione di Ortega che l'arte è gioco di prestigio sublime e trasformismo geniale. Inoltre, ho dimostrato che furono esaurite empiricamente tutte le combinazioni possibili per comporli. Tutto questo fa che l'Alhambra abbia oggi per i

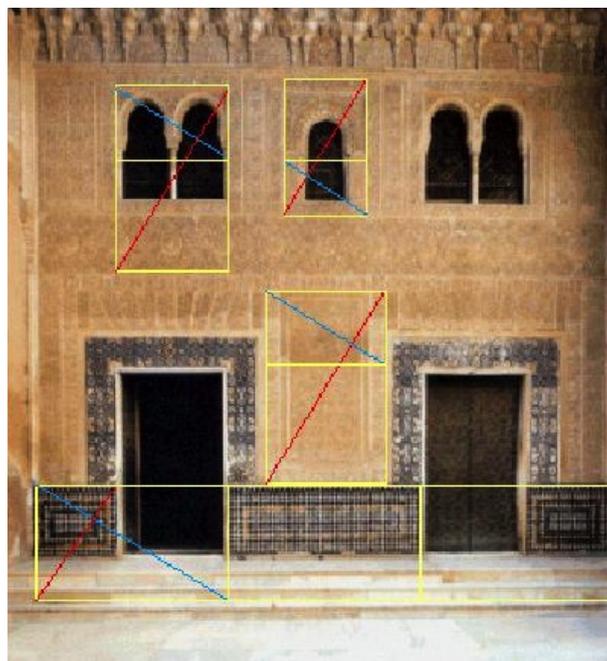


Figura 1: Alhambra, facciata del Palazzo de Comares.

antes me refería, ya que los tracistas andalusíes-granadinos pusieron de manifiesto con su trabajo una forma de abordar el trabajo científico mediante la cual indios y arábigo-parlantes desarrollaron la ciencia hasta la Edad Media: la búsqueda de nuevas ideas desde el ejercicio libre y audaz del método heurístico. Aquellos tracistas granadinos fueron capaces de desarrollar, en los mosaicos de la Alhambra, las 17 posibilidades que hoy conocemos desde el descubrimiento de los rayos X y la Teoría de Grupos Cristallográficos Planos. Es más, la Alhambra es, actualmente, el único monumento construido antes del descubrimiento de la teoría de grupos que cuenta con al menos un ejemplo de cada uno de los grupos cristallográficos planos.

17 Grupos cristallográficos planos

Es cierto que cada mosaico podría definir una página para un tratado de Geometría con regla y compás (con juegos de cartabones y compases rígidos [1], para ser más precisos). Mas no es esta la intención de estos párrafos. Cuando admiramos un astrolabio, un manuscrito, una herramienta propia de la Medicina o de la Ingeniería, valoramos el nivel de conocimiento de

matemáticos l'interesse speciale per i matematici al quale mi riferivo, poiché i progettisti andalusí-granadini resero manifesta con il loro lavoro una forma di affrontare il lavoro scientifico mediante la quale indiani e arabo-parlanti svilupparono la scienza fino al medio evo: la ricerca di nuove idee e pratiche dall'esercizio libero e audace del metodo euristico. Quei progettisti granadini furono capaci di sviluppare nei mosaici dell'Alhambra le 17 possibilità che oggi conosciamo dalla scoperta dei raggi X e la teoria dei gruppi cristallografici piani. Di più, l'Alhambra è, attualmente, il solo monumento costruito prima della scoperta della teoria dei gruppi che possiede almeno un esempio di ognuno dei gruppi cristallografici piani.

17 Gruppi cristallografici piani

Sicuramente ogni mosaico potrebbe essere usato come la pagina di un trattato di geometria da farsi con righello e compasso [1] (con riga e compassi rigidi per essere più preciso). Non è questa l'intenzione di questi paragrafi. Quando ammiriamo un astrolabio, un manoscritto, uno strumento di medicina o dell'ingegneria, apprezziamo il livello di conoscenza dei suoi creatori e

sus creadores y nos admiramos por ello, pero no serían los objetos más adecuados, pongamos por caso, para explicar teorías sobre cirugía con rayos láser. Los mosaicos de la Alhambra de Granada constituyen, hoy en día, la mejor, y más valiosa, colección de ejemplos de la moderna teoría matemática de Grupos, de ahí que constituyan un auténtico legado cultural con valor científico actual.

Es normal encontrar en libros de Historia de la Ciencia afirmaciones acerca de que la producción en lengua árabe en el campo de la fundamentación científica fue escasa. En particular, en cuanto a Matemáticas se refiere, se dice que se limitaron casi exclusivamente a traducir obras griegas e indias que fueron introducidas en Europa, algunas a través de al-Andalus. Sin embargo, ante la evidencia de los manuscritos que se han ido descubriendo, hay que aceptar que su aportación al conocimiento matemático no se limitó a la mera labor de puente. Es cierto que sus mayores y más conocidas aportaciones de carácter matemático giran alrededor de la Aritmética y del álgebra y, dentro de esta última, de la Trigonometría. Las razones para que se produjese tal situación estriban en que el hecho investigador no es un fenómeno aislado como algunos pueden creer, sino que está fuertemente condicionado por la sociedad en la cual se desarrolla. Desde este punto de vista, en la cultura árabe, en general, y en la de al-Andalus, en particular, los asuntos relacionados con la organización sociopolítica y religiosa marcan una línea de desarrollo práctico de la investigación matemática. Este pueblo concebía las Matemáticas como herramienta para otras ciencias -astronomía, astrología, óptica y medicina (a través de la astrología)- y como de utilidad social inmediata. Matemáticas para rezar, para navegar, para repartir herencias, medir tierras, etc. En suma, concebían las Matemáticas como hecho cultural. Lo anterior no significa que no hubiera avances en los fundamentos de las Matemáticas. Ciñéndonos al legado científico andalusí, sabemos [2] que al-Mu'taman, el rey sabio de Zaragoza durante el periodo de los Reinos de Taifas (1031-1086), escribió en su Kitab al-Istikmal ('Libro de la Perfección') una demostración del teorema atribuido casi siete siglos después al italiano Giovanni Ceva (1648-1738); que Qalasaki (s. XV), o Alcasavi o el Bastí según diferentes

lo admiramos por questo, ma non sono gli oggetti più adeguati, ad esempio, per spiegare teorie sulla chirurgia con raggi LASER. I mosaici dell'Alhambra di Granada costituiscono, a tutt'oggi, la migliore, e più completa, collezione di esempi della moderna teoria dei gruppi, e, per questo, costituiscono un autentico retaggio culturale con valore scientifico di attualità.

È normale incontrare nei libri di storia della scienza affermazioni che indicano che la produzione in lingua araba nel campo dei fondamenti della scienza fu scarsa. In particolare, per quanto riguarda la matematica, si dice che gli arabi si limitarono a tradurre opere greche e indiane che furono introdotte in Europa, qualcuna attraverso al-Andalus. Tuttavia, davanti all'evidenza dei manoscritti che sono stati man mano ritrovati, bisogna accettare che il loro contributo alla conoscenza matematica non si limitò al puro lavoro di trasferimento di conoscenza. È certo che i maggiori, e più noti, contributi di carattere matematico riguardano l'aritmetica e l'algebra e, all'interno di quest'ultima, la trigonometria. Le ragioni per le quali si produsse questa situazione sono legate al fatto che l'attività scientifica non è un fenomeno isolato, come qualcuno potrebbe credere, ma è fortemente condizionata dalla società all'interno della quale si sviluppa. Da questo punto di vista, nella cultura araba, in generale, e in quella di al-Andalus, in particolare, le relazioni con l'organizzazione socio-politica e religiosa segnano una linea di sviluppo pragmatico della ricerca matematica. Questo popolo concepiva la matematica come strumento per altre scienze - astronomia, astrologia, ottica e medicina (attraverso l'astrologia) - e per la sua utilità sociale immediata. Matematica per pregare, per navigare, per dividere eredità, misurare terre, ecc.. In sintesi concepivano la matematica come fatto culturale. Questo non significa che non ci furono dei progressi nei fondamenti della matematica. Limitandoci al contributo scientifico andalusí, sappiamo [2] che al-Mu'taman, il re sapiente di Saragozza durante il periodo del Regno di Taifa (1031-1086), scrisse nel suo Kitab al-Istikmal (Libro della perfezione) una dimostrazione del teorema attribuito quasi sette secoli più tardi all'italiano Giovanni Ceva (1648-1738); che Qalasaki (secolo XV), o Alcasavi o il Bastí se-

transcripciones, oriundo de la ciudad granadina de Baza y posible profesor de la Madraza en Granada mandada construir por Yusuf I, introdujo una notación simbólica algebraica similar a la que finalmente fue extendida por la escuela italiana durante el Renacimiento; o el granadino Abulcasim Asbag Abenmohamed (s. X), conocido vulgarmente por el Muhandis ('Geómetra'), que si bien es más conocido por sus trabajos en Astronomía, escribió una obra de comentarios a los Elementos de Euclides en forma de introducción a las Matemáticas; Benabixácar (s. XIII) fue gran conocedor de la ciencia griega, tal y como lo demuestran sus libros sobre la obra de Euclides y las crónicas de Apolonio; etc. No limitándonos en este comentario a al-Andalus, lo concluiremos mencionando a Ibn Sina, al que normalmente se le asigna erróneamente procedencia andalusí, que [2] además de traducir e interpretar como pocos a Aristóteles, analizó la axiomática de Euclides para la Geometría estableciendo su propio conjunto de axiomas. Cabe mayor fundamentación teórica?

Entonces, por qué se siguen haciendo uso de los tópicos antes aludidos? En pleno desarrollo islámico en la Península Ibérica, se destruyeron miles de volúmenes que estaban depositados en importantes bibliotecas. Como ejemplo baste citar la destrucción iniciada por Almanzor de la famosísima Biblioteca de al-Hakam II, en Córdoba, que contaba con más de 400.000 ejemplares, o la quema de libros escritos en árabe ordenada por el Cardenal Cisneros en la plaza de Bibrambla de Granada. Tras la expulsión de moriscos y judíos, la cultura andalusí fue desapareciendo rápidamente de la sociedad medieval de la Península Ibérica. Las Casas de la Sabiduría se tornaron en Bibliotecas de Monasterios y Universidades incipientes que utilizaban el latín como lenguaje científico. Así, esta civilización dejó un devastado legado científico al cual difícilmente puede accederse si no es de la mano de arabistas expertos. En resumen, no es que los andalusíes -al igual que el resto de los científicos del Islam Medieval- abandonasen la investigación en el terreno de los fundamentos, sino que sería más prudente afirmar que realmente no sabemos cuál fue su verdadera producción científica.

Sin embargo, esta civilización que goza de merecida fama por su gran sentido estético y gusto

condo trascrizioni differenti, oriundo della città granadina di Baza e forse professore della Madraza in Granada fatta costruire da Yusuf I, introdusse una notazione simbolica algebraica simile a quella che fu infine estesa dalla scuola italiana durante il rinascimento; o il granadino Abulcasim Asbag Abenmohamed (secolo X), conosciuto vulgarmente come el Muhandis (Geometra), più conosciuto per i suoi lavori in Astronomia, scrisse un'opera di commenti sugli Elementi di Euclide nella forma di introduzione alla matematica; Benabixácar (sec. XIII) fu un grande conoscitore della scienza greca, così come dimostrano i suoi libri sull'opera di Euclide e le cronache di Apollonio; ecc. Senza limitarci in questo commento ad al-Andalus, lo concluderemo menzionando Ibn Sina, al quale si assegna erroneamente origine andalusí, che [2] oltre a tradurre ed interpretare come pochi Aristotele, analizzò l'assiomatica di Euclide per la geometria stabilendone il suo insieme di assiomi. Ci può essere maggiore base teorica?

Quindi perché si continuano ad utilizzare gli stereotipi a cui abbiamo alluso in precedenza? In pieno sviluppo islamico nella penisola Iberica, si distrussero migliaia di volumi che erano conservati in importanti biblioteche. Come esempio, basta citare la distruzione iniziata da Almanzor della famosissima biblioteca di al-Hakam, in Cordoba, che contava più di 400.000 esemplari, o il rogo dei libri scritti in arabo ordinato dal cardinale Cisneros nella piazza Bibrambla di Granada. Con l'espulsione dei mori e degli ebrei, la cultura andalusí scomparve rapidamente dalla società medioevale della penisola Iberica. Le Case della Conoscenza divennero le biblioteche di monasteri e le nascenti università che utilizzavano il latino come linguaggio scientifico. Così questa civiltà lasciò un'eredità scientifica devastata alla quale si può accedere soltanto per il tramite di esperti arabisti. Riassumendo, non è che gli andalusí, così come per il resto degli scienziati dell'Islam medioevale, avessero abbandonato il terreno dei fondamenti, ma sarebbe più prudente affermare che realmente non sappiamo quale fu la loro vera produzione scientifica.

Tuttavia, questa civiltà che gode di meritata fama per il suo grande senso estetico e per il gu-

por los caminos indirectos para presentar sus pensamientos, nos dejó libros escritos sobre los muros de sus monumentos que, por desconocimiento de los conquistadores, nunca fueron objeto de atentado alguno antes más bien de elogio y admiración sin poder intuir la existencia de un tesoro oculto: la expresión de sus creencias en un bello mensaje altamente codificado. No se trata de la poesía epigráfica, sino de la decoración geométrica de la Arquitectura Islámica, en general, y de la Alhambra, en particular. Tal y como hoy sabemos, en la decoración geométrica de este monumento se esconde un tratado sobre el método heurístico de investigación cuyos resultados han necesitado del transcurso de varios siglos para que se produjesen los descubrimientos científicos necesarios que permitan clasificar el conocimiento allí desarrollado intuitivamente. Nunca fueron mejor aplicadas las Matemáticas, en este caso la Geometría, ya que sirvieron para manifestar sus creencias, como a continuación se verá, en un bello alarde de ingenio que se traduce en una creatividad sin precedentes en el diseño de los mosaicos de la Alhambra.

El caso singular que sigue pone de manifiesto una forma de abordar el trabajo científico con la cual indios y árabe-parlantes desarrollaron la ciencia: la búsqueda de nuevas ideas desde el ejercicio libre y audaz del método heurístico. Las justificaciones lógicas vienen a posteriori, una vez que hay algo que fundamentar. En el caso que nos ocupa, estas fundamentaciones tardaron cinco siglos en llegar. Los diseñadores de sus trazados fueron capaces de desarrollar, de modo empírico, todas las posibilidades que hoy nos demuestra la Teoría de Grupos Cristallográficos Planos, elaborada por cristalógrafos a partir del descubrimiento de los rayos-X. En un intento de divulgación, diré que hay 17 Grupos Cristallográficos Planos [3]; son objetos matemáticos abstractos mediante los cuales pueden ser clasificados los mosaicos periódicos planos según su simetría.

Fueron Fedorov, Schoenflies y Barlow quienes descubrieron por separado que en dimensión 3 hay 230 grupos cristallográficos que dan explicación a la estructura de las materias cristalinas. G. Polya y P. Niggli, ya en nuestro siglo, demostraron la existencia de los 17 grupos de isometrías del plano. Desde entonces se ha comenzado

sto di presentare i propri pensieri con percorsi indiretti, ci ha lasciato libri scritti sui muri dei suoi monumenti, che, per ignoranza dei conquistatori, non furono mai oggetto di alcun attentato anzi di elogio e ammirazione senza che si intuisse l'esistenza di un tesoro oculto: l'espressione delle proprie credenze religiose in un bel messaggio altamente codificato. Non si tratta di poesia epigrafica, bensì della decorazione geometrica dell'architettura islamica, in generale, e dell'Alhambra, in particolare. Per quanto ne sappiamo oggi, nella decorazione geometrica di questo monumento si nasconde un trattato sul metodo euristico di studio i cui risultati hanno richiesto il passare di vari secoli perché si producessero le scoperte scientifiche necessarie per permettere di classificare le conoscenze che lì sono state sviluppate intuitivamente. Mai fu applicata meglio la matematica, in questo caso la geometria, che servì per manifestare le proprie credenze religiose, come vedremo in seguito, in una bella ostentazione di ingegno che si traduce in creatività senza precedenti nel disegno dei mosaici dell'Alhambra.

Il caso singolare che segue presenta in maniera chiara un modo di affrontare il lavoro scientifico con il quale indiani e arabo-parlanti svilupparono la scienza: la ricerca di nuove idee usando l'esercizio libero e audace del metodo euristico. Le giustificazioni logiche vengono a posteriori, una volta che si trova qualche cosa che necessita di avere basi fondamentali. Nel caso di cui ci occupiamo, questi fondamenti tardarono cinque secoli prima di arrivare. I disegnatori di questi tracciati furono capaci di sviluppare, in modo empirico, tutte le possibilità che oggi ci dimostra la teoria dei gruppi cristallografici piani, elaborata dai ricercatori di cristallografia a partire dalla scoperta dei raggi X. In termini divulgativi posso affermare che esistono 17 gruppi cristallografici piani [3]; sono oggetti matematici astratti mediante i quali possono essere classificati i mosaici periodici piani secondo la loro simmetria.

Furono Fedorov, Schoenflies y Barlow che, separatamente, scoprirono che in tre dimensioni ci sono 230 gruppi cristallografici che spiegano la struttura della materia cristallina. Già nel nostro secolo, G. Polya e P. Niggli, dimostrarono l'esistenza dei 17 gruppi di isometria nel piano. Da quel momento è cominciata la ricerca, fatta dai

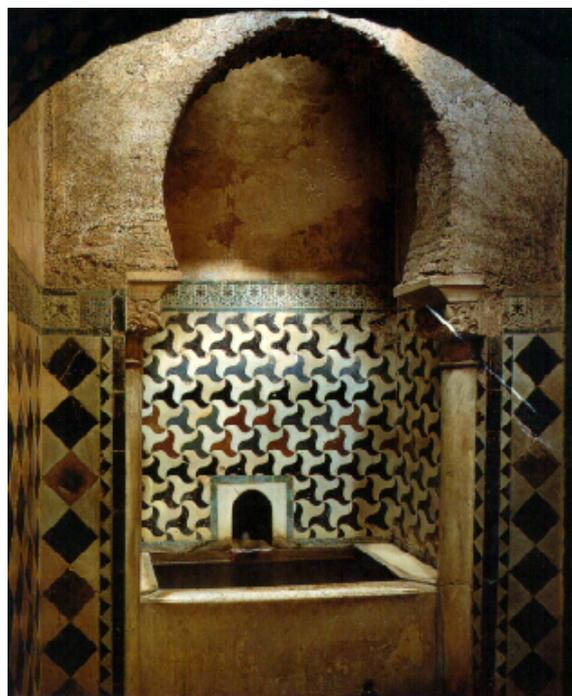


Figura 2: Alhambra, bagno nel Palazzo de Comares.

la búsqueda, hecha por matemáticos, de decoraciones periódicas del plano en obras de arte de determinadas culturas que han destacado en estas realizaciones. La publicación de los resultados obtenidos por unos y otros han dado origen a controversias. Por ejemplo, H. Weyl asegura en su obra *Simetría* [4], considerada como un clásico en el tema, que las 17 posibilidades eran conocidas por los artesanos del viejo Egipto; Fejes-Tóth en *Regular Figures*, [5], asegura que en la Alhambra de Granada hay una representación geométrica de cada uno de los 17 modelos posibles; ... Por contra, B. Grünbaum [6], sostiene que los egipcios sólo utilizaron 12 posibilidades y que los constructores de la Alhambra llegaron a obtener 13 variantes. En 1985 se cerró definitivamente tal discusión publicando los resultados de una labor de búsqueda exhaustiva en el monumento [7] mostrando ejemplos de los grupos ausentes. De este modo se descubre la existencia de una teoría ingenua de grupos en el diseño de decoraciones islámicas medievales ya que, sin conocerse el concepto matemático de grupo, se utilizaron las 17 estructuras básicas posibles para la creación de ciertos diseños simétricos planos, hoy en día interpretados como mosaicos periódicos. A esta etapa la denominamos Prehistoria de la Teoría de Grupos.

matemáticos, di decorazioni periodiche del piano in opere d'arte di varie culture che hanno primeggiato in queste realizzazioni. La pubblicazione dei risultati ottenuti da vari ricercatori ha dato origine a controversie. Per esempio H. Weyl assicura nella sua opera *Simmetria* [4], considerata un classico dell'argomento, che le 17 possibilità erano conosciute dagli artigiani dell'antico Egitto; Fejes-Tóth, in *Regular Figures* [5], assicura che nell'Alhambra di Granada c'è la rappresentazione geometrica di ognuno dei 17 modelli possibili. Al contrario, B. Grünbaum [6], sostiene che gli egiziani utilizzarono soltanto 12 possibilità e che i costruttori dell'Alhambra arrivarono ad ottenere 13 varianti. Nel 1985 si chiuse definitivamente la discussione pubblicando i risultati di un lavoro di ricerca esauriente nel monumento [7] mostrando esempi dei gruppi assenti. In questo modo si scopre l'esistenza di una teoria ingenua dei gruppi nel disegno di decorazioni islamiche medievali dato che, senza conoscere il concetto matematico di gruppo, si usano le 17 strutture di base possibili per la creazione di definiti disegni simmetrici piani, oggi interpretati come mosaici periodici. A questo stadio è stato assegnato il nome di preistoria della teoria dei gruppi.

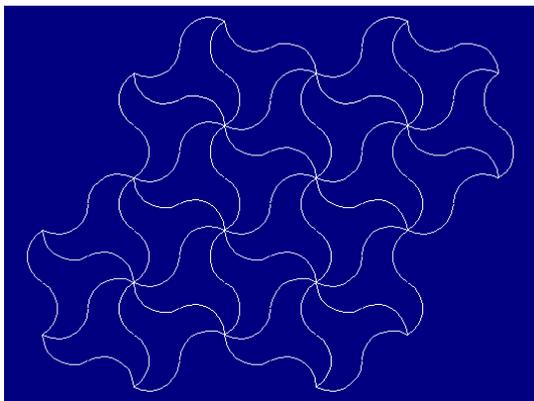


Figura 3

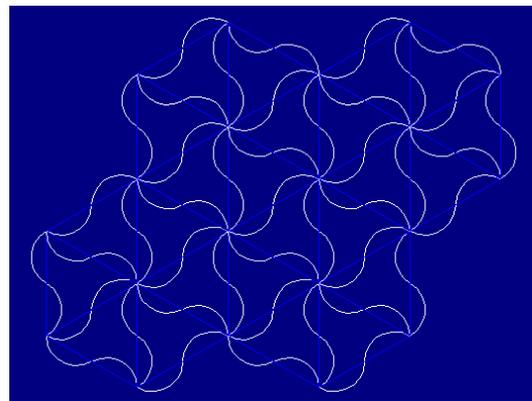


Figura 4

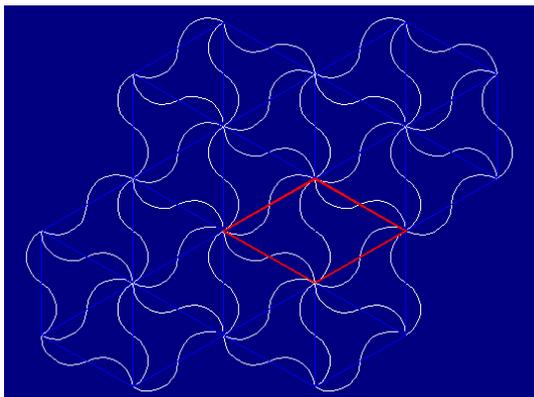


Figura 5

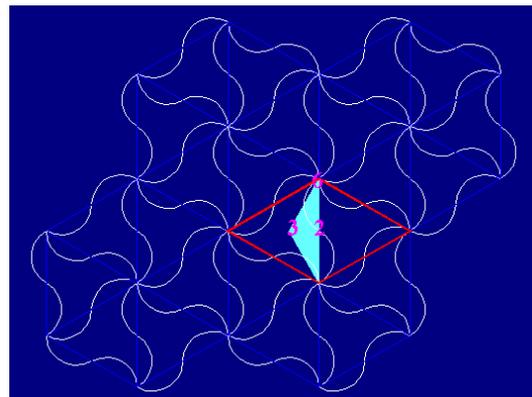


Figura 6

En efecto, en un deseo de manifestar con el lenguaje de la Geometría su creencia en la existencia de la Unidad (Allah) dentro de la multiplicidad, agotaron las estructuras geométricas planas posibles. Esta es la razón primera de la decoración geométrica de la Alhambra. Al existir la prohibición coránica de hacer figuraciones de Allah, se recurre a un lenguaje abstracto de formas geométricas para la decoración tanto en la arquitectura religiosa como en la del poder, clase a la que pertenece la Alhambra. En ella se usa la llamada tesela básica (como unidad) que, utilizando nuestro lenguaje actual, es extendida a todo el plano mediante los elementos de simetría del grupo cristalográfico correspondiente. Haciendo uso del alicatado que se encuentra en los baños del Palacio de Comares, Figura 2, y tomando su diseño básico (es decir, sin tener en cuenta el color), Figura 3, podemos observar que hay una malla de triángulos equiláteros que le dan estructura, Figura 4. Tomando un rombo como el que se muestra en la Figura 5, se puede observar que si lo trasladamos por todo el plano, vamos generando el diseño del alicatado elegi-

In effetti, nel desiderio di manifestare con il linguaggio della geometria il loro credo nell'esistenza dell'Unità (Allah) nella molteplicità, esaurirono le strutture geometriche piane possibili. Questa è la motivazione principale della decorazione geometrica dell'Alhambra. Data la proibizione coranica di rappresentare Allah, si ricorre ad un linguaggio astratto di forme geometriche per la decorazione tanto dell'architettura religiosa quanto come quella del potere, classe alla quale appartiene l'Alhambra. In quest'ultima si usa una tassellatura di base (come unità) che, usando il nostro linguaggio attuale, è estesa a tutto il piano mediante gli elementi di simmetria del gruppo cristallografico corrispondente. Utilizzando la piastrellatura che si incontra nei bagni del Palazzo de Comares (Figura 2), e prendendo il suo disegno di base (questo vuol dire, senza considerare il colore), Figura 3, possiamo osservare che c'è una maglia di triangoli equilateri che formano la struttura, Figura 4. Prendendo un rombo, come quello che di mostra nella Figura 5, si può osservare che, se lo trasliamo per tutto il piano, generiamo il disegno della tassell-

do; es decir, existe un paralelogramo tal que sus lados determinan las direcciones de dos traslaciones independientes con las que se puede generar el mosaico completo; se le llama región generatriz unidad y, en su interior, podemos destacar la existencia de un tesela básica (in triángulo isósceles), ver Figura 6, en la cual se encuentra el diseño mínimo necesario para reproducir el mosaico completo (la multiplicidad) someténdola a las transformaciones del grupo cristalográfico plano del tipo $p6$, cuyos generadores pueden ser dos rotaciones de órdenes 2 y 3, respectivamente, y con distinto centro. Como anécdota, cabe decir que, hasta hoy, la Alhambra es el único monumento en el mundo, construido antes del desarrollo de la Teoría de Grupos, que presenta tal singularidad.

Epílogo

Después de cinco siglos, los mosaicos de la Alhambra sirven para que artistas de la talla de Gaudí, Escher, Robinson, Falla, Debussy, ... que se inspiraron en ella para crear parte de su arte. Están sirviendo también para que los estudiantes de Matemáticas del mundo entero tengan los más bellos ejemplos de objetos tan abstractos como las isometrías, los grupos o las superficies, elementos de Geometría Euclídea, álgebra o Geometría Diferencial, que cobran vida. ¡Impresionante! Pero, ¿qué Alhambra podrían hacer hoy nuestros actuales tracistas. Si pensamos que en la Puerta del Vino hay un botella de Klein, alguien puede pensar que Klein la habrá dejado olvidada; igual sucede con las cintas de Moëbius de las tacas que hay a la entrada del Salón del Trono; pero oír que en los baños de Comares hay una esfera topológica, puede extrañar a muchos. ¿Se imagina estos objetos matemáticos? ¿Tienen interés para la Arquitectura y el Diseño? Con las nuevas geometrías y herramientas de nuestro siglo, mucho puede aportarse al mundo del Arte, en general. Será esta una puerta abierta por la que seguimos avanzando y ofreciendo nuevas visiones de la Alhambra.

Por último, es importante inculcar en nuestro alumnado que desde las Matemáticas se puede acceder a disfrutar de la belleza, ya que, según

latura scelta; questo significa che esiste un parallelogramma tale che i suoi lati determinano le direzioni di due traslazioni indipendenti con le quali si può generare il mosaico completo; si chiama regione generatrice unità e, al suo interno, possiamo identificare l'esistenza di una tassellatura di base (in triangoli isosceli), vedi Figura 6, nella quale si identifica il disegno minimo necessario per riprodurre il mosaico completo (la molteplicità) sottomettendola alla trasformazioni del gruppo cristalografico piano di tipo $p6$, i cui generatori possono essere due rotazioni di ordine 2 e 3, rispettivamente, e con centro distinto. Come aneddoto rimane da dire che, fino ad oggi, l'Alhambra risulta essere l'unico monumento al mondo, costruito prima dello sviluppo della teoria dei gruppi, che presenta questa caratteristica.

Epílogo

Dopo cinque secoli, i mosaici dell'Alhambra servono perché artisti del livello di Gaudí, Escher, Robinson, Falla, Debussy, ... si ispirino a questi per creare parte della loro arte. Stanno servendo anche perché gli studenti di matematica del mondo intero abbiano gli esempi più belli di oggetti tanto astratti come isometria, gruppi o superfici, elementi della geometria euclidea, algebra, geometria differenziale che diventano vivi. Impresionante! Però, che Alhambra potrebbero fare oggi i nostri progettisti? Se pensiamo che nella Porta del Vino c'è una bottiglia di Klein, qualcuno potrebbe pensare che Klein l'avrà dimenticata; succede anche con le striscie di Moëbius che ci sono nelle edicole che ci trovano all'entrata del Salone del Trono; però sentire che nel bagno de Comares c'è una sfera topologica, può sorprendere molti. Si immaginano questi oggetti matematici? Sono interessanti per l'architettura e il disegno? Con le nuove geometrie e strumenti del nostro secolo, molto può essere offerto al mondo dell'arte, in generale. Sarà questa una porta aperta per la quale continuiamo a progredire e offrire nuove visioni dell'Alhambra.

Infine, è importante inculcare ai nostri alunni che dalla matematica si può arrivare a godere della bellezza, in modo tale che, secondo le pa-

palabras de A.E. Pérez Sánchez, es intuita por todo el mundo, y se goza, pero siempre se desea alcanzar un nivel que la simple visión ingenua no alcanza. El espectador sensible quiere aprender a mirar; quiere recorrer - con alguien primero, solo después - el largo camino que el arte le sugiere. Las Matemáticas son una gran herramienta para lograrlo.

role di A. E. Pérez Sánchez, è intuita da tutti, e si gode, però sempre si desidera raggiungere un livello tale che la semplice visione ingenua non riesce a raggiungere. Lo spettatore sensibile desidera imparare ad osservare; desidera percorrere, prima insieme a qualcuno, e poi da solo, il lungo camino che suggerisce l'arte. La matematica è un grande strumento per arrivare a questo.



Ringraziamo A. M. Lallena per l'aiuto nella traduzione.



- [1] E. NUERE: *Tratado de la Carpintería de lo Blanco*, Ed. Ministerio de Cultura, Madrid (1982).
- [2] ALI A. AL-DAFFA- JOHN J. STROYLS: *Studies in the Exact Sciences in Medieval Islam*, Ed. Wiley and Sons (1984).
- [3] Y. BOSSARD: *Rosaces, frises et pavages (vol. 1 y 2)*, Ed. CEDIC, Paris (1979).
- [4] H. WEYL: *La Simetría*, Princeton University Press (1952).
- [5] L. FEJES-TOTH: *Regular Figures*, Ed. Pergamon Press, New York (1964).
- [6] B. GRÜNBAUM, G. S. SHEPARD: *Tilings and patterns*, Freeman, San Francisco (1984).
- [7] R. PÉREZ GÓMEZ: "The four regular mosaics missing in The Alhambra", *Comp. and Math. with Appls.* **14**, **2** (1987) 133-137.



R. Pérez Gómez: Doctor en Matemáticas, profesor del Departamento de Matemática Aplicada de la Universidad de Granada que imparte docencia en las Escuelas Técnicas Superiores de Arquitectura e Ingeniería de Caminos, Canales y Puertos. Ha escrito 9 libros de carácter científico-profesional, 10 capítulos de libros y 8 prólogos de libro. Sobre Matemáticas en el Arte, ha escrito dos libros sobre monumentos que forman parte del Patrimonio de la Humanidad: 7 paseos por la Alhambra (Ed. Proyecto Sur, 2007) y El Escorial. Historia, Arquitectura, Ciencia y Matemáticas (Ed. Servicio de Publicaciones del Ministerio de Educación).

R. Pérez Gómez: Dottore in Matematica, professore del Dipartimento di Matematica Applicata dell'Università di Granada, insegna presso la Scuola Tecnica Superiore di Architettura e Ingegneria stradale, di canali e dei porti. Ha scritto 9 libri di carattere scientifico-professionale, 10 capitoli di libri e 8 introduzioni di libri. Sulla Matematica e Arte, ha scritto due libri su monumenti che formano parte del patrimonio dell'umanità: 7 paseos por la Alhambra (Ed. Proyecto Sur, 2007) y El Escorial. Historia, Arquitectura, Ciencia y Matemáticas (Ed. Servicio de Publicaciones del Ministerio de Educación).

La fisica della musica

Giampaolo Co'

Dipartimento di Matematica e Fisica "Ennio De Giorgi" - Università del Salento

La musica è composta da sequenze di suoni le cui frequenze hanno rapporti semplici e ben definiti. Storicamente, nella musica occidentale è stata selezionata una sequenza particolare di frequenze che definiscono le note della scala musicale temperata.

Introduzione

Fruire di un'opera d'arte significa provare sensazioni che esulano da quelle solite, ordinarie. Questa affermazione, che spero possa essere largamente condivisa, non vuole essere una definizione di arte, ma è utile per identificare suoni che definisco musicali. In questa visione, ascoltare musica non significa semplicemente udire dei suoni, ma implica che i suoni uditi abbiano un ordine preciso, straordinario, ed, eventualmente, che producano anche piacere all'ascoltatore.

I suoni musicali si evolvono nel tempo con regolarità e gradevolezza. Ritmo, melodia ed armonia differenziano i suoni musicali da quelli comuni, ordinari. In questo articolo intendo legare queste tre caratteristiche dei suoni musicali alle grandezze fisiche che caratterizzano un generico suono e che, normalmente, sono identificate in durata, intensità e frequenza. Cercherò di illustrare come i suoni diventino musicali solo quando queste grandezze fisiche si mescolano in quantità ben definite.

Nel corso dei secoli la musica occidentale ha selezionato una determinata scala sonora composta da dodici suoni. Si è trattato di un processo puramente culturale, a cui sono associati l'evoluzione e lo sviluppo dei ritmi, delle melo-

die e delle armonie che caratterizzano l'ampia varietà di stili musicali prodotti dalla musica occidentale. Tuttavia, l'analisi fisica di questa scala musicale, e quindi delle armonie e melodie che si generano utilizzandola, mostra come esistano delle relazioni relativamente semplici tra le quantità fisiche che caratterizzano i suoni della scala dodecafonica occidentale.

Il suono

Dal punto di vista sensoriale, la musica è una forma d'arte selettiva, nel senso che attiva soltanto uno dei nostri cinque sensi, l'udito. Dal punto di vista scientifico questo semplifica molto lo studio della musica, perché si tratta di conoscere, e comprendere, l'unico fenomeno fisico percepito dall'orecchio, il suono.

Il suono è generato dalla vibrazione di un corpo. Questa vibrazione produce una perturbazione della pressione e della densità del mezzo nel quale il corpo vibrante si trova. La perturbazione si propaga attraverso questo mezzo per collisione delle sue molecole, e trasferisce energia ed impulso senza che ci sia un effettivo spostamento della materia. Questo significa che le molecole che compongono il mezzo oscillano attorno ad un punto di equilibrio, ma il loro spostamento, mediato su tempi molto più lunghi rispetto a quelli in cui si attiva la perturbazione, è nullo. Il trasferimento di energia e impulso senza trasferimento di materia è quello che caratterizza il fenomeno fisico delle onde, ecco perché si parla di onda sonora.

I fenomeni sonori di interesse per la musica sono prodotti da variazioni di densità e pressio-

ne i cui valori sono molto più piccoli dei valori che queste due quantità hanno quando non sono perturbate, ovvero quando sono in equilibrio. In questo caso le onde sonore sono in un regime lineare, e vale il principio di sovrapposizione, ovvero l'onda sonora risultante dall'incontro tra due, o più onde, è data dalla loro somma algebrica, che può essere sia costruttiva che distruttiva.

In questo regime di linearità la velocità del suono è data da

$$v = \sqrt{\left(-\frac{\Delta P}{\Delta V/V}\right) \frac{1}{\rho_0}} \quad (1)$$

Il termine tra parentesi tonde indica il rapporto tra la variazione di pressione ΔP e la variazione relativa del volume, V , del mezzo. Si tratta, per definizione, di quello che in inglese viene chiamato *bulk modulus*, che, per definizione, corrisponde all'inverso della compressibilità del materiale. Il termine ρ_0 rappresenta la densità d'equilibrio. Nell'equazione (1) non ci sono termini legati alla sorgente sonora, quindi la velocità del suono è una caratteristica del mezzo nel quale si propaga.

Nella musica, la produzione dei suoni avviene nel regime lineare sopra descritto, e, inoltre, la maggior parte delle sorgenti sonore produce suoni in maniera armonica. La vibrazione di una corda pizzicata, percossa o strofinata da un archetto, è descrivibile in termini di vibrazioni armoniche. Nello stesso modo i suoni prodotti dagli strumenti a fiato, e in generale da tutti i tubi sonori, e anche le vibrazioni di membrane, sono vibrazioni armoniche. Questi suoni possono essere descritti come sovrapposizione di onde armoniche.

Se consideriamo, ad esempio, il moto di una molecola d'aria che si trova nella posizione x al tempo zero, possiamo identificare la sua posizione al tempo t come

$$s(x, t) = A \sin k(x - vt) \quad (2)$$

dove A è l'ampiezza di oscillazione, k è detto numero d'onda e ha le dimensioni dell'inverso di una lunghezza, v è la velocità di propagazione dell'onda. La lunghezza d'onda è la distanza tra i due massimi, o più in generale tra due punti con lo stesso valore, $\lambda = 2\pi/k$. Il periodo T è il tempo necessario alla molecola per ritornare

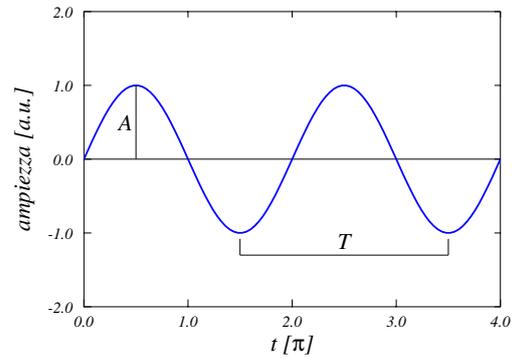


Figura 1: *Onda sinusoidale in funzione del tempo t , in unità di π . L'ampiezza dell'onda è indicata con A . Il periodo, T , è il tempo necessario per compiere un ciclo. Se avessimo disegnato l'onda in funzione dello spazio, al posto di T , ci sarebbe la lunghezza d'onda λ , che è la distanza percorsa in un periodo T . La frequenza ν è il numero di cicli effettuati nell'unità di tempo, quindi l'inverso del periodo.*

alla posizione iniziale. La frequenza ν , definita come inverso del periodo $\nu = 1/T$ è legata alla lunghezza d'onda secondo la relazione $\nu = v/\lambda$. Fissato il valore della velocità di propagazione, più piccola è la lunghezza d'onda, maggiore è la frequenza. L'unità di misura della frequenza, che corrisponde all'inverso di un tempo è chiamata Hertz (Hz).

Il trasporto di energia da parte di un'onda è normalmente descritto in termini dell'intensità dell'onda, definita come l'energia media trasferita nell'unità di tempo nell'unità di area, quindi, dato che energia per unità di tempo è potenza, come potenza media per unità di area $I = P/A$.

Se una sorgente puntiforme emette un suono in tutte le direzioni, l'intensità diminuirà con il quadrato della distanza dalla sorgente, poiché la potenza verrà distribuita su una superficie sferica, quindi $I = P/4\pi r^2$. Se si fa l'ipotesi che il suono non venga assorbito dal suolo si ha che la superficie è un'emisfera. Nella progettazione di teatri, si considerano le geometrie ed i materiali adatti per ottenere della buona acustica, ovvero per correggere al meglio la perdita di intensità legata al quadrato della distanza dalla sorgente.

Per un'onda armonica, è possibile legare l'intensità dell'onda all'ampiezza di oscillazione A dell'equazione (2). Per un'onda sonora alla pressione e alla densità di equilibrio P_0 e ρ_0

rispettivamente si ottiene [1]

$$I = \frac{1}{2} \frac{P_0^2}{\rho_0 v} \quad (3)$$

L'intensità del suono è legata al quadrato della pressione. Vedremo come la sensibilità dell'orecchio umano sia legata sia alle frequenze che alle intensità dei suoni percepiti.

Le sorgenti sonore

Normalmente, i suoni musicali sono sovrapposizione di onde armoniche di varie frequenze. Consideriamo, ad esempio, il caso di una corda vibrante con le due estremità bloccate, come indicato nella Figura 2. Questo è il caso delle corde

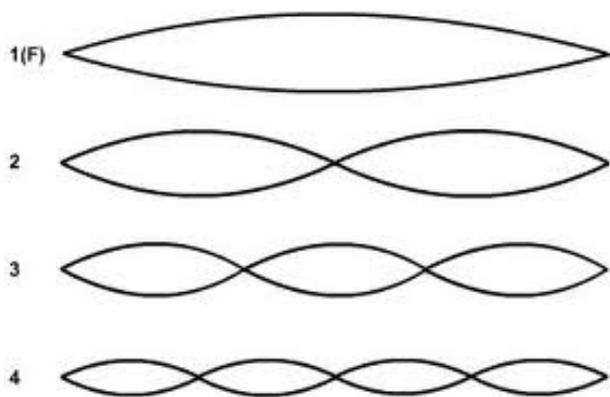


Figura 2: Vibrazioni di una corda vincolata ad entrambi gli estremi. In alto si mostra la frequenza principale, prima armonica, poi le armoniche successive.

del pianoforte, di quelle della chitarra e di tutti gli strumenti ad arco. La vibrazione principale della corda è indicata nella parte in alto della Figura 2. Questa vibrazione, detta principale, anche perché normalmente è quella che ha l'ampiezza maggiore, sfrutta tutta la lunghezza della corda. Quando si fa vibrare una corda, questo non è il solo modo di vibrazione. In aggiunta alla vibrazione principale, ci sono altri modi di vibrazione nei quali alcuni punti rimangono fissi. Questi punti sono detti nodi. Nel secondo modo di vibrazione indicato dalla Figura 2, la corda vibra sfruttando solo la metà della sua lunghezza, e ha un nodo. Come abbiamo visto, la frequenza è inversamente proporzionale alla lunghezza d'onda, quindi dimezzando la lunghezza d'onda si hanno moti oscillatori di frequenza doppia. C'è anche un moto nel quale appaiono due nodi.

In questo caso si hanno vibrazioni che sfruttano un terzo della lunghezza della corda, quindi, per questi modi di vibrazione la frequenza è tripla. Una sorgente armonica sonora genera una

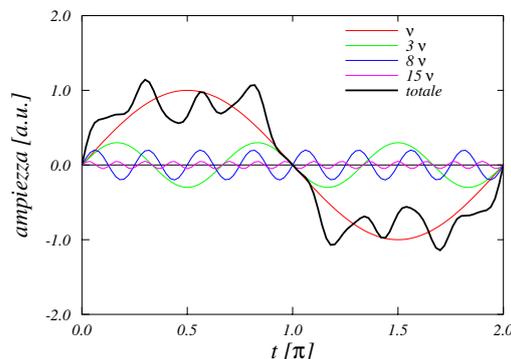


Figura 3: Sovrapposizione di onde sinusoidali con diverse frequenze multiple della frequenza principale ν .

sequenza di vibrazioni che hanno frequenze con semplici rapporti numerici che le contraddistinguono. Il suono finale è dato dalla somma di tutte queste semplici vibrazioni armoniche. In questa somma, le diverse armoniche sono pesate in maniera differente. L'armonica con la frequenza principale è quantitativamente la più importante, seguita poi dalle armoniche con frequenze più alte il cui contributo diminuisce con l'aumentare della frequenza. Questo è vero per i suoni prodotti da strumenti musicali, quindi per vibrazioni armoniche.

La Figura 3 mostra il risultato della somma di onde che all'armonica principale ν aggiungono, con ampiezze decrescenti, le armoniche con frequenze 3, 8, e 15 volte maggiori. La curva finale, indicata dalla linea nera, è piuttosto sconnessa nell'aspetto se confrontata con l'andamento dell'onda sinusoidale, ma è dominata dalla regolarità e periodicità della frequenza principale.

In ogni strumento musicale la frequenza principale è quella che definisce la nota, e accordare gli strumenti, significa fare in modo che le frequenze principali di tutti gli strumenti siano identiche. Le armoniche superiori, che si aggiungono a quella principale, caratterizzano il suono di ogni strumento. Per questo motivo la stessa nota suonata da strumenti diversi ha un suono differente. Ogni strumento, e qui mi riferisco al singolo oggetto, ha un suo timbro, una sua voce,

che lo distingue.

Alcuni strumenti musicali sono costruiti per produrre solo un suono, come ad esempio tamburi, campane. La maggior parte degli strumenti può produrre note differenti. Questo può avvenire perché lo strumento ha in dotazione un grande numero di elementi vibranti, come succede per il pianoforte, il clavicembalo, l'arpa e l'organo a canne. Molto più spesso la produzione di diverse armoniche dominanti avviene cambiando le dimensioni dell'entità vibrante. Negli strumenti a corda, questo avviene premendo la corda, quindi riducendo la lunghezza della parte vibrante. Negli strumenti a fiato si cambia nota aprendo le chiavi, permettendo all'aria di uscire da fori laterali e quindi modificando, riducendola, la lunghezza del tubo vibrante.

Qualsiasi suono può essere decomposto in termini di suoni armonici sinusoidali. Da questo punto di vista, ci sono due caratteristiche che distinguono i suoni ordinari dai suoni gradevoli prodotti dagli strumenti musicali. Una prima differenza è che non esiste una frequenza dominante rispetto alla quale le armoniche di frequenza superiore hanno ampiezze decrescenti. La seconda differenza consiste nel fatto che le varie armoniche non sono coerenti. Nei suoni musicali le vibrazioni sono nulle agli estremi del corpo vibrante, nei suoni non musicali questo non avviene, e le varie armoniche che si sommano sono fuori fase. La produzione di un suono da parte di uno strumento musicale genera una sequenza di onde armoniche che si sommano in fase e che hanno ampiezze decrescenti aumentando il valore delle frequenze. Queste due caratteristiche sono quelle che ci fanno percepire come gradevole il suono prodotto da uno strumento musicale.

La percezione del suono

Mentre le sorgenti sonore sono varie e multiformi, il rivelatore di suono è unico, ed è un sistema composto da una parte che converte il segnale meccanico in segnale elettrico, l'orecchio, e da un sistema che analizza quest'ultimo segnale, il cervello.

Nella Figura 4 è schematizzata l'anatomia dell'orecchio. Le variazioni di pressione, generate dall'onda sonora, penetrano nel condotto uditivo,

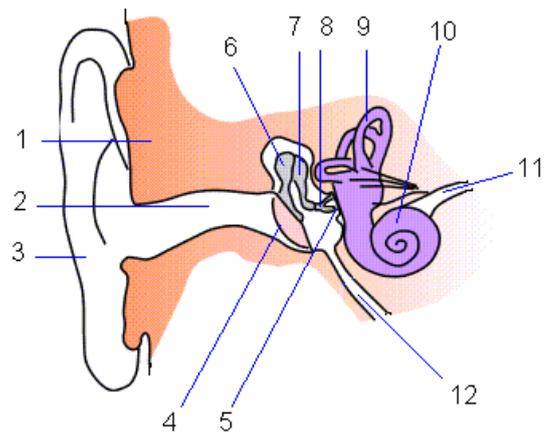


Figura 4: Orecchio esterno: 1 pericondrio, 2 condotto uditivo, 3 padiglione auricolare. Orecchio medio: 4 timpano, 5 finestra ovale, 6 martello, 7 incudine, 8 staffa. Orecchio interno: 9 canali semicircolari, 10 coclea, 11 nervo acustico, 12 tromba di Eustachio.

vo, fanno vibrare la membrana del timpano, che a sua volta mette in vibrazione i tre ossicini con i nomi di martello, incudine e staffa. La vibrazione si trasmette all'orecchio interno e mette in moto il liquido contenuto nella coclea dove delle cellule trasformano le oscillazioni del liquido in un segnale elettrico che attiva il nervo acustico. A questo punto il segnale viene portato al cervello che elabora l'informazione.

Questa sequenza di processi viene attivata se l'intensità dell'onda sonora supera un valore di soglia I_{\min} di circa 10^{-12} W/m² corrispondente ad una pressione di circa 3×10^{-5} Pa. Data la grande estensione di intensità sonore a cui è sensibile l'orecchio umano, si utilizza una scala logaritmica per descrivere le intensità sonore. In questa scala misurata in decibel (dB), il livello di intensità β è dato dall'espressione

$$\beta = 10 \log \frac{I}{I_{\min}} \quad (4)$$

In questa scala, la soglia di udibilità è data da $\beta = 0$ dB. Normalmente si ha una sensazione di dolore per intensità sonore di circa 1 W/m², corrispondenti a circa 120 dB.

L'altra caratteristica dell'onda sonora è la frequenza, e, anche in questo caso la sensibilità dell'orecchio è limitata ad una finestra di frequenze comprese tra 16 Hz a 20000 Hz (20 kHz). La pres-

sione dell'onda sonora si trasforma nella sensazione di quanto sia intenso il suono percepito. La frequenza invece ci produce la sensazione di quanto il suono sia acuto. I suoni di bassa frequenza sono quelli che classifichiamo gravi, mentre quelli di alta frequenza sono i suoni acuti.

La conformazione dell'orecchio è tale che un'onda sonora di media intensità mette in vibrazione varie parti che generano a loro volta onde sonore. Questo significa che, anche se il suono prodotto è composto da una sola frequenza, l'orecchio produce suoni con frequenze doppie, triple, quadruple eccetera della frequenza principale. Questi suoni di frequenza superiore a quella fondamentale hanno intensità inferiori con l'aumentare della frequenza.

Se il suono viene prodotto contemporaneamente da più sorgenti, con frequenze solo leggermente differenti, il cervello registra un solo suono con frequenza intermedia tra le due. Se la differenza tra le frequenze di emissione è abbastanza grande, il cervello distingue i suoni differenti. Se le differenze tra le frequenze sono frazioni semplici delle frequenze coinvolte si ha una sensazione di generica gradevolezza, altrimenti si ha una sensazione spiacevole, di stonatura.

Le note e la scala musicale

Sin dall'antichità ci si accorse che era possibile far suonare contemporaneamente più di uno strumento producendo comunque suoni globalmente gradevoli. La necessità di mettere d'accordo, da cui accordare, strumenti diversi perché suonando insieme non fossero dissonanti, spinse a formalizzare la scelta delle note da utilizzare. Da questa esigenza nacquero le varie scale musicali che si sono succedute nella storia della musica.

Quasi tutte le scale musicali sono basate su intervalli di ottava. Due note che hanno una frequenza doppia rispetto all'altra suonano particolarmente bene al nostro orecchio. Le percepiamo come fossero la stessa nota, una più alta rispetto all'altra. Molti ritengono che questa peculiarità abbia una base evolutiva. Nella produzione naturale dei suoni insieme ad una frequenza principale si producono altre frequenze con intensità decrescenti, tra le quali quella con frequenza doppia è quella più intensa. Il cervello permette di considerare uniche le note separate

da ottave evitando di cercare sorgenti differenti. Un altro intervallo particolarmente piacevole è quello di quinta, anche questo legato alla produzione naturale di ipertoni, prodotti da sorgenti naturali.

Già i filosofi greci di scuola pitagorica si accorsero dell'esistenza di relazioni semplici tra le frequenze di suoni gradevoli [2]. Sulla base di queste semplici relazioni si definirono nei secoli varie sequenze di suoni scelti per produrre musica. L'insieme di questi suoni, detti note musicali, viene definita come scala musicale. La musica di diversi contesti sociali si differenzia anche per l'uso di diverse scale musicali.

Non è mia intenzione descrivere l'evoluzione storica delle scale musicali che ha portato alla selezione della scala utilizzata oggi nella musica occidentale. Vorrei, invece, presentare i motivi che hanno portato al fallimento di tutti i tentativi fatti per costruire scale musicali su una base matematica o su una base fisica. A questo scopo, partendo dagli intervalli di quinta e ottava costruirò prima una scala musicale, che chiamerò Pitagorica, basata su rapporti semplici, razionali, tra i valori delle frequenze. Metterò in evidenza una difficoltà intrinseca nella scala Pitagorica, e indicherò la soluzione adottata che ha portato alla scala tonale moderna utilizzata nella musica occidentale. Questa scala musicale è dovuta ad una convenzione che definirei sociale, anche se si discosta di poco dalla scala Pitagorica.

La scala Pitagorica

La scala Pitagorica che descriverò non è la scala musicale utilizzata dagli antichi greci [2], ma si tratta di un esempio di scala musicale ideale costruita su base matematica.

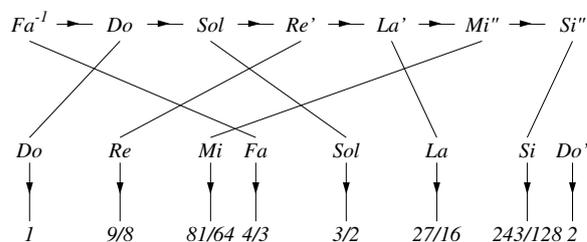


Figura 5: Schema della costruzione della scala pitagorica

Affrontiamo il problema della costruzione di una scala musicale utilizzando le informazioni

fisiche date finora. Sappiamo che due note le cui frequenze sono una il doppio dell'altra risultano produrre un suono gradevole se suonate contemporaneamente. Questo succede anche quando le due frequenze hanno un rapporto di $3/2$ come avviene, per esempio, tra la frequenza di Sol e quella del Do precedente.

La Figura 5 indica lo schema della costruzione della scala Pitagorica. Partiamo dal Do e, raddoppiando la frequenza, abbiamo un altro Do, che chiamiamo Do'. Moltiplicando per $3/2$ la frequenza del Do di partenza, alla quale assegnamo valore 1 nelle unità fisiche preferite dal lettore, otteniamo un Sol. Moltiplicando per $3/2$ la frequenza di Sol otteniamo la frequenza di $9/4=(3/2)^2$ che corrisponde ad un Re'. Il valore di questa frequenza è maggiore di 2, quindi questo Re è più alto del Do'. Dividendo per 2 la frequenza del Re' otteniamo $9/8 = 3^2/2^3 < 2$ che corrisponde alla frequenza del Re. Proseguendo in questo modo, moltiplicando Re' per $3/2$ otteniamo un La', quindi un Mi'' e un Si''. Dividendo per 2, o per 4, riportiamo tutte le frequenze tra 1 e 2. Otteniamo il Fa dividendo la frequenza di Do per $3/2$, e poi moltiplicando per due. In questo modo abbiamo costruito una scala di sette note i cui rapporti di frequenza sono indicati nella Figura 5. Queste sette note sono suonate dai tasti bianchi della tastiera del pianoforte. Bisogna notare come tutte le frequenze prodotte siano identificate da rapporti tra potenze dei numeri 3 e 2.

La scala temperata

La difficoltà d'uso della scala Pitagorica emerge quando si cerca di traslare uno scritto musicale da un tono all'altro. Abbiamo messo in evidenza come la gradevolezza di un brano musicale sia legata al rapporto tra le note che lo compongono, e non al valore assoluto delle loro frequenze, purché queste rimangano nei limiti dell'udibilità. È capitato a tutti di cantare una canzone partendo da una nota ed incontrando difficoltà a riprodurre la melodia, perché l'estensione della voce non riusciva a raggiungere certe note, di ricominciare a cantare la melodia partendo da una nota più bassa o più alta. Questo significa che lo stesso brano musicale risulta egualmente gradevole se suonato spostando tutte le note che lo

compongono dello stesso fattore moltiplicativo. Si può, ad esempio pensare di trasformare una

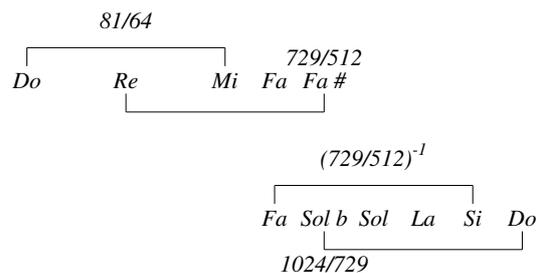


Figura 6: Rapporti di frequenze tra Do e Mi e tra Si e Fa

melodia che parte in Do facendola partire dal Re. Supponiamo che le prime note della melodia siano Do e Mi. Il rapporto tra queste frequenze è di $81/64$, come indicato nella Figura 6. Moltiplicando la frequenza di Re di $9/8$ per $81/64$ si ottiene $729/512$, che è la frequenza di una nota che non si trova nella scala definita precedentemente. Chiameremo Fa# questa nuova nota la cui frequenza si trova tra il Fa e il Sol. Queste note, sulla tastiera del pianoforte vengono suonate premendo i tasti neri.

Supponiamo che la melodia che stiamo trasportando di un tono abbia un passaggio tra Si e Fa che dà $(729/512)^{-1}$ in termini di rapporto di frequenza. Questo passaggio fatto partendo dal Do', di frequenza 2, porta ad una nota con frequenza $1024/729$, che chiameremo Solb. Anche questa nota si trova tra Fa e Sol **ma ha frequenza differente da quella del Fa#!** D'altra parte sul pianoforte c'è un solo tasto nero tra i tasti bianchi del Fa e del Sol.

Questo esempio mette in evidenza la difficoltà intrinseca nell'uso della scala pitagorica. Ogni cambiamento di tono produce nuove note le cui frequenze non sono presenti nelle scale precedentemente definite. Se si usasse la scala Pitagorica per uno strumento a tastiera, tra due tasti bianchi ci sarebbero numerosi tasti neri, ognuno costruito per rappresentare una nota ottenuta partendo da una tonalità differente.

Il problema è stato ampiamente discusso nei secoli scorsi fino ad arrivare ad un accordo nel quale si è fatta la scelta più semplice. Si è deciso di suddividere l'ottava in dodici parti uguali, quindi il rapporto di frequenza tra due note successive è $2^{1/12}$. Il risultato di questa operazione è presentato nella Tabella 1 dove confronto i

Nota	Pitagorica	Temperata
Do	1.00000	1.00000
Re	1.12500	1.12246
Mi	1.26563	1.25992
Fa	1.33333	1.33484
Sol	1.50000	1.49831
La	1.68750	1.68179
Si	1.89844	1.88775
Do'	2.00000	2.00000

Tabella 1: Frequenze delle sette note principali nella scala Pitagorica e in quella Temperata.

rapporti tra le frequenze della scala Pitagorica, mostrati anche nella Figura 5, con quelli della scala Temperata. Come si vede, le differenze tra le frequenze delle due scale è dell'ordine di qualche percento. In questo modo si è raggiunto un ragionevole compromesso tra l'armonia pitagorica dei numeri razionali e la necessità pratica di avere note che permettono di trasportare la musica in tonalità differenti.

Una nobile testimonianza della polemica legata a questo problema è la raccolta BWV 846-893, divisa in due libri, di preludi e fughe in tutte le 12 tonalità, in modo maggiore e minore, composta da Johann Sebastian Bach (1685-1750) e da lui nominata *Das wohltemperierte Clavier*, tradotta in *Il Clavicembalo ben temperato*, anche se *Clavier* ai tempi di Bach indicava genericamente la tastiera.

Ritmo, melodia ed armonia

Dal punto di vista fisico possiamo definire la musica come una sequenza temporale di suoni selezionati tra quelli che formano la scala musicale. Se si produce soltanto un suono per ogni istante di tempo si ha una monofonia, e l'evoluzione temporale di questi suoni produce la melodia. La produzione contemporanea di vari suoni viene normalmente chiamata polifonia, e l'effetto dei vari suoni suonati contemporaneamente genera l'armonia. Il terzo elemento che forma la musica è il ritmo ed è legato a come melodia ed armonia si ripetono nel tempo. La definizione della scala musicale, quindi, sta alla musica come l'alfabeto alla poesia. Bisogna comporre le

note per produrre parole, e parole per formare frasi.

A questo punto la fisica ha assolto il suo compito. Per capire perché una sequenza di suoni risulti essere musicalmente attraente, è necessario considerare l'intreccio tra la fisica dell'onda sonora e la fisiologia della ricezione del suono, la psicologia dell'interpretazione del segnale sonoro, e, ultima ma non per questo meno importante, la sociologia, ovvero il contesto sociale e culturale dell'ascoltatore [3].

I tentativi compositivi di andare oltre i limiti della scala tonale hanno prodotto della musica difficilmente ascoltabile. Questa evoluzione della musica occidentale era basata più su posizioni ideologiche dei compositori che sulla sensibilità musicale.

Conclusioni

Il piacere all'ascolto della musica è legato al fatto che le note sono definite da una scala nella quale le loro frequenze hanno relazioni semplici. Abitudini all'ascolto e culture differenti possono portare a godere di combinazioni leggermente differenti di questi suoni, ma rimane il forte vincolo di una base fisiologica e psicologica che non fa discostare di molto le varie scale musicali da quella usata nella musica tonale occidentale.

Da quanto detto, sembrerebbe che le atonalità, le note fuori scala, siano da eliminare in una composizione musicale sulla base di considerazioni di tipo fisico, neurologico e psicologico. In realtà la situazione è molto più sfumata. Per utilizzare una metafora culinaria, l'uso delle note fuori scala è un po' come quello delle spezie in cucina. Sono molto utili per accendere l'interesse su qualche cosa che altrimenti sarebbe banale, ovvia, prevedibile. Ogni forma musicale può essere più o meno speziata secondo le abitudini e la cultura di chi la produce e di chi ne fruisce, ma così come è difficile cucinare piatti gustosi utilizzando soltanto spezie, è difficile costruire brani musicali gradevoli utilizzando esclusivamente suoni non inseriti in alcuna scala musicale.



- [1] P. A. TIPLER: *Physics*, Worth, New York (1976).
- [2] S. SAMBURSKY: *Il mondo fisico dei greci*, Feltrinelli, Milano (1973).
- [3] P. BALL: *L'istinto musicale*, Dedalo, Bari (2010).

Giampaolo Co': Professore associato di Fisica Nucleare presso l'Università del Salento si occupa di teorie a molti corpi applicate allo studio della struttura dei nuclei atomici.



Una geometrizzazione dello spazio degli accordi

Musica est exercitium arithmeticae occultum nescientis se numerare animi.

————— **Gottfried Wilhelm von Leibniz (1646 -1716)**

Mattia G. Bergomi

IRCAM - UPMC - CNRS - LIM - Università degli Studi di Milano

Riccardo D. Jadanza

DISMA - Politecnico di Torino

Alessandro Portaluri

DISAFA- Università degli Studi di Torino

Si può trovare un'equazione per descrivere un brano musicale, o addirittura per prevedere se sarà gradevole all'ascolto? Sarebbe possibile modellare una sinfonia con un certo numero di equazioni? La risposta a tutte queste domande è facilmente intuibile da chiunque addetto e non addetto ai lavori ed è invariabilmente un enfatico **“no!”**. Tuttavia si possono riconoscere alcune strutture matematiche insite in tutte le opere musicali, anche se non sono espresse attraverso equazioni. Il nostro viaggio attraverso **“curiose iper-superfici singolari”** inizia proprio da qui. . .

Musica e Matematica sono strettamente legate. Le connessioni tra questi due mondi sono ovunque: le corde di un'arpa vibrano ad una certa frequenza e hanno lunghezze specifiche regolate da rapporti ben precisi, le onde sonore possono essere descritte da equazioni matematiche, la forma della cassa del violoncello è studiata per farlo

risuonare con quel timbro caldo e suadente che ben conosciamo. La Matematica, in fondo, è il linguaggio che i fisici, i chimici e gli ingegneri (tra gli altri) usano per descrivere il mondo reale e in alcuni casi anche per prevedere il futuro a volte lontano! Nel campo della Musica è un valido strumento per cercare di comprendere, se esistono, alcune strutture astratte sottostanti al metalinguaggio e aiutare così la Teoria Musicale nel suo intento di comunicazione e di educazione all'ascolto. Chiunque abbia mai ascoltato il **“Klaviersstück”** di Stockhausen, infatti, capisce bene che questo non è sempre facile da fare!

La Teoria Musicale ci fornisce categorie concettuali per organizzare e capire la musica, e attraverso di esse le nostre impressioni auditive diventano idee vivaci. I musicologi attingono spesso alla grande potenza della Matematica per la creazione di tali categorie: i numeri interi, ad esempio, sono particolarmente adatti per l'etichettatura delle note o dei tasti del pianoforte, mentre una disciplina matematica che va sotto il nome di Combinatoria permette di contare i molti modi in cui le note possono essere asso-

ciate tra loro. Questo dà luogo a tassonomie e permette di classificare i diversi scenari che si presentano. La Teoria dei Gruppi, poi, altra importante area della Matematica, descrive il modo in cui le note sono legate e come si passa dall'una all'altra.

Un pianista può suonare migliaia di note in un concerto, senza bisogno di leggere lo spartito. Ma come ci riesce? Sicuramente è necessaria una buona dose di memoria, ma esistono anche alcuni trucchi armonici... Matematicamente parlando, è sufficiente riconoscere una semplicissima *foliazione di un orbifold* (Figura 1)! Ma iniziamo da...



Figura 1: Foliazione di un orbifold?!

Una chiacchierata sugli spazi quoziente

Uno *spazio quoziente* si costruisce identificando, cioè "incollando" fra loro, tutti i punti di uno spazio di partenza (detto *spazio genitore*) secondo una ben determinata famiglia di operazioni \mathcal{F} , che fornisce opportune "regole di attaccamento" che stabiliscono quali punti debbano essere accoppiati con quali altri. Un *dominio fondamentale* per \mathcal{F} è una regione nello spazio genitore avente le seguenti proprietà:

- ogni punto dello spazio genitore è legato ad ogni punto del dominio fondamentale tramite qualche operazione di \mathcal{F} ;
- nessuna coppia di punti del dominio fondamentale è legata tramite una qualsivoglia operazione di \mathcal{F} .

Intuitivamente, un dominio fondamentale può essere pensato come una singola tessera di un mosaico che permette di riprodurre tutto il disegno.

Ad esempio, il primo quadrante chiuso del piano cartesiano (quello delle ascisse e ordinate non negative, $x, y \geq 0$) è il dominio fondamentale per la famiglia di simmetrie definita dalle rotazioni

di 180° . È facile rendersene conto: ogni punto del piano è infatti legato tramite una rotazione di 180° a un punto di questo quadrante (v. Figura 2).

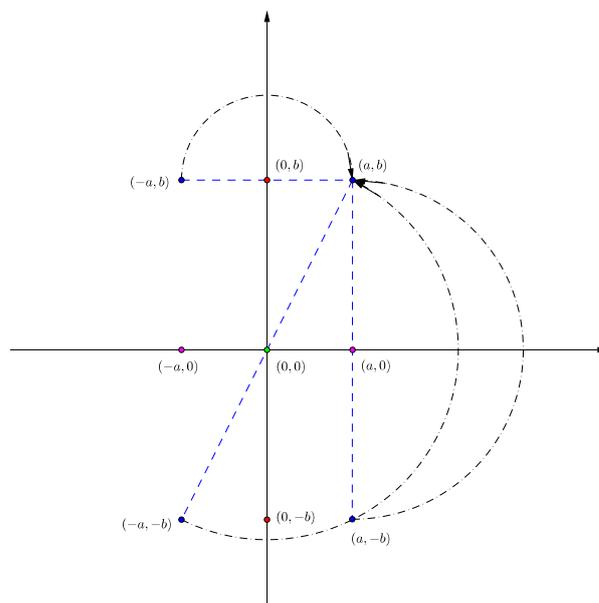


Figura 2: Regola di incollamento sullo spazio genitore.

Si osserva facilmente che che il semiasse negativo delle ascisse deve essere incollato al semiasse positivo in modo che -1 coincida con 1 , $-\pi$ con π e in generale ogni $-x \in \mathbb{R}^-$ con il suo opposto $x \in \mathbb{R}^+$. Il punto $(0, 0)$ invece si identifica con se stesso (v. Figura 3).

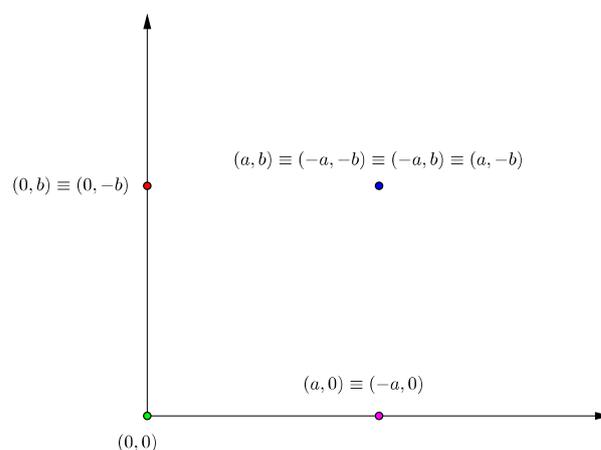


Figura 3: Spazio quoziente con identificazioni.

Note, accordi e orbifolds

Il *temperamento equabile* è il sistema musicale per la costruzione della scala fondato sulla suddivisione dell'ottava in intervalli uguali tra loro. Nell'uso più frequente l'ottava è suddivisa in 12

Notazione anglosassone

Una notazione utile per indicare le note pure (cioè senza alterazioni) è la seguente:

$$A = \text{La}, \quad B = \text{Si}, \quad C = \text{Do}, \quad D = \text{Re}, \\ E = \text{Mi}, \quad F = \text{Fa}, \quad G = \text{Sol}.$$

Un pedice numerico indica l'ottava a cui la nota appartiene. Ad esempio A_4 è il La sopra al Do centrale e ha una frequenza fondamentale pari a 440 Hz.

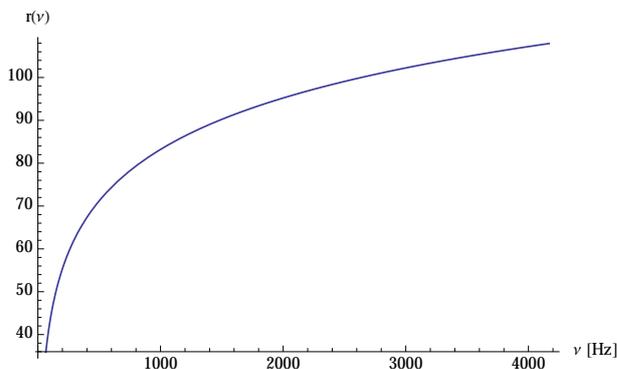


Figura 4: La funzione r per frequenze comprese tra 65 Hz e 4186 Hz, ovvero da C_2 a C_8 .

parti uguali, i *semitoni*, e un buon modo per modellarla matematicamente è usare l'insieme dei numeri reali, che indicheremo con \mathbb{R} .

Dallo spazio lineare delle note...

Un modo di associare ad ogni nota un numero reale è quello di utilizzare il logaritmo in base 2. Se ν rappresenta la frequenza fondamentale di una nota, il numero reale ad essa corrispondente è definito come

$$r(\nu) := c_1 + c_2 \log_2 \left(\frac{\nu}{440} \right),$$

dove c_1 rappresenta l'*origine* (ossia il numero corrispondente alla frequenza fondamentale 440 Hz, caratteristica della nota A_4) e c_2 è il numero di semitoni in un'ottava. Il vantaggio è che se si vuole calcolare la distanza tra due note r e s è sufficiente considerare la quantità

$$|r - s|$$

anziché che il rapporto ν_r/ν_s tra le loro frequenze fondamentali.

In musica classica si pone di regola

$$c_1 := 69, \quad c_2 := 12.$$

In questo modo al Do centrale C_4 corrisponde il numero 60, mentre per la nota A_5 dell'ottava superiore, che ha una frequenza fondamentale di 880 Hz, si ha $r(880) = 69 + 12 = 81$.

Osserviamo (Figura 4) che la funzione r è sempre strettamente crescente all'aumentare di ν e assume valori positivi nello spettro delle frequenze udibili dall'orecchio umano (da 20 Hz a

20 000 Hz). Si annulla infatti per un valore della frequenza pari a $\nu_0 := 2^{-69/12} \cdot 440 \approx 8.18$ Hz.

... allo spazio curvo delle classi di note

Nello spazio lineare delle note le ottave non rivestono alcun ruolo privilegiato, perché la distanza di 12 semitoni non ha un significato particolare. Tuttavia noi percepiamo le note che differiscono di un'ottava come aventi la stessa "qualità", lo stesso "colore"! È infatti evidente a chiunque abbia un minimo di esperienza musicale che alcuni intervalli sono "speciali", almeno per quanto riguarda la dissonanza.

Diventa allora del tutto naturale ripartire le note in *classi modulo ottava*: ogni classe è un punto dello spazio quoziente¹

$$\mathbb{T}^1 := \mathbb{R}/12\mathbb{Z},$$

che non è nient'altro che una *circonferenza*. Dunque, ricapitolando, le note "vivono" su una retta (Figura 5), mentre le *classi di note* "abitano" su una circonferenza e le loro "case" sono i vertici di un dodecagono regolare inscritto (Figura 6).

Gli elementi di \mathbb{T}^1 sono *insiemi di numeri reali* della forma

$$\{ r + 12k \mid k \in \mathbb{Z} \},$$

dove r rappresenta una qualsiasi nota in una data classe di note. Possiamo etichettare questi insiemi utilizzando l'intervallo

$$[0, 12) \subset \mathbb{R},$$

¹La scrittura $12\mathbb{Z}$ indica l'insieme dei numeri interi multipli di 12.

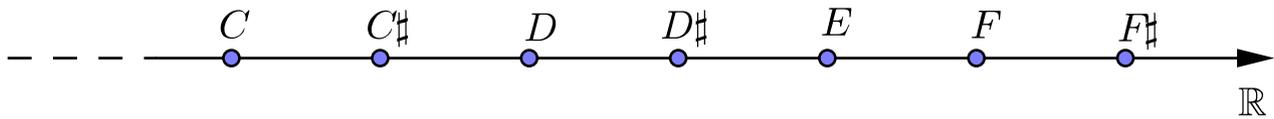


Figura 5: Rappresentazione grafica delle note sulla retta dei numeri reali.

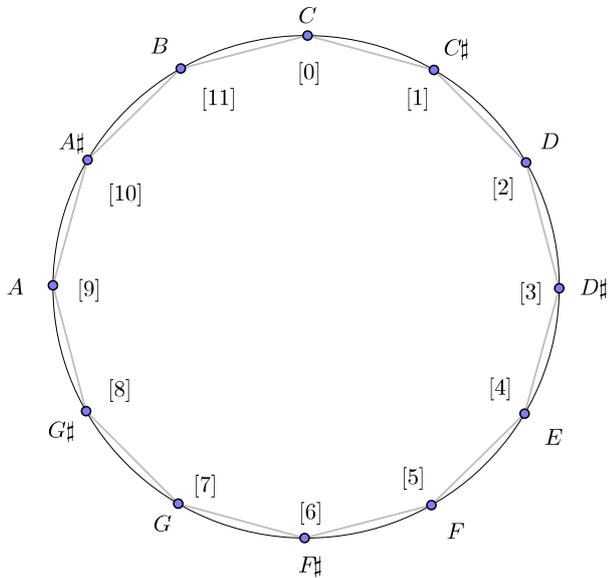


Figura 6: Rappresentazione grafica delle classi di note modulo ottava.

che è il *dominio fondamentale* dell'azione di traslazione del gruppo $12\mathbb{Z}$ sullo spazio genitore \mathbb{R} .

Diadi

E se considerassimo accordi di due note, detti anche *bicordi* o *diadi*? In questo caso la questione inizia a complicarsi un po'. Ingenuamente si potrebbe pensare che un bicordo sia un punto del *toro* bidimensionale

$$\mathbb{T}^2 := \mathbb{R}^2 / (12\mathbb{Z})^2$$

rappresentato in Figura 7, o più precisamente un punto del dodecagono bidimensionale (una sorta di "prodotto" di un dodecagono per se stesso) contenuto nel toro \mathbb{T}^2 . È infatti naturale dedurre che se ogni classe di note è rappresentata da un punto di \mathbb{T}^1 , allora una coppia di note deve essere un punto del prodotto $\mathbb{T}^1 \times \mathbb{T}^1 = \mathbb{T}^2$!

In realtà, un'analisi più attenta suggerisce che si sia trascurato un fatto importante. Infatti, considerando le classi di note modulo ottava, non siamo in grado di stabilire l'*ordine* tra le note che

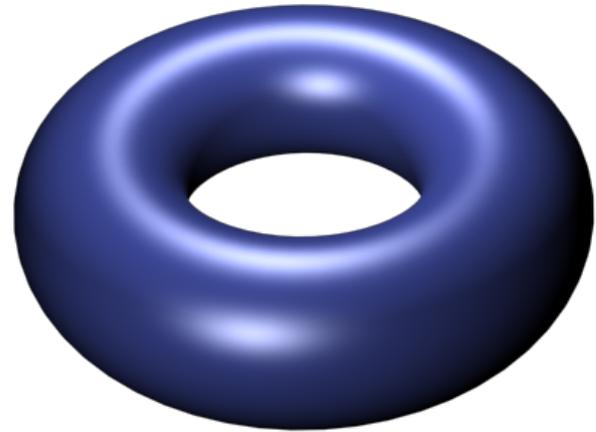


Figura 7: Rappresentazione grafica di un toro bidimensionale. (Fonte: <http://it.wikipedia.org>).

compongono ogni intervallo: il bicordo (E_4, C_5) , ad esempio, è indistinguibile da (E_4, C_4) perché C_4 e C_5 sono elementi della stessa classe di note C modulo ottava. Inoltre, musicalmente parlando, (E_4, C_4) è identico a (C_4, E_4) : se si eseguono questi accordi su un pianoforte, i tasti che si premono sono il Do centrale e il Mi sopra al Do centrale in entrambi i casi!

Occorre allora "rinunciare all'ordine" e passare dallo spazio delle coppie *ordinate* di classi di note (rappresentato geometricamente dal toro bidimensionale \mathbb{T}^2) allo spazio delle coppie *non ordinate*; per farlo si ricorre a una tecnica tipica della Topologia della dimensione bassa nota col nome di *taglia e cucì*.

Un toro si può costruire a partire da un quadrato $\mathcal{Q} := [0, 1] \times [0, 1] \subset \mathbb{R}^2$ incollandone i lati a due a due opposti con la stessa orientazione (Figura 8): una prima identificazione genera un cilindro; successivamente si attaccano tra loro le due circonferenze di base del cilindro ottenuto (che sono in realtà gli altri due lati del quadrato di partenza) e si ottiene la "ciambella" voluta. Utilizzando questa rappresentazione del toro, identifichiamo le coppie (x, y) e (y, x) in \mathcal{Q} : ciò si effettua ripiegando il quadrato a metà lungo la

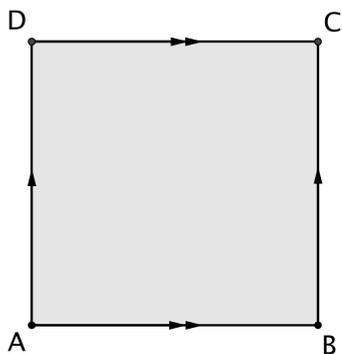


Figura 8: Costruzione di un toro attraverso l'identificazione dei lati opposti di un quadrato.

diagonale che giace sulla bisettrice del primo e terzo quadrante nel piano cartesiano e incollandolo. Si ottiene così un triangolo rettangolo. A questo punto bisogna procedere con gli attaccamenti dei lati opposti secondo le direzioni precedentemente fissate: tagliando il triangolo lungo la bisettrice dell'angolo retto e segnando i segmenti divisi dal taglio si formano altri due triangoli rettangoli più piccoli, che hanno le ipotenuse e un cateto identificati a coppie. Incollando fra loro prima le ipotenuse e poi i cateti (secondo le orientazioni fissate) si arriva allo spazio cercato:

il nastro di Möbius.

Il procedimento è riportato in Figura 11, il risultato invece in Figura 9.



Figura 9: Geometrizzazione dello spazio delle diadi: il nastro di Möbius. (Fonte: <http://it.wikipedia.org>).

Lo spazio delle diadi corretto è allora lo spazio quoziente

$$\mathbb{X}_2 := \mathbb{T}^2 / \mathcal{S}_2,$$

dove \mathcal{S}_2 denota il gruppo delle simmetrie degli insiemi di due elementi e rappresenta matemati-

camente l'azione di identificazione delle coppie (x, y) e (y, x) , rendendo in questo modo l'ordine irrilevante (Figura 10). Questa volta quindi lo spazio che si ottiene ha un'importante differenza rispetto alla circonferenza (oltre alla dimensione, naturalmente, che prima era 1 e adesso è 2): è *non orientabile*, cioè non si può distinguere un "interno" e un "esterno". Se infatti si prova a seguire idealmente con un dito la superficie raffigurata in Figura 9 ci si accorge che esiste una sola faccia!

Questo non è certo l'unico modello possibile di spazio quoziente che si può usare per geometrizzare lo spazio degli accordi! Sarebbe infatti del tutto legittimo modellare lo spazio delle classi di note (modulo ottava) nel temperamento equabile con il gruppo ciclico

$$\mathbb{Z}_{12} := \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\},$$

in cui 12 viene identificato con 0, 13 con 1 e così via. Con questa scelta, lo spazio delle diadi sarebbe allora lo spazio quoziente

$$\mathbb{Y}_2 := (\mathbb{Z}_{12})^2 / \mathcal{S}_2$$

rappresentato in Figura 10.

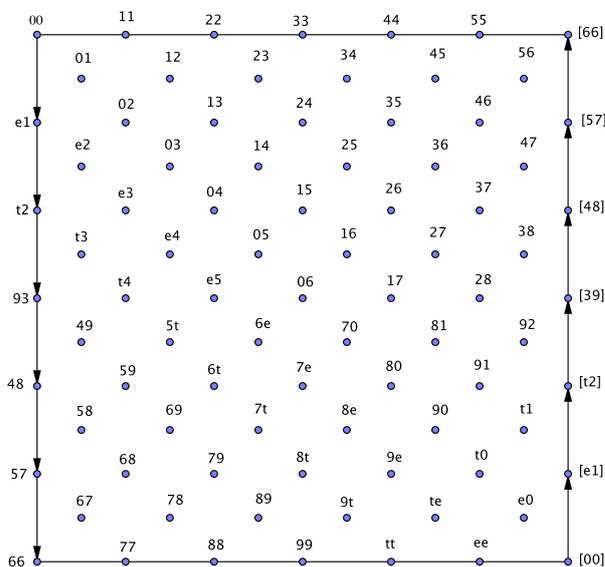


Figura 10: Lo spazio reticolare \mathbb{Y}_2 immerso in \mathbb{X}_2 . I numeri che etichettano ogni punto sono degli intervalli, le cifre (e lettere) che li compongono corrispondono alle note della rappresentazione circolare di Figura 6, dove $10 = t$ e $11 = e$.

Consideriamo ora le diadi

$$\{D, E\}, \quad \{D\#, E\}$$

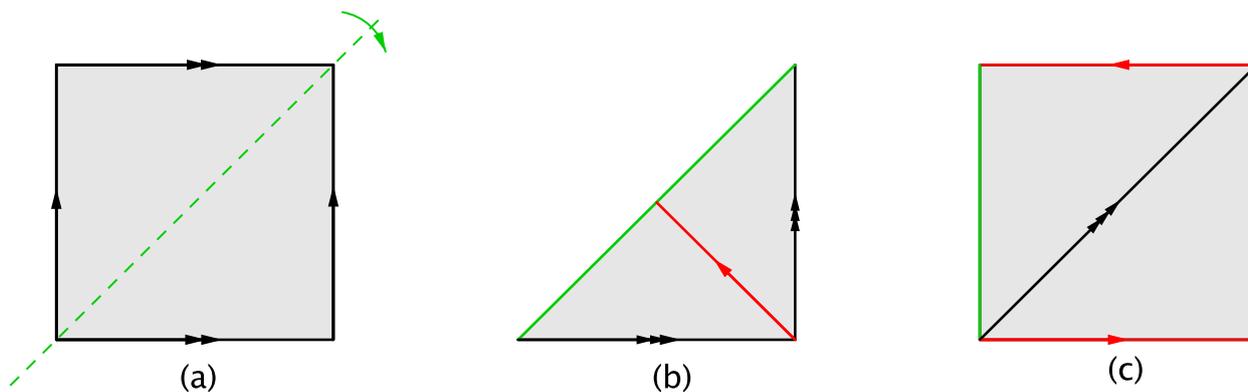


Figura 11: I tre passi necessari per realizzare lo spazio delle diadi \mathbb{X}_2 : (a) piegare e incollare il quadrato lungo la diagonale, (b) tagliare il triangolo ottenuto lungo la bisettrice dell'angolo retto, (c) incollare prima le ipotenuse nel giusto verso e poi identificare gli altri due cateti (in rosso) seguendo l'orientazione fissata. Il bordo del nastro di Möbius è rappresentato dai segmenti verdi durante tutto il procedimento.

(le parentesi graffe stanno a significare che stiamo prendendo punti dello spazio \mathbb{X}_2 , rappresentati da coppie *non ordinate*) e le due progressioni

$$p_1 : (D, E) \mapsto (D\sharp, E), \quad p_2 : (D, E) \mapsto (E, D\sharp),$$

completamente diverse tra loro dal punto di vista armonico. In \mathbb{X}_2 queste ultime sono rappresentate da due cammini distinti:

- p_1 si muove direttamente da $\{D, E\}$ a $\{D\sharp, E\}$;
- p_2 si inizia in $\{D, E\}$ e termina in $\{D\sharp, E\}$ ma segue un percorso diverso: “rimbalza” sul bordo del quadrato (nello spazio genitore), ovvero compie un giro sul nastro di Möbius.

Se considerassimo lo spazio discreto \mathbb{Y}_2 queste due progressioni risulterebbero “invisibili” (v. Figura 12 più avanti, a pagina 39): nello spazio discreto infatti non esistono segmenti veri e propri, ma solo un reticolo di punti!

Inoltre osserviamo che nella definizione di \mathbb{X}_2 e di \mathbb{Y}_2 non abbiamo escluso gli elementi del tipo $\{C, C\}$, che sono diadi fittizie: se, di nuovo, provassimo a suonarle su un pianoforte premeremo solamente un tasto! Questa “degenerazione” è in realtà molto utile per la ragione seguente: se considerassimo il nastro di Möbius non compatto (cioè “senza bordo”, ottenibile facilmente dalla costruzione di Figura 11 rimuovendo i due lati verticali dal quadrato (c)), che rappresenta lo spazio delle diadi con due classi di note distinte,

non avremmo alcun bordo su cui “rimbalzare” per poter distinguere due progressioni diverse! Questo fatto fornisce una motivazione in più per preferire il modello continuo a quello discreto: il cammino (geodetico) interseca il bordo singolare in un punto a coordinate non intere! Quindi, anche se anche riuscissimo a definire in \mathbb{Y}_2 la nozione analoga di “segmento”, sarebbe a dir poco scomodo modellare questo cammino.

Negli ultimi decenni i musicologi e gli studiosi di teoria musicale hanno introdotto vari spazi quoziente, utili a definire in maniera appropriata lo spazio degli accordi, tenendo conto della necessità di descrivere al meglio la condotta delle parti. Tra le rappresentazioni più rilevanti in questa direzione ricordiamo la teoria degli accordi di Allen Forte descritta in [8].

Triadi, quadriadi e orbifolds

Ovviamente quanto costruito finora è del tutto insufficiente in Musica! In Armonia si parla infatti di *triadi* e *quadriadi*, ovvero accordi costituiti da tre o quattro note rispettivamente, ciascuna delle quali è assegnata a una specifica voce o parte. Ovviamente esistono e si usano accordi composti anche da un numero maggiore di note! Eccezion fatta per una certa difficoltà d’immaginazione (e di rappresentazione grafica!), il modello di spazio degli accordi rimane però, *mutatis mutandis*, quello che abbiamo presentato per le diadi!

Nel paragrafo sugli spazi quoziente abbiamo introdotto la nozione di dominio fondamentale e di famiglia \mathcal{F} di operazioni (o simmetrie). Quali

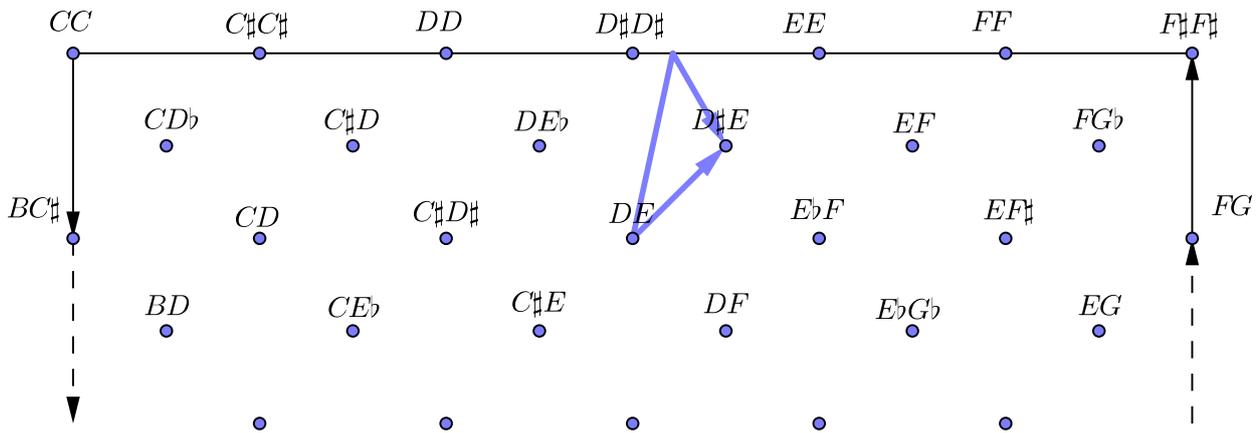


Figura 12: Sostituendo i numeri di Figura 10 con i nomi delle note, è facile vedere come \mathbb{X}_2 sia adatto a rappresentare le condotte delle parti: le frecce infatti rappresentano i due diversi percorsi p_1 e p_2 utilizzati per muoversi dall'intervallo (D, E) a $(D\sharp, E)$.

sono le simmetrie in questo caso, e quali i loro domini fondamentali?

Procedendo per ordine, iniziamo dallo spazio euclideo \mathbb{R}^n dotato del sistema di coordinate (x_1, \dots, x_n) . Per fissare le idee si può prendere $n = 3$ e immaginare lo spazio tridimensionale a cui siamo abituati, con $(x_1, x_2, x_3) = (x, y, z)$. Un punto di \mathbb{R}^n può dunque descrivere, dopo un'opportuna sequenza di identificazioni, un accordo di n note. Le operazioni di cui abbiamo bisogno sono analoghe a quelle descritte in precedenza e sono presentate nel seguito.

- 1) Consideriamo dapprima il passaggio \mathcal{O} dalla singola nota alla classe corrispondente modulo ottava. L'azione di $\mathcal{O} := (12\mathbb{Z})^n$ su \mathbb{R}^n produce lo spazio quoziente

$$\mathbb{T}^n := \mathbb{R}^n / (12\mathbb{Z})^n,$$

detto n -toro. Come dominio fondamentale $\mathcal{D}(\mathcal{O})$ sarebbe naturale scegliere $[0, 12)^n$, tuttavia si preferisce utilizzarne un altro, matematicamente inusuale ma che ha una notevole importanza in Musica: l'insieme di tutti i punti di \mathbb{R}^n che verificano le seguenti due condizioni:

- i) $\max\{x_1, \dots, x_n\} \leq \min\{x_1, \dots, x_n\} + 12$;
- ii) $0 \leq \sum_{i=1}^n x_i \leq 12$.

Il primo vincolo assicura che tutte le note dell'accordo siano comprese in un'ottava, men-

tre il secondo determina la posizione globale dell'accordo nello spazio genitore.

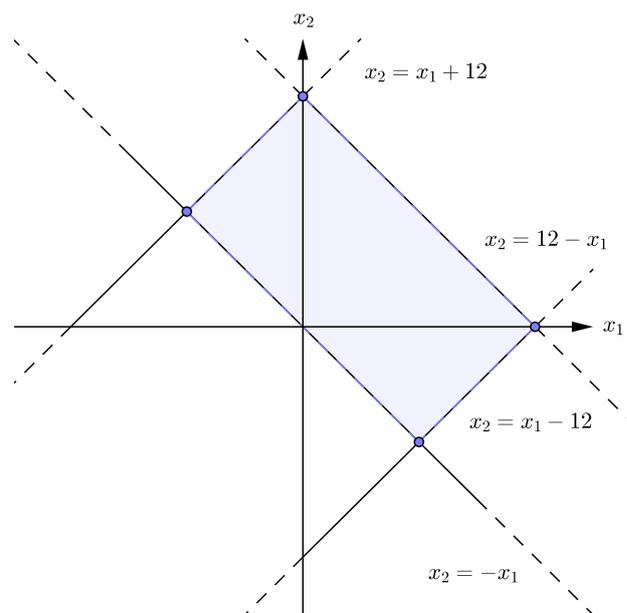


Figura 13: Rappresentazione del dominio fondamentale $\mathcal{D}(\mathcal{O})$ per $n = 2$.

Intuitivamente $\mathcal{D}(\mathcal{O})$ è un prisma le cui facce opposte vengono identificate con la stessa orientazione. La base del prisma è la proiezione (cioè l'"ombra") di un ipercubo n -dimensionale sul piano $(n - 1)$ -dimensionale che contiene i punti la cui somma delle coordinate è uguale a 0. Nel caso $n = 2$ il dominio fondamentale si riduce al rettangolo

delimitato dalle quattro rette di equazione

$$\begin{aligned} x_1 + x_2 = 0, & & x_1 - x_2 = 12, \\ x_1 + x_2 = 12, & & x_1 - x_2 = -12; \end{aligned}$$

in questo caso la base è un segmento (Figura 13). In dimensione 3 il prisma è a base esagonale, mentre in dimensione 4 la base è un rombododecaedro (o dodecaedro rombico). L'utilità di questo dominio fondamentale risiede nel fatto che la direzione dell'altezza del suddetto prisma coincide con la direzione della *trasposizione* (questa è rappresentata infatti da una traslazione individuata da un vettore con tutte le componenti uguali).

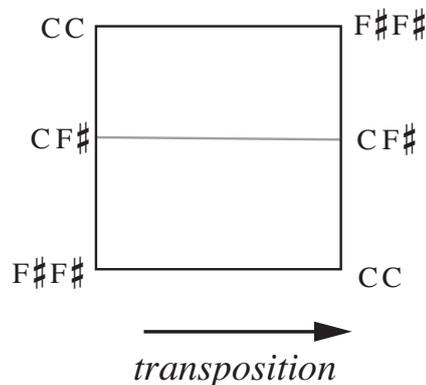


Figura 14: Lo spazio delle diadi \mathbb{X}_2 come spazio quoziente. Prima di identificare i due lati verticali è necessario fare combaciare le etichette poste sui vertici. L'intervallo di quarta eccedente, che divide l'ottava in due parti identiche, è al centro dello spazio. Questa figura è contenuta in [2].

- 2) La seconda operazione di cui abbiamo bisogno è il passaggio \mathcal{P} dagli insiemi ordinati di classi di note (che rappresentano gli accordi) agli stessi insiemi "non ordinati", ossia vogliamo "dimenticare" l'ordine delle n -uple.² In termini matematici consideriamo l'azione del gruppo \mathcal{S}_n delle permutazioni su un insieme di n elementi. Il dominio fondamentale $\mathcal{D}(\mathcal{P})$ è definito da tutti i punti di \mathbb{R}^n le cui coordinate soddisfano la condizione

$$x_1 \leq x_2 \leq \dots \leq x_n.$$

Componendo le precedenti due operazioni otteniamo l'azione \mathcal{OP} , il cui dominio fondamentale è un prisma che ha come base un simpleso³ $(n-1)$ -dimensionale. Lo spazio quoziente risultante è

$$\mathbb{T}^n / \mathcal{S}_n,$$

ottenuto a partire dal dominio fondamentale identificando i punti della base con quelli della faccia opposta. Ma come si effettua l'incollamento? Definiamo la funzione \mathbb{O} che associa al punto di coordinate (x_1, \dots, x_n) il punto

$$\left(x_2 - \frac{12}{n}, x_3 - \frac{12}{n}, \dots, x_n - \frac{12}{n}, x_1 + 12 - \frac{12}{n} \right).$$

²Una n -upla è una sequenza ordinata di n oggetti. Con la notazione precedente (x_1, \dots, x_n) è una n -upla di note.

³Un *simpleso* è la generalizzazione del concetto di triangolo in più dimensioni. Per esempio, un simpleso tridimensionale è un tetraedro.

È facile vedere che \mathbb{O} è un *automorfismo* (ossia una funzione biiettiva da un certo spazio in se stesso) che permuta ciclicamente i vertici di ciascun simpleso ottenuto sezionando il prisma con il piano $\sum_{i=1}^n x_i = c$, al variare di c in \mathbb{R} . Applicando ripetutamente questa mappa a un certo accordo J otteniamo gli accordi $\mathbb{O}(J), \mathbb{O}^2(J), \dots, \mathbb{O}^{n-1}(J)$, tutti legati tra di loro da una *trasposizione* e con le classi di note che hanno somma costante. Nello spazio euclideo \mathbb{O} rappresenta una rotazione quando il prisma ha dimensione dispari, mentre è una composizione di una rotazione e una riflessione se la dimensione è pari.

- 3) Poiché l'insieme dei vertici del simpleso che hanno somma costante c è *invariante* per l'azione della mappa \mathbb{O} (nel senso che i vertici vengono mandati l'uno nell'altro, non al di fuori del simpleso), questa induce una *partizione* dell'insieme dei vertici in *classi di equivalenza*. È quindi naturale introdurre una nuova azione

$$\mathcal{T} : (x_1, \dots, x_n) \mapsto \sum_{i=1}^n x_i.$$

Per quanto appena detto, questa mappa è ben definita. Il suo dominio fondamentale $\mathcal{D}(\mathcal{T})$ è rappresentato dai punti di \mathbb{R}^n tali che

$$\sum_{i=1}^n x_i = 0,$$

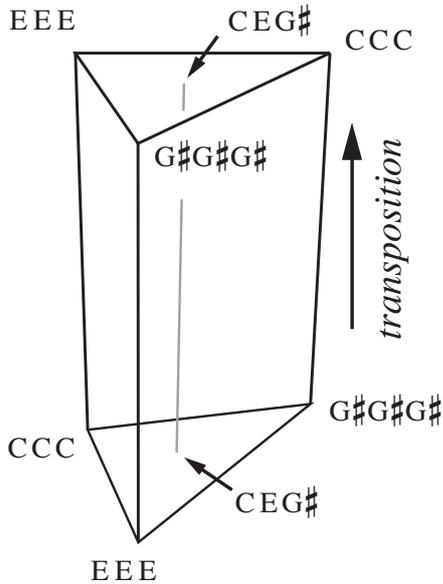


Figura 15: Lo spazio delle triadi \mathbb{X}_3 come spazio quoziente. Prima di incollare le facce del prisma è necessario ruotarne una di 120° , in modo che gli unisoni sui vertici coincidano. Al centro dello spazio troviamo le triadi aumentate. Questa figura è contenuta in [2].

cioè la faccia del prisma corrispondente alla base. Lo spazio quoziente è isomorfo (come spazio vettoriale) a \mathbb{R}^{n-1} .

Considerando la composizione OPT delle operazioni finora definite, lo spazio quoziente risultante è

$$\mathbb{T}^{n-1}/\mathcal{S}_n,$$

ossia la proiezione ortogonale dello spazio quoziente $\mathbb{T}^n/\mathcal{S}_n$ lungo la direzione della trasposizione: si tratta di un simpleso $(n-1)$ -dimensionale modulo "twist". Questo spazio può essere visualizzato anche come un cono che si appoggia sul quoziente di una sfera $(n-2)$ -dimensionale. Infatti un simpleso è omeomorfo ad una palla piena (ossia può essere deformato con continuità in una sfera, per vederlo basta "gonfiarlo") e quindi ad un cono sopra la sfera che racchiude la palla. Quando n è un numero primo il gruppo generato dalla permutazione ciclica non ha punti fissi e lo spazio quoziente risultante $\mathbb{T}^{n-1}/\mathcal{S}_n$ è uno spazio lenticolare.

Una delle proprietà matematiche più interessanti legate alla trasposizione è che questa induce una foliazione (cioè una stratificazione) dello spazio degli accordi $\mathbb{T}^n/\mathcal{S}_n$, in cui i

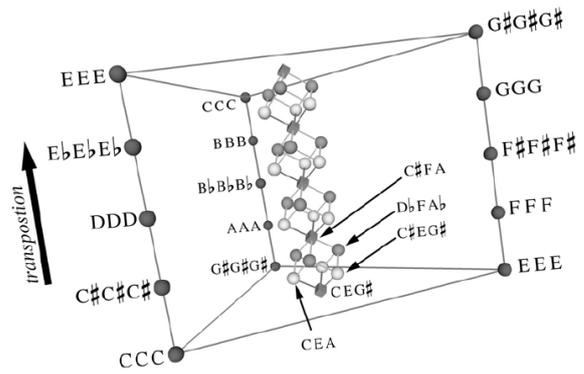


Figura 16: Lo spazio delle triadi \mathbb{X}_3 . In questa raffigurazione si mostra come i diversi generi di triadi occupino diverse posizioni nello spazio. Le triadi aumentate al centro sono rappresentate da cubi, le triadi minori da sfere chiare e quelle maggiori da sfere scure.

fogli, o gli strati, corrispondono ai sottospazi $\mathbb{T}^{n-1}/\mathcal{S}_n$ che identificano lo stesso tipo di accordo (nel senso della trasposizione).

In conclusione, quindi, nel caso speciale delle triadi e delle quadriadi gli spazi corrispondenti risultano essere rispettivamente i seguenti:

$$\mathbb{X}_3 := \mathbb{T}^3/\mathcal{S}_3 \quad \text{e} \quad \mathbb{X}_4 := \mathbb{T}^4/\mathcal{S}_4.$$

Se invece si considerano accordi di n note in generale avremo

$$\mathbb{X}_n := \mathbb{T}^n/\mathcal{S}_n.$$

Osserviamo che per ogni $n \geq 2$ si ha $\mathbb{T}^n = \mathbb{R}^n/(12\mathbb{Z})^n$, e quindi lo spazio quoziente $\mathbb{T}^n/\mathcal{S}_n$ corrisponde allo spazio delle orbite dell'azione del gruppo non abeliano $(12\mathbb{Z})^n \rtimes \mathcal{S}_n$ sullo spazio euclideo n -dimensionale \mathbb{R}^n .

Per $n = 1, 2, 3$ lo spazio \mathbb{X}_n è una varietà differenziabile, mentre se $n \geq 3$ è qualcosa di più generale:

un orbifold.⁴

⁴Ecco l'origine del termine orbifold: "Poco prima dell'inizio del suo corso, nel 1976, Bill Thurston volle introdurre un termine che sostituisse la V -manifold di Satake. La sua prima scelta cadde su manifolded, al fine di distinguerla da manifold (varietà differenziabile). Tuttavia nella comunicazione orale si creavano diversi problemi in quanto risultava difficile distinguere le due parole. L'idea successiva fu quella di foldimani, ma molti matematici si dimostrarono ben poco entusiasti di questa scelta. A un certo punto Bill decise di indire un'elezio-

Tabella riassuntiva: simmetrie, spazi e domini fondamentali

Simmetria	Spazio	Dominio fondamentale
\mathcal{O}	\mathbb{T}^n	$\max\{x_1, \dots, x_n\} \leq \min\{x_1, \dots, x_n\} + 12,$ $0 \leq \sum_{i=1}^n x_i \leq 12$
\mathcal{P}	$\mathbb{R}^n / \mathcal{S}_n$	$x_1 \leq x_2 \leq \dots \leq x_n$
\mathcal{T}	\mathbb{R}^{n-1}	$\sum_{i=1}^n x_i = 0$
\mathcal{OP}	$\mathbb{T}^n / \mathcal{S}_n$	$x_1 \leq x_2 \leq \dots \leq x_n,$ $0 \leq \sum_{i=1}^n x_i \leq 12$
\mathcal{OT}	\mathbb{T}^{n-1}	$\min\{x_1, \dots, x_n\} = x_1,$ $\max\{x_1, \dots, x_n\} \leq x_1 + 12,$ $\sum_{i=1}^n x_i = 0$
\mathcal{PT}	$\mathbb{R}^{n-1} / \mathcal{S}_n$	$x_1 \leq x_2 \leq \dots \leq x_n,$ $\sum_{i=1}^n x_i = 0$
\mathcal{OPT}	$\mathbb{T}^{n-1} / \mathcal{S}_n$	$x_n \leq x_1 + 12,$ $\sum_{i=1}^n x_i = 0,$ $x_1 + 12 - x_n \leq x_{i+1} - x_i \quad \forall 1 \leq i < n$

Gli orbifold in Musica

Gli spazi che abbiamo costruito nelle sezioni precedenti rappresentano un ottimo ambiente per l'analisi del *voice leading*, o *condotta delle parti* (cfr. [1], [3], [5]), cioè della maniera in cui un compositore passa da un accordo all'altro in un dato brano. Un primo esempio, già discusso in precedenza dal punto di vista matematico, di come sia possibile visualizzare questo processo sugli orbifold è riportato in Figura 12. Le frecce indicano due possibili percorsi per muoversi dall'in-

ne "democratica" fra una rosa di parole proposte da alcuni matematici. Tra gli altri, Chuck Giffen suggerì il nome di *origami*, Dennis Sullivan quello di *spatial dollop* (grumo spaziale) e Bill Browder quello di *orbifold*. Dopo un discreto numero di tornate elettorali, in cui si erano eliminate tutte le proposte che avevano totalizzato il più basso numero di voti, ne erano sopravvissute solo quattro: *origami*, *orbifold*, *foldimani* e *V-manifold*. Sfortunatamente anche *orbifold* fu in seguito eliminata. Michael W. Davies prese allora la parola e fece un discorso appassionato, convincendo l'uditorio che gli altri due nomi rimasti erano troppo ridicoli. La parola *orbifold* rientrò quindi in lista e con la votazione successiva si aggiudicò la vittoria." (Cfr. [4, pagg. 5-6]).

tervallo $\{D, E\}$ a $\{D\sharp, E\}$. Percorrere la freccia che non rimbalza sul bordo significa rappresentare la progressione $(D, E) \mapsto (D\sharp, E)$, mentre muoversi tra i due punti del nastro di Möbius rimbalzando sul bordo singolare implica che la progressione seguita sia $(D, E) \mapsto (E, D\sharp)$.



Figura 17: Due *voice leading* musicalmente diversi: $(D, E) \mapsto (D\sharp, E)$ e $(D, E) \mapsto (E, D\sharp)$.

A livello musicale la differenza tra i due percorsi è evidente: quello che non tocca il bordo singolare dello spazio corrisponde a una condotta delle parti dove queste non si scambiano, mentre il secondo caso prevede un incrocio di voci, che chiamiamo *crossed voice leading* (Figura 17). Dunque gli orbifold che abbiamo definito sono effettivamente in grado di distinguere tra queste due scelte. Sebbene creare *voice leading* in cui le voci si incrociano non sia vietato, è certo

sconsigliabile in Musica! Da qui la scelta naturale di “pesare” in maniera differente le due scelte: anche chi non è esperto di Musica riconoscerà che una condotta delle parti dove non vi siano intersezioni risulti più gradevole ed “efficiente” di una che invece presenta degli scambi. Si immagina un brano dove improvvisamente il contralto inizi a cantare la linea melodica del soprano e viceversa: è una scelta possibile, ma sicuramente più ardita e meno immediata del suo *alter ego* libero da intersezioni! Si veda [1] per maggiori dettagli.

L’applicazione di queste rappresentazioni non si limita all’analisi di brani già composti, ma è molto utile anche alla vera e propria speculazione teorica. In Composizione, utilizzare *voice leading* passando in maniera efficiente da un accordo all’altro significa che le note seguono linee melodiche individuali, in modo tale che nessuna voce si sposti esageratamente in seguito ad un cambio armonico (cfr. [6]).

È naturale chiedersi quali siano i *voice leading* più efficienti possibili e cercare di rappresentarli negli spazi che abbiamo definito. In Figura 18 sono rappresentati i *voice leading* di massima efficienza tra quinte diatoniche in *C* maggiore: $(C, G) \mapsto (G, D) \mapsto (D, A) \mapsto \dots$. Le condotte delle parti che massimizzano l’efficienza nel passaggio tra un intervallo e l’altro sono un percorso che passa attraverso il centro del nastro di Möbius. Gli accordi (in questo caso intervalli) non sono legati non in base a una distanza data da proprietà acustiche o altri criteri di tipo fisico, seguono bensì ragioni contrappuntistiche: $\{C, G\}$ è “vicino” a $\{G, D\}$ nella circonferenza di Figura 18, perché per passare da una diade all’altra è sufficiente muovere *C* di un passo diatonico.

Il percorso formato dai *voice leading* di massima efficienza nello spazio delle triadi è rappresentato invece nella Figura 19. Anche in questo caso il percorso occupa interamente il centro dell’*orbifold*. La circonferenza permette di vedere come i *voice leading* siano pensati muovendo diatonicamente una sola nota alla volta: si passa ad esempio da *C* maggiore ad *A* minore (*a* nella notazione della figura) nel modo seguente:

$$\{C, E, G\} \mapsto \{C, E, A\}$$

ossia trasformando il *G* contenuto in *C* maggiore

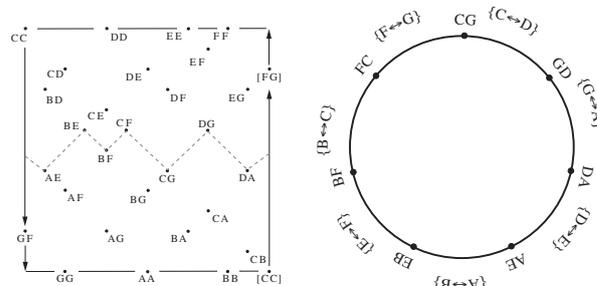


Figura 18: *Voice leading* per quinte diatoniche nello spazio \mathbb{X}_2 . Lo “zig-zag” è irregolare perché lo spazio è costruito utilizzando come unità il semitono; sarebbe regolare se avessimo usato il passo diatonico, cioè le note della scala maggiore.

in *A*. Le triadi “vicine” distano due passi diatonici (per passare da *C* a *A* è necessario muoversi di due note sulla scala di *C* maggiore).

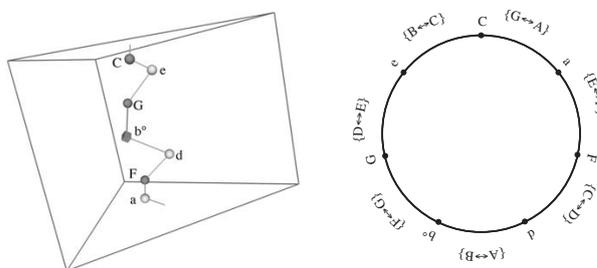


Figura 19: *Voice leading* per quinte efficienti nello spazio \mathbb{X}_3 .

Conclusioni e sviluppi futuri

Abbiamo spiegato come gli *orbifold* siano un modello efficiente per la descrizione degli accordi e al contempo come diverse traiettorie in questi spazi descrivano scelte differenti a livello contrappuntistico, permettendo la convivenza tra l’analisi verticale (accordi) e orizzontale (melodia) delle parti.

Queste geometrizzazioni nascono dall’esigenza di classificare e misurare i *voice leading*, ma è ovviamente possibile considerare altri modelli: un’altra rappresentazione della condotta delle parti è realizzabile attraverso le *trecce*, come la cadenza perfetta in Figura 20. In questo caso l’obiettivo non è quello di misurare la distanza tra *voice leading* di tipo differente, ma piuttosto quello di cercare degli invarianti topologici che permettano di calcolare un indice di complessità

per una quadriade fissata (cfr. [7] per maggiori dettagli).

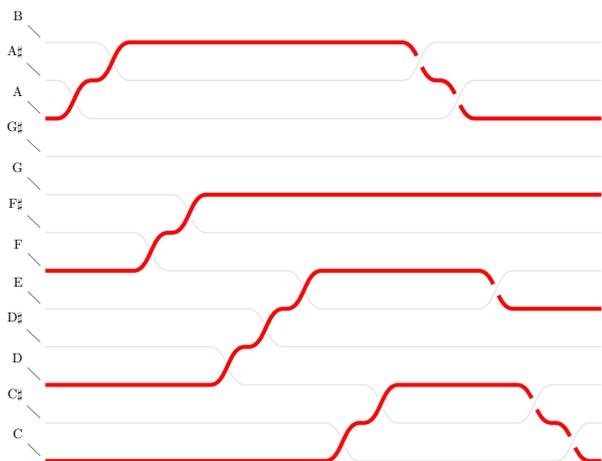


Figura 20: Una rappresentazione della cadenza perfetta $D_7 \mapsto G_7 \mapsto C_7$ come proiezione bidimensionale di una treccia.

In tutt'altra direzione vanno modelli discreti come il *tonnetz*, che ha il vantaggio di generare l'intero spazio delle note su un toro, il cui rivestimento universale è un rettangolo dove le triadi minori e maggiori sono rappresentate da triangoli equilateri (Figura 21).

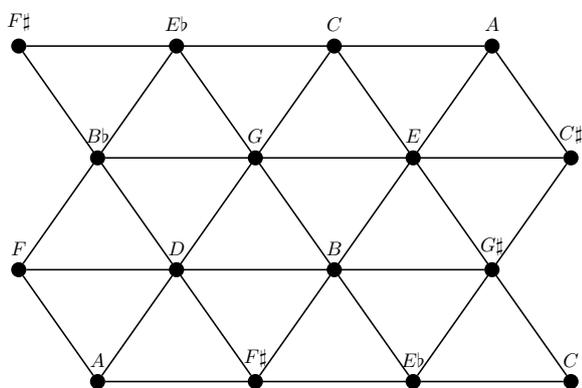


Figura 21: Il tonnetz di Oettingen-Riemann.

Dunque gli orbifold, per quanto siano oggetti estremamente astratti e complessi per dimensioni più grandi di 2, modellano in maniera del tutto naturale lo spazio degli accordi e permettono di distinguere tra *voice leading* differenti.

Lo svantaggio di questo tipo di rappresentazioni è che non esistono direzioni preferenziali: il temperamento equabile obbliga le note ad avere la stessa "massa", cioè non è possibile distinguere a priori un C e un $C\sharp$, data la natura simbolica di questo tipo di analisi.

Un possibile sviluppo potrebbe essere quello di realizzare un modello che, rinunciando ad alcune identificazioni, permetta di descrivere lo spazio delle note e degli accordi come una struttura in cui ogni oggetto musicale sia pesato e conseguentemente generi direzioni preferenziali. In questo modo sarebbe possibile studiare alcune caratteristiche qualitative indotte dall'anisotropia. Chiaramente un modello di questo tipo sarà necessariamente in dimensione bassa, al fine di garantire una rappresentazione comprensibile. Proprio questa è la direzione in cui ci si è mossi rappresentando successioni armoniche attraverso le trecce, che permettono di operare in sole 2 o 3 dimensioni, a prescindere dal numero di note degli accordi che le compongono.



- [1] CLIFTON CALLENDER, IAN QUINN, DMITRI TYMOCZKO: "Generalized Voice-Leading Spaces", *Science* **320** (2008) 346–348.
- [2] DMITRI TYMOCZKO: "The geometry of musical chords", *Science* **313** (2006) 72–74.
- [3] DMITRI TYMOCZKO: *A geometry of music. Harmony and counterpoint in the extended common practice*, Oxford Studies in Music Theory. Oxford University Press, Oxford (2011).
- [4] MICHAEL W. DAVIS: *Lectures on orbifolds and reflection groups*, Transformation groups and moduli spaces of curves (2011). 63–93. Adv. Lect. Math. (ALM), 16. Int. Press, Somerville, MA.
- [5] PISTON WALTER, MARK DeVOTO, ARTHUR JANNERY: *Harmony*, London (Gollancz). 1978.
- [6] DMITRI TYMOCZKO: "Scale theory, serial theory and voice leading", *Music Analysis* **27.1** (2008) 1–49.
- [7] MATTIA G. BERGOMI, ALESSANDRO PORTALURI: *Modes in modern music from a topological viewpoint*, arXiv preprint: 1309.0687 (2013).
- [8] ERIC REGENER: *On Allen Forte's theory of chords*, Perspectives of New Music (1974). 191–212.



Mattia G. Bergomi: Ha conseguito la Laurea Magistrale in Matematica nel 2011 presso l'Università di Milano-Bicocca. Attualmente è dottorando presso l'Università degli Studi di Milano (Informatica) e l'Université Pierre et Marie Curie (Informatica con menzione di specialità in Matematica), sotto la guida di Goffredo Haus e

Moreno Andreatta rispettivamente. I suoi interessi di ricerca consistono principalmente nella creazione di modelli matematici dinamici per l'analisi musicale, utilizzando strumenti tipici della topologia algebrica e dalla topologia algebrica computazionale. Nel contempo si occupa dello studio di alcuni aspetti cognitivi, legati alla percezione della sovrapposizione di stimoli audio e video.

Riccardo D. Jadanza: Ha conseguito la Laurea in Ingegneria Energetica (2007) presso il Politecnico di Torino, la Laurea in Matematica (2009) e la Laurea Magistrale in Matematica (2011) presso l'Università degli Studi di Torino. Attualmente è dottorando in Matematica per le Scienze dell'Ingegneria presso il Politecnico di Torino, sotto la supervisione di Susanna Terracini, Alessandro Portaluri e Vivina Barutello. Sta svolgendo attività di ricerca sul tema della stabilità lineare di una classe particolare di orbite periodiche in sistemi dinamici singolari tipo n corpi, utilizzando tecniche di analisi lineare e non lineare, geometria simplettica e invarianti topologici quali il flusso spettrale e l'indice di Maslov.

Alessandro Portaluri: Ha conseguito la laurea in Matematica nel 1999 presso l'Università di Pisa e il Dottorato in Matematica ed Applicazioni presso l'Università di Genova nel Consorzio Università di Genova-Politecnico di Torino. Dal 2008 al 2012 ha ricoperto la posizione di ricercatore in Analisi Matematica presso l'Università del Salento e dal 2012 ad oggi presso l'Università degli Studi di Torino. I principali interessi di ricerca sono legati alla relazione tra le proprietà geometriche dello spazio delle soluzioni di alcune equazioni differenziali lineari alle derivate ordinarie e parziali e le proprietà spettrali di questi operatori differenziali; ai teoremi indice di tipo Morse-Sturm-Liouville e Atiyah-Singer, all'analisi globale ed in particolare all'omologia di Floer ed infine allo studio di invarianti simplettici di tipo Maslov. Recentemente si è occupato dello studio di sistemi dinamici singolari di molte particelle sviluppando delle tecniche di de-singularizzazione alla McGehee e introducendo delle teorie indice nello studio della stabilità lineare di alcune particolari soluzioni periodiche.

Bosoni, Fermioni e ...

Luigi Martina

Dipartimento di Matematica & Fisica "Ennio De Giorgi" - Università del Salento

La nozione di spin è basilare per comprendere il mondo che ci circonda. Essa emerge in maniera naturale dalle proprietà di simmetria che modellano le teorie fisiche. Tuttavia una certa classe di fenomeni sembra aggirare le loro costrizioni. Solo un'attenta analisi delle condizioni geometriche e della loro realizzazione fisica spiega come cambiare le regole del gioco.

Spin

L'avvento della Meccanica Quantistica (MQ) negli anni '20-'30 del secolo scorso introdusse il concetto di quantizzazione del momento angolare a partire da principi primi molto generali, non più in maniera euristica come aveva fatto Bohr. Come si ricorderà tale impostazione derivava dalla necessità di interpretare i dati della spettroscopia atomica [1]. Dallo studio dell'equazione di Schrödinger per la particella libera nello spazio tridimensionale, oppure per problemi a potenziali centrali, derivava in maniera diretta che lo spettro del momento angolare quadro dovesse essere della forma $\ell(\ell + 1)\hbar^2$, con $\ell \in \mathbb{N}_0$, ovvero ℓ è un intero positivo o nullo. Inoltre la sua terza componente deve assumere solo valori pari ad $m\hbar$, con $-\ell \leq m \leq \ell$, dove si è utilizzata la costante di Planck ridotta $\hbar = \frac{h}{2\pi} = 1.05457 \times 10^{-34} \text{ J sec}$ (esistono numerose trattazioni sull'argomento, a titolo indicativo si fa riferimento all'opera [2]).

Un passo ulteriore venne compiuto postulando, da un punto di vista fenomenologico, che l'elettrone può essere dotato di momento angolare

intrinseco (lo *spin*) semi-intero (sempre in unità di \hbar), il quale non poteva evidentemente essere interpretato come un momento angolare orbitale. Esso fu introdotto da G. Uhlenbeck e S. Goudsmit [3], partendo dalla struttura fine delle righe spettrali, dall'effetto Zeeman anomalo e dal celebre esperimento dei fasci atomici di Stern e Gerlach [4]. Ma ben presto si riconobbe l'esistenza di altre particelle di momento angolare intrinseco semi-intero [5], ma anche di spin intero. Una per tutte si ricordi il *fotone*, cioè il quanto del campo elettromagnetico, oppure l'atomo (o il nucleo) di ${}^4\text{He}$. Infine si collegò il comportamento di particelle identiche di spin intero/semi-intero alle proprietà statistiche di loro insiemi, con l'introduzione delle categorie dei *bosoni* (spin intero) e dei *fermioni* (spin semi-intero). Questi ultimi non possono coesistere nello stesso stato quantico (il celebre *principio di esclusione* di Pauli [5]), dando in tal modo una interpretazione di molti spettacolari fenomeni naturali, quali la stabilità della configurazione elettronica degli elementi atomici, che si esprime nelle loro proprietà chimiche, o il comportamento dei metalli e dei semi-conduttori, alla base di ogni tecnologia sviluppata dall'uomo.

In seguito [6] il principio di esclusione divenne un teorema nell'ambito della più generale Teoria Quantistica dei Campi [7]. Poiché questa teoria possiede intrinsecamente la capacità di descrivere stati con un numero arbitrario di particelle, la corrispondenza tra spin e statistica costituisce uno dei suoi maggiori successi e, in ultima analisi, dà una ragione della stabilità della materia.

In questa sede passeremo in rassegna alcuni

degli argomenti che legano la fenomenologia dello spin delle particelle, e di insiemi di esse, alle idee fondamentali di spazio e tempo (in particolare in ambito non relativistico). Esplorando geometrie meno convenzionali si possono ottenere risultati sorprendenti, nuovi concetti che sono attualmente sottoposti a verifica sperimentale. Faremo un breve resoconto di questi aspetti.

La dittatura della Geometria

L'esame delle proprietà di un certo sistema fisico S può essere affrontata, in primo luogo, chiedendosi in che modo due osservatori distinti (nel senso che sono solidali a due diversi sistemi di riferimento) possono confrontare le loro misure su di esso. La questione è fondamentale, poiché fa riferimento all'oggettività della descrizione dei fenomeni fisici. La risposta consiste nel ritenere due osservatori equivalenti, se esiste una corrispondenza biunivoca tra tutti i rispettivi osservabili su S , i possibili valori delle loro misure, nonché l'uguaglianza delle corrispondenti distribuzioni di probabilità. Se una tale corrispondenza esiste essa è detta *trasformazione di simmetria* per il sistema fisico. Questa idea è generalizzabile ad altri tipi di trasformazioni non strettamente geometriche (con linguaggio gergale i fisici parlano in questo caso di trasformazioni di *gauge*).

Quali siano le trasformazioni geometriche ammissibili può essere una questione estremamente delicata, dipendente dall'ambito di validità che vogliamo assegnare al nostro modello di Fisica. Tuttavia, una sterminata mole di esperienze ci dice che, idealmente lontani da masse e da campi di forze di ogni tipo, le traslazioni spaziali e temporali, le rotazioni e le trasformazioni tra sistemi di riferimento in moto relativo uniforme costituiscono le simmetrie per il sistema formato da una particella, libera appunto. Nel caso della Relatività Speciale le trasformazioni tra sistemi in moto uniforme sono dettate da quelle di Lorentz proprie, piuttosto che da quelle di Galilei in Meccanica Classica. In Relatività Generale, infine, dovremo considerare l'insieme molto più ampio dei diffeomorfismi dello spazio-tempo. La scelta *a priori* delle trasformazioni spazio-temporali determina il tipo di Fisica che potremo sviluppare [8]. È familiare che: 1) ognuna delle trasfor-

mazioni menzionate in precedenza dipende con continuità e differenziabilmente da un certo numero di parametri (ad esempio: 3 parametri per le traslazioni, 3 per le rotazioni, 3 per le traslazioni dipendenti dal tempo ed 1 per le traslazioni temporali); 2) l'applicazione consecutiva di due di tali trasformazioni ne produce una nuova, i cui parametri sono funzioni differenziabili dei parametri delle trasformazioni componenti; 3) tale composizione è associativa, 4) esiste la trasformazione identica, 5) per ogni trasformazione si può trovare la sua inversa. In sintesi esse formano un gruppo di Lie (dal nome del matematico norvegese che studiò la classe dei gruppi di trasformazioni continue). Nel caso di un sistema relativistico il gruppo è detto di Poincaré, di Galilei nella cinematica non relativistica. Una rotazione agisce sulle coordinate dello spazio come una matrice ortogonale di determinante 1, che denotiamo con $R_{\hat{n}}(\theta)$, per indicare che è definita con continuità da un asse di versore \hat{n} , passante per l'origine del sistema di riferimento 0 e da un'ampiezza angolare θ . Le rotazioni costituiscono un esempio di sottogruppo (denotato con $SO(3)$) dei più ampi gruppi di Poincaré o di Galilei.

A rigore la Fisica è fortemente determinata dall'esistenza di gruppi di trasformazioni discrete, quali l'inversione temporale e l'inversione spaziale rispetto ad un polo, che non discuteremo in questa sede.

Ora ricordiamo per sommi capi che la MQ descrive un sistema fisico S in termini di vettori di stato (indicati con Ψ), l'insieme dei quali possiede la struttura di uno spazio vettoriale (di Hilbert) \mathcal{H} sui numeri complessi \mathbb{C} . In tal modo le combinazioni lineari di più vettori di stato sono dei nuovi stati. Da ogni stato Ψ è possibile estrarre tutte, e soltanto, le distribuzioni di probabilità circa gli esiti delle misure di un qualsivoglia osservabile O su S . Gli osservabili sono operatori lineari auto-aggiunti di $\mathcal{H} \rightarrow \mathcal{H}$. Dette probabilità sono proporzionali al modulo quadro del prodotto scalare (definito in \mathcal{H}) dello stato Ψ con i possibili auto-stati dell'osservabile O in questione [2]. Affinché tali distribuzioni di probabilità rimangano le stesse a seguito di una trasformazione di simmetria, la corrispondenza invertibile degli stati $\Psi \rightarrow \Psi'$ e degli osservabili $O \rightarrow O'$ deve essere tale da mantenere invariante

i moduli quadri dei prodotti scalari tra gli stati. Questo equivale ad affermare che ad ogni trasformazione geometrica continua corrisponde un operatore unitario su \mathcal{H} , che realizza quanto sopra.

Esemplificando nel caso delle rotazioni, si avrà la corrispondenza $R_{\hat{n}}(\theta) \rightarrow \mathcal{U}_{\hat{n}}(\theta)$, tale che in ogni punto dello spazio il nuovo stato Ψ' , nel sistema di coordinate ruotato $\vec{r}' = R_{\hat{n}}(\theta) \vec{r}$, sarà dato dalla relazione

$$\Psi'(\vec{r}') = \mathcal{U}_{\hat{n}}(\theta) \Psi(\vec{r}) = D_{\hat{n}}(\theta) \Psi\left(R_{\hat{n}}(\theta)^{-1} \vec{r}\right),$$

con $D_{\hat{n}}(\theta)$ una matrice unitaria di dimensione finita. Nel caso particolare in cui Ψ assuma valori in \mathbb{C} (caso *scalare*), esso viene semplicemente moltiplicata per un fattore di fase globale.

Tecnicamente è molto più conveniente introdurre la nozione di generatori infinitesimi delle rotazioni, i quali costituiscono il piano tangente al gruppo nell'identità. Questo è uno spazio lineare, dotato di un prodotto di commutazione interno antisimmetrico (indicato con $[a, b]$), che gode dell'identità di Jacobi $[a, [b, c]] + [c, [a, b]] + [b, [c, a]] = 0$ per tutti gli elementi a, b, c dell'insieme. Una tale struttura è un'algebra di Lie. Nel caso in considerazione essa coincide con quella del gruppo delle matrici unitarie 2×2 e possiede tre generatori di base, indicati qui con J_i , $i = 1, 2, 3$. Ogni elemento del gruppo è ottenuto *esponenziando* una generica combinazione lineare a coefficienti reali degli elementi J_i . Il significato fisico delle J_i è quello di componenti del momento angolare [2] [8]. Per un sistema costituito da una particella puntiforme con gradi di libertà interni si adotta come *sistema irriducibile di osservabili* quello degli operatori di posizione \vec{Q} , del momento \vec{P} e dello *Spin* \vec{S} . L'analisi delle relazioni di commutazione tra tutti i generatori delle trasformazioni geometriche condurrà alla decomposizione del momento angolare in parte orbitale e di spin

$$\vec{J} = \vec{Q} \times \vec{P} + \vec{S}.$$

Il lettore interessato può reperire trattazioni dettagliate di questa, e altre questioni collegate, nell'imponente letteratura, per esempio [9] e [10] sull'argomento, o in forma semplificata in tutti i manuali di MQ [2]. In sintesi, usando solo le proprietà algebriche del momento angolare, è possi-

Algebra del Momento Angolare

Le componenti del momento angolare J_i con $i = 1, 2, 3$ sono i generatori infinitesimi delle rotazioni. Essi costituiscono una base per uno spazio vettoriale reale 3-dimensionale, dotato di un prodotto interno di commutazione definito dalla relazione

$$[J_i, J_j] = i \epsilon_{ijk} J_k,$$

dove ϵ_{ijk} assume il valore 1 o -1 a seconda che i, j, k siano in una permutazione pari o dispari $(1, 2, 3)$, altrimenti 0. Metodi algebrici consentono di determinare tutte le proprietà del momento angolare e delle rotazioni a partire dalla precedente relazione.

bile affermare che ogni rappresentazione (irriducibile) è determinata dagli autovalori $j(j+1)\hbar^2$ del momento angolare quadro \vec{J}^2 , con $j \in \frac{1}{2}\mathbb{N}_0$, cioè un naturale o un semi-intero positivo. Esso fissa la dimensione dello spazio di rappresentazione a $2j+1$. Esiste inoltre un unico generatore diagonale, che ha autovalori $-j\hbar \leq m\hbar \leq j\hbar$. Infine, per le rotazioni finite corrispondenti alla rappresentazione di momento j , vale la notevole proprietà $\mathcal{U}_{\hat{n}}^{(j)}(2\pi) = (-1)^{2j}$. Essa separa gli stati di momento angolare intero da quelli di momento semi-intero, senza che alcuna combinazione di elementi tra questi due sottospazi sia possibile. Con una tale *regola di super-selezione* si esclude l'interferenza tra stati di spin intero e stati spin semi-intero per un singolo sistema. Infine, per i sistemi costituiti da due sottosistemi, il prodotto (tensoriale) di due rappresentazioni di momenti angolari j_1 e j_2 è decomponibile nella somma diretta delle rappresentazioni di indici $|j_1 - j_2| \leq j \leq j_1 + j_2$, ciascuna presa una sola volta. Conseguentemente ogni rappresentazione si può costruire a partire da quella fondamentale (o spinoriale) con $j = 1/2$.

Tuttavia, quando si costruiscono sistemi di più particelle identiche, cioè tutte con le stesse proprietà intrinseche (massa, spin, carica elettrica ...), la misura di un insieme completo di osservabili compatibili di singola particella non consente di

determinare univocamente lo stato del sistema. Questa ambiguità si chiama *degenerazione di scambio*. Essa si origina dalla possibilità di permutare tra loro le particelle, senza che si osservi alcuna conseguenza nei risultati delle misure sul sistema complessivo. In linea di principio, in Meccanica Classica si ha sempre la possibilità di *seguire* ogni singola particella lungo la propria traiettoria e quindi distinguere gli stati finali. Il concetto di particelle indistinguibili è invece familiare in Meccanica Statistica. Un erroneo computo degli stati relativi al mescolamento di due gas ideali identici conduce al ben noto paradosso di Gibbs, nel calcolo della variazione di entropia. Assunto un tale *principio di indistinguibilità* in MQ, la sua conseguenza è che tutti gli osservabili debbono essere invarianti rispetto a permutazioni. In particolare lo deve essere l'Hamiltoniano, cioè il generatore dell'evoluzione temporale. Questo comporta che la classificazione di uno stato rispetto alla simmetria per permutazioni si deve conservare. Tuttavia queste regole di selezione non sono sufficienti a eliminare dalle varie possibilità certi stati non osservati in Natura. Pertanto si ricorre ad una aggiuntiva regola di *superselezione*, che stabilisce che non sono osservabili effetti di interferenza tra stati di differente simmetria sotto permutazione. Questa affermazione risulta equivalente al più forte postulato *postulato di simmetrizzazione*, che si può sintetizzare affermando

- a) particelle con spin j intero formano soltanto stati simmetrici (esse sono chiamate *bosoni*);
- b) particelle con spin j semi-intero formano soltanto stati anti-simmetrici (esse sono chiamate *fermioni*);
- c) non esistono stati che non appartengano alle categorie precedenti.

In altri termini, indicato con $\Psi(1, 2)$ un vettore di stato per due particelle identiche di spin j , deve valere la relazione

$$\Psi(2, 1) = (-1)^{2j} \Psi(1, 2). \quad (1)$$

L'enunciato, sostenuto e confermato da una enorme mole di osservazioni, implica per i fermioni il menzionato principio di esclusione di Pauli. Esso non riguarda solo gli elettroni ma, come accennato nell'introduzione, anche protoni e neutroni,

così come i quark, che li costituiscono. Atomi e nuclei atomici possono essere trattati come una singola particella, che si comportano da bosoni, o da fermioni, in funzione del numero dei loro costituenti, in accordo alle regole definite precedentemente per la somma di momenti angolari. Ad esempio, Il nucleo dell'atomo di elio si presenta in due specie isotopiche: ${}^4\text{He}$ è un bosone, ${}^3\text{He}$ un fermione. Il loro comportamento a basse temperature è radicalmente diverso [11], a causa del diverso spin. Analoghi eclatanti effetti, dovuti alla natura bosonica dei componenti, si manifestano nei fasci di luce laser, nei superconduttori, nei superfluidi e negli ormai celebri BEC (Bose-Einstein-Condensates) [12][13].

In altri termini, le proprietà spinoriali microscopiche hanno un riflesso macroscopico, riassunto nelle ben note distribuzioni di Fermi - Dirac e di Bose - Einstein. Esse stabiliscono che per un sistema a contatto con un termostato a temperatura T e potenziale chimico μ , la probabilità di trovarlo in uno stato quantistico Ψ_i di energia E_i è dato da

$$n_i = \frac{g_i}{e^{\beta(E_i - \mu)} \pm 1}, \quad (2)$$

(+ per i fermioni, - per i bosoni) dove g_i è il grado di degenerazione dello stato e $\beta = \frac{1}{kT}$, con k la costante di Boltzmann [2].

Nonostante la relazione spin - statistica abbia un impatto così profondo in MQ, per una sua spiegazione che prescindendo da postulati *ad hoc* si è dovuto aspettare lo sviluppo della teoria di campo quantistica relativistica (QFT) (cioè coerente sia con MQ che con Relatività Speciale). I lavori di Fierz e Pauli prima [6], e poi quelli di Lüders e Zumino e di Burgoyne alla fine degli anni '50 [14] [15], dimostrano che, affinché la teoria sia priva di incongruenze, i campi debbono soddisfare a delle regole di anti-commutazione se descrivono particelle di spin semi-intero, invece di commutazione se di spin intero [7]. Tali restrizioni sono equivalenti al principio di Pauli e quindi essi si collocano come uno dei maggiori successi della teoria QFT. Aspetti tecnici e storici su questa parte della Fisica sono magistralmente narrati nel volume di Duck e Sudarshan [16].

Naturalmente il passaggio all'ambito relativistico comporta notevoli mutamenti matematici ed interpretativi. Infatti le rappresentazioni uni-

tarie irriducibili del gruppo di Poincaré esistono solo su spazi di stati di dimensione infinita. Esse sono indicizzate da due numeri quantici: la massa $m \in \mathbb{R}_0^+$, definita dall'invarianza del quadrimpulso, e lo spin $j \in \frac{1}{2}\mathbb{N}_0$, in analogia a quanto visto sopra. Da questo punto di vista esse individuano ogni possibile particella elementare. Se in particolare la particella ha massa $m = 0$, solo la componente dello spin nella direzione del moto può essere definita (*l'elicità*) [7]. Questo fatto ha varie importanti conseguenze di carattere cinematico e dinamico, ma qui preme sottolineare il diverso ruolo svolto dall'osservabile massa tra i due gruppi (Galilei e Poincaré) considerati. Questo rende delicata ogni procedura di limite non relativistico ($c \rightarrow \infty$). La soluzione consiste nell'estendere l'algebra di Galilei introducendo un elemento apposito, l'operatore di massa M , il quale commuta con tutti gli altri elementi. Inoltre, indicando con G_i e P_i i generatori delle trasformazioni di Galileo proprie e delle traslazioni, rispettivamente, l'operatore di massa interviene nella relazione $[G_i, P_j] = i\delta_{ij} M$. A tal proposito V. Bargmann aveva fatto notare [17] che in dimensioni ≥ 3 questa è l'unica possibile estensione centrale per il gruppo di Galilei. Inoltre le rappresentazioni unitarie irriducibili del gruppo di Galilei dipendono da un fattore di fase legato alla massa e per questo sono dette *proiettive*. Questa peculiarità impone una regola di *superselezione*, che vieta l'esistenza di sovrapposizioni di stati con massa differente e, quindi, l'impossibilità per la MQ non relativistica di descrivere particelle instabili. Per poterlo fare, ancora una volta, bisogna ricorrere alla QFT, ma anche si riafferma che i metodi geometrici determinano cosa le nostre teorie possano effettivamente descrivere.

D'altra parte la ricerca di una equazione relativisticamente invariante in forma, che sostituisce quella di Schrödinger, condusse Dirac a introdurre una particolare rappresentazione del gruppo di Poincaré (detta degli spinori di Dirac) a quattro componenti, di massa non nulla, di spin $1/2$ e quindi fermionica, invariante per inversione spaziale. La corrispondente *equazione di Dirac* è invariante anche per un'altra trasformazione discreta, detta *coniugazione di carica*. Essa consente di risolvere il problema degli stati di energia negativa (che affliggono le equazioni relativisticamente invarianti), in quanto essi sono interpre-

tati come stati di energia positiva per particelle con la stessa massa e spin, ma con carica elettrica opposta. Tale corrispondenza permette di introdurre la nozione di *antiparticella*. Poiché le leggi dell'Elettrodinamica sono invarianti rispetto alle trasformazioni di parità e coniugazione di carica (oltre ovviamente a quelle di Poincaré), il formalismo più naturale per sviluppare la teoria quantistica, ad esempio, dell'elettrone e del positrone (sua antiparticella) è quello degli spinori di Dirac. Se in aggiunta si impone l'invarianza dello spinore rispetto alla coniugazione di carica, quindi la particella è neutra e coincide con la sua antiparticella, si parla di spinore di Majorana, il quale per primo elaborò l'idea che i neutrini possano essere descritti da un tale tipo di fermioni. Gli esperimenti al Large Hadron Collider sono nella posizione per verificare, in un prossimo futuro, queste ipotesi [18], nel contesto teorico più ampio delle teorie di campo supersimmetriche sviluppatasi a partire degli anni '70. Invece nell'ambito della Fisica della Materia Condensata esiste una fervida ricerca di fermioni di Majorana in una vasta gamma di sistemi, motivata sia da ragioni di fisica fondamentale che da applicazioni in computazione quantistica [19]. Qui ci basti ricordare che le particelle studiate nello stato solido non sono, ovviamente, elementari, potendo ridurle in elettroni e ioni (ulteriormente frazionabili) del reticolo cristallino. In realtà si tratta di eccitazioni (meta)stabili di sistemi a multi-corpi, chiamate spesso quasi-particelle. Le loro proprietà fisiche (massa, spin carica elettrica) possono però essere ancora catalogate con i metodi grupali descritti in precedenza. In definitiva, si è visto che i postulati della MQ e la struttura geometrica dei gruppi dinamici codificano i tratti essenziali di gran parte della Fisica alla quale siamo interessati. In particolare, sia in un contesto relativistico, che non-relativistico, la nozione di spin e la sua quantizzazione emergono dai principi primi. Tuttavia la relazione tra valori interi (o semi-interi) dello spin con le proprietà degli stati di bosoni/fermioni richiede un'attenzione particolare, strettamente legata all'invarianza relativistica. A tal proposito Feynman osservò [20]: *Sembra che questo sia uno dei pochi casi [...] in cui un risultato possa essere enunciato molto semplicemente, mentre nessuno è riuscito a trovarne la spiegazione in termini semplici ed immediati. ... Questo significa*

probabilmente che non abbiamo una comprensione completa del principio fondamentale che sta alla base del risultato.

Dinamica in 2 dimensioni

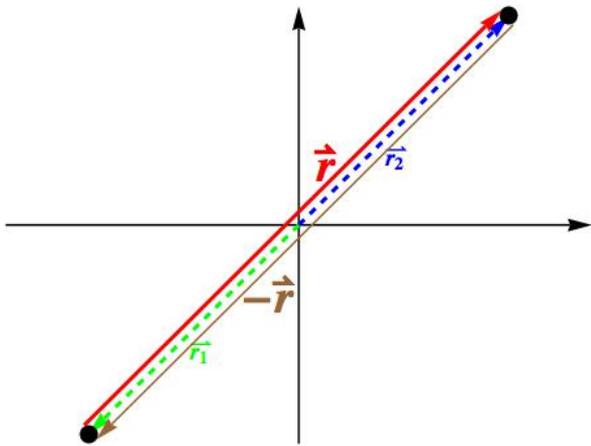


Figura 1: Per un sistema quantistico costituito da due particelle identiche non c'è possibilità di distinguere con alcuna misura quale delle due si trovi nel punto \vec{r}_1 e quale in \vec{r}_2 . I vettori posizione relativa \vec{r} e $-\vec{r}$ rappresentano la stessa configurazione e per la descrizione matematica dello stato del sistema è sufficiente assegnare solo uno di essi.

L'argomento spin-statistica sembrava definitivamente chiuso con l'avvento della QFT, quando Levy-Leblond notò [21] che in 2 dimensioni spaziali le relazioni di consistenza dell'algebra di Galilei ammettevano una seconda estensione centrale, per il minor numero di restrizioni imposte dalle identità di Jacobi. In particolare vale la relazione $[G_1, G_2] = i \kappa \mathbf{1}$, con $\kappa \in \mathbb{R}$ arbitrario, piuttosto che esattamente 0, come nel caso a 3 dimensioni. Per molto tempo questo risultato rimase tra le curiosità matematiche ma poneva il problema, oltre che del significato fisico di κ , quello più generale di una Fisica in 2 dimensioni e degli osservabili in essa. A tal proposito si osservi che il gruppo delle rotazioni nel piano ($SO(2)$) è abeliano, quindi lo spettro dello spin è continuo. D'altra parte Leinaas e Myrheim [22] e, indipendentemente, Wilczek [23] hanno compiuto una riflessione critica sul principio di esclusione, osservando una sua dipendenza dalle dimensioni spaziali.

Infatti, ritornando a riflettere sulle procedure concrete di permutazione, il primo passo è

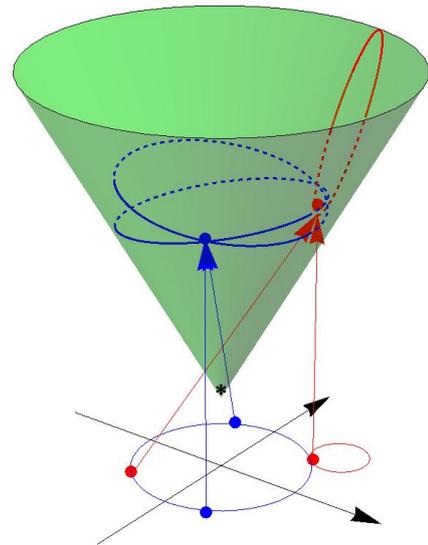


Figura 2: Una rappresentazione di \mathbb{R}^2/S_2 : punti dello stesso colore nel piano sono identificati sul cono. Differentemente dalle curve chiuse che non includono l'origine (rosse), quelle che la racchiudono almeno una volta sono doppiamente mappate sul cono. Poiché il vertice è escluso, le curve di questa classe non possono essere deformate con continuità in quelle della prima, né tra di loro se hanno un differente grado di omotopia.

identificare lo spazio delle configurazioni accessibile ad un sistema di due particelle identiche. Se si è in d dimensioni, lo spazio complessivo è $\mathbb{R}^d \times \mathbb{R}^d$, dal quale possiamo fattorizzare un termine \mathbb{R}^d riservato al centro di massa del sistema. Nel restante \mathbb{R}^d delle variabili relative i punti $\vec{r} = \vec{r}_1 - \vec{r}_2$ e $-\vec{r}$ corrispondono alla stessa configurazione fisica, visto che non è possibile distinguere in alcun modo le due particelle (si veda Figura 1). In altri termini deve esistere una funzione continua $p : \mathbb{R}^d \rightarrow \mathbb{R}^d/S_2 \subset \mathbb{R}^d$, dove S_2 è il gruppo delle permutazioni di ordine 2. Essa associa una stessa immagine ad ogni coppia di punti simmetrici. Generalizzazioni a N particelle portano allo spazio \mathbb{R}^d/S_N . Nel caso estremo mono-dimensionale ($d = 1$) la funzione che si sta cercando è $r \rightarrow |r|$, quindi lo spazio effettivo è $\mathbb{R}/S_2 = \mathbb{R}_0^+$: la semiretta reale positiva incluso lo 0. Passando a $d = 2$, per ogni punto \vec{r} si può dare la distanza dall'origine $\vec{r} \rightarrow r = |\vec{r}|$ e poi assegnare la direzione della retta che lo

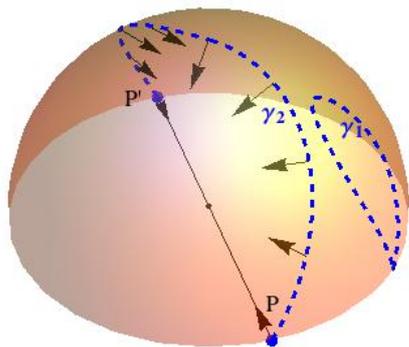


Figura 3: Rappresentazione dello spazio $\mathbb{P}\mathbb{R}^2$, i punti sull'equatore debbono essere fatti coincidere. La curva γ_1 si può contrarre con continuità in un punto, la curva γ_2 no, in quanto nei punti equatoriali diametralmente opposti la normale è diretta in versi opposti. Bisogna percorrere la stessa curva un'altra volta per ritornare alla condizione di partenza. Questa curva doppia è contraibile in un punto.

congiunge al simmetrico. Per ogni assegnato r , l'insieme $\frac{\vec{r}}{r}$ è una circonferenza unitaria centrata nell'origine, dove la funzione p ha la stessa immagine su punti opposti. Costruttivamente si fanno coincidere i punti della semicirconferenza inferiore con quelli opposti della superiore, compresi i punti estremi della stessa. Questa è ancora una circonferenza, che rappresenta l'insieme di tutte le direzioni del piano ed è indicato con il simbolo $\mathbb{P}\mathbb{R}^1$, per tener memoria del suo significato. In conclusione lo spazio disponibile per il sistema di due particelle nel piano è $\mathbb{R}_0^+ \times \mathbb{P}\mathbb{R}^1$, quindi un cono (si veda Figura 2) con il vertice in corrispondenza della configurazione di coincidenza $\vec{r} \equiv \mathbf{0}$. Il cono è localmente isometrico a \mathbb{R}^2 , ma nel suo vertice ha curvatura infinita. Escludere tale singolarità equivale a considerare particelle impenetrabili, ipotesi che adotteremo nel seguito. Ripetendo lo stesso tipo di costruzione nel caso $d = 3$, lo spazio delle configurazioni relative è dato da $\mathbb{R}^3/S_2 = \mathbb{R}_0^+ \times \mathbb{P}\mathbb{R}^2$, dove $\mathbb{P}\mathbb{R}^2$ è l'insieme delle direzioni nello spazio. La sua costruzione è simile alla precedente: si consideri la sfera unitaria, si facciano coincidere tutti i punti dell'emisfero inferiore con gli opposti di quello superiore, si proceda analogamente con

i punti dell'equatore diametralmente opposti e, infine, con i punti estremi di questa semicirconferenza (Figura 3). Ancora una volta $\vec{r} \equiv \mathbf{0}$ è un punto di singolarità per la curvatura dello spazio e, pertanto, sarà rimosso. La geometria di $\mathbb{R}_0^+ \times \mathbb{P}\mathbb{R}^1$ (o di $\mathbb{R}_0^+ \times \mathbb{P}\mathbb{R}^2$) è meglio specificata se si tracciano delle curve chiuse su di esso e si cerca di deformarle con continuità le une nelle altre. Si verifica facilmente che questa operazione non è sempre possibile. Se si costruisce una curva chiusa sul cono, che si avvolge n volte attorno al vertice, potrà essere deformata in un'altra con lo stesso numero di spire (o dello stesso grado di omotopia), ma mai in un'altra con un numero inferiore (o superiore) di spire senza che si attraversi il vertice. Poiché n è un intero arbitrario, si dice che $\mathbb{R}_0^+ \times \mathbb{P}\mathbb{R}^1$ è connesso infinite volte. Differentemente $\mathbb{R}_0^+ \times \mathbb{P}\mathbb{R}^2$ è solo doppiamente connesso, cioè esistono curve che sono contraibili con continuità in punto, oppure altre che circondando almeno una volta la singolarità. Tuttavia, per il carattere non orientabile di $\mathbb{P}\mathbb{R}^2$, si dimostra che percorrendo due volte una stessa curva, si ottiene di nuovo una curva contraibile. In tal modo si costruiscono curve contraibili ad un punto in tutto lo spazio delle configurazioni per la coppia di particelle identiche. Questa conoscenza delle proprietà geometriche è importante per procedere alla quantizzazione, in quanto la funzione d'onda su tali spazi dipende sensibilmente da esse. Questo è detto nell'ottica della rappresentazione di Schrödinger della MQ. Quella di Heisenberg è stata studiata in vari lavori [24] e, inoltre, si è dimostrata l'equivalenza dei risultati ottenuti con il metodo dell'Integrale Funzionale [25].

Nel caso mono-dimensionale si può imporre la condizione che la densità di corrente di probabilità sia nulla nell'origine. Questo equivale ad imporre che $\partial_r \Psi|_{r=0} = \eta \Psi|_{r=0}$, essendo η una costante arbitraria. È facile verificare che se $\eta = 0$ si deve avere $\Psi(x) = \Psi(-x)$, quindi simmetrica (caso bosonico). Invece per $1/\eta = 0$ si ha $\Psi(x) = -\Psi(-x)$ (caso fermionico). Variando η , queste situazioni estreme sono interpolate da un continuo di simmetrie intermedie. Almeno in questo particolare caso la dicotomia bosoni/fermioni non sussiste. Inoltre si è appreso che la simmetria della funzione d'onda può essere codificata da opportune condizioni imposte nel

punto di singolarità dello spazio.

Sulla scorta di questo esempio, si deve scrivere l'appropriata equazione di Schrödinger sugli spazi curvi bi- e tri-dimensionali effettivamente accessibili e con le singolarità precedentemente determinate. Per tener conto di questo fatto è chiaro che i generatori delle traslazioni, proporzionali alle componenti del momento lineare nel caso euclideo, devono essere opportunamente modificati. Daremo qualche idea in proposito, senza essere rigorosi, ma sapendo che le procedure matematicamente corrette portano ai risultati che esporremo.

Dai manuali di MQ [2] si ricorderà che de Broglie associò ad ogni particella libera di quantità di moto ben definita un'onda piana, secondo la relazione

$$\vec{p} \rightarrow \psi = e^{i\frac{\vec{p}}{\hbar} \cdot \vec{r}}. \quad (3)$$

Effettuando una traslazione spaziale lungo uno spostamento $\Delta\vec{r}$, l'onda si trasformerà in

$$T_{\Delta\vec{r}} \psi(\vec{r}) = e^{i\frac{\vec{p}}{\hbar} \cdot (\vec{r} + \Delta\vec{r})} = \psi(\vec{r}) e^{i\frac{\vec{p}}{\hbar} \cdot \Delta\vec{r}}. \quad (4)$$

Ma se piuttosto che essere nell'abituale spazio euclideo, la particella si muovesse in uno spazio curvo, la formula precedente per la traslazione deve venir corretta. Ispirandosi agli esempi di propagazione ondulatoria (meccanica o ottica) tra mezzi continui diversi, si può facilmente arguire che la prima modifica da apportare è nella fase dell'onda. Quindi per $\Delta\vec{r}$ sufficientemente piccolo (al limite infinitesimo) la (4) viene modificata in

$$T'_{\Delta\vec{r}} \psi(\vec{r}) = \psi(\vec{r}) e^{i(\frac{\vec{p}}{\hbar} - \vec{a}) \cdot \Delta\vec{r}} = \psi(\vec{r} + \Delta\vec{r}) e^{-i\vec{a} \cdot \Delta\vec{r}}, \quad (5)$$

dove il vettore $\vec{a}(\vec{r})$ è una funzione continua e derivabile di \vec{r} , da determinarsi in relazione alla geometria scelta. A scanso di equivoci, si osservi che la quantità introdotta non è un potenziale nel senso abituale del termine: 1) non è una quantità scalare, ma ha tante componenti quante le dimensioni dello spazio in cui si opera; 2) agisce direttamente sulla derivata prima della funzione d'onda e non sulla sua derivata seconda. In altri termini \vec{a} descrive la maniera con cui viene effettuato il *trasporto parallelo* di un vettore sullo spazio curvo scelto. Per esempio, traslando un vettore \vec{v} tangente ad una sfera di una quantità infinitesima $\Delta\vec{r}$, anch'essa tangente alla sfera, è

facile accorgersi che esso acquista una componente ortogonale alla superficie. Tale componente va compensata con una opportuna proiezione sulla sfera, proporzionale al vettore stesso. Questa operazione di proiezione è regolata da un opportuno campo \vec{a} . A livello infinitesimo il trasporto parallelo è descritto da una generalizzazione della derivata direzionale, che prende il nome di *derivata covariante* e assume la forma $D_i = \partial_{r_i} - \iota a_i(\vec{r})$ nella direzione i -esima [26]. Questi operatori costituiscono i generatori infinitesimi delle traslazioni sullo spazio curvo, in analogia con quanto avviene nello spazio euclideo abituale. Pertanto la trasformazione infinitesima indicata in (5) si può estendere lungo qualunque curva continua γ nello spazio considerato, ottenendo

$$T'_\gamma \psi(\vec{r}) = \psi'(\vec{r}) e^{-\iota \int_\gamma \vec{a} \cdot d\vec{r}}, \quad (6)$$

dove $\psi'(\vec{r})$ è la funzione d'onda traslata lungo γ , senza tener conto della curvatura dello spazio. Questa funzione quindi è a un solo valore (monodroma) nello spazio, esattamente come lo era nello spazio euclideo di pari dimensioni.

Se ora si eseguono due trasporti paralleli infinitesimi in successione lungo due direzioni diverse e, poi, si procede nello stesso ordine con i loro opposti, è facile vedere che non necessariamente si torna nel punto di partenza, come nel caso euclideo, ma in generale si ottiene una deviazione infinitesima pari a

$$f_{ij} = \iota [D_i, D_j] = \partial_{r_i} a_j - \partial_{r_j} a_i \quad (7)$$

dove f_{ij} è detto *tensore di curvatura* dello spazio. Esso è l'analogo di un campo magnetico $\vec{B} = \nabla \times \vec{a}$, associato ad un certo potenziale vettore \vec{a} . In effetti, il *potenziale* \vec{a} dà significato all'idea che una permutazione di particelle identiche (una trasformazione discreta) sia equivalente ad uno scambio di posizione lungo cammini continui in spazi curvi. In Geometria Differenziale il campo \vec{a} è chiamato *connessione*.

Si è visto che la curvatura f_{ij} deve essere nulla ovunque eccetto che nel/nei punti di singolarità, dove non è definita. Tale richiesta corrisponde a considerare campi \vec{a} che siano localmente irrotazionali, ma non conservativi. L'esempio familiare è preso ancora una volta dall'elettromagnetismo classico, ricordando che il campo magnetico

generato da una corrente filiforme soddisfa la relazione $\nabla \times \vec{B} = 0$ in tutti i punti eccetto che sul filo, dove diverge. Ispirandosi quindi alla legge di Biot - Savart (ora $\vec{a} \equiv \vec{B}$), nel caso bi-dimensionale si può porre $\vec{a} = \frac{\Phi}{2\pi r} \hat{\tau}$, dove $\hat{\tau}$ è il versore ortogonale al raggio vettore e preso nel senso anti-orario, mentre Φ è un parametro reale arbitrario, che fissa il valore della circuitazione. La corrispondente curvatura è $f_{1,2} = \Phi \delta(\vec{r})$ e, naturalmente, si ha $\int f_{1,2} dr^2 = \Phi$. Se ora si percorre un'arbitraria curva chiusa Γ sul cono, la funzione d'onda per una coppia di particelle identiche acquista un fattore di fase della forma

$$\exp\left(-i \int_{\Gamma} \vec{a} \cdot d\vec{r}\right) = \exp(-2\pi i n_{\Gamma} \Phi), \quad (8)$$

dove n_{Γ} è un intero che conta quante volte la curva concatena il vertice. Inoltre è evidente che il risultato è indipendente dal cammino scelto, ma solo dall'ordine di concatenazione (o grado di omotopia) della curva: essa può essere deformata arbitrariamente con continuità, purché non attraversi il vertice. Quindi Φ è un parametro dipendente solo dalla specie particellare ed il fattore di fase (8) è ineliminabile dalla teoria. Per $\Phi = 0$ la fase vale 1 e quindi la funzione d'onda è simmetrica: essa descrive bosoni. Per $\Phi = 1/2$ la funzione è antisimmetrica e descrive lo scambio di fermioni. Ma non vi è alcuna ragione per limitarsi a tali valori e quindi anche in 2 dimensioni si può realizzare un continuo di simmetrie intermedie. Per come è stata introdotta Φ è un caso particolare di *fase geometrica* o *fase di Berry* [27] [28]. Senza entrare in dettagli qui si ricorda che l'esempio più spettacolare di fase geometrica è costituita dal cosiddetto *Effetto Aharonov - Bohm*, osservato sperimentalmente da A. Tonomura nel 1986 [29]. Esso ha permesso di verificare che in MQ il flusso magnetico è un osservabile fisico indipendente dall'effettiva azione del campo magnetico sulle cariche elettriche. Esattamente lo stesso effetto della singolarità conica discussa in precedenza.

Tuttavia, il potenziale \vec{a} non è stato determinato in maniera univoca e, in effetti, si è lavorato solo su un esempio che permette di produrre la forma del tensore di curvatura $f_{1,2}$ desiderata, ma non sappiamo se esso sia unico. Ancora una volta ci soccorre l'elettromagnetismo, il quale afferma che a parità di \vec{B} (qui identificato con

$f_{i,j}$) il potenziale vettore \vec{a} è noto a meno di una *trasformazione di gauge*, aggiungendo ad esso il gradiente di una arbitraria funzione scalare differenziabile. Per l'esempio considerato, scelta la funzione $\Lambda = \frac{\Phi\phi}{2\pi}$ dove ϕ è la variabile azimutale nel piano, si ottiene la trasformazione $a'_i = a_i - \partial_i \Lambda \equiv 0$, per $\vec{r} \neq 0$. L'equazione di Schrödinger non contiene più potenziali \vec{a} . Ma se si crede che essa sia una legge fondamentale della Natura e che debba descrivere la stessa Fisica indipendentemente dalla gauge scelta, allora è chiaro che un cambiamento di \vec{a} deve comportare necessariamente una trasformazione della funzione d'onda ψ . Specificatamente, nel caso visto appena sopra, la trasformazione corrispondente è $\psi'(\vec{r}) = \psi(\vec{r}) \exp\left[i\frac{\Phi\phi}{2\pi}\right]$, che ora è polidroma in ϕ . Quindi il prezzo, del richiedere che il sistema ridotto soddisfi l'equazione di Schrödinger libera, è dato da condizioni al bordo inusuali. Tuttavia, queste soluzioni consentono di calcolare la terza componente di momento angolare con i metodi abituali, trovandola pari a $m = \text{intero} + \frac{\Phi}{2\pi}$. In ogni caso non riusciamo a sopprimere il contributo statistico dato da Φ .

Ripetendo gli stessi argomenti in 3 dimensioni spaziali, e ricordando quanto detto a proposito della deformabilità continua di ogni curva chiusa in $\mathbb{P}R^2$ ad un singolo punto, possiamo concludere che l'integrale curvilineo in (8) è sempre nullo. Questo significa che la fase acquistata dalla funzione d'onda è 1, dopo che una delle due particelle abbia compiuto una sola rotazione o al più due attorno all'altra. Quindi a seguito di una loro permutazione la fase acquistata dalla loro funzione d'onda è pari a ± 1 . Si ritrova in questo modo il risultato classico: in 3 dimensioni i vettori di stato per sistemi di particelle identiche sono soltanto o simmetrici (bosoni), o antisimmetrici (fermioni). Con tutte le conseguenze illustrate in precedenza.

Si osservi che il campo \vec{a} è strettamente connesso con il sistema delle due particelle identiche ed evolve con esso. In altri termini \vec{a} non è un campo che si propaga autonomamente come quello elettromagnetico, ma rimane strettamente collegato al sistema e, quindi, è caratteristico della specie componente. Questo fatto ha suggerito che in 2 dimensioni spaziali possa esistere un oggetto composto *particella - flusso*, che interagisce con le altre in maniera analoga con quanto avviene

nell'effetto Aharonov-Bohm. Allora, la funzione d'onda di multi-particella acquista un fattore di fase caratteristico della specie Φ , ma per il resto arbitrario, a seguito dello scambio di due di esse. Poiché tale parametro assume valori continui entro due limiti estremi, che corrispondono alle fasi per bosoni e fermioni, rispettivamente, esso caratterizza particelle battezzate *anyoni* da Wilczek [23], i quali godono di una *statistica anyonica*, a volte detta anche frazionaria. Inoltre lo stesso autore ha suggerito come implementare la statistica frazionaria nell'ambito di una teoria di campo locale, abeliana e a lungo range detta di Chern-Simons, che interagisce in modo minimale con campi di materia bosonici. Questa costruzione ha dato luogo a numerosi lavori teorici e varie generalizzazioni [30].

Inoltre, utilizzando i metodi della quantizzazione geometrica [8] si è dimostrato che la relazione scoperta da Levy-Leblond per il gruppo di Galilei in 2 dimensioni consente di costruire dei modelli classici di particelle cariche in campi magnetici, che sono l'analogo degli anyoni. In particolare il parametro di deformazione κ è proporzionale alla fase Φ e si comporta in maniera duale rispetto al campo magnetico [31][32]. Questi modelli sono utilizzati per studiare l'Effetto Hall Anomalo [33], che si realizza in materiali con proprietà magnetiche peculiari, l'Effetto Hall di Spin [34], nel trasporto di correnti di spin, e nell'effetto Hall ottico [35], che concerne la rifrazione trasversale rispetto al piano formato dalla direzione di propagazione di un fascio di luce e la normale alla superficie del mezzo su cui incide.

Fisica degli anyoni

Passando a studiare la dinamica di due, o più, anyoni identici, le loro proprietà statistiche possono essere convenientemente descritte proiettando le loro traiettorie in un piano (spazio, tempo), con l'unica informazione aggiuntiva di quale, tra due traiettorie, una *passi sotto l'altra*. Il linguaggio matematico appropriato per descrivere tali proiezioni è quello detto del *gruppo di treccia* [36] [37]. L'operazione elementare del gruppo σ_j corrisponde a prendere la traiettoria j e scambiarla con la $j+1$ *passandole sopra*, lasciando tutte le altre invariate. Il prodotto tra le operazioni elementari è definito eseguendole sequenzialmente.

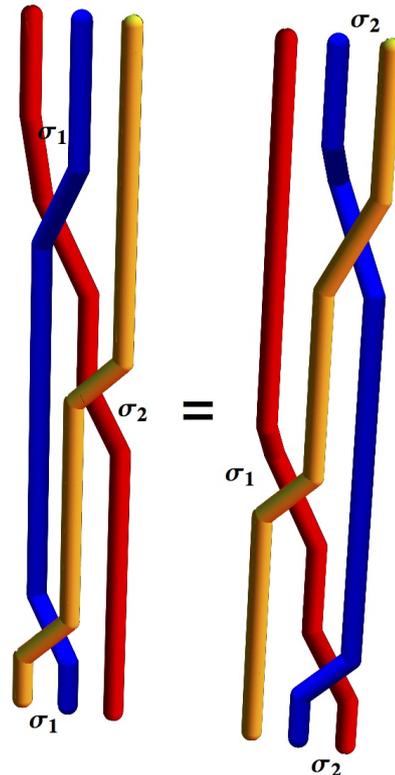


Figura 4: Uguaglianza dei prodotti $\sigma_1\sigma_2\sigma_1$ e $\sigma_2\sigma_1\sigma_2$

Si ottengono allora delle identità molto semplici:

$$\sigma_j\sigma_k = \sigma_k\sigma_j \quad |j - k| \geq 2$$

e

$$\sigma_j\sigma_{j+1}\sigma_j = \sigma_{j+1}\sigma_j\sigma_{j+1},$$

quest'ultima detta equazione di Yang-Baxter (si veda Figura 4). Come abbiamo già osservato sopra, se il moto di queste particelle avvenisse in 3 dimensioni spaziali, allora deve sussistere la relazione

$$\sigma_j^2 = 1 \quad \forall i,$$

che corrisponde al fatto che lo scambio di due particelle produce al più un cambio di segno del suo stato. L'effetto di queste relazioni aggiuntive consiste nel ridurre il gruppo a quello abituale delle permutazioni S_N di N particelle identiche. Se invece ci limitiamo ad una dinamica in 2 dimensioni spaziali, solo le prime due identità debbono essere considerate. Se, come di consueto, si cercano delle rappresentazioni unitarie del grup-

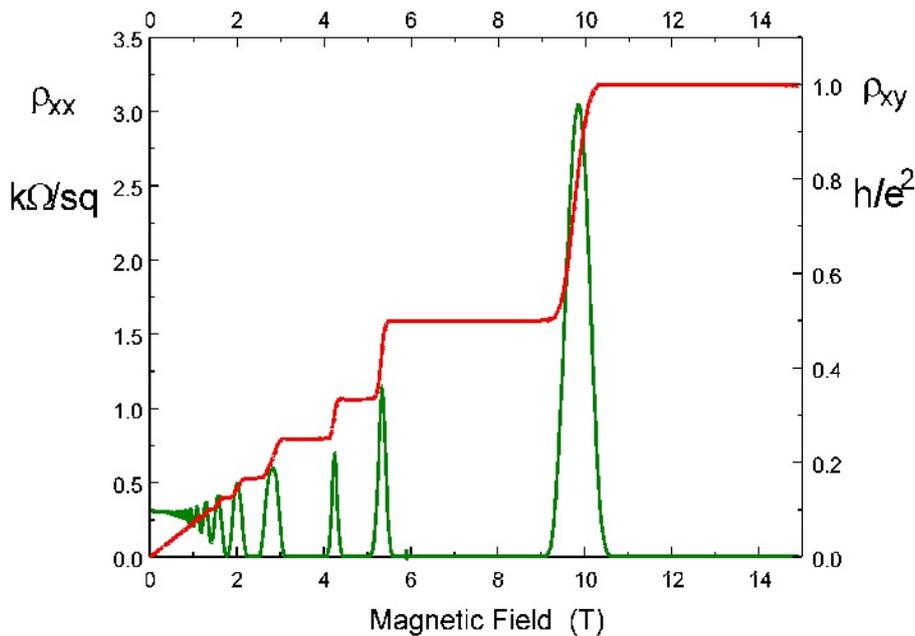


Figura 5: Grafico della conducibilità trasversa (o Hall) (tr. rossa) e longitudinale nell'IQHE (tr. verde). I dati sono riportati in unità della costante di Klitzing $R_K = h/e^2 = 25\,812.807557(18)\,\Omega$. Al variare del campo magnetico si osservano dei plateaux in corrispondenza di $1, 1/2, 1/3, \dots$. La resistività longitudinale è sempre trascurabile, eccetto che per i valori di B in cui il Livello di Landau transita attraverso l'energia di Fermi. Contemporaneamente si ha un brusco passaggio da un plateau all'altro, corrispondente alla diminuzione del numero dei livelli di Landau occupati da elettroni, con densità di stati proporzionale a $\Phi(B)/\Phi_0$, dove $\Phi_0 = h/(2e)$ è il quanto di flusso magnetico.

po da esse generato, allora la più semplice non banale è fornita da $\sigma_i = e^{i\theta}$, per ogni j e $\theta \in \mathbb{R}$. Ci ritroviamo a dover ammettere quindi che è possibile l'esistenza di particelle, il cui vettore di stato cambia per un fattore di fase arbitrario (θ) per effetto dello scambio di due di loro. Un esempio teorico di tale fenomeno è costituito dai sistemi composti particelle-flusso discussi poco sopra [39]. Tuttavia è facile vedere che non esistono semplici regole per stabilire gli stati di occupazione per particelle indipendenti, né è ovvio come generalizzare le distribuzioni statistiche di Bose-Einstein o Fermi-Dirac al caso degli anyoni. Pertanto il problema di determinare lo spettro di Hamiltoniani che coinvolgono anyoni interagenti è difficile [40]. Inoltre la loro meccanica statistica è ancora in fase di sviluppo [41]. In breve, si è seguita l'idea che il comportamento statistico di un gas ideale di bosoni/fermioni possa essere analogo a quello di un gas di particelle classiche interagenti con un potenziale fittizio a due corpi attrattivo/repulsivo, monotonicamente crescente/decescente. Per gli anyoni tale *potenziale statistico* dipende dal parametro θ , che interpola con continuità i precedenti, ma perdendo in generale la monotonia. Quindi emergono fasi quasi-bosoniche, che possono essere trasmutate in fasi quasi-fermioniche. Nel caso di anyoni non abeliani, questo quadro è ulteriormente complicato dall'accoppiamento del potenziale con gli stati di isospin [42].

Naturalmente si vorrebbe trovare una concreta realizzazione fisica di tali particelle. A questo scopo si dovrebbe ricordare che alcuni sistemi quantistici possono effettivamente ridurre la loro dimensionalità, se la loro dinamica è vincolata dalla presenza di gap (salti) di energia tra lo stato fondamentale e quelli eccitati troppo grandi rispetto a possibili eccitazioni esterne. Sotto opportune condizioni questo vincolo può condurre ad una riduzione della dinamica ad uno spazio di configurazioni di dimensioni ridotte.

L'esempio fisico concreto di questa situazione è costituita dalla scoperta, negli anni '80, dapprima dell'Effetto Hall Quantistico Intero (IQHE) in un gas bidimensionale di elettroni (nel 1980 da K. Von Klitzing, Nobel per la Fisica 1985) [43] e poi, molto più pertinente alla nostra discussione, dell'Effetto Hall Quantistico Frazionario (FQHE) (nel 1982 da Tsui, Stormer, Nobel nel 1998 con R. Laughlin, Gossard) [44] (si veda la Figura 5 per l'IQHE e la 6 per il FQHE). Come è noto questi fenomeni si manifestano quando un gas bidimensionale di cariche elettriche, tipicamente gli elettroni presenti all'interfaccia tra due semiconduttori opportunamente drogati, è sottoposto all'azione ortogonale di intensi campi magnetici esterni e a temperature sufficientemente basse.

La dinamica rilevante di questi sistemi si svolge nel piano ortogonale al campo magnetico, tra elettroni fortemente interagenti coulombianamente. Il campo magnetico confina gli elettroni

Integer and Fractional Quantum Hall Effects

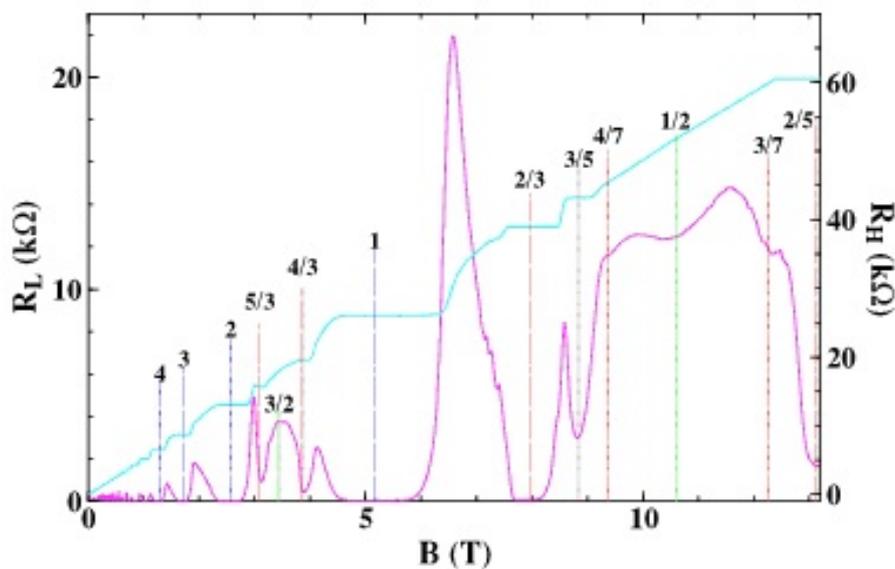


Figura 6: Grafico della resistività trasversa Hall (celeste) e longitudinale (rosa) nell'FQHE. Quando la resistenza Hall è misurata in funzione del campo magnetico, vengono osservati dei plateaux non solo in corrispondenza di valori multipli interi di h/e^2 , ma anche frazionari. In corrispondenza dei plateaux la resistenza longitudinale si annulla. La misura deve essere condotta a valori del campo magnetico più elevati rispetto al caso intero, ma soprattutto a temperature molto più basse dell'ordine delle decimesime di Kelvin. La figura è ripresa da <http://www.sp.phy.cam.ac.uk/>.

in livelli discreti di energia, detti *Livelli di Landau* [2]. Eventuali effetti dovuti al disordine possono essere riassunti in un potenziale aggiuntivo di singola particella. Nel limite di grandi campi magnetici e basse temperature gli elettroni occupano il più basso livello di Landau (LLL) e hanno tutti lo stesso spin. Se gli effetti di disordine sono trascurabili, l'Hamiltoniano si riduce alle pure interazioni coulombiane. Questo processo si chiama di proiezione su LLL e genera dinamicamente una geometria a due dimensioni. Tutti gli altri livelli sono a energie praticamente infinite: già il gap dovuto all'effetto Zeeman è proibitivo per tali elettroni. Questo non è un problema perturbativo, in quanto non esistono scale di energia rispetto ad un livello fondamentale. Inoltre il grado di degenerazione dello stato fondamentale è infinito. La presenza del campo magnetico comporta che la funzione d'onda a molti elettroni assuma tanti zeri per unità di area, quanti sono i quanti di flusso magnetico che la penetrano. In tal modo si creano una sorta di vortici, che presentano un deficit di carica elettronica pari a νe associati a dei quanti di flusso magnetico incorporati rigidamente nello strato elettronico. D'altra parte vortici ed elettroni tendono ad attrarsi e l'energia dovuta all'interazione elettrostatica è significativamente ridotta se essi vengono strettamente legati. La situazione ricorda la schermatura di cariche elettriche nei conduttori ordinari, ma ora essa è quantizzata in unità di flusso magnetico. Laughlin propose [45] una funzione

d'onda collettiva per gli elettroni. Essa permette di descrivere uno stato FQHE contenente ν vortici nella distribuzione di elettroni, ciascuno portatore di una unità di flusso magnetico, carica elettrica e/ν e spin $1/\nu$. Questa funzione d'onda fu trovata usando metodi variazionali di minimizzazione dell'energia, in un ottimo accordo con i dati sperimentali. La sua distribuzione di probabilità può essere interpretata come quella di un plasma bi-dimensionale di una sola specie ionica: i composti elettrone-vortice, che hanno carica e spin frazionari, cioè sono anyoni. Quindi la funzione d'onda di Laughlin descrive la condensazione di particelle anyoniche in un nuovo stato della materia, che, in analogia con il IQHE, dà luogo al FQHE [38, 39]. Si può dimostrare che qualunque deviazione dagli interi dispari nei rapporti tra numero di elettroni e vortici porta a considerevoli costi energetici per il sistema complessivo. La creazione di vortici addizionali, per esempio a causa di un incremento del campo magnetico, è energeticamente sfavorita e il plasma si comporta come un fluido incompressibile. Da qui il fatto che possano esistere particelle di spin $1/3$, $2/5$, $3/7$ e moltissimi altri valori, come si può arguire dalla Figura 6. Come queste eccitazioni si combinino in varie gerarchie di particelle è interpretato da vari modelli, tra i quali citiamo quello dei Fermioni Composti [46]. Comunque la formazione di gap di energia, la presenza di quasiparticelle cariche e la loro localizzazione in potenziali residuali costituiscono gli ingredienti

per spiegare le proprietà di trasporto illustrate in Figura 6. Infine va menzionata l'intensa attività sperimentale per verificare queste ipotesi, in particolare la natura anyonica delle quasi-particelle qui introdotte. In particolare alcuni autori [47] hanno sviluppato un dispositivo interferometrico che sembra dimostrare l'esistenza di tali particelle, anche se la questione è tuttora controversa.

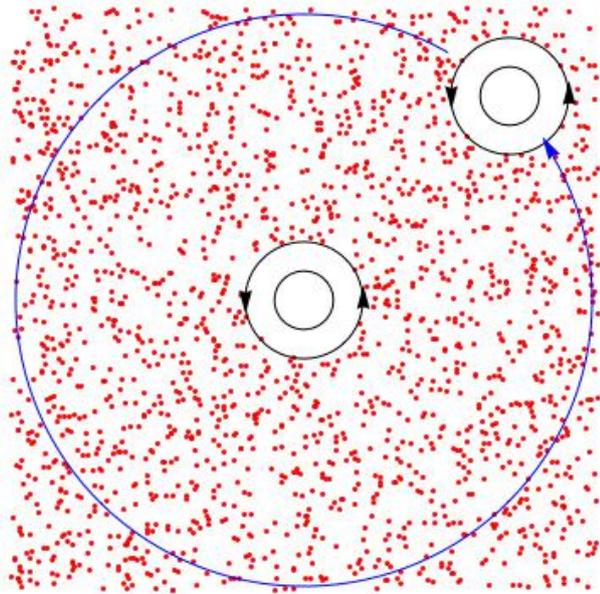


Figura 7: Nel mare di elettroni fortemente interagenti si generano dei vortici associati a quanti di flusso magnetico e con un deficit di carica elettrica media. Essi sono descritti dalle funzioni d'onda di Laughlin e si possono pensare come particelle di carica frazionaria e/ν . Se un vortice gira attorno ad un altro, la sua funzione d'onda acquista una fase di Berry proporzionale al deficit di carica e si comporta come una particella con spin $1/\nu$.

Conclusioni

Per chiudere questa breve rassegna ricordiamo che le particelle in natura si presentano con proprietà fortemente dettate dalla geometria. In particolare lo spin di una particella è strettamente legato alle rappresentazioni unitarie irriducibili del gruppo di Poincaré, o di Galilei nel caso non si volesse tener conto degli effetti relativistici. La conclusione è che nello spazio ordinario esistono solo bosoni e fermioni. Questo fatto era già da tempo condiviso dalla comunità dei fisici, anche sulla scorta dei lavori dei matematici, Weyl e Wi-

gner in primis, durante la prima metà del secolo scorso. Tuttavia scendere di dimensionalità spaziale ha significato esplorare nuove possibilità e non, semplicemente, ridurre la difficoltà di analisi dei problemi. Spesso aspetti del tutto nuovi e sorprendenti si sono rivelati in dimensioni ridotte, sia nella Matematica, che nella Fisica e, per darne un quadro sufficientemente chiaro, è stato necessario abbandonare pregiudizi consolidati. Infine, bisogna ancora sottolineare che spesso raffinati concetti matematici trovano la loro realizzazione nei più disparati ambiti della Fisica con modalità imprevedibili: difficilmente nella prima metà del '900 Emil Artin avrebbe potuto immaginare che il suo gruppo di treccia, ispirato dai lavori a maglia della moglie, potesse essere di una qualche utilità nella Fisica della Materia e nella Meccanica Quantistica.



- [1] M BORN: *Atomic Physics*, Dover Pub. (1969).
- [2] C COHEN-TANNOUDJI, B DUI, F LALOË: *Quantum Mechanics*, Wiley, New York (1977).
- [3] G E UHLENBECK, S GOUDSMIT *Naturwissenschaften* **47** (1925), 953.
- [4] W GERLACH, O STERN: "Der Experimentelle Nachweis der Richtungsquantelung im Magnetfeld", *Zeitschrift für Physik* **9** (1922) 349.
- [5] W Pauli Ueber den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren *Z. Physik* 311925765
- [6] W PAULI: "The connection between spin and statistics", *Physical Review* **58** (1940) 716.
- [7] S WEINBERG : *The quantum theory of fields - I*, Cambridge University Press (1996).
- [8] J M SOURIAU : *Structure of dynamical systems. A symplectic view of physics*, *Progress in Mathematics*, Birkhauser , Boston (1997).
- [9] R GILMORE: *Lie Groups, Lie Algebras and some of their applications*, John Wiley & Sons Inc. (1974).
- [10] J A DE AZCARRAGA, J M IZQUIERDO: *Lie Ggroups, Lie algebras, cohomology and some applications in Physics*, Cambridge Univ. Press (1998).
- [11] A J LEGGETT: " Theory of a superfluid Fermi Liquid I. General Formalism and static properties", *Physical Review* **A140** (1965) 1869.
- [12] 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", *Physical Review Letters* **75** (1995) 3969.

- [13] 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** (1995) 198.
- [14] G LÜDERS, GERHART AND B ZUMINO: "Connection between Spin and Statistics", *Physical Review* **110** (1958) 1450.
- [15] N BURGOYNE: "On the connection of spin with statistics", *Nuovo Cimento* **8** (1958) 607.
- [16] I DUCK, E C G SUDARSHAN: *Pauli and the Spin-Statistics Theorem*, World Scientific (1998).
- [17] V BARGMANN: "On unitary ray representations of continuous groups", *Annals of Mathematics* **59** (1954) 1.
- [18] D E MORRISSEY, T PLEHN, T M P TAIT: "Physics searches at the LHC", *Physics Reports* **515** (2012) 1.
- [19] C W J BEENAKKER: "Search for Majorana Fermions in Superconductors", *Annual Review of Condensed Matter Physics* **4** (2013) 113.
- [20] R P FEYNMAN, R B LEIGHTON, M SANDS: *The Feynman Lectures on Physics*, Addison-Wesley Pub. C. (1970). III, Cap 4
- [21] J M LEVY-LEBLOND : " Galilei group and Galilean invariance", *Group Theory and Applications II* (Academic Press (1971)) 221.
- [22] J M LEINAAS, J MYRHEIM: "On the theory of identical particles", *Il Nuovo Cimento* **37 B** (1977) 1.
- [23] F WILCZEK: "Quantum Mechanics of Fractional-Spin Particles", *Physical Review Letters* **49** (1982) 957.
- [24] J M LEINAAS, J MYRHEIM: "Quantum Theory for Identical Particles", *Int. J. Modern Physics A* **5** (1991) 2573.
- [25] Y. -S. WU: "General Theory for Quantum Statistics in Two Dimensions", *Phys. Rev. Lett.* **52** (1984) 2103.
- [26] S P NOVIKOV, A T FOMENKO B A DUBROVIN: *Geometria Contemporanea*, Editori Riuniti (1994).Vol I
- [27] M V BERRY: "Quantal Phase Factors Accompanying Adiabatic Changes", *Proc. R. Soc. Lond. A* **392** (1984) 45.
- [28] A BOHM, A MOSTAFAZADEH, H KOIZUMI, Q NIU, J ZWANZIGER: *The Geometric Phase in Quantum Systems*, Springer (2003).
- [29] A TONOMURA: *The Quantum World Unveiled by Electron Waves*, World Scientific (1998).
- [30] G DUNNE: *Self-Dual Chern-Simons Theories*, Springer (1995).
- [31] C DUVAL, P A HORVATHY: "Exotic Galilean symmetry in the non-commutative plane and the Hall effect", *J. Phys. A: Math. Gen.* **34** (2001) 10097.
- [32] P HORVATHY, L. MARTINA, P. STICHEL: "Exotic galilean symmetry and non-commutative mechanics ", *SIGMA* **6** (2010) 60.
- [33] T JUNGWIRTH, Q NIU, A H MACDONALD: "Anomalous Hall Effect in Ferromagnetic semiconductors", *Phys. Rev. Lett.* **90** (2002) 207208.
- [34] S MURAKAMI, N NAGAOSA, S -C ZHANG: "Dissipationless Quantum Spin Current at Room Temperature", *Science* **301** (2003) 1348.
- [35] K Y BLOKH: "Geometrodynamics of polarized light: Berry phase and spin Hall effect in a gradient-index medium", *J. OPTICS A: Pure and Applied Optics* **11** (2009) 094009. arXiv:0903.1910 , physics.optics.
- [36] L H. KAUFFMAN: *Knots and Physics*, World Scientific (1991).
- [37] C NAYAK, F WILCZEK: "2n-quasihole states realize 2n+1-dimensional spinor braiding statistics in paired quantum Hall states", *Nuclear Physics B* **479** (1996) 529.
- [38] WILCZEK F.: *Fractional Statistics and Anyon Superconductivity*, World Scientific (1990).
- [39] STERN A: "Anyons and the quantum Hall effect: A pedagogical review", *Annals of Physics* **323** (2008) 204.
- [40] A FEIGUIN, S TREBST, A W W LUDWIG, M TROYER, A KITAEV, Z WANG, M H FREEDMAN: "Interacting Anyons in Topological Quantum Liquids: The Golden Chain", *Physical Review Letters* **98** (2007) 160409.
- [41] A KHARE: *Fractional statistics and quantum theory*, World Scientific (2005).
- [42] F MANCARELLA, A TROMBETTONI, G MUSSARDO: "Statistical mechanics of an ideal gas of non-Abelian anyons", *Nuclear Physics B* **867** (2013) 950.
- [43] K v KLITZING, G DORDA, M PEPPER : "New Method for High-Accuracy Determination of the Fine-Structure Constant Based on Quantized Hall Resistance", *Physical Review Letters* **45** (1980) 494.
- [44] D.C. TSUI, H.L. STORMER, A.C. GOSSARD: "Two-Dimensional Magnetotransport in the Extreme Quantum Limit", *Physical Review Letters* **48** (1982) 1559.
- [45] LAUGHLIN R.B.: " Anomalous quantum Hall effect: an incompressible quantum fluid with fractionally charged excitations", *Phys. Rev. Lett.* **50** (1983) 1395-1398.
- [46] O HEINONEN: *Composite Fermions*, World Scientific (2003).
- [47] F CAMINO, W ZHOU, V GOLDMAN: "Realization of a Laughlin quasiparticle interferometer: Observation of fractional statistics", *Physical Review B* **72** (2005) 075342.

Luigi Martina: Luigi Martina è professore associato di fisica Teorica presso l'Università del Salento. La sua attività di ricerca riguarda l'applicazione di metodi gruppali per lo studio di sistemi fisici nonlineari.

Fermioni di Majorana in sistemi a due dimensioni?

Pasquale Scarlino

Kavli Institute of Nanoscience - TU Delft (NL)

I fermioni di Majorana sono descritti da soluzioni dell'equazione di Dirac per cui vi è una identificazione tra particella ed antiparticella. A livello fondamentale mancano verifiche sperimentali della loro esistenza in qualità di particelle elementari. Lavori teorici hanno però suggerito la presenza di un loro analogo in sistemi allo stato solido. Recentemente, un primo importante segno della loro esistenza sembra esser stato raggiunto.

Origine dei Fermioni di Majorana

In meccanica quantistica relativistica, un fermione, come l'elettrone, è descritto da soluzioni a valori complessi dell'equazione di Dirac. Tuttavia nel 1937, Ettore Majorana, guidato da un principio di semplicità ed eleganza matematica, propose un'equazione modificata, contenente solo quantità reali ed avente soluzioni anch'esse reali (i cosiddetti Fermioni di Majorana). A prima vista, potrebbe sembrare una sottile discussione matematica, ma queste considerazioni hanno profonde implicazioni fisiche. Infatti ad ogni particella corrisponde una specifica antiparticella il cui stato è descritto applicando l'operatore

di coniugazione di carica alla funzione d'onda ψ che la trasforma nella sua complessa coniugata ψ^* . Segue, quindi, che i fermioni proposti da Majorana, essendo descritti da soluzioni reali, sono particelle coincidenti con le proprie antiparticelle ed i relativi operatori di creazione e annichilazione devono essere anch'essi coincidenti e quindi Hermitiani ($\gamma = \gamma^\dagger$).

A questo punto, alcune domande sorgono spontanee. Esistono questi fermioni di Majorana, perché sono così interessanti, e cosa hanno di diverso rispetto ai fermioni normali, detti di Dirac? Per spiegarlo è necessaria una breve premessa. Un'antiparticella è caratterizzata da medesima massa, stesso spin ma opposta carica elettrica rispetto alla sua controparte. Di conseguenza, coincidendo con la propria antiparticella, un eventuale fermione di Majorana dovrebbe necessariamente avere carica nulla. Tuttavia questo non è sufficiente, ed infatti il neutrone ha una sua antiparticella, l'antineutrone. Inizialmente Majorana suggerì l'applicazione della sua equazione per descrivere i neutrini, a quel tempo ancora non identificati, ma l'esistenza di una netta distinzione tra neutrino ed antineutrino è ancora controversa.

Sorprendentemente, negli ultimi anni, i fermioni di Majorana hanno suscitato notevole interesse

Quasiparticelle

In sistemi di stato solido accade spesso che gruppi di elettroni correlino il loro comportamento in seguito ad interazioni reciproche o con l'ambiente circostante legate alla loro carica o spin. Tali interazioni danno vita a nuove "particelle" effettive, spesso chiamate "quasiparticelle" poiché caratterizzate da numeri quantici ben definiti e specifiche caratteristiche fisiche che possono differire di gran lunga da quelle dei loro costituenti fondamentali. Sono esempi di quasiparticelle i fononi, quanti di vibrazione nei cristalli, come anche i plasmoni nei metalli, gli eccitoni nei semiconduttori, i polaritoni nelle cavità, i magnoni nei materiali magnetici, le coppie di Cooper nei superconduttori, ecc.

in un settore totalmente differente rispetto alla fisica delle particelle elementari, la fisica teorica dello stato solido. In sistemi materiali oggi sperimentalmente realizzabili, esisterebbero infatti fermioni di Majorana non come particelle fondamentali, bensì come entità effettive (quasiparticelle) descriventi il comportamento globale di elettroni correlati. Il concetto di eccitazione coincidente con la propria antiparticella non è totalmente estraneo al settore. Un tipico esempio è infatti costituito dall'eccitone, una quasiparticella costituita da un elettrone ed una lacuna legati assieme che soddisfa però una statistica di tipo bosonico.

Fermioni di Majorana in sistemi a stato solido

Come può un sistema allo stato solido, costituito da fermioni tradizionali, contribuire ad originare fermioni di Majorana, e cosa serve per poterli osservare? Gli elementi richiesti sono essenzialmente due: superconduttività e magnetismo.

In un superconduttore, in seguito all'interazione con i fononi circostanti, gli elettroni risultano essere appaiati nelle cosiddette coppie di

Fermioni senza spin?

Fermioni e bosoni sono generalmente associati a particelle aventi rispettivamente spin semintero o intero. Com'è quindi possibile che si parli di fermioni di Majorana riferendosi a particelle prive di spin? L'utilizzo del termine fermione in questo caso differisce rispetto alla fisica particellare. Infatti, nei sistemi materiali considerati, il termine deriva dalle regole di anticommutazione $\{\gamma_j, \gamma_k^\dagger\} \equiv \gamma_j \gamma_k^\dagger + \gamma_k^\dagger \gamma_j = \delta_{jk}$ che continuano a valere. Tuttavia questi fermioni di Majorana risultano essere privi non solo di carica elettrica ma anche di massa e spin. Inoltre $\gamma^2 = 1$, invece del tradizionale $\gamma^2 = 0$ valido per i fermioni tradizionali (Principio di esclusione di Pauli).

Riguardo alla statistica, non è poi quella fermionica classica ma si ha una statistica non-abeliana. A tal riguardo, va detto che in semiconduttori a bassa dimensionalità, gli elettroni in moto in presenza di un alto campo magnetico possono cooperare alla creazione di quasiparticelle aventi carica frazionaria e statistica intermedia tra quella fermionica e la bosonica. Per tale motivo ci si riferisce ad essi col termine di "anioni", a loro volta distinti in "abeliani" o "non abeliani", in base alla commutabilità o meno delle varie operazioni di scambio di posizione.

Cooper. Attorno al livello di Fermi ($E = 0$, cioè nel mezzo della gap superconduttiva) le eccitazioni (quasiparticelle) risultano essere sovrapposizioni (neutre) di elettroni e lacune, soddisfacendo così la nostra prima richiesta, che però è condizione necessaria ma non sufficiente.

Non è detto infatti che l'operatore di annichilazione di queste quasiparticelle sia hermitiano e coincidente con il rispettivo operatore di creazione ($\gamma = uc_\sigma^\dagger + u^*c_\sigma = \gamma^\dagger$). Il motivo è che, escluse rare eccezioni, la superconduttività ha tipicamente luogo con appaiamento in "simmetria di onda s", cioè tra elettroni con spin opposto che formano un singoletto. Per ragioni di simme-

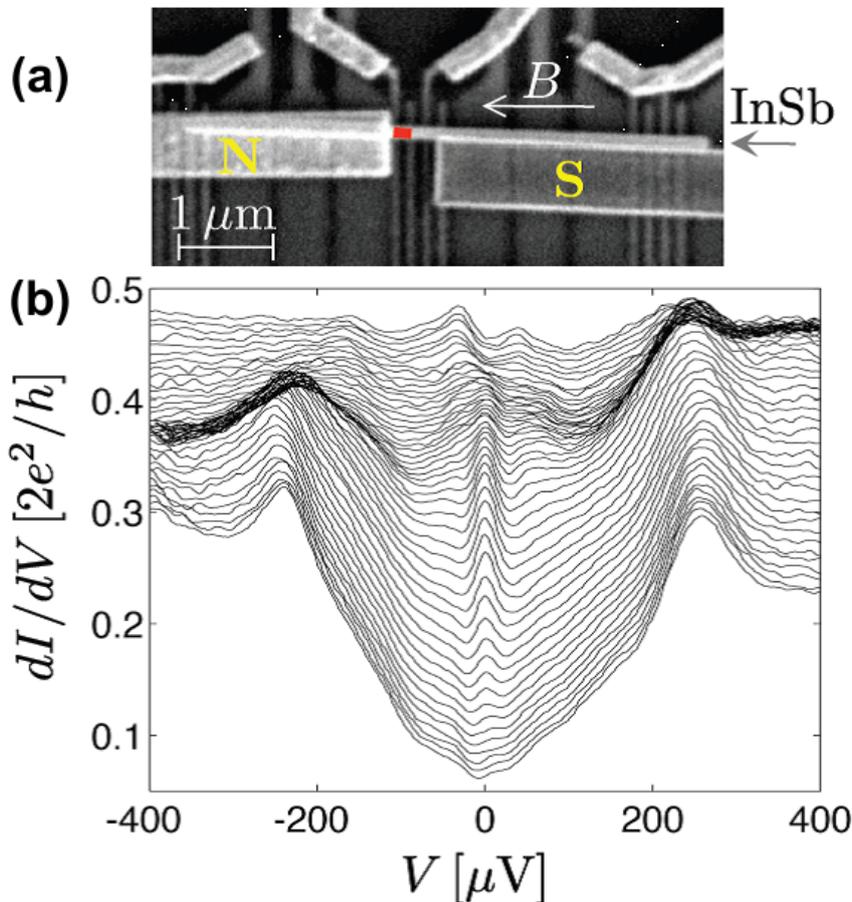


Figura 1: a) Immagine al microscopio elettronico del sistema utilizzato a Delft per evidenziare fermioni di Majorana. Un nanowire di InSb (alto spin-orbita) è ricoperto a sinistra da Ti/Au (contatto metallico-N) ed a destra da NiTiN (superconduttore-S), mentre nella zona intermedia vi sono contatti metallici più sottili, utilizzati per modificare localmente il potenziale elettrochimico (e quindi la densità elettronica) nella porzione di nanowire sovrastante. b) Conduttanza differenziale registrata per campi magnetici tra 0 mT e 500 mT in step da 10 mT. Si possono nettamente discernere due picchi attorno a $\pm 250 \mu\text{V}$ che rappresentano la gap superconduttiva indotta per effetto di prossimità nel nanowire. Un altro picco è visibile attorno a $0 \mu\text{V}$ per campi magnetici compresi tra circa 100 mT e 400 mT; tale picco indica la creazione di una coppia di quasiparticelle di Majorana localizzate all'estremità del nanowire. Adattata da [4].

tria è indispensabile invece che l'appaiamento di Cooper abbia luogo tra elettroni e lacune col medesimo spin (stato di tripletto) e funzione d'onda orbitalica antisimmetrica, di tipo *onda p*. Sfortunatamente però non si possiede tuttora una verifica sperimentale della concreta esistenza di un superconduttore di questo tipo, ed inoltre pare che nei sistemi candidati questo appaiamento sia particolarmente fragile rispetto al disordine. Ad ogni modo il messaggio fondamentale è che serve un meccanismo che forzi gli elettroni ad avere medesimo spin, rimuovendo la degenerazione di spin.

La svolta decisiva nel settore si è avuta nel 2008, ad opera di due teorici dell'università di Pennsylvania, Fu e Kane, che suggerirono la combinazione di un materiale con alto effetto spin-orbita (un cosiddetto "isolante topologico") ed un comune superconduttore di onda-s per realizzare un'interazione di appaiamento equivalente, dal punto di vista formale, a quella di tipo-p [1]. La prossimità dei materiali può determinare infatti una "superconduzione indotta" quando coppie di Cooper provenienti dal superconduttore attraversa-

no l'interfaccia conservandosi integre; ma l'ingrediente fondamentale della ricetta proposta da Fu e Kane è costituito dalla forte interazione spin-orbita che rompe la simmetria tra stati con diverso spin favorendo un appaiamento parallelo di tipo-p. Sfortunatamente, anche questa volta sussiste un problema. Pur avendo effettuato enormi passi avanti, la ricerca sperimentale sugli isolanti topologici non risulta infatti essere ancora abbastanza matura da consentire una realizzazione sperimentale di questa proposta teorica.

Un ulteriore passo in avanti è stato però conseguito nel 2010 dal gruppo teorico di Das Sarma, dell'Università del Maryland, che fece notare come nella ricetta proposta da Fu e Kane fosse possibile sostituire l'isolante topologico semplicemente con un semiconduttore ad alto potenziale di spin-orbita in presenza di un opportuno campo magnetico [2]. Il punto di forza della più recente proposta di Jason Alicea, del Californian Institute of Technology, consiste invece nel ricercarli in un sistema ibrido facilmente realizzabile dal punto di vista sperimentale: un nanowire (cioè un nanofilo monodimensionale) fatto di semicon-

Effetto spin-orbita

In fisica atomica, l'interazione spin-orbita contribuisce a spiegare la struttura fine dei livelli energetici atomici. La sua origine è legata al moto di un elettrone attorno al nucleo. Nel sistema di riferimento dell'elettrone, infatti, il campo elettrico nucleare \mathbf{E} dà luogo ad un campo magnetico effettivo proporzionale a $\mathbf{p} \times \mathbf{E}$, dove \mathbf{p} è il momento dell'elettrone.

Più in generale un elettrone immerso in un potenziale elettrico $V(\mathbf{r})$ risente di un'interazione spin-orbita della forma: $H_{SO} = -\frac{e\hbar}{4mc^2} \boldsymbol{\sigma} \cdot (\mathbf{p} \times (\nabla V))$, dove m è la massa dell'elettrone e $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ sono le matrici di Pauli. Questo contributo energetico può essere visto come l'effetto di un campo magnetico interno effettivo che, agendo sullo spin elettronico comporta uno splitting energetico dei due stati di spin, anche in assenza di campo magnetico esterno.

In un sistema allo stato solido, un elettrone è soggetto al campo elettrico determinato dal reticolo cristallino e conseguentemente si ha un accoppiamento spin-orbita che comporta una modifica della relazione di dispersione ("struttura a bande") che descrive lo spettro energetico dell'elettrone nel cristallo in funzione del suo momento. Per elettroni confinati nel gas bidimensionale di un'eterostruttura i campi elettrici generanti l'effetto spin-orbita presentano due principali origini. Una connessa direttamente alla possibile natura asimmetrica della cella cristallina (che comporta la rottura della simmetria di inversione spaziale) e che dà origine al contributo detto di "Dresselhaus"; l'altra dovuta ad un'asimmetria, lungo la direzione di crescita, dell'eterostruttura nel suo complesso e che dà origine al contributo di spin-orbita detto di "Rashba". Di conseguenza, l'entità dell'interazione spin-orbita risulta dipendere fortemente dal particolare materiale considerato.

duttori con alto spin-orbita quali InSb o InAs e posto in prossimità di un superconduttore in presenza di un campo magnetico non troppo elevato applicato lungo l'asse del nanowire [3].

Prima evidenza sperimentale

Al momento della pubblicazione di questi lavori teorici, il gruppo di Kouwenhoven, dell'Università di Delft, possedeva tutta la tecnologia necessaria per una realizzazione sperimentale dell'ultimo sistema suggerito. Difatti già nel 2008 un esperimento sull'interazione tra nanowires di InAs e superconduttori era stato effettuato nel medesimo gruppo, che successivamente aveva studiato anche nanowire di InSb.

La difficile questione da affrontare riguardava invece quale fosse il modo migliore per identificare la presenza dei fermioni di Majorana in qualità di stati aventi energia, carica elettrica e spin nulli. In altre parole stati "*as close to Nothing as anything can be*", come è solito sottolineare Sergey Frolov, postdoc nel gruppo di Delft. L'immediata conseguenza è che campi elettrici o magnetici, non troppo elevati, non risultano avere alcuna

influenza, sull'energia totale di tali eccitazioni, che rimane ancorata attorno allo zero energetico, rendendo alquanto complicata una loro identificazione sperimentale.

La soluzione ideata dal gruppo di Kouwenhoven risulta particolarmente semplice e geniale al contempo. Consiste nel connettere il nanowire tra un normale contatto metallico ed un contatto superconduttivo (Fig.1a). Iniettando elettroni dal contatto metallico nel nanowire fino a raggiungere la porzione con superconduzione indotta, si può così risalire alla presenza o meno delle quasiparticelle ricercate in quanto le previsioni teoriche associano loro una caratteristica riduzione della resistenza del sistema complessivo. Di conseguenza ci si aspetta di osservare un picco stabile e riproducibile nel grafico della conduttanza in condizioni di assenza di differenza di potenziale (cioè nell'intorno del livello di Fermi) al variare di parametri sperimentali quali potenziale elettrochimico, campo magnetico, temperatura. Questo corrisponde esattamente a quanto osservato dal gruppo di Kouwenhoven (Fig.1b) [4].

Sebbene a rigor di logica, questa evidenza sperimentale risulti essere ancora indiretta, la meti-

colosità e completezza dello studio dello spazio dei parametri evidenzia come sia effettivamente necessaria la presenza di ciascuno degli ingredienti suggeriti dalla ricetta di Alicea e Das Sarma (spin-orbita, campo magnetico, superconduzione indotta) affinché lo stato ad energia zero misurato sperimentalmente, ed identificato con un singolo fermione di Majorana, possa esistere. Non resta che un'ultima domanda. Se confermata la loro esistenza, cosa fare di questi fermioni così speciali? Ebbene, la natura topologica di questi stati conferisce loro speciali proprietà di stabilità rispetto a perturbazioni locali rendendoli sistemi hardware ideali per il controllo e l'immagazzinamento di informazione quantistica. Potrebbero quindi aprire la strada ad un'implementazione reale di futuri computer quantistici con hardware topologicamente protetto.



- [1] L. FU, C. L. KANE: "Superconducting Proximity Effect and Majorana Fermions at the Surface of a Topological Insulator", *Phys. Rev. Lett.* **100** (2008) 096407.
- [2] J. D. SAU ET AL.: "Generic New Platform for Topological Quantum Computation Using Semiconductor Heterostructures", *Phys. Rev. Lett.* **104** (2010) 040502.
- [3] J. ALICEA: "Majorana fermions in a tunable semiconductor device", *Phys. Rev. B* **81** (2010) 125318.
- [4] V. MOURIK ET AL.: "Signature of Majorana Fermions in Hybrid Superconductor-Semiconductor Nanowire Devices", *Science* **336** (2012) 1003.



Pasquale Scarlino: ha conseguito la laurea magistrale in Fisica presso l'Università del Salento nell'A.A. 2009-2010 con una tesi dal titolo "Radio-Frequency QPC for Spin-QuBits Read-Out" svolta presso il Kavli Institute of NanoScience, Università di Delft dove attualmente svolge il dottorato di ricerca nel gruppo del Prof. Lieven M. K. Vandersypen.

Breve storia dei numeri primi

Alessandro Zaccagnini

Dipartimento di Matematica e Informatica, Università di Parma, Parma.

Facciamo una breve panoramica dei più importanti risultati sulla distribuzione dei numeri primi. Andremo sostanzialmente in ordine cronologico presentando i risultati dei matematici più importanti del XIX secolo. Di seguito parleremo della Congettura di Riemann ed infine faremo il punto della situazione sui quattro problemi proposti da Landau nel 1912. In diversi casi, daremo dimostrazioni moderne dei risultati che presenteremo, pur rispettando essenzialmente il lavoro dei grandi matematici di cui parleremo.

La storia fino al XIX secolo

Euclide

Euclide per primo ha dimostrato che l'insieme dei numeri primi è infinito. Qui diamo una dimostrazione leggermente diversa da quella che la tradizione gli attribuisce.

Teorema 1 (Euclide). *Esistono infiniti numeri primi.*

Dim. Consideriamo un numero intero positivo N ed osserviamo che

$$p \mid N! + 1 \quad \implies \quad p > N$$

Dunque l'insieme dei numeri primi è illimitato superiormente.

Per esempio il più piccolo numero primo p che divide $5! + 1 = 121 = 11^2$ è necessariamente > 5 . Questa dimostrazione ha il pregio di fornire automaticamente una limitazione dal basso per *qualsiasi* divisore non banale di $N! + 1$, non necessariamente primo. Vedi la nota 1.

Eulero

Molti secoli dopo, Eulero ha dato una dimostrazione alternativa dello stesso fatto, molto importante per almeno due motivi: in primo luogo, perché per la prima volta intervengono concetti non strettamente aritmetici nelle considerazioni sui numeri primi, e in secondo luogo perché questa dimostrazione dà, come sottoprodotto, un'indicazione della *densità* dei numeri primi nell'insieme dei numeri naturali.

Teorema 2 (Eulero). *La serie*

$$\sum_p \frac{1}{p} \quad \text{è divergente}$$

dove la notazione indica che la somma è fatta solo sui numeri primi.

Il Teorema di Eulero non solo implica che i numeri primi sono infiniti, ma anche che sono piuttosto densi, più dei quadrati perfetti solo per

fare un esempio, altrimenti la serie dell'enunciato non divergerebbe. Questo risultato, dunque, è estremamente più forte del Teorema di Euclide. La dimostrazione è relativamente semplice e fornisce una buona illustrazione di come si possono far interagire al meglio concetti di natura aritmetica e altri di natura analitica; in altre parole, "discreto" e "continuo" possono tranquillamente e proficuamente coesistere.

Lemma 3. Per $N \rightarrow +\infty$ si ha

$$\prod_{p \leq N} \left(1 - \frac{1}{p}\right)^{-1} \geq \sum_{n \leq N} \frac{1}{n} \\ = \int_1^N \frac{dx}{x} + \mathcal{O}(1) = \log(N) + \mathcal{O}(1)$$

Dim. L'uguaglianza a destra nell'enunciato deriva dal fatto che stiamo essenzialmente confrontando l'integrale di una funzione decrescente con una opportuna *somma di Riemann*. Sviluppando ciascun fattore a sinistra in serie geometrica di ragione $1/p$, otteniamo

$$\prod_{p \leq N} \left(1 - \frac{1}{p}\right)^{-1} \\ = \prod_{p \leq N} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \frac{1}{p^3} + \dots\right) \\ = \sum_{n \in \mathcal{A}(N)} \frac{1}{n}$$

dove

$$\mathcal{A}(N) = \{n \in \mathbb{N}^* : p \mid n \Rightarrow p \leq N\} \\ \supset \{1, 2, \dots, N\}$$

Per esempio, per $N \geq 11$ l'addendo $1/198$ compare nella forma $(1/2) \cdot (1/3^2) \cdot 1 \cdot 1 \cdot (1/11)$ poiché $198 = 2 \cdot 3^2 \cdot 11$. In generale, possiamo moltiplicare fra loro un numero finito di serie assolutamente convergenti e riordinarne i termini. In altre parole, $\mathcal{A}(N)$ è il *semigrupp* moltiplicativo di \mathbb{N}^* generato dai numeri primi che non superano N .

Prendendo il logaritmo ed usando la formula di Taylor al primo ordine otteniamo

$$\log \prod_{p \leq N} \left(1 - \frac{1}{p}\right)^{-1} \\ = \sum_{p \leq N} \frac{1}{p} + \mathcal{O}(1) \geq \log \log(N) + \mathcal{O}(1)$$

che è una versione quantitativamente forte del Teorema di Eulero.

Definizione 4 (Funzione zeta). Per s reale, $s > 1$, poniamo

$$\zeta(s) = \sum_{n \geq 1} \frac{1}{n^s} = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1} \quad (1)$$

L'uguaglianza a destra (*identità di Eulero*) si dimostra come nel Lemma precedente, prima moltiplicando fra loro le serie geometriche assolutamente convergenti di ragione p^{-s} relative ai numeri primi p fino ad un certo valore N , e poi prendendo il limite per N che tende all'infinito. È la versione analitica del *Teorema Fondamentale dell'Aritmetica*. Da questa deduciamo ancora una volta il Teorema di Euclide: il *criterio integrale* per le serie ci dà

$$\zeta(s) \geq \int_1^{+\infty} \frac{dt}{t^s} = \frac{1}{s-1} \\ \Rightarrow \lim_{s \rightarrow 1^+} \zeta(s) = +\infty$$

e quindi esistono infiniti numeri primi. Infatti, se ne esistessero solo un numero finito, l'identità di Eulero implicherebbe un valore finito per quest'ultimo limite, perché il prodotto avrebbe solo un numero finito di fattori. Vedi la nota 2.

Gauss

Gauss capì che la chiave per capire la distribuzione dei numeri primi risiede nello studiare la loro densità, cioè quanti numeri primi, approssimativamente, cadono nell'intervallo $[1, N]$ quando N tende all'infinito.

Definizione 5.

$$\pi(N) = |\{p \leq N\}| = \sum_{p \leq N} 1$$

Gauss aveva l'abitudine di calcolare i numeri primi in intervalli di 1000 interi consecutivi, come passatempo. Basandosi sui dati così raccolti, formulò la seguente

Congettura 6 (Gauss). Per $N \rightarrow +\infty$ si ha

$$\pi(N) \approx \text{li}(N) = \int_0^N \frac{dt}{\log(t)}$$

dove il simbolo \approx indica un'uguaglianza approssimata.

Integrando $n + 1$ volte per parti e stimando l'integrale che resta, si vede facilmente che

$$\text{li}(N) = \frac{N}{\log N} \sum_{k=0}^n \frac{k!}{(\log N)^k} + \mathcal{O}_n\left(\frac{N}{(\log N)^{n+2}}\right) \quad (2)$$

per ogni $n \in \mathbb{N}$ fissato. Se prendiamo $n = 0$ nella (2) otteniamo in particolare che $\text{li}(N) \sim N/\log(N)$, dove $f \sim g$ significa che $f/g \rightarrow 1$ quando $N \rightarrow +\infty$.

Congettura 7 (Forma debole).

$$\lim_{N \rightarrow +\infty} \frac{\pi(N)}{\text{li}(N)} = \lim_{N \rightarrow +\infty} \frac{\pi(N) \log(N)}{N} = 1 \quad (3)$$

Al primo ordine, dunque, non c'è molta differenza fra $\text{li}(N)$ ed $N/\log(N)$, ma dal punto di vista numerico, come Gauss aveva intuito, $\text{li}(N)$ è un'approssimazione di $\pi(N)$ molto più accurata: si veda la Figura 1. Prendendo $n = 1$ nella (2) e trascurando errori di ordine più piccolo si ritrova, essenzialmente, la formula proposta da Legendre sulla base di dati che oggi possiamo giudicare del tutto insufficienti, e cioè $\pi(N) \sim N/(\log(N) - 1)$. Questa è leggermente più accurata, dal punto di vista numerico, della relazione $\pi(N) \sim N/\log(N)$, ma evidentemente equivale alle altre due contenute nella Congettura 7.

Sia ora p_n l' n -esimo numero primo. Le due funzioni $\pi(N)$ e p_n sono sostanzialmente l'una l'inversa dell'altra, nel senso preciso che $\pi(p_n) = n$. La forma debole della Congettura di Gauss equivale alla formula asintotica $p_n \sim n \log(n)$. Questo è pienamente compatibile con il Teorema di Eulero visto sopra. Vedi la nota 3.

Dirichlet

Facciamo una brevissima digressione per ricordare un importante risultato di Dirichlet, che dimostrò l'esistenza di infiniti numeri primi in tutte le progressioni aritmetiche che ne possono contenere almeno due. Dirichlet riutilizzò le tecniche di Eulero e per primo introdusse l'analisi

N	$\pi(N)$	$\text{li}(N) - \pi(N)$
10	4	2
10^2	25	5
10^3	168	10
10^4	1 229	17
10^5	9 592	38
10^6	78 498	130
10^7	664 579	339
10^8	5 761 455	754
10^9	50 847 534	1 701
10^{10}	455 052 511	3 104
10^{11}	4 118 054 813	11 588
10^{12}	37 607 912 018	38 263
10^{13}	346 065 536 839	108 971
10^{14}	3 204 941 750 802	314 890
10^{15}	29 844 570 422 669	1 052 619
10^{16}	279 238 341 033 925	3 214 632
10^{17}	2 623 557 157 654 233	7 956 589
10^{18}	24 739 954 287 740 860	21 949 555
10^{19}	234 057 667 276 344 607	99 877 775
10^{20}	2 220 819 602 560 918 840	222 744 644
10^{21}	21 127 269 486 018 731 928	597 394 254
10^{22}	201 467 286 689 315 906 290	1 932 355 208
10^{23}	1 925 320 391 606 803 968 923	7 250 186 216
10^{24}	18 435 599 767 349 200 867 866	17 146 907 278
10^{25}	176 846 309 399 143 769 411 680	55 160 980 939

Figura 1: Confronto tra la quantità di numeri primi e la formula di Gauss. Si noti che $\pi(N) < \text{li}(N)$ per tutti questi valori. Si noti inoltre che la colonna a destra è ampia circa la metà di quella centrale.

complessa nella Teoria dei Numeri, prima ancora di Riemann, anche se in una forma che oggi appare rudimentale.

Teorema 8 (Dirichlet, 1837). Per ogni $q \in \mathbb{N}^*$ ed ogni $a \in \mathbb{Z}$ tali che $(q, a) = 1$ esistono infiniti numeri primi $p \equiv a \pmod{q}$.

Non c'è lo spazio per approfondire i dettagli, e quindi citiamo brevemente due importanti questioni lasciate aperte dal lavoro di Dirichlet. La prima riguarda il fatto che Dirichlet teneva q fissato, dimostrando una forma appropriata del Teorema di Eulero 2 dove la somma è fatta solo sui numeri primi $p \equiv a \pmod{q}$, mentre è ora riconosciuta l'importanza di avere risultati *uniformi*, cioè che valgono per un gran numero di progressioni simultaneamente. Inoltre, per le applicazioni (citiamo, fra le tante, solo quella al problema di Goldbach) è necessario avere, allo stesso tempo, anche relazioni *quantitative* dello stesso tipo della Congettura 6 di Gauss, se possibile con un termine d'errore esplicito. Vedi la nota 4.

Chebyshev

Chebyshev fu il primo a fornire un'evidenza quantitativa alla Congettura di Gauss.

Definizione 9. Poniamo

$$\lambda = \liminf_{N \rightarrow +\infty} \frac{\pi(N) \log(N)}{N}$$

e

$$\Lambda = \limsup_{N \rightarrow +\infty} \frac{\pi(N) \log(N)}{N}$$

Teorema 10 (Chebyshev). Si ha $0.92 \dots \leq \lambda \leq \Lambda \leq 1.105 \dots$

Questo significa che il vero ordine di grandezza quando $N \rightarrow +\infty$ della funzione $\pi(N)$ è effettivamente $N/\log(N)$.

Teorema 11 (Chebyshev). Se $\lambda = \Lambda$ allora $\lambda = \Lambda = 1$.

In altre parole, se il limite nella Congettura 7 esiste, allora vale 1. Chebyshev introdusse due funzioni per "contare" i numeri primi, oltre alla funzione π definita qui sopra: una di queste, chiamata ψ , ha un'interpretazione aritmetica del tutto naturale.

Definizione 12 (Funzione ψ di Chebyshev). $\psi(N) = \log \text{mcm}\{1, 2, 3, \dots, N\}$

Non daremo l'intera dimostrazione del Teorema di Chebyshev 10, ma solo le idee necessarie per ottenere la minorazione con il valore $\log(2)$ al posto di $0.92 \dots$. Cominciamo da una disuguaglianza fondamentale.

Lemma 13.

$$\psi(N) = \sum_{p \leq N} \left\lfloor \frac{\log(N)}{\log(p)} \right\rfloor \log(p) \leq \pi(N) \log(N)$$

Inoltre $\psi(N) \geq (1 + o(1))\pi(N) \log(N) + \mathcal{O}(N(\log N)^{-1})$

Dim. Per dimostrare l'uguaglianza a sinistra è sufficiente osservare che, dato un numero primo $p \leq N$, la sua massima potenza che non supera N è quella con esponente $a_p = \lfloor \log(N)/\log(p) \rfloor = \max\{m \in \mathbb{N} : p^m \leq N\}$, per poi ricordare la caratterizzazione elementare del minimo comune multiplo. Inoltre, per ogni $y \in (1, N]$ abbiamo

$$\psi(N) \geq \sum_{y < p \leq N} \log p \geq (\pi(N) - \pi(y)) \log y$$

La disuguaglianza cercata segue scegliendo $y = N(\log N)^{-2}$, poiché $\pi(y) \leq y$.

Dimostrazione (Teorema di Chebyshev). Per $m \in \mathbb{N}$ poniamo

$$\begin{aligned} I_m &= \int_0^1 x^m (1-x)^m dx \\ &= B(m+1, m+1) = \frac{m!^2}{(2m+1)!} \end{aligned}$$

Per i nostri scopi presenti non è strettamente necessario osservare che I_m è un valore della funzione Beta di Eulero. Ci basta la disuguaglianza

$$0 < I_m \leq \left(\max_{x \in [0,1]} x(1-x) \right)^m = \frac{1}{4^m}$$

La funzione integranda è un polinomio a coefficienti interi di grado $2m$: dunque

$$\int x^m (1-x)^m dx = (-1)^m \frac{x^{2m+1}}{2m+1} + \dots \in \mathbb{Q}[x] + \mathbb{R}$$

e inoltre, a parte il valore della costante di integrazione, i coefficienti di questa primitiva sono numeri razionali con denominatore $\leq 2m+1$. Valutiamo una primitiva in 0 e in 1, e poi moltiplichiamo il risultato per il minimo comune multiplo degli interi fra 1 e $2m+1$: il risultato è un intero *positivo*; quindi

$$\begin{aligned} I_m \exp \psi(2m+1) &\in \mathbb{N}^* \\ \implies \psi(2m+1) &\geq \log(I_m^{-1}) \geq 2m \log(2) \end{aligned}$$

Usando il Lemma 13 otteniamo $\pi(2m+1) \geq 2m \log(2) / \log(2m+1)$. Osservando infine che $\pi(2m) \geq \pi(2m+1) - 1$ per $m \geq 1$, con qualche breve calcolo si ottiene la tesi.

Vedi la nota 5.

Mertens

Mertens riuscì a semplificare la dimostrazione dei Teoremi di Chebyshev e a rendere più precisa la stima del risultato del Teorema di Eulero 2. L'idea da cui si può partire è il calcolo di $\log(N!)$ effettuato in due modi diversi. Da una parte la *formula di Stirling*, in una versione non

particolarmente precisa, ci dà

$$\begin{aligned} \log(N!) &= \sum_{n \leq N} \log(n) \\ &= \int_1^N \log(t) dt + \mathcal{O}(\log(N)) \\ &= N \log(N) + \mathcal{O}(N) \end{aligned}$$

dove anche questa volta confrontiamo l'integrale di una funzione monotona con una sua somma di Riemann.

Il secondo passo è quello di scomporre in fattori primi $N!$, tenendo conto del fatto che solo i numeri primi $p \leq N$ possono comparire nella fattorizzazione:

$$\begin{aligned} &\log(N!) \\ &= \sum_{p \leq N} \left(\left\lfloor \frac{N}{p} \right\rfloor + \left\lfloor \frac{N}{p^2} \right\rfloor + \left\lfloor \frac{N}{p^3} \right\rfloor + \dots \right) \log(p) \\ &= N \sum_{p \leq N} \frac{\log(p)}{p} + \mathcal{O}(N). \end{aligned} \quad (4)$$

L'uguaglianza a destra si dimostra trascurando le parti intere e il contributo delle potenze perfette. Per ottenere l'uguaglianza a sinistra è sufficiente osservare che, scelto un numero primo $p \leq N$, vi sono esattamente $\lfloor N/p \rfloor$ suoi multipli che non superano N , e questi contribuiscono un'unità all'esponente di p nella fattorizzazione di $N!$. Inoltre, vi sono $\lfloor N/p^2 \rfloor$ multipli di p^2 che non superano N , che contribuiscono un'ulteriore unità, e così via. La somma qui sopra al centro è solo apparentemente infinita poiché $\lfloor N/p^k \rfloor = 0$ per k sufficientemente grande. Per esempio, la massima potenza di 2 che divide $100!$ è quella con esponente $50 + 25 + 12 + 6 + 3 + 1 = 97$, poiché nell'intervallo $[1, 100]$ ci sono 50 numeri pari, 25 multipli di 4, 12 multipli di 8, 6 multipli di 16, 3 multipli di 32 ed infine un solo multiplo di 64. Ciascuno di questi contribuisce all'esponente in questione come affermato. Abbiamo dunque dimostrato il seguente risultato:

Teorema 14 (Mertens). *Per $N \rightarrow +\infty$ si ha*

$$\sum_{p \leq N} \frac{\log(p)}{p} = \log(N) + \mathcal{O}(1)$$

Una conseguenza è la formula asintotica, che

precisa il Teorema di Eulero 2,

$$\sum_{p \leq N} \frac{1}{p} = \log \log(N) + \mathcal{O}(1) \quad (5)$$

che si deduce dalla precedente mediante uno strumento della stessa sofisticazione della formula di integrazione per parti, detta *Formula di sommazione parziale*; si veda l'Appendice. Da questa formula asintotica segue una dimostrazione molto semplice del Teorema di Chebyshev 11: infatti, se $\pi(N) \sim LN/\log(N)$ per qualche $L \in \mathbb{R}^+$, allora la somma a primo membro della (5) vale $(L + o(1)) \log \log(N)$. In altre parole, se esiste il limite nella Congettura 7, Eq. (3), allora vale 1. Resta sempre il problema di dimostrare che questo limite esiste. La dimostrazione originale di Chebyshev, precedente rispetto alle scoperte di Mertens, utilizzava le proprietà analitiche della funzione ζ di Riemann, di cui parleremo sotto.

Naturalmente i Teoremi di Chebyshev 11 e di Mertens 14 seguono immediatamente dalla Congettura di Gauss 6, con la formula di sommazione parziale, ma possono essere ricavati direttamente ed elementarmente, come abbiamo, seppur solo in parte, visto. Vedi la nota 6.

Riemann

Riemann scrisse un solo breve articolo [1] sulla Teoria dei Numeri, nel quale indicò la strada che portò, una quarantina d'anni più tardi, alla dimostrazione della Congettura di Gauss, che oggi si chiama *Teorema dei Numeri Primi*. La chiave di tutto il ragionamento di Riemann è quella di considerare la funzione ζ con argomento complesso. Possiamo riassumere i punti cruciali del suo lavoro come segue, in un certo senso, scattando 6 istantanee del procedimento suggerito da Riemann, e portato a termine dai suoi successori, per indicare la plausibilità della sua argomentazione. È opportuno ricordare che all'epoca della pubblicazione del lavoro citato (1859) l'analisi complessa come la conosciamo oggi quasi *non esisteva* e quindi, ciò che a noi appare ragionevole, oltre 150 anni dopo, era allora straordinariamente innovativo.

1. Dimostrazione dell'Identità di Eulero (1) per tutti i numeri complessi $s = \sigma + it$ con $\sigma > 1$.

2. Equazione funzionale e prolungamento analitico della funzione ζ a $\mathbb{C} \setminus \{1\}$.
3. Espressione di ζ come *prodotto di Weierstrass* sugli zeri.
4. Stima del numero degli zeri di ζ nella *striscia critica* $0 \leq \sigma \leq 1$.
5. Espressione di $\psi(N)$ mediante un opportuno integrale complesso su un cammino illimitato contenuto nel semipiano $\sigma > 1$.
6. Deformazione del cammino di integrazione: connessione fra ψ e gli zeri della funzione ζ (la *formula esplicita*).

Del primo punto ci siamo, sostanzialmente, già occupati sopra, poiché $|p^{-s}| = p^{-\sigma}$, e $\sigma > 1$ per ipotesi, e quindi si può ripetere la stessa dimostrazione del Lemma 3. Per la precisione, Riemann lavorava con la funzione π , e non con la funzione ψ che risulta tecnicamente più comoda.

Teorema 15 (Equazione funzionale). *Posto*

$$\begin{aligned} \xi(s) &= \frac{1}{2}s(s-1)\pi^{-s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s) \\ &= \pi^{-s/2}\Gamma\left(\frac{s}{2}+1\right)(s-1)\zeta(s) \end{aligned}$$

si ha che ξ è olomorfa su \mathbb{C} e soddisfa la relazione $\xi(s) = \xi(1-s)$.

Poiché le proprietà di ζ sono note nel semipiano $\sigma > 1$, l'equazione funzionale permette di dedurle nel semipiano $\sigma < 0$. In particolare, $\zeta(s) \neq 0$ in $\sigma > 1$ e quindi la funzione ζ non si annulla nemmeno in $\sigma < 0$, a parte i cosiddetti zeri "banali" ai poli di Γ , situati in $-2, -4, -6, \dots$. Si noti che ξ è reale sull'asse reale e sulla retta di equazione $\sigma = \frac{1}{2}$. Questo è cruciale per dimostrare numericamente che gli zeri complessi della funzione zeta (tutti quelli noti fino ad oggi) si trovano sulla retta $\sigma = \frac{1}{2}$: si veda la discussione sulla Congettura di Riemann, più avanti.

La regione del piano complesso con $0 \leq \sigma \leq 1$ si dice "striscia critica": sono gli zeri in questa zona che determinano la bontà delle approssimazioni di $\pi(N)$.

Lasciamo da parte il prodotto di Weierstrass per la funzione ζ (più precisamente, per la funzione ξ) che è un argomento strettamente legato

all'analisi complessa e ci costringerebbe ad una digressione. Ci limitiamo ad osservare che è un ingrediente fondamentale nella dimostrazione, fra le altre cose, della Formula di Riemann-von Mangoldt 16 qui sotto e della regione libera da zeri. Per l'equazione funzionale e il principio di riflessione, gli zeri complessi $\rho = \beta + i\gamma$ della funzione zeta sono disposti simmetricamente rispetto all'asse reale e rispetto al punto $s = \frac{1}{2}$. Per la dimostrazione del Teorema dei Numeri Primi 18 è indispensabile avere informazioni più precise sulla loro distribuzione, da almeno due punti di vista. Il primo è la loro *densità*.

Teorema 16 (Formula di Riemann-von Mangoldt). *Per $T \rightarrow +\infty$ si ha*

$$\begin{aligned} N(T) &= |\{\rho = \beta + i\gamma: \zeta(\rho) = 0, \\ &\quad \beta \in (0, 1), \gamma \in [0, T]\}| \\ &= \frac{T}{2\pi} \log \frac{T}{2\pi} - \frac{T}{2\pi} + \mathcal{O}(\log T) \quad (6) \end{aligned}$$

La dimostrazione è un'applicazione non banale del principio dell'argomento alla funzione ξ . Il termine principale proviene dalla *formula di Stirling* generalizzata per la funzione Γ di Eulero presente nell'equazione funzionale di ξ . Il termine d'errore proviene essenzialmente da $\arg(\zeta(\frac{1}{2} + iT))$.

Teorema 17 (Formula esplicita). *Per $N \rightarrow +\infty$ vale la relazione*

$$\begin{aligned} \psi(N) &= \log \text{mcm}\{1, 2, 3, \dots, N\} \\ &= N - \sum_{\rho} \frac{N^{\rho}}{\rho} + \mathcal{O}(\log N) \end{aligned}$$

dove la somma è fatta su tutti gli zeri $\rho = \beta + i\gamma$ con $\beta \in (0, 1)$ ed è intesa in senso simmetrico, con ρ e $\bar{\rho}$ presi insieme.

La dimostrazione dipende dalla relazione fondamentale valida per $c > 1$

$$\psi(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} -\frac{\zeta'(s)}{\zeta(s)} x^s \frac{ds}{s} \quad (7)$$

dove $x > 1$ ed $x \notin \mathbb{N}$. Questa trasformazione integrale, di cui vedremo più avanti l'inversa, si chiama trasformata di Mellin ed è formalmente simile alla trasformata di Fourier. Vedi la nota 7.

Hadamard — de la Vallée Poussin

Sono stati necessari quasi 40 anni per dimostrare rigorosamente tutte le affermazioni di Riemann (tranne quella sulla famosa Congettura, naturalmente): si può dire, senza esagerare troppo, che l'analisi complessa sia stata creata nella seconda metà del XIX secolo proprio per raggiungere questo obiettivo. Il risultato finale, il Teorema dei Numeri Primi, è stato dimostrato, indipendentemente e quasi simultaneamente, da Hadamard e de la Vallée Poussin nel 1896.

Teorema 18 (dei Numeri Primi). *Esiste una costante positiva c tale che*

$$\pi(x) = \text{li}(x) + \mathcal{O}\left(x \exp(-c(\log x)^{3/5}(\log \log x)^{-1/5})\right)$$

Questo equivale alla relazione $\psi(x) = x + \mathcal{O}(x \exp(-c_1(\log x)^{3/5}(\log \log x)^{-1/5}))$, che è più facile da dimostrare grazie alla (7). Si veda una spiegazione nella didascalia della Figura 2. L'ingrediente mancante, che spiega anche la curiosa funzione nel *termine d'errore*, è un'altra informazione sulla distribuzione degli zeri della funzione ζ . Per la precisione, questo è l'unico punto del programma attribuito sopra a Riemann che non è presente nell'originale, visto che è sostanzialmente inutile per dimostrare semplicemente che il limite nella Congettura di Gauss 6 esiste, mentre è indispensabile per valutare la bontà dell'approssimazione. L'apparente anacronismo dei due risultati citati è dovuto al fatto che il Teorema dei Numeri Primi realmente dimostrato da Hadamard e de la Vallée Poussin era leggermente più debole di quello enunciato qui perché, ovviamente, non avevano a disposizione il risultato seguente, ma il meccanismo della loro dimostrazione non cambia se non marginalmente inserendo la versione della regione libera da zeri trovata da Vinogradov e Korobov, indipendentemente, nel 1958.

Teorema 19 (Regione libera da zeri, Vinogradov e Korobov). *Esiste una costante $c > 0$ tale che se $\zeta(\beta + i\gamma) = 0$ con $\beta \in [0, 1]$ e $\gamma > 0$ allora*

$$\beta < 1 - \frac{c}{(\log \gamma)^{2/3}(\log \log \gamma)^{1/3}}$$

Questa regione libera da zeri si usa insieme alla Formula Esplicita nella sua versione troncata

di cui si parla nella didascalia della Figura 2. In effetti si ha

$$\psi(x) = x - \sum_{\substack{\rho=\beta+i\gamma \\ |\gamma|\leq T}} \frac{x^\rho}{\rho} + \mathcal{O}\left(\frac{x}{T}(\log(xT))^2 + \log x\right) \quad (8)$$

La scelta ottimale per T è quella che rende il termine sugli zeri (una somma *finita*) dello stesso ordine di grandezza, essenzialmente, di x/T . Si ha

$$\psi(x) = x + \mathcal{O}\left(\left\{\max_{0<\gamma\leq T} x^\beta\right\} \sum_{0<\gamma\leq T} \frac{1}{\gamma} + \frac{x}{T}(\log xT)^2\right) \quad (9)$$

dove abbiamo scritto implicitamente $\rho = \beta + i\gamma$ per il generico zero non banale di zeta. Si ricordi che gli zeri sono disposti simmetricamente rispetto all'asse reale. Il massimo può essere stimato usando la regione libera da zeri fornita dal Teorema 19, mentre per la somma utilizziamo la stima per il numero degli zeri della funzione zeta con parte immaginaria $|\gamma| \leq T$ data dal Teorema 16, con la sommazione parziale. In definitiva, possiamo scrivere

$$\psi(x) = x + \mathcal{O}\left(x(\log T)^2 \exp\left\{-c \frac{\log x}{(\log T)^{2/3}(\log \log T)^{1/3}}\right\} + \frac{x}{T}(\log xT)^2\right) \quad (10)$$

Ora scegliamo T come funzione di x in modo che

$$\frac{1}{T} \approx \exp\left\{-c \frac{\log x}{(\log T)^{2/3}(\log \log T)^{1/3}}\right\}$$

per rendere approssimativamente uguali i due termini a destra nella (10). Questo accade se $T = T(x)$ soddisfa $(\log T)^{5/3}(\log \log T)^{1/3} \approx \log x$: invertendo la relazione funzionale si trova $\log T \approx (\log x)^{3/5}(\log \log x)^{-1/5}$. Infine, sostituendo e semplificando otteniamo

$$\psi(x) = x + \mathcal{O}\left(x \exp\{-c_1(\log x)^{3/5}(\log \log x)^{-1/5}\}\right) \quad (11)$$

dove c_1 è un'opportuna costante positiva. Come si vede, c'è una relazione "meccanica" fra l'ampiezza della regione libera da zeri e la dimensione del termine d'errore nella (11).

Vedi la nota 8.

Selberg — Erdős

La dimostrazione del Teorema dei Numeri Primi data da Hadamard e de la Vallée Poussin indipendentemente nel 1896, originata dalle idee di Riemann, è tanto intrinsecamente legata alle proprietà della funzione zeta come funzione oloomorfa, da far considerare improbabile, se non proprio impossibile, dimostrare lo stesso risultato usando solo l'analisi reale. In effetti, per alcuni decenni è stata informalmente osservata l'equivalenza espressa dal seguente teorema, dimostrato poi rigorosamente da Wiener nel 1927.

Teorema 20 (Wiener).

$$\zeta(1 + it) \neq 0$$

per ogni $t \in \mathbb{R} \setminus \{0\}$

$$\iff \pi(N) \sim \frac{N}{\log(N)}$$

Con una certa sorpresa, dunque, nel 1948 i Teorici dei Numeri accolsero la scoperta, da parte di Selberg ed Erdős, della cosiddetta dimostrazione elementare, che non fa uso dell'analisi complessa. Il suo cuore è il risultato qui sotto.

Teorema 21 (Formula di Selberg).

$$\begin{aligned} \psi(N) \log(N) + \sum_{p \leq N} \psi\left(\frac{N}{p}\right) \log(p) \\ = 2N \log(N) + \mathcal{O}(N) \end{aligned}$$

Una conseguenza immediata di questa e delle formule di Mertens viste sopra è il prossimo Corollario, che fornisce una dimostrazione alternativa del Teorema di Chebyshev 11, poiché implica che $\lambda \leq 1 \leq \Lambda$.

Corollario 22. Si ha $\lambda + \Lambda = 2$.

Per ottenere che $\lambda = \Lambda = 1$ è necessaria una lunga e complicata analisi di tipo *tauberiano* che qui omettiamo. Nelle applicazioni del Teorema dei Numeri primi è però di fondamentale importanza avere delle buone stime per il termine

d'errore $\psi(N) - N$: in questo campo, le tecniche "elementari" appaiono più deboli delle tecniche analitiche. Vedi la nota 9.

La Conggettura di Riemann

La Conggettura e le sue forme equivalenti elementari

Conggettura 23 (Riemann). *Tutti gli zeri complessi di ζ della forma $\beta + i\gamma$ con $\gamma \neq 0$ hanno $\beta = \frac{1}{2}$.*

Questo equivale a dire che

$$\pi(N) = \text{li}(N) + \mathcal{O}\left(N^{1/2} \log N\right) \quad (12)$$

$$\psi(N) = N + \mathcal{O}\left(N^{1/2}(\log N)^2\right)$$

Un'implicazione è conseguenza diretta della Formula Esplicita del Teorema 17, o meglio della sua versione "troncata" (8) descritta nella didascalia della Figura 2, e della Formula di Riemann-von Mangoldt del Teorema 16: basta scegliere $T = x^{1/2}$ nella (9). L'altra dipende dalla trasformazione integrale che lega ζ a ψ

$$\begin{aligned} -\frac{\zeta'}{\zeta}(s) &= s \int_1^{+\infty} \frac{\psi(x)}{x^{s+1}} dx \\ &= \frac{s}{s-1} + s \int_1^{+\infty} \frac{\psi(x) - x}{x^{s+1}} dx \quad (13) \end{aligned}$$

valida per $\sigma > 1$, che è l'inversa della (7). La connessione fra π e ζ , data originariamente da Riemann nel suo articolo, è più complicata e indiretta, e, poiché coinvolge esplicitamente il logaritmo della funzione ζ , è più difficile da trattare. Vedi la nota 10.

Una versione debole della Conggettura di Riemann

Sia

$$\Theta = \sup\{\beta: \exists \gamma \in \mathbb{R} \text{ tale che } \zeta(\beta + i\gamma) = 0\}$$

La simmetria degli zeri rispetto ad $s = \frac{1}{2}$ implica che $\frac{1}{2} \leq \Theta \leq 1$. Sia

$$\Theta' = \inf\{\alpha \in \mathbb{R}: \pi(N) = \text{li}(N) + \mathcal{O}(N^\alpha)\}$$

Teorema 24. Si ha $\Theta = \Theta'$.

La formula (13) mostra esplicitamente l'importanza di maggiorazioni forti per $R(x) = \psi(x) - x$.

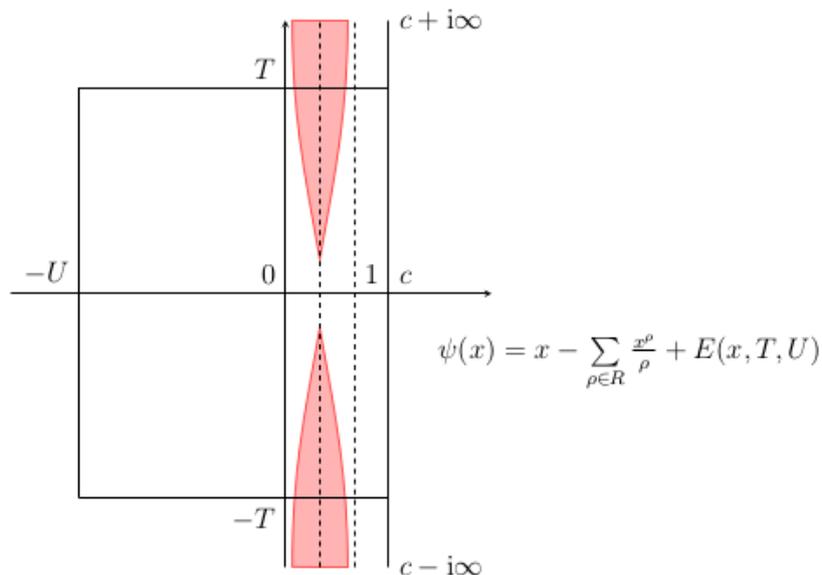


Figura 2: Il cammino d'integrazione nella (7), la retta verticale dei numeri complessi di parte reale c , viene deformato in un rettangolo R , con vertici in $c \pm iT$ e $-U \pm iT$, più le due "code" verticali, da $c + iT$ a $c + i\infty$ e la sua simmetrica, per poter usare il Teorema dei Residui sul rettangolo R . Si deve scegliere T con cura (non deve coincidere con l'ordinata di qualche zero di ζ), mentre U è sostanzialmente irrilevante. Il contributo del polo della funzione integranda in $s = 1$ dà il termine principale della Formula Esplicita nel Teorema 17, e gli zeri di ζ compresi nel rettangolo danno il termine secondario. Vi sono naturalmente dei termini d'errore, provenienti dalle code verticali, dai due segmenti ad altezza $\pm T$ e dal segmento verticale sull'ascissa $-U$ che devono essere accuratamente stimati. Facendo tendere separatamente U e T a $+\infty$ si ottiene la Formula Esplicita. In realtà, per le applicazioni è di gran lunga più utile la versione della Formula Esplicita in cui compare T , che poi può essere scelto in funzione di x : questa scelta dipende dall'ampiezza della regione libera da zeri disponibile. In definitiva, il procedimento illustrato sopra permette di scrivere $\psi(x) = x - \sum_{\rho \in R} x^\rho / \rho + E(x, T, U)$ dove E indica un "errore" che è funzione di x , T ed U e che raccoglie i contributi indicati sopra. Gli zeri non banali della funzione ζ sono confinati all'interno della regione colorata. Riemann ha congetturato che tutti gli zeri complessi di ζ (esclusi dunque quelli sull'asse reale negativo), siano sulla retta $\sigma = \frac{1}{2}$.

Il più semplice risultato in questo direzione è proprio il Teorema appena enunciato: in un verso si dimostra come abbiamo detto sopra, prendendo $T = x^{1-\Theta}$ nella (9); questo implica che $\Theta' \leq \Theta$. L'altra disuguaglianza si ottiene osservando che se $R(x) = \mathcal{O}(x^{\Theta'+\varepsilon})$ allora la funzione definita implicitamente dall'integrale all'estrema destra della (13) può essere prolungata analiticamente al semipiano $\sigma > \Theta'$: dunque la funzione a primo membro, in questo semipiano, non può avere poli diversi da $s = 1$ e di conseguenza ζ non ha zeri in questa regione. In sostegno della Congettura di Riemann si noti la sua connessione con l'ampiezza delle colonne al centro e a destra nella Figura 1: la seconda è, come previsto, ampia circa la metà della prima. Vedi la nota 11.

Risultati numerici

Ci limitiamo a qualche breve cenno ai più importanti risultati numerici.

1. I "primi" 10^{13} zeri sono sulla retta $\sigma = \frac{1}{2}$: Gourdon [2].
2. Almeno il 40% degli zeri di ζ (in senso asintotico) è sulla retta critica: Conrey [3].
3. Gli zeri al di fuori della retta critica (sempre se ne esistono) sono relativamente pochi, in un senso quantitativo preciso: Huxley [4].
4. Per tutti gli $x > 3$ per cui è stato calcolato il valore esatto di $\pi(x)$ si ha $\pi(x) < \text{li}(x)$, ma Littlewood [5] ha dimostrato nel 1914 che $\text{li}(x) - \pi(x)$ cambia segno infinite volte quando $x \rightarrow +\infty$.

Vedi la nota 12.

Il Teorema di Beurling

In questo e nei prossimi paragrafi discutiamo di alcune equivalenze, in qualche caso parziali, fra la Congettura di Riemann e risultati di altri campi della matematica. Alcune di queste equivalenze sono "elementari" nel senso che gli enunciati che vedremo non fanno intervenire l'analisi complessa o altra matematica avanzata: è uno dei tanti esempi di enunciati semplici e comprensibili che nascondono difficoltà molto profonde, di cui la Teoria dei Numeri è così ricca. A parte i problemi di Landau, con cui concluderemo questa trattazione, ci basti ricordare il Teorema di Fermat. Cominciamo senz'altro con il Teorema di Beurling [6].

Teorema 25 (Beurling, 1955). *Sia C lo spazio vettoriale generato da $\{\rho_\theta: 0 < \theta \leq 1\}$ definite da*

$$\rho_\theta(x) = \rho(\theta/x) - \theta\rho(1/x) \quad \text{per } x \in (0, 1]$$

dove $\rho(x) = x - [x]$ è la parte frazionaria di x . Sia C^p la chiusura di C in $L^p([0, 1])$. Allora ζ non ha zeri con parte reale $\beta > 1/p$ se e solo se $C^p = L^p([0, 1])$.

Non è del tutto sorprendente che la funzione ζ sia legata alla funzione parte frazionaria. La formula di sommazione parziale dà, dopo un breve calcolo, la relazione

$$\zeta(s) = \frac{s}{s-1} - s \int_1^{+\infty} \frac{\rho(t)}{t^{s+1}} dt \quad (14)$$

Questa formula fornisce, in modo del tutto elementare, il prolungamento analitico della funzione ζ al semipiano dei numeri complessi di parte reale positiva, privato, evidentemente, del punto $s = 1$, poiché ρ è una funzione limitata. Vedi la nota 13.

La Congettura di Mertens

Passiamo ora a descrivere alcune equivalenze elementari, nel senso spiegato qui sopra.

Definizione 26 (Funzione di Möbius). *Sia*

$$\mu(n) = \begin{cases} 1 & \text{se } n = 1, \\ 0 & \text{se esiste un primo } p \\ & \text{tale che } p^2 \mid n, \\ (-1)^k & \text{se } n = p_1 \cdots p_k \\ & \text{con } p_1 < \cdots < p_k \end{cases}$$

Poniamo

$$M(N) = \sum_{n \leq N} \mu(n) \quad (15)$$

e ricordiamo la stima elementare

$$\sum_{n \leq N} \mu^2(n) = \frac{6}{\pi^2} N + \mathcal{O}(N^{1/2})$$

Se un intero generico n ha un numero pari o dispari di fattori primi distinti con uguale "probabilità," allora ci si può aspettare che $M(N) = \mathcal{O}(N^{1/2} \log^2 N)$. Questo equivale alla Congettura di Riemann. Infatti, M è la somma dei coefficienti della serie di Dirichlet

$$\sum_{n \geq 1} \frac{\mu(n)}{n^s} = \frac{1}{\zeta(s)}$$

e la stima qui sopra implica che $1/\zeta$ è olomorfa nel semipiano $\sigma > \frac{1}{2}$, e quindi ζ non si annulla nello stesso semipiano. Le due relazioni

$$\sum_{n=1}^{\infty} \frac{\mu(n)}{n} = 0 \quad \text{e} \quad \sum_{n \leq N} \mu(n) = o(N) \quad (16)$$

sono equivalenti al Teorema dei Numeri Primi nella forma congetturata da Gauss. La prima delle due è stata enunciata per la prima volta da Eulero e dimostrata da Landau: la dimostrazione dell'equivalenza non è semplice. Le due relazioni fondamentali soddisfatte dalla funzione μ sono

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{se } n = 1, \\ 0 & \text{se } n > 1 \end{cases} \quad (17)$$

$$\sum_{n \leq x} \mu(n) \left[\frac{x}{n} \right] = \begin{cases} 0 & \text{se } x \in (0, 1) \\ 1 & \text{se } x \geq 1 \end{cases} \quad (18)$$

Vedi la nota 14.

Il Teorema di Franel e Landau

Passiamo ad una cosa ancor più elementare: parliamo di frazioni con denominatori limitati.

Definizione 27 (Frazioni di Farey). *Sia N un intero ≥ 1 . Poniamo*

$$\begin{aligned} \mathcal{F}_N &= \left\{ \frac{a}{q} : 0 < a \leq q \leq N \text{ e } (a, q) = 1 \right\} \\ &= \{x_1 < x_2 < \cdots < x_{\Phi(N)} = 1\} \end{aligned}$$

Per esempio

$$\mathcal{F}_5 = \left\{ \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{2}{5}, \frac{1}{2}, \frac{3}{5}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, 1 \right\}$$

Notiamo che $\Phi(N) = |\mathcal{F}_N| \sim 3N^2/\pi^2$ quando $N \rightarrow +\infty$.

Teorema 28 (Franel e Landau). *La Congettura di Riemann vale se e solo se per tutti gli $\varepsilon > 0$*

$$\Delta(N) = \sum_{n=1}^{\Phi(N)} \left| x_n - \frac{n}{\Phi(N)} \right| = \mathcal{O}\left(N^{1/2+\varepsilon}\right)$$

Ricordando la definizione (15), la chiave della dimostrazione è la relazione

$$\sum_{n=1}^{\Phi(N)} f(x_n) = \sum_{k \geq 1} M\left(\frac{N}{k}\right) \sum_{j=1}^k f\left(\frac{j}{k}\right) \quad (19)$$

valida per ogni funzione f di periodo 1. Posto $x_n = n/\Phi(N) + \delta_n$ per $n = 1, \dots, \Phi(N)$, valutiamo la (19) con $f(x) = e^{2\pi i x}$, osservando che, per una nota proprietà delle radici k -esime dell'unità, la somma interna a destra della (19) vale 0 quando $k > 1$ e vale 1 quando $k = 1$. In definitiva, il secondo membro vale $M(N)$. Dunque

$$\begin{aligned} M(N) &= \sum_{n=1}^{\Phi(N)} e^{2\pi i(n/\Phi(N) + \delta_n)} \\ &= \sum_{n=1}^{\Phi(N)} e^{2\pi i n/\Phi(N)} (e^{2\pi i \delta_n} - 1) + \sum_{n=1}^{\Phi(N)} e^{2\pi i n/\Phi(N)} \end{aligned}$$

L'addendo all'estrema destra vale 0 se $N > 1$, mentre per l'altro possiamo usare la disuguaglianza $|e^{2\pi i \delta_n} - 1| = 2|\sin(\pi \delta_n)| \leq 2\pi|\delta_n|$, ottenendo in definitiva $|M(N)| \leq 2\pi\Delta(N)$, che dimostra una delle due implicazioni. L'altra implicazione è più complessa da dimostrare, ma si basa anch'essa sulla (19). Vedi la nota 15.

Il Teorema di Robin

Sorprendentemente, la Congettura di Riemann può essere espressa in modo assolutamente elementare come segue.

Definizione 29 (Numero armonico). *Sia n un intero ≥ 1 . Poniamo*

$$H_n = \sum_{i=1}^n \frac{1}{i}$$

Teorema 30 (Robin). *La Congettura di Riemann è vera se e solo se per ogni $n \geq 1$ si ha*

$$\sigma(n) = \sum_{d|n} d \leq H_n + \exp(H_n) \log(H_n) \quad (20)$$

Informalmente, se la Congettura di Riemann fosse falsa esisterebbe una successione divergente $(x_j)_{j \in \mathbb{N}}$ tale che $\pi(x_j) > \text{li}(x_j) + x_j^{1/2+\delta}$, dove $\delta > 0$ è una quantità fissata. Usando i numeri primi in $[1, x_j]$ si potrebbe costruire un intero n_j con un valore $\sigma(n_j)$ più grande della norma e tale da falsificare, seppur di poco, la (20).

I problemi di Landau

Nel 1912 il grande matematico tedesco Edmund Landau propose una lista di importanti problemi aperti all'International Congress in Mathematics, a Cambridge.

Congettura 31. *È vero che $n^2 + 1$ è primo per infiniti n ?*

Congettura 32. *È vero che $n = p_1 + p_2$ ha sempre soluzione se $n \geq 4$ è pari?*

Congettura 33. *È vero che $p_{n+1} - p_n = 2$ per infiniti n ?*

Congettura 34. *È vero che esiste un numero primo tra n^2 ed $(n+1)^2$ per ogni n ?*

Vedi la nota 16.

Il primo problema di Landau: valori primi di polinomi

Hardy & Littlewood [7] hanno congetturato che

$$|\{n \leq N : n^2 + 1 \text{ è primo}\}| \sim C \frac{N}{\log(N)}$$

per una certa costante esplicita $C > 0$. Fermat ha osservato che il polinomio $X^2 + Y^2$ rappresenta $p = 2$ e tutti i numeri primi $p \equiv 1 \pmod{4}$. Friedlander & Iwaniec [8] ed Heath-Brown [9] hanno dimostrato che i polinomi $X^2 + Y^4$ e $X^3 + 2Y^3$ assumono valori primi per infiniti valori interi delle variabili. L'unico caso noto ad oggi di polinomio in una variabile che assume infiniti valori primi è quello del Teorema di Dirichlet 8. Vedi la nota 17.

Il secondo problema di Landau, o problema di Goldbach

Hardy & Littlewood [7] hanno congetturato che

$$r_2(n) = |\{(p_1, p_2): p_1 + p_2 = n\}| \\ \sim C \prod_{\substack{p|n \\ p>2}} \frac{p-1}{p-2} \frac{n}{(\log(n))^2}$$

per una certa costante esplicita $C > 0$, se n è pari e grande. Oggi è noto che il numero degli $n \leq N$ pari per cui questa formula non è accurata non supera $N/(\log N)^A$ per ogni $A > 0$ fissato ed $N \geq N_0(A)$. Pintz [10] ha dimostrato che

$$|\{n \leq N: n \text{ è pari e } r_2(n) = 0\}| \leq N^{2/3}$$

per N sufficientemente grande. Vinogradov [11] ha dimostrato che *tutti* i numeri dispari $n \geq N_0$ si possono rappresentare come somma di 3 numeri primi. Helfgott [12] ha recentemente chiuso la questione dimostrando che $N_0 = 7$. In questo particolare problema, la Congettura di Riemann non è utile. Vedi la nota 17.

Il terzo problema di Landau, o problema dei primi gemelli

Hardy & Littlewood [7] hanno congetturato che, se $h > 0$ è un intero pari fissato, allora

$$|\{p \leq N: p+h \text{ è primo}\}| \sim \\ C \prod_{\substack{p|h \\ p>2}} \frac{p-1}{p-2} \frac{N}{(\log(N))^2}$$

per $N \rightarrow +\infty$, e per una certa costante esplicita $C > 0$. Partendo da risultati rivoluzionari di Goldston, Pintz & Yıldırım (2006–2013), Zhang [13] ha dimostrato recentemente la congettura

$$\liminf_{n \rightarrow +\infty} (p_{n+1} - p_n) \leq C < +\infty$$

Zhang ha fornito la maggiorazione $C \leq 7 \cdot 10^7$. Maynard [14] (19 novembre 2013) dà $C \leq 600$. Secondo Maynard il limite del metodo è $C = 12$. Siamo dunque lontani dall'aver $C = 2$, e lontanissimi dalla versione quantitativa congetturata da Hardy & Littlewood. Vedi la nota 19.

Il quarto problema di Landau: numeri primi fra quadrati

La Congettura di Riemann implica che

$$p_{n+1} - p_n = \mathcal{O}\left(p_n^{1/2} (\log p_n)^2\right)$$

Infatti, dalla (12) valutata in x ed in $x+y$ segue che $\pi(x+y) - \pi(x) > 0$ se $y > Cx^{1/2}(\log x)^2$ dove C è una costante sufficientemente grande. Si congettura (Cramér, 1932) che

$$\limsup_{n \rightarrow +\infty} \frac{p_{n+1} - p_n}{(\log p_n)^2} < +\infty$$

Baker, Harman & Pintz [15] hanno dimostrato che

$$p_{n+1} - p_n = \mathcal{O}(p_n^{0.525})$$

Si può congetturare che $\pi(x+y) - \pi(x) \sim y/\log x$ se $y \leq x$, purché y non sia troppo piccolo. Questa relazione è probabilmente vera, per esempio, se $y = x^\alpha$, dove $\alpha > 0$ è fissato; in questo caso si avrebbe $p_{n+1} - p_n = \mathcal{O}(p_n^\alpha)$.

La formula di sommazione parziale

Diamo la formula di sommazione parziale con un breve cenno della sua dimostrazione. Se $\phi \in C^1([1, x])$ ed $(a_n)_{n \in \mathbb{N}}$ è una successione di numeri complessi qualsiasi, allora

$$\sum_{n \leq x} a_n \phi(n) = A(x)\phi(x) - \int_1^x A(t)\phi'(t) dt$$

dove

$$A(x) = \sum_{n \leq x} a_n \quad (21)$$

Posto $N = [x]$ e usando il fatto che A è costante fra due interi consecutivi, si ha

$$\int_1^x A(t)\phi'(t) dt \\ = \sum_{n=1}^{N-1} \int_n^{n+1} A(t)\phi'(t) dt + \int_N^x A(t)\phi'(t) dt \\ = \sum_{n=1}^{N-1} A(n) \int_n^{n+1} \phi'(t) dt + A(N) \int_N^x \phi'(t) dt$$

Si integra e si trova una somma “telescopica” dove si sfrutta il fatto che $A(n+1) - A(n) = a_{n+1}$. Riordinando la somma si trova la (21).

Note e spunti per letture ulteriori

Introduzione storica

Si veda il bell’articolo di Bateman & Diamond [16].

Nota 1

La nostra dimostrazione del Teorema di Euclide è diversa da quella data di solito, e ricorda il Teorema di Wilson: $p \geq 2$ è un numero primo se e solo se $p \mid (p-1)! + 1$. Per esempio, $6! + 1 = 721 = 7 \cdot 103$, mentre $10! + 1 = 3628801 = 11 \cdot 329891$. Si tratta di una delle più semplici, e inutili, caratterizzazioni dei numeri primi.

Nota 2

La funzione ζ era stata già introdotta in casi particolari già un secolo prima di Eulero: ricordiamo solo il cosiddetto Problema di Mengoli del 1644, che, nel linguaggio moderno, corrisponde ad esprimere il valore di $\zeta(2) = \pi^2/6$ in termini finiti.

Nota 3

Si osservi che la (2) è una *serie asintotica*: si può prendere n arbitrariamente grande, ma non esiste il limite per $n \rightarrow +\infty$ del secondo membro per il Teorema di Hadamard sul raggio di convergenza delle serie di potenze. Una discussione dei metodi per il calcolo numerico esatto di $\pi(N)$ si trova in Deléglise & Rivat [17], a cui si deve il valore di $\pi(10^{18})$: naturalmente, questi numeri primi non sono conosciuti individualmente. Il valore esatto di $\pi(10^{25})$ dato nella Figura 1 è stato annunciato nel maggio del 2013 da J. Buethe, J. Franke, A. Jost e T. Kleinjung.

Sono note disuguaglianze precise per p_n : qui diamo due delle più semplici.

$$\begin{aligned} & n(\log(n) + \log(\log(n)) - 1) \\ & \leq p_n \leq n(\log(n) + \log(\log(n))) \end{aligned}$$

dove quella a sinistra è valida per $n \geq 2$ e l’altra per $n \geq 6$. Sono dovute a Dusart [18] e Rosser

[19] rispettivamente. La formula di Legendre è discussa in Pintz [20].

Nota 4

La distribuzione dei numeri primi nelle progressioni aritmetiche è trattata in grande dettaglio in Davenport [21] con particolare attenzione agli aspetti quantitativi, inclusa l’uniformità in q , nonché alla sua rilevanza per applicazioni come il Problema di Goldbach.

Nota 5

La definizione usuale della funzione ψ di Chebyshev è

$$\psi(x) = \sum_{n \leq x} \Lambda(n)$$

dove

$$\Lambda(n) = \begin{cases} \log(p) & \text{se esiste } m \in \mathbb{N}^* \\ & \text{tale che } n = p^m, p \text{ primo,} \\ 0 & \text{altrimenti.} \end{cases}$$

Λ è detta funzione di von Mangoldt. Quest’ultima, a sua volta, può essere definita mediante l’identità $-\zeta'/\zeta(s) = \sum_{n \geq 1} \Lambda(n)/n^s$, che si ottiene derivando il logaritmo dell’Identità di Eulero (1). Una importante proprietà di Λ è espressa dall’identità $\sum_{d|n} \Lambda(d) = \log(n)$.

La dimostrazione della minorazione nel Teorema 10 è quella di Nair [22], dove si trova anche la maggiorazione corrispondente.

Nota 6

Si noti che l’uguaglianza a sinistra nella (4) può essere riscritta come

$$\log(N!) = \sum_{n \leq N} \psi\left(\frac{N}{n}\right)$$

la quale a sua volta suggerisce che $\psi(N) \sim N$, equivalente alla Congettura di Gauss 6.

Legendre [23] ha dimostrato che l’esponente $\nu_p(N)$ cercato nella dimostrazione del Teorema 14, e cioè la massima potenza del numero primo p che divide $N!$, vale esattamente $(N - s_p(N))/(p-1)$, dove $s_p(N)$ indica la somma delle cifre di N quando è scritto in base p .

Infatti, è sufficiente scrivere $N = N_1 p + r$ con $r \in [0, p - 1]$ e procedere per induzione, supponendo che la formula di Legendre valga per N_1 e osservando infine che questa è banale se $N < p$ cioè se $N_1 = 0$, poiché $\nu_p(m) = 0$ se $m < p$. In alternativa, si può sfruttare il fatto che la j -esima cifra di N scritto in base p è $\lfloor N/p^j \rfloor - p \lfloor N/p^{j+1} \rfloor$.

Nota 7

L'articolo originale di Riemann [1] è tradotto e dettagliatamente commentato in Edwards [24]; si veda anche Davenport [21], §8. È certo che Eulero conosceva l'equazione funzionale (almeno in casi particolari): infatti quest'ultima si può ottenere formalmente manipolando la definizione di ζ ed usando con cautela proprietà non ovvie della funzione Γ . Il suo punto di partenza può essere stata la formula (14): si sostituisce lo sviluppo in serie di Fourier per ρ , integrando poi la serie termine a termine. Per i dettagli si veda Titchmarsh [25], p. 15 oppure Hardy [26], §2.2.

Per la teoria generale della funzione ζ , oltre alle monografie di Edwards e Titchmarsh appena menzionate, si veda anche Ivić [27]. Una discussione della formula esplicita per la funzione π si trova in Zagier [28], del quale consigliamo in particolare i grafici che rendono evidente la bontà dell'approssimazione scoperta da Riemann.

Osserviamo che le relazioni (7) e (13) rappresentano una coppia trasformata e anti-trasformata di Mellin, che formalmente sono trasformazioni dello stesso tipo di quella di Fourier, e il cui esempio più noto è la coppia e^{-x} , $\Gamma(s)$. Infatti, ricordiamo che per $\sigma = \Re(s) > 0$ e $c > 0$ si ha

$$\begin{aligned} \Gamma(s) &= \int_0^{+\infty} e^{-x} x^{s-1} dx \iff \\ e^{-x} &= \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \Gamma(s) x^{-s} ds \end{aligned}$$

e, in generale, sotto opportune condizioni di sommabilità per f si ha

$$\begin{aligned} \mathcal{M}(f, s) &= \int_0^{+\infty} f(x) x^{s-1} dx \iff \\ f(x) &= \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \mathcal{M}(f, s) x^{-s} ds \end{aligned}$$

Prendendo formalmente $y = \log x$ nell'integrale

a sinistra si trova la trasformazione di Fourier. Per la precisione, $\psi(x^{-1})$ è l'anti-trasformata di Mellin di $-s^{-1} \zeta' / \zeta(s)$.

Nota 8

La dimostrazione dettagliata del Teorema dei Numeri Primi 18 si trova in Davenport [21].

Nota 9

La dimostrazione elementare del Teorema dei Numeri Primi è all'origine di una controversia sulla priorità fra Selberg, che ha certamente scoperto le formule che portano il suo nome, ed Erdős, che le ha usate per dedurne, probabilmente per primo, il Teorema dei Numeri Primi. Per i dettagli e per informazioni di prima mano, si veda Goldfeld [29]. La dimostrazione elementare è presentata nel Capitolo 22 di Hardy & Wright [30] e in Levinson [31]. È anche istruttivo leggere la recensione che Ingham [32] scrisse dei due lavori di Selberg e di Erdős, nella quale mette in luce le analogie strutturali tra la dimostrazione elementare e quella analitica. Ulteriori dettagli sull'equivalenza dimostrata da Wiener si trovano nel libro di Ingham [33], §2.11. Una semplice dimostrazione del Teorema dei Numeri Primi che sfrutta solo proprietà della funzione zeta in una regione limitata del semipiano $\sigma \geq 1$, e dunque nello spirito di quella originale di Wiener, si trova in Newman [34]. Si veda anche Zagier [35].

La Congettura di Riemann

Nota 10

Si vedano Bombieri [36] e Conrey [37] per delle ampie panoramiche in cui la Congettura di Riemann è inquadrata nel suo contesto generale.

È opportuno notare che non soltanto $\pi(x) - \text{li}(x)$ cambia segno infinite volte, ma, più in generale, $\pi(x) - \text{li}(x) = \Omega_{\pm}(x^{1/2}(\log \log \log x) / \log x)$ (Littlewood [5]), dove $f = \Omega_{\pm}(g)$ significa che $\liminf_{x \rightarrow +\infty} f(x)/g(x) < 0$ e $\limsup_{x \rightarrow +\infty} f(x)/g(x) > 0$. Un risultato più forte vale se la Congettura di Riemann è *falsa*. Vedi anche Ingham [33] §§4.8–4.9.

Nota 11

Si veda la discussione nel Capitolo 4 di Ingham [33].

Nota 12

Si veda Gourdon [2]. La funzione ξ è reale sulla retta $\sigma = \frac{1}{2}$ ed esiste una formula accurata per la funzione $N(T)$ definita nella (6): si contano i cambiamenti di segno di ξ fra $s = \frac{1}{2}$ ed $s = \frac{1}{2} + iT$. Si veda Conrey [3] per il secondo risultato enunciato. Si può dimostrare che il numero di zeri $\rho = \beta + i\gamma$ con $\gamma \in [0, T]$ e $\beta \neq \frac{1}{2}$ è $o(N(T))$. I cosiddetti *Teoremi di densità* forniscono stime molto precise per il loro numero: naturalmente, se è vera la Congettura di Riemann non ve ne sono affatto. Bays & Hudson [38] hanno dimostrato nel 2000 che $\pi(x) > \text{li}(x)$ per qualche $x \leq 1.39822 \cdot 10^{316}$.

Nota 13

Questa caratterizzazione della Congettura di Riemann è oggetto di risultati recenti di numerosi autori, tra cui citiamo Baez-Duarte & Balazard.

Nota 14

La stima elementare citata è il Teorema 333 di Hardy & Wright [30]. La dimostrazione della prima delle (16) si trova nel §5.6 di Edwards [24]. La Congettura di Mertens originale è falsa: ne abbiamo dato una versione leggermente indebolita.

Diamo un cenno alla dimostrazione delle formule (17) e (18): la prima si dimostra osservando che $\mu(pm) = -\mu(m)$ se p è un numero primo che non divide m . Usando iterativamente questa proprietà, se $n > 1$ il primo membro può essere riscritto come $(1 + \mu(p_1)) \cdots (1 + \mu(p_k))$ dove p_1, \dots, p_k sono i fattori primi *distinti* di n , e questa quantità vale evidentemente 0. Sia ora $f(x)$ il primo membro della (18) e osserviamo che f è costante a tratti e ovunque continua a destra su $\mathbb{R}^+ \setminus \mathbb{N}$ (poiché $[x/n] = \lfloor [x]/n \rfloor$ per ogni $x \geq 1$ e per ogni $n \in \mathbb{N}^*$), e che su \mathbb{N}^* può avere discontinuità di salto. Per concludere è dunque sufficiente dimostrare che f è continua a sinistra in tutti gli N naturali con $N \geq 2$. Scelto $\varepsilon > 0$

sufficientemente piccolo, osserviamo che

$$\begin{aligned} & f(N) - f(N - \varepsilon) \\ &= \sum_{n \leq N - \varepsilon} \left(\left\lfloor \frac{N}{n} \right\rfloor - \left\lfloor \frac{N - \varepsilon}{n} \right\rfloor \right) \mu(n) + \mu(N) \\ &= \sum_{n|N} \mu(n) = 0 \end{aligned}$$

e dunque la formula (18) segue dalla (17). Il passaggio cruciale della catena di uguaglianze qui sopra è quello in cui si riconosce che se $N \geq 2$ è intero ed $\varepsilon \in (0, 1)$, allora $\lfloor N/n \rfloor = \lfloor (N - \varepsilon)/n \rfloor$ se n non è un divisore di N .

Una dimostrazione alternativa richiede di scrivere $[x/n] = \sum_{m \leq x/n} 1$ nella definizione di f e di procedere con lo scambio delle somme ed un riordinamento degli addendi, un trucco standard della Teoria dei Numeri. Concludiamo notando che la (17) può essere interpretata come una sorta di ortogonalità rispetto al prodotto scalare (n, m) , massimo comun divisore fra n ed m .

Nota 15

La formula (19) segue dalla (18) mediante scambio della somma, usando il fatto che

$$\begin{aligned} \sum_{m \leq x} M\left(\frac{x}{m}\right) &= \sum_{m \leq x} \sum_{n \leq x/m} \mu(n) \\ &= \sum_{n \leq x} \mu(n) \sum_{m \leq x/n} 1 = \sum_{m \leq x} \mu(m) \left\lfloor \frac{x}{m} \right\rfloor \end{aligned}$$

Infatti, se $x_n = a/q$ con a e q primi fra loro e $q \leq N$, allora $f(x_n)$ compare al secondo membro della (19) con un peso $M(N/q) + M(N/2q) + M(N/3q) + \dots$, dove la somma è finita perché $M(x) = 0$ per $x < 1$, e si usa la relazione qui sopra con $x = N/q$. Tutti i dettagli si possono trovare in Edwards [24] §12.2.

Nota 16

Per i dettagli si veda Lagarias [39].

I problemi di Landau

I quattro problemi di Landau sono discussi dettagliatamente in Pintz [40].

Nota 17

L'osservazione di Fermat è il Teorema 251 di Hardy & Wright [30]. La Congettura di Hardy & Littlewood è un caso particolare della Congettura di Schinzel & Sierpiński sui valori primi assunti da polinomi in una variabile.

Congettura 35 (Schinzel & Sierpiński). Dato $f \in \mathbb{Z}[n]$ di grado $d \geq 1$ sia

$$\omega(p) = \text{card}(\{f(\mathbb{N}) \bmod p\}) = \text{card}(\{h \bmod p : \exists n \in \mathbb{Z} \text{ t.c. } f(n) \equiv h \bmod p\})$$

Se f è irriducibile su \mathbb{Z} ed $\omega(p) < p$ per tutti i numeri primi p , allora f assume valore primo per infiniti $n \in \mathbb{N}$.

Le due condizioni su f sono evidentemente necessarie, e la seconda è automaticamente verificata per $p > d$. In generale, Bateman & Horn hanno addirittura congetturato che, se sono soddisfatte le ipotesi della congettura di Schinzel & Sierpiński ed f ha primo coefficiente positivo, allora

$$\text{card}(\{n \in \mathbb{N} : f(n) \leq x \text{ ed è primo}\}) \sim C(f) \frac{x^{1/d}}{\log x} \quad \text{per } x \rightarrow +\infty,$$

dove $C(f) > 0$ è un'opportuna costante esplicita che può essere definita mediante un prodotto infinito sui numeri primi che dipende essenzialmente dai valori $\omega(p)$. Si vedano le note per il Capitolo introduttivo di Halberstam & Richert [41].

Nota 18

Un'argomentazione euristica elementare in sostegno delle formule asintotiche congetturate da Hardy & Littlewood nel secondo e terzo problema di Landau si può trovare in Zaccagnini [42].

Nota 19

Citiamo i risultati in norma L^2 per $\psi(x+y) - \psi(x) - y$: poniamo

$$J(x, y) = \int_x^{2x} |\psi(t+y) - \psi(t) - y|^2 dt$$

una quantità che potremmo chiamare varianza dei numeri primi. Selberg [43] ha dimostrato che la Congettura di Riemann implica che $J(x, y) = \mathcal{O}(xy(\log(2x/y))^2)$, uniformemente per $1 \leq y \leq x$. Questo significa che la formula asintotica vale "quasi ovunque" anche per y molto piccolo rispetto ad x , se è vera la Congettura di Riemann. In particolare, la relazione $\psi(t+y) - \psi(t) \sim y$ vale per quasi tutti i $t \in [x, 2x]$ non appena $y \geq (\log x)^{2+\varepsilon}$.



- [1] G. F. B. RIEMANN: Über die Anzahl der Primzahlen unter einer gegebenen Grösse. *Monatsber. Königl. Preuss. Akad. Wiss. Berlin*, pages 671–680, 1859. In "Gesammelte Mathematische Werke" (ed. H. Weber), Dover reprint 1953.
- [2] X. GOURDON: "The 10^{13} first zeros of the Riemann zeta-function, and zeros computation at very large height", (2004). Available from: <http://numbers.computation.free.fr/Constants/Miscellaneous/zetazeros1e13-1e24.pdf>.
- [3] J. B. CONREY: "More than two fifths of the zeros of the Riemann zeta function are on the critical line", *J. reine angew. Math.* **399** (1989) 1–26.
- [4] M. HUXLEY: *The Distribution of Prime Numbers*, Clarendon Press, Oxford (1972).
- [5] J. E. LITTLEWOOD: "Sur la distribution des nombres premiers", *C. R. Acad. Sc. Paris* **158** (1914) 1869–1872.
- [6] A. BEURLING: "A closure problem related to the Riemann zeta-function", *Proc. Acad. Sci. U.S.A.* **41** (1955) 312–314.
- [7] G. H. HARDY AND J. E. LITTLEWOOD: "Some problems in 'Partitio Numerorum'; III. On the expression of a number as a sum of primes", *Acta Math.* **44** (1923) 1–70.
- [8] J. FRIEDLANDER AND H. IWANIEC: "The polynomial $x^2 + y^4$ captures its primes", *Ann. Math.* **148** (1998) 945–1040.
- [9] D. R. HEATH-BROWN: "Primes represented by $x^3 + 2y^3$ ", *Acta Math.* **186** (2001) 1–84.
- [10] J. PINTZ: *Recent results on the Goldbach conjecture*, In W. Schwarz and J. Stending, editors, *Proceedings of the ELAZ Conference, May 24–28, 2004*, pages 220–254, Stuttgart, Franz Steiner Verlag (2006).
- [11] I. M. VINogradov: "Some theorems concerning the theory of primes", *Mat. Sb.* **N.S.2** (1937) 179–195.
- [12] H. A. HELFGOTT: "Major arcs for Goldbach's problem", *Arxiv preprint* (2013) 1305.2897.
- [13] Y. ZHANG: "Bounded gaps between primes", *Ann. Math.* (2013) accettato per la pubblicazione.
- [14] J. MAYNARD: "Small gaps between primes", *Arxiv preprint* (2013) 1311.4600.

- [15] R. C. BAKER, G. HARMAN, AND J. PINTZ: "The difference between consecutive primes, II", *Proc. London Math. Soc.* (3) **83** (2001) 532–562.
- [16] P. T. BATEMAN AND H. G. DIAMOND: "A hundred years of prime numbers", *Amer. Math. Monthly* **103** (1996) 729–741.
- [17] M. DELÉGLISE AND J. RIVAT: "Computing $\pi(x)$: The Meissel, Lehmer, Lagarias, Miller, Odlyzko method", *Math. Comp.* **65** (1996) 235–245.
- [18] P. DUSART: "The k -th prime is greater than $k(\ln k + \ln \ln k - 1)$ for $k \geq 2$ ", *Math. Comp.* **68(225)** (1999) 411–415.
- [19] J. B. ROSSER: "Explicit bounds for some functions of prime numbers", *Amer. J. Math.* **63** (1941) 211–232.
- [20] J. PINTZ: "On Legendre's prime number formula", *Amer. Math. Monthly* **87** (1980) 733–735.
- [21] H. DAVENPORT: *Multiplicative Number Theory*, Graduate Texts in Mathematics, v.74, III ed, Springer (2000).
- [22] M. NAIR: "On Chebyshev-type inequalities for primes", *Amer. Math. Monthly* **89** (1982) 126–129.
- [23] A. M. LEGENDRE: *Essai sur la Théorie des Nombres*, Paris, III ed (1830).
- [24] H. M. EDWARDS: *Riemann's Zeta Function*, Dover, ristampa dell'edizione Academic Press, 1974 (2001).
- [25] E. C. TITCHMARSH: *The Theory of the Riemann Zeta-Function*, Oxford University Press, Oxford, II ed (1986).
- [26] G. H. HARDY: *Divergent Series*, Chelsea, New York, II ed (1991).
- [27] A. IVIĆ: *The Theory of the Riemann Zeta-Function*, J. Wiley, New York (1985).
- [28] D. ZAGIER: "The first 50 million prime numbers", *The Mathematical Intelligencer* **0** (1977) 7–19.
- [29] D. GOLDFELD: *The elementary proof of the prime number theorem: an historical perspective*, In "Number theory (New York, 2003)", p. 179–192, Springer, New York (2004).
- [30] G. H. HARDY AND E. M. WRIGHT: *An Introduction to the Theory of Numbers*, Oxford Science Publications, Oxford, V ed (1979).
- [31] N. LEVINSON: "A motivated account of an elementary proof of the Prime Number Theorem", *Amer. Math. Monthly* **76** (225–245) 1969.
- [32] A. E. INGHAM: "Review", *Mathematical Reviews* **10** (1949) 595–596.
- [33] A. E. INGHAM: *The Distribution of Prime Numbers*, Cambridge University Press, Cambridge (1990).
- [34] D. J. NEWMAN: "Simple analytic proof of the Prime Number Theorem", *Amer. Math. Monthly* **87** (1980) 693–696.
- [35] D. ZAGIER: "Newman's short proof of the prime number theorem", *Amer. Math. Monthly* **104** (1997) 705–708.
- [36] E. Bombieri. "Problems of the Millennium: the Riemann Hypothesis" (2000). Si veda: http://www.claymath.org/prize_problems/index.html.
- [37] J. B. CONREY: "The Riemann Hypothesis", *Notices Amer. Math. Soc.* **50** (2003) 341–353.
- [38] C. BAYS AND R. H. HUDSON: "A new bound for the smallest x with $\pi(x) > \text{li}(x)$ ", *Math. Comp.* **69(231)** (2000) 1285–1296.
- [39] J. LAGARIAS: "An elementary problem equivalent to the Riemann Hypothesis", *Amer. Math. Monthly* **109** (2002) 534–543.
- [40] J. PINTZ: "Landau's problems on primes", *Journal de théorie des nombres de Bordeaux* **21** (2009) 357–404.
- [41] H. HALBERSTAM AND H.-E. RICHERT: *Sieve Methods*, Academic Press, London (1974).
- [42] A. ZACCAGNINI: "Variazioni Goldbach: problemi con numeri primi", *L'Educazione Matematica, Anno XXI, Serie VI* **2** (2000) 47–57. http://people.math.unipr.it/alessandro.zaccagnini/psfiles/papers/Goldbach_I.pdf
- [43] A. SELBERG: "On the normal density of primes in small intervals, and the difference between consecutive primes", *Arch. Math. Naturvid.* **47** (1943) 87–105.

Alessandro Zaccagnini: Professore associato di Analisi Matematica all'Università di Parma, si occupa di Teoria Analitica dei Numeri. Ha pubblicato principalmente risultati su problemi additivi del tipo di quello di Goldbach e sue varianti, e sulle connessioni fini tra posizione degli zeri della funzione zeta di Riemann e il termine di resto in varie forme del Teorema dei Numeri Primi. Ha anche pubblicato diversi saggi e articoli divulgativi.

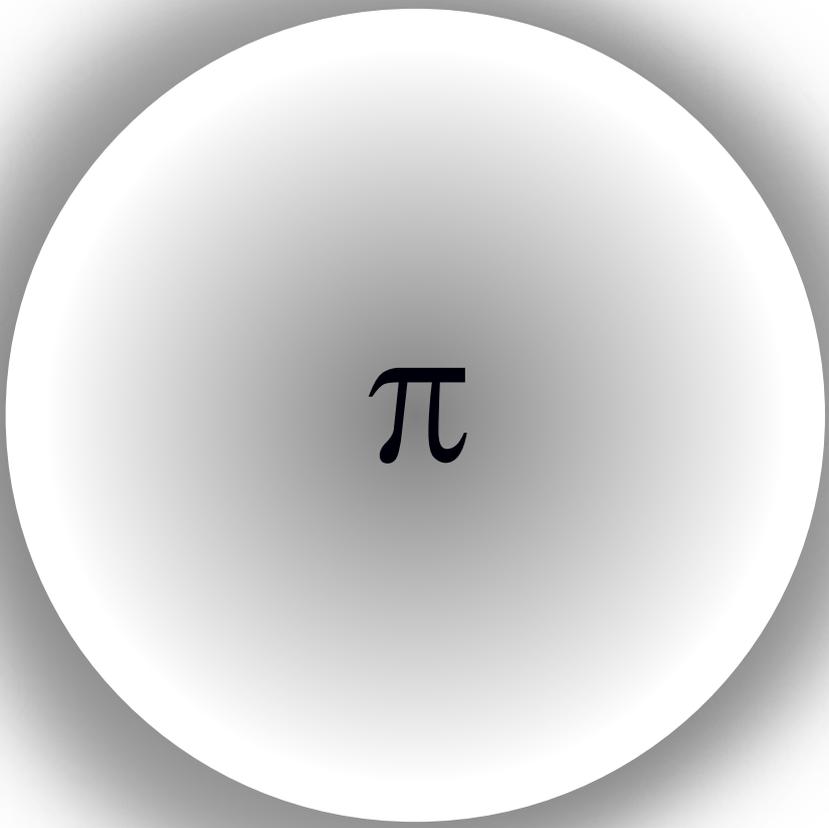
Numero III Anno 2014

Viaggio nella Scienza

Ithaca



Arte e Scienza



π