

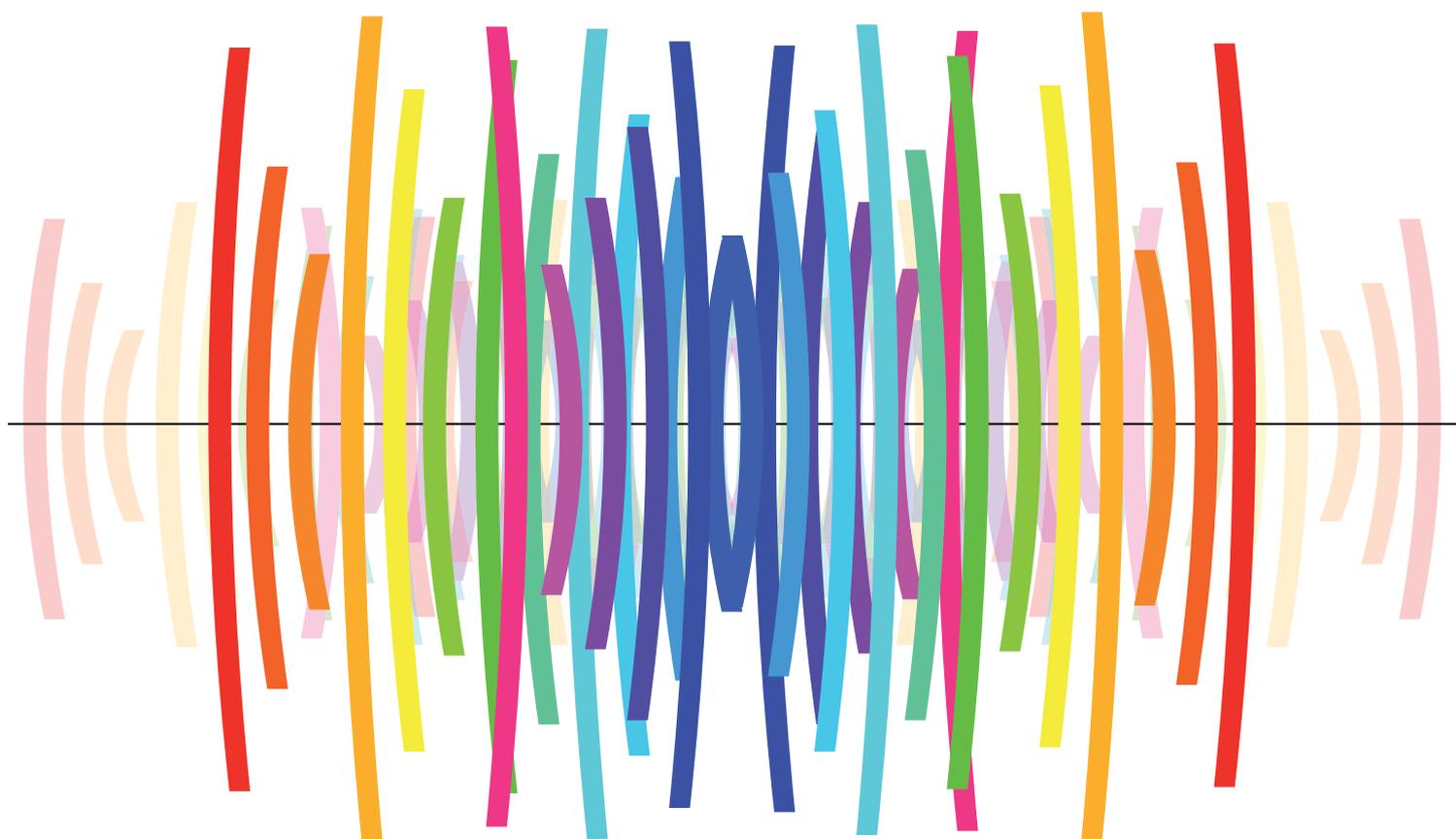
Numero VII
Anno 2016


DIPARTIMENTO DI MATEMATICA E FISICA
"Ennio De Giorgi"


UNIVERSITÀ
DEL SALENTO

Viaggio nella Scienza

Ithaca



Luce

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,
Paolo Ciafaloni,
Maria Luisa De Giorgi,
Vincenzo Flaminio,
Luigi Martina,
Giuseppe Maruccio,
Marco Mazzeo,
Francesco Paparella,
Carlo Sempi.

Segreteria di Redazione
Daniela Dell’Anna.

© 2013-2016 Dipartimento di Matematica e Fisica “*Ennio de Giorgi*”.

© 2016 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

Ithaca

Viaggio nella Scienza

VII 2016

3 In questo numero

Fotonica

5 Il lato oscuro della luce

Marco Mazzeo

59 Iride

Luigi Martina

83 Luce e Quantum Dot

Vittorianna Tasco, Luigi Carbone

99 Organic Light-Emitting Diodes: la nuova era della luce

Fabrizio Mariano

119 I Metamateriali: Invisibilità e Rifrazione Negativa in Mezzi Nanochirali

Marco Esposito, Francesco Todisco

131 Strong Coupling: l'entanglement tra luce e materia

Armando Genco

147 Neutrini e Cosmologia
Gianpiero Mangano

167 Ascoltando l'Universo di onde gravitazionali:
implicazioni e prospettive dopo la scoperta
Paola Leaci

In questo numero

Il 2015 è stato proclamato dall'Assemblea Generale delle Nazioni Unite, in collaborazione con l'UNESCO e con un gran numero di enti di ricerca, *Anno della Luce*. Per questo motivo, abbiamo pensato di scegliere la *Luce* come argomento principale del primo numero di *Ithaca* del 2016. L'articolo di Marco Mazzeo è una rassegna sulle attuali conoscenze ed interpretazioni, fisiche e filosofiche, dei fenomeni luminosi, sia in un contesto classico che quantistico. L'articolo di Luigi Martina presenta storia, fisica, leggenda e filosofia di uno dei fenomeni naturali che affascina tutti gli esseri umani: l'arcobaleno. Gli altri articoli sulla luce mettono in evidenza aspetti della scienza e tecnologia moderne legati all'uso della luce. L'articolo di Vittorianna Tasco e Luigi Carbone discute di come i materiali si comportano in termini di emissione e assorbimento della luce quando la loro dimensionalità viene modificata. L'articolo di Fabrizio Mariano descrive gli OLED, diodi organici emettitori di luce, oggetti ampiamente utilizzati nella moderna attività di ricerca, e presto utilizzati anche in prodotti industriali ad alta tecnologia. Marco Esposito e Francesco Todisco raccontano cosa siano i metamateriali e come, mediante la loro interazione con la luce, possano consentire di violare la legge di Snell con ricadute applicative come l'invisibilità. L'articolo di Armando Genco presenta fenomeni di interazione luce-materia in regime di "accoppiamento forte" in cui possono essere generati nuovi stati quantici *entangled*, detti polaritoni, in cui materia e luce sono miscelati.

Gli ultimi due articoli di questo numero si discostano dal tema principale. L'articolo di Gianpiero Mangano, a completamento del numero VI di *Ithaca*, parla del ruolo dei neutrini in cosmo-

logia. L'articolo di Paola Leaci è strettamente legato all'attualità scientifica e presenta la recente scoperta delle Onde Gravitazionali.

Buona lettura,
il Comitato di Redazione

Il lato oscuro della luce

Marco Mazzeo

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

A causa del suo carattere ineffabile, etereo è il caso di dire, e della sua importanza nel mostrarci il mondo esterno attraverso quelle finestre biologiche chiamate occhi, la luce ha da sempre suscitato nell'animo umano un senso di mistero. Probabilmente nessun ente fisico ha posto così tanti interrogativi come la luce. Se qualcuno dovesse chiedermi "cosa è la luce?" tutt'ora non saprei rispondere. L'unica risposta sensata sarebbe infatti "non sappiamo nella realtà cosa la luce sia ma abbiamo compreso parte del suo comportamento". Come vedremo tuttavia il comportamento fisico della luce è alquanto bizzarro e il suo studio ci ha costretto a rinunciare a visioni sulla realtà che si sono rivelate inadeguate e semplicistiche. Nel corso della storia della fisica sono state sviluppate almeno quattro teorie della luce, ciascuna che rendeva conto di fatti sperimentali ben precisi, e ognuna che inglobava quella precedente. Le teorie sono schematizzate in Fig.1. In questo lavoro esploreremo le teorie e gli esperimenti chiave effettuati sulla luce e le loro conseguenze nella nostra comprensione della realtà fisica.

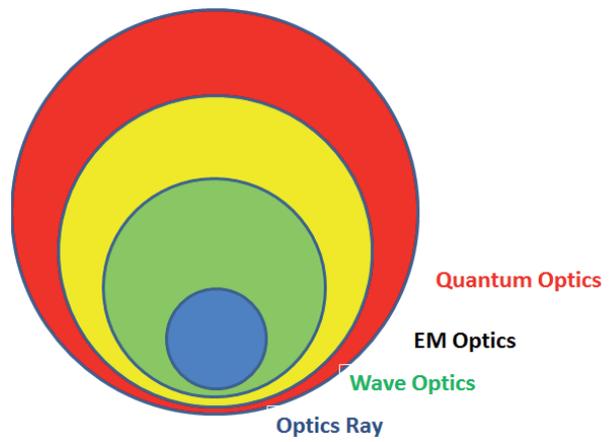


Figura 1: Le quattro teorie fisiche sulla luce: ottica geometrica, ottica ondulatoria, ottica elettromagnetica (EM) e ottica quantistica.

Parte I: teorie classiche della luce

La Luce è fatta di raggi: l'ottica geometrica

Sono in auto, appena partito da Lecce per Taranto, ma con l'intenzione di fare una tappa a Manduria al ghetto ebraico. In tempi come questi rivisitare la storia tra due tradizioni religiose monoteiste, che per quasi duemila anni si sono scontrate, non fa mai male. Non ricordo tuttavia il quartiere e imposto da subito il GPS, sperando che quei satelliti lassù non mi facciano sbagliare

strada. Nel frattempo sento una gran voglia di ascoltare i Pink-Floyd lungo il tragitto. Prendo il CD dal cruscotto dell'auto. Di fronte a me il Sole ha iniziato il suo declino verso l'orizzonte, illuminando l'asfalto della superstrada verso Brindisi. La sua luce, dopo aver colpito il CD, tinge l'abitacolo dei colori dell'arcobaleno. Inserisco il CD nel lettore e il laser inizia a leggere i fori micrometrici del CD convertendo il tutto in una sfilza di 1 e 0 che saranno poi trasdotti in musica. Mi metto comodo ma la luce del Sole riflessa dalla superstrada mi abbaglia e decido così di indossare i miei occhiali da Sole polarizzati. Ora il comfort è notevole e posso proseguire dritto. Mentre "Money" risuona nell'auto mi viene in mente che tutto ciò che ho fatto ha a che fare con l'ottica e con le quattro teorie sulla luce sviluppate in quasi duemilatrecento anni. Il lettore laser "legge" i fori del CD grazie all'ottica geometrica, i colori dell'arcobaleno del CD sono apparsi nell'abitacolo seguendo le leggi dell'ottica ondulatoria, il riflesso sull'asfalto viene bloccato dagli occhiali mediante meccanismi che seguono l'ottica elettromagnetica, e infine il GPS mi avverte di girare a sinistra localizzandomi con gran precisione grazie alla comprensione delle leggi dello spaziotempo locale e della fisica quantistica degli orologi atomici, entrambe teorie che hanno la luce come gran parte del loro oggetto di studio.

Il lettore CD della nostra auto si basa anzitutto su un diodo Laser che invia un fascio di luce che si propaga seguendo percorsi rettilinei fino a colpire uno specchio ed essere riflesso verso il CD da leggere. Il sistema non funzionerebbe se la luce non si propagasse seguendo linee rette e la prima teoria sviluppata dalla fisica sul comportamento della luce ha proprio a che vedere con i percorsi dei raggi di luce. Fu formulata nei suoi principi basilari già da Archimede di Siracusa prima e dall'arabo Alhazen in seguito. Il secondo tuttavia scoprì, contrariamente a ciò che credevano i greci, che i raggi di luce non sono prodotti dagli occhi come fossero delle propaggini sensoriali, bensì essi si muovono dall'oggetto agli occhi, seguendo percorsi rettilinei. Il principio che generalizza questo comportamento e che spiega gran parte di ciò che un raggio di luce fa è detto principio di Fermat,

un raggio di luce segue sempre il percorso di tempo

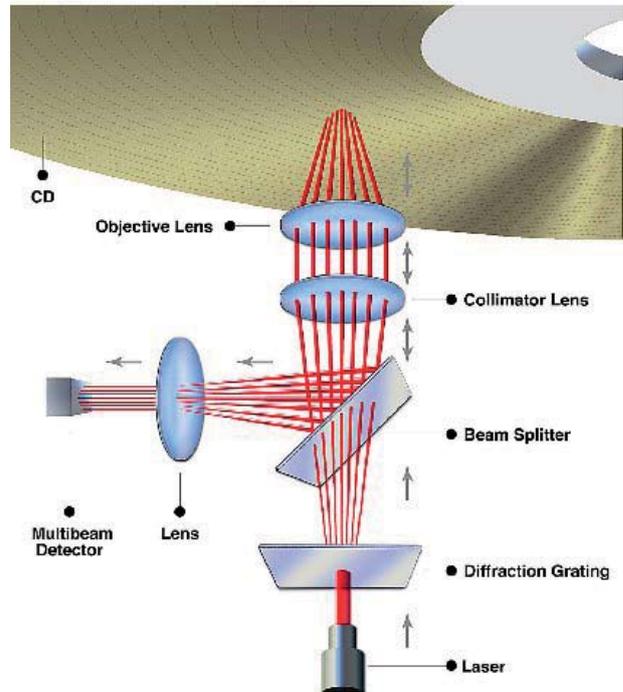


Figura 2: Principio di funzionamento di un lettore CD.

più breve

Dal punto di vista matematico è possibile esprimere questo principio dicendo che il seguente integrale

$$\int_A^B n(\mathbf{r}) d\mathbf{l} \quad (1)$$

è un *estremale* attorno al cammino che la luce percorrerà nella realtà. Estremale vuol dire che l'integrale cambia poco se lo calcoliamo lungo un cammino diverso ma vicino a quello reale. Il coefficiente $n(\mathbf{r})$ è detto indice di rifrazione (superiore a 1) ed è un parametro definito come il rapporto tra la velocità della luce c nel vuoto e la velocità della luce nel mezzo v , ovvero c/v , e che tiene conto del rallentamento della luce in un mezzo rispetto al vuoto.

Il principio di Fermat, è alla base dell'ottica geometrica e spiega molti fenomeni conosciuti, come ad esempio il perché la luce si propaghi in linea retta nel vuoto. Dati due punti A e B nello spazio la traiettoria di minimo percorso coincide con quello di minimo tempo e non può che essere la retta passante per quei due punti. Questo accade anche in qualsiasi altro mezzo a patto però che l'indice di rifrazione non dipenda dalla posizione r . Se n infatti è costante il principio del minimo tempo diventa quello di mini-

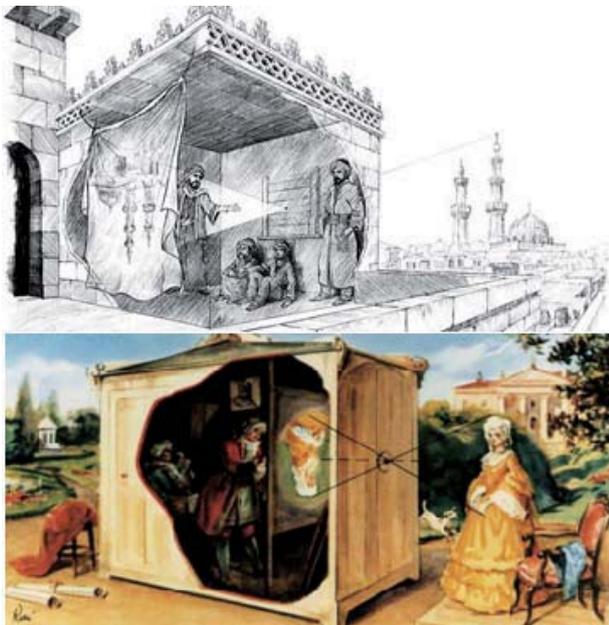


Figura 3: Principio di funzionamento della camera oscura, compreso da Alhazen e poi adoperato da molti pittori in occidente per riprodurre su tela immagini esterne di gran precisione.

mo percorso in quanto la velocità della luce non cambia nello spazio. In altre parole possiamo portare n fuori dall'integrale di (1) e l'estremale sarà effettuato sul percorso fisico e non su quello pesato dall'indice di rifrazione, fornendo come risultato un percorso rettilineo. Questo semplice principio spiega come mai le ombre sono nette e perché si formano i coni d'ombra durante una eclisse. Se tutto ciò accade è solo perché la luce descrive delle linee rette nello spazio. Lo sapevano benissimo appunto il musulmano Alhazen e il cristiano Leonardo da Vinci, quando usavano entrambi una camera forata per proiettare una scena reale su un pannello. La scienza supera i confini imposti dalle tradizioni. Se si pratica un forellino in una stanza buia e attraverso di esso passano i raggi luminosi provenienti dall'esterno sulla parete opposta della stanza si vedrà l'ambiente di fuori capovolto. La spiegazione risiede proprio nel principio di cui sopra, come potete vedere dalla figura 3.

Anche la prospettiva si basa sul principio di Fermat. Alla base di questa tecnica pittorica c'è il fatto che un corpo che viene distanziato del doppio apparirà proporzionalmente più piccolo della metà, il che, ovviamente, è una derivazione del moto rettilineo dei raggi luminosi in un mezzo otticamente isotropo. Questo principio è

alla base anche delle eclissi di Sole. Già Aristarco di Samo aveva compreso questo, come potete osservare dalla figura 4a tratta dal libro "sulle dimensioni e le distanze del Sole e della Luna".

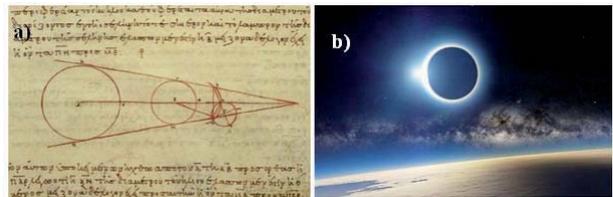


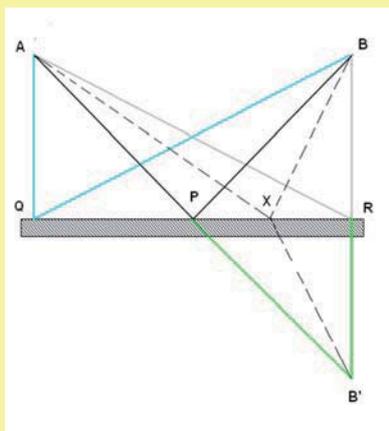
Figura 4: a) Schema delle eclissi riportato in una copia dell'unico testo di Aristarco di Samo sopravvissuto alla distruzione della biblioteca di Alessandria "Sulle dimensioni e le distanze del Sole e della Luna". I raggi dimostrano che Aristarco conosceva benissimo l'ottica geometrica. b) L'eclisse di Sole funziona a causa del principio di Fermat e dell'evento casuale per cui la luna è 390 volte più vicina del Sole alla Terra, ma anche 390 volte più piccola.

Nel lettore CD rom di fig. 2 la luce viene sia riflessa da degli specchi e sia rifratta da delle lenti. Come mai? La spiegazione risiede ancora nel principio di Fermat, fornita per entrambi i fenomeni rispettivamente nei Box 1 e 2.

Il principio del tempo minimo descrive perfettamente il perché la lente ha la forma appunto di "lenticchia". Infatti ci si può chiedere: esiste un congegno fatto di un pezzo di materia come il vetro che riesce a deviare i raggi di luce in modo da concentrarli tutti in un punto per accumulare lì l'energia ad esempio dei raggi solari? Perché ciò accada è ovvio che se da A si diramano tanti raggi di luce che devono convergere tutti nel punto B allora essi dovranno impiegare lo stesso tempo. Un corpo avente spessore costante non sarebbe adatto in quanto al suo interno la luce impiegherebbe lo stesso tempo, ma al di fuori di esso i raggi incidenti percorrerebbero tratti più o meno inclinati a partire dalla sorgente in A e perciò impiegheranno tempi diversi. E' necessario allora che i tragitti più corti siano rallentati maggiormente di quelli più lunghi. Il raggio più breve, quello assiale, dovrà attraversare il tratto di vetro maggiore per essere rallentato nella giusta misura. Una soluzione possibile è una lente biconvessa, ovvero un disco di vetro bombato al centro, a forma appunto di lenticchia. Nel lettore CD le lenti servono a focalizzare il fascio laser in poche centinaia di micron quadrati, riuscen-

Box 1: riflessione nell'ottica geometrica

Prediamo una superficie piana metallica come l'argento, di cui sono fatti appunto gli specchi, e analizziamo come mai la luce si riflette. Inviama un raggio di luce dal punto A al punto B mediante una riflessione su uno specchio e chiediamoci: qual'è il percorso AXB più breve? E' quello in cui X è uguale a P ed è proprio al centro? Oppure è quello in cui X=Q, o quello in cui X=R, o uno intermedio? Ebbene non è semplice rispondere a questo quesito perché, se ad esempio spostiamo il punto X verso destra rispetto a P, (come nel caso del percorso tratteggiato) il tratto AX aumenta mentre XB diminuisce. Non possiamo dire a priori che AXB è più lungo o più breve di qualsiasi altro, ma un metodo c'è. Proiettiamo il punto B in B' in modo tale cioè che B' sia alla stessa distanza di B dalla superficie superiore dello specchio, cioè $BR=B'R$. Ora prendiamo un punto X qualsiasi sullo specchio. E' facile provare che i triangoli rettangoli BXR e B'XR sono uguali avendo gli stessi cateti BR e B'R e il cateto in comune XR. Allora anche le ipotenuse XB e XB' saranno uguali. Quindi, essendo AX comune a tutti i percorsi, chiedersi quale percorso AXB sia il più breve equivale a chiedersi quale percorso AXB' sia il più breve. E' ovvio che di tutti questi percorsi il più breve sarà il percorso APB', che è una retta. X cioè deve stare al centro, ovvero in P. Ma se APB' è una retta allora i due angoli APQ e B'PR sono uguali in quanto opposti al vertice P. Infine essendo i triangoli B'PR e BPR uguali anche gli angoli corrispondenti B'PR e BPR saranno uguali tra loro. In definitiva concludiamo che l'angolo $APQ=BPR$. Cioè l'angolo di incidenza è uguale a quello di riflessione. Che è la ben nota legge degli specchi.



do a localizzare un punto da leggere sul CD con maggior precisione, garantendo così la possibilità di introdurre più dati leggibili sulla stessa superficie. Ma la luce non è solo un tragitto. E' anche colore, intensità. Come possono essere spiegate queste caratteristiche? Circa l'intensità essa non è altro se non la densità di raggi che attraversa una data superficie. Laddove i raggi si addensano l'intensità diventa grande. Il discorso è molto simile a ciò che accade nel caso del rapporto tra campo elettrico e linee di forza. Laddove le linee si addensano il campo è maggiore. L'associazione non è casuale. Infatti un

altro modo equivalente di esprimere il principio di Fermat è di introdurre la superficie S. Tale superficie corrisponde al fronte d'onda ovvero il luogo dei punti che vibrano in fase e che risulta perpendicolare ai raggi luminosi. E' possibile dimostrare in modo molto semplice che il principio di Fermat è equivalente a dire che

$$|\nabla S| = n(\mathbf{r}) \quad (2)$$

Questa equazione è detta *equazione Iconale*, e mostra come l'indice di rifrazione determini la direzione dei raggi luminosi, che è espressa appunto dalla direzione del gradiente del fronte

Box 2: rifrazione nell'ottica geometrica

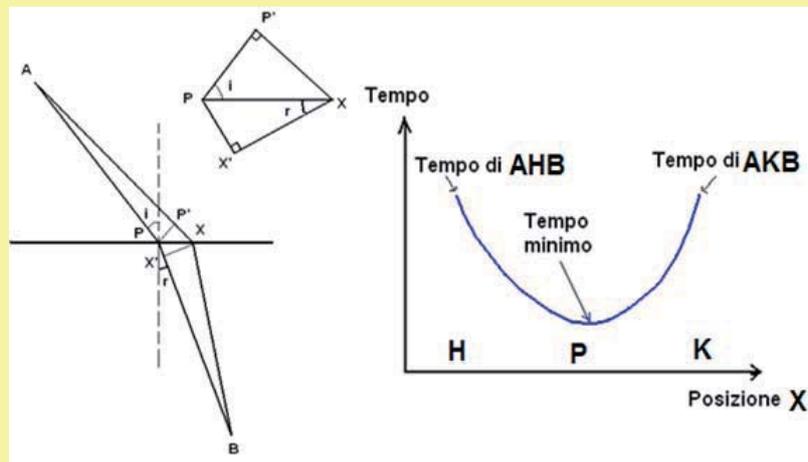
Se disegnassimo il tempo di percorrenza di ogni percorso in funzione della posizione del punto B dovremmo trovare un grafico del tipo mostrato in Fig.B2.b in cui il tempo richiesto non cambia tanto attorno al tragitto reale (in P). Prendiamo pertanto la curva reale che la luce segue, ovvero APB. Immaginiamo di scostarci di poco e seguiamo il percorso AXB. Poiché X è molto vicino a P il tempo di percorrenza sarà simile per quanto detto prima. Se proiettiamo il punto P sul raggio AX trovando P' e lo stesso facciamo con X sul raggio PB trovando X' individueremo due triangoli rettangoli, PP'X e PXX' che disegniamo ingranditi a parte. Ora poiché i tempi di percorrenza saranno simili, ed essendo AP e AP' praticamente uguali così come BX e BX', ecco che anche P'X e PX' saranno percorsi praticamente nello stesso tempo, ovvero

$$\frac{P'X}{c} = \frac{PX'}{v}$$

Però osservando i triangoli disegnati a parte osserviamo che $P'X = PX \sin(i)$ e $PX' = PX \sin(r)$, quella uguaglianza diventa

$$\sin(i) = \frac{c}{v} \sin(r)$$

che è proprio la legge di Snell a patto di individuare in c/v proprio l'indice di rifrazione n .



d'onda S. Sappiamo che il campo elettrico è esprimibile a meno di un segno come il gradiente del potenziale elettrico, per cui possiamo dire che formalmente S sta al potenziale elettrico come n sta al modulo del campo elettrico. Ma cosa è il colore nella teoria dell'ottica geometrica? Ebbene poiché ogni colore viene rifratto ad angoli diversi possiamo dire che il colore esprime il grado di rifrazione di un raggio, ma nulla di più, come propose Newton [1]. La teoria dell'ottica geometrica è formidabile ma ci sono fenomeni che non si riescono davvero a spiegare.

La luce è fatta di “raggi di onde”: l'ottica ondulatoria

La luce si muove in linea retta se il mezzo non cambia indice di rifrazione. E' un principio così semplice che potremmo domandarci: e cos'altro avrebbe potuto fare la luce? Eppure il principio di minimo percorso non spiega un fenomeno comune che spesso accade e che osserviamo ma di cui spesso non ci domandiamo la causa. Se siete seduti di sera vicino ad una finestra dotata di una zanzariera o di una tenda e riuscite ad osservare un lampione acceso vedrete una strana figura (Figura 6), come se il lampione venisse

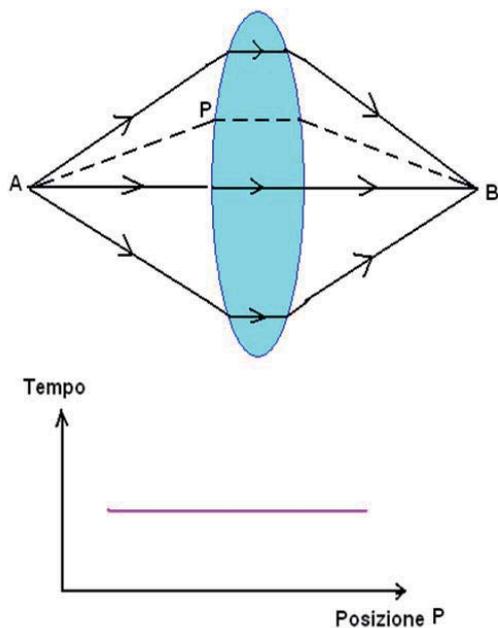


Figura 5: Principio di funzionamento di una lente biconvessa. Il tempo che impiegano i raggi a focalizzarsi è lo stesso.

riprodotto ai lati formando una sorta di croce con zone luminose e buie. Rimossa la tenda o la zanzariera vedrete solo il lampione. Vi siete mai chiesti perché la luce del lampione si comporta in questo modo?

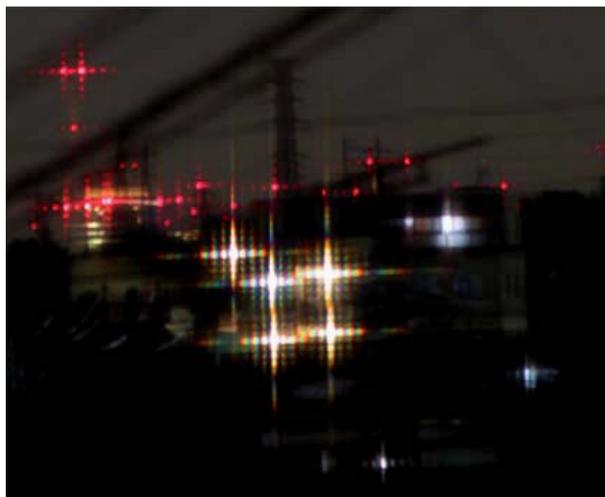


Figura 6: Interferenza e diffrazione della luce emessa da un lampione attraverso una tenda.

Il principio di minimo percorso non spiega questo fenomeno, chiamato interferenza. Se riuscite a procurarvi un laser potete fare degli esperimenti che vi lasceranno sconcertati e che sono collegati a quella figura che la luce del lampione

crea passando attraverso la tenda. Se accendete il Laser vedrete un raggio luminoso perfettamente rettilineo. Segue il principio di minima percorrenza, nulla di nuovo. Fate però attraversare il raggio laser attraverso la fessura creata tra due dita unite tra loro, ad esempio l'indice e il medio. Tale fessura cambia di larghezza perché in alcuni punti le dita combaciano meglio e in altri meno. Muovendo il raggio lungo la fessura vi accorgete che fintanto che la fessura è larga vi apparirà un punto luminoso sulla parete, ma quando diventa piccola quel punto diventa slargato e appaiono zone consecutive luminose e buie. Se vi procurate un capello e lo disponete in verticale e gli inviate contro il raggio laser ecco che non vi apparirà un'ombra rettilinea, corrispondente al capello, ma delle zone chiare e scure come in Fig.7. Come mai la zona luminosa si espande e appaiono zone chiare e scure?

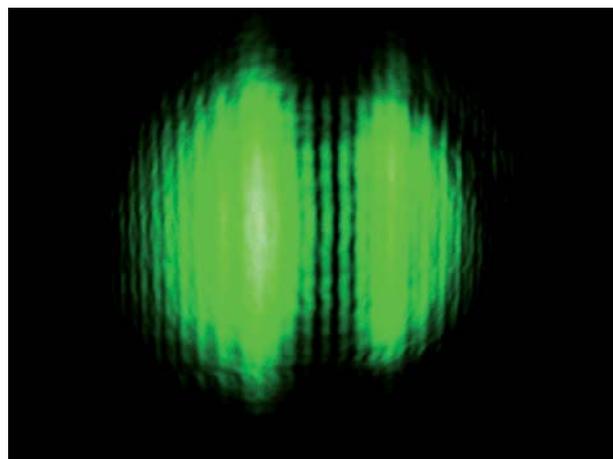


Figura 7: Interferenza e diffrazione della luce laser diffusa da un capello.

Inizia a venirci un dubbio: forse avevamo torto nell'ammettere che la luce segue il percorso più breve o di minimo tempo? Dovremmo forse modificare questa semplice idea che pure spiegava tantissimi fenomeni attorno a noi? Come spiegare quella periodicità nelle frange luminose e buie? Hans Christian Huygens e dopo di lui Thomas Young proposero che la luce fosse costituita da onde [2]. Huygens ipotizzò che la luce fosse costituita da un'onda in quanto se fosse fatta di corpuscoli che viaggiano di moto rettilineo nell'aria due raggi che si incrociano si perturberebbero a vicenda. E invece ciò che osserviamo è che essi continuano a percorrere due traiettorie rettilinee. Ma cosa è un'onda? Dal punto di vista

matematico un'onda $U(\mathbf{r},t)$ è una funzione dello spazio e del tempo che soddisfa la seguente equazione, detta appunto equazione d'onda

$$\nabla^2 U(\mathbf{r},t) = \frac{1}{v^2} \frac{\partial^2 U(\mathbf{r},t)}{\partial t^2} \quad (3)$$

dove v è la velocità di propagazione dell'onda. La lunghezza d'onda è la distanza tra due creste vicine dell'onda ed è indicata con la lettera greca λ . Abbiamo visto come nella teoria geometrica l'intensità fosse la densità spaziale dei raggi e il colore il grado di rifrazione. Nella teoria ondulatoria della luce invece il colore non è altro che la lunghezza d'onda: il blu si aggira attorno a lunghezze d'onda di 400nm, il verde di 500nm e il rosso di 600nm; l'intensità invece, l'energia dell'onda luminosa, non è altro se non il modulo quadro dell'ampiezza, ovvero $I = |U|^2$, come in qualsiasi onda che si rispetti. L'energia trasportata da un'onda di mare ad esempio sarà proporzionale al quadrato della sua ampiezza.

Un aspetto importante della equazione (3) è che, prese due soluzioni U_1 e U_2 , una qualsiasi loro combinazione lineare è ancora soluzione. Questo è il ben noto *principio di sovrapposizione*. Possiamo adoperare questa matematica per descrivere i fenomeni osservati attraverso la zanzariera? E come si inquadra il principio di Fermat in questa nuova teoria? Thomas Young propose il seguente esperimento della doppia fenditura [2], divenuto ormai famoso. Prendiamo una superficie, praticiamo due fenditure molto strette e vicine tra loro e inviamo un fascio luminoso. Cosa ci aspettiamo di osservare su uno schermo posto dall'altra parte? Si potrebbe pensare: due strisce luminose! E del resto questa sarebbe la risposta corretta se la luce seguisse sempre il principio di minima percorrenza. E questo è in effetti ciò che osserveremmo con fenditure troppo distanti e larghe (da qualche millimetro in su).

Ma le cose stanno diversamente se proviamo ad avvicinare le fenditure sino a valore di qualche decimo di millimetro. Quel che osserveremo è mostrato in fig. 9, ovvero della frange di interferenza periodiche chiare e scure distribuite sul pannello.

Una teoria geometrica non può spiegare nulla di tutto questo ma una teoria ondulatoria sì. Vediamo come. Le due onde U_1 e U_2 (di intensità

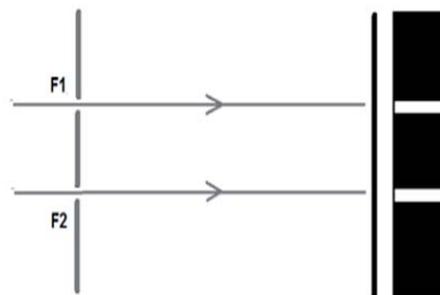


Figura 8: Due raggi di luce laser che attraversano due fenditure molto grandi e lontane produrranno due strisce luminose su uno schermo.

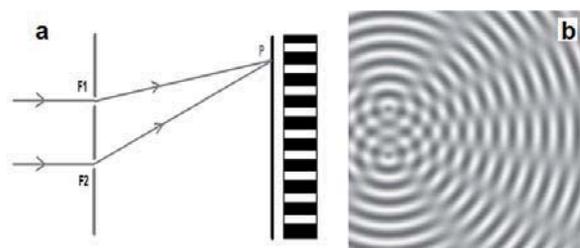


Figura 9: Due raggi di luce laser che attraversano due fenditure molto piccole e vicine produrranno una figura di interferenza come due sorgenti che creano delle increspature su una superficie di acqua.

I_1 e I_2) partono dalle fenditure F_1 ed F_2 rispettivamente (Fig.8 e 9) e là dove si intersecano esse si sommeranno algebricamente dando origine ad un'onda pari a U_1+U_2 . Pertanto avremo che l'intensità totale I , nel punto P della figura, è pari a

$$I = |U_1 + U_2|^2 = I_1 + I_2 + 2U_1 * U_2 \quad (4)$$

Come si vede luce+luce non sempre fa luce, a volte fa buio! La causa? Il principio di sovrapposizione e la proporzionalità quadratica dell'intensità dall'ampiezza. Infatti se il terzo addendo, detto *termine di interferenza*, è negativo l'intensità totale è inferiore alla somma delle due intensità, a tal punto che può anche annullarla. Se invece il termine è positivo l'intensità luminosa finale può perfino essere superiore a I_1+I_2 . Quando accade ciò? Se i tratti F_2P ed F_1P differiscono tra loro di un tratto che è un semplice multiplo di una lunghezza d'onda le onde si rafforzeranno dando

una interferenza costruttiva e più luce; se invece differiranno di mezza lunghezza d'onda la cresta di un'onda si sommerà con la valle dell'altra dando somma nulla e producendo interferenza distruttiva (le zone buie). Tutto dipende pertanto dalla differenza del cammino dei due raggi, dalla distanza tra le fenditure, dalla lunghezza d'onda, e ovviamente dall'ampiezza delle due onde. In particolare avremo i massimi ($U_1 \cdot U_2$ positivo) quando

$$F_2P - F_1P = m \frac{\lambda}{2} \quad (5)$$

mentre i minimi ($U_1 \cdot U_2$ negativo) in mezzo. La periodicità delle zone chiare e buie dipende dalla presenza di quell'intero m , che vale 0, 1, 2 ecc. L'intero trae origine a sua volta dalla periodicità intrinseca al fenomeno ondulatorio. Per la luce blu questa differenza è più piccola che per luce verde che a sua volta è più piccola che per la luce rossa. Ecco perché la lunghezza d'onda è un parametro che tiene conto del colore e che diventa sempre più piccolo quanto più ci spostiamo dal rosso (600nm) al blu (400nm) attraverso lo spettro, traslando nei fatti la posizione delle frange. La situazione è del tutto identica a ciò che accade all'interferenza prodotta da delle onde d'acqua in fig. 9b. Eppure c'è di più. La formula scoperta sopra ci sta dicendo che il luogo dei punti P dello spazio che distano da due punti F_1 ed F_2 (che rappresentano due sorgenti puntiformi) è una quantità costante (dipendente solo dal colore della luce).

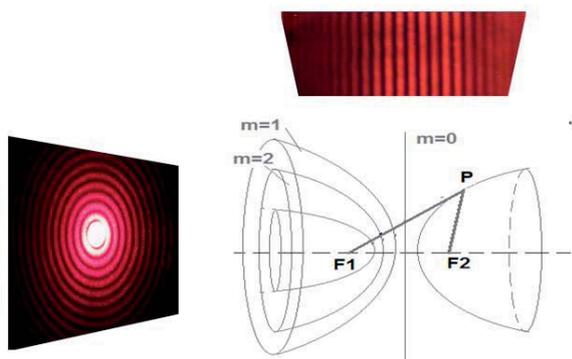


Figura 10: Due sorgenti puntiformi coerenti di luce raggi di luce produrranno degli iperboloidi di rotazione con i fuochi disposti nelle sorgenti. A seconda di dove collochiamo gli schermi ci sarà interferenza circolare o frange pressappoco rettilinee.

Ogni coppia di iperboloidi è data una volta che m sia fissato. Più grande è m e più curvi risultano gli iperboloidi. Per $m=0$ invece essi collassano in un piano. In Fig. 10 ho rappresentato entrambi i due rami di iperboloidi corrispondenti a $m=2$. Ecco allora un bel modo per ricavare la forma dei massimi e dei minimi su un piano qualsiasi. Se prendiamo un piano parallelo ma molto distante dalle due sorgenti esso intersecherà tutti gli iperboloidi meno curvi e i massimi appariranno come delle strisce perpendicolari alla direzione F_1 - F_2 . Supponete adesso di prendere un piano perpendicolare alla retta che passa per i fuochi F_1 ed F_2 . L'intersezione di questo con gli iperboloidi a sinistra della figura darà origine a delle frange circolari. Questa previsione è confermata costruendo interferometri, facendo riflettere uno stesso raggio di luce su uno specchio semiriflettente che produce così un raggio sdoppiato e "simulando" la presenza di due sorgenti. Se invece di una sorgente monocromatica (cioè di un solo colore) inviamo contro le fenditure luce bianca, come quella del Sole, dovremmo trovare i massimi del rosso, del verde e del blue in posizioni leggermente spostate tra loro. Quindi potremmo non vedere più i minimi in quanto nel minimo di un colore cadrebbe il massimo di un altro. Insomma dall'altra parte dello schermo osserveremmo tante frange colorate invece di minimi e massimi. Ed è proprio quel che si osserva provando a osservare il cielo nelle regioni vicino al Sole e chiudendo le palpebre dei vostri occhi di modo da sovrapporre le ciglia, come quando, stando al mare stesi sulla spiaggia, chiudete gli occhi perché il sole vi dà fastidio. Provateci dunque e vedrete attorno alle vostre ciglia tutti i colori dell'arcobaleno. Ebbene le fessure tra le ciglia si stanno comportando come piccole fenditure che vi scompongono la luce solare nei vari colori di cui è fatta. Resta tuttavia una domanda: se la luce è un'onda come si spiega il principio di Fermat? Ebbene l'onda può sempre essere espressa a meno di un fattore di fase che contiene S .

$$U(r, t) = U(r) e^{i\omega t} e^{ikS} \quad (6)$$

Sostituendo questa soluzione nella equazione d'onda vien fuori che l'equazione iconale è solo una approssimazione di una equazione più generale, che tiene conto dell'effetto ondulatorio

$$|\nabla S|^2 = n(\mathbf{r})^2 + \frac{\nabla^2 U(\mathbf{r})}{U(\mathbf{r})} \left(\frac{\lambda}{2\pi} \right)^2 \quad (7)$$

L'addendo aggiuntivo diventa determinante per valori di lunghezza d'onda grandi rispetto alle dimensioni degli oggetti investiti dall'onda. Nel caso della luce visibile dimensioni di alcuni micron sono già sufficientemente piccole per vedere deviazioni della equazione iconale e dunque processi di interferenza o diffrazione. Questa equazione ci fa capire come l'ottica geometrica sia una sottoclasse dell'ottica ondulatoria. Ad esempio se inviamo luce contro un palo la lunghezza d'onda è molto piccola rispetto alle dimensioni del palo e si formerà un'ombra netta in quanto il secondo addendo nella 7 è trascurabile. Ma se le dimensioni diventano piccole come un capello il termine ondulatorio inizia a crescere e vedremo effetti sia di interferenza che di diffrazione. Cosa c'entra questo con un CD? Ebbene la quantità di dati inserita nel CD dipende da dei solchi che vengono scavati comunemente mediante un processo di litografia, che prevede l'uso di una maschera e di luce ultravioletta. Questi solchi saranno netti solo se risulteranno sufficientemente grandi rispetto alla lunghezza d'onda della luce usata di modo che il termine ondulatorio nella equazione iconale non diventi grande. Un CD è molto simile ad un disco in vinile ma mentre il primo viene letto da una puntina il secondo è letto da un Laser che riesce a percepire la presenza o meno di solchi a loro volta generati da altra luce durante il processo litografico. La luce ultravioletta pone pertanto un limite di dimensioni minime ai solchi e dunque alla quantità di dati inserita nel CD. Questo è un problema non da poco nell'industria dell'elettronica integrata. I Chip infatti vengono prodotti usando tecniche fotolitografiche, che però hanno un limite intrinseco dato dal termine ondulatorio dell'equazione (7). La miniaturizzazione, alla base del prodigioso sviluppo dei personal computers, non può più proseguire pertanto seguendo vie convenzionali. Uno dei compiti delle nanotecnologie e della fotonica è proprio quello di trovare vie non convenzionali alla realizzazione di dispositivi che collochino ed elaborino più dati a parità di superficie.

La teoria ondulatoria della luce parrebbe quella definitiva. Eppure manca ancora qualcosa.

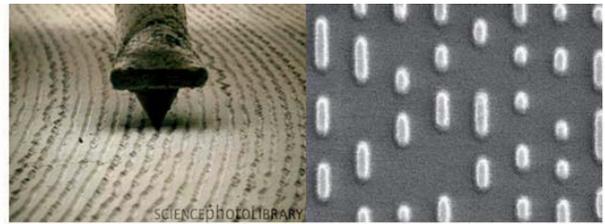


Figura 11: Solchi in un disco di vinile e in un CD-ROM. Le dimensioni dei solchi in un CD sono ai limiti del termine ondulatorio della equazione iconale.

Infatti processi semplici come l'intensità della luce riflessa da un pezzo di vetro non possono essere spiegati da una semplice teoria ondulatoria. E' quello che vedremo ora con la teoria elettromagnetica della luce. E partiremo da una domanda che ci condurrà fino alla relatività: *se la luce è un'onda, allora cosa è che ondeggia?*

Maxwell: La luce è un'onda elettromagnetica

Il XVIII e il XIX secolo sono stati di importanza capitale per la comprensione delle proprietà dell'elettricità e del magnetismo. Due teorie in lotta tra loro, e che troppo spesso vengono raccontate come fossero la stessa teoria nei libri scolastici e universitari, si contendevano lo status di spiegazione ultima di tali fenomeni. Da un lato vi erano i vari Coulomb e Ampère, propositori della *teoria dell'azione a distanza*, una teoria che era stata sviluppata a immagine e somiglianza della teoria della gravitazione universale di Newton. Dall'altra vi erano Faraday e Maxwell, i quali proposero una visione nuova: *la teoria di campo*. La differenza non era solo formale ma sostanziale. I primi vedevano cariche e correnti che interagivano *nel* e *attraverso* il vuoto, concepito come l'assenza di tutto tranne che dell'estensione geometrica. Le forze con cui cariche e correnti interagivano erano alla base dei fenomeni elettrici e magnetici ed erano governate da leggi a distanza come la legge di Coulomb tra cariche elettriche. Faraday e Maxwell, invece, nello spazio tra le cariche o tra le correnti, vedevano un mezzo che chiamarono *etere*, il quale era responsabile della trasmissione della forza elettrica e magnetica. Questo mezzo fu poi battezzato più propriamente *campo elettromagnetico*. In questa teoria una carica Q_A (o una

corrente I_A) non interagisce a distanza e all'istante con una carica Q_B (o una corrente I_B) ma con il campo elettrico (magnetico) prodotta da questa. Le due scuole vedevano gli stessi fenomeni ma li interpretavano in modi opposti. Tuttavia la teoria di campo di Faraday e Maxwell faceva una previsione: una carica o una corrente può agire su un'altra carica o corrente solo dopo un tempo dato dalla velocità della luce. Tali equazioni furono pubblicate per la prima volta da Maxwell nel 1865, ben 150 anni fa, nel *"Treatise on Electricity and Magnetism"*, che è per l'elettromagnetismo ciò che i *"Principia Philosophiae naturalis"* furono per la meccanica e decretarono la fine della teoria dell'azione a distanza e l'inizio di una nuova era. Eccole riportate in fig.12.

And God said:

$$\begin{cases} \nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \\ \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \\ \nabla \cdot \mathbf{B} = 0 \\ \nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \epsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t} \end{cases}$$

And there was light.



Figura 12: Le quattro equazioni di Maxwell. In figura è riportata la frase della genesi "E dio disse sia la luce e la luce fu" a memoria del fatto che le equazioni di Maxwell prevedono l'esistenza della luce come fenomeno elettromagnetico.

La prima equazione ammette l'esistenza di monopoli elettrici, ovvero l'esistenza di cariche elettriche (ρ nella equazione esprime la densità volumetrica di carica elettrica), che sono le responsabili dell'esistenza dei campi elettrici \mathbf{E} . La seconda equazione è la legge di induzione di Faraday. Essa esprime un fenomeno molto interessante e che è alla base dei moderni generatori di corrente: un campo magnetico \mathbf{B} variabile nel tempo è anche esso sorgente di un campo elettrico \mathbf{E} . Questa è una delle equazioni più potenti ed eleganti della storia della fisica perché esprime l'idea secondo cui un campo può generare un altro campo. La terza esprime l'idea che non esistono i monopoli magnetici, ovvero che suddividendo una calamita non si può mai ottenere un singolo polo magnetico, ma si otterranno sempre due calamite, con un polo nord e un polo sud. Infine la quarta equazione, detta di Ampere-Maxwell esprime l'i-

dea secondo cui una densità di corrente \mathbf{J} crea un campo magnetico \mathbf{B} (esperimento di Oersted), ma anche che una variazione temporale di un campo elettrico \mathbf{E} può essere responsabile di un campo magnetico \mathbf{B} . Queste equazioni descrivevano esattamente ciò che le contemporanee teorie di azione a distanza descrivevano nel limite statico, ma in modo molto più semplice, intuitivo e visualizzabile, mediante il concetto intermedio di linee di forza. Ma esse contenevano qualcosa in più. Manipolando infatti le equazioni, Maxwell provò che, in assenza di sorgenti fisiche di cariche elettriche e correnti, esse soddisfacevano una equazione d'onda del tipo

$$\nabla^2 \mathbf{E} = \epsilon \mu \frac{\partial^2 \mathbf{E}}{\partial t^2}; \nabla^2 \mathbf{B} = \epsilon \mu \frac{\partial^2 \mathbf{B}}{\partial t^2} \quad (8)$$

Dove ϵ e μ sono la permittività e suscettibilità elettrica e magnetica del mezzo, incluso il vuoto. Ovviamente queste erano equazioni d'onda a patto di sostenere che la velocità dell'onda fosse, in accordo con l'equazione (3)

$$v = c = \frac{1}{\sqrt{\epsilon \mu}} \quad (9)$$

Misurando le costanti di cui sopra con semplici esperimenti di elettrostatica e magnetostatica, si misurò che l'onda in questione viaggia nel vuoto alla velocità di 300000km/sec, la velocità della luce! Questa incredibile scoperta teorica permetteva di prevedere l'esistenza di onde di elettricità e magnetismo, dette onde elettromagnetiche, che viaggiavano proprio alla velocità della luce! Ce n'era abbastanza per iniziare a sospettare che l'onda di luce di Huygens e Young fosse un'onda di elettricità e magnetismo. Come provarlo? C'è un aspetto da non trascurare. Queste onde avevano infatti due caratteristiche molto importanti: anzitutto erano onde trasversali, i campi elettrico e magnetico oscillavano cioè perpendicolarmente alla direzione di propagazione, al contrario del suono, che, nell'aria, è un'onda longitudinale. Secondo, non esisteva un'onda elettrica senza una magnetica. Campo elettrico e magnetico infatti oscillavano sempre insieme, in fase, ed erano per giunta perpendicolari l'uno all'altro. In figura 13 si riporta come si presentano le onde elettromagnetiche piane.

Come provare sperimentalmente l'esistenza delle onde elettromagnetiche e soprattutto come

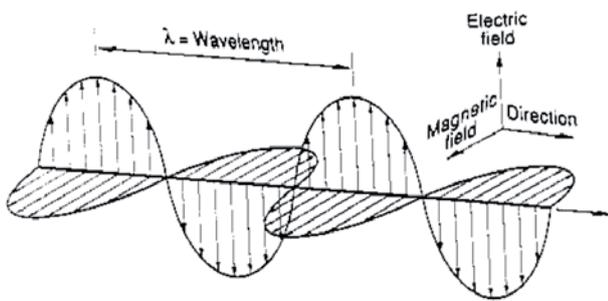


Figura 13: Rappresentazione di un'onda elettromagnetica piana.

provare in via definitiva che la luce è un'onda elettromagnetica? Fu Heinrich Hertz a condurre il primo esperimento risolutivo nel 1886.

Le prove: telecomunicazioni e tramonti

Hertz fece scorrere in una sorta di antenna (un oscillatore) una corrente alternata con una certa frequenza (fig. 14), in modo da produrre, stando alle leggi di Maxwell, un'onda elettromagnetica con una polarizzazione del campo elettrico (la direzione di oscillazione) parallela all'antenna. Dall'altra parte, a parecchi metri di distanza dall'oscillatore, dispose uno specchio metallico che consentisse alle onde elettromagnetiche, stando alle previsioni delle equazioni suddette, di essere riflesse indietro e generare onde stazionarie. E' un po' come la corda di chitarra che è tenuta ferma ai due nodi estremi e vien fatta oscillare. Tuttavia Hertz sapeva benissimo che, data la distanza tra antenna trasmittente e specchio, se si fossero formate onde stazionarie vi sarebbero stati punti dello spazio tra l'oscillatore e lo specchio in cui l'onda non oscillava affatto (i cosiddetti nodi), e altri (detti antinodi) in cui le oscillazioni del campo elettrico erano massime. Notate che nei nodi anche il campo magnetico **B** (perpendicolare a **E**) non oscilla affatto ed è nullo, mentre negli antinodi è massimo, a causa del fatto che **E** e **B** oscillano, come detto, in fase. Le onde non sono visibili, l'unica cosa che vediamo sono l'oscillatore, ovvero l'antenna, e uno specchio. Come provare l'esistenza di cose invisibili? Attraverso l'effetto che producono. E Hertz fu un genio in questo. Conosceva la frequenza ν di oscillazione delle cariche (in quanto

aveva impostato lui tale frequenza), e sapeva che la frequenza delle onde prodotte, stando alle leggi di Maxwell, avrebbe dovuto essere la stessa, per cui i nodi avrebbero dovuto trovarsi ad una distanza dallo specchio di un multiplo di mezza lunghezza d'onda (ricordo che la lunghezza d'onda è pari alla velocità della luce diviso la frequenza dell'onda, $\lambda=c/\nu$), ovvero nei punti distanti $\lambda/2$, λ , $3\lambda/2$ dallo specchio. Gli antinodi avrebbero invece dovuto trovarsi tra questi. Ad esempio per una frequenza di 37,5 MHz (la frequenza usata da Hertz) la lunghezza d'onda è pari a 8 metri. Quindi il primo antinodo era a 2 metri dallo specchio, a 4 metri vi era un nodo, a 6 metri vi era un altro antinodo e così via. Disponendo un anello circolare conduttore nei nodi, non ci sarebbe stata alcuna oscillazione del campo magnetico. Dunque, per la legge di Faraday (un campo magnetico variabile crea una corrente in un conduttore chiuso), non potrà esserci nessuna corrente indotta nell'anello. Disponendo invece tale anello nei massimi, ovvero negli antinodi, ci sarebbe stato un campo magnetico oscillante (cioè variabile nel tempo) intenso. A patto di disporre l'anello perpendicolarmente a **B**, e quindi nel piano in cui oscilla **E**, si sarebbe vista una corrente indotta. Ma Hertz invece di usare un anello chiuso ne usò uno con un piccolo taglio. La corrente indotta avrebbe così generato un accumulo di cariche alle estremità tali da provocare una scintilla. Immaginate l'effetto visivo: disponete tanti anelli recanti un piccolo taglio come in figura 14, in corrispondenza di quei valori che vi ho detto e perpendicolarmente alla direzione del supposto campo magnetico. Inviare nell'oscillatore una corrente oscillante per produrre le famigerate e invisibili onde e vedrete un anello che non produce nulla, uno che produce una vistosa scarica, quello successivo nulla, quello adiacente a questo ancora una scintilla e via dicendo! E' la conferma che le onde elettromagnetiche della teoria di Maxwell esistono per davvero!

Va precisato però che Hertz non provò affatto che la luce fosse un'onda elettromagnetica, ma al contrario che le onde elettromagnetiche esistevano e che, indipendentemente dalla loro frequenza, si comportavano seguendo le leggi di un raggio di luce ordinario. Provò infatti che le onde da lui prodotte viaggiavano alla velocità

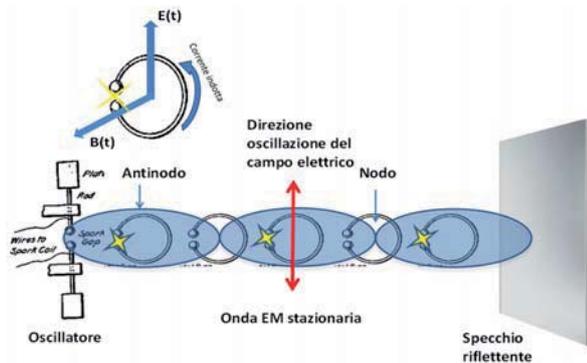


Figura 14: *Esperimento di Hertz sulle onde elettromagnetiche. L'oscillatore produce un'onda di pari frequenza, la quale, riflettendosi su di uno specchio riflettente di onde lunghe, genera onde stazionarie presentanti dei nodi localizzati a distanze di mezza lunghezza d'onda. Collocando degli anelli metallici nei nodi non si osserva nulla. Ma traslandoli negli antinodi si osserva corrente indotta (o una scarica nel caso in cui gli anelli rechino una fenditura). Questo avviene purché gli anelli giacciono nel piano di oscillazione del campo elettrico, di modo che il campo magnetico, perpendicolare ad esso, produca corrente indotta a causa della legge di Faraday (figura in alto).*

della luce (come la luce), che si riflettevano sui metalli (come la luce), che si rifrangevano attraverso mezzi più densi seguendo la legge di Snell (come la luce) e che era possibile polarizzarle (come la luce). Del resto avrebbe potuto essere una coincidenza. In fondo anche le onde gravitazionali (appena scoperte) viaggiano alla velocità della luce e sono onde trasversali (polarizzate) ma non ci sogneremmo mai di dire che la luce è un'onda gravitazionale! La dimostrazione che la luce è un'onda trasversale di elettricità e magnetismo vien fuori analizzando come cambiano i campi nel passaggio da un mezzo ad un altro, come aria e acqua o vetro. Usando le equazioni di Maxwell e riconoscendo che l'indice di rifrazione diventa nella teoria di Maxwell la radice quadrata della costante dielettrica relativa del mezzo ($n = \sqrt{\epsilon_r}$), è possibile calcolare il rapporto r tra il modulo quadro del campo elettrico riflesso e il quello del campo incidente nonché il rapporto t tra il modulo quadro del campo elettrico trasmesso (o rifratto) e quello incidente, tra due mezzi diversamente densi. Si può fare questo conto (ed è in effetti l'aspetto più interessante) anche in funzione dell'angolo di incidenza del fascio.

Quelle che si ottengono sono le cosiddette leggi di Fresnel per le onde elettromagnetiche. Queste equazioni in particolare tengono conto anche della polarizzazione del campo elettrico.

In Fig. 15 un'onda elettromagnetica colpisce la superficie di separazione tra aria e un dielettrico di indice di rifrazione n producendo un raggio riflesso e uno rifratto. Le due componenti polarizzate perpendicolarmente e parallelamente al piano di incidenza saranno riflesse e trasmesse in modi diversi a causa delle leggi di Maxwell calcolate tra due mezzi con diversa costante dielettrica (*boundary conditions*). L'intensità dell'onda riflessa e trasmessa dipenderà dall'angolo di incidenza. In fig. 15 b potete osservare come variano i coefficienti di riflessione al variare dell'angolo di incidenza per i due campi con oscillazione parallela al piano di incidenza (p) e perpendicolare a questa (s). Come vedete per piccoli angoli di incidenza (minori di 30°) la riflettività resta pressoché costante. Se il mezzo su cui l'onda si riflette è il vetro tale percentuale è pari a circa il 4% per svariate lunghezze d'onda. Dopo di che quella dovuta al campo parallelo al piano di incidenza (p), indicato con le frecce in figura 15a, diminuisce fino ad annullarsi per poi risalire, mentre quella dovuta al campo perpendicolare, indicato con i puntini in Fig. 15a, aumenta sempre con l'angolo di incidenza. L'angolo a cui il campo parallelo riflesso si annulla si chiama angolo di Brewster e si ha quando

$$\operatorname{tg}\theta_B = n \quad (10)$$

per una riflessione dall'aria al mezzo dielettrico, e

$$\operatorname{tg}\theta_B = 1/n \quad (11)$$

per una riflessione dal mezzo dielettrico all'aria. Bisogna fare esperimenti sofisticati per vedere all'opera queste leggi? No. Infatti sono le stesse leggi che la luce segue. E questo dimostra in modo incontrovertibile che la luce è un'onda elettromagnetica. Osservate la luce riflessa sul pavimento di casa da una finestra (Fig. 16a) e vi renderete conto che, avvicinandovi alla finestra, la luce diminuirà repentinamente di intensità in quanto l'angolo di incidenza sul pavimento diminuirà dello stesso grado. In particolare accadrà che la luce apparirà polarizzata all'ango-

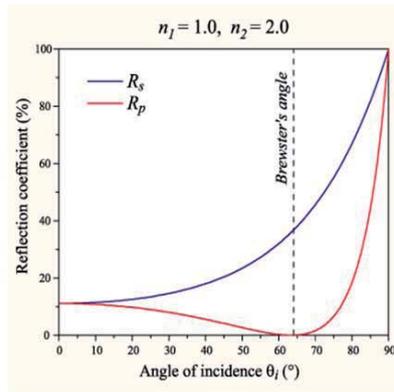
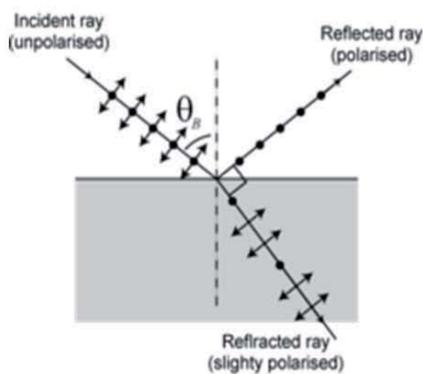


Figura 15: a) A sinistra, raggi incidente, riflesso e trasmesso (o rifratto) su uno strato di materiale dielettrico. All'angolo di Brewster (θ_B) delle due polarizzazioni perpendicolare (puntini) e parallelo (freccie) al piano di incidenza individuato dal piano di giacenza dei raggi subiscono delle modifiche sostanziali. Infatti la luce riflessa sarà polarizzata solo perpendicolarmente al piano di incidenza. b) A destra, coefficienti di Fresnel in funzione dell'angolo di incidenza. R_p indica l'intensità relativa della luce riflessa con polarizzazione parallela al piano di incidenza. Con R_s è indicata l'intensità relativa della luce riflessa e polarizzata perpendicolarmente.

lo di Brewster, come potrete verificare con un semplice filtro polaroid.

Un secondo fenomeno è la riflessione del sole sulla superficie del mare al tramonto (Fig. 16b). La luce sembra aumentare di intensità man mano che il riflesso si avvicina all'orizzonte. Questo accade per due motivi: affinché i raggi paralleli provenienti dal sole giungano nell'occhio dell'osservatore gli angoli di incidenza aumenteranno man mano che si volge lo sguardo verso il Sole, e le leggi di Fresnel dicono che all'aumentare di tale angolo aumenta la luce riflessa totale. Il secondo motivo, concomitante al primo, è che ci sono delle increspature periodiche sulla superficie del mare (le onde) che appariranno più dense verso l'orizzonte e più distanti vicino a noi per un effetto prospettico. Curioso che la maggior parte delle risposte al perché l'intensità luminosa sembra aumentare volgendo lo sguardo verso l'orizzonte siano quasi tutte del tipo "perché ci avviciniamo alla sorgente, il sole" senza specificarne il senso. Invece dipende dalla prospettiva e dalle leggi di Fresnel, dirette conseguenze delle equazioni di Maxwell. Anche un arcobaleno apparirà polarizzato perché l'angolo a cui la luce è dispersa e riflessa dalle gocce si aggira intorno ai 42° rispetto al centro dell'arco. L'angolo di Brewster per la riflessione interna all'acqua è a circa 37° , come potete calcolare dalla equazione (11), un valore molto vicino al precedente, per cui ci aspettiamo che la luce dell'arcobaleno sia sufficientemente polarizzata

nella direzione ad esso tangente. L'effetto è visibile in Fig. 16c e d. La prossima volta procuratevi un filtro polaroid o un paio di occhiali polarizzati pronti all'uso. Tutte queste sono prove che la luce è un'onda elettromagnetica.

Parte II: Olismo classico.

Etere o non etere, questo è il problema.

Poiché ogni onda si propaga in un mezzo come il suono nell'aria, la luce, per Maxwell, non poteva fare eccezioni. Maxwell ipotizzò che lo spazio vuoto non fosse realmente tale ma che fosse riempito di un mezzo battezzato *etere luminifero*. Nelle equazioni di Maxwell compariva la velocità della luce c , che dipendeva dalle costanti dielettriche e magnetiche del vuoto. In effetti questo era una sorta di indizio che il vuoto fosse in realtà pieno di qualcosa, come una specie di dielettrico. La motivazione per introdurre l'etere come supporto della propagazione delle onde elettromagnetiche derivava non solo da una mentalità "meccanicistica", non tanto newtoniana quanto cartesiana, ma anche dal fatto che nelle equazioni di Maxwell compariva la velocità



Figura 16: a) Alto sinistra, riflessione dal pavimento della luce proveniente da una finestra. b) Alto destra, tramonto del sole sulla superficie del mare. La riduzione della intensità della luce riflessa man mano che la direzione in cui si osserva si abbassa deriva dalle leggi di Fresnel, secondo cui riducendo l'angolo di incidenza (e dunque di riflessione) l'intensità diminuisce assestandosi attorno ad un valore costante. c) basso sinistra e d) basso destra: polarizzazione della luce riflessa da un arcobaleno. A causa dell'angolo di Brewster la luce riflessa in una goccia viene polarizzata perpendicolarmente all'angolo di incidenza. Tale angolo è molto vicino a quello di riflessione critica della luce, per cui l'arcobaleno apparirà polarizzato lungo la direzione ad esso tangente in ogni punto, come è possibile osservare mediante un filtro polarizzatore orientato parallelamente (c) e perpendicolarmente (d) a tale direzione.

della luce c , come fosse un parametro assoluto e non relativo al sistema di riferimento. Ne derivava che le equazioni di Maxwell erano valide solo per sistemi immobili rispetto all'etere. Ma qui si poneva la questione: muovendoci rispetto all'etere con una velocità v le equazioni di Maxwell andavano modificate in modo da sostituire a c una combinazione delle due? Tutto dipendeva se l'etere veniva o meno trascinato dal moto del riferimento. Facciamo un esempio relativo al suono. Se mi trovo nel vagone di un treno in moto ad una velocità pari a v_T l'aria viene completamente trascinata al suo interno, per cui un suono si propagherà nella cabina alla velocità v_s solita del suono, e non mi accorgerei di essere in moto. Se andassi invece sopra il vagone del

treno in corsa e stessi in piedi avvertirei un vento sul mio viso in quanto l'aria a qualche metro di altezza dal vagone non sarebbe affatto trascinata, sicché un suono sarebbe percepito prima o dopo a seconda della direzione del treno. Posso in questo caso stabilire che il treno è in corsa? La risposta è no in quanto esiste anche il vento vero e proprio che trascina più o meno efficacemente un suono, per cui non potrei stabilire se è il treno a creare l'effetto del vento o viceversa vi è un vento a prescindere dal moto del treno. Tuttavia se mi sedessi sul vagone il vento che percepirei sarebbe meno intenso, in quanto l'aria verrebbe parzialmente trascinata a causa dell'attrito con il tetto del vagone. Ora sostituiamo al treno in corsa la Terra, all'aria l'etere che ora permea il

vuoto attorno alla terra e al suono la luce. In perfetta analogia con il treno ci sono tre possibili interazioni tra il nostro pianeta e l'etere:

- 1) La Terra attraversa l'etere senza trascinarlo (come l'aria ben al di sopra del vagone).
- 2) La Terra trascina parzialmente l'etere con sé (come l'aria in prossimità del tetto del vagone).
- 3) La Terra trascina con sé l'etere completamente (come nel caso dell'aria all'interno del vagone).

Queste le tre ipotesi possibili ma contrapposte e contraddittorie. Come discernere? Attraverso l'esperimento, direbbe qualcuno, consapevole che l'unica guida per la verità sia l'esperienza e l'osservazione. Ma proprio qui capiamo come i dati sperimentali nulla ci dicono circa la natura se non attraverso un modello mentale già preconstituito attraverso cui interpretare i fenomeni osservati. Infatti ognuna delle tre ipotesi precedenti riesce a spiegare un preciso fenomeno.

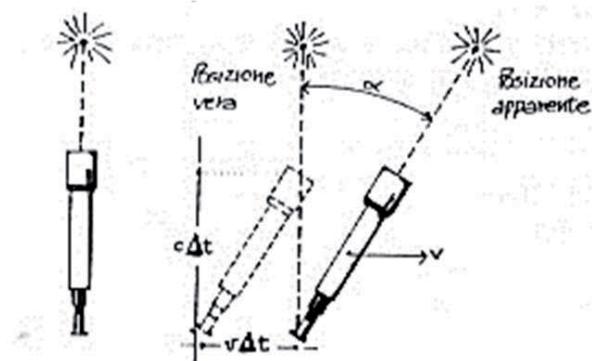


Figura 17: Spiegazione dell'aberrazione stellare mediante la composizione delle velocità di Galileo. A sinistra un cannocchiale punta nella direzione della stella in un riferimento immobile rispetto ad essa. Ne caso a destra il sistema di riferimento dell'osservatore è in moto con velocità v . Per puntare la stella il cannocchiale dovrà essere inclinato di un angolo dato dall'arcotangente del rapporto tra v e c . Infatti nello stesso intervallo temporale Δt la terra si sarà spostata di $v\Delta t$ mentre la luce da una stella si sarà propagata per un tratto $c\Delta t$. Alla fine sembrerà che la stella abbia percorso un tragitto inclinato.

Esperimento 1

L'ipotesi 1 riesce a spiegare il fenomeno della aberrazione stellare (Fig. 17). Brevemente, supponete di osservare una stella dritto sopra la vostra testa. Ebbene la stella non è "realmente" lì.

Potremmo affermarlo se la luce si muovesse con velocità infinita, ma poiché la velocità della luce è di "soli" 300.000km/s e la Terra si muove nell'etere immobile "a ben" 30km/s, noi vedremo la stella in una posizione differente data dalla combinazione delle velocità. L'angolo di aberrazione (deviazione) sarà piccolo ma misurabile con tecniche di precisione.

$$\operatorname{tg}\alpha = \frac{v}{c} = 0.0001 \quad (12)$$

Il risultato di questo fenomeno torna con i valori previsti nel caso in cui l'etere sia immobile.

Esperimento 2

L'ipotesi del trascinamento parziale dell'etere da parte della Terra conduce a spiegare un fenomeno alquanto bizzarro. Supponete di inviare un raggio di luce attraverso un tubo cilindrico pieno d'acqua in movimento come in figura 18. L'acqua potrebbe trascinare l'etere e l'entità di tale trascinamento essere misurata. In tal caso la velocità della luce osservata in laboratorio sarebbe

$$w = \frac{c}{n} \pm fv \quad (13)$$

a seconda che l'etere abbia la stessa direzione (+; tubo inferiore) o direzione opposta (-; tubo superiore) al raggio di luce. Se la frazione è $f=0$, allora l'etere resta immobile e la velocità della luce è solamente c/n , dove n è il solito indice di rifrazione. Se $f=1$, l'etere ha una velocità pari proprio a v , la velocità dell'acqua, per cui la velocità finale sarà $c/n+v$. Ma se l'etere venisse trascinato parzialmente, e dunque con una velocità inferiore a v ma non nulla, ecco che dovremmo avere la formula di prima, intermedia tra i due casi estremi. Possiamo determinare f , e dunque capire il risultato? Prendete in considerazione l'apparato riportato in Fig.18, in cui viene mostrato un interferometro modificato di tipo particolare.

Una sorgente S emette due raggi di luce nello stesso istante ed essi si incontrano nel detector O grazie alla riflessione prodotta da due specchi $M1$ ed $M2$, dopo aver percorso due tubi in cui viene fatta scorrere acqua con velocità v ma in direzioni opposte, come in figura. Il raggio superiore sarà rallentato dal flusso opposto di acqua, mentre quello inferiore si muoverà più velocemente, entrambi di una frazione fv rispet-

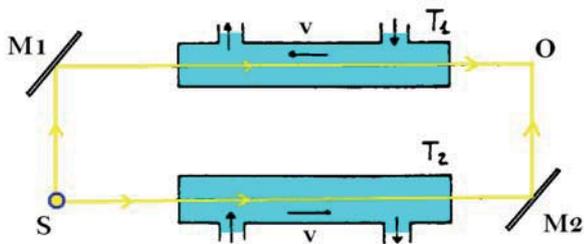


Figura 18: Esperimento di Fizeau sul trascinamento dell'etere tramite flusso d'acqua. La sorgente S emette due raggi di luce che si incontrano in O dopo essere stati riflessi da due specchi M_1 ed M_2 ed aver percorso dei contenitori in cui scorre acqua in direzione opposta.

to all'aria. Poiché si ricongiungeranno in O con un certo sfasamento produrranno delle frange di interferenza la cui posizione dipenderà proprio dal ritardo. Se variamo quest'ultimo, cambiando la velocità del flusso d'acqua, vedremo variare la posizione delle frange e dedurre quanto vale f . Se ad esempio non dovessimo vedere alcun cambiamento nelle frange, pur variando v , è ovvio che f dovrà essere zero. Louis Fizeau, che condusse un simile esperimento, giunse alla conclusione che quel fattore f fosse pari a

$$f = 1 - \frac{1}{n^2} \quad (14)$$

dove n è il solito indice di rifrazione del mezzo. Nel caso dell'acqua ad esempio, che ha $n=1.33$, f era uguale a 0.43, un valore quindi a metà strada tra zero e uno. Se ne deducono due conclusioni:

1. L'etere viene trascinato del 43% dal flusso di acqua.
2. L'entità del trascinamento dipende dal mezzo.

Esperimento 3

Qui si applica un interferometro come nel caso di Fizeau ma non al moto di un flusso d'acqua ma a quello della Terra attorno al Sole. Supponiamo di inviare un raggio di luce da una sorgente ad uno specchio fisso M_2 nella stessa direzione del moto della Terra. Al centro tra la sorgente ed M_2 vi è uno specchio semiriflettente M_s che divide il fascio in due tragitti perpendicolari tra loro. Uno prosegue nella direzione del moto della Terra percorrendo un tragitto di andata e ritorno pari

a L_2 e l'altro, perpendicolare a questo, prosegue verso un secondo specchio M_1 , distante L_1 dal divisore di fascio, e che verrà da questo riflesso indietro.

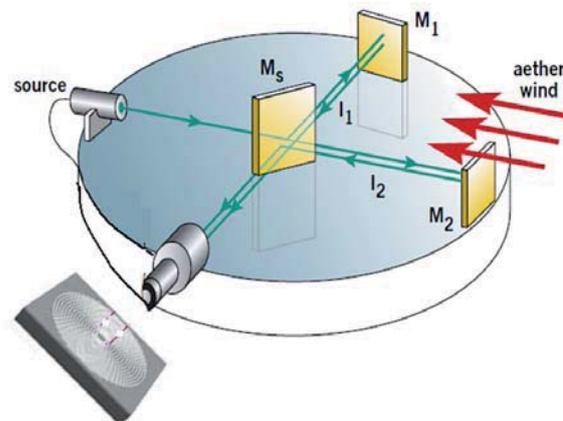


Figura 19: Esperimento di Michelson e Morley per determinare la velocità della Terra rispetto all'etere. Una sorgente colpisce uno specchio semiriflettente che divide il fascio in due raggi perpendicolari che proseguono uno verso uno specchio riflettente M_1 e l'altro verso M_2 per poi far ritorno e interferire.

Sia kv la velocità relativa della Terra rispetto all'etere dove k tiene conto del trascinamento dell'etere. Se $k=0$ la velocità relativa è nulla e l'etere viene completamente trascinato dal moto terrestre attorno al Sole, se $k=1$ la velocità relativa è proprio pari a quella della Terra, v , attorno al Sole, per cui l'etere non è affatto trascinato. Quanto tempo impiegherà ciascun raggio per effettuare il moto di andata e ritorno lungo i due tragitti? Introducendo la quantità $\gamma = 1/\sqrt{1 - \frac{(kv)^2}{c^2}}$

nel caso del tragitto $M_s \rightarrow M_2 \rightarrow M_s$ avremo che

$$T_2 = \frac{L_2}{c - kv} + \frac{L_2}{c + kv} = \frac{2L_2}{c} \gamma^2 \quad (15)$$

Notate che il tempo sarà maggiore del tempo impiegato per percorrere lo stesso tragitto, se l'apparato fosse in quiete rispetto all'etere, di una quantità pari a γ^2 . E' ovvio che quella frazione è molto vicina a 1 essendo $(fv)^2/c^2$, una quantità pari o inferiore a 0.00000001. Così piccola che Maxwell era convinto che non si potesse misurare. Ma non aveva fatto i conti con gli interferometri, in particolare quello di Michelson e Morley riportato in figura 19. Calcoliamo ora il

tempo relativo al tragitto $M_S \rightarrow M_1 \rightarrow M_S$. Un semplice calcolo prova che tale tempo, è pari a

$$T_1 = \frac{2L_1}{c} \gamma \quad (16)$$

quindi non più proporzionale a γ^2 ma solo a γ . Tenendo conto che i percorsi hanno la stessa lunghezza ($L_1=L_2=D$) e dell'approssimazione per cui il termine in γ pari a $(kv/c)^2$ è dell'ordine di 10^{-8} allora le lunghezze reali dei percorsi H_1 e H_2 dipenderanno dal moto relativo in prima approssimazione come

$$H_1 = 2D\gamma \cong 2D \left(1 + \frac{k^2 v^2}{2c^2} \right) \quad (17)$$

$$H_2 = 2D\gamma^2 \cong 2D \left(1 + \frac{k^2 v^2}{c^2} \right) \quad (18)$$

Sicché la differenza di cammino ottico reale è pari a

$$\Delta = H_2 - H_1 \cong D \left(\frac{kv}{c} \right)^2 \quad (19)$$

Se questa differenza è pari ad un multiplo intero di mezza lunghezza d'onda si avrà interferenza distruttiva nel detector, altrimenti se è pari a un multiplo intero di lunghezza d'onda l'interferenza sarà costruttiva. Grazie a riflessioni multiple Michelson e Morley resero l'esperimento più preciso, facendo in modo che D diventasse pari a circa ben 10 metri. Pertanto Δ sarà dell'ordine di 10^{-7} m, cioè 100 nm. Usando luce gialla ($\lambda = 500$ nm) Δ è dell'ordine di $0,2 \lambda$, una quantità misurabile. Ovviamente se orientiamo l'interferometro, ad esempio ruotandolo di 45° rispetto alla situazione appena discussa, ci aspettiamo che Δ dipenderà da tale angolo α e oscillerà tra 0 e il valore massimo di $0,2\lambda$. In figura 20 sono riportati i risultati delle misure eseguite dai due scienziati.

I valori misurati in questo esperimento sono indicati dai tratti continui in Fig. 20 mentre quelli punteggiati indicano l'andamento previsto. Faccio notare che le scale lungo le ordinate si riferiscono ai valori misurati e che le curve teoriche sono state compresse perché fuori scala di circa un fattore 10. Leggiamo le parole dei due scienziati: "lo spostamento reale era certamente inferiore alla ventesima parte di tale spostamento previsto, e probabilmente inferiore alla sua

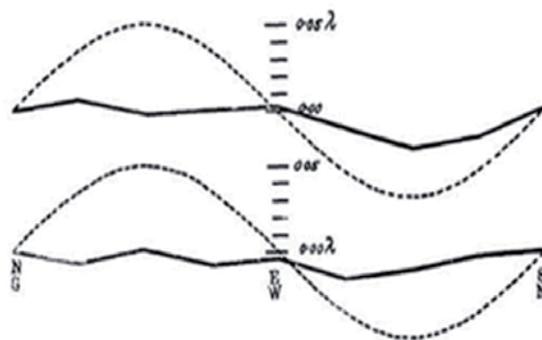


Figura 20: Risultati dell'esperimento di Michelson e Morley (linea continua) rispetto alle previsioni fatte (linea tratteggiata). Lungo l'asse delle ordinate compare lo sfasamento previsto delle frange in unità di lunghezza d'onda, lungo l'asse delle ascisse vi è la direzione di orientamento dell'interferometro rispetto ai punti cardinali.

quarantesima parte". Insomma se un effetto c'era era enormemente inferiore a quello previsto. L'unica spiegazione possibile era che il fattore k fosse nullo. Morale: se la Terra si muove attorno al Sole ma la sua velocità è nulla rispetto all'etere, allora quest'ultimo partecipa completamente del moto della Terra e ne viene completamente trascinato.

Siamo allora arrivati alla seguente situazione :

1. La spiegazione dell'aberrazione implica che l'etere *non partecipa* al moto dei corpi.
2. L'effetto di Fizeau trova la spiegazione in un etere che *partecipa parzialmente* al moto.
3. La spiegazione dell'esperimento di Michelson e Morley infine è che l'etere *partecipa completamente* al moto dei corpi.

Insomma i tre punti sono in contraddizione reciproca. Che diavolo sta succedendo? Qui la cosa si fa interessante e ci mostra come un dato esperimento non solo non può essere risolutivo per dirimere le questioni scientifiche, ma addirittura è l'indizio di come la teoria trova un limite quando, se applicata ai dati asettici di un esperimento, arriva ad una contraddizione interna. Come già detto, i dati non dicono nulla del mondo se non sono già filtrati mediante uno schema teorico. Ma come risolvere l'enigma? Una via di uscita c'è, ma occorre pagare un prezzo enorme: la simultaneità temporale di due eventi non è un

fatto associato per tutti gli osservatori. Entra in scena Albert Einstein.

L'illusione della simultaneità

Il 1905 è definito l'*annus mirabilis* di Einstein. In quell'anno quest'uomo straordinario pubblicò, lavorando all'ufficio brevetti di Berna, quattro lavori di cui due fondarono la genesi rispettivamente della teoria della relatività speciale e della meccanica quantistica. L'aspetto interessante dei due lavori è che emerge uno scienziato assolutamente privo di briglie. Infatti mentre il primo articolo parte dall'ipotesi che le leggi di Maxwell siano corrette, il secondo parte dall'assunto che tali leggi sono sbagliate. Nonostante ciò Einstein fonda due visioni del mondo assolutamente nuove e internamente coerenti, sebbene tuttora in contraddizione tra loro almeno negli assunti epistemologici. Nel primo lavoro Einstein parte da due postulati:

- P1) Nessun esperimento né di meccanica e né di elettromagnetismo consentiranno di capire se ci troviamo in un sistema di riferimento fermo o in moto rettilineo uniforme
- P2) La velocità della luce, c , è indipendente dalla velocità dell'osservatore o della sorgente

Perché arrivò a tali conclusioni? Anzitutto Einstein rilevò una grave contraddizione tra la teoria di Maxwell e le leggi del moto di Galilei-Newton. Nella quarta legge di Maxwell compare la velocità della luce c (sotto forma del prodotto tra le costanti dielettrica e magnetica). Invece nella seconda legge di Newton compare l'accelerazione a . Ora mentre la relatività di Galileo è compatibile con la legge di Newton, non lo è con quelle di Maxwell. Infatti l'accelerazione non dipende dall'osservatore ma la velocità sì. In quanto le velocità si combinano sommandosi tra loro. Se un oggetto si muove rispetto ad un osservatore con velocità w , allora un altro osservatore che si muove rispetto al primo con velocità costante v vedrà il moto dell'oggetto avvenire con una velocità data dalla legge di composizione di Galileo

$$w' = w \pm v \quad (20)$$

Se deriviamo entrambi i membri rispetto al tempo otteniamo che le accelerazioni a' e a del

corpo viste dai due osservatori saranno uguali, in quanto v è costante. Pertanto nella seconda legge di Newton $F=ma$ compaiono la massa e l'accelerazione, entrambe quantità invarianti. Invece l'elettromagnetismo sembra che prediliga una velocità nelle sue equazioni: c . Ma allora questo vuol dire una e una cosa sola: al contrario della legge di Newton le leggi di Maxwell non sono invarianti, esse cambiano da un riferimento all'altro in moto relativo rispetto al primo. Questa idea è compatibile col trascinarsi dell'etere. Eppure abbiamo già visto a quali contraddizioni l'idea di etere porta. Ad Einstein però non sfuggì un fatto. Se abbiamo un magnete e lo muoviamo rispetto ad un anello conduttore ecco che vediamo una corrente indotta, senza dover applicare all'anello alcuna batteria. E' l'esperimento di Faraday, ben descritto dalle leggi di Maxwell. Cosa accade se però, al contrario, muoviamo l'anello rispetto al magnete? Continueremo a vedere corrente, proprio come prima. Come spiega questo fatto la teoria di Maxwell-Lorentz? Con la forza di Lorentz. Il conduttore si sta muovendo nel campo magnetico fisso con velocità v , ma il conduttore è fatto di elettroni che si stanno muovendo con la velocità v del conduttore. Allora essi saranno sottoposti ad una forza di Lorentz

$$\mathbf{F} = q\mathbf{v} \times \mathbf{B} \quad (21)$$

e verranno messi in moto nel conduttore producendo una corrente. Perciò se da un lato è vero che si fa ricorso a due leggi diverse: quella di Faraday nel caso in cui sia il magnete a muoversi, e quella di Lorentz nel caso in cui sia il conduttore a muoversi, è pur vero che tali leggi cospirano in modo da non farci capire chi si muove. Copernico e Galileo possono dormire sogni tranquilli. Di fatto ne deriva che i fenomeni elettromagnetici sono indipendenti dal moto dell'osservatore proprio come il primo postulato di Einstein affermava. Ma allora come mai compare la velocità c e non una accelerazione nelle equazioni di Maxwell? Ecco la contraddizione nelle due teorie. Bene, Einstein ipotizzò arbitrariamente nel suo secondo postulato che quella quantità c non può cambiare da un riferimento all'altro. E' l'unico modo per rendere invarianti le leggi di Maxwell. L'aspetto più importante è che tutti e tre gli esperimenti visti prima vengo-

no spiegati in un colpo solo. Vediamo il perché. L'esperimento di Michelson e Morley non mostra nulla in quanto non c'è nulla da vedere. Se non possiamo infatti capire con esperimenti di nessun tipo se la Terra è in moto (postulato P1) non possiamo pretendere di trovare una differenza nei cammini ottici tra i due raggi di luce, esattamente come non lo avremmo trovato se la Terra fosse ferma. La luce si deve muovere alla stessa velocità sia lungo il percorso parallelo al moto terrestre che perpendicolarmente ad esso. Ma se le cose stanno così allora la legge di composizione delle velocità deve cambiare. Einstein provò dai postulati P1 e P2 che le leggi di composizione dei moti divenivano le seguenti

$$w' = \frac{w \pm v}{1 + \frac{vw}{c^2}} \quad (22)$$

Per piccole velocità rispetto a c la 22 ridiventa uguale alla 20, ridandoci la legge di Galileo, che pertanto è un limite per basse velocità della 22. Da questa relazione si evince inoltre che il limite massimo per la velocità di un segnale è proprio c . Come mai la legge di composizione delle velocità viene modificata da quei due postulati? Il punto essenziale è che alla base della legge di composizione galileiana delle velocità vi è il fatto che il tempo è una quantità indipendente dall'osservatore. Invece nella relatività speciale ciò che è indipendente dall'osservatore è proprio c . In questo caso la costanza della velocità della luce nel vuoto impone che il tempo (e anche lo spazio) sia relativo al moto dell'osservatore. In un certo senso nella fisica di Galileo le velocità si adattano per far sì che spazio e tempo restino invarianti, mentre nella fisica di Einstein spazio e tempo si adattano affinché c resti invariante. Vediamo meglio tale aspetto. Supponiamo di prendere un orologio ottico (Fig. 21), costituito da due specchi tra i quali rimbalza un raggio di luce. La distanza percorsa in mezzo ciclo è pari a $c\Delta t/2$. Un osservatore che si muove con velocità v vedrà invece la luce muoversi lungo la diagonale percorrendo un tratto che, a causa della costanza della velocità della luce, sarà pari a $c\Delta t'/2$. Ovviamente il tempo non potrà essere lo stesso perché la velocità della luce è la stessa ma il percorso differente.

Usando il teorema di Pitagora al triangolo rettangolo ABH abbiamo che

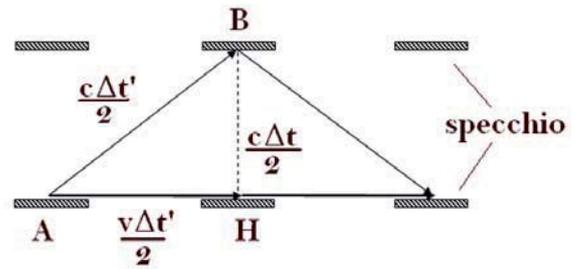


Figura 21: Orologio ottico di Einstein. A causa della costanza della velocità della luce il tempo si adatta dilatandosi.

$$(c\Delta t')^2 = (v\Delta t')^2 + (c\Delta t)^2 \quad (23)$$

Da cui ricaviamo che il battito di tempo visto dall'osservatore in moto sarà

$$\Delta t' = \gamma \Delta t \quad (24)$$

Questa uguaglianza ci dice appunto che il tempo si adatta affinché c resti invariata. Più in dettaglio, il tempo scorre più lentamente se un osservatore è in moto, e questo rallentamento sarà tanto più evidente quanto più la velocità v è grande rispetto alla velocità della luce! L'effetto è piccolissimo per le velocità v a cui siamo abituati ma è stato misurato mediante orologi atomici. A velocità comparabili a quella della luce, invece, il ritardo diventa enorme. Adesso però abbiamo tutti gli strumenti per poter comprendere gli altri due esperimenti sull'etere. Einstein spiega in modo molto semplice l'effetto dell'aberrazione della luce proveniente da una stella con la legge di composizione dei moti. Resta l'esperimento di Fizeau, davvero interessante se visto nell'ottica della relatività. Applicando anche qui la legge di composizione delle velocità w' e per piccoli valori rispetto a c , otteniamo che

$$w' = \frac{c/n + v}{1 + \frac{vc/n}{c^2}} \cong \frac{c}{n} + \left(1 - \frac{1}{n^2}\right)v - \frac{v^2}{nc} \quad (25)$$

dove c/n è la velocità della luce in un mezzo con indice di rifrazione n . Se ora consideriamo che l'ultimo termine è ancora più piccolo e può essere trascurato rispetto agli altri termini ritroviamo che la velocità w' è in pratica

$$w' = \frac{c}{n} + \left(1 - \frac{1}{n^2}\right)v \quad (26)$$

Che è la formula di Fizeau. Tutti e tre gli esperimenti entrano in un'unica cornice. Una cornice in cui l'etere non esiste più, non è necessario. Capirete però adesso che se l'etere è stato seppellito nel cimitero delle teorie defunte, ne deriva che la luce è il primo caso di un'onda che si propaga in assenza di un mezzo "meccanico". La luce è una perturbazione del campo elettromagnetico, che è uno stato del vuoto. Punto. Dunque il vuoto ha stati, come la materia? Sì. Il vuoto non è nulla, è qualcosa, ma questo qualcosa non si comporta affatto come un mezzo materiale. L'eliminazione dell'etere ci catapulta in una realtà contro-intuitiva, fatta di un tempo e uno spazio che ritenevamo assoluti e che invece dipendono dallo stato di moto dell'osservatore, per cui non sono enti oggettivi. L'unica realtà oggettiva e non dipendente dallo stato dell'osservatore è la luce e l'esistenza di qualcosa che da allora in avanti sarà chiamato spazio-tempo, in cui eventi simultanei per un osservatore possono non esserlo più per altri in moto relativo ma in cui concetti come la velocità della luce o la continuità topologica dello spazio (non esistono discontinuità atomiche nello spazio) e dello scorrere del tempo (non ci sono salti temporali) sono assoluti. Una realtà in cui c diventa un limite invalicabile, nessun messaggio può essere spedito a velocità superluminali.

La scienza della causa e dell'effetto

Concludiamo questa prima parte del lavoro con una breve digressione sul concetto di causa-effetto (causalità) che emerge dal quadro teorico osservato e dal concetto di spazio-tempo. Supponiamo di osservare (Fig. 22a) tre sequenze di un evento in un diagramma spaziotemporale. Faccio notare che una dimensione spaziale è stata rimossa per consentire la visualizzazione dei fenomeni spaziali che descriveremo. Ora ciascuna fetta rappresenta una collezione di eventi in tre istanti temporali t_1 , t_2 e t_3 . In questa sequenza il Sole è fermo rispetto all'osservatore mentre sia il satellite che la Terra sono in moto. All'istante t_1 il satellite emette un segnale rappresentato da un cerchio (un'onda sferica nella realtà di uno spazio tridimensionale) che si espande sempre più formando un cono detto *cono-luce*.

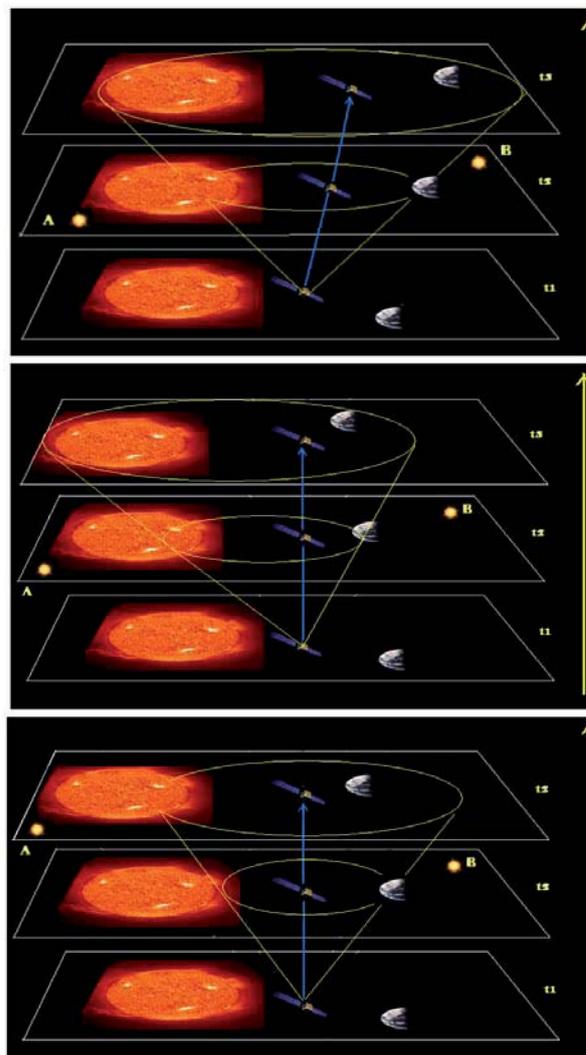


Figura 22: a) Alto, sequenza spaziotemporale di un evento (emissione di un segnale sferico di luce da parte di un satellite) rispetto a un osservatore solidale col Sole. I due eventi A e B sono in questo sistema di riferimento simultanei. b) Centro, stessa sequenza vista da un osservatore solidale col satellite ma in uno spaziotempo newtoniano in cui il tempo è assoluto e la simultaneità di due eventi è indipendente dallo stato di moto dell'osservatore. c) Basso, stessa sequenza ma vista in uno spaziotempo relativistico in cui il tempo è relativo. Notate che mentre il cono-luce si deforma nello spaziotempo newtoniano per consentire agli eventi di essere ancora simultanei, gli eventi sono ordinati in una sequenza temporale differente per permettere al cono-luce di essere indipendente dallo stato di moto dell'osservatore.

Ora all'istante t_2 il segnale di luce raggiunge *simultaneamente* sia il Sole che la Terra. In altre parole i due eventi *la luce raggiunge la superficie del Sole* (evento A) e *la luce raggiunge la superficie terrestre* (evento B) sono *simultanei*. Lo spazio-tempo è un concetto che ha senso anche nella fisica di Newton, ma in tale spazio tutti gli eventi che sono simultanei per un osservatore O lo sono anche per un altro O' in moto relativo uniforme per cui è possibile dare la definizione di un tempo oggettivo e comune a tutti. E' possibile quindi, in modo assoluto, oggettivo e universale, cioè valido per tutti gli osservatori in moto relativo,

sezionare lo spazio-tempo in una sequenza temporale di spazi contenenti gli eventi che sono simultanei per tutti gli osservatori. Il tempo è oggettivo.

Per la fisica di Newton, collocandoci nel sistema di riferimento solidale al satellite, la sequenza di eventi resta fissa ma il cono-luce cambierà come in figura 22b. Al contrario, nella fisica di Einstein, il cono luce resterà immutato, ma i due eventi A e B non saranno più solidali per un osservatore solidale col satellite. Questo è quanto. Ciò che resta però immutato tra un osservatore e l'altro è anche la sequenza dei due eventi causa effetto. In entrambi i sistemi di riferimento (Fig.22a e c) la causa che farà attivare un sensore posto sia sul Sole che sulla Terra ha la sua origine nel passato, ovvero nell'evento in cui il satellite invia un impulso collocato all'istante t_1 per entrambi gli osservatori in moto relativo. Faccio notare inoltre che i moti del satellite, della Terra e del Sole in entrambe le figure a e c sono individuati da traiettorie che sono meno inclinate di quella della luce rispetto all'asse verticale del tempo. Essi cioè si muovono più lentamente della luce. Possiamo anche esprimere le stesse cose in un modo più incisivo sostenendo che tutti gli eventi fisici connessi causalmente con l'evento iniziale in t_1 avranno linee di universo dentro il cono luce. Faccio notare infine che nessun segnale può connettere istantaneamente il satellite e la Terra. Se ciò fosse possibile la linea di universo di questo segnale che connette il satellite alla Terra sarebbe fuori dal cono-luce. La conclusione è che ognuno ha a che fare con un universo differente e tuttavia lo spaziotempo non è separabile. Non ha senso "affettare" lo spaziotempo in sequenze di eventi assoluti. E'

come se avessimo a disposizione un bel pane che rappresenta lo spazio-tempo e tutto ciò che c'è in esso (ovvero l'universo e tutti i suoi eventi passati, presenti e futuri). Ciascuno di noi, a seconda del proprio moto, lo affetta con tagli più o meno inclinati. C'è chi lo affetta tagliandolo verticalmente, chi lo affetta tagliando fette inclinate di 45 gradi e via dicendo. Ogni fetta rappresenta tutti gli eventi che sono simultanei per quell'osservatore. In effetti Poincarè provò che cambiare riferimento inerziale equivale a deformare (o affettare) lo spaziotempo come accadrebbe ad un quadrato che venisse stirato dai lati per diventare un rombo. Lo spaziotempo cioè si comporta come un foglio piano di gomma che è possibile tirare da un lato e dall'altro senza però cambiare il cono-luce. Queste operazioni corrispondono alle trasformazioni di Lorentz ovvero ai passaggi da un riferimento inerziale all'altro. Osservate la figura 23, dove ho rimosso per semplicità di visualizzazione, una ulteriore dimensione spaziale, di modo che lo spazio e il tempo formano insieme uno spaziotempo bidimensionale.

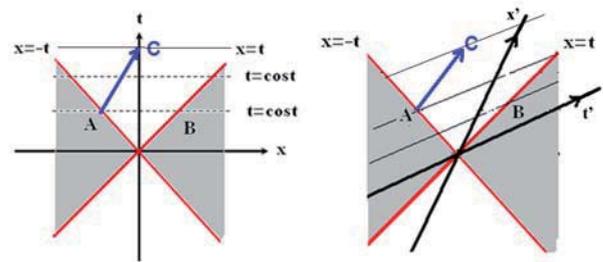


Figura 23: a) Sinistra, spaziotempo per un osservatore per il quale due eventi A e B sono simultanei. b) Destra, spaziotempo per un osservatore in moto relativo al primo, verso la direzione in cui accade B. In questo spaziotempo B viene prima di A. La relatività degli eventi (A prima di B o viceversa) deriva dal fatto che A non è la causa di B e viceversa. Pertanto possiamo sempre trovare degli stati di moto che non concorderanno sulla sequenza temporale di A e B. Se però A è la causa di un altro evento, C, allora A verrebbe temporalmente prima di C per entrambi gli osservatori e non esisterebbe nessuno stato di moto per cui A e C sono simultanei. L'unica differenza tra i due osservatori è che l'intervallo temporale che separa la causa A dall'effetto C è più breve per l'osservatore in moto, come se questo procedesse più speditamente verso il futuro (notate che il collegamento causale tra A e C-la linea blue-è dentro il cono luce che ha origine in A).

Immaginiamo che gli eventi A (Alice accende un laser) e B (Bob spegne un laser) avvengano simultaneamente per l'osservatore Charlie, come ho raffigurato a sinistra nella figura. Questo significa che essi sono due punti attraverso cui passa una retta parallela all'asse spaziale, ovvero definiscono un unico istante, essendo appunto simultanei. Charlie ha affettato il pane spazio-temporale con fette parallele al piano orizzontale. Se un quarto osservatore, che chiameremo Michelle, invece è in moto verso Bob (come nel satellite dell'esempio precedente) è come se l'asse temporale e quello spaziale, rispetto all'altro grafico, si siano inclinati verso un lato del cono-luce, formando un angolo più piccolo. Quali saranno per Michelle gli eventi simultanei, ovvero come affetterà il pane spaziotemporale? In questo caso, come nel precedente, gli spazi simultanei sono quelli paralleli all'asse spaziale, qui x' , ed essi individuano gli eventi simultanei, quelli a $t'=\text{cost}$. Come potete vedere però i due eventi A e B sul cono luce verranno visti in momenti differenti nel nuovo riferimento perché si troveranno su spazi simultanei differenti, come deve essere. In particolare l'evento B accadrà prima dell'evento A. Se invece Michelle si muove verso Alice vedrà questa sequenza di eventi ribaltata, prima accadrà A e solo dopo accadrà B. Occorre specificare, a costo di essere pedanti, che A e B non sono in questo caso l'uno la causa dell'altro. Se per Charlie A è la causa di B, ovvero se l'evento B accade solo perché nel passato A ha scatenato una sequenza di eventi che ha causato B, allora A sarà la causa di B anche per Michelle, Alice e Bob. La sequenza causa effetto cioè resta invariata e non può essere invertita. Se fosse così addio razionalità nel nostro universo. Pertanto sebbene lo spazio-tempo non si può separare in spazio e tempo in modo oggettivo, la sequenza causale degli eventi nel nostro universo è oggettiva. Prima di congedarmi da questa seconda parte vorrei volare un po' con la fantasia per arrivare a chiederci: come apparirebbe il mondo a cavallo di un raggio di luce? Gli assi del tempo e dello spazio si inclineranno sempre più man mano che la velocità dell'osservatore aumenta, seguendo la deformazione di Poincarè che abbiamo visto sopra, fin quando, raggiunta la velocità c , i due assi del tempo e dello spazio coincideranno e l'angolo di separazione degli assi spaziotempo-

rali dal cono-luce diviene zero. In tal caso per un tale osservatore luminale tutti gli eventi saranno simultanei! Per un raggio di luce il tempo non scorre affatto e lo spazio, dal suo punto di vista, viene percorso all'istante! La luce vivrebbe la sua intera storia, da quando il raggio viene creato a quando viene distrutto, in un attimo. Pensate ad un raggio di luce che parte dalla galassia di Andromeda e giunge ad eccitare la vostra retina. La luce, dal vostro punto di vista, avrà impiegato due milioni di anni per giungere nei vostri occhi, ma dal suo punto di vista, da quando essa è stata emessa a quando è stata assorbita eccitando i fotorecettori del vostro occhio, è passato un istante, anzi nemmeno quello! Un raggio che fosse stato prodotto ai tempi del big bang e che venisse distrutto in un eventuale big crunch finale vedrebbe l'intera storia dell'universo compressa in un istante. È come se la luce saltasse letteralmente l'intervallo temporale intermedio misurato da noi. Questo implica che per osservatori che viaggiano a velocità inferiori a c rispetto alle stelle fisse, la storia dell'universo (ovvero della evoluzione stellare e di tutto ciò che accade nei loro dintorni) apparirà comunque più breve fintanto che saranno in moto. Questi viaggiatori in poche parole corrono verso il futuro più rapidamente di un osservatore immobile rispetto alle stelle "fisse". Il viaggio nel tempo, e più precisamente verso il futuro, esiste. Viaggiamo tutti nel tempo futuro quando viaggiamo nello spazio.

Parte III: grani di luce

Un mondo statistico?

Quello presentato nelle parti precedenti è il quadro offertoci dalle teorie classiche della luce. Il 1905 fu l'anno in cui accanto alla teoria della relatività una nuova visione dell'universo prese piede con la nascita del concetto di grano di luce: il fotone. Era nata l'ottica quantistica, una teoria che ci spingerà in un territorio dalle regole "folli" ma coerenti con tutto ciò che sappiamo sulla luce. Fu sempre Einstein ad introdurre l'idea, ma, contrariamente all'articolo sulla relatività in cui egli

presuppose la validità delle leggi classiche dell'elettromagnetismo di Maxwell, l'articolo sulla natura quantistica della luce faceva leva sui limiti della teoria maxwelliana, e propose che la luce, più che propagarsi nello spazio distribuendo l'energia in modo continuo, si propagasse a pacchetti di energia che battezzò, appunto, fotoni. In questo Einstein dimostra la sua natura di giovane opportunista, nel senso di lavorare a briglie sciolte senza lasciarsi incastrare nella visione accademica a lui contemporanea e sfruttando tutto ciò che proveniva da altre branche, come la teoria dei quanti di Planck. Perché introdurre questa idea? A causa di un fenomeno che non si lasciava inquadrare: l'effetto fotoelettrico.

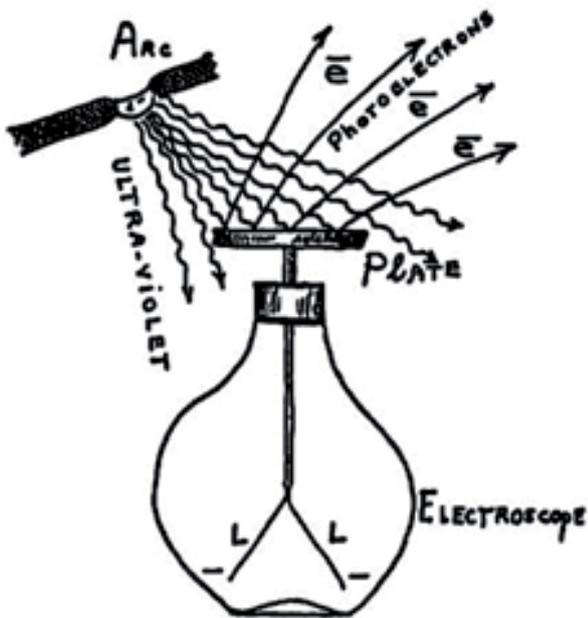


Figura 24: Effetto fotoelettrico.

L'effetto è il seguente: illuminando con luce ultravioletta una placca di metallo collegata ad un elettroscopio si osserva che le foglioline si caricano (Fig.24).

Questo indica che la luce estrae dal piatto delle cariche elettriche, gli elettroni, caricando le foglioline d'oro dell'elettroscopio. Possono le teorie classiche spiegare questo fenomeno? Fino a un certo punto sì, usando il concetto proprio di questa teoria secondo cui la luce cede energia al piatto proporzionalmente al modulo quadro del campo elettrico. Se si usa luce infrarossa, ovvero più lunga della luce ultravioletta, la fisica classica prevede che comunque l'elettrone sarà scalzato e che il tempo necessario affinché

ciò accada dipenderà dalla intensità della luce stessa, ovvero dall'ampiezza del campo elettrico. E invece ciò che si osserva è che se si usa luce infrarossa, indipendentemente dalla intensità e dal tempo atteso, non verrà estratto nemmeno un elettrone. Come si spiega? Questo e altri fenomeni come la radiazione di corpo nero non trovavano alcuna soluzione in seno alla teoria di Maxwell. Ma Einstein ebbe una "illuminazione" e, partendo dall'assunto che la luce non si propaghi nello spazio vuoto come un'onda, ma come dei corpuscoli, che battezzò fotoni, risolse l'enigma. L'ipotesi fu che l'energia di un raggio di luce non dipende dall'ampiezza del campo, ma dal numero N di proiettili, i fotoni appunto, che colpiscono la piastra e dall'energia del singolo fotone, che, secondo la teoria di Planck è proporzionale alla frequenza ν della luce. Pertanto l'intensità I sarà data da

$$I = Nh\nu \quad (27)$$

dove h è la costante di Planck. In questo modo Einstein interpretò l'effetto fotoelettrico come un urto tra due biglie: un elettrone e un fotone. L'elettrone ha una energia minima data dalla energia di legame col metallo ($e\phi$), per cui dalla conservazione dell'energia ricaviamo che l'energia cinetica di rilascio dell'elettrone è data dalla differenza tra l'energia del singolo fotone, $h\nu$, e tale energia di legame

$$\frac{1}{2}mv^2 = h\nu - e\phi \quad (28)$$

Siccome l'energia cinetica è sempre positiva ne risulta che

$$\nu \geq e\phi/h \quad (29)$$

Ovvero che esiste una frequenza di soglia al di sotto della quale non è possibile estrarre alcunché, a meno di non avere più fotoni che, simultaneamente, colpiscono l'elettrone, evento altamente improbabile. Una teoria del genere spiegava benissimo questi fenomeni che risultavano incomprensibili mediante una teoria elettromagnetica della luce.

L'esistenza dei fotoni come mattoni fondamentali di un raggio di luce pone delle questioni enormi circa la natura della realtà. Tutti gli esperimenti visti prima infatti (la riflessione, l'interfe-

renza e la polarizzazione) devono essere descritti nuovamente usando il concetto discreto di corpuscolo al posto di quello continuo di onda. Come è possibile che un fotone, particella di luce, possa mostrare interferenza, concetto tipico di un'onda? Chiariamo meglio questo concetto. Si potrebbe pensare ad esempio che il fotone in realtà si sparpagli dappertutto come un gomitolo di lana che si srotola, ma ciò non spiegherebbe un fatto sperimentale ormai assodato: abbiamo visto sopra che l'intensità luminosa di un raggio di luce è proporzionale al numero dei fotoni, sicché riducendo l'intensità è possibile mettersi nelle condizioni di avere un fotone alla volta nello spazio. Come viene individuato un fotone? Come una entità delocalizzata e spalmata spazialmente o come un granulo localizzato? Ebbene grazie a detector ultrasensibili si è scoperto che i fotoni si presentano sempre come pacchetti indivisibili e localizzati, ovvero come proiettili. Non si rivelerà mai né delocalizzato né diviso. Difficile pertanto credere che, nello spazio tra sorgente e rivelatore, il fotone non abbia una esistenza "puntiforme".

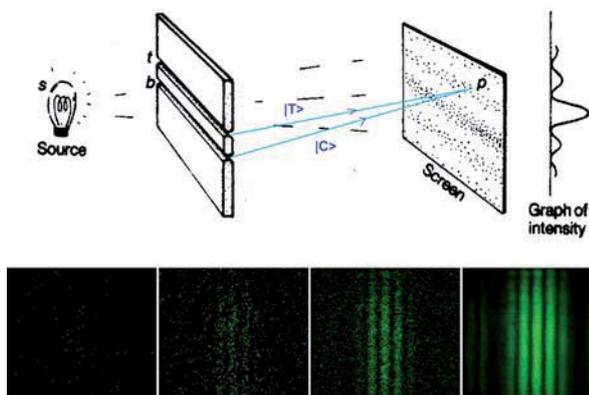


Figura 25: Interferenza di fotoni che passano tra due fenditure. I fotoni si accumulano sullo schermo in modo casuale e scorrelato ma producendo alla fine un pattern.

Ecco allora il paradosso: se la luce si individua come pacchetti localizzati, come spiegare la formazione della figura di interferenza in un tipico esperimento a due fenditure? Non avendo più onde delocalizzate ma entità localizzate come spiegare tutti quegli effetti ondulatori tipici di un'onda elettromagnetica? Se si fa passare ciascun fotone dalle fenditure si osserva che col passare del tempo i singoli impatti si distribuiscono, all'inizio, apparentemente in modo casuale,

per poi formare, dopo molti eventi di impatto, un pattern, proprio quello tipico dell'interferenza (Fig.25). I fotoni insomma vengono sempre rilevati come proiettili e l'aspetto periodico delle frange (che prima attribuivamo alla periodicità di un'onda fisica) emerge dall'accumulo statistico degli impatti. Come è possibile che da una serie di impatti casuali emerga lo schema ordinato delle frange di interferenza?

Testa o Croce?

Il caso in fisica classica, e in generale nella scienza, è un animale strano. Esso denota l'ignoranza che abbiamo dei dettagli. In meccanica il caso è la nostra non conoscenza perfetta delle condizioni iniziali con cui, ad esempio, viene lanciata una moneta. Immaginiamo di porre una moneta in una scatola e di agitarla. Ovviamente non sapremo se, quando la scatola verrà aperta, la moneta mostrerà testa (T) o croce (C). Di fatto essa è un sistema fisico a due stati, T o C, mutuamente esclusivi ma non possiamo dire nulla sullo stato della moneta prima dell'apertura della scatola. La nostra ignoranza è totale. La moneta si troverà ovviamente in uno dei due stati ma noi non sappiamo quale. Se apriamo la scatola e troviamo ad esempio T non sapremmo inoltre prevedere cosa accadrà la volta successiva che ripetiamo l'esperimento. La moneta non ha memoria degli esiti precedenti. Tuttavia, nonostante questa ignoranza nelle previsioni, qualcosa possiamo dire: pur non potendo fare previsioni sugli esiti singoli possiamo fare previsioni statistiche, cioè su un gran numero di prove. Ripetiamo l'esperimento un migliaio di volte, e sappiamo infatti che, per la "legge" dei grandi numeri, dopo molti lanci avremo in media il 50% degli esiti che danno T e il 50% che danno C. In altre parole mentre per l'esito di ogni prova siamo in balia dell'imprevedibilità e dunque del caso, dopo migliaia di lanci possiamo essere sicuri che per lo meno la distribuzione statistica rientrerà nel dominio della prevedibilità. Formalmente posso indicare col simbolo $|T\rangle$ e $|C\rangle$ i due stati della moneta e raffigurarli in uno spazio bidimensionale vettoriale. Definisco gli stati testa e croce come stati ortogonali, asserendo con questo termine che, dopo il lancio, la moneta si presenterà o in uno stato o nell'altro, ma dove l'uno esclude

l'altro. Ad ogni apertura della scatola non potrà mai trovare lo stato della moneta testa e croce. Non vale la congiunzione *e* ma la congiunzione *o*: o Testa o Croce. Per indicare questa proprietà di ortogonalità diremo che la proiezione di uno stato sull'altro, ovvero il prodotto scalare tra questi vettori, è *nullo*. Indicheremo questo prodotto scalare mediante i simboli

$$\langle C|T \rangle = \langle T|C \rangle = 0$$

$$\langle C|C \rangle = \langle T|T \rangle = 1 \quad (30)$$

Come possiamo descrivere la nostra ignoranza (o meglio il nostro grado di conoscenza) sullo stato del sistema prima dell'apertura della scatola? Mediante lo stato indeterminato (per noi ma non per chi fosse dentro la scatola come la moneta stessa)

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|T\rangle + |C\rangle) \quad (31)$$

indicando con la somma che, a causa della nostra ignoranza sulle condizioni di lancio, l'esito potrà essere indifferentemente $|T\rangle$ o $|C\rangle$. Dico indifferentemente perché ciascuno stato è pesato con lo stesso coefficiente $1/\sqrt{2}$. Ma perché proprio questo coefficiente? La componente dello stato su uno dei due vettori unitari, ovvero il prodotto scalare, è esattamente un segmento lungo $1/\sqrt{2}$. In poche parole il prodotto scalare $\langle T|\psi\rangle$ è pari a $1/\sqrt{2}$. Elevando al quadrato questo risultato otteniamo proprio la probabilità P statistica di avere Testa, cioè $\frac{1}{2}$.

$$P = |\langle T|\psi\rangle|^2 = |\langle C|\psi\rangle|^2 = \frac{1}{2} \quad (32)$$

Faccio notare che tale probabilità è di natura *epistemologica* e non *ontologica*, ovvero riguarda la nostra conoscenza e non la realtà esistenziale della moneta. Questo, in quanto *in fisica classica il puro caso non esiste*, riguarda solo la nostra ignoranza sulle condizioni al contorno. Le teorie classiche sono infatti deterministiche, nonostante un sistema fisico non sia il più delle volte prevedibile. Sottolineo, a costo di sembrare pedante, che la moneta si trova comunque in uno dei due stati $|T\rangle$ o $|C\rangle$ prima dell'apertura della scatola e che la formula 31 esprime solo la nostra ignoranza. Dopo l'apertura della scatola acquisiamo

informazione sullo stato "reale" della moneta e possiamo dire che l'apertura coincide con una proiezione dello stato di informazione $|\psi\rangle$ in uno dei due stati, ad esempio $|T\rangle$.

$$|\psi\rangle \Rightarrow |T\rangle \quad (33)$$

La freccia indica l'apertura della scatola e l'acquisto di informazione dell'osservatore. La sua ignoranza sullo stato posseduto oggettivamente dalla moneta è sparita istantaneamente.

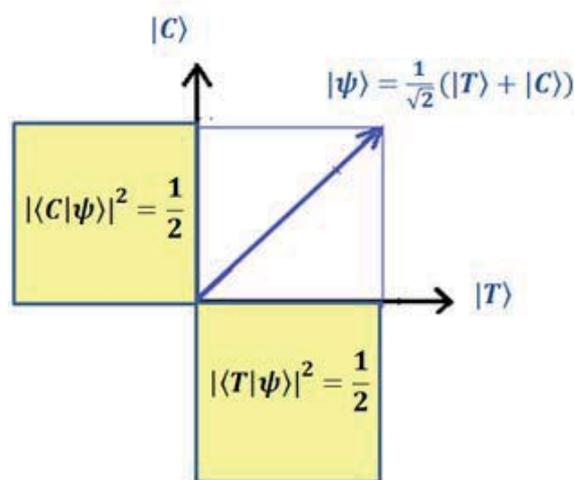


Figura 26: Regole quantistiche di sovrapposizione degli stati e collasso della funzione d'onda.

Supponiamo ora di voler vedere cosa accade a due monete nella scatola, e chiediamoci qual è la probabilità di trovare testa per la moneta A e croce per la moneta B. Lo stato del sistema adesso è descritto dal cosiddetto prodotto tensoriale, $|\psi_A, \psi_B\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$ che ci avverte che lo stato del sistema è dato dalla presenza di due sottosistemi (le due monete). Il prodotto tensoriale tra stati è molto semplice da descrivere:

$$|\psi_A, \psi_B\rangle = \frac{1}{\sqrt{2}}(|T\rangle + |C\rangle)_A \otimes \frac{1}{\sqrt{2}}(|T\rangle + |C\rangle)_B$$

$$= \frac{1}{2}(|T, T\rangle + |T, C\rangle + |C, T\rangle + |C, C\rangle) \quad (34)$$

dove per semplicità abbiamo eliminato gli indici A e B all'ultimo membro. Il formalismo ci avverte di una cosa sensata: le due monete si troveranno in uno dei quattro stati in parentesi, ma noi non sappiamo quale finché non apriremo la scatola. Stando alle regole del prodotto scalare, la probabilità di ottenere simultaneamente testa

per la prima moneta e croce per la seconda è data da

$$P(T, C) = |\langle T, C | \psi_A, \psi_B \rangle|^2 = \frac{1}{4} = 25\% \quad (35)$$

come deve essere. Il formalismo funziona. Notiamo che:

a) adoperiamo il prodotto \otimes nel caso in cui ci poniamo domande del tipo: cosa accade ad A e simultaneamente a B?

b) lo stato composto $|\psi_A, \psi_B\rangle$ adesso "vive" in uno spazio vettoriale a quattro dimensioni dove $|T, T\rangle; |T, C\rangle; |C, T\rangle; |C, C\rangle$ rappresentano gli esiti possibili, ovvero i vettori ortogonali di base.

Prima di congedarci da questo aspetto formale vorrei evidenziare come le probabilità, almeno nel caso della moneta, non dipendono da un eventuale fattore di fase complesso e unitario. Avremmo potuto anche scrivere lo stato della moneta prima dell'apertura della scatola in questo modo

$$|\psi\rangle = \frac{1}{\sqrt{2}}(e^{i\varphi}|T\rangle + |C\rangle) \quad (36)$$

senza alterare le probabilità finali. Il fattore di fase ruota di un angolo φ lo stato $|T\rangle$. Infine possiamo sempre identificare i due vettori $|T\rangle$ e $|C\rangle$ con due vettori del tipo $(1; 0)$ e $(0; 1)$ (vedi Box3). Lo stato $|\psi\rangle$ diventa allora $(e^{i\varphi}/\sqrt{2}; 1/\sqrt{2})$. Se l'angolo è 90° , allora $|\psi\rangle = (i/\sqrt{2}; 1/\sqrt{2})$ ma la probabilità che, aprendo la scatola, troviamo testa è comunque $|i/\sqrt{2}|^2 = 1/2$. Adesso torniamo ai fotoni sullo schermo.

La polarizzazione dei fotoni

Cosa determina l'emergere del *pattern* di interferenza dalla casualità con cui avvengono gli impatti? Come può il caso generare ordine? Poiché il fotone o passa dalla fenditura A o passa dalla fenditura B, nei fatti si comporta come un sistema a due stati. I fisici hanno sviluppato un formalismo molto elegante e generale per un tale sistema. Tale formalismo è esattamente quello della moneta vista sopra, in cui lo stato $|\psi\rangle$ è dato da

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|A\rangle + |B\rangle) \quad (37)$$

dove A e B sono le due alternative. Tuttavia, come vedremo ora, vi è una differenza notevole tra quanto asserito per una moneta e un oggetto quantistico a due stati. Invece di trattare subito le alternative di traiettoria di un fotone, valuteremo le due alternative di polarizzazione. Proviamo pertanto a descrivere lo stato di polarizzazione di un fotone, una proprietà che si descrive mediante due stati (ad esempio polarizzazione verticale V oppure orizzontale O) e rientra nella cornice formale appena sviluppata. Prima però chiediamoci cosa significa classicamente che la luce è polarizzata V? Vuol dire che se inviamo la luce verso un polarizzatore orientato V la luce passa tutta. Se invece la inviamo contro un polarizzatore orizzontale essa sarà bloccata. Se invece la inviamo contro un polarizzatore orientato lungo un angolo θ rispetto alla verticale allora, seguendo la ben nota legge di Malus, l'intensità uscente sarà ridotta di un fattore $(\cos\theta)^2$.

Dal punto di vista classico questi esperimenti si spiegano ammettendo che la polarizzazione sia la direzione di oscillazione del campo elettrico che, essendo di natura vettoriale, può essere scomposto lungo due versori tra loro ortogonali qualsiasi $|\theta\rangle$ e $|\frac{\pi}{2} + \theta\rangle$ nel seguente modo

$$\begin{aligned} \vec{E}_{tot} &= \vec{E}_\theta + \vec{E}_{\frac{\pi}{2} + \theta} \\ &= \cos\theta E_{tot} |\theta\rangle - \sin\theta E_{tot} |\frac{\pi}{2} + \theta\rangle \end{aligned} \quad (38)$$

dove intendiamo con $|\theta\rangle$ e $|\frac{\pi}{2} + \theta\rangle$ rispettivamente il versore parallelo e ortogonale alla direzione di polarizzazione del polarizzatore. L'intensità dell'onda uscente dal polarizzatore è proporzionale al modulo quadro della componente del campo, ovvero

$$\frac{I_\theta}{I_{tot}} = \left| \frac{E_\theta}{E_{tot}} \right|^2 = (\cos\theta)^2 \quad (39)$$

come detto prima. Osservando il riflesso della luce sulla strada, che è polarizzato "O", esso può essere eliminato con occhiali polaroid, che sono polarizzati "V", in quanto $\cos 90^\circ = 0$. Questo se siamo in condizioni di fasci intensi. Ma se abbassiamo l'intensità possiamo descrivere la trasmissione della luce mediante singoli fotoni che attraversano il polarizzatore. Ma che significa che un fotone è polarizzato "V"? Vuol

Box 3: Il prodotto tensoriale tra stati

Se lo stato di una moneta è dato da $|T\rangle$; $|C\rangle$ e il generico stato descritto dall'equazione (31) è $|\psi\rangle$, questi possiamo definirli mediante le componenti vettoriali di uno spazio cartesiano (ricordiamo però che gli stati in generale sono complessi) in modo da scrivere

$$|T\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \quad |C\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}; \quad |\psi\rangle = \begin{pmatrix} \psi_T \\ \psi_C \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Gli stati di due monete "vivono" in due spazi separati. Il prodotto tensoriale produce uno spazio comune definito da:

$$|\psi_A, \psi_B\rangle = |\psi_A\rangle \otimes |\psi_B\rangle = \begin{pmatrix} a_T \\ a_C \end{pmatrix} \otimes \begin{pmatrix} b_T \\ b_C \end{pmatrix} = \begin{pmatrix} a_T \begin{pmatrix} b_T \\ b_C \end{pmatrix} \\ a_C \begin{pmatrix} b_T \\ b_C \end{pmatrix} \end{pmatrix} = \begin{pmatrix} a_T b_T \\ a_T b_C \\ a_C b_T \\ a_C b_C \end{pmatrix}$$

Ad esempio

$$|T, C\rangle = |T\rangle \otimes |C\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ 0 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Pertanto da due vettori bidimensionali ne otteniamo di quadridimensionali. Lo spazio prodotto ha la dimensione data dal prodotto delle dimensioni dei singoli spazi.

dire che se inviamo questo fotone contro un filtro polarizzatore orientato verticalmente esso lo supererà sempre (Fig. 27), e non supererà mai un polarizzazione orizzontale (Fig. 28). Pertanto le probabilità di trasmissione sono

$$P_V = |\langle V|V\rangle|^2 = 1$$

$$P_O = |\langle V|O\rangle|^2 = 0 \quad (40)$$

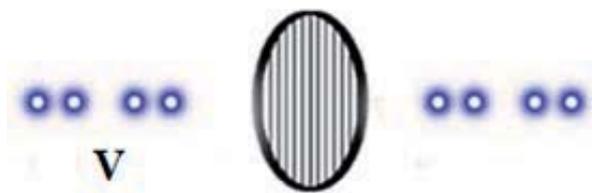


Figura 27: Esperimento di polarizzazione di fotoni polarizzati verticalmente attraverso un polarizzatore verticale. Tutti i fotoni passano.

Questo implica che un fotone polarizzato a 45° , che indicheremo con il simbolo $|+\rangle$, attraverserà con certezza un filtro polarizzato a

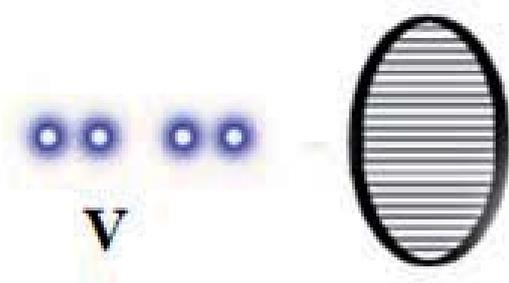


Figura 28: Esperimento di polarizzazione di fotoni polarizzati verticalmente attraverso un polarizzatore orizzontale. Non passa nessun fotone.

45° (Fig 29) e non ne attraverserà uno polarizzato perpendicolarmente a tale direzione, cioè a 135° .

$$P = |\langle 45^\circ|+\rangle|^2 = 1 - |\langle 135^\circ|+\rangle|^2 = 1 \quad (41)$$

Cosa accade però se tale fotone $|+\rangle$ viene spedito contro un filtro polarizzatore verticale V? Ebbene gli esperimenti a singolo fotone indicano che esso passerà o meno il test in modo assolutamente casuale, col 50% delle probabilità. In altre

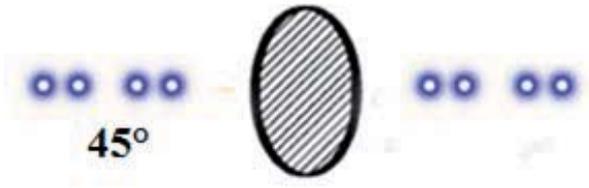


Figura 29: Esperimento di polarizzazione di fotoni polarizzati a 45° attraverso un polarizzatore a 45° . Passano tutti i fotoni.

parole, spedendo un gran numero di fotoni nello stato $|+\rangle$ (sappiamo che sono così perché hanno superato un precedente test di polarizzazione a 45°) contro tale filtro “verticale V”, allora metà di questi, in modo assolutamente arbitrario e casuale, passeranno il test, e metà no, proprio come per la moneta.

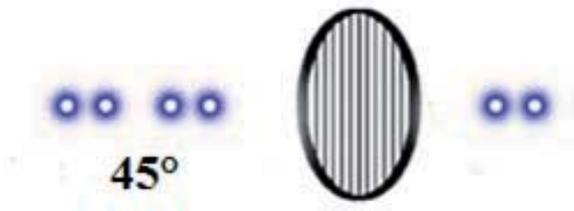


Figura 30: Esperimento di polarizzazione di fotoni polarizzati a 45° attraverso un polarizzatore verticale. Solo il 50% dei fotoni passa.

Prima della misura possiamo dire allora che lo stato del fotone può essere descritto nel modo seguente

$$|+\rangle = \frac{1}{\sqrt{2}}(|V\rangle + |O\rangle) \quad (42)$$

Il significato è il solito: se effettuiamo un test di polarizzazione verticale la probabilità che il fotone superi il test sarà

$$P = |\langle V|+\rangle|^2 = \frac{1}{2} \quad (43)$$

Come possiamo vedere, il formalismo sviluppato nel caso della moneta è assolutamente in grado di descrivere cosa accade con fotoni inviati contro filtri polarizzatori. Ricordo però che nel caso della moneta classica lo stato $|\psi\rangle$ rappresentava solo il nostro grado di conoscenza dello stato della moneta. Ad esso non poteva essere associato alcun elemento di realtà, mentre gli stati $|T\rangle$ e $|C\rangle$ erano oggettivi. Nel caso del fotone invece lo stato $|+\rangle$ ha senso reale, anzi è il solo senso reale che possiamo attribuire

al fotone in sovrapposizione $\frac{1}{\sqrt{2}}(|V\rangle + |O\rangle)$, in quanto un fotone preparato in questo stato supera con certezza un test a 45° . Dunque, essendo in tale stato in modo oggettivo, non può possedere oggettivamente *anche* lo stato $|V\rangle$ oppure lo stato $|O\rangle$, proprio il contrario della moneta! A meno di non immaginare una qualche struttura “nascosta” particolare del fotone. Ecco qui l’apparente paradosso: se accettiamo l’interpretazione quantistica degli stati data sopra anche per la moneta avremmo che essa è sia testa che croce (interpretazione alla Feynman) o non si trova in nessuno dei due stati (interpretazione alla Bohr), ma è possibile trovare un test che dia senso alla sovrapposizione. Ovviamente per sistemi macroscopici come monete tutto questo non ha alcun senso, mentre per fotoni, elettroni, atomi e molecole è proprio quello che accade. A causa di ciò nel formalismo quantistico il processo indicato dalla freccia

$$|+\rangle \implies |V\rangle. \quad (44)$$

adesso non indica più che la nostra conoscenza è aumentata all’improvviso, ma che il sistema improvvisamente è collassato in V a seguito della misura di polarizzazione (corrispondente all’apertura della scatola).

La misurazione ha creato, in modo genuinamente casuale, l’elemento di realtà V.

Infatti se dopo questo test ne facciamo un altro sullo stesso fotone collassato, esso supererà il test di polarizzazione con certezza. In linea generale se inviamo un fotone polarizzato verticalmente contro un polarizzatore orientato con un angolo θ generico allora lo stato del fotone può essere scritto nel modo seguente rispetto al primo

$$|V\rangle = \cos\theta |\theta\rangle - \sin\theta \left| \frac{\pi}{2} + \theta \right\rangle \quad (45)$$

e la probabilità che il detector disposto dopo il polarizzatore suoni sarà

$$P = |\langle \theta|V\rangle|^2 = (\cos\theta)^2 \quad (46)$$

che è la legge di Malus. Quanto detto ci porta a concludere in via del tutto generale che

Se un fotone (o un generico sistema quantistico) ha due possibili alternative, esso vivrà in una sovrapposizione di queste due possibilità finché, a seguito di una interazione con un apparato di misura classico

e macroscopico, esso non collassa in una delle due realtà.

Abbiamo visto come nel caso della “moneta classica” nella scatola, lo stato rappresenti la nostra conoscenza (o il suo complementare -l’ignoranza-) sull’esito della “misura”, ma anche la certezza statistica che, eseguendo molte misurazioni, avremo per il 50% teste e per il 50% croci. La moneta tuttavia si trova in uno stato ben definito nella scatola indipendentemente se la apriamo o meno. Possiamo enunciare un criterio di realismo classico (RC) che vale per tutte le teorie classiche standard della fisica.

RC: Lo stato $|\psi\rangle$ esprime una conoscenza incompleta circa lo stato della moneta e con la misura (apertura della scatola) possiamo aumentarla; la moneta era oggettivamente in uno dei due stati anche prima dell’apertura della scatola; l’apertura della scatola fa collassare l’incertezza della nostra conoscenza.

L’ipotesi RC è ciò che rende “classica” una descrizione della realtà. L’interpretazione della $|\psi\rangle$ in fisica quantistica, nella sua versione alla Bohr, è però totalmente diversa. Essa non solo esprime ciò che sappiamo della moneta nella scatola ma essa esprime in modo completo il suo stato “reale”: la moneta quantistica si trova realmente e in modo oggettivo in uno stato indeterminato di testa e croce, ipotesi che potremmo definire “realismo quantistico (RQ)”.

RQ: la nostra conoscenza del sistema è completa e in nessun modo possiamo aumentarla; la moneta è oggettivamente in sovrapposizione di stati (come se non esistesse in nessuna delle due situazioni separate o esistesse in entrambe le alternative); l’apertura della scatola fa collassare lo stato quantistico in uno stato classicamente sensato che prima non esisteva in senso classico.

Questa visione RQ va sotto il nome di *interpretazione di Copenaghen* dovuta principalmente a Bohr (principio di complementarità), Heisenberg (principio di indeterminazione) e Born (interpretazione probabilistica). In RQ il coefficiente davanti agli stati non esprime la probabilità che la moneta sia nello stato testa prima di aprire la scatola, ma la probabilità che *aprendo* la scatola, all’atto cioè della misura, lo stato indeterminato della moneta collassi nello stato testa, cioè si *oggettivi*. Torniamo ora alle nostre considerazioni sui fotoni e sulle monete. Fa qualche differenza pratica aderire a RC o a RQ? Per la moneta

no, in quanto aprendo la scatola, la troveremo sempre in uno dei due stati indipendentemente se prima era già nello stato osservato (RC) o in sovrapposizione indeterminata (RQ). Infatti in entrambe le interpretazioni dello stato $|\psi\rangle$ tutto ciò che è possibile fare è prevedere l’esito statistico delle misure. La faccenda è un po’ simile al caso in cui, uscendo dalla stanza dopo aver scritto questa pagina sul mio PC (il tempo di prendere un caffè), rientrando ritrovo il mio PC sul tavolo in standby. Cosa è successo al computer nell’intervallo di tempo in cui ero fuori dalla stanza? Nel caso della RC diremmo che era andato in modalità standby prima di rientrare nella stanza. Nel caso della RQ diremmo che proprio l’apertura della stanza ha determinato il collasso del PC in quello stato, mentre prima esso era in una sovrapposizione non oggettiva di acceso e spento. Non possiamo verificare quale delle due ipotesi sia vera in quanto gli esiti sono sempre gli stessi. E allora perché preoccuparsi di questi accademismi? Lo vedremo tra qualche riga.

Sovrapposizione di due cammini

Finché si parla di sovrapposizioni di stati si potrebbe essere tentati ad ammettere che questa interpretazione non epistemologica (non inerente cioè alla nostra ignoranza) ma ontologica (inerente alla realtà dell’oggetto studiato), ovvero l’interpretazione di Copenaghen, sia legata solo a problematiche di polarizzazione. Invece è qualcosa di molto generale e connesso perfino alla propagazione nello spazio. E’ possibile infatti associare a stati di polarizzazione dei percorsi spaziali o traiettorie e vedere cosa accade. Prendiamo un cristallo birifrangente come in figura 31.

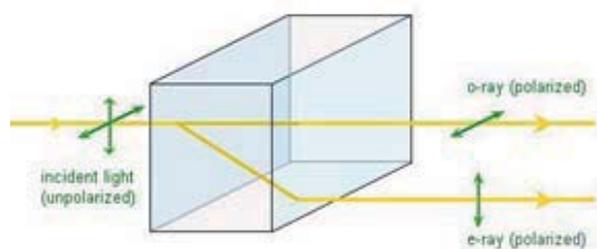


Figura 31: Cristallo birifrangente.

Un cristallo birifrangente separa un fascio di luce non polarizzato in due fasci paralleli polariz-

zati ortogonalmente. Il fascio ordinario (ovvero quello che segue la nota legge di Snell della rifrazione) risulterà polarizzato O, mentre quello straordinario sarà polarizzato V. Ovviamente se il fascio incidente è polarizzato già dall'inizio lungo V allora avremo in uscita solo il raggio straordinario. Se è polarizzato O già dall'inizio avremo in uscita solo il raggio ordinario. Se inviamo infine un raggio polarizzato a 45° avremo ovviamente due fasci di pari intensità polarizzati uno O e uno V. Questa situazione è schematizzata in figura 32.

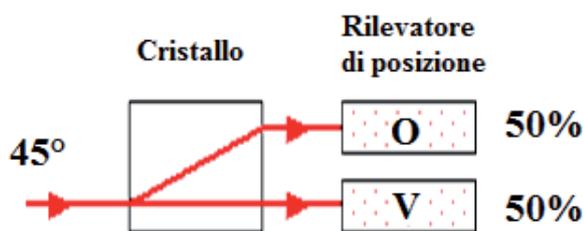


Figura 32: Esperimento di separazione di un fascio di fotoni polarizzati a 45° attraverso un cristallo birifrangente.

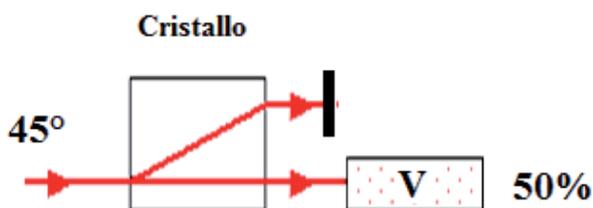


Figura 33: Esperimento di separazione di un fascio di fotoni polarizzati a 45° attraverso un cristallo birifrangente dove un fascio è bloccato.

Ci chiediamo cosa accade se inviamo un fotone alla volta. Ovviamente stando alla sovrapposizione e al fatto che lo stato di polarizzazione O è identificato col percorso superiore (su) e che lo stato di polarizzazione V è identificato col percorso inferiore (giù) avremo che

$$|+\rangle = \frac{1}{\sqrt{2}}(|V\rangle + |O\rangle) = \frac{1}{\sqrt{2}}(|giù\rangle + |su\rangle) \quad (47)$$

Questo implica che il rivelatore inferiore (o superiore) suonerà il 50% dei casi (il quadrato del coefficiente davanti allo stato $|giù\rangle$). Poniamoci una domanda: quando il contatore inferiore fa

click il fotone polarizzato a 45° è passato per il raggio inferiore o superiore? Ovviamente diremo inferiore. Infatti se blocchiamo il primo fascio il contatore inferiore farà click il 50% delle volte. Quindi il fotone a 45°, concludiamo, a volte passa per il percorso superiore, altre inferiore, a caso. La sovrapposizione di percorsi non è una situazione "reale" più di quanto non lo sia quella della moneta in sovrapposizione testa e croce. Eppure osservate ora cosa accade disponendo un altro cristallo ma ribaltato e disponiamo oltre questo cristallo un filtro polarizzatore a 45° e un contatore di fotoni (Fig. 34F1). Nel primo esperimento blocchiamo il fascio superiore. Quanti fotoni attraverseranno il polarizzatore? Ovviamente la probabilità che il fotone in ingresso $|+\rangle$ prenda il percorso giù è pari al 50% ($P_1 = |\langle V|+\rangle|^2 = \frac{1}{2}$). Una volta però che questo percorso sarà stato preso, il fotone sarà polarizzato $|V\rangle$, per cui quando incontrerà il filtro avrà una probabilità del 50% di superare il test ($P_2 = |\langle 45^\circ|V\rangle|^2 = \frac{1}{2}$). La probabilità risultante sarà il prodotto delle probabilità, ovvero il 25% e l'intensità del fascio in ingresso sarà ridotto di $\frac{1}{4}$. È ciò che si osserva, un fotone dopo l'altro (Figura 34F1).

Nel secondo esperimento blocchiamo il fascio inferiore e otterremo, analogamente a quanto visto nel primo esperimento, sempre il 25% (Figura 34F2).

Nel terzo esperimento ((Figura 34F3) non blocchiamo nulla. Se il fotone in ingresso prendesse o un percorso o un altro (ovvero se collassasse o in V o in O) durante il tempo di volo, ci aspetteremmo un risultato dato dalla somma delle due situazioni viste prima, cioè il contatore suonerà il 50% delle volte. E' davvero così? No: ciò che accade è che *nel terzo esperimento accade che il contatore suonerà il 100% delle volte e non solo la metà dei casi!* Il formalismo ci aiuta a capire questo aspetto. Infatti calcoliamo la probabilità che il fotone in ingresso superi il test di polarizzazione finale a 45°. Tenendo conto dell'equazione (47)

$$P = |\langle 45^\circ|+\rangle|^2 = 1 \quad (48)$$

La fisica quantistica fa le previsioni corrette! Ne concludiamo che l'interpretazione meramente statistica ed epistemologica non regge e il fotone davvero non si trova a possedere in modo oggettivo lo stato V ed O, come nel caso delle mo-

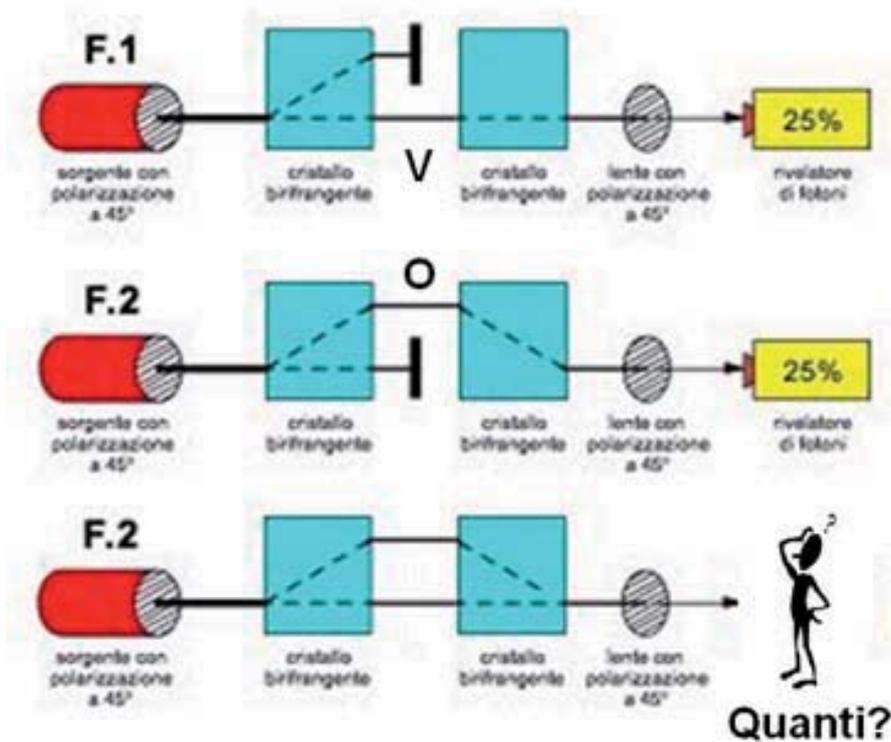


Figura 34: Successione di esperimenti con singoli fotoni attraverso due cristalli birifrangenti. I fotoni incidenti sono polarizzati a 45° . In coda all'apparato vi è un filtro polarizzato a 45° e un contatore di fotoni. Nel primo esperimento il fascio superiore è bloccato e vengono conteggiati il 25% dei fotoni incidenti. Nel secondo è bloccato il fascio inferiore e vengono conteggiati il 25% dei fotoni incidenti. Questi due esperimenti si spiegano in modo classico immaginando che il 50% dei fotoni incidenti sono polarizzati oggettivamente V e gli altri O. Nel terzo esperimenti nessun percorso è bloccato. Se le polarizzazioni dei fotoni incidenti fossero oggettivamente una miscela statistica di V e O otterremo che il contatore misura il 50% dei fotoni incidenti. Invece non è così. La risposta alla domanda "quanti?" che lo sperimentatore si pone è nel testo e spiega come mai i "quanti" conteggiati non trovano spiegazione classica in una miscela statistica.

nete. Ma questo implica che non ha nemmeno un percorso oggettivo. L'elemento di realtà *percorso giù* o *percorso su* non è inerente al fotone in questione. E se d'altro canto vogliamo individuare quale percorso il fotone sceglie bloccandone uno ecco che invece di avere il 50% di risultati otterremo solo il 25%, la misura altera i risultati. In effetti possiamo dire che lo stato a 45° è uno stato di interferenza dei fotoni O e V. Ma anche O e V sono stati di interferenza degli stati a 45° e 135° . Solo che queste coppie di stati (O,V) e (45° , 135°) sono, per dirla alla Bohr, complementari. Non possiamo misurare entrambe le coppie di stati senza alterare la misura. Proprio come non possiamo misurare posizione e velocità di una particella quantistica, in quanto proprietà complementari.

Interferenza di cammini

Ora siamo in grado di rispondere alla domanda: come può un fotone, che è un proiettile, fare interferenza, che è la proprietà tipica di un'onda? Schematizzo quello che succede se inviamo on-

de, sassi e fotoni contro una doppia fenditura in figura 35.

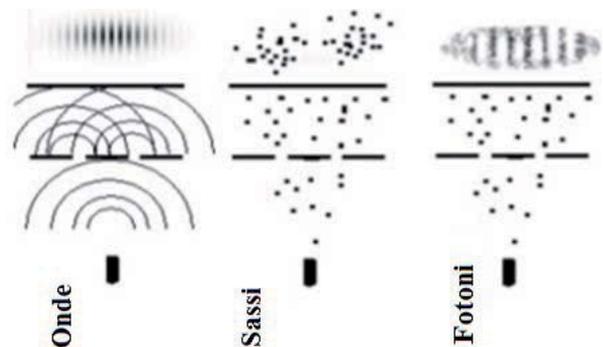


Figura 35: Esperimento di interferenza attraverso due fenditure mediante onde (a sinistra), sassi (al centro) e fotoni (a destra).

Come è possibile che compaia, dunque, la figura di interferenza con sassi quantistici? Prima cerchiamo di calcolare cosa accade se concepiamo la luce come un'onda classica elettromagnetica. Abbiamo detto che il campo elettrico, infatti, è un'utile strumento per fornirci la probabilità di trovare un fotone. In questo senso possiamo usare la propagazione del campo elettrico attra-

verso le due fenditure come fosse una semplice onda trasversale (figura 35 a sinistra). Ricavati i valori dell'intensità finale essa rappresenterà una quantità proporzionale alla probabilità di impatto del fotone sullo schermo. Ecco perché inviando un fotone alla volta esso si comporta casualmente ma poi emerge un "pattern". La meccanica quantistica ci fornisce, mediante il concetto di campo elettrico, la statistica di distribuzione dei fotoni. Nella teoria ondulatoria il fenomeno era facilmente spiegabile come una sovrapposizione dei campi elettrici. In particolare l'intensità I era proporzionale al modulo quadro del campo totale:

$$I \sim |E_1 + E_2|^2 = |E_1|^2 + |E_2|^2 + \{E_1^* E_2 + E_1 E_2^*\} \quad (49)$$

I termini tra parentesi sono i termini di interferenza. Volendo scrivere un'onda come un numero complesso del tipo

$$E_1(x, t) = E_0 e^{i\delta_1} = E_0 e^{i(kx_1 - \omega t)}; \\ E_2(x, t) = E_0 e^{i\delta_2} = E_0 e^{i(kx_2 - \omega t)} \quad (50)$$

dove il numero d'onda è inversamente proporzionale alla lunghezza d'onda ($k=2\pi/\lambda$) e la pulsazione è proporzionale alla frequenza ($\omega=2\pi\nu$), l'equazione (49) diventa

$$I \sim \{1 + \cos(\delta_2 - \delta_1)\} \quad (51)$$

L'intensità della luce diventa nulla se l'argomento del coseno è un multiplo intero dispari di π , ovvero nel caso in cui

$$x_2 - x_1 = m \frac{\lambda}{2} \quad (52)$$

che è la formula (5) trovata nella parte precedente. I minimi si trovano in tutti quei punti le cui posizioni distano dalle fenditure un multiplo di mezza lunghezza d'onda.

Ma, visto che abbiamo due fenditure, possiamo eludere il concetto di campo elettrico ed effettuare il calcolo come un sistema a due stati? Sì. Lo stato corrisponde all'alternativa *il fotone descrive la traiettoria sorgente-fenditura top-punto x (traiettoria T)* oppure quello corrispondente al-

l'alternativa *il fotone descrive la traiettoria sorgente-fenditura bottom-punto x (traiettoria B)*. (Figura 36) Come abbiamo visto per la polarizzazione e per i cristalli birifrangenti, se il fotone ha due alternative di percorso, si pone in una sovrapposizione di percorsi (che non normalizzeremo per semplicità)

$$|\psi\rangle \sim |T\rangle + |B\rangle \quad (53)$$

La probabilità che il fotone giunga in x è allora proporzionale a

$$P = |\langle x|\psi\rangle|^2 \sim |\langle x|T\rangle + \langle x|B\rangle|^2 \quad (54)$$

Notate che questa sovrapposizione tra ampiezze è analoga a quella della catena di uguaglianze vista per i cristalli birifrangenti. Quanto valgono le ampiezze $\langle x|T\rangle$ o $\langle x|B\rangle$? Ebbene l'aspetto ondulatorio che abbiamo nella trattazione del problema col campo elettrico viene qui ripreso ammettendo che i due stati sono diversi tra loro solo a causa di una fase in quanto i cammini sono equiprobabili. Questo aspetto potrebbe sembrare alquanto strano in quanto i cammini più probabili dovrebbero essere quelli rettilinei che collegano la sorgente a ciascuna fenditura come rappresentati e invece gli altri dovrebbero essere trascurabili, proprio come per i sassi (figura 35).

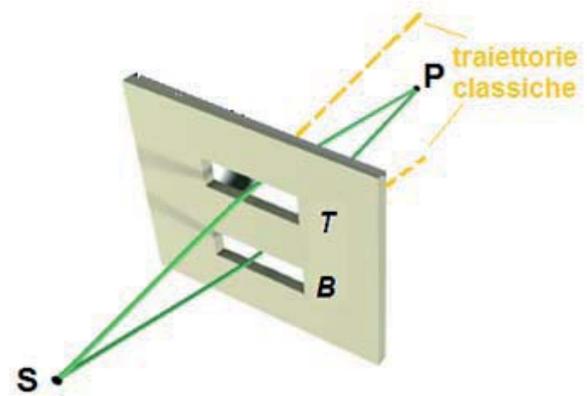


Figura 36: Esperimento di interferenza attraverso due fenditure mediante l'idea dei percorsi in sovrapposizione. Nel caso di traiettorie classiche i fotoni seguirebbero le leggi dell'ottica geometrica e il punto P non potrebbe mai essere raggiunto.

Tuttavia questo sarebbe contrario a ciò che si sperimenta, ovvero l'emergere del pattern di interferenza. Se ammettiamo invece che *tutti i cam-*

mini sono equiprobabili allora lo saranno anche quelli che collegano la sorgente a un punto P (in posizione x) qualsiasi dello schermo attraverso le fenditure (figura 36). In linea di principio, classicamente, nessuna "pallina" potrebbe abbandonare S, passare per B e deviare verso P. Classicamente se la luce non fosse un'onda ma fosse fatta di particelle, nella parte di schermo corrispondente allo spazio tra le due fenditure, dovremmo vedere buio. Invece vediamo luce. Ecco perché abbiamo dovuto supporre che la luce fosse un'onda e non costituita da proiettili. Ora invece stiamo asserendo che essa è fatta di proiettili, ma questi sono alquanto strani: si pongono in una sovrapposizione di traiettorie e riescono ad arrivare in P. Tutte le traiettorie sono semplicemente sfasate di un fattore complesso che non altera la probabilità del singolo cammino, per cui

$$\langle x|T\rangle = e^{i\delta_T}; \langle x|B\rangle = e^{i\delta_B} \quad (55)$$

Ora questo implica che la probabilità P che il fotone impatti sul punto P passando per uno o per l'altro dei cammini sarà

$$P = |\langle x|\psi\rangle|^2 \sim |e^{i\delta_T} + e^{i\delta_B}|^2 \sim 1 + \cos(\delta_B - \delta_T) \quad (56)$$

Che fornisce la figura di interferenza. Cosa esprime fisicamente la differenza di fase $\delta_B - \delta_T$? Ovviamente sarà $k(x_T - x_B)$ cioè il numero d'onda k moltiplicato la differenza di cammino ottico dalle fenditure al punto P sullo schermo, proprio come nell'equazione (51). Abbiamo provato che anche l'interferenza può essere spiegata usando il concetto di fotone e di sovrapposizione degli stati. Ma se gli stati sono cammini cosa significa concretamente sovrapposizione dei cammini? Come spiegare il fatto che un fotone, che non si può dividere, sente che entrambe le fenditure sono aperte? Come interpretare cioè la sovrapposizione di stati di un fotone?

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|\text{alternativa A}\rangle + |\text{alternativa B}\rangle) \quad (57)$$

Ebbene esporrò qui brevemente le posizioni tipiche, eccetto quella della decoerenza che

discuterò alla fine del lavoro.

Non porti troppe domande: l'interpretazione di Copenaghen

In questa interpretazione la funzione d'onda Ψ descrive in modo completo lo stato di un oggetto quantistico. La Ψ è un ente matematico che vive in uno spazio astratto e fornisce la probabilità di un evento. La sovrapposizione di alternative indica che il sistema si trova in una sorta di realtà sospesa, nebulosa, che evolve seguendo leggi deterministiche (come l'equazione di Shrodinger) ma che collassa in un elemento di realtà classica in modo genuinamente casuale quando il sistema viene osservato o misurato da un apparato classico. Necessariamente il mondo sarà pertanto diviso in oggetto osservato quantistico e apparato di misura classico e solo alcuni elementi di realtà sono posseduti oggettivamente da un sistema quantistico, altri elementi di realtà sono creati mediante le misure con apparati classici. Ci sono poi elementi di realtà complementari. Impostare un esperimento per misurare un aspetto della realtà non permette di mostrare quello ad esso complementare, se non alterando lo stato del sistema in modo irreversibile. Ad esempio o vediamo frange di interferenza (proprietà ondulatorie) o vediamo due strisce in corrispondenza delle fenditure (proprietà corpuscolari) a seconda se non ci chiediamo o se ci chiediamo da quale fenditura è passato il fotone. Le due domande implicano apparati sperimentali mutuamente esclusivi (principio di complementarità di Bohr) e non è possibile rispondere a tali questioni simultaneamente. La situazione può essere rappresentata dal drago nebuloso di Wheeler che esprime l'intrinseco velo quantistico che parrebbe emergere dalla interpretazione di Copenaghen prima della misura.

Gli svantaggi della interpretazione di Copenaghen è che non risponde a due domande

1. come è possibile avere una descrizione unitaria del mondo se la teoria divide la realtà in un sistema quantistico (il fotone) e un sistema classico (il polarizzatore e il contatore)?
2. Se un oggetto macroscopico è fatto esso stesso da atomi non dovremmo usare la

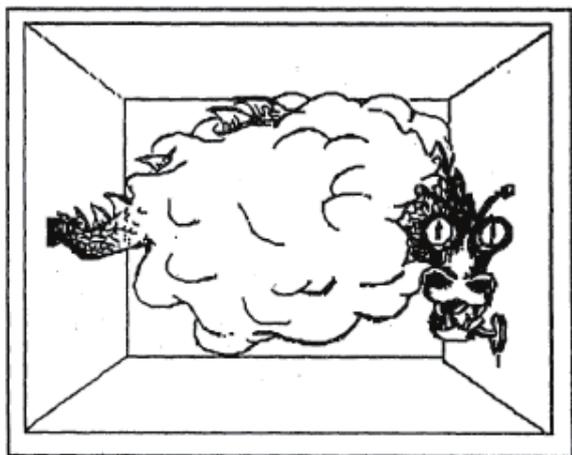


Figura 37: *Il drago nebuloso di Wheeler. Nella interpretazione di Copenaghen la realtà è sospesa prima della misura in stati non identificabili con situazioni classiche oggettive.*

stessa descrizione quantistica usata per il fotone? Ma allora come mai non vediamo sovrapposizioni di stati per oggetti macroscopici?

Circa la prima domanda nei fatti l'interpretazione di Copenaghen fa leva su due descrizioni della realtà: una deterministica (l'evoluzione dello stato mediante l'equazione di Shroedinger) e l'altra basata sul collasso casuale dello stato causato dalla interazione del sistema quantistico con l'apparato di misura classico. Punto. Per buona pace di chi vuole adoperare la fisica quantistica anche in cosmologia per descrivere la funzione d'onda di tutto l'universo al momento del Big Bang. In tal caso chi è l'apparato di misura? Alla seconda domanda in genere si risponde in modo molto ambiguo. Ad esempio dicendo che i corpi classici hanno una lunghezza d'onda di De Broglie molto piccola data la massa per cui tutti gli effetti quantistici sono piccoli. Tuttavia la teoria non fornisce alcun modello di transizione dal microscopico al macroscopico in situazioni più delicate come l'entanglement tra un oggetto micro e un oggetto macro, come nel caso del famoso gatto di Shroedinger.

Quante storie! I cammini di Feynman

Questa dei cammini simultanei più che essere una visione diversa da quella di Copenaghen è

un diverso punto di vista che val la pena però affrontare. Piuttosto che interpretare la sovrapposizione come una sospensione della realtà in uno stato nebuloso di elementi di realtà (non ha senso chiedersi se il fotone ha uno stato A o B prima della misura al contrario della moneta) Feynman interpreta la sovrapposizione come storie alternative che hanno luogo simultaneamente (il fotone esiste in entrambi gli stati A e B e fa le due cose simultaneamente senza dividersi!). Fu proprio il principio di minima azione (o di Fermat) a spingere Feynman a concepire questo punto di vista che trova riscontro soprattutto nella elettrodinamica quantistica, in cui ad ogni ampiezza di probabilità di un evento, calcolata con metodi matematici rigorosi ma complessi, si associa un grafico che ridona una minima visualizzazione geometrica del processo. Un esempio classico: supponiamo di voler calcolare la probabilità che due elettroni in moto l'uno verso l'altro, interagendo per via della carica, siano deflessi all'indietro. Ebbene in una dinamica spaziotemporale (spazio 1D+tempo) il processo si sviluppa in un piano e gli elettroni si scambiano un fotone.

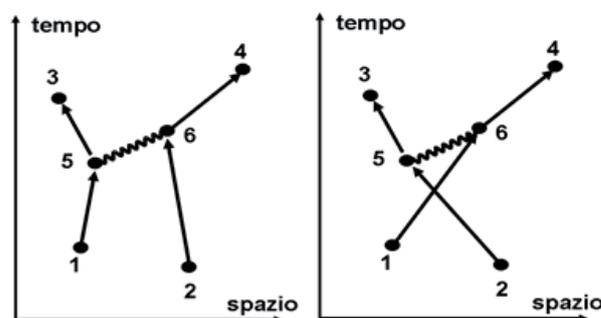


Figura 38: *Processo di interazione tra due elettroni secondo Feynman: i due elettroni 1 e 2 interagiscono scambiandosi un fotone (la linea ondulata tra 5 e 6) e subiscono un rinculo per portarsi in 3 e 4. Data l'indistinguibilità del processo questa storia non è distinguibile nei risultati da quella in cui i due elettroni 1 e 2 si scambiano (a destra). Entrambe le storie vanno conteggiate come reali.*

Il fotone è disegnato inclinato perché segue il cono-luce (vedi seconda parte del lavoro). Tuttavia anche il grafico a destra è consentito e contempla la situazione in cui i due elettroni sono scambiati. A causa delle proprietà fermioniche degli elettroni (ovvero due elettroni non possono avere gli stessi stati quantici) che seguono il principio di esclusione di Pauli e della loro in-

distinguibilità non possiamo a priori capire se ciò che è accaduto è il processo di figura 38 sinistra o 38 destra. Allora, per Feynman, entrambe le storie sono vissute dagli elettroni, e ciascuna storia rappresenta l'ampiezza di probabilità del processo, cioè rappresenta un numero. Poiché le ampiezze e non le probabilità si sommano in MQ, anche qui sommeremo i grafici e ne faremo il modulo quadro per ottenere la probabilità. La potenza di questo metodo è che se da un lato rappresenta un interessante ed efficace strumento matematico per calcolare le probabilità quantistiche d'altro canto ci riconsegna una visualizzazione dei processi, sebbene dobbiamo considerarli tutti e sommarli. Difficilmente possiamo dire che la possibilità è sospesa in questo quadro finché non la si osserva. Piuttosto essa è una sovrapposizione di tutte quante queste realtà. Purtroppo le storie di figura 38 non sono gli unici processi da tenere in conto se vogliamo descrivere in modo preciso i risultati sperimentali. Il processo in questione è molto più complesso e se vogliamo trovare un accordo molto preciso tra teoria ed esperimento bisogna sommare tutti i possibili eventi, come quelli rappresentati sul famoso furgone con cui Feynman girava con la sua famiglia, in figura 39.

Dobbiamo cioè considerare processi, che per far convergere i conti, anche esotici e poco probabili, e che richiedono alla fine dei conti un intervento esterno di chi calcola per eliminare gli infiniti che spesso vengono fuori dalle somme (procedura di rinormalizzazione). Insomma se saltano degli infiniti metterli sotto il tappeto e tacere. Ovviamente se la somma converge nessun problema. Ma se diverge sono guai perché l'intero risultato darà una probabilità infinita, che non ha senso. Per convergere è necessario che il parametro di accoppiamento (α nel nostro esempio, che vale $1/137$ ed è chiamata costante fine) della interazione tra particella "materiale" (l'elettrone nel nostro caso) e la particella di "scambio" (il fotone nel nostro caso) sia minore dell'unità. Il termine di accoppiamento è moltiplicato per l'ampiezza un numero di volte pari a quanti vertici ci sono nel grafico. Negli eventi del primo ordine mostrati in figura avremo pertanto che ogni storia sarà moltiplicata per α^2 . Negli eventi di ordine superiore sarà α^N . La serie converge perché il parametro di accoppiamento è

inferiore a uno e tutto va bene, a tal punto che l'elettrodinamica quantistica è la teoria più precisa mai inventata dall'uomo. Ma in altri modelli come nella Cromodinamica Quantistica (QCD) il termine di accoppiamento della interazione è superiore a uno e la somma va troncata in modo artificiale per evitare esplosioni! Non è un quadro confortante per la fisica teorica. Faccio notare, a costo di essere pedante, che in questo approccio la probabilità non è stata cancellata, siamo ancora in pieno regime di interpretazione standard ma con un punto di vista diverso e degno di un premio Nobel, come fu conferito giustamente a Feynman per aver risolto gran parte dei problemi legati alla quantizzazione dell'elettromagnetismo. Un altro aspetto, a mio avviso molto interessante, è che il vuoto assume uno status di entità dinamica, perché ha un effetto sui processi molto importante. Esso non è assenza di processi ma, al contrario, pullula di eventi e fluttuazioni anche molto violente. Un esempio si ha quando disponiamo un atomo nello stato eccitato stazionario. La MQ standard e classica (cioè non relativistica) prevede che l'elettrone resterà per sempre in quello stato. Eppure esso dopo un po' ricade nello stato fondamentale emettendo un fotone in modo spontaneo. Come mai? In realtà l'elettrodinamica quantistica (QED) ha mostrato che l'emissione non avviene in modo spontaneo ma indotto dalle fluttuazioni del vuoto che generano quanti di campo elettromagnetico (fotoni virtuali) i quali stimolano l'elettrone a decadere. Questi quanti vivono per un tempo molto breve dato dal principio di indeterminazione di Heisenberg $\Delta t \sim \hbar/\Delta E$ dove ΔE rappresenta l'energia presa in prestito dal vuoto (violando la conservazione dell'energia per quel tempo minuscolo) per creare le fluttuazioni. Restano sospese le problematiche già viste nella interpretazione di Copenaghen, da cui tutto sommato questa visione non si discosta, ovvero come mai non si vedono sovrapposizioni di storie nel mondo macroscopico.

Il rasoio di Occam? Mai sentito!

Nel 1957 Hugh Everett propose una diversa interpretazione detta a "molti mondi" in cui tenta di risolvere definitivamente la dicotomia tra il collasso genuinamente casuale dello stato dov-

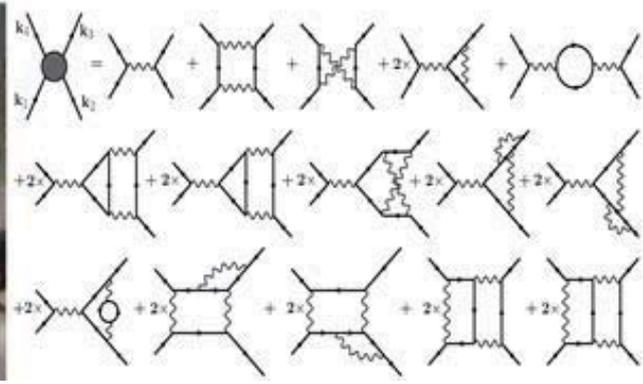


Figura 39: Furgone di Feynman dove sono rappresentati alcune delle storie possibili della interazione tra due fermioni come due elettroni. A destra storie di ordine inferiore ma da conteggiare per avere un riscontro con l'esperimento.

to alla misura e la sua evoluzione deterministica dovuta alla equazione di Shroedinger. La soluzione è che il *collasso casuale* è solo illusorio. Infatti il sistema evolve in modo deterministico e in sovrapposizione di stati. Lo stato A vive in un universo A e lo stato B in un universo B, e questi interagiscono tra loro durante la sovrapposizione. Quando avviene la misura accade che tutto l'universo che interagisce con lo stato quantico si divide esso stesso in due copie, una in cui la particella è collassata in A e una in cui essa è collassata in B. In ciascun universo ci sarà anche l'osservatore che in un universo misurerà lo stato A e nell'altro B. L'osservatore in A perde consapevolezza della sua copia in B. In questo modo il determinismo è ancora vivo e vegeto e la casualità è solo illusoria perché perdiamo conoscenza dell'altro universo. A ben rifletterci quindi mentre in una interpretazione del tutto classica il collasso della funzione d'onda vuol dire prendere conoscenza dello stato reale del sistema, nella interpretazione di Copenaghen il collasso è una oggettivazione dello stato del sistema, nella interpretazione a molti mondi il collasso è altro se non la perdita di informazione perché ci sdoppiamo e noi abbiamo coscienza solo di questo universo A e non dell'altro, come la nostra copia del resto ha consapevolezza solo dell'universo B e non dell'universo A.

Questa teoria ha il vantaggio di rimuovere il problema dell'osservatore e risponde alla domanda: se è necessario un osservatore per oggettivare un sistema quantistico chi osservava l'universo alla sua nascita? I vantaggi sono indubbi per i cosmologi, perché si elude il problema della misura in modo semplice, ma ha come svantag-

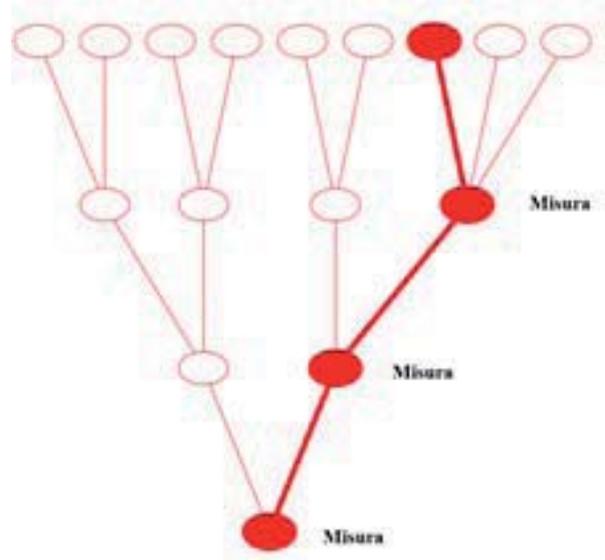


Figura 40: Interpretazione a molti mondi di Everett. Noi siamo consci solo in una delle storie parallele sviluppatesi a seguito delle misure, ma ci sono copie di noi, altrettanto consci, in tutti gli altri mondi. La funzione d'onda descrive il tutto indiviso della pluralità di mondi e il collasso casuale è dato solo dal fatto che le nostre esperienze coscienti vivono in una "traiettoria" (quella rossa) e non abbiamo esperienza delle altre situazioni (come le nostre copie non hanno esperienza cosciente delle nostre percezioni in questo mondo).

gio il dover postulare infinite copie di sé stessi. Per uno sperimentale come il sottoscritto però non si può non pensare al povero Occam e alla sua frase: "non considerare la pluralità se non è necessario". Everett risponderebbe che è stato necessario. Peccato che è una tesi impossibile da dimostrare oggi. Chissà, magari ci riusciremo con un computer quantistico.

L'assenza di una prova non è la prova dell'assenza: le teorie a variabili nascoste di Bohm

Di fronte ad un panorama come quello descritto molti fisici gettano la spugna e sostengono che alla fine la fisica debba solo fare previsioni e non modelli di realtà. In fondo ciò che si richiede in genere ad una teoria fisica è che funzioni e la fisica quantistica funziona da Dio (che, parafrasando il famoso motto di Einstein, parrebbe giocare ai dadi) indipendentemente dalle interpretazioni. Il formalismo è neutro rispetto alle interpretazioni e al più ogni interpretazione ci fa vedere aspetti che l'altra non ci fa notare. Ma senza interpretare nulla i conti tornano lo stesso. Insomma, senza fare qualunquismo, è un po' come le politiche economiche impostate dalla destra e dalla sinistra oggi. Indipendentemente dal colore il risultato è lo stesso in quanto è a livello globale che si prendono le vere decisioni e tutto funziona come deve funzionare con questo modello di libero mercato in mente (giudicate voi se bene o male) che è neutro alle interpretazioni circa la natura del capitale. Questa visione è la sindrome del "zitto e calcola e non chiederti altro". Non è la prima volta che accade. In fondo non era così anche la teoria dell'azione a distanza di Newton in cui il suo "non fingo ipotesi" esprimeva un disagio dello stesso padre della meccanica gravitazionale attorno al problema della trasmissione della forza attraverso il vuoto? Eppure la teoria funzionava benissimo nel fare previsioni a tal punto che grazie ad essa fu scoperto il pianeta Nettuno da delle perturbazioni del moto orbitale di Urano. Alcuni tuttavia non si accontentano del zitto e calcola. E direi fortunatamente, perché altrimenti non avremmo oggi la teoria della relatività generale, che spiega il *meccanismo di trasmissione* della forza di gravità come teoria geometrica di campo. La fisica quantistica appartiene nei fatti a quelle teorie che non spiegano i "meccanismi" ma fornisce le regole generali. Ma se ciò fosse vero allora dovremmo rinunciare a dare una spiegazione razionale del mondo, in quanto se un fotone collassa in uno stato di polarizzazione definito lo fa non a seguito di un meccanismo ma in virtù di una "decisione" che non ha cause. Il fotone collassa perché collassa così come io al mattino vado sempre a bere

un cappuccino a un bar A e un giorno appresso al bar B, a caso. Chi mi osservasse dall'esterno direbbe: domani prevedo che quel tizio andrà a prendere il cappuccino al bar A col 50% di probabilità. Perché? Un fisico quantistico ortodosso direbbe: perché così ha deciso. Un fisico però che si interroga sui meccanismi delle cose crede che ci sia dell'altro dietro. Infatti mentre scrivo queste righe sono andato al bar A perché pioveva e ho trovato più conveniente andare in quel bar e non nell'altro per non bagnarmi. Vi è cioè una variabile nascosta (odio bagnarmi durante la pioggia) che spiega come mai ho preso quella decisione. A volte questa variabile non dipende da agenti esterni e misurabili ma dal mio stato d'animo inaccessibile perfino al sottoscritto. Ma questo vuol dire che non vi è una causa nella decisione, per quanto inconscia? Ebbene la posizione di un fisico quantistico ortodosso è proprio questa: non vi è una causa al collasso nello stato A o B. E il contrario di assenze di CAUSA è il CASO, parole simili ma profondamente diverse. Il punto è che da quando la scienza è nata noi interpretiamo gli eventi in termini di causa-effetto. Introdurre il caso ha per molti il sapore di una non spiegazione. E David Bohm era uno di questi. Nel 1952 Bohm scrisse un articolo in due parti in cui iniziò a porsi la questione se non fosse "il caso" di iniziare a pensare che il problema risiedesse nella interpretazione probabilistica della funzione d'onda e non nella realtà ontologica dei quanti (di luce o di materia). Nel suo articolo Bohm esplicita il suo pensiero in poche e semplici mosse: partendo dall'equazione di Shroedinger

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi + U(x)\psi \quad (58)$$

ed esprimendo la funzione d'onda mediante un'ampiezza R e una fase S

$$\psi = \langle x | \psi \rangle = R e^{i\frac{S}{\hbar}} \quad (59)$$

si ottiene una coppia di equazioni reali

$$\frac{\partial R}{\partial t} = -\frac{1}{2m} [R \nabla^2 S + 2 \nabla R \bullet \nabla S] \quad (60)$$

$$-\frac{\partial S}{\partial t} = \frac{(\nabla S)^2}{2m} + U(x) - \frac{\hbar^2}{2m} \frac{\nabla^2 R}{R} \quad (61)$$

L'equazione (61) può essere vista come l'equazione di Hamilton Jacobi (HJ) di un sistema classico quando \hbar va a zero o la massa dell'oggetto diventa grande (cioè per strutture macroscopiche). Non vi è quindi una separazione netta tra realtà quantistica e classica. Primo vantaggio. Nella equazione di HJ se si associa alla quantità ∇S la velocità v ($v = \nabla S/m$) della particella, il primo termine a secondo membro non è altro che l'energia cinetica. Per questo motivo la prima equazione, stando alla interpretazione classica in cui la probabilità P di un processo è data dal quadrato dell'ampiezza dell'onda ($P = R^2$) può essere scritta in un altro modo ovvero come un'equazione di continuità per la densità di probabilità dell'insieme di particelle dell'esperimento. Riscriviamo pertanto la coppia di equazioni

$$\frac{\partial P}{\partial t} + \nabla \bullet (Pv) = 0 \quad (62)$$

$$\frac{\partial S}{\partial t} + \frac{(\nabla S)^2}{2m} + U(x) - \frac{\hbar^2}{2m} \frac{\nabla^2 R}{R} = 0 \quad (63)$$

La seconda ci indica che la particella quantistica è sottoposta a due tipi di potenziale: quello classico U e un potenziale quantistico che chiameremo Q

$$Q = -\frac{\hbar^2}{2m} \frac{\nabla^2 R}{R} \quad (64)$$

che dipenderà dalla scala del fenomeno, dalla massa del sistema e dalla ampiezza della funzione d'onda. La seconda equazione è analoga a quanto visto nella prima parte di questo lavoro a proposito della equazione *iconale* modificata per tener conto degli effetti di diffrazione e interferenza della luce e che per comodità riportiamo qui sotto per un onda elettromagnetica in cui la fase S è indipendente dal tempo

$$|\nabla S|^2 = n^2 + \left(\frac{\lambda}{2\pi}\right)^2 \frac{\nabla^2 E_0}{E_0} \quad (65)$$

In questo caso l'indice di rifrazione prende il posto della energia potenziale U , l'ampiezza del campo elettrico il posto dell'ampiezza della funzione d'onda (in accordo con la nostra interpretazione attuale del campo elettrico), e la lunghezza d'onda ha a che vedere con la massa della particella. In sostanza la particella si muove sotto un

campo di forze non dovuto solo all'interazione della particella con un altro sistema "materiale" (o nel caso del fotone in mezzi con un dato indice di rifrazione), ma è determinato anche da questo potenziale quantico che dipende dalla funzione d'onda o dal campo elettrico per un elettrone e per un fotone rispettivamente. In questa interpretazione la funzione d'onda rappresenta un vero e proprio campo e non una interpretazione statistica del sistema. Un tale potenziale non è osservabile su grandi scale in quanto la costante di Planck è troppo piccola o la massa del sistema troppo grande. Un esperimento a doppia fenditura viene perciò interpretato da Bohm come un processo in cui la particella ha una esistenza in un punto oggettivo dello spazio e del tempo e viene "guidata" dall'onda. La particella non è incastonata come propose De Broglie ma è la manifestazione dell'onda stessa.

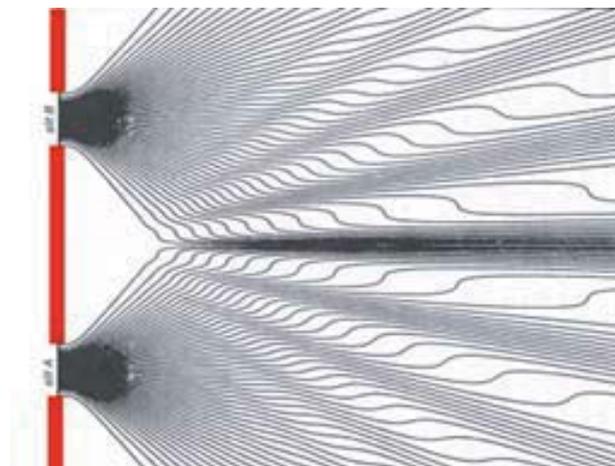


Figura 41: *Onda pilota di Bohm a seguito della interferenza da due fenditure. La particella segue, come manifestazione dell'onda stessa, una traiettoria di tipo oggettivo.*

In questo caso il fotone, una manifestazione del campo elettrico, è pilotato da questo lungo una delle possibili traiettorie, come un serfista su un'onda marina. A seconda della sua posizione il fotone può prendere una direzione o l'altra e impattare sullo schermo nelle regioni dove l'onda ha la massima ampiezza. Il caso non esiste ma è solo apparente in quanto noi non conosciamo esattamente da quale punto della fenditura il fotone è passato, né da quale fenditura. Se proviamo a chiederci infatti quale fessura il fotone ha attraversato dovremmo inserire un detector subito dopo una delle due fenditure ma questo

altera la modalità di diffusione dell'onda producendo le due tipiche strisce luminose sullo schermo. Perché un modello così semplice non è stato preso in considerazione in modo serio? Va detto che probabilmente la ragione risiede nel fatto che postulare variabili inaccessibili (come la posizione del fotone) ad alcuni sembra inutile. Ma è anche vero che in altri campi questo si fa senza problemi, come postulare l'esistenza dei quark che non sono direttamente accessibili a causa della proprietà del confinamento. E' questa una proprietà oggettiva o piuttosto un modo per risolvere il problema del perché i quark non sono stati mai osservati se non nei loro effetti? Se è così anche le variabili nascoste di posizione di Bohm hanno degli effetti: l'apparente casualità di un processo di misura. Va detto però, come vedremo ora, che ci sono altri aspetti che complicano la partita: i fenomeni di entanglement da un lato e dall'altra il fatto che tali variabili sono contestuali, cioè dipendono dalle variabili del resto dell'universo.

Parte IV: Olismo quantistico

Il dado è tratto

L'entanglement è uno dei fenomeni più distintivi dell'approccio quantistico alla luce (e a tutte le altre particelle fondamentali). Prima di procedere nella disamina di questo comportamento tutto quantistico e che non trova alcun corrispettivo classico, cerchiamo di rispondere ad una domanda: dopo aver capito il comportamento di una singola particella quantistica come un singolo fotone, come descrivere un sistema composto? Questo è un aspetto importantissimo e che pilota la nostra ricerca nel tentativo di rispondere alle questioni aperte sugli aspetti macroscopici non quantistici della realtà. Supponiamo di prendere una moneta ed un dado. La moneta "vive" in uno spazio a due stati distinti, se siamo interessati allo stato testa o croce, mentre il dado vive in uno spazio a sei stati, se siamo interessati al numero che compare sulla sua faccia. Lo spazio degli stati della moneta è uno spazio di

Hilbert bidimensionale mentre quello del dado è uno spazio di Hilbert esadimensionale. Possiamo chiamare gli stati della moneta $|T\rangle$ e $|C\rangle$ come prima e quelli dei dadi $|1\rangle; |2\rangle; \dots |6\rangle$. Il sistema combinato si descrive mediante lo spazio prodotto. Siccome possiamo avere dodici combinazioni possibili, ad esempio $|T; 3\rangle$ indicando con T lo stato della moneta e con 3 lo stato del dado, lo spazio in cui vive il sistema combinato *moneta+dado* è dodecadimensionale. Questo è lo stato prodotto discusso nel box 3. Ora eseguiremo tre esperimenti con una moneta e un dado.

Esperimento 1

Uno sperimentatore di nome Charlie inserisce in una scatola la moneta e in un'altra il dado per poi consegnarli in due scatole separate a due colleghi di nome Alice e Bob. Poi questi partono per due destinazioni molto lontane, ad esempio possiamo immaginare che Bob vada su una colonia vicino a Saturno mentre Alice resta sulla Terra. Prima di partire i due personaggi hanno sincronizzato il loro orologi e si sono accordati che alle 12 ore terrestri dell'anno 2025 (il viaggio verso Saturno è lungo come testimonia il film di Chris Nolan *Interstellar*) apriranno le scatole. Cosa scopriranno? Se Alice trova una moneta nella sua scatola Bob troverà un dado e viceversa. Cosa c'è di strano in questo? Nulla! Infatti i sistemi sono separati. Ma attenzione: il fatto che Alice prima di aprire la scatola non sapesse cosa ci fosse dentro e né tantomeno cosa Bob avesse nella scatola, non vuol dire che dentro la scatola non ci sia una moneta o un dado prima di aprirla. Lo abbiamo già visto nel caso della moneta: in fisica classica vale il principio RC. L'informazione che dentro la scatola di Alice ci fosse una moneta e dentro la scatola di Bob un dado era già contenuta nel cono luce passato di entrambi, quando Alice e Bob hanno ricevuto la scatola da Charlie. Semplicemente la loro ignoranza sullo stato interno della scatole è svanita immediatamente appena aperta la scatola, ma non c'è nulla di metafisico in ciò.

Il sistema prima dell'apertura delle scatole è descrivibile in questo modo

Esperimento 2

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|Moneta; Dado\rangle + |Dado; Moneta\rangle) \quad (66)$$

Questo vuol dire che vi è una probabilità che, quando le scatole vengono aperte, Alice trovi una moneta e Bob un dado e viceversa. Nel primo caso il sistema collassa in

$$|\psi\rangle \Rightarrow |Moneta; Dado\rangle \quad (67)$$

dove la parte sinistra dello stato collassato riguarda cosa trova Alice e la parte destra cosa trova Bob. Questo collasso non è come quello della polarizzazione del fotone, ma come quello della moneta singola ed esprime solo il fatto che aprendo la scatola Alice e Bob rileveranno istantaneamente una informazione. In altre parole se prepariamo l'esperimento N volte avremo sempre che, quando Alice troverà una moneta Bob troverà un dado e viceversa. Data la casualità presunta nella distribuzione delle scatole ci si aspetta che sia Alice che Bob troveranno la moneta o il dado il 50% dei casi, come indicato dal formalismo

$$|\langle Moneta; Dado | \psi \rangle|^2 = \frac{1}{2} \quad (68)$$

Supponiamo di associare alla moneta il valore +1 e al dado il valore -1. Queste in effetti sono delle misure che indicheremo con σ^A nel caso di Alice e σ^B nel caso di Bob. La media dei risultati sarà ovviamente 0 per entrambi, in quanto Alice e Bob troveranno casualmente a volte una moneta (+1) e volte un dado (-1) cioè

$$\langle \sigma^A \rangle = \langle \sigma^B \rangle = 0 \quad (69)$$

sicché la somma diviso il numero N di prove (la media) darà zero. Tuttavia se proviamo a calcolare la media del prodotto delle misure otterremo

$$\langle \sigma^A \sigma^B \rangle = -1 \quad (70)$$

cioè le misure sono *anticorrelate*. Questo significa che ogni volta che Alice trova un dado ($\sigma^A = -1$) Bob troverà una moneta ($\sigma^B = +1$) sicché il prodotto è -1 e il valor medio dei prodotti calcolato su tutte le coppie di scatole sarà parimenti -1. Ma, lo ripeto, nulla da eccepire e nulla di strano.

Supponiamo ora di cambiare esperimento e di interessarci a rispondere alla domanda: quale faccia Alice e Bob troveranno esposta nella scatola? Supponiamo che uno dei due riceva sempre una moneta (Alice) e l'altro sempre il dado (Bob), e che ora siamo interessati a capire quale faccia esposta essi troveranno, se una testa o una croce nel caso di Alice, oppure un numero pari (p) o dispari (d) nel caso di Bob. Immaginiamo anche di formalizzare il test come prima. Se Alice trova testa associa un +1, se trova croce vi associa -1. Invece se Bob trova una faccia dispari vi associa un +1 e se trova una faccia pari vi associa -1. La probabilità che Alice trovi testa sarà del 50% (una faccia su due possibili), così come la probabilità che Bob trovi un numero pari è anch'essa del 50% (tre facce pari su sei facce possibili). La probabilità che Alice trovi testa e Bob pari simultaneamente sarà dato dal prodotto delle probabilità, cioè il 25%. Formalmente lo stato sarà il seguente

$$|\psi\rangle = \frac{1}{2} (|T; p\rangle + |T; d\rangle + |C; p\rangle + |C; d\rangle) \quad (71)$$

Facciamo un esempio e supponiamo di calcolare la probabilità che Alice trovi Testa e Bob un numero pari. La probabilità sarà data dal quadrato del coefficiente davanti allo stato $|T; p\rangle$ che è $\frac{1}{4}$, ovvero 25% come ci aspettavamo. Cosa ci aspettiamo che accada alla media delle misure nel caso in cui Charlie abbia distribuito ad Alice e Bob N coppie di pacchi? Ci aspettiamo che, se la preparazione dei pacchi da parte di Charlie è avvenuta in modo casuale, troveremo ancora una volta

$$\langle \sigma^A \rangle = \langle \sigma^B \rangle = 0 \quad (72)$$

(fate la media dei valori casuali +1 e -1 ottenuti nelle singole misurazioni e troverete zero). Ma cosa accadrà alla media dei prodotti? Cioè cosa accade se confrontiamo le misure per ciascuna coppia? Poiché ora, contrariamente a prima, sia Alice che Bob troveranno casualmente +1 e -1 nelle loro rispettive misure, il prodotto potrà dare sia +1 che -1 (mentre prima forniva sempre -1) e il valor medio sarà nullo, ovvero

$$\langle \sigma^A \sigma^B \rangle = 0 \quad (73)$$

Le misure in questo caso sono assolutamente *scorrelate*. Anche qui nulla di strano.

Esperimento 3

Stesso esperimento del 2 ma ora supponiamo che, per qualche strana ragione, ogni volta che Alice trova testa (+1) Bob trovi *sempre* un numero pari (-1) mentre ogni volta che Alice trova croce (-1) Bob trova *sempre* dispari (+1). Il sistema sarà formalmente descritto da

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|T; p\rangle + |C; d\rangle) \quad (74)$$

La media delle misure per ciascuno sarà ancora nulla, ovvero

$$\langle \sigma^A \rangle = \langle \sigma^B \rangle = 0 \quad (75)$$

ma il prodotto medio delle loro misure non sarà più nullo in quanto le misure mostreranno un schema di anti-correlazione.

$$\langle \sigma^A \sigma^B \rangle = -1 \quad (76)$$

Cosa penseremmo? Come minimo che il gioco è truccato o stiamo assistendo a pura magia. E' ovvio infatti che abbiamo solo tre possibili spiegazioni:

1. *E' magia mentale*: Alice comunica a Bob il risultato telepaticamente e istantaneamente e Bob altera il suo risultato con qualche meccanismo non appena viene a conoscenza (telepaticamente) del risultato di Alice.
2. *E' fisica quantistica*: gli stati della moneta e del dado non esistono prima che le scatole vengano aperte. Non appena Alice apre la scatola la moneta "collassa" nello stato testa e il dado di Bob nello stato "pari". La realtà non è definita prima della misura.
3. *E' frode classica* (o meglio è la classica frode): Charlie ha barato e ha preparato la moneta e il dado già a monte.

Quale alternativa scegliamo? Dipende. Ebbene, vi sorprenderà sapere che la gente a volte opta per la prima ipotesi (basta andare in libreria al reparto pseudoscienza, e ahimè spesso anche al reparto scienza, per trovare libri a riguardo). Se

apparteniamo pertanto alla schiera di coloro che credono nel paranormale o nei *poteri telepatici* del cervello opteremo per la prima. La seconda ipotesi invece richiederebbe che la moneta e il dado non hanno una realtà oggettiva prima che Alice o Bob aprano il pacco. Se accettiamo questa ipotesi apparteniamo alla schiera di coloro che sostengono che noi creiamo la realtà classica aprendo le scatole. I più ragionevoli tra noi tuttavia opteranno per la terza ipotesi. Per quale motivo? Perché sotto sotto crediamo nella oggettività del mondo reale (lo stato della moneta "testa" e il numero pari "4" del dado erano presenti nelle rispettive scatole già prima di aprirle). La realtà segue il principio RC e non il principio della "magia" quantistica. Ma non basta: infatti non crediamo alla telepatia, dunque non crediamo che l'informazione possa attraversare i cervelli e soprattutto che lo faccia all'istante (infatti ricordo che Alice e Bob si trovano a 45 minuti luce l'uno dall'altro). Se crediamo che l'informazione non possa viaggiare più velocemente della luce, allora crediamo nel principio di località di Einstein (un evento ne può influenzare un altro che si trova nel suo cono-luce futuro). Del resto se vediamo una persona che si arrabbia e mostra una faccia contrita in realtà stiamo osservando com'era la sua faccia qualche frazione di secondo fa, giusto il tempo che la sua immagine percorra la distanza che ci separa. Lo stato mentale del suo cervello è stato comunicato al mio cervello attraverso i muscoli facciali ma in tempi che sono compatibili con la relatività di Einstein e non mediante telepatia. Voglio far notare che la teoria gravitazionale di Newton, in cui due pianeti si attraggono mediante un'azione a distanza, non soddisfa il criterio 2 ma soddisfa una specie di criterio 1. Infatti molti la contestarono definendola un ritorno alla magia e ai concetti di affinità (concetti che sono rimasti almeno nella terminologia in chimica). Eppure la teoria di Newton è crollata sotto la sferza della relatività lasciando il posto alla curvatura dello spazio-tempo, una teoria compatibile col principio di località, a causa del quale sia il criterio 1 che il 2 sono impossibili da accettare senza creare paradossi. Da notare che tutte le teorie classiche del mondo (anche quelle sulla luce) aderiscono al criterio 3, e pertanto sia al criterio di realismo classico (RC) che di località (L). Quando entrambi questi principi

vengono considerati insieme parleremo di principio RL (*realismo locale*). Fin qui monete e dadi truccati. Ma la realtà dei fotoni qual è?

Entanglement: la “spettrale” azione a distanza

Se la teoria elettromagnetica poteva spiegare molti fenomeni purché non si abbassava troppo l'intensità luminosa, di certo l'entanglement è proprio un fenomeno impossibile da descrivere con criteri classici. Supponiamo di prendere due fotoni, ciascuno descritto mediante il suo stato di polarizzazione (verticale V od orizzontale O). Ciascuno di essi vive in un spazio di Hilbert bidimensionale. Abbiamo detto che se il fotone non ha alcuna proprietà del tipo O o V di sicuro una qualche proprietà oggettiva che l'avrà: ad esempio la polarizzazione a 45° (|+>) o a 135° (|->) (Fig. 42).

$$|\pm\rangle = \frac{1}{\sqrt{2}} (|O\rangle \pm |V\rangle) \quad (77)$$

La MQ asserisce che il fotone in questione, pur non avendo alcuna proprietà oggettiva inerente alla polarizzazione orizzontale o verticale, una qualche realtà oggettiva di sicuro ce l'ha: la polarizzazione a 45° o a 135°. Non solo: supponiamo che il fotone superi un test di polarizzazione V, allora in uscita lo stato del fotone sarà |V>. Ebbene se disponiamo un filtro polarizzatore a 45° subito possiamo prevedere cosa accadrà? Dalla coppia di equazioni precedente possiamo vedere che (Fig. 42)

$$|V\rangle = \frac{1}{\sqrt{2}} (|+\rangle + |-\rangle) \quad (78)$$

Quindi il fotone ha una proprietà oggettiva (possiede polarizzazione V) ma non ha alcuna proprietà oggettiva (secondo l'interpretazione RQ) dello stato di polarizzazione a 45° o 135°. Il fotone supererà un test a 45° il 50% delle volte.

Consideriamo ora il sistema “coppia di fotoni”. Lo stato che descrive questo sistema vivrà in uno spazio di Hilbert dato dallo spazio prodotto dei due, le cui basi saranno non più |V> o |O>, come per i fotoni separati, ma |OO>; |OV>; |VO>; |VV>. Ora, stando alla MQ, se è possibile produrre una coppia di due fotoni in ciascuno di questi quattro stati allora è possibile impostare

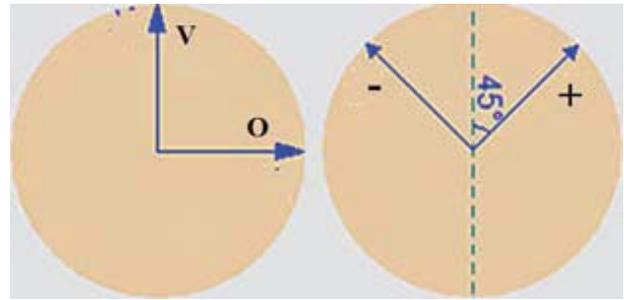


Figura 42: Stati (V ; O) e (+; -) corrispondenti a misure di polarizzazione (verticale, orizzontale) e (45°, 135°) rispettivamente.

un esperimento per il quale lo stato $|\psi\rangle$ della coppia è una sovrapposizione lineare delle quattro basi

$$|\psi\rangle = a |OO\rangle + b |OV\rangle + c |VO\rangle + d |VV\rangle \quad (79)$$

dove a, b, c e d sono numeri complessi tali che la somma dei loro quadrati è pari all'unità per la normalizzazione delle probabilità. Prenderemo, come esempio, tre tipi di coppie per semplificare.

$$|F\rangle = \frac{1}{\sqrt{2}} (|OO\rangle + |VO\rangle) \quad (80)$$

$$|S\rangle = \frac{1}{\sqrt{2}} (|OV\rangle - |VO\rangle) \quad (81)$$

$$|T\rangle = \frac{1}{\sqrt{2}} (|OO\rangle + |VV\rangle) \quad (82)$$

Nonostante questi tre stati siano stati ammissibili nello spazio di Hilbert prodotto, c'è una grande differenza tra il primo stato e gli altri due. Il primo stato si dice *fattorizzato* in quanto possiamo “mettere in evidenza” lo stato |O> a destra e scrivere lo stato |F> come il prodotto tra uno stato che “vive” nello stato di Hilbert del primo fotone (spazio di Alice) e uno stato che vive nello spazio di Hilbert del secondo fotone (spazio di Bob):

$$|F\rangle = \frac{1}{\sqrt{2}} (|O\rangle + |V\rangle)_A \otimes |O\rangle_B \quad (83)$$

Questo invece non è possibile per gli stati |S> e |T>, (chiamati stato di Singoletto e Tripletto) i quali saranno detti *stati entangled*. In sostanza negli stati entangled non è possibile separare i sottosistemi, lo stato sarà indivisibile anche concettualmente e non potremo intenderlo come la semplice somma dei due sottosistemi, come invece accade per lo stato fattorizzato. Semplice-

mente i due fotoni che costituiscono il sistema totale non possiedono proprietà indipendenti. Quando è possibile fattorizzare il sistema, ovvero scriverlo come un prodotto di un vettore appartenente allo spazio di Alice per un vettore appartenente allo spazio di Bob, cioè ai due stati di partenza considerati separati? Lo stato generico $|\psi\rangle$ è fattorizzabile solo se siamo in grado di scrivere questo stato come $|\psi\rangle_A \otimes |\psi\rangle_B$ altrimenti sarà entangled.

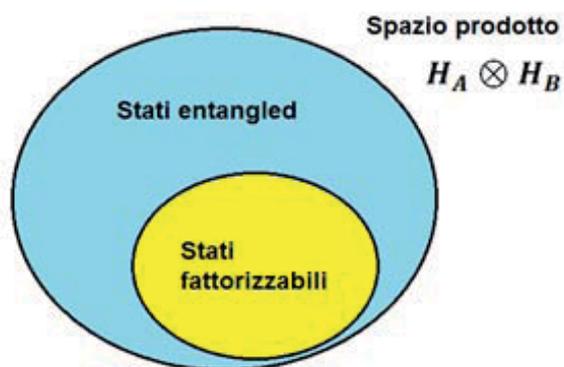


Figura 43: Spazio di Hilbert di stati prodotto. Gli stati si suddividono in stati fattorizzabili e stati entangled.

Questi stati esistono davvero in natura? Vediamo qual è la loro natura stando al formalismo quantistico. Supponiamo che Alice e Bob compiano le stesse misure come per il caso della moneta e del dado in due laboratori diversi, uno sulla Terra e l'altro su Saturno. Immaginiamo che Charlie sia riuscito a produrre coppie di fotoni nello stato $|F\rangle$ e che Alice e Bob decidano di effettuare ad un certo istante (dopo essersi sincronizzati) una misura di polarizzazione verticale e orizzontale rispettivamente. Il formalismo quantistico del secondo membro dello stato $|F\rangle$ indica che Alice avrà il 50% di probabilità che il suo fotone passi il test e altrettanto che non lo passi. Supponiamo che il fotone superi il test. Allora per Alice lo stato sarà collassato da $|F\rangle$ a $|VO\rangle$. Cosa accadrà a Bob? Poiché lo stato è collassato in $|VO\rangle$ Bob troverà lo stato del suo fotone in O ed esso anche supererà il test di Bob (polarizzazione orizzontale). Cosa c'è di strano? Nulla. Infatti possiamo riscrivere lo stato $|F\rangle$ in un modo più chiaro

$$|F\rangle = \left[\frac{1}{\sqrt{2}} |O\rangle + \frac{1}{\sqrt{2}} |V\rangle \right]_A \otimes [1 |O\rangle + 0 |V\rangle]_B \quad (84)$$

Essendo separati, il formalismo ci sta dicendo che Alice ha il 50% di probabilità di vedere il proprio fotone superare il test di polarizzazione verticale, mentre Bob il 100% di vedere il suo fotone superare il test di polarizzazione orizzontale. Perciò il fotone di Alice non ha alcuna proprietà definita di polarizzazione orizzontale o verticale prima della misura mentre il fotone di Bob è certamente polarizzato orizzontalmente. Questo significa che il fotone di Alice, al contrario di quello di Bob, non ha alcuna polarizzazione? Certo che sì. Ricordiamo che finora abbiamo trovato che un sistema ha sempre uno stato oggettivo, pur non potendo definire tutti gli stati come posseduti oggettivamente. Infatti lo stato del fotone di Alice è proprio quello di polarizzazione a 45° e lo stato della coppia si può anche scrivere

$$|F\rangle = [1 |+\rangle + 0 |-\rangle]_A \otimes [1 |O\rangle + 0 |V\rangle]_B = |+, O\rangle \quad (85)$$

che equivale a dire che sia il fotone di Alice che quello di Bob hanno uno stato di polarizzazione ben definito, cioè la polarizzazione a 45° e quella orizzontale sono due proprietà possedute oggettivamente dai due fotoni rispettivamente di Alice e Bob, indipendentemente se la misura viene o meno eseguita. Questo caso è identico al primo esempio della moneta e del dado. In un certo senso i fotoni esistono nel loro stato oggettivo di polarizzazione prima della misura proprio come la moneta e il dado esistono nella loro rispettiva configurazione prima che le scatole vengano aperte (RC). L'unica differenza tutta quantistica è che il fotone di Alice, se rappresentato in una base $|O\rangle$ e $|V\rangle$ non ha uno stato orizzontale o verticale definito, avendolo definito a 45° .

Analizziamo ora uno dei due stati entangled, ad esempio $|S\rangle$. Questo è simile agli stati 66 e 74. In questo caso, se Alice trova che il suo fotone ha superato il test, allora lo stato della coppia è collassato in $|VO\rangle$. Se Alice trova V allora necessariamente Bob trova O e il suo fotone supera il suo test (che ricordiamo essere orizzontale). Ma cosa accade se il fotone di Alice non supera il test

di polarizzazione verticale? Accade che lo stato collassa in $|OV\rangle$. Contrariamente allo stato fattorizzato se Alice cioè trova O allora Bob troverà V. Cioè se il fotone non supera il test *polarizzazione verticale* di Alice nemmeno il fotone di Bob supererà il suo test di *polarizzazione orizzontale*. In altre parole, il sistema si comporterà, *indipendentemente dalla distanza*, come il terzo caso della moneta e del dado. Ciò che fa un fotone farà l'altro, e ciò che non fa un fotone non farà l'altro. In altre parole, ciascun fotone ha una probabilità del 50% di superare o meno il relativo test, proprio come Alice poteva trovare testa con una probabilità del 50% e Bob un numero pari con probabilità del 50%. Ora tuttavia la coppia di fotoni si comporta proprio come il caso in cui ogni volta che Alice trova testa Bob trova un numero pari (2, 4, 6) e ogni volta che Alice trova Croce Bob trova un numero dispari (1,3,5). Eliminando l'ipotesi della telepatia le possibili spiegazioni sono due

1. *Fisica quantistica*: Gli stati dei due fotoni non esistono prima che le misure di polarizzazione vengano effettuate (ovvero prima che le scatole vengano aperte). Non appena Alice misura V il fotone di Bob "collassa" nello stato O. In questo senso la realtà non è definita prima della misura di Alice.
2. *Frode classica*: l'apparato che produce i fotoni (Charlie) "ha barato" e ha preparato i fotoni ad esempio nello stato $|VO\rangle$ già a monte. (In realtà non sarebbe frode ma fisica classica).

Ovviamente ora siamo svezzati per poter accettare la "magia quantistica" almeno per i fotoni e non pensare alla frode classica. Tuttavia qualcosa non torna. Se accettiamo la prima ipotesi, infatti, il collasso del fotone di Alice si propagherà all'istante dalla Terra fino a Saturno, superando i limiti imposti dalla velocità della luce c e violando la relatività, o meglio il principio di località L. Se invece accettiamo la seconda ipotesi allora stiamo supponendo che i fotoni stessero già nello stato $|VO\rangle$ proprio come la moneta era già nello stato "Testa" e il dado nello stato "numero pari" anche prima dell'apertura delle scatole e il collasso dello stato è realmente un'istantanea presa di conoscenza dello stato nella scatola. In tal caso durante il viaggio dei due fotoni o delle due scatole fino alle postazioni di Alice e Bob

gli stati dei sottosistemi erano già predeterminati oggettivamente e non avverrebbe così alcuna violazione di L. Insomma la spiegazione 1 viola il realismo locale RL, mentre la 2 rispetta RL.

Essere o non essere

Il punto cruciale per accettare una ipotesi o l'altra è capire cosa sia davvero uno stato entangled come lo stato di singoletto. Ricorderete che quand'anche la MQ asseriva che un fotone in sovrapposizione non possedeva oggettivamente nessuno dei due stati $|O\rangle$ o $|V\rangle$ ne possedeva tuttavia certamente uno (ad esempio $|+\rangle$) Ebbene lo stato di singoletto ha una proprietà molto strana:

contrariamente a particelle separate, le due particelle dello stato entangled di singoletto S, se prese separatamente, non hanno alcuna proprietà di polarizzazione di alcun tipo.

Sviluppiamo infatti questo stato nella base $|+\rangle$, $|-\rangle$ che corrispondono agli stati a 45° e 135° .

$$\begin{aligned}
 |S\rangle &= \frac{1}{\sqrt{2}} |O\rangle \otimes |V\rangle - \frac{1}{\sqrt{2}} |V\rangle \otimes |O\rangle \\
 &= \frac{1}{\sqrt{2}} \left(\left(\frac{1}{\sqrt{2}} |+\rangle - \frac{1}{\sqrt{2}} |-\rangle \right) \otimes \left(\frac{1}{\sqrt{2}} |+\rangle + \frac{1}{\sqrt{2}} |-\rangle \right) \right. \\
 &\quad \left. - \frac{1}{\sqrt{2}} \left(\left(\frac{1}{\sqrt{2}} |+\rangle + \frac{1}{\sqrt{2}} |-\rangle \right) \otimes \left(\frac{1}{\sqrt{2}} |+\rangle - \frac{1}{\sqrt{2}} |-\rangle \right) \right) \right) \\
 &= \frac{1}{2\sqrt{2}} (|+,+\rangle + |+,-\rangle - |-,+\rangle - |-, -\rangle) \\
 &\quad - |+,+\rangle + |+,-\rangle - |-,+\rangle + |-, -\rangle) \\
 &= \frac{1}{\sqrt{2}} |+,-\rangle - \frac{1}{\sqrt{2}} |-,+\rangle \quad (86)
 \end{aligned}$$

Ho evidenziato i termini che si elidono con lo stesso colore. In sostanza lo stato di singoletto assume la stessa forma nelle basi complementari dei test a 45° e 135° . In generale si può provare che, date due direzioni qualsiasi di polarizzazione perpendicolari tra loro $\theta = n$ e $\theta + \frac{\pi}{2} = p$, lo stato di singoletto può sempre scriversi come

$$|S\rangle = \frac{1}{\sqrt{2}} |n;p\rangle - \frac{1}{\sqrt{2}} |p;n\rangle \quad (87)$$

per qualsiasi coppia di polarizzazioni ortogonali tra loro. La conseguenza è drammatica: per la MQ il sistema si troverà sempre in una sovrapposizione di stati indipendentemente da quale misura di polarizzazione scegliamo, ovvero non vi

è alcun test di polarizzazione del tipo θ e $\pi/2 + \theta$ per il quale il fotone di Alice o di Bob supererà col 100% delle probabilità, ma accadrà sempre che il fotone passerà il 50% delle volte in modo casuale. Pertanto, contrariamente al caso di fotoni separabili, i fotoni *entangled* presi singolarmente non possiederanno alcuna proprietà di polarizzazione oggettiva. Solo la misura di Alice (o di Bob) oggettiverà il sistema a seconda del tipo di test che si vorrà effettuare. Ne derivano delle conseguenze di portata enorme:

1. *In fisica classica possiamo determinare lo stato di un sistema composto solo conoscendo in modo completo lo stato (oggettivo) di ciascuno dei suoi sottosistemi. Se non conosciamo tali sottostati possiamo solo dare una descrizione statistica del sistema. In fisica quantistica possiamo conoscere in modo completo la proprietà totale del sistema senza conoscere le proprietà dei sottosistemi in quanto, per l'interpretazione RQ, questi non hanno alcuna proprietà di polarizzazione oggettiva.*
2. *lo stato di Bob crollerà istantaneamente a seguito della misura di Alice e diverrà oggettivo all'istante, indipendentemente dalla distanza e indipendentemente dagli angoli di polarizzazione dei polarizzatori scelti.*

Queste due considerazioni ci portano ad alcune conseguenze. Anzitutto solo la totalità indivisa ha senso oggettivo (olismo), non avendo le sottoparti alcuna proprietà fisica. Pertanto il *riduzionismo* ontologico, cioè la visione secondo cui le proprietà di un sistema sono date dalle proprietà possedute oggettivamente dai suoi componenti, viene scosso dalle fondamenta. Inoltre nessuna proprietà può essere attribuita ai sottosistemi di un sistema *entangled*. Quando, a causa della misura, uno di essi acquista un elemento di realtà, il formalismo quantistico asserisce che l'altro sottosistema immediatamente acquisterà la proprietà opposta, senza che la distanza ritardi l'effetto di collasso dello stato. Questa parrebbe una seria violazione del principio di Località L. Non solo la meccanica quantistica, se completa, viola il presupposto RC, ma violerebbe anche L. Quindi la MQ non soddisfa RL. O la realtà non è oggettiva, o non è locale o entrambe. Come ebbe a dire Paolo di Tarso, tutto ciò fu scandalo per alcuni, follia

per altri, incluso Einstein. Nel 1935 Einstein insieme ai colleghi Podolski e Rosen (EPR), scrisse un lavoro dal titolo "*can the quantum-mechanical description of physical reality be considered complete?*" [8] da cui si evince che la fisica quantistica non può essere completa, a meno di non rinunciare alla località o al realismo classico oggettivo. Se infatti è da completare allora

1. i sottosistemi possiedono delle proprietà oggettive anche prima della misura,
2. non c'è alcun segnale superluminale da inviare perché non c'è alcun collasso, ma solo una presa di conoscenza degli stati dei sottosistemi.

Se non fossero state eseguite misure su tali stati io personalmente avrei concluso che gli stati *entangled* sono soluzioni della teoria da scartare (soluzioni spurie-SOSP). Quando una SOSP appare viene il SOSPetto che essa sia da scartare in quanto fisicamente non sostenibile. E' già accaduto con la soluzione dell'atomo di idrogeno mediante l'equazione di Shroedinger. Alcune soluzioni esplodono all'infinito esponenzialmente e vanno scartate perché prive di significato fisico. Altre volte tuttavia scartare tali soluzioni si è rivelato un grave errore, come nel caso delle soluzioni ad energia negativa che indicano l'esistenza dell'antimateria. Nel caso delle soluzioni *entangled*, data la follia dei risultati, ci verrebbe da dire, per salvare sia la MQ che la relatività, che non sono soluzioni accettabili. Il punto è però che tali soluzioni non sono come gli esponenziali infiniti di probabilità dell'equazione di Shroedinger ma come quelle ad energia negativa dell'antimateria. Esse cioè sono state verificate in laboratorio. Coppie di fotoni *entangled* in polarizzazione possono essere prodotte oramai quotidianamente nei laboratori irraggiando un cristallo non lineare con fotoni ultravioletti. Se i fotoni *entangled* esistono (ed esistono) ha ragione la MQ oppure è possibile completare la MQ con una teoria in cui la descrizione dello stato quantico è solo da intendere come descrizione statistica e legata alla nostra conoscenza (o ignoranza)? EPR concludono il lavoro così

"We left open the question of whether or not such a description exists. We believe, however, that such a theory is possible".

E' possibile capire se una siffatta teoria, oggettiva e locale, in sostituzione della MQ, esiste?

La disuguaglianza di Bell e gli esperimenti di Aspect

Un passo in avanti decisivo si ebbe nel 1962, quando il fisico John Bell pubblicò [10] un lavoro in cui provò che le teorie locali e oggettive (RL) ponevano un limite al grado di correlazione tra due particelle. Il suo teorema, riformulato poi da Clauser, Horne, Shimony e Holt (teorema SHSH), [11] fa leva sulla misura congiunta di varie situazioni. Supponiamo di avere due basi $|O\rangle, |V\rangle$ per Alice e $|+\rangle, |-\rangle$ per Bob e di misurare congiuntamente le probabilità per i due fotoni $P(O,+), P(O,-), P(V,+), P(V,-)$. Come potete vedere in questo caso i due polarizzatori di Alice e Bob non usano la stessa base ma basi complementari (Fig. 42). Ebbene la disuguaglianza di Bell asserisce che, se i due fotoni hanno polarizzazioni oggettive e nessun segnale può essere trasmesso più velocemente della luce (RL), allora vale la seguente disuguaglianza

$$|P(O,+) - P(O,-)| + |P(V,+) + P(V,-)| \leq 2 \quad (88)$$

In sostanza la probabilità che Alice misuri "O" e la probabilità che Bob misuri "+" tranne la probabilità che Alice misuri "O" e Bob misuri "-" oppure la probabilità che Alice misuri "V" e Bob misuri "+" oppure che Alice misuri "V" e Bob misuri "-" sarà minore di 2. Fate la prova sostituendo a "O" lo stato "pantalone di Jeans" e a "+" lo stato "occhiali". V sta ad indicare "nessun pantalone di jeans" e "-" indica "nessun paio di occhiali". Questi sono stati classici. Ovviamente infatti nessuno si sognerebbe di credere che lo stato del pantalone che indossate sia indeterminato prima di osservarlo. Prendete un po' di persone statisticamente significative e verificate che la seguente disuguaglianza sia soddisfatta

$$|n(\text{Jeans, Occh.}) - n(\text{Jeans, No Occh.})| + |n(\text{No Jeans, Occh.}) + n(\text{No Jeans, No Occh.})| \leq 2 \quad (89)$$

dove n rappresenta il numero che conteggiate con quella coppia di caratteristiche in parentesi diviso il numero totale di persone. Torniamo ai fotoni. Se è possibile trovare una teoria locale e oggettiva che ha la pretesa di spiegare i risultati quantistici allora essa prevede che il massimo grado di correlazione tra queste probabilità congiunte oscillerà tra -2 e +2.

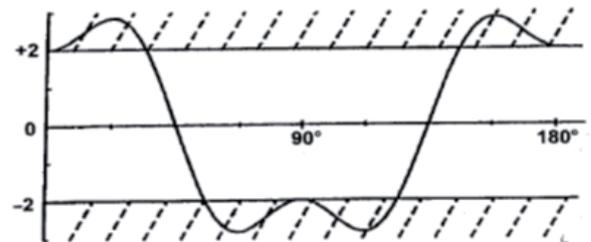


Figura 44: Le teorie classiche del mondo che rispondono al requisito RL (realismo locale) fanno previsioni probabilistiche congiunte alla Bell sono nell'intervallo tra -2 e 2. I fenomeni di entanglement previsti dalla MQ violano la disuguaglianza (linee della curva nelle zone tratteggiate). In figura sono mostrate le previsioni teoriche di probabilità congiunte per esperimenti di polarizzazione di fotoni. Gli angoli indicano la direzione reciproca di polarizzazione dei filtri di polarizzazione.

Quali previsioni da la MQ? Essa prevede che misure del genere congiunte, ovvero $\langle \sigma^A \bullet \sigma^B \rangle$, diano per uno stato di singoletto il valore di $-\cos(\theta)$ dove θ è l'angolo tra le polarizzazioni dei polarizzatori di Alice e Bob, cioè

$$\langle \sigma^A \bullet \sigma^B \rangle = -\cos \theta_{AB} \quad (90)$$

Infatti se tale angolo è nullo vi è anticorrelazione, $\langle \sigma^A \bullet \sigma^B \rangle = -1$, come abbiamo potuto appurare.

Pertanto la MQ per gli stati di cui sopra fornisce

$$|P(O,+) - P(O,-)| + |P(V,+) + P(V,-)| = |-\cos \pi/4 - \cos 3\pi/4| + |-\cos \pi/4 - \cos \pi/4| = 2\sqrt{2} \quad (91)$$

Che è palesemente superiore a 2! La MQ, pertanto, fa la previsione che esisterebbero stati che violerebbero le disuguaglianze di Bell per alcuni angoli, come riportato in fig.44. Questo implica che gli stati entangled non esistono nella real-

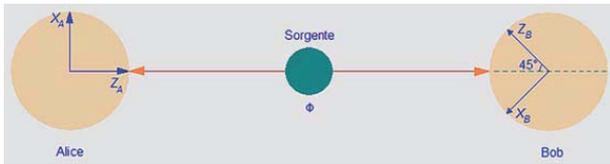


Figura 45: Tipica configurazione di un esperimento con fotoni entangled per trovare gli effetti di correlazione massima che violano le disuguaglianze di Bell.

tà? Oppure che esistono ma non possono essere descritti da alcuna teoria locale e oggettiva che abbia la pretesa di completare la MQ?

Nel 1982 Alain Aspect (dopo altri tentativi precedenti ma meno convincenti) dimostrò sperimentalmente all'istituto di ottica teorica e applicata di Parigi che era non solo possibile produrre coppie entangled ma che tali sistemi si comportavano violando le disuguaglianze di Bell. L'apparato è mostrato in fig. 45 e 46. Una sorgente di atomi di calcio, a seguito di eccitazione laser, produceva a cascata coppie di fotoni entangled in polarizzazione i quali erano inviati verso due detector A e B. Tuttavia Aspect decise di rendere inefficace una improbabile ma potenziale via di comunicazione luminale tra i rivelatori, che avrebbe impedito di accertare che la località fosse stata violata. L'idea fu geniale: spedire i due fotoni verso due apparati in cui uno specchio cambiava orientazione in modo casuale mentre i fotoni erano ancora in volo. In questo modo ciascun fotone poteva essere o meno riflesso, a caso, verso uno di due polarizzatori. Era come scegliere all'ultimo secondo quale polarizzatore usare (i due polarizzatori, per ciascun fotone, sono indicati con +1 e -1 in figura). Ebbene grazie a questi accorgimenti gli esperimenti di Aspect provarono che le disuguaglianze di Bell erano violate con uno scarto superiore a quaranta deviazioni standard. Per dare un'idea dell'efficacia dei risultati basti pensare che oggi il Bosone di Higgs è stato rilevato e confermato con uno scarto di sole 5 deviazioni standard!

Cosa ne deduciamo? La conclusione è che il comportamento dei fotoni suggerisce che la realtà non può essere simultaneamente *oggettiva e locale*. Il realismo locale è morto! L'aspetto che non dobbiamo trascurare è che è stata la MQ a consentirci di effettuare tali esperimenti, in quanto è proprio la MQ ad aver fatto la previsione di

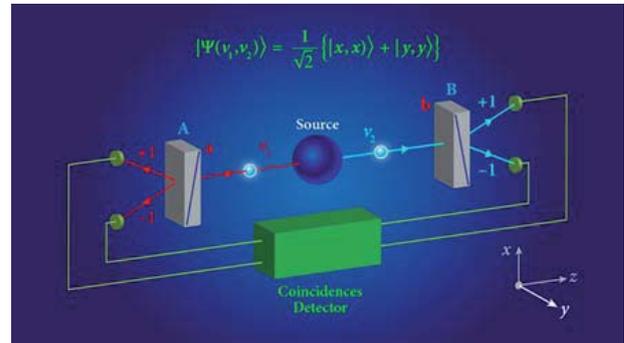


Figura 46: Esperimento di Aspect (1982). La sorgente emette coppie EPR di fotoni polarizzati. Alice (a sinistra) e Bob (a destra) possiedono due rivelatori di polarizzazione che segnano +1 e -1 come autovalori.

esistenza delle coppie entangled. Ma se ora ribaltassimo lo sviluppo storico e dicessimo di realizzare un esperimento *à la Aspect* senza preconcetti teorici, dopo aver visto le correlazioni in atto cosa dedurremmo? Ne dedurremmo che qualsiasi teoria (incluso la MQ) che debba descrivere tali esperimenti non potrà soddisfare simultaneamente i criteri RC e L in quanto i fotoni entangled sono *troppo correlati* rispetto al massimo grado di correlazione previsto da una qualsiasi possibile teoria oggettiva e locale.

Il senso della non località

Vorrei spendere due parole sullo status finale della descrizione quantistica della realtà degli stati entangled che abbiamo visto sopra, ovvero qual è il senso della non località. Spesso si sente dire che gli esperimenti con i fotoni condotti da Aspect avrebbero decretato la vittoria della meccanica quantistica contro le teorie a variabili nascoste. La frase va assolutamente precisata. Anzitutto, come già più volte ripetuto, gli esperimenti di Aspect, e altri simili, hanno provato che *nessuna teoria* seria può essere *simultaneamente locale e oggettiva* in senso classico. Punto. Una volta che sia stata fatta piazza pulita di teorie di questo tipo, questo lascia tuttavia spazio sia a teorie *non oggettive e non locali* (la trattazione ortodossa della MQ nelle varie interpretazioni) e sia a teorie *oggettive e non locali* (come la teoria di Bohm). La vittoria è di un insieme di versioni della MQ (Copenaghen, Bohm, Teoria degli stati relativi ecc) su teorie oggettive e locali, ormai

defunte. Il punto ora è scegliere tra le diverse teorie MQ. In quell'insieme le problematiche sono più sottili ed è sbagliato dire che la MQ coincide con la versione alla Bohr, ovvero con l'interpretazione di Copenaghen. Sennò non si chiamerebbe interpretazione. Riassumo le posizioni alla Copenaghen (che va per la maggiore) e alla Bohm (a mio avviso l'unica alternativa seria alla interpretazione standard) per ciò che concerne l'entanglement

1. *nella interpretazione di Copenaghen i sottosistemi che formano lo stato di singoletto non hanno alcuna realtà oggettiva finché non lo si osserva. Una volta che ciò sia accaduto l'oggettivazione della realtà del sottosistema A determina all'istante l'oggettivazione del sottosistema B violando la località.*
2. *Nella interpretazione di Bohm i sottosistemi che formano lo stato di singoletto hanno una realtà oggettiva indipendentemente se li si osserva o meno. Tuttavia una misura sul sistema A ne può perturbare lo stato e all'istante l'onda di Bohm (il potenziale quantico) comunica al sottosistema B tale perturbazione, violando la località. La teoria presuppone un tipo di campo non locale ad hoc (l'ampiezza della funzione d'onda) che non trasporta energia, mentre le variabili nascoste non sono accessibili.*

Rispetto a tali interpretazioni dovremmo forse essere più pragmatici ed accettare uno o l'altro modello a seconda del potere predittivo che hanno e della capacità di consentire uno sviluppo ulteriore della ricerca. L'interpretazione statistica di Copenaghen, ad esempio, permette di poter effettuare calcoli su situazioni diverse senza postulare alcun modello strutturale delle particelle elementari. Tuttavia è possibile che una tale interpretazione vincoli la ricerca ad una visione limitata, mentre, come ha fatto rilevare Bohm, probabilmente la funzione d'onda ha un contenuto più ricco e ci sta informando che lo spazio o il tempo sono strutture ben più esotiche di quanto immaginiamo. E una interpretazione o l'altra può fare la differenza sulla direzione da dare alla ricerca. Ma alla fine, la realtà quantistica, al di là delle interpretazioni, è in conflitto o no con la relatività? La risposta è sottile e richiede una spiegazione precisa.

Nella interpretazione ortodossa di Copenaghen la relatività è fatta salva perché i risultati di Alice e Bob sono genuinamente casuali. Non c'è possibilità per Bob, da una lettura delle sue proprie misure i cui esiti sono casuali, capire se Alice ha o meno condotto l'esperimento. Supponiamo che Alice conduca l'esperimento ogni volta che è raggiunta da un fotone della coppia, e che ottenga su quindici esperimenti la seguente stringa di dati [010001011001011] (per semplicità al risultato 1 abbiamo associato 1, e al risultato -1 abbiamo associato 0). Allora Bob otterrà una stringa di risultati casuale correlato, ovvero [101110100110100]. Se Alice non avesse eseguito tutti gli esperimenti (o addirittura nessuno) in modo da far collassare lo stato del fotone di Bob, allora quest'ultimo avrebbe ottenuto una stringa diversa dalla prima ma sempre casuale, ad esempio del tipo [110110001010100], a causa questa volta del collasso dovuto alla propria misura. A causa della casualità di entrambe le misure, Bob non può dedurre dalla semplice lettura dei suoi esiti, se Alice ha condotto o meno la misura. In altre parole non potrà capire se la causa del collasso del suo fotone B sia stata la misura effettuata col polarizzatore B o se la causa sia da attribuire al collasso del fotone A mediante la misura che Alice ha compiuto col suo polarizzatore A. L'unica cosa che Bob può fare è "telefonare" ad Alice e confrontare gli esiti e le modalità delle misure, ma questo include l'uso di un canale non superluminale classico di comunicazione. Grazie alla casualità, nonostante la non località, non è possibile inviare informazione o materia o energia a velocità superluminali. In questo senso la località è salva. Il punto cruciale è che nel caso delle misure di Alice e Bob non ha senso nemmeno dire chi è la causa di chi. Riprendiamo l'esempio fatto nella seconda parte di questo lavoro quando, a proposito della relatività, abbiamo discusso come appaiono eventi simultanei e causalmente connessi quando sono visti da due osservatori in moto relativo.

Supponiamo che al centro degli assi coordinati vi sia la sorgente EPR che invia i due fotoni verso Alice e Bob lungo il cono luce (in quanto i fotoni viaggiano a velocità c). Quando i due fotoni intercettano le linee universo di Alice e Bob essi vengono misurati simultaneamente e nessuno è la causa di nessuno. Tuttavia possiamo sempre

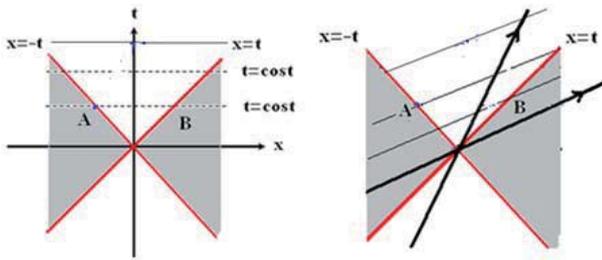


Figura 47: Esperimento di Aspect in una prospettiva spaziotemporale relativistica. A sinistra l'osservatore è fermo rispetto ad Alice e Bob. A destra l'osservatore è in moto verso Bob. Ciò che prima era simultaneo ora è una sequenza di eventi in cui la misura di Bob viene prima di quella di Alice.

scegliere un riferimento in moto verso B (figura a destra) in cui l'evento B avviene prima di A.

In questo sistema di riferimento concludiamo che B è la causa di A. Se invece ci poniamo in un riferimento verso Alice allora l'evento che descrive la misura fatta da Alice avverrebbe sicuramente prima della misura di Bob e concluderemmo che la misura di Alice è la causa del risultato che ottiene Bob. Ma sbagliaremmo perché semplicemente il concetto di prima e dopo causale è relativo solo ad eventi connessi all'interno del cono-luce, mentre questi eventi A e B sono fuori dal cono-luce di ciascuno (B si trova fuori dal cono Luce futuro e passato di Alice e viceversa). Ecco perché dire che uno è la causa dell'altro è sbagliato. Possiamo solo dire che gli eventi sono correlati ma non causalmente connessi.

Nella teoria di Bohm accade qualcosa di simile. Come interpreta questa teoria gli stati entangled? Se due particelle di massa m_1 e m_2 sono inviate contro due detector la loro velocità sarà in relazione con la fase S , ovvero

$$m_1 v_1 = \nabla_1 S(1, 2); m_2 v_2 = \nabla_2 S(1, 2) \quad (92)$$

E' ovvio che se si altera la velocità di una particella allora anche S sarà alterata e la velocità della seconda particella subirà una variazione istantanea. In questo caso è evidente come la funzione d'onda sia non locale. Nonostante la non località della funzione d'onda, non è possibile tuttavia nemmeno in questo caso inviare messaggi superluminali perché, per farlo, bisognerebbe accedere alle variabili nascoste e modificarle per cambiare la casualità epistemologica

degli esiti delle misure. Se Alice accedesse a tali variabili, Bob, misurando il suo fotone, potrebbe trovare una stringa palesemente non casuale, del tipo [10101010101010]. Sarebbe un po' come lanciare una moneta e avere successivamente Testa, Croce, Testa, Croce e così via. Si potrebbe così inviare, grazie alla non località della funzione d'onda, messaggi superluminali binari. Ma ciò non è possibile proprio perché le variabili nascoste restano nascoste. Né tanto meno la funzione d'onda potrebbe trasportare energia e materia a velocità superluminali. E allora come mai spesso si dice che le teorie a variabili nascoste (non quelle RL ma quella alla Bohm, oggettive e non locali) sono roba da cimitero delle cose vecchie e non si insegnano nei corsi universitari? La risposta richiederebbe un lavoro a cavallo tra scienza e sociologia e non ci addentreremo nella disamina. Tuttavia la teoria di Bohm non viene tipicamente considerata nei corsi accademici. Purtroppo anche l'oggettività di alcuni scienziati (a volte tra i più quotati) spesso non è garantita, sebbene ne sia garantita la località in riviste prestigiose che trascurano i *border lines*. Per lealtà nei confronti del lettore, farò la mia professione di fede. Credo che la teoria di Bohm sia la più razionale tra tutte quelle che siano state prese in considerazione e che forse non accettarla a volte deriva dal fatto che essa ruba un po' di fascino a ipotesi fantascientifiche come quella degli universi paralleli, oggi di moda. Se accettiamo la teoria di Bohm, però, dobbiamo accettare che la funzione d'onda non sia una astrazione matematica vincolata necessariamente a una descrizione probabilistica, ma fa parte della struttura stessa della realtà. E dovremmo piuttosto domandarci: quale tipo di spazio-tempo può ammettere funzioni d'onda alla Bohm, ovvero non locali e tali da non trasportare energia? Ci sono molti elementi a mio avviso che provano che seguire questa traccia sia conveniente. Se il campo elettrico ad esempio (o il potenziale vettore) è la funzione d'onda del fotone, come mai attribuiamo una realtà fisica al campo ma non alla funzione d'onda dell'elettrone, che ne è il suo analogo? Se accettiamo ciò si aprirebbero nuove prospettive di indagine nei confronti delle quali l'interpretazione probabilistica di Copenaghen è assolutamente neutra. Alcune di queste ipotesi potrebbero dar ragione dell'esistenza dello spin in termini di struttura

delle particelle elementari, della loro connessione con altre strutture come lo spazio-tempo, nonché della gravità quantistica. Sono solo ipotesi di lavoro, ma degne di essere perlustrate. La situazione a mio avviso è simile a quando si usava a proposito della nascita del cosmo e della vita l'ipotesi di Dio creatore che, anche se sembrava spiegare la complessità dell'esistente biologico e non possiamo escluderla a priori, tuttavia costituiva un ostacolo metodologico alla scoperta di altre spiegazioni naturalistiche e oggi, forse, non avremmo avuto la meravigliosa teoria di Darwin o la meno bella ma pur sempre affascinante teoria dell'universo inflazionario. C'è un'ultima osservazione da fare. In un certo senso, qualsiasi descrizione usiate (alla Bohr o alla Bohm), la meccanica quantistica salva la relatività ristretta (mediante la casualità alla Bohr o la non accessibilità delle variabili nascoste alla Bohm) e viceversa la relatività ristretta protegge la causalità nella meccanica quantistica. La casualità regge la causalità e questo senza che nella discussione quantomeccanica si sia fatto uso di una ipotesi relativistica come nelle teorie quantistiche relativistiche di campo. Forse, le due strutture teoriche sono più connesse di quanto immaginiamo. E' un po' come la teoria di Darwin e la genetica di Mendel. L'una può esistere senza l'altra, ma l'una salva l'altra e la connessione tra le due è il DNA che spiega sia le leggi di Mendel dell'ereditarietà e sia l'evoluzione delle specie. Manca una teoria fondamentale per far luce su questo quadro singolare delle due teorie? A mio avviso sì, ma occorre indagare meglio il campo di Bohm.

Imbarazzi e gatti

Toccheremo adesso un aspetto cruciale della fisica quantistica che ha a che vedere con gli stati entangled: la descrizione dei sistemi macroscopici. Nei fatti ogni fisico divide il mondo in due: una realtà "sensibile" e macroscopica fatta di sassi, apparati di misura e gatti e una realtà fatta di fotoni, elettroni e altre particelle elementari. Il primo mondo segue comportamenti classici fatti di traiettorie ben definite, localizzazioni nel teatro dello spaziotempo e viene chiamato mondo classico, a cui appartengono anche le nostre percezioni di esseri macroscopici. Il secondo mondo segue comportamenti quantistici, fatti di sovrapposizioni di traiettorie, bilocazione ed elementi di realtà non oggettive.

Il primo è legato a leggi di causa-effetto reali (come la seconda legge di Newton) e probabilità epistemologiche (come nel caso della meccanica statistica), il secondo a leggi di causa-effetto di funzioni d'onda che vivono in spazi matematici astratti (come l'equazione di Shroedinger) e a probabilità ontologiche (come il collasso della funzione d'onda). Insomma non potrebbero essere mondi così diversi. Eppure tra questi due mondi vi è un doppio legame. Anzitutto il collasso della funzione d'onda di un oggetto quantistico è determinato da un apparato di misura, che per definizione è classico. Anche il nostro occhio è un apparato classico. Pertanto per poter accedere sperimentalmente al mondo quantistico occorre postulare il mondo classico. E tuttavia molte proprietà classiche, come la fluidità, possono essere comprese appieno solo postulando proprietà quantistiche. Le due visioni teoriche sembrano interconnesse e non potrebbero sembrare più diverse. Il punto è che gli oggetti classici, che servono per descrivere il mondo quantistico come viene da noi percepito, sono una combinazione di oggetti quantistici (le particelle elementari) ma si comportano in modo assolutamente non quantistico. Viceversa per descrivere il mondo quantistico mettiamo in combinazione storie classiche che spesso si comportano in modi strani (stati discreti, simultaneità di eventi, effetti non locali). Questi aspetti e i paradossi che ne emergono furono chiariti da Shroedinger nel suo famoso argomento del gatto. In questo esperimento ideale e macabro, che modificheremo, si predispose un laser a singolo fotone che spara fotoni polarizzati a 45° verso un cristallo birifrangente dietro cui vi sono due detector collegati a un martello posizionato sopra una fiala di veleno. Il tutto è collocato in una scatola. Ora viene la parte macabra. Prenderemo il nostro gatto e lo infileremo nella scatola la quale verrà poi chiusa. A un certo punto decidiamo di accendere il laser il quale invierà il singolo fotone. Ora se il fotone percorrerà il raggio ordinario il detector superiore suonerà e non attiverà il martello e il gatto vivrà. Se invece percorrerà il tratto inferiore (il raggio straordinario) esso giungerà nel detector inferiore facendo attivare il martello il quale romperà la fiala il cui veleno sarà respirato dal gatto il quale morirà subito dopo.

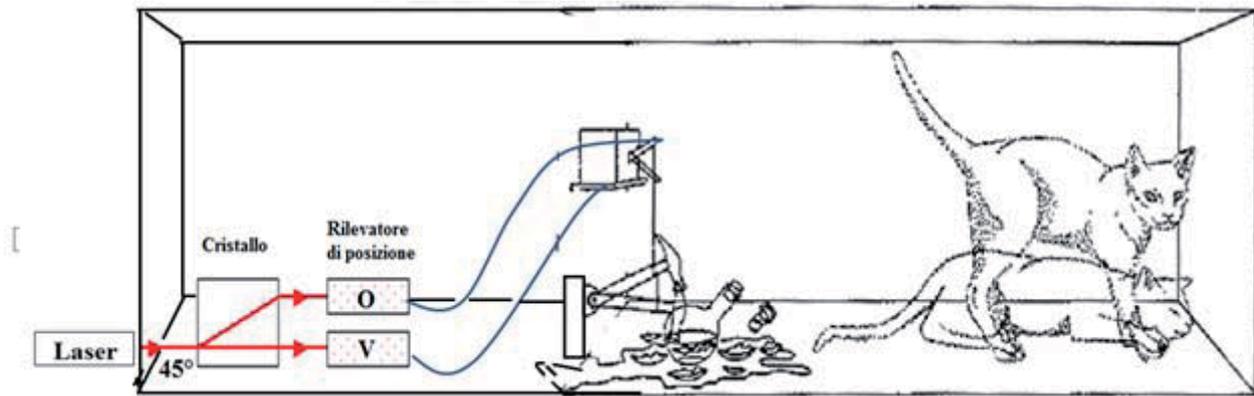


Figura 48: Paradosso del gatto di Shroedinger con fotoni.

Classicamente il fotone prenderebbe una strada o l'altra dando vita a due storie alternative

$$|fotone\ passa\ gi\grave{u};\ detector\ V\ clicca;\ gatto\ morto\rangle \quad (93)$$

$$|fotone\ passa\ su;\ detector\ O\ clicca;\ gatto\ vivo\rangle \quad (94)$$

Noi non sappiamo quale delle due alternative si manifesta nella scatola chiusa finché non la osserviamo. Tuttavia, stando al formalismo standard, ciascun sistema (fotone o gatto che sia) potrà essere descritto in una sovrapposizione di questo tipo:

$$|S\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \quad (95)$$

Nel caso del fotone e del gatto i due stati $|0\rangle$ e $|1\rangle$ sono rispettivamente $|V\rangle$ e $|O\rangle$ e $|gatto\ vivo\rangle$ e $|gatto\ morto\rangle$. Il punto però è che i due sistemi sono descritti dai fisici in modo completamente differente perché il fotone è un oggetto appartenente al mondo quantistico e quella sovrapposizione rappresenta un effetto reale e misurabile (l'interferenza), mentre il gatto è un oggetto classico e quella sovrapposizione rappresenta solo una manifestazione della nostra ignoranza prima dell'apertura della scatola. Se i due sistemi fossero separati da una barriera invalicabile tutto andrebbe bene e potremmo accontentarci di due descrizioni alternative del mondo. Tuttavia il gatto è fatto di organi, questi di cellule, queste di molecole che sono fatte di atomi, i quali sono costituiti da particelle elementari ciascuno dei quali si comporta in modo quantistico. Pertanto questa barriera non dovrebbe esistere e la descri-

zione dovrebbe essere unica. Ma qui avviene il paradosso. Se la descrizione è unica dobbiamo cambiare l'interpretazione dell'eq. (95) del fotone rendendo la sovrapposizione di natura epistemologica come accade per descrivere il gatto, o al contrario dobbiamo cambiare l'interpretazione dell'eq. (95) del gatto rendendo gli stati in sovrapposizione vivo e morto reali e non oggettivi? La faccenda è resa ancora più paradossale proprio dall'entanglement e dal fatto che gli oggetti possono essere composti mediante il prodotto tensoriale. Se usiamo il formalismo della fisica quantistica, poiché ogni sistema (fotone e gatto) vive in uno spazio a due stati, abbiamo che lo spazio prodotto in cui vive il sistema composto è fornito dal prodotto delle seguenti basi

$$|V; vivo\rangle; |O; morto\rangle; |O; vivo\rangle; |V; morto\rangle \quad (96)$$

Tuttavia stati come $|V; vivo\rangle$ e $|O; morto\rangle$ non sono stati da prendere in considerazione perché non sono causalmente connessi. Se il fotone è V allora a scattare non è il detector sù e il gatto non può restare vivo. Se il fotone è O, a scattare non è il detector giù e il gatto non può essere morto. Gli stati causalmente connessi sono solo quelli in cui il fotone è O e il gatto vivo, e il fotone è V e il gatto morto. Gli altri due stati sono una collezione di eventi non connessi e non permessi. Ora il punto è che, se sono permessi quei due stati e accettiamo una descrizione quantistica dell'intero sistema, allora lo stato seguente è permesso

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|V; morto\rangle + |O; vivo\rangle) \quad (97)$$

Come possiamo descrivere questo sistema entangled tra un sistema microscopico e uno macroscopico? Mediante una interpretazione epistemologica od ontologica? Se è epistemologica allora non si capisce perché i fotoni vivono realmente in sovrapposizione e mostrano interferenza, se è ontologica non si capisce perché non vediamo mai gatti in sovrapposizione. Per uscire dalla scappatoia potremmo pensare che uno stato descritto dalla equazione (97) non sia permesso. Tuttavia il punto è che un sistema quantistico in breve tempo entra in stato entangled col suo ambiente e con tutto ciò con cui interagisce (vedi i cammini di Feynman per le interazioni). Insomma in men che non si dica ogni oggetto macroscopico dovrebbe essere entangled con tutto. L'universo intero apparirebbe una mescolanza di stati in sovrapposizione non fattorizzabili. Ecco il paradosso. E per uscirne o occorre completare la meccanica quantistica introducendo elementi di oggettività o occorre capire come la realtà classica emerga dalla realtà quantistica a causa di qualche ingrediente che manca in entrambe le descrizioni a qualche livello intermedio tra un fotone e un gatto. Abbiamo già visto che completare la fisica quantistica in modo oggettivo implica l'introduzione di variabili nascoste non accessibili e una non località della funzione d'onda che però non spiega il perché tali non località non si manifestino a livello macroscopico. L'ampiezza della funzione d'onda R nella teoria di Bohm rappresenta il potenziale quantistico che descrive come viene pilotata la particella nello spazio, mentre la fase S determina le correlazioni descritte dalla eq. (92) tra le particelle. Mentre si comprende come mai nel mondo macroscopico (per grandi masse) il potenziale quantico è trascurabile, non si capisce come gli effetti di entanglement possano essere evitati. D'altro canto tutti gli studi volti a identificare un qualche livello intermedio in cui si ha una sorta di transizione tra realtà quantistica e realtà classica non hanno mostrato alcun successo, anche se la strada è ancora lunga.

Una interpretazione FAPP: la decoerenza

Recentemente per spiegare come mai non vediamo allo stesso tempo gatti vivi e morti è stata proposta la teoria della decoerenza ambientale.[12] Purtroppo lo spazio non è sufficiente per poterci addentrare nella disamina di questo modello che sarà affrontata altrove. Il punto cruciale è proprio che, man mano che il sistema interagisce con strutture sempre più grandi, la sovrapposizione tende a mescolarsi (entanglement) con il grande numero di stati di cui è composto un macrosistema finché non è più possibile osservare la sovrapposizione dello stato quantico che appare invece in una miscela statistica classica. La matrice densità è un utile strumento per tale descrizione. Per la situazione 97 la matrice densità è la seguente

$$\begin{aligned} \rho &= (|\Psi\rangle \langle\Psi|) \\ &= \frac{1}{2} (|V; morto\rangle \langle V; morto| + |O; vivo\rangle \langle O; vivo|) \\ &\quad + \frac{1}{2} (|V; morto\rangle \langle O; vivo| + |O; vivo\rangle \langle V; morto|) \end{aligned} \quad (98)$$

Nella base $|V; morto\rangle; |V; vivo\rangle; |O; morto\rangle; |O; vivo\rangle$ la matrice diventa

$$\rho = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (99)$$

La matrice densità contiene termini non nulli non solo lungo la diagonale (gli stati classici) ma anche fuori dalla diagonale (termini puramente quantistici di interferenza detti di coerenza). Con la matrice densità possiamo evidenziare i termini puramente di interferenza separandoli da quelli puramente classici. Ora il punto è che se questo sistema interagisce con un ambiente a molti stati (ad esempio l'aria nella scatola che il gatto respira e con cui interagisce termicamente) il fotone, il gatto e l'ambiente dovrebbero entrare in entanglement dopo un certo tempo

$$|\Psi\rangle = \frac{1}{\sqrt{2}} (|V; morto\rangle |A_1\rangle + |O; vivo\rangle |A_2\rangle) \quad (100)$$

dove $|A_1\rangle$ ed $|A_2\rangle$ sono stati dell'ambiente ortogonali ma non distinguibili sperimentalmente. In tal caso l'idea, poiché non possiamo distinguere A_1 da A_2 data la complessità ambientale, è quella di sommare sugli stati indistinguibili e questo ci porta alla fine, a causa dell'enorme numero di molecole di cui è costituito sia il gatto che l'ambiente, ad avere solo una miscela statistica analoga alla diagonalizzazione della matrice

$$\sum_{statik=1,2} |\Psi_k\rangle \langle \Psi_k| = \frac{1}{2} (|V; morto\rangle \langle V; morto| + |O; vivo\rangle \langle O; vivo|) \quad (101)$$

Insomma non vediamo sovrapposizioni di gatti perché tali effetti risultano così mescolati all'ambiente che non sono sperimentalmente distinguibili da miscele statistiche. Per poter evidenziare gli effetti di coerenza quantistica o l'interferenza in un ambiente così complesso occorrerebbero dei sistemi di misura troppo precisi. E' come se vedessimo in un certo senso una realtà sfocata che ci impedisce di osservare le sovrapposizioni. Va detto che questa interpretazione non è considerata soddisfacente da molti autori, i quali la reputano solo una "diavoleria" formale per nascondere sotto il tappeto problemi seri.

Conclusioni

Abbiamo mostrato in questo lungo viaggio a cavallo di un raggio di luce e attraverso le teorie classiche e quantistiche le idee fondamentali che abbiamo sviluppato sul mondo. Siamo passati dalla teoria geometrica a quella ondulatoria ed elettromagnetica. Abbiamo esplorato la relatività e visto i grandi enigmi lasciati aperti dalla fisica quantistica: entrambe teorie esatte del mondo, ma quanto sono diverse! Qual è oggi l'immagine del mondo che abbiamo noi fisici? E' una immagine coerente? Nel suo conciso ma profondissimo libro "sette brevi lezioni di fisica"

Carlo Rovelli dà una immagine molto concreta di qual è la situazione oggi: [13]

"Uno studente universitario che assista alle lezioni di relatività generale il mattino e a quelle di meccanica quantistica il pomeriggio non può che concludere che i professori sono citrulli, o hanno dimenticato di parlarsi da un secolo: gli stanno insegnando due immagini del mondo in completa contraddizione. La mattina il mondo è uno spazio curvo dove tutto è continuo; il pomeriggio, il mondo è uno spazio piatto dove saltano quanti di energia. Il paradosso è che entrambe le teorie funzionano terribilmente bene. La Natura si sta comportando con noi come quell'anziano rabbino da cui erano andati due uomini per dirimere una contesa. Ascoltato il primo, il rabbino dice «hai ragione». Il secondo insiste di essere ascoltato, il rabbino lo ascolta, e gli dice «hai ragione anche tu». Allora la moglie del rabbino, che orecchiava dalla cucina, urla: «ma non possono avere ragione entrambi!». Il rabbino ci pensa, annuisce, e conclude «hai ragione anche tu»."

Parole che dipingono in modo schietto e lucido lo status attuale della fisica. Tuttavia dalla fisica classica a quella quantistica possiamo concludere che il tutto indiviso è probabilmente la vera natura delle cose. Un tutto indiviso fatto di parti che non possono essere scorrelate tra di loro. C'è ancora tanto da fare e tanto da capire. In fondo, siamo solo agli inizi.



- [1] I. NEWTON, A CURA DI F. GIUDICE: *Scritti sulla luce e i colori*. BUR, Milano (2006).
- [2] J. C. MAXWELL: *A Treatise on Electricity and Magnetism*. Dover Publications, New York (1873).
- [3] A. EINSTEIN: "Zur Elektrodynamik bewegter Körper", *Annalen der Physik und Chemie* **17** (1905) 891-921.
- [4] A. EINSTEIN: *Emissione e trasformazione della luce, da un punto di vista euristico, tratto da Teoria dei quanti di Luce*. Edizioni Newton Compton, Roma (1972).
- [5] P. R. FEYNMAN: *QED: la strana teoria della luce e della materia*. Adelphi, Milano (2010).
- [6] H. EVERETT: "Relative state, formulation quantum mechanics", *Rev.Mod.Phys* **29** (1957) .
- [7] D. BOHM: "A suggested interpretation of the quantum theory in terms of hidden variables", *Phys. Rev.* **85** (1952) 166.
- [8] A. EINSTEIN, B. PODOLSKI, N. ROSEN: "Can quantum-mechanical description of reality be considered complete?", *Phys. Rev.* **47** (1935) 777.
- [9] J. S. BELL: "On the Einstein- Podolski – Rosen paradox", *Physics* **1** (1964) 195.

- [10] J. F. CLAUSER, M. A. HORNE, A. SHIMONY, R. A. HOLT: “Proposed experiment to test local hidden-variable theories”, *Phys. Rev. Lett.* **23** (1969) 880.
- [11] A. ASPECT, J. DALIBARD AND G. ROGER: “Experimental Test of Bell’s Inequalities Using Time-Varying Analyzers”, *Physical Review Letters* **49(25)** (1982) 1804.
- [12] W. H. ZUREK: “Decoherence and the Transition from Quantum to Classical—Revisited”, *Los Alamos Science* **27** (2002) 1-25. <http://vukuz.ru/books/zurek.pdf>
- [13] C. ROVELLI: *Sette brevi lezioni di fisica*. Piccola Biblioteca Adelphi, Milano (2014).



Marco Mazzeo: è ricercatore in Fisica sperimentale presso il Dipartimento di Matematica e Fisica “E. De Giorgi” dell’Università del Salento. Si occupa di nanofotonica e sviluppo di dispositivi molecolari che vanno dalla generazione di luce incoerente (OLED) e coerente (Laser) allo studio di condensati di Bose Einstein a temperatura ambiente in composti organici. Appassionato di storia della scienza e storia delle religioni.

Iride

Luigi Martina

Dipartimento di Matematica & Fisica "Ennio De Giorgi" - Università del Salento, INFN Sezione di Lecce

Senza grandi eccezioni, l'arcobaleno sembra essere fonte di emozioni positive: un'esperienza che rallegra, rassicura e ispira il senso del bello. Ma da dove traggono origine queste concezioni? Ed è sempre stato così? E come si sono evolute le interpretazioni scientifiche e quanto hanno inciso nel sentimento comune e nella cultura? Infine, quanto la spiegazione del fenomeno particolare ha fornito alla scienza strumenti e spunti per altri temi di ricerca?



Figura 1: Arco primario. Si notino i deboli archi sovrannumerari al di sotto del violetto

διὰ γὰρ τὸ θαυμάζειν οἱ ἄνθρωποι καὶ νῦν
καὶ τὸ πρῶτον ἠρξάντο φιλοσοφεῖν

In effetti, ora come in origine, gli uomini hanno iniziato a filosofare per la meraviglia.

(Aristotele, *Methaphysica*, A 2, 982b)

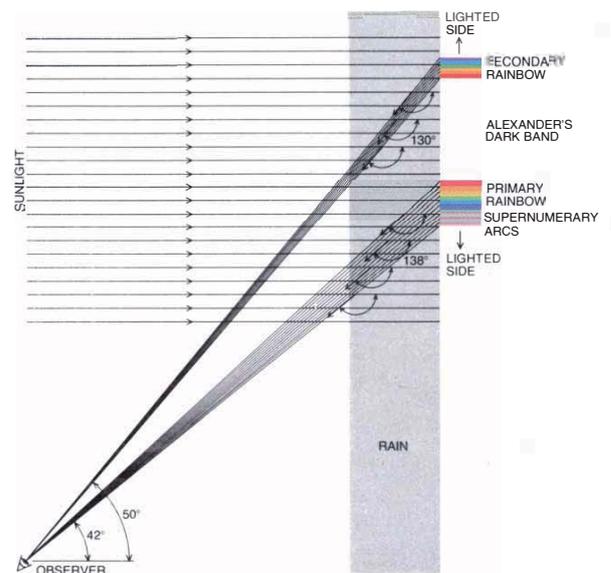


Figura 2: La geometria e le principali caratteristiche morfologiche dell'arcobaleno.



Figura 3: Arco primario e secondario. Gli archi sovrannumerari sono molto deboli, mentre è evidente la bassissima luminosità tra i due archi. Tale banda fu per la prima volta descritta dal filosofo Alessandro di Afrodisia, attorno 200 d.C., nei suoi Commentari alla Meteorologia di Aristotele.

Introduzione

If you want the rainbow, you have to deal with the rain.



Figura 4: Una iride senza pioggia, ma dovuta a minute goccioline o cristalli di ghiaccio che formano la nube

(Se vuoi l'arcobaleno devi sopportare la pioggia)

Hazel Grace Lancaster in *The Fault in Our Stars*
- John Green (2011)

Bambino, ancora impaurito dall'ultimo tuono, mi aggrappavo al vestito di mamma sull'uscio di casa. *Attento alle pozze, non bagnarti i piedi!*, ma il mio pensiero ormai correva lungo il profilo del blu, che saliva sempre più in alto, incurvandosi e sciogliersi assieme al rosso e al verde nel bianco sfolgorante di una nube, che ancora ribolliva in una battaglia tra Dragoni Neri e Candidi Pegasi.

Ancora oggi cerco di immaginare lo stupore di un bimbo (ma anche di un adulto) di 1.000 o 10.000 anni fa, o ancor prima, nelle savane africane, nelle foreste del Borneo o nella steppa asiatica mentre contempla la fine della tempesta con il distendersi tra le nubi di un'iride: cos'è, chi è, dov'è, cosa vuol dire? E' la stessa ricchezza di sfumature dell'arcobaleno fisico che diventano il paradigma delle sue interpretazioni e rappresentazioni. Chiedendo al vostro vicino cosa sappia sull'arcobaleno, molto probabilmente vi risponderà che compare quando piove nello stesso momento in cui il sole splende e che ha sette colori. Le risposte possono sicuramente variare e, a volte, ci possono essere spiegazioni più o meno lunghe sulla rifrazione e/o riflessione dei raggi solari da parte delle gocce di pioggia. In genere l'arcobaleno sembra essere fonte di emozioni positive: un'esperienza che allietta il cuore ed un simbolo di bellezza e pacificazione, anche se esistono significative discordanze

interpretative in alcune mitologie nordiche e mesopotamiche. Molto frequentemente l'immagine è associata a un'idea di varietà nell'unità. Tutte idee che vengono da molto lontano, nel tempo e nello spazio, e culture diverse le hanno declinate in maniera molto differente, ma sempre con quel fascino profondo, che ci ha ispirato ed è molto più dettagliatamente analizzato nell'opera di Lee e Fraser [1].

Certamente, dopo un acquazzone estivo, il fatto sensibile più notevole è l'arco primario, con l'attraente parata dei suoi colori, di intensità e nitidezza piuttosto variabile, ma sempre nello stesso ordine: violetto il più interno, che vira nel blu e gradualmente passa al verde, giallo, arancio e rosso all'esterno. Altre caratteristiche sono più delicate e, in effetti, non sono sempre presenti. Più alto nel cielo può comparire l'arco secondario: più debole in intensità e con l'ordine di esposizione dei colori invertito. I due archi delimitano una regione significativamente più oscura, detta banda di Alessandro. Altra caratteristica è la poco visibile serie di archi, in genere rosa e verdi, alternativamente, nella parte interna del primario. Tuttavia la loro spiegazione ha avuto una grande influenza nello sviluppo di una teoria dell'arcobaleno. La descrizione scientifica più accurata dell'arcobaleno non è però solo un problema di geometria, ma una sua teoria quantitativa soddisfacentemente completa è stata sviluppata, come si vedrà, solo in epoche relativamente recenti [2][3][4]. Essa attinge a tutto ciò che sappiamo sulla natura della luce, tramite la manifestazione delle sue proprietà ondulatorie, quali l'interferenza, la diffrazione e la polarizzazione, ma anche ricorrendo alle sue caratteristiche particellari, quali la quantità di moto trasportata da un fascio di luce. Per tale trattazione sono quindi necessari i più potenti mezzi della fisica matematica, tanto che l'arcobaleno è paradigmatico nella costruzione di vari modelli, quali la determinazione della distribuzione della luce nel cielo in ottica atmosferica [5][6], quanto nella descrizione di analoghi fenomeni di diffusione in fisica atomica e nucleare [7] [8].

Il Ponte di Dio

Negli ultimi esempi della sezione precedente diventa evidente l'afflato spirituale-religioso che si vuole trasmettere, anche se non necessariamente esso è incanalato nelle espressioni specifiche dei vari credi storicamente sviluppati. Quindi l'origine di questo connubio tra il fenomeno meteorologico e il suo significato metafisico deve essere ricondotto, almeno per quanto riguarda la tradizione giudaico-cristiana, al passo biblico [9]

Dio disse:

Questo è il segno dell'alleanza, che io pongo tra me e voi e ogni essere vivente che è con voi, per tutte le generazioni future.

Pongo il mio arco sulle nubi, perché sia il segno dell'alleanza tra me e la terra.

Quando ammasserò le nubi sulla terra e apparirà l'arco sulle nubi,

ricorderò la mia alleanza che è tra me e voi e ogni essere che vive in ogni carne, e non ci saranno più le acque per il diluvio, per distruggere ogni carne.

L'arco sarà sulle nubi, e io lo guarderò per ricordare l'alleanza eterna tra Dio e ogni essere che vive in ogni carne che è sulla terra.

Disse Dio a Noé: Questo è il segno dell'alleanza che io ho stabilito tra me e ogni carne che è sulla terra. (Gen. 8, 12-17)

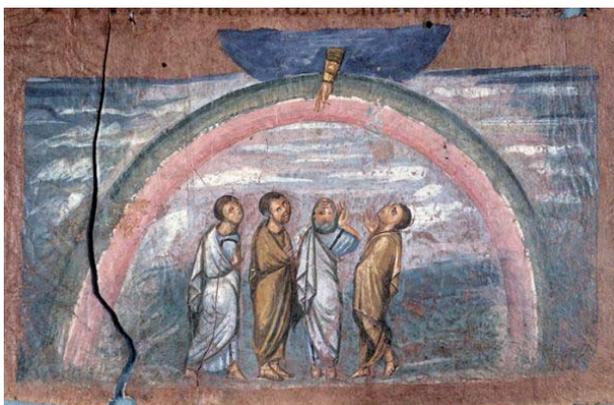


Figura 9: Noé e i suoi tre figli in piedi, a bocca aperta e addirittura contorti dalla meraviglia, osservano l'arcobaleno creato da Dio come segno di alleanza. Dal codice denominato Vienna Genesis (Österreichische Nationalbibliothek, cod. theol. gr. 31, Vienna, Austria). È il più antico codice biblico illustrato che si sia ben preservato. Si tratta di un frammento della versione greca dei Settanta, probabilmente redatto in Siria nella prima metà del VI secolo.

Questa speciale relazione tra arco e divino è ribadita in vari passi dell'Antico Testamento: *Osserva l'arcobaleno e benedici colui che lo ha fatto: bello nel suo splendore! Avvolge il cielo con un cerchio di gloria, lo hanno teso le mani dell'Altissimo. (Sir. 42,11-12),*

... Come sole sfolgorante sul tempio dell'Altissimo, come arcobaleno splendente fra nubi di gloria. (Sir. 40,5).

e infine

... Simile a quello dell'arcobaleno fra le nubi in un giorno di pioggia. Così percepii in visione la gloria del Signore. Quando la vidi, caddi con la faccia a terra e udii la voce di uno che parlava. (Ez 1,28).

Nell'Apocalisse il simbolo diventa una specifica caratteristica del divino e l'annuncio del regno di Dio:

Colui che stava seduto era simile nell'aspetto a diaspro e cornalina. Un arcobaleno simile nell'aspetto a smeraldo avvolgeva il trono. (Ap 4,3)

E vidi un altro angelo, possente, discendere dal cielo, avvolto in una nube; l'arcobaleno era sul suo capo e il suo volto era come il sole e le sue gambe come colonne di fuoco. (Ap 10,1)

Tuttavia, nella Bibbia non vi si trova alcun riferimento circa i colori dell'arcobaleno o al numero di essi, ma ovviamente permea gran parte della elaborazione teologica successiva. Certamente le citazioni apocalittiche possono spiegare il ruolo piuttosto popolare dell'arcobaleno nella pittura tardo gotica. Nell'arte romanica è possibile associare l'arcobaleno con la mandorla, o alone, che circonda Dio o un santo. Nell'arte gotica Cristo è spesso raffigurato seduto nella mandorla iridescente. Il più celebre esempio si ritrova nel Giudizio Universale di Giotto, nel quale il decoro del trono del Signore costituisce un tutt'uno con la Maiestas Domini. Pertanto tutti i personaggi sono rivolti verso il cerchio di luce iridata nel quale siede Gesù, che risulta il centro focale del dipinto e che con un semplice gesto della mano manifesta il suo giudizio, conseguenziale alla libera scelta dell'uomo. Sette secoli più tardi la scena è citata nel *Decameron* di Pasolini, dove ancora la mandorla, ora solo luminosa, racchiude la divinità e la separa dagli altri personaggi. Ma la peculiare religiosità di Pasolini, sostituisce Cristo con la Madonna (Silvana Mangano) col Bambino: il senso del giudizio si sostituisce con quello della pietà.



Figura 10: Giotto da Bondone: *Giudizio Universale*, Cappella degli Scrovegni, Padova

Probabilmente il più moderno e spettacolare riferimento al significato mistico dell'arcobaleno si ritrova nelle vetrate realizzate da J. Vila-Grau alla basilica della Sagrada Família di Gaudí a Barcellona.



Figura 11: La luce che penetra dalle vetrate della Sagrada Família si diffonde nelle navate di destra, con colorazioni che seguono l'ordine dell'iride.

Il Ponte del Tempo

*Come si volgon per tenera nube
due archi paralleli e concolori,
quando Iunone a sua ancella iube,*

*nascendo di quel d'entro quel di fori,
a guisa del parlar di quella vaga
ch' amor consunse come sol vapori;*

*e fanno qui la gente esser presaga,
per lo patto che Dio con Noé puose,
del mondo che già mai più non s' allaga:*

*così di quelle sempiterne rose
volgiensi circa noi le due ghirlande,
e sì l'estrema a l'intima rispuose.*

(Dante: *La Divina Commedia*, *Paradiso* XII, 10-21)

Come abbiamo visto l'arcobaleno è un ponte non solo tra gli uomini e Dio, ma inevitabilmente anche tra culture diverse: artisti e scienziati si sono sfidati a lungo per darne descrizioni diverse, se non a volte incompatibili. La conoscenza scientifica non sempre è stata accolta senza riserve. Goethe scrisse che l'analisi dei colori dell'arcobaleno data da Newton *paralizzerebbe il cuore della Natura*. Un sentimento simile fu espresso da J. Keats in *Lamia* del 1820. Nella *Teoria del colore* [10] la critica goethiana a Newton sottolineava la mancanza della fondamentale relazione tra luce e occhio dell'osservatore, che è la sede prima di trasformazione e assimilazione delle qualità luminose. Tale aspetto, per così dire, psicologico della visione è l'oggetto di maggiore interesse per il poeta. Di conseguenza la sua ricerca conduce ad una prospettiva ben diversa rispetto a quella quantitativa. Eppure gli scienziati che hanno contribuito alla la teoria dell'arcobaleno non erano affatto insensibili alla sua bellezza. D'altro canto Newton ebbe ammiratori anche tra gli artisti, in particolare J M W Turner, J F Overbeck e J Costable, che dipinsero arcobaleni con occhio fenomenologico, giudicandolo non solo il più bel fenomeno luminoso, ma un punto di partenza per lo sviluppo di una teoria del colore applicabile all'arte.

Il primo tentativo di fornire una spiegazione razionale dell'arcobaleno fu, probabilmente quella di Aristotele nel libro III della *Meteorologia*. Egli



Figura 12: P. P. Rubens: *Panorama con Arcobaleno*, 1636 circa, *The Wallace Collection, Londra*. L'arcobaleno è piuttosto innaturale come angolazione, anche se molto accurato per altri dettagli osservativi. Rubens abbandona i temi mistici per sottolineare la grande bellezza del creato e, allo stesso tempo, incoronare la semplicità della vita rurale. Certamente studiò in dettaglio il fenomeno naturale, ma lo ha trattato come un elemento puramente allegorico e compositivo

propose che l'arcobaleno fosse una speciale forma di riflessione della luce da parte delle nubi. Infatti la luce viene riflessa ad un angolo fissato, dando luogo ad un cono circolare di raggi. In questo modo Aristotele spiegò correttamente la forma circolare dell'arco, facendo comprendere inoltre che non si tratta di un oggetto con una ben definita posizione e dimensione, ma piuttosto un insieme di direzioni, lungo le quali la luce è fortemente diffusa verso gli occhi dell'osservatore. Quindi perchè l'iride si possa osservare deve esistere una precisa relazione tra la posizione relativa del Sole con le nubi e l'osservatore. Lo studio fu inoltre l'occasione per l'elaborazione di una prima teoria dei colori. Qui ci basti ricordare che sul numero dei colori dell'arcobaleno egli affermò essere tre: rosso, verde e blu (talvolta menziona il giallo) e la loro origine è fatta risalire alla mescolanza di luce bianca (omogenea e pura) e di buio. Ma nel complesso la scienza classica non era particolarmente interessata all'ottica, sebbene Euclide scrisse una *Ottica*, di Archimede si sa che scrisse una *Catottica*, e Tolomeo nella sua *Ottica* descrisse quantitativamente il fenomeno della rifrazione [11]. In effetti i fenomeni ottici dipendono fortemente dallo stato di illuminazione dell'oggetto, dalle condizioni di osservazione, dalla interposizione di mezzi diversi lungo il cammino ottico. Tutto in contrasto

con la ricerca di una verità ultima delle cose e quindi non un argomento consono agli interessi del filosofo. Solo il passaggio attraverso la cultura araba ed un nuovo interesse verso la natura e le tecniche permise, durante il tardo Medioevo, il rifiorire di interessi e metodi di carattere più osservativo della Natura [12].

Siamo di fronte ad un viaggio nel tempo di un corpo di osservazioni e conoscenze, che passa da Archimede agli arabi, poi a Witelo (≈ 1230 , ≈ 1280) e a Ruggero Bacone (≈ 1214 , 1294) [2][12]. In particolare un importante sviluppo all'ottica fu dato da al-Kindi, il principale esponente della scuola aristotelica araba (801 - 873) e Ibn al-Haytham (più noto come Alhazem, Bassora, 965 circa - Il Cairo, 1039). Quest'ultimo rivoluzionò il campo con i suoi studi sulla fisiologia dell'occhio, l'invenzione della camera oscura e la prima descrizione teorica della rifrazione. Infine sostenne l'abbandono totale del meccanismo di visione degli $\epsilon\iota\delta\omega\lambda\alpha$ aristotelici. In particolare compose un *Trattato sull' arcobaleno e l'alone* nel quale riporta l'osservazione *Poiché il loro soggetto è l'aria, la loro investigazione deve essere fisica, e essendo la loro forma circolare, devono essere studiate matematicamente*, che risuonano in qualche modo con le idee galileiane. Nel suo trattato *Ottica* [13] tratta il fenomeno della rifrazione con grande dettaglio, riportando misure delle relazioni angolari sia all'interfaccia aria-acqua, che aria-vetro. Realizzò anche il primo studio di come un fascio di luce viene deflesso da un globo di vetro riempito di acqua. Tuttavia non sembra che egli abbia collegato questo esperimento con un modello di goccia illuminata dal Sole. Queste opere ebbero difficoltosa diffusione anche all'interno del mondo islamico, ma tramite l'opera di traduttori sia dal greco che dall'arabo, per tutti si ricordi Gerardo da Cremona (Cremona 1114, Toledo 1187), gli eruditi europei recuperarono entusiasmo per lo studio dei fenomeni naturali, tra i quali l'arcobaleno occupava un posto non secondario, divenendo un esempio sul quale provare e confrontare teorie ottiche. Il tortuoso cammino che fu percorso fino ad una trattazione moderna dell'argomento suggerisce come le credenze e le personalità possono giocare un ruolo cruciale nello sviluppo delle Scienze. Da questo punto di vista l'arcobaleno costituisce effettivamente un ponte che lega epoche differenti.

Ispirato dagli studi di Witelo [13], attorno al 1266 Ruggero Bacone misurò l'angolo di vista dell'arco primario rispetto alla luce solare incidente: circa 42° . Il secondario si trova a 8° più in alto. Attualmente è più comune riferirsi ad essi in termini di *angoli di deflessione* del fascio luminoso, quindi 138° e 130° rispettivamente.

Nel 1304 nel *De iride* il monaco e teologo Teodorico di Friburgo suggerì che ogni singola goccia di pioggia contribuisce alla formazione dell'arcobaleno, mostrando sperimentalmente che poteva riprodurlo tramite una bottiglia sferica riempita di acqua. Mostrò che la luce del primario aveva subito una sola riflessione interna, mentre il secondario due. Notando inoltre che lungo ogni direzione di vista solo un colore alla volta può essere osservato, concluse che ciascuno dei colori dell'iride proviene da un differente insieme di gocce di pioggia. Tuttavia, pose le gocce su una supposta sfera meteorologica. Chi volesse

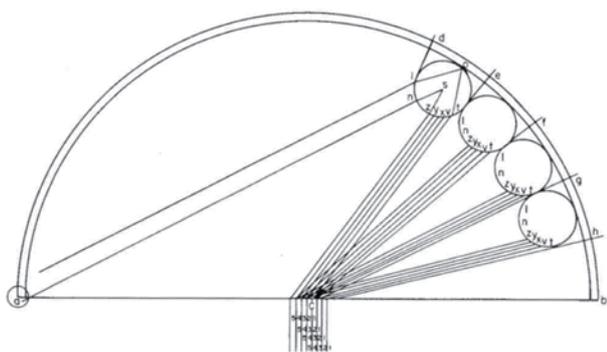


Figura 13: Il modello di arcobaleno di Teodorico di Friburgo

oggi riprodurre a livello scolastico l'esperienza di Teodorico potrebbe seguire le indicazioni di R Bagnolesi [15]. Gran parte di questo lavoro dovette essere riscoperto più volte nei trecento anni successivi, con un lento ma progressivo arricchimento di osservazioni. In particolare si ricordino quelle, allo stesso tempo scientifiche e pittoriche, di Leonardo da Vinci [1][2][11][12][14][17]. Il problema teorico originario non fu mai completamente abbandonato, ma ad intermittenza ricevette significativi contributi alla sua soluzione da parte di notevoli scienziati, tra i quali il matematico messinese Francesco Maurolico (1494 - 1575), noto per i calcoli di geometria delle masse, Marco Antonio De Dominis (1566-1624) (*De radiis visus et lucis* [19], unica, e sorprendente citazione di Newton in *Optiks*), e J. Kepler (1571

- 1630) [14][16]. Il primo introdusse in *Photismi* [18] una teoria dell'arcobaleno basata solo su riflessioni multiple all'interno di una goccia d'acqua (evidentemente Alhazen e Teodorico erano stati dimenticati). Le gocce per produrre l'arco dovevano però trovarsi su un preciso piano. Chiaramente il modello non poteva che condurre ad artificiose spiegazioni sul valore degli angoli di deflessione e sulla teoria dei colori (che per Maurolico erano sette). Keplero, nei suoi studi di ottica raccolti in *Paralipomena ad Vitellionem* prima e in *Dioptrice* poi [20], era arrivato ad una buona conoscenza quantitativa della rifrazione. I suoi risultati erano noti a Galileo, che li utilizzò nel *Nuncius Sidereus* [21] per la costruzione del cannocchiale. Lo spirito misticheggiante di Keplero lo condusse a formulare una spiegazione dell'arcobaleno basata sull'idea che solo i raggi luminosi tangenti alla superficie possono assumere colori prismatici. Usando una forma approssimata della legge di rifrazione calcolò l'angolo di deflessione dei raggi solari che formano l'arco primario a 135° , che è abbastanza vicino al vero, ma comunque ancora insoddisfacente per lo stesso Keplero.

Con Cartesio [22] *l'arcobaleno è una tale notevole meraviglia della Natura . . . che non riuscivo a scegliere un esempio più adatto per l'applicazione del mio metodo.*

Ovviamente si riferiva al metodo sviluppato nel *Discours de la méthodes* e da lui applicato nel trattatello *De l'arc-en-ciel* (1637), facente parte de *La dioptrique*. Il punto chiave è un enunciato corretto della legge della rifrazione in termini di costanza del rapporto dei seno dell'angolo di incidenza con quello di rifrazione.

Rilevando che solo le gocce che si trovano in una ben definita relazione angolare con l'osservatore determinano il colore che eventualmente si può vedere, fanno superare tutte le difficoltà nella comprensione della geometria dell'arcobaleno, quali erano ancora in Teodorico e in Maurolico. Ma il punto cruciale ancora non è raggiunto, in quanto raggi luminosi che incidono sulla sfera d'acqua in punti differenti debbono essere difusi ad angoli differenti, quindi cosa hanno di speciale gli angoli di 42° e 50° e i loro vicini? Cartesio intraprende una lunga serie di calcoli che gli permettono di studiare i cammini di molti raggi paralleli che intersecano una sfera e scrive

La legge di rifrazione

Riflessione e rifrazione all'interfaccia tra aria e acqua sono i fenomeni fondamentali nella creazione di un arcobaleno. Nella riflessione l'angolo di incidenza è uguale all'angolo di riflessione. Nella rifrazione l'angolo del raggio trasmesso è determinata dalle proprietà del mezzo, in quanto caratterizzati da suo indice di rifrazione. La luce entrante in un mezzo con indice più alto è piegato verso l'asse normale alla superficie. La formulazione matematica dovuta a Cartesio (ma già ottenuta da Snell) è

$$\frac{\sin i}{\sin r} = n,$$

dove n è chiamato *indice di rifrazione relativo* ai due mezzi. Esso può essere espresso come rapporto tra due indici *assoluti*, cioè riferentesi alla rifrazione della luce dal vuoto ad un dato mezzo. Nel passaggio aria - acqua $n_{A \rightarrow Acq} \approx 1.33$, mentre nell'attraversamento inverso $n_{Acq \rightarrow A} = \frac{1}{n_{A \rightarrow Acq}} \approx 0.75$.

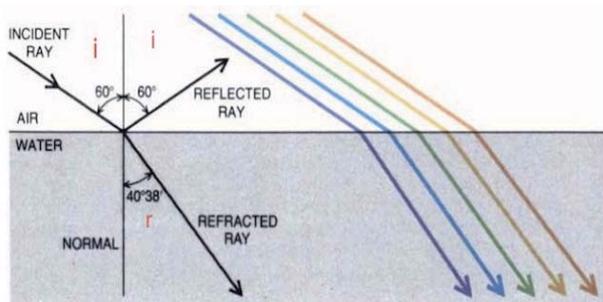


Figura 14: La rifrazione della luce tra aria e acqua. Poiché l'indice di rifrazione diminuisce nel passare dal violetto al rosso (vedi Figura 14), i raggi del primo colore sono più deviati di quelli del secondo.

che molti più possono essere visti entro un angolo tra 41° e 42° di quanti ad angoli minori, e che nessuno può essere visto ad angoli maggiori. Analogo calcolo esegue per l'arco secondario, trovando l'angolo di visuale rispetto all'asse antisolare tra 51° e 52° . Oggi sappiamo che T. Harriot aveva eseguito calcoli simili un decennio prima, ma non aveva mai pubblicato nulla in proposito [2]: una vicenda molto simile a quella di Snell, che

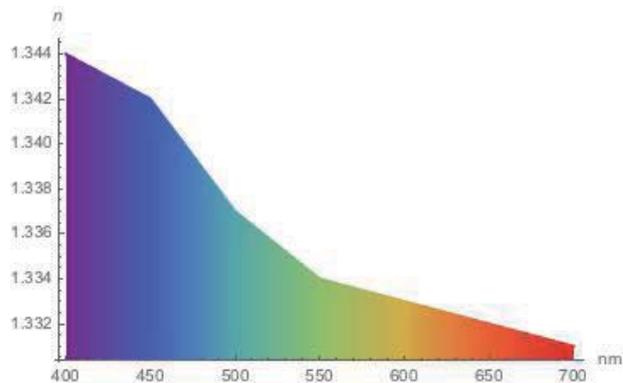


Figura 15: La dipendenza dell'indice di rifrazione dell'acqua dalla lunghezza d'onda nella regione del visibile. Valori dell'indice di rifrazione e della sua componente complessa su uno spettro molto ampio si possono ricavare in <http://refractiveindex.info>

può far riflettere sulla dinamica delle scoperte scientifiche e la loro attribuzione.

I lunghi e faticosi calcoli furono grandemente semplificati da Newton, grazie al suo potente metodo di calcolo dei massimi e minimi di una funzione. Come si può osservare dalle Figure 15 - 17 l'angolo di deflessione (cioè l'angolo tra le semirette definite dal fascio incidente e da quello rifratto) di un raggio che subisce p riflessioni interne ha il valore

$$D_p = 2i - 2(p + 1)r + p\pi.$$

Chiaramente un osservatore con il Sole alle spalle vedrà una maggiore intensità luminosa se esistono famiglie di raggi solari, che subiscono pressoché una medesima deflessione $\frac{\pi}{2} \leq D_p \pmod{2\pi} \leq \frac{3\pi}{2}$. In altri termini la funzione D_p deve variare poco attorno a certi valori dell'angolo di incidenza sulla goccia, il che avviene nei punti di massimo o di minimo per D_p . In corrispondenza di questi angoli si concentra più luce che altrove e, per questo motivo, le superfici (coni) da essi definite vengono chiamate *caustiche*. Una familiare caustica è la brillante curva cuspidale che si forma in una tazza, quando la luce del sole si riflette sulle sue pareti interne. L'osservatore dell'arcobaleno vedrà questa maggiore luminosità a certi angoli di vista V_p , che per $p = 1, 2$ sono $V_1 = 4r - 2i$ e $V_2 = 2i - 6r + \pi$, come si ricava dalle Figure 19-20.

Se la goccia è una sfera, allora dalla Figura 18 è facile vedere che l'angolo di incidenza di un

La dispersione della luce

Newton [23] scoprì che raggi di colore diverso sono differentemente rifrangibili, propagandosi da un mezzo all'altro. In termini moderni la *dispersione cromatica* è causata dalla dipendenza dell'indice di rifrazione dei mezzi ottici dal colore della luce. Dalla legge di rifrazione si ricava che, a parità di angolo di incidenza i , l'angolo di rifrazione r diminuisce all'aumentare dell'indice n .

Nell'ambito della teoria ondulatoria, la nozione di colore viene sostituita dalla grandezza fisica *lunghezza d'onda* λ o, tramite la *legge di dispersione della luce nel vuoto*, dalla frequenza $\nu = \frac{c}{\lambda}$, $c =$ velocità della luce.

Con Maxwell, l'ottica della luce visibile rientra tra i fenomeni elettromagnetici, che si collocano nell'intervallo di lunghezze d'onda $\sim 400 - \sim 700$ nm dello spettro (Figura 16).

L'indice di rifrazione assoluto di un mezzo è $n(\lambda) = \frac{c}{v} = \frac{\text{velocità della luce}}{\text{velocità della luce nel mezzo}}$, che è detta *la legge di dispersione nel mezzo*. Questa funzione dipende sensibilmente dalle proprietà molecolari del mezzo, spesso con una serie di *picchi di risonanza*. Formule (di Sellmeier) semi-empiriche per $n(\lambda)$ si derivano abbastanza semplicemente [25], ma calcoli *ab initio* richiedono l'uso della *Meccanica Quantistica*

raggio luminoso su di essa è dato da $\sin i = \frac{b}{R}$, detto parametro d'urto essendo $0 \leq b \leq R$ la distanza del punto di incidenza dall'asse della sfera ed R il suo raggio. Dalla legge di rifrazione si ha che l'angolo di vista V_1 è legato al parametro d'urto dalla relazione

$$V_1(b) = 4 \arcsin \frac{b}{nR} - 2 \arcsin \frac{b}{R}.$$

Lo studio di questa funzione può essere condotto con i metodi abituali dell'analisi, riconoscendo che essa ha un solo massimo assoluto in

$$\frac{b_{max}}{R} = \sqrt{\frac{4 - n^2}{3}},$$

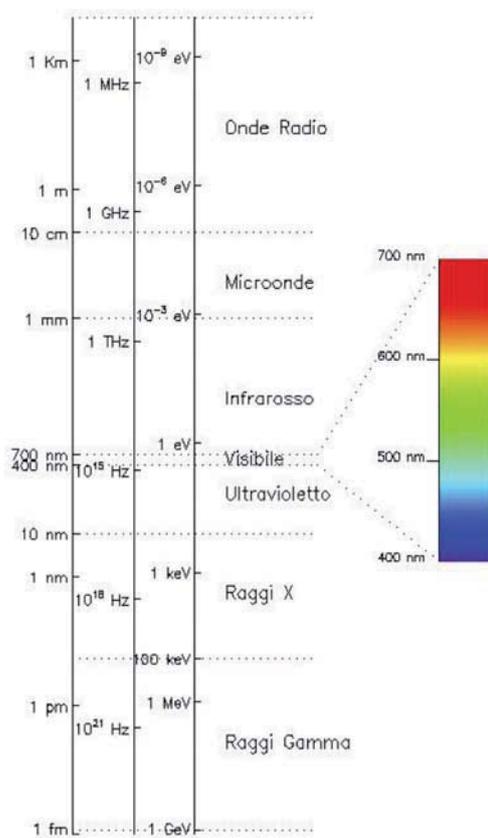


Figura 16: Lo spettro della radiazione elettromagnetica.

che per l'indice di rifrazione dell'acqua $n \approx 1.33$ fornisce il valore dell'angolo di vista $V_{1max} = 42.5^\circ$. Calcoli simili per l'arco secondario conducono a $V_{2max} = 50^\circ$.

Inoltre, poiché non ci sono raggi che, subendo 1 o 2 riflessioni interne, si diffondono tra V_{1max} e V_{2max} , anche la banda oscura di Alessandro viene spiegata (Figura 21), essendo così delimitata dai bordi oscuri di due caustiche. In generale l'intensità luminosa cresce continuamente da una parte di una caustica, per poi cadere bruscamente.

In linea di principio si potrebbero trovare infiniti arcobaleni, ma la loro intensità decresce esponenzialmente, su una base di circa 0.02, con il numero delle riflessioni interne. Inoltre può accadere, come ad esempio per $p = 3, 4$, che relativamente alla direzione di incidenza l'angolo di diffusione sia piccolo ($-\frac{\pi}{2} \leq D_p \pmod{2\pi} \leq \frac{\pi}{2}$), pertanto l'arcobaleno circonda il Sole, rendendolo ancora più difficilmente visibile (<http://www.atoptics.co.uk/rainbows/orders.htm>).

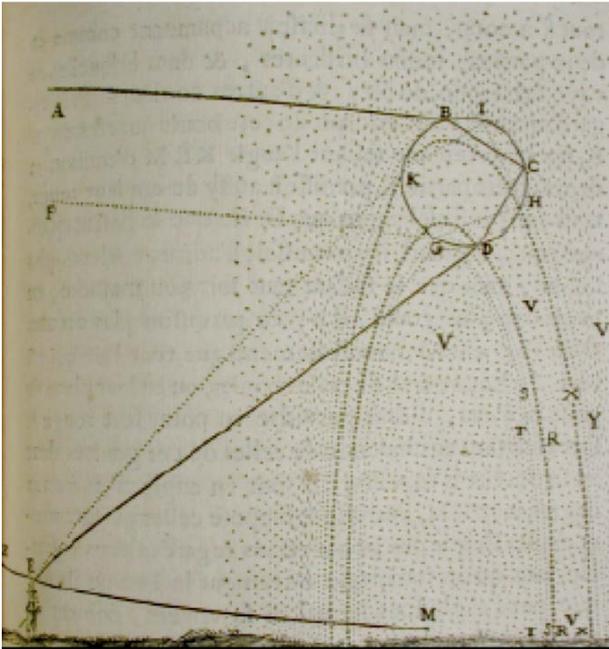


Figura 17: Il modello di Cartesio in *De l'arc-en-ciel*. Ogni singola goccia di pioggia illuminata dal sole, dopo un percorso dovuto a rifrazione-riflessione-rifrazione, devia la luce verso l'osservatore ad un particolare angolo, indipendentemente dalla posizione nello spazio. La simmetria circolare nel processo di diffusione della luce spiega la forma dell'arco attorno al punto antisolare M. Inoltre l'arco secondario segue un percorso con due riflessioni interne.

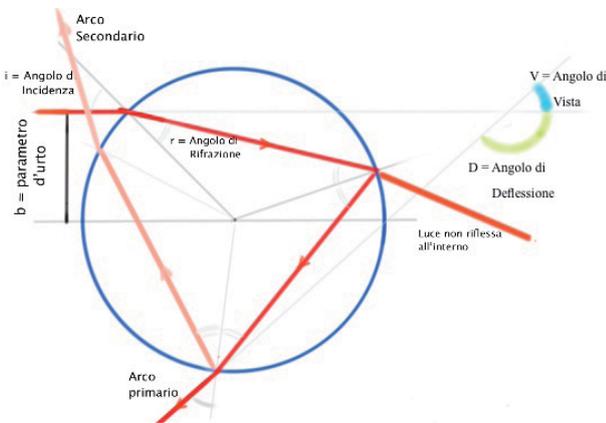


Figura 18: Lo schema di diffusione di un fascio di luce all'interno di una goccia.

Il Ponte della Conoscenza

La teoria di Cartesio è completamente concentrata sull'aspetto geometrico: il suo è un arcobaleno senza colori. Cartesio ne era consapevole, ma la sua teoria dei colori non ebbe la forza di sostituire quella aristotelica nella quale, comunque, l'origine dei colori era piuttosto confusa, riconducendo-

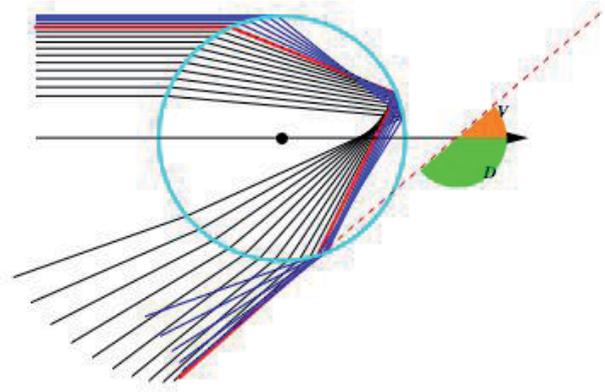


Figura 19: La costruzione di Cartesio per la determinazione dell'arco primario. In rosso è indicato il raggio limite di formazione della caustica. La figura indica chiaramente che dalla parte inferiore della caustica c'è oscurità completa.



Figura 20: Caustica sul fondo di una tazza.

la a mescolanze di luce bianca e di oscurità. Tra i colori si distinguevano tre *primari*, perchè con essi i pittori ne possono produrre gli altri, secondo i tre metodi: accostamento, giustapposizione e mescolamento. La diatriba sul numero e la natura dei colori proseguì per tutto il Medioevo, ponendoli in relazione addirittura con la Trinità oppure con gli Elementi empedoclei. Anche la scelta di Newton di sette colori primari era determinata da motivazioni non strettamente fisiche, ma di analogia con le note musicali. Ma già nella sua lettera al *Phylosophical Transactions of the Royal Society* 1671-72 sulla Teoria della Luce e dei colori, era chiarissimo che i colori fossero *i costituenti* della luce bianca e che quelli identificabili non si limitano ai più eminenti, ma anche tutte le loro gradazioni intermedie ... *cosicché tutti colori intermedi, in una successione continua, appartengono a gradi di rifrangibilità intermedi* [23, 24]. Quattro aspetti

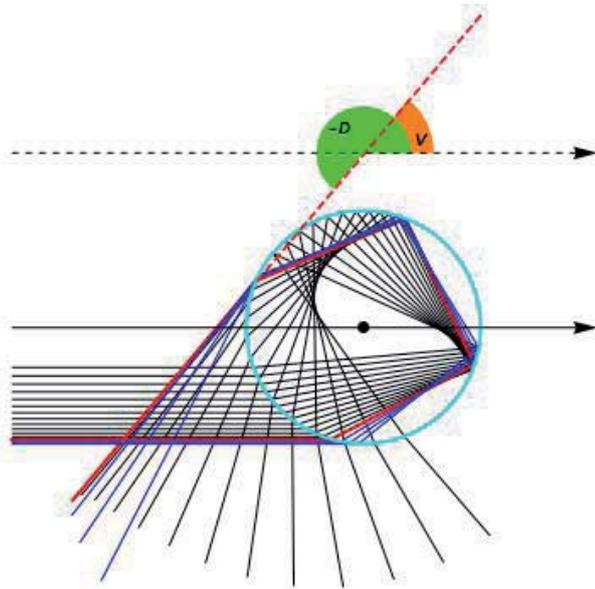


Figura 21: La costruzione di Cartesio per la determinazione dell'arco secondario

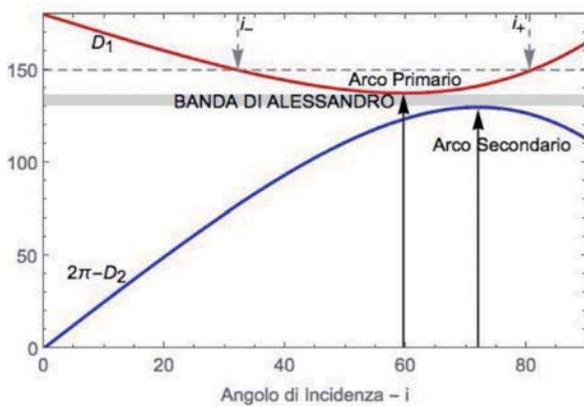


Figura 22: Gli angoli di deflessione di un raggio incidente sulla goccia con angolo i . L'angolo di deflessione per il secondario è stato traslato di 2π . Gli angoli di deflessione assumono un minimo/massimo in corrispondenza della formazione degli arcobaleni primario/secondario, facendo convergere infiniti raggi incidenti in quelle particolari direzioni, dette caustiche. Si osservi che nessun raggio del primario o del secondario è deflesso nell'intervallo $129^\circ - 137^\circ$, dando luogo alla banda di Alessandro. Infine i_-, i_+ sono una coppia di angoli di incidenza che hanno lo stesso angolo di deflessione: i corrispondenti raggi si propagano parallelamente.

appaiono evidenti nell'approccio newtoniano: 1) i colori spettrali (o primari) sono una connaturale caratteristica continua della luce, essi non possono essere ridotti ulteriormente; 2) i colori primari possono essere separati usando la rifrazione della luce, ad esempio usando un prisma



Figura 23: P. P. Rubens: *Giunone e Argo*, 1610-11, Colonia, Wallraf - Richartz Museum. Di ritorno dall'Italia il giovane Rubens compone un'opera pedagogica, applicando la teoria del colore di Aristotele alla pittura barocca. La dea è aiutata dalla sua serwa Iride, nel ritrovare i cento occhi del volto del fedele guardiano Argo, ucciso da Mercurio, e collocarli sulla coda del pavone. Iride è rappresentata sia come fenomeno meteorologico che personificata.

costituito da materiale dispersivo (quindi anche da una goccia di pioggia), ma non possono essere ulteriormente mutati da riflessioni o rifrazioni; 3) i colori sono infiniti, anche se per una ipotetica relazione con le strutture musicali Newton ne sottolinea sette: rosso, arancio, giallo, verde, blu, violetto-porpora e indaco, 4) da due o più colori primari si possono ottenere colori composti, che non appartengono allo spettro, in particolare il bianco. Tutte proprietà in netto contrasto con la fisica aristotelica, che applicate direttamente al caso dell'arcobaleno, non solo ne spiegano l'origine, ma danno una ragione della larghezza angolare degli archi. Infatti usando i due diversi valori (Figura 15) dell'indice di rifrazione per il violetto (il colore più deflesso) e quello per il rosso (il meno deflesso) in V_1 , si ottiene la larghezza angolare $\Delta D_1 = 1^\circ 45'$. Ammettendo $\frac{1}{2}^\circ$ per il diametro apparente del Sole, Newton ottenne $\hat{\Delta}D_1 = 2^\circ 15'$ per il primario, in accordo con le sue osservazioni. Naturalmente, questa è l'apertura angolare delle caustiche di differenti colori che emanano da una singola goccia. Quindi un singolo osservatore viene raggiunto da caustiche di colore diverso provenienti da gocce diverse, viste nell'apertura angolare calcolata sopra. L'ar-

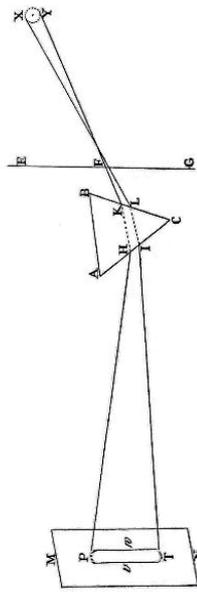


Figura 24: Disegno di Newton che illustra la scoperta della scomposizione spettrale della luce (*Opticks, Part II, prop. III Prob. I.*)

cobaleno ha molto di un mosaico, le tessere del quale sono una miriade di goccioline, che riflettono e rifrangono la luce. Questo fatto è molto ben visibile se si realizza un arcobaleno casalingo con lo spruzzo di un irrigatore da giardino. Una immagine così costituita si dice che è ottenuta per *mescolamento additivo di colori*. Esso si può realizzare 1) sovrapponendo luci di colore diverso, 2) colorando piccole aree giustapposte con colori diversi. Esattamente la procedura adottata negli schermi televisivi o dei computer. Il colore che si ottiene è sempre *più brillante* dei suoi componenti, perché l'occhio, appunto, somma la brillantezza delle diverse sorgenti di luce. A complemento, c'è il cosiddetto *mescolamento sottrattivo*, l'unico concretamente realizzabile secondo Aristotele attraverso la miscela di pigmenti diversi. In questo caso, ogni mistura di colori è *più scura* dei suoi componenti e alla fine, facilmente, si raggiunge il nero. Osservazioni queste che hanno enormemente influenzato il dibattito sulla costruzione di una moderna teoria del colore. Per oltre un secolo, dalla comparsa dell'*Opticks*, artisti e

teorici dell'arte hanno entusiasticamente accolto, appassionatamente respinto, dettagliatamente criticato e, talvolta, incompreso le provocatorie idee di Newton sui colori [26]. Ma il progresso più significativo non vennero dalle polemiche, né da eleganti schemi pittorici, ma dal lavoro di fisici e fisiologi, quali Thomas Young (1773-1829), Hermann von Helmholtz (1821-94) e James Clerk Maxwell (1831-79). Essi iniziarono a riconciliare l'esperienza visiva artistica [27], ma anche quotidiana, con il modello fisico della luce e a formare quella comprensione che abbiamo oggi della percezione dei colori [28]. Ma non ci soffermiamo di più su un argomento, che ci porterebbe troppo lontano.



Figura 25: G Seurat (1859-91): *Una domenica pomeriggio alla Grande Jatte* (1884-86), Art Institute of Chicago. Seurat, teorizzò e esercitò la "peinture optique", ridefinendo scientificamente il suo obiettivo artistico al fine di ottenere particolari effetti ottici. La sua arte è una felice combinazione di premeditata costruzione accademica e di Impressionismo. In questa opera Seurat sviluppa completamente la tecnica del puntinismo, che consiste nel dipingere una serie di piccole macchie, brevi tratti e pennellate separate. Differentemente dalla tecnica della stampa a colori (piccoli punti di colore, regolari in dimensione e posizione), quella di Seurat crea una complessa miscela di tratti di vario colore, dimensione ($\approx 2\text{ mm}$) e direzione. Tanto che ogni parte del dipinto sembra vibrare o debolmente luccicare. E' ipotizzabile che in questo intervenga non solo il fenomeno della giustapposizione additiva, ma che si instauri un disequilibrio nella risoluzione dei tratti di colore, a causa della differente acuità visiva tra centro e periferia dell'occhio umano.

Il Ponte di Onde

Newton non procedette nella direzione di una teoria ondulatoria della luce propugnata, invece, da Huygens [29], ma confermata definitivamente solo dall'esperienza della doppia fenditura di Young [30] [31]. Da un punto di vista ondulatorio la luce può essere descritta da una combinazione (sovrapposizione) lineare di onde sinusoidali. A ciascuna lunghezza d'onda, cioè la distanza tra due creste consecutive (o comunque tra due punti con la stessa fase oscillatoria), corrisponde un singolo colore spettrale. Come si evidenzia nella Figura 16, le lunghezze d'onda della luce visibile sono comprese tra 400 nm (violetto) e 700 nm (rosso), con una buona dose di approssimazione. Misure di queste lunghezze furono portate ad una grande accuratezza solo dopo l'invenzione dello spettroscopio e del reticolo di diffrazione da parte di J Fraunhofer negli anni 1814-20 [32].

D'altra parte, con l'esperimento della doppia fenditura egli aveva chiaramente dimostrato che la luce subisce interferenza. Questo fenomeno, grazie alla sovrapposizione di onde in relazione di fase costante, consiste nel rafforzamento delle creste, oppure la loro cancellazione, con la formazione di figure regolari di frange luminose alternate da regioni di buio. Young dimostrò l'interferenza di onde di luce, facendo passare un singolo fascio di luce monocromatica attraverso due fori in uno schermo e osservando le frange prodotte al di là dello stesso schermo.

Ma, in particolare, proprio applicando il modello ondulatorio della luce [31] egli riuscì a dare una spiegazione degli archi sovrannumerari. Questi sono effetti di complemento all'arco primario, ai quali nell'arte e nel mito, non ci si era molto interessati (come del resto all'arco secondario o agli archi monocolori). Rare descrizioni di questi fenomeni risalgono a Witelo e all'attento abate Teodorico di Friburgo.

Young osservò che le coppie di raggi di luce (i_- , i_+ nella Figura 21), diffusi nella stessa direzione D_1 da una goccia di pioggia, sono equivalenti ai due raggi prodotti dai due fori del suo esperimento originario. Ad angoli molto vicini a quello della caustica i due cammini differiscono di molto poco, così i due raggi interferiscono costruttivamente, perché i contributi al campo luminoso si sommano. Al crescere della diffe-

renza angolare, i due raggi seguono cammini di lunghezza ben diversa. Quando la differenza di cammino ottico uguaglia la metà della lunghezza d'onda, l'interferenza è completamente distruttiva, in quanto ad ogni cresta in uno dei raggi corrisponde una valle nell'altro, producendo un campo totale nullo. Ad angoli ancora maggiori i due raggi riprendono a rinforzarsi di nuovo, fino a raggiungere il massimo di intensità per una differenza di cammino pari ad una lunghezza d'onda. Il risultato è una periodica oscillazione angolare dell'intensità della luce diffusa, i massimi della quale corrispondono a differenze di cammini ottici pari a multipli interi della lunghezza d'onda. Di conseguenza la figura de-

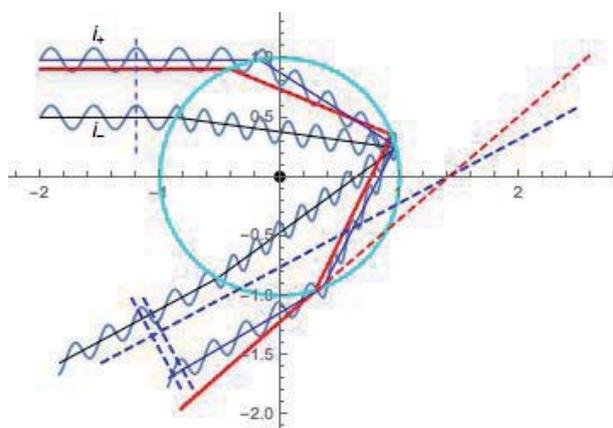


Figura 26: Schema della spiegazione di Young degli archi sovrannumerari. Benché l'onda incidente abbia una ben definita fase, il diverso cammino ottico, relativo a ciascuno dei due raggi con diverso angolo di incidenza, sfasa le due componenti ondulatorie. Lo sfasamento è costruttivo solo se è un multiplo intero della lunghezza d'onda della luce incidente. Altrimenti l'ampiezza è minore, fino ad annullarsi quando lo sfasamento corrisponde a mezza lunghezza d'onda.

gli archi sovrannumerari così formati dipende significativamente dalle dimensioni delle gocce. Questo perché la differenza di cammino ottico cresce rapidamente con il raggio della sfera, oltre che dipendere dal parametro d'urto. Quindi più sono grandi le gocce, più è stretta la separazione angolare tra gli archi, che diventano difficilmente distinguibili se le gocce hanno $R \geq 1$ mm. La sovrapposizione dei colori tende così a espandersi su tutto l'arco. Questo spiega pure perché essi si possono vedere più facilmente verso il culmine dell'arco, quando le gocce sono ancora piccole,

mentre si dissolvono in basso, tendendo le gocce a ingrandirsi nella caduta.

Quindi sembrava che Young riuscisse a spiegare tutte le principali caratteristiche dell'arcobaleno con la teoria dell'interferenza della luce. Ma, come la teoria di Descartes - Newton, anche quella di Young prevedeva la completa oscurità della banda di Alessandro, il che non solo è falso, ma un tale brusco passaggio non è coerente con la stessa teoria ondulatoria. Questa prevede che bordi netti tra luce e ombra siano sempre sfumati a causa del fenomeno della diffrazione.

Questo fatto, la più familiare manifestazione del quale è il cambiamento di direzione di propagazione della luce (ma anche del suono e di ogni altro tipo di onda) in prossimità del bordo di un ostacolo, fu studiato per la prima volta da F M

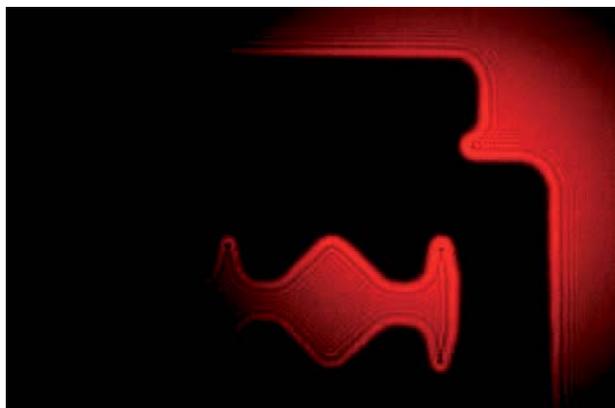


Figura 27:

Grimaldi [33], ripreso dallo stesso Newton e poi, nell'ambito della teoria ondulatoria, approfondito da Young ed espresso in forma matematica moderna da Fresnel [34]. Nell'arcobaleno non ci sono ostacoli, ma attorno alle caustiche, che definiscono l'arco primario e il secondario, si dovrebbero manifestare degli effetti diffrattivi, con il corrispondente smussamento delle variazioni nell'intensità della luce.

Nel 1838 il noto astronomo inglese G B Airy fece il tentativo [37] di determinare la distribuzione dell'intensità luminosa in prossimità dell'arcobaleno, basandosi sul cosiddetto *principio di Huygens*, poi elaborato da A J Fresnel [34] (Figura 26). Successivamente Kirchhoff derivò il principio di Huygens-Fresnel dalle equazioni di Maxwell [35], almeno limitatamente alla teoria scalare. L'intero argomento dell'ottica come fenomeno elettromagnetico è autorevolmen-

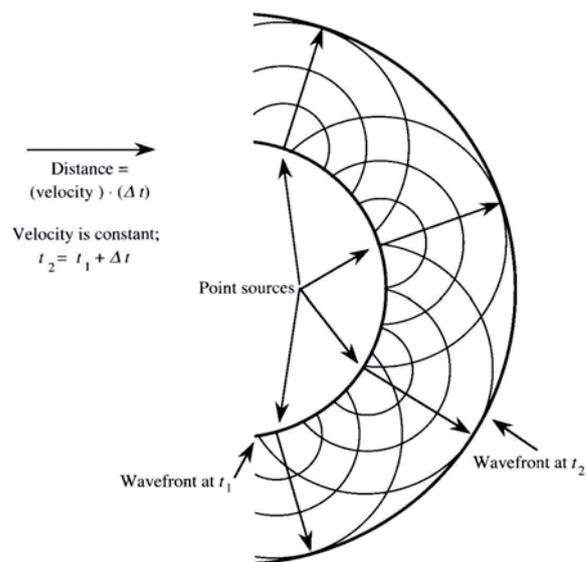


Figura 28: Schematizzazione del Principio di Huygens. In un mezzo omogeneo e isotropo, all'istante t_1 ciascun punto di un fronte d'onda (luogo dei punti di uguale fase) si comporta come una sorgente puntiforme secondaria di onde sferiche di opportuna ampiezza. Trascorso un intervallo di tempo Δt , il nuovo fronte è dato dall'involuppo di tutti i fronti secondari sferici di raggio $\Delta r = v \Delta t$ e l'ampiezza dell'onda si ottiene sommando tutte le ampiezze delle singole perturbazioni. In un punto di intersezione tra il fronte d'onda e l'interfaccia tra mezzi con indice di rifrazione differente, emergono onde secondarie sferiche, che nelle due bande si propagano con le rispettive velocità.

te trattato da Born e Wolf [36]. Il principio di Huygens consentirebbe di calcolare l'ampiezza della radiazione in ogni punto dello spazio, partendo dalla conoscenza dei suoi valori su uno specifico fronte d'onda. Così l'intero arcobaleno potrebbe essere descritto rigorosamente, se si conoscesse l'intensità luminosa su un particolare fronte d'onda entro una singola goccia. Sfortunatamente, questo è impossibile, ma alcune ipotesi ragionevoli potrebbero fornire una buona approssimazione. Airy scelse una superficie ortogonale a tutti i raggi all'interno della goccia vicini quello di Descartes, proseguendo in maniera simmetrica al di là di esso con un punto di inflessione (cioè un cambio nel senso della curvatura). Questo è un *fronte virtuale*, nel senso che non è quello effettivo all'interno della goccia, ma è quello che in aria e dalla parte dell'arco primario produrrebbe due fronti che si

sovrappongono e interferiscono. Orientando gli

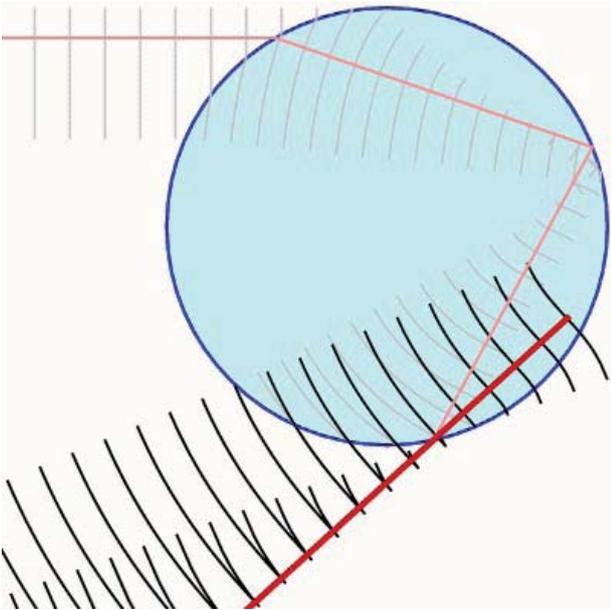


Figura 29: Il fronte d'onda di Airy e la sovrapposizione a sinistra di due componenti ondulatorie che interferiscono attorno alla caustica.

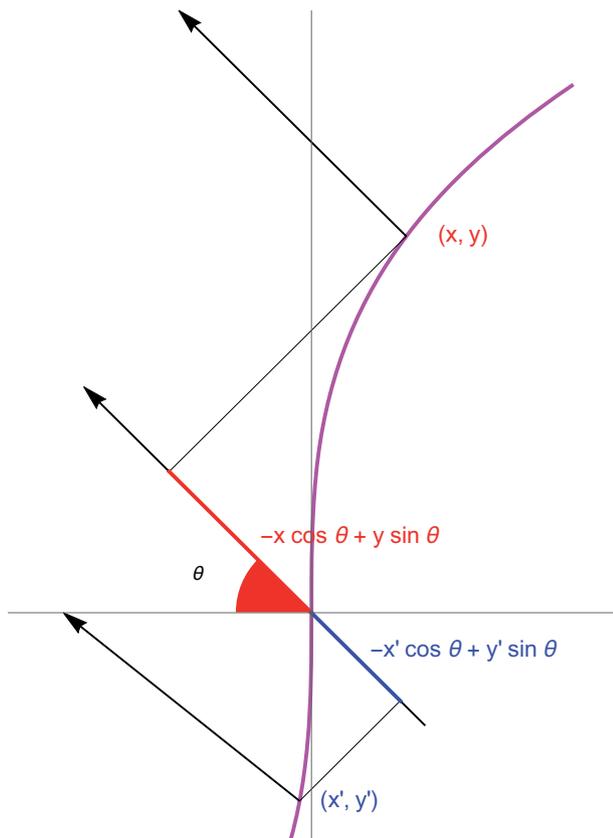


Figura 30: Calcolo degli sfasamenti lungo il fronte d'onda virtuale di Airy

assi coordinati in modo che la caustica sia l'asse \hat{x} negativo, il fronte può essere espresso in buona

approssimazione dal $x = \frac{\kappa^2}{3R^2} y^3$, dove κ è una costante adimensionale dipendente dall'indice di rifrazione e dall'ordine dell'arco. Supponendo che l'intensità della luce sia costante lungo il fronte, limitandosi al piano e concordemente ad Huygens, l'ampiezza cercata è data dall'integrale dei contributi provenienti da ogni punto (x, y) della cubica (vedi Figura 29). Se si pensa di determinarla in un punto lontano, nella direzione individuata dall'angolo θ , bisogna tener conto dei diversi ritardi di fase delle onde sferiche secondarie, per esempio, rispetto al punto di riferimento $(0, 0)$, che è pari a $\frac{-x \cos \theta + y \sin \theta}{\lambda}$. Attorno all'angolo di Descartes, cioè $\theta \approx 0$, lo sfasamento prende la forma limite $-\frac{\kappa^2}{3\lambda R^2} y^3 + \frac{\theta}{\lambda} y$. Questo porta, dopo vari calcoli ed approssimazioni, all'intensità della luce proporzionale al quadrato dell'espressione integrale

$$Ai(-m\theta) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \cos \left[\frac{s^3}{3} - m\theta s \right] ds,$$

con $m \propto \left(\frac{R}{\kappa\lambda}\right)^{2/3}$. Airy calcolò numericamente il suo integrale, oggi chiamato funzione di Airy, per diversi valori di $m\theta$: un lavoro pesante senza la moderna tecnologia dei computer. Ma poco dopo, su suggerimento di De Morgan, utilizzò lo sviluppo in serie attorno a $\theta = 0$. Come si vede dal grafico in Fig 30, la teoria di Airy prevede un picco principale, corrispondente all'arco primario, una serie di picchi secondari corrispondenti agli archi sovrannumerari per $\theta > 0$ di ampiezza decrescente, mentre per $\theta < 0$ l'arco si dissolve nell'ombra in maniera esponenzialmente rapida.

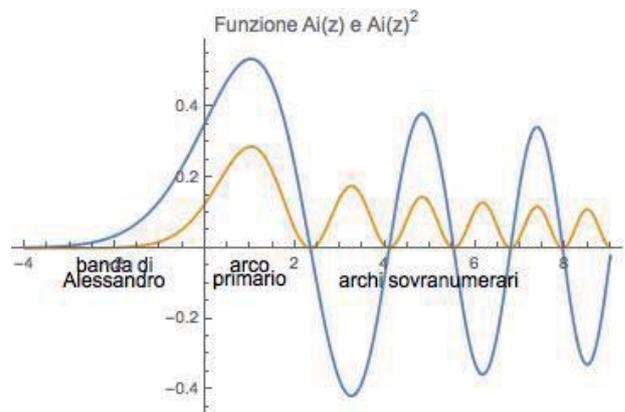


Figura 31:

La distribuzione dell'intensità luminosa pre-

detta da Airy è analoga alla figura di diffrazione che compare nell'ombra di un bordo rettilineo, ma la posizione e le altezze dei picchi differiscono alquanto rispetto a quelli previsti dalla teoria dell'interferenza di Young. Un'altra significativa distinzione della teoria di Airy è che il massimo di intensità dell'arco cade ad un angolo un po' più grande di quello calcolato da Descartes, infine l'intensità non diventa mai infinita in alcun punto.

I calcoli di Airy furono fatti per una luce monocromatica, quindi se si vuole applicare il suo metodo all'arco prodotto dalla luce solare, uno deve sovrapporre le varie componenti cromatiche. Procedendo in maniera analoga alla discussione fatta a suo tempo da Newton, per descrivere l'immagine effettivamente percepita bisogna ricorrere alla teoria della visione dei colori.

La purezza dei colori dell'arcobaleno è determinata dall'estensione con la quale i colori si sovrappongono, il che a sua volta dipende dalle dimensioni delle gocce. Grandi gocce uguali ($R \approx 1$ mm) generalmente danno archi brillanti con colori puri. Con gocce molto piccole ($R \approx 0.01$ mm) la sovrapposizione dei colori è così grande che la luce risultante è quasi bianca. Cosicché è opportuno introdurre il rapporto della circonferenza della goccia rispetto alla lunghezza d'onda della luce come il *parametro di taglia*

$$\beta = \frac{2\pi R}{\lambda} = k R.$$

Questo parametro varia da 100, per le gocce di nebbia, a diverse migliaia per quelle grandi. L'approssimazione di Airy è accettabile per gocce con $\beta \approx 5000$.

V. Khare, dell'Università di Rochester, fece un confronto dettagliato delle tre teorie dell'arcobaleno: l'approssimazione di Airy, l'approssimazione esatta, ottenuta sommando al computer le onde parziali e i termini di arcobaleno nel metodo del momento angolare complesso, associato con la collisione dei punti sella. Per polarizzazione perpendicolare la teoria di Airy richiede solo piccole correzioni entro l'arco primario: i suoi errori diventano apprezzabili solo nella regione dei sovrannumerari. Per la luce polarizzata parallelamente, la teoria di Airy fallisce pesantemente. Per gli archi sovrannumerari la soluzione esatta mostra minimi dove la teoria di Airy ha

dei massimi e viceversa. Questi seri fallimenti sono un risultato indiretto della quasi coincidenza tra l'angolo di riflessione interna e l'angolo di Brewster. All'angolo di Brewster l'ampiezza del raggio riflesso cambia segno, cosa della quale non si tiene conto nella teoria di Airy. Di conseguenza l'interferenza lungo le direzioni corrispondenti ai picchi nella soluzione di Airy è distruttiva invece che costruttiva.

In prospettiva Airy ha aperto nuove strade, ma non sembra che egli lo abbia colto completamente, tanto che non ne fa menzione nella sua autobiografia. In primo luogo, nell'ambito dell'ottica si è abbandonata definitivamente una descrizione puramente geometrica, per passare ad una teoria analitica di campo. Il problema matematico diventa quello delle *equazioni differenziali alle derivate parziali* con condizioni assegnate. Questa classe di problemi, già nota ai fisici-matematici dell'epoca, presentava metodi risolutivi lontani da essere consolidati in una teoria unitaria. Tuttavia, proprio in quegli anni Poisson aveva sviluppato formule integrali per le equazioni di d'Alembert, che erano interpretabili in termini di principio di Huygens-Fresnel. Altro campo della Matematica collegato alla funzione di Airy fu quello delle equazioni differenziali ordinarie del II ordine lineari a coefficienti non costanti, la loro classificazione e soluzione in termini di *funzioni speciali classiche*. Si ricordi a tal proposito il legame con le *funzioni di Bessel* e relative equazioni. Inoltre molti ricorderanno il ruolo centrale svolto dalla funzione di Airy nel *metodo di approssimazione semiclassica*, noto anche come *Wentzel-Kramers-Brillouin* e sue varianti. In parallelo il suo studio sviluppò ulteriormente la teoria delle funzioni analitiche e lo studio delle singolarità delle funzioni nel piano complesso. In particolare Stokes ricavò formule di approssimazioni per $|m| \gg 1$ in termini di *serie asintotiche*, avviandone lo studio, e la scoperta delle *linee di Stokes*, che separano le regioni di validità dei suddetti sviluppi [38]. Approfondimenti su questi temi si possono trovare in innumerevoli manuali, tra i quali, più per motivi storici, qui menzioniamo il celebre trattato di Jeffreys [39].

Il Ponte delle Risonanze

Una importante proprietà della luce che abbiamo fin qui ignorato è la trasversalità delle sue oscillazioni: esse sono perpendicolari alla direzione di propagazione (differentemente, il suono è un'onda longitudinale). L'orientazione delle oscillazioni trasversali può essere decomposta lungo due assi perpendicolari tra loro, che definiscono stati di polarizzazione indipendenti. La luce solare non è polarizzata, perché è una miscela incoerente dei due stati in eguali proporzioni, ma le riflessioni in una goccia di pioggia possono modificare questo rapporto e, quindi anche quello della luce in un arcobaleno. Il legame quantitativo di quale frazione di luce incidente all'interfaccia tra due mezzi sia riflessa (oppure trasmessa) è fornito dalle ben note relazioni di Fresnel [36], la forma delle quali è sensibile allo stato di polarizzazione. Intanto, se fissiamo il

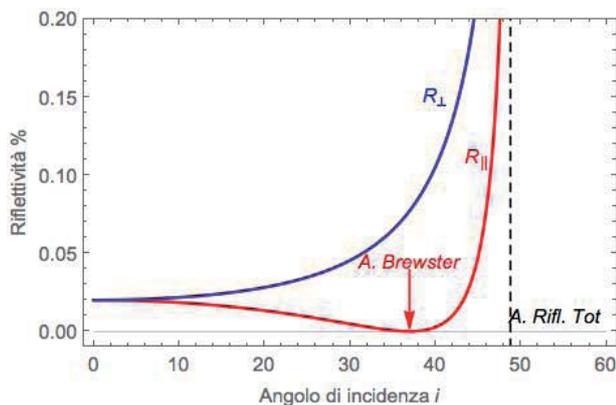


Figura 32: La riflettività della luce che si propaga dall'acqua verso l'aria. R_{\parallel} e R_{\perp} si riferiscono al caso di polarizzazione parallela o perpendicolare, rispettivamente, rispetto al piano di riflessione.

piano contenete sia un raggio luminoso che il suo riflesso sulla superficie interna di una goccia d'acqua, gli stati di polarizzazione da considerare sono quello parallelo e quello perpendicolare a tale piano. Le formule di Fresnel forniscono la riflettività R_{\parallel} e R_{\perp} per entrambe le polarizzazioni, come indicato nel grafico di Figura 31. La riflettività della superficie è piccola ad angoli di incidenza prossimi alla perpendicolare ($i \approx 0$), ma vicino all'angolo di riflessione totale, il cui valore è determinato dall'indice di rifrazione, essi divergono, fino a descrivere la riflessione totale del raggio per angoli maggiori. Ad an-

goli intermedi, tuttavia, la riflettività dipende sensibilmente dalla polarizzazione: la componente perpendicolare è maggiormente riflessa, mentre per la perpendicolare, addirittura viene completamente trasmessa al cosiddetto l'angolo di Brewster, da D Brewster, che ne discusse il significato nel 1815. Poiché l'angolo interno di incidenza ad una goccia di pioggia è molto vicino all'angolo di Brewster, la maggior parte delle componenti parallele sfuggono come raggi trasmessi, lasciando una preponderanza della componente perpendicolare della luce. Di conseguenza la luce dell'arcobaleno è quasi completamente polarizzata, come si può osservare ruotando delle lenti Polaroid attorno alla linea di vista.

Per riassumere, sebbene di molto anteriori alla teoria elettromagnetica di Maxwell, l'ipotesi ondulatoria della luce e la scoperta della polarizzazione riescono a dare una serie di risposte, approssimate ancora, al problema posto dall'arcobaleno. Tuttavia solo nell'ambito teorico molto più ricco e complesso dell'Elettromagnetismo (si veda, nello stesso numero di Ithaca, M. Mazzeo: *Il lato oscuro della luce*) si riesce a trovare una soluzione esatta, anche se ben presto ci accorgeremo di dover pagare un prezzo.

Una precisa formulazione matematica del problema dell'arcobaleno ottico consiste nel calcolare l'ampiezza dell'onda diffusa da una sfera di indice di rifrazione costante n (anche complesso), sulla quale incide un'onda piana monocromatica. Nel caso delle onde sonore la soluzione a tale problema fu discusso da Rayleigh [40], mentre per l'analogo problema elettromagnetico fu trovata nel 1908 da G Mie [41], anche se precedenti passi [42][43] verso la sua formulazione e furono apportati miglioramenti successivi [44] [45] [46]. La soluzione viene espressa in una serie infinita di termini, chiamati *onde parziali*. E' chiaro che il prezzo al quale si alludeva consiste nel controllo della convergenza e nella somma della serie. Questa forma proviene dalla necessità di trovare una soluzione stazionaria delle equazioni di Maxwell, cioè dalle quali è possibile fattorizzare la dipendenza temporale della forma $e^{-i\omega t}$, in regioni sferiche concentriche, distinte dal diverso indice di rifrazione. Equazioni di questo tipo

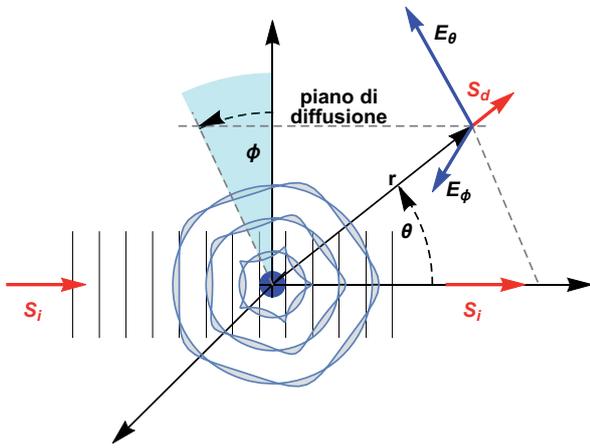


Figura 33: L'onda piana incidente viene diffusa in tutte le direzioni dal centro (goccia di pioggia). L'onda così prodotta è espressa dalla somma di infinite onde sferiche concentriche, con ampiezza modulata sia da oscillazioni radiali, che da oscillazioni angolari. L'interferenza di tutti questi possibili stati di oscillazione determina l'ampiezza complessiva dell'onda, e quindi di ogni altra quantità fisica. Per le onde e.m. il vettore di Poynting \vec{S} esprime il flusso di potenza in ogni punto dello spazio. Le componenti del c. elettrico con polarizzazione indipendente sono \vec{E}_ϕ (perpendicolare) e \vec{E}_θ (parallela), con riferimento al piano di diffusione.

si riducono alla ben nota equazione di Helmholtz

$$\nabla^2 U + n^2 k^2 U = 0,$$

dove k assume valori diversi nella sfera e al suo esterno e U è, in generale, una funzione vettoriale a valori complessi. Questa è una sorta di madre delle equazioni della fisica-matematica del XIX e XX secolo [39]. La difficoltà nel trovare una sua soluzione nel caso considerato consiste nel raccordare con continuità, sino alle derivate prime, le soluzioni alla superficie di separazione tra i due mezzi. Supponendo di far incidere sulla sfera un fascio di luce monocromatica polarizzata linearmente, a grande distanza dalla goccia ($r \rightarrow \infty$) le componenti del c. elettrico della radiazione diffusa assumono la forma

$$E_\phi = \frac{e^{i k r}}{i k r} \sin \phi S_\perp(\theta; \beta), \quad E_\theta = \frac{e^{i k r}}{i k r} \cos \phi S_\parallel(\theta; \beta).$$

Il primo fattore di queste espressioni esprime l'andamento di un'onda sferica con numero d'onda $k = \frac{\omega}{c} = |\vec{k}_{inc}|$, il secondo dipende solo dall'angolo azimutale che definisce il piano di dif-

fusione, il terzo (S_\perp o S_\parallel) contiene l'angolo di diffusione e il parametro di taglia, quindi codifica tutta l'informazione relativa al tipo di goccia e di luce che vi facciamo incidere. Perciò essi vengono chiamati *ampiezze di diffusione* e determinano completamente l'intensità luminosa diffusa (relativamente all'intensità della luce incidente)

$$I_{tot}(\theta, \phi) = |S_\perp|^2 \sin^2 \phi + |S_\parallel|^2 \cos^2 \phi,$$

che i fisici abitualmente chiamano *sezione d'urto differenziale*.

Ora l'espressione di $S_{\perp, \parallel}$ è data, come annunciato, sommando la serie di infiniti termini

$$S_j = \sum_{\ell=1}^{\infty} \left\{ a_\ell^{(j)}(\beta) \pi_\ell(\cos \theta) + b_\ell^{(j)}(\beta) \tau_\ell(\cos \theta) \right\}, \quad (1)$$

dove $j = \perp, \parallel$, π_ℓ e τ_ℓ determinano la distribuzione angolare dell'intensità luminosa (Figura 35) e sono dati in termini di funzioni di Legendre (box), mentre i coefficienti di Mie $a_\ell^{(j)}$ e $b_\ell^{(j)}$ sono espressioni razionali di funzioni di Bessel sferiche (box), dipendenti solo dal parametro di taglia β e dall'indice di rifrazione. Su di esse esiste una letteratura matematica vastissima [39], ma vale ancora una volta la pena di sottolineare il profondo legame tra i problemi di diffusione della luce, la teoria delle equazioni differenziali lineari del secondo ordine e la teoria delle funzioni di variabile complessa. Campo di ricerca centrale per l'Analisi Matematica, a partire da Cauchy, Riemann e Weierstrass. Non senza trascurare quel senso estetico insito nella Matematica ispirata dall'arcobaleno.

A questo punto potrebbe sembrare facile determinare il preciso carattere dell'arcobaleno, se si riuscisse a sommare la serie di onde parziali, il che non è possibile analiticamente. D'altra parte una soluzione approssimata stabile rispetto all'aggiunta di termini ulteriori si ottiene quando si somma almeno fino $\ell_{max} = \beta + 2\beta^{1/3} + 2$, in quanto essa converge lentamente.

Perciò la serie in onde parziali si può ridurre ad un solo termine quando $\beta \ll 1$, cioè quando le dimensioni del corpo diffusore sono molto piccole rispetto alle lunghezze d'onda coinvolte (Figura 38). Questo regime di diffusione, detto di Rayleigh, si verifica quando le molecole di aria diffondono la luce visibile. In tal caso il fenome-

Le funzioni di Legendre

L'equazione di Legendre per la funzione $y = y(x)$ è

$$(1 - x^2) \frac{d^2 y}{dx^2} - 2x \frac{dy}{dx} + \ell(\ell + 1)y = 0.$$

Come tutte le equazioni lineari del 2° ordine, essa ha due soluzioni linearmente indipendenti: $P_\ell(x)$ regolare nei punti singolari ± 1 e le Q_ℓ singolare in essi. Se il parametro ℓ è un intero, come nel caso della serie di $S_{\perp, \parallel}$, le P_ℓ sono polinomi. Essi hanno notevolissime proprietà analitiche e algebriche e costituiscono il prototipo dei cosiddetti *polinomi ortogonali*. Una loro generalizzazione è costituita dalle *funzioni associate* di Legendre $P_\ell^m = (1 - x^2)^{\frac{m}{2}} \frac{d^m P_\ell}{dx^m}$, che sono pertinenti al nostro caso essendo

$$\pi_\ell = \frac{P_\ell^1(\cos \theta)}{\sin \theta}, \quad \tau_\ell = \frac{d P_\ell^1(\cos \theta)}{d \theta}.$$

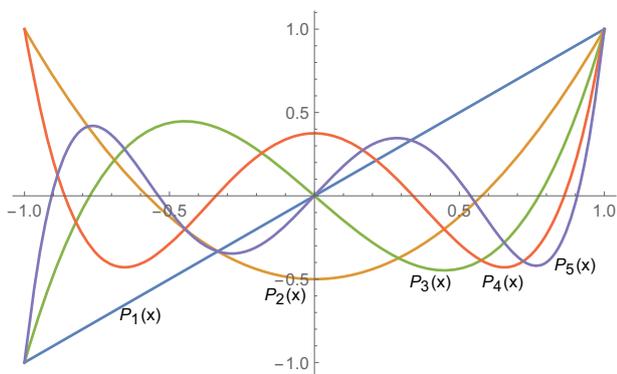


Figura 34: I primi cinque polinomi di Legendre.

no più eclatante è l'azzurro del cielo, poiché la sezione d'urto differenziale di Rayleigh è circa 16 volte più grande per il blu rispetto al rosso, dipendendo da $\approx \lambda^{-4}$.

Ma per gocce di pioggia illuminate dal Sole si ha $\beta \approx 10^3$, quindi una buona approssimazione della soluzione di diffusione comporta il calcolo di migliaia di termini complicati e rapidamente variabili in β e θ (si vedano le figure 39 e 40). Anche se al giorno d'oggi esistono potenti strumenti e metodi di calcolo numerico [6][47][48], anche on-line [49], il costo computazionale può

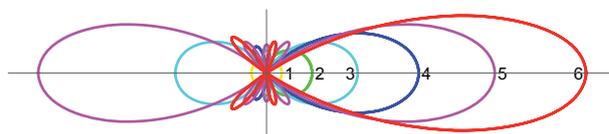


Figura 35: Rappresentazione polare delle funzioni π_ℓ , per $1 \leq \ell \leq 6$, nelle ampiezze S_i della radiazione diffusa.

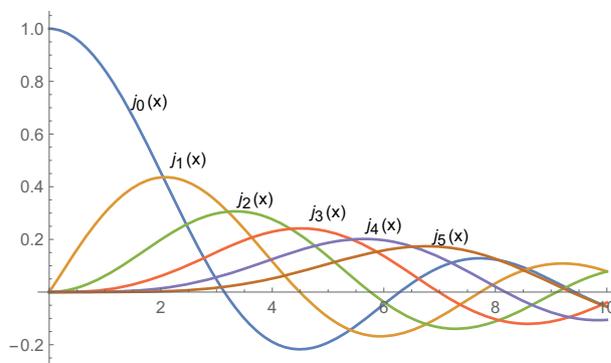


Figura 36: Le prime sei funzioni di Bessel sferiche di prima specie.

diventare ben presto proibitivo. D'altra parte il calcolo numerico offre una certa intuizione sulla fisica coinvolta solo estraendola dal risultato finale, mentre è meno chiaro come un certo risultato emerga. Ci troviamo allora nella imbarazzante situazione di conoscere la forma della soluzione esatta, ma di essere ancora incapaci di comprendere completamente i meccanismi che la determinano. E' come se la speciale miscela di effetti diffrattivi che determinano l'arcobaleno fossero sepolti nelle formule della soluzione esatta di Mie.

Verso un chiarimento di questo paradosso H Poincaré e G Watson, all'inizio del XX secolo, trovarono un metodo per trasformare la serie lentamente convergente delle onde parziali in una espressione integrale molto più maneggevole, nota come trasformazione di Watson o del momento angolare complesso.

Che il momento angolare intervenga nel problema dell'arcobaleno deriva dall'interpretazione corpuscolare, nella quale un fascio di luce è visto come un flusso di fotoni. Anche se i fotoni non hanno massa, essi trasportano energia e quantità di moto, quindi anche momento angolare, che è il prodotto tra il momento lineare ed il parametro d'urto. Quando la luce subisce un numero di riflessioni interne ad una goccia di acqua, equivale a dire che un fascio di fotoni in

Le funzioni di Bessel Sferiche

L'equazione di Bessel sferica per la funzione $z = z(r)$ è

$$r^2 \frac{d^2 z}{dr^2} + r \frac{dz}{dr} + \left[r^2 - \left(\ell + \frac{1}{2} \right)^2 \right] z = 0, \ell \in \mathbb{N}.$$

Anche in questo caso ogni soluzione è la combinazione di due soluzioni linearmente indipendenti. In particolare quelle di prima specie j_ℓ sono regolari come x^ℓ nel punto singolare $r = 0$, mentre quelle di seconda specie n_ℓ sono ivi singolari, come $r^{-\ell-1}$. Per $r \rightarrow \infty$ il comportamento di queste funzioni è decrescente come $\frac{1}{r}$ modulato da $\sin r$ o $\cos r$ alternativamente per ℓ pari o dispari. Le speciali combinazioni lineari $h_\ell^{(\pm)} = j_\ell \pm i n_\ell$ sono chiamate funzioni di Hankel. Il loro comportamento asintotico a $r \rightarrow \infty$ è quello $\frac{\exp[\pm i r]}{r}$ delle onde sferiche uscenti/entranti.

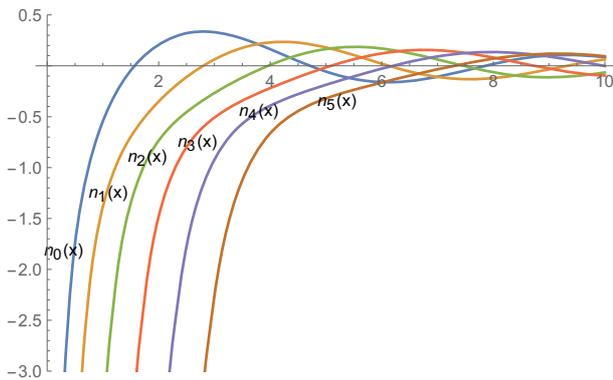


Figura 37: Le prime sei funzioni di Bessel sferiche di seconda specie.

effetti sta orbitando attorno ad centro di attrazione specifico. Ma la meccanica (quantistica) pone dei vincoli a questi processi elementari, richiedendo che il momento angolare assuma soltanto valori multipli interi di \hbar , che è il momento angolare fondamentale. Pertanto ogni termine della somma in ℓ dello sviluppo in onde parziali per (1) corrisponde all'ampiezza di uno stato di onda diffusa avente uno specifico valore del momento angolare. Al contempo il parametro d'urto non svolge alcun ruolo, manifestando così il suo carattere di osservabile incompatibile con il mo-

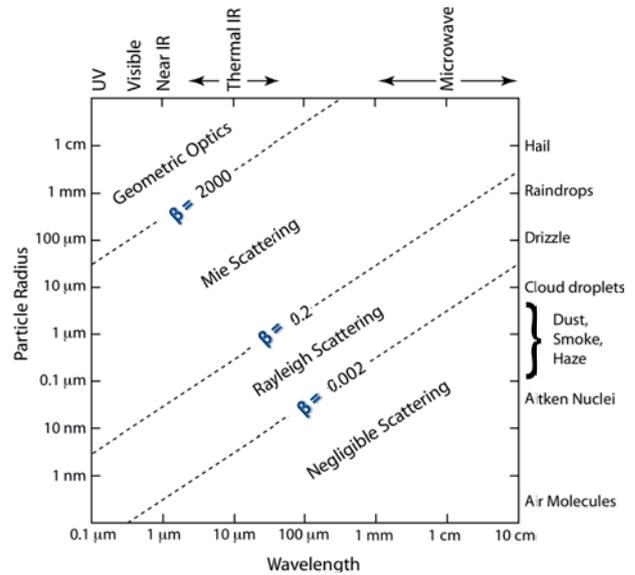


Figura 38: Particelle sferiche isolate puramente dielettriche diffondono la luce secondo vari regimi, definiti dal parametro di taglia β , e che corrispondono ad approssimazioni differenti della soluzione esatta.

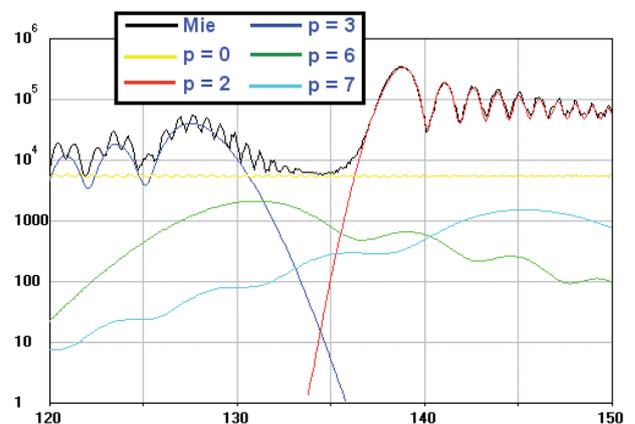


Figura 39: Diffusione di luce (rossa) non polarizzata di lunghezza d'onda $\lambda = 0.65 \mu\text{m}$ da parte di una goccia d'acqua sferica di raggio $r = 100 \mu\text{m}$. Sono evidenziati i contributi che provengono dai raggi che subiscono $p - 1$ riflessioni interne. Pertanto $p = 2$ corrisponde alla diffusione che contribuisce principalmente all'arco primario, mentre $p = 3$ produce l'arco secondario. [48]

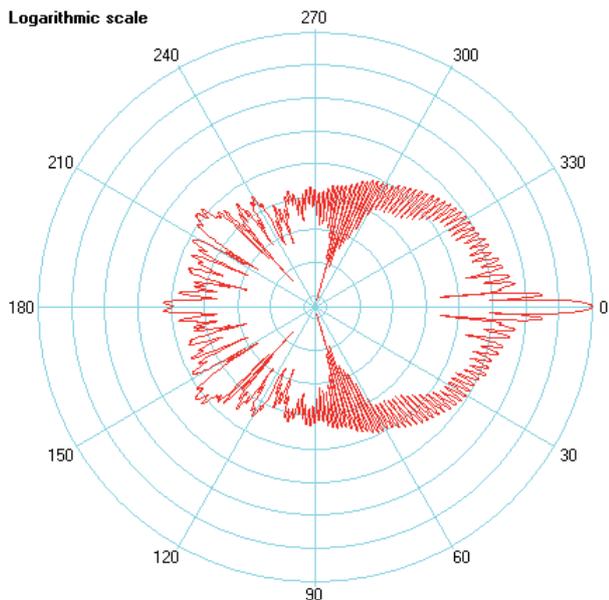


Figura 40: Diagramma polare della diffusione di luce polarizzata perpendicolarmente con $\lambda = 0.65 \mu\text{m}$, da parte di una goccia d'acqua di raggio $r = 10 \mu\text{m}$. Si noti che la scala delle ampiezze è logaritmica, per cui il picco in avanti è circa 10^4 volte più intenso dei picchi dell'arco primario e 10^5 volte rispetto al secondario. Si noti anche il pronunciato picco all'indietro a 180° . [48]

mento angolare (un esempio di applicazione del principio di indeterminazione della meccanica quantistica).

A questo punto l'astrazione matematica ci permette di pensare alla variabile momento angolare come ad una quantità continua $\ell \rightarrow z$, addirittura che vari nel campo complessi. Corrispondentemente i coefficienti di Mie e le altre espressioni presenti nella formula (1) si interpretano come i valori di certe funzioni complesse in z , che interpolano i valori ottenuti in corrispondenza dei valori discreti $z = \ell$. Queste funzioni si assumono, ad eccezione di alcune singolarità, analitiche in tutto il piano complesso, in particolare lo sono in prossimità del semiasse reale positivo di z . Questa proprietà è senz'altro verificata nei casi in cui il *potenziale di diffusione* ha un raggio di azione limitato, come appunto gocce o bolle di liquido, all'esterno delle quali tutto si svolge come nello spazio libero. Per una funzione interpolante ϕ del tipo descritto e usando il *teorema dei residui*, Watson [50] ottenne la formula

integrale

$$\frac{1}{2} \int_C \phi(z, x) \frac{\exp(i\pi z)}{\cos(\pi z)} dz = \sum_{\ell=0}^{\infty} \phi(\ell + \frac{1}{2}, x),$$

dove il cammino di integrazione C è riportato in Figura 41, che si può dimostrare equivalente ad una analoga formula ottenuta in precedenza da Poisson. In virtù dell'analiticità della funzione

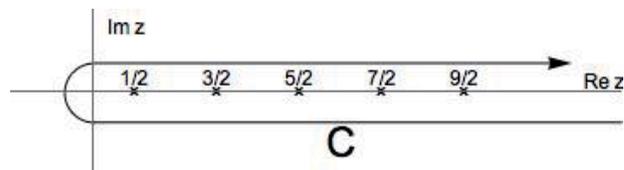


Figura 41: Nel piano del momento angolare complesso la funzione integranda nella trasformata di Watson possiede poli semplici solo nei punti della forma $\ell + \frac{1}{2}$.

integranda nel piano complesso, il cammino C può essere deformato con continuità, facendolo passare in particolare per punti di fase stazionaria (punti sella) [39]. Infatti i contributi all'integrale negli intorno di questi punti sono quelli che maggiormente contribuiscono all'integrale in questione. Pertanto si possono agevolmente ottenere espansioni asintotiche, che risommano moltissimi termini dello sviluppo in onde parziali. In particolare i contributi provenienti da punti sella reali sono associati con i raggi luminosi diffratti/deflessi fin qui considerati. I punti sella in corrispondenza di valori complessi del momento devono descrivere un qualche smorzamento dell'onda. Un simile comportamento si ha sicuramente nel caso di riflessione interna all'interfaccia acqua-aria, nel qual caso l'intensità diventa trascurabile in uno strato di profondità dell'ordine una lunghezza d'onda in aria. Un analogo fenomeno si verifica nella descrizione del lato buio di una caustica. La loro matematica è simile a quella che descrive l'effetto tunnel in Meccanica Quantistica.

In questo linguaggio l'origine dell'arcobaleno, vista nella semplice analisi di Descartes come la confluenza di più raggi nell'unico di deflessione minima, viene sostituita dalla confluenza di due punti sella reali nel piano complesso del momento angolare al variare di β . Superato tale angolo di arcobaleno, i punti non spariscono, ma diventano complessi, assumendo parti im-

maginarie, corrispondenti alla diffrazione nella banda oscura di Alessandro. Essa è un esempio di raggio complesso nel lato oscuro della curva di una caustica.

D'altra parte le ampiezze di diffusione possono presentare anche delle singolarità polari nel piano complesso del momento angolare, note come *poli di Regge*, dal nome del celebre fisico italiano che contribuì fortemente al loro studio nell'ambito della Fisica delle Alte Energie [51].

I poli di Regge descrivono onde luminose associate a raggi incidenti tangenzialmente la superficie acqua-aria. Esse possono eccitare un'onda superficiale, che si propaga lungo una geodetica. Ad un certo punto può sfuggire alla superficie, ma essa ha una certa probabilità di penetrare nella sfera all'angolo critico, compiendo una o più riflessioni interne, prima di riemergere ancora in superficie (vedi Figura 42). Qualcosa di molto

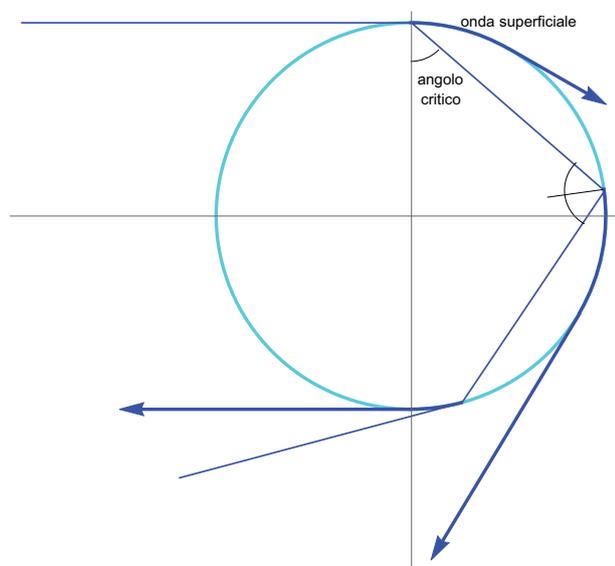


Figura 42: I raggi incidenti tangenzialmente alla goccia possono propagarsi come onde superficiali e poi essere riemesse tangenzialmente. Tuttavia possono anche essere rifratte all'interno all'angolo critico, subendo un numero finito di riflessioni interne, prima di ritornare in superficie.

simile alla congettura di Keplero, ma abbiamo visto che esso non contribuisce all'arcobaleno. Piuttosto l'effetto diventa dominante a grandi angoli fino alla direzione anti-solare, dove si è già osservato un importante picco relativo di intensità luminosa. Questi raggi sono quindi responsabili dell'affascinante fenomeno meteorologico del *gloria* (Figura 43). Esso appare come un alone di

colori spettrali, che circonda l'ombra di un osservatore messo sulle nubi o nella nebbia. Esso è più facilmente osservabile da un aeroplano che vola sulle nubi.



Figura 43: Un gloria solare che circonda il cosiddetto Spettro di Brocken, *TheBrockenInaGlory*, <https://sites.google.com/site/thebrockeninglory/>

Si dovrebbe notare che l'adozione della teoria del momento angolare complesso non implica che le precedenti soluzioni siano errate. La spiegazione di Descartes dell'arco primario come raggio di minima deflessione non è sbagliata e, ancora, gli archi sovranumerari possono essere visti come effetto dell'interferenza di Young. Tuttavia il metodo del momento angolare complesso fornisce una catalogazione più completa dei cammini che la luce / i fotoni possono seguire e una più accurata valutazione delle loro ampiezze.

Conclusioni

Abbiamo iniziato dall'osservazione infantile di un fenomeno naturale e dopo un lungo percorso, culturale e scientifico, siamo approdati ad una teoria matematica astratta e, in un certo senso, piuttosto lontana dall'intuizione. Tuttavia non è stato un esercizio privo di frutti. In primo luogo abbiamo adottato l'arcobaleno a mo' di banco da prova dei nostri modelli scientifico/culturali. Siamo inoltre riusciti a dedurre dalle nostre teorie sulla luce tanti dettagli sull'arcobaleno, che non li avremmo mai notati senza uno sguardo, altrimenti motivato, indirizzato a cercarli. Infine le idee e i metodi che abbiamo sviluppato su questo particolare effetto sono diventati paradigmatici nell'ambito della fisica atomica e nucleare,

dove arcobaleni atomici e nucleari si possono studiare con i metodi discussi sopra. Per tutti basti citare il lavoro sviluppato in Fisica Nucleare dal compianto Raimondo Anni e dalla sua *Rainbow Company* <http://www.dmf.unisalento.it/anni/Rainbow/Rainbow.htm>, al quale moltissimo devo per la mia formazione di fisico.

Queste idee e questi metodi saranno sicuramente fecondi in molti altri ambiti della scienza pura ed applicata e colgo l'occasione per scusarmi con i lettori per essermi troppo dilungato su un tema forse marginale ma dal caleidoscopico fascino, il quale ancora una volta mi ha trascinato lontano, quasi annegando in una vera pioggia (per dirlo con Hazel Grace nell'Introduzione) di ricordi, esperienze, riflessioni e idee.



[1] R. L. LEE, A. B. FRASER: *The Rainbow Bridge: Rainbows in Art, Myth, and Science*. Penn State Press, University Park (2001).

[2] C. B. BOYER: *The Rainbow: from Myth to Mathematics*. Princeton Univ. Press, Princeton (1987).

[3] J. A. ADAM: "Like a bridge over colored water", *AMS Notices* **49** (2002) 1360–1371.

[4] V. KHARE, H. M. NUSSENZVEIG: "Theory of the rainbow", *Physical Review Letters* **33** (1974) 976–980.

[5] D. K. LINCHE, W. LIVINGSTON: *Color and Light in Nature*. Cambridge University Press, Cambridge (2001).

[6] Si suggerisce di visitare <http://www.atoptics.co.uk> e <http://www.philiplaven.com>

[7] J. A. ADAM: "The mathematical physics of rainbows and glories", *Phys. Rep.* **356** (2002) 229–365.

[8] H. M. NUSSENZVEIG: *Diffraction Effects in Semiclassical Scattering*. Cambridge University Press, Cambridge (2006).

[9] "Bibbia", Conferenza Episcopale Italiana (2008).

[10] J. W. VON GOETHE: *Zur Farbenlehre*. J.G. Cotta'schen Buchhandlung, Tubinga (1810). J. W. VON GOETHE: *La teoria del colore*. Il Saggiatore, Milano (2008).

[11] L. RUSSO: *La rivoluzione dimenticata*. Universale Economica Feltrinelli, Milano (2003).

[12] E. GRANT: *A Source Book in Medieval Science*. Harvard University Press, Harvard (1974).

[13] F. RISNER: *Opticae thesaurus, Alhazeni Arabis libri septem. Vitelloni Thuringopoloni opticae libri X*. Episcopios, Basilea (1572).

[14] L. RUSSO: *Flussi e riflussi*. Feltrinelli, Milano (2003).

[15] R. BAGNOLESI: "Un modello per capire meglio la formazione dell'arcobaleno", *La Fisica nella Scuola* **XIX**, **1** (1986) 38.

[16] V. RONCHI: *Scritti di Ottica*. Il Polifilo, Milano (1968).

[17] LEONARDO DA VINCI: *Trattato della pittura*. Giunti-Demetra, Firenze (2006).

[18] F. MAUROLICO: *Photismi de Lumine*. Tarquinio Longo, Napoli (1611).

[19] M. A. DE DOMINIS: *De radiis visus et lucis in vitris perspectivis et iride*. T Baglioni, Venezia (1611).

[20] J. KEPLER: *Ad Vitellionem Paralipomena, quibus Astronomiae Pars Optica Traditur*. C. Marnium & I. Aubrii, Francoforte (1604). J KEPLER: *Dioptrice*. David Francius, Augusta (1611).

[21] G. GALILEI: *Sidereus Nuncius*. Tommaso Baglioni, Venezia (1610).

[22] R. DESCARTES: *Discours de la Méthode pour bien conduire sa raison et chercher la vérité dans les Sciences*. Ian Maire, Leyden (1637).

[23] I. NEWTON: "Letter containing his New Theory about Light and Colors", *Philosophical Transactions* **6** (1671) 3075-3087.

[24] I. NEWTON: *Opticks, or a Treatise of the Reflections, Refractions, Inflections and Colours of Light*. W. Innys, London (1730).

[25] R. P. FEYNMAN, R. B. LEIGHTON, M. SANDS: *The Feynman Lectures on Physics*. Addison-Wesley Pub. Co., London (1969).

[26] M. KEMP: *The Science of Art*. Yale University Press, New Haven (1990).

[27] M. LIVINGSTONE: *Vision and Art*. Harry & Abrams inc., New York (2008).

[28] International Commission on Illumination <http://www.cie.co.at>

[29] C. HUYGENS: *Traité de la lumière*. P. Vander, Leida (1690).

[30] T. YOUNG: "On the Theory of Light and Colours", *Phil. Trans. R. Soc. Lond.* **92** (1802) 12-48.

[31] T. YOUNG: "Experiments and Calculations Relative to Physical Optics", *Phil. Trans. R. Soc. Lond.* **94** (1804) 1-16.

[32] J. FRAUNHOFER: "Bestimmung des Brechungs- und des Farben-Zerstreungs - Vermögens verschiedener Glasarten, in Bezug auf die Vervollkommnung achromatischer Fernröhre", *Denksch. K. Ak. der Wiss. München* **5** (1814) 193-226.

[33] F. M. GRIMALDI: *Physico-mathesis de lunime, coloribus et iride*. V. Benati, Bologna (1665).

[34] A.-J. FRESNEL: "Memoire sur la diffraction de la lumière", *Mém. Académie des Sciences* **5** (1826) 33-475.

[35] G. KIRCHHOFF: "Zur Theorie der Lichtstrahlen", *Ann. d. Physick* **254** (1883) 663-695.

[36] M. BORN, E. WOLF: *Principles of Optics*. Pergamon, New York (1970).

[37] G. B. AIRY: "On the intensity of light in the neighbourhood of a caustic", *Trans. Cambridge Philos. Soc.* **6** (1838) 379-403.

- [38] G. G. STOKES: "On the numerical calculation of a class of definite integrals and infinite series", *Trans. Cambridge Phil. Soc.* **9** (1856) 166-188. G. G. STOKES: "On the discontinuity of arbitrary constants which appear in divergent developments", *Trans. Cambridge Phil. Soc.* **10** (1864) 105-124.
- [39] H. JEFFREYS, B. JEFFREYS: *Methods of Mathematical Physics*. Cambridge Univ. Press, Cambridge (1956).
- [40] J. W. STRUTT B. RAYLEIGH: *The Theory of sound*. Macmillan, London (1896).
- [41] G. MIE: "Beiträge zur Optik trüber Medien, speziell kolloidaler Metallosungen", *Ann. Phys.* **25** (1908) 377,445. Royal Aircraft Establishment library translation, 1976
- [42] A. CLEBSCH: "Ueber die Reflexion an einer Kugelfläche", *Z. fur Math.* **61** (1863) 195.
- [43] L. V. LORENZ: "Sur la lumiere reflechie et refractee par une sphere transparente", *Vidensk. selsk. Skrifter* **6** (1890) 1..
- [44] P. DEBYE: "Der Lichtdruck auf Kugeln von beliebigem Material", *Ann. d. Phys.* **30** (1909) 57.
- [45] M. KERKER: *The Scattering of Light*. Academic Press, New York (1969).
- [46] W. T. GRANDY: *Scattering of Waves from Large Spheres*. Cambridge Univ. Press, Cambridge (2000).
- [47] W. WISCOMBE: "Improved Mie scattering algorithms", *Appl. Opt.* **19** (1980) 1505.
- [48] P. LAVEN: "Simulation of rainbows, coronas, and glories by use of Mie theory", *Applied Optics* **42** (3) 2003.436
- [49] <http://omlc.org/software/mie/>
- [50] G. N. WATSON: "The diffraction of electric waves by the earth", *Proc. Roy. Soc. (London)* **A95** (1918) 83.
- [51] V. DE ALFARO, T. REGGE: *Potential Theory*. North Holland, Amsterdam (1965).



Luigi Martina: Luigi Martina è professore associato di Fisica Teorica presso l'Università del Salento. La sua attività di ricerca è incentrata nello studio di sistemi fisici nonlineari con metodi analitici e algebrici.

Luce e Quantum Dot

Vittorianna Tasco

CNR Nanotec, Istituto di Nanotecnologia - Lecce

Luigi Carbone

CNR Nanotec, Istituto di Nanotecnologia - Lecce

La definizione di materiale nanostrutturato prevede che questo cambi in modo consistente le sue proprietà ottiche o elettroniche quando le sue dimensioni caratteristiche sono dell'ordine di decine o centinaia di nanometri. L'importanza di questi materiali sta nel fatto che permettono di passare da sistemi di volume, o *bulk*, a sistemi fondamentali, come quelli atomici o molecolari, aprendo strade completamente nuove per applicazioni in elettronica, optoelettronica e biologia. Negli ultimi vent'anni, l'attenzione della ricerca nell'ambito della fisica della materia è stata notevolmente focalizzata sulle proprietà optoelettroniche di semiconduttori nanostrutturati, detti *quantum dot*, dove svariate proprietà fondamentali diventano *size-dependent* e possono essere ingegnerizzate tramite processi di sintesi e di deposizione con livello di controllo su scala atomica.

1 Introduzione e Origini

Un punto quantico (in inglese quantum dot - Qdot) è una piccola struttura di un materiale posta all'interno di una matrice con diversa composizione, e con dimensioni caratteristiche dell'ordine del nanometro (fino a poche centinaia di nm). Cambiando le dimensioni della nanostrut-

tura, proprietà fondamentali come luminescenza, assorbimento ottico e conduzione elettrica, possono essere ingegnerizzate a un livello non raggiungibile con i materiali tradizionali. La prima realizzazione di Qdot nella storia risale a più di 2000 anni fa, quando fu messo a punto un processo per sintetizzare particelle di solfuro di piombo usando ossido di piombo, idrossido di calcio (comunemente detto calce spenta) ed acqua; questo materiale veniva usato da Greci e Romani per colorare i capelli. Più recentemente, all'inizio del '900, si riportava il controllo di precipitati di CdS o CdSe in vetri silicati per modificarne il colore, e ancor prima particelle di semiconduttore venivano usate per drogare il vetro e realizzare così dei filtri ottici. All'inizio degli anni '80, la scienza e tecnologia dei semiconduttori era focalizzata sulle proprietà di multistrati di film sottili (eterostrutture a semiconduttore) e sullo studio degli effetti quantistici in tali sistemi per l'elettronica e l'optoelettronica. In questo contesto, due scienziati giapponesi, Arakawa e Sakaki [1], proposero l'utilizzo del confinamento quantistico dei portatori di carica nella zona attiva di un dispositivo laser a semiconduttore per migliorarne le prestazioni, in termini di correnti di soglia e temperatura caratteristica. In questo lavoro, il concetto di Qdot è ancora primordiale, ma la condizione essenziale era che questo sistema dovesse avere dimensioni caratteristiche comparabili con la lunghezza d'onda di De Broglie degli elettroni del materiale, in modo che i portatori di carica (elettroni o lacune) al suo interno potessero trovarsi in regime di confinamento quantistico. Sempre nel 1982, A. L.

Efros ed A. L. Efros [2] postularono che la variazione delle proprietà opto-elettroniche con le dimensioni (cosiddetto effetto di *quantum size*) potevano essere usati insieme alla sua composizione per controllare il colore dei vetri. Da lì in poi si aprì la strada verso lo sviluppo di metodi di sintesi o di fabbricazione in grado di realizzare e manipolare queste nanostrutture così promettenti. Negli ultimi vent'anni, la ricerca teorica e sperimentale su questi sistemi ha visto una crescita esponenziale, permettendo sia di esplorare proprietà fondamentali della materia precedentemente inaccessibili, sia di trovare applicazioni che sono diventate realtà commerciali, come il caso dei display televisivi a Qdot, o il laser a Qdot come sorgente per le telecomunicazioni.

2 Struttura dei Qdots

I Qdots sono nanostrutture di materiale semiconduttore con una banda di energia minore di quella della matrice in cui si trovano. I portatori di carica elettrica (elettroni e lacune), che vengono intrappolati al loro interno tramite iniezione elettrica o ottica, sono confinati lungo tutte e tre le direzioni dello spazio; quindi i Qdots sono strutture zero-dimensionali dal punto di vista elettronico. La scala di lunghezze su cui avviene questo confinamento è minore della lunghezza d'onda di de Broglie di tali portatori, e questo ne implica la condizione di confinamento quantistico: gli stati di energia possibili per elettroni e lacune nei dot sono completamente quantizzati e separati tra loro da differenze in energia tipicamente maggiori della loro energia termica (rappresentata dal prodotto kBT , dove kB è la costante di Boltzman e T la temperatura in gradi Kelvin; per fare un esempio, a temperatura ambiente questo fattore è pari a 26 meV). La prima importante conseguenza di ciò è che i livelli energetici (quantizzati) disponibili per elettroni e lacune possono essere ingegnerizzati modificando dimensione, forma e composizione chimica delle nanostrutture. Le proprietà ottiche di emissione o assorbimento possono essere quindi modulate tramite questi parametri strutturali (Figura 1). Il secondo effetto è che la funzione densità degli stati, che rappresenta come sono distribuiti in energia gli stati di energia disponibili per i portatori di carica, viene profondamente modificata,

con importanti implicazioni e potenzialità dal punto di vista applicativo in dispositivi optoelettronici come light emitting devices (LED), celle solari e marker biologici. Date le svariate analogie con il comportamento dei singoli atomi o delle singole molecole, i Qdots vengono spesso descritti come atomi artificiali. Basti pensare allo spettro di luminescenza che, come per gli atomi, è caratterizzato da linee spettrali distinte e molto strette.

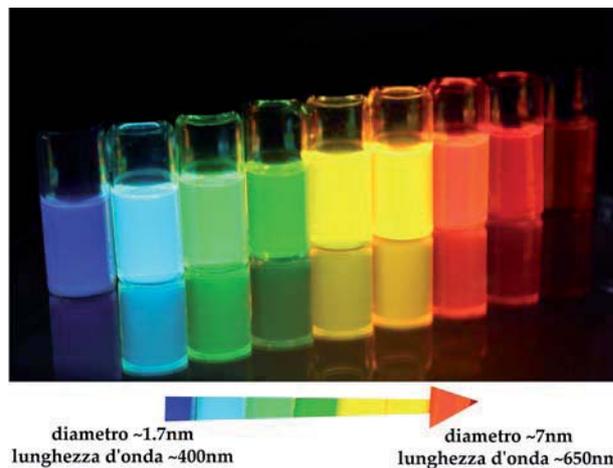


Figura 1: Fotoluminescenza al variare delle dimensioni di Qdots di CdSe eccitati da radiazione UV 350 nm e realizzati presso i laboratori del CNR-NANOTEC.

Confinamento Quantistico e Densità degli Stati

La proprietà principale dei Qdots, da cui deriva il loro nome, è il confinamento quantistico, che porta a una modifica della funzione densità degli stati vicino ai bordi di banda. Il diagramma schematico in Figura 2 mostra come evolve la funzione densità degli stati quando si riduce progressivamente la dimensionalità di un materiale semiconduttore. Se consideriamo un blocco voluminoso di materiale semiconduttore (chiamato comunemente bulk), dalle leggi della fisica statistica si ricava una funzione densità degli stati che dipende dalla radice quadrata dell'energia ($E^{1/2}$, Figura 2-a), e che quindi è una funzione continua con andamento parabolico. Se limitiamo una sola dimensione del materiale a lunghezze su scale nanometriche, dal bulk passiamo alla buca quantica (quantum well): in questo caso il movimento dei portatori di carica è limitato solo in quella direzione e la funzione densità degli

stati assume un andamento a gradino (Figura 2-b). Aumentando a due le direzioni di confinamento, si ottiene un filo quantico (quantum wire) e in queste condizioni, la funzione densità degli stati dipende da $E^{-1/2}$ (Figura 2-c). Nei Qdots il confinamento dei portatori su scala nanometrica avviene in tutte e tre le direzioni dello spazio (Figura 2-d): il risultato è che la funzione densità degli stati ha un andamento estremamente diverso. Assume infatti un comportamento tipico della funzione detta di Delta di Dirac, che assume un valore sempre nullo eccetto per determinati valori di energia, corrispondenti ai livelli energetici permessi nel sistema e, quindi, nel caso dei Qdots, quantizzati. Questo fa sì che gli stati elettronici di ogni livello energetico mostrino funzioni d'onda simili a quelle degli atomi. Le soluzioni dell'equazione d'onda di Schrödinger nei Qdots sono molto simili a quelle degli elettroni legati al nucleo di un atomo, dando così origine a picchi di emissione estremamente stretti e ben definiti. La spaziatura tra i livelli energetici disponibili all'interno delle singole bande diventa così dell'ordine di 10-100 meV.

Il confinamento quantistico risulta generalmente in un allargamento della banda proibita di energia al diminuire delle dimensioni dei Qdots. La banda proibita o band gap di un materiale semiconduttore è l'energia richiesta per creare un elettrone e una lacuna a riposo (cioè con energia cinetica nulla), ad una distanza sufficiente da rendere trascurabile la loro interazione Coulombiana. Quando un elettrone e una lacuna si avvicinano uno all'altro possono formare uno stato legato elettrone-lacuna che è detto eccitone. Questo eccitone si comporta come un atomo di idrogeno, l'unica differenza è che invece di un protone c'è una lacuna a formare il nucleo. Ovviamente, la massa di una lacuna è molto più piccola di quella di un protone, e questo influenza le soluzioni dell'equazione d'onda di Schrödinger. La distanza tra elettrone e lacuna è detta raggio di Bohr dell'eccitone (r_B). Se la massa dell'elettrone e quella della lacuna sono quelle efficaci, cioè tengono conto della struttura del semiconduttore in cui si trovano, r_B per semiconduttori bulk può essere espresso dall'equazione seguente:

$$r_B = \frac{(\hbar^2 \epsilon)}{e^2(1/m_e + 1/m_h)} \quad (1)$$

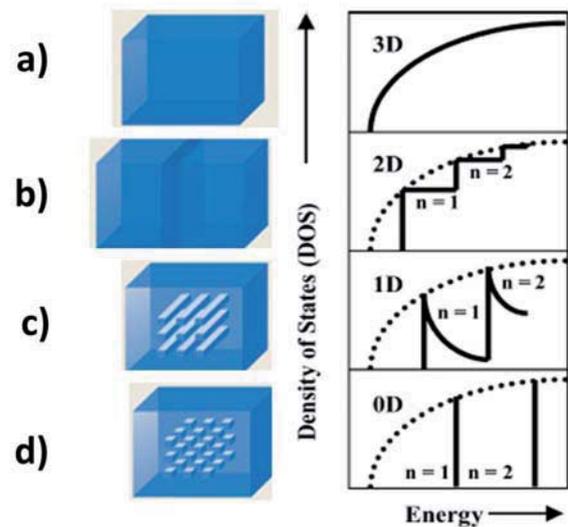


Figura 2: Evoluzione della funzione densità degli stati di energia in funzione della dimensionalità di un materiale semiconduttore: materiale bulk o 3D (a); quantum well (struttura 2D) dove il confinamento avviene in una sola direzione) (b); quantum wire (o struttura 1D dove i portatori sono confinati in due direzioni dello spazio) (c); quantum dot (o struttura 0D), in cui la relativa funzione assume un andamento discreto, a delta di Dirac, con valori diversi da zero solo per determinate energie rappresentate dai livelli ($n=1,2,\dots$) energetici ammissibili per tali sistemi.

dove ϵ , \hbar , ed e sono, rispettivamente, la costante dielettrica ottica del materiale, la costante di Planck ridotta e la carica dell'elettrone. Se invece consideriamo un Qdot di raggio R vicino ad r_B nel materiale, il movimento di elettroni e lacune è confinato spazialmente nel dot, portando ad un aumento del band gap del dot e conseguentemente dell'energia di transizione eccitonica, con spostamento della relativa luminescenza verso lunghezze d'onda minori. Il raggio dell'eccitone di Bohr è un valore di soglia e l'effetto di confinamento diventa sempre più forte se le dimensioni dei dots si riducono ad di sotto di tale limite. Uno strumento importante per determinare le proprietà dell'eccitone in un Qdot è rappresentato dai modelli teorici che sono stati sviluppati a questo proposito. Anche questo aspetto è stato oggetto di ricerca negli ultimi due decenni, portando a modelli sempre più accurati e precisi. Il modello più semplice che permette di visualizzare cosa succede ai livelli energetici di lacune ed elettroni confinati in un dot è il modello di approssimazione della massa efficace (effective

mass approximation-EMA), basato sul concetto 'Particle-in-Box Model' che, come dice il nome, considera una particella (cioè elettrone o lacuna) all'interno di una buca di potenziale dall'altezza di barriera infinita, che corrisponde al semiconduttore/matrice all'interno del quale si trova il dot. Questo modello, pur partendo da ipotesi molto semplificate, aiuta a chiarire l'origine fisica dei livelli energetici disponibili per le transizioni dei portatori di carica. Se consideriamo una particella libera nel nostro Qdot (box), la relazione tra la sua energia E e il suo vettore d'onda k risulta:

$$E = \frac{\hbar^2 k^2}{2m^*} \quad (2)$$

Il confinamento dell'eccitone in un Qdot con diametro R induce una variazione dell'energia di band gap (ΔE_g) rispetto a (2) pari a:

$$\Delta E_g = \frac{\hbar^2 \pi^2}{2\mu R^2} - \frac{1.8e^2}{\epsilon R} \quad (3)$$

dove μ è la massa ridotta di una coppia elettrone-lacuna. Il primo termine di questa equazione fornisce una relazione tra l'energia di confinamento e il raggio (R) del Qdot, mentre il secondo termine tiene conto dell'interazione Coulombiana e mostra conseguentemente una dipendenza da R^{-1} . Da questa relazione si vede come, l'energia della transizione principale per l'eccitone nel Qdot aumenti al diminuire delle dimensioni della nanostruttura. Modelli sempre più accurati e complessi sono stati sviluppati nel tempo, come quello della combinazione lineare degli orbitali atomici/molecolari, o quello dello pseudo-potenziale, dove al posto del potenziale reale dovuto agli ioni e agli elettroni di core del materiale si considera un potenziale equivalente (pseudo-potenziale, appunto) tale da originare un comportamento simile della funzione d'onda dell'elettrone nella regione compresa tra gli ioni e lontano dal core. Il trend in continua crescita delle capacità computazionali ha portato più di recente allo sviluppo di modelli atomistici che tengono conto anche della tipologia di interfaccia tra dot e barriera circostante.

Luminescenza nei Qdots

Quando arriva sul dot uno stimolo esterno quale un fascio di fotoni ad alta energia o un campo elettrico, l'elettrone e la lacuna nel suo interno si portano ad energie superiori rispetto a quelle a riposo. L'elettrone può ricombinarsi con la lacuna e riportarsi a uno stato energetico più basso, fino a raggiungere il livello fondamentale. L'energia in eccesso che risulta da questi processi può portare all'emissione di fotoni e quindi a luminescenza, oppure può essere trasferita al reticolo cristallino del materiale. Mentre nei semiconduttori bulk la fotoluminescenza è legata direttamente al profilo parabolico delle bande di energia nello spazio del vettore d'onda, nei Qdot questo processo è significativamente più complesso, in quanto condizionato dagli effetti di confinamento quantistico sui livelli energetici in generale, e da dinamiche dei portatori di carica intrinsecamente dipendenti dalle proprietà di materiali utilizzati per realizzare i Qdot stessi. Questo ha reso lo studio di tali nanostrutture in funzione di applicazioni pratiche come dispositivi elettroluminescenti, estremamente interessante dal punto di vista scientifico, e svariati esempi possono essere descritti a riguardo. A titolo di esempio, si riporta qui il fenomeno di blinking, cioè della intermittenza casuale nella luminescenza osservata per la prima volta in Qdots colloidali di CdSe a temperatura ambiente da Nirman *et al.* [3] (vedi Figura 3). Si tratta di un effetto che compromette tutti gli utilizzi dei Qdot in dispositivi per la conversione di energia, o nei diodi, nei laser, o in applicazioni di bio-imaging ed ha pertanto stimolato intensi studi per individuarne le due principali cause. La prima è legata al fatto che la luce induce dei processi di acquisizione di carica e perdita di carica da parte del core del nanocristallo: a seguito dell'irradiazione, i Qdots elettrostaticamente neutri, rilassano dallo stato eccitato emettendo fotoni, e quindi la nanoparticella si accende. Le particelle cariche invece appaiono dark (cioè non emettono) in quanto rilassano emettendo elettroni in un processo noto come emissione Auger. Accanto a questo meccanismo, però, il blinking è stato associato anche al caricamento e alla perdita di carica degli stati trappola presenti in superficie. Se questi stati non sono occupati, possono inter-

mettere elettroni ad alta energia, impedendo loro di rilassare verso uno stato a energia minore con conseguente emissione di luce [4].

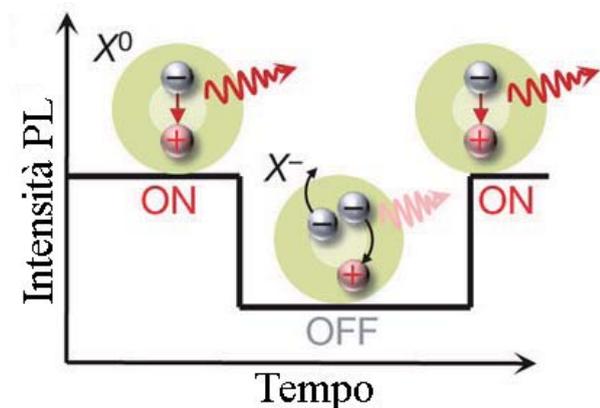


Figura 3: Descrizione schematica del meccanismo di blinking di fotoluminescenza (PL). Periodi di ON ed OFF corrispondono rispettivamente ad un Qdot in uno stato neutro (X^0) o in uno stato carico (X^-). Il modello riportato descrive un Qdot carico negativamente. Immagine adattata da [4].

3 Procedure di Fabbricazione

Per poter usufruire di queste proprietà sono necessari processi di sintesi chimica o di crescita epitassiale in grado di controllare le proprietà intrinseche di dimensione, forma, cristallinità e presenza di difetti o impurezze. Diversi sono gli approcci finora impiegati per la sintesi di Qdots. Oramai è prassi comune nella comunità scientifica distinguere tali procedure di crescita in due grandi categorie: il tipo *top-down* e il tipo *bottom-up*. La prima tipologia consiste nel partire da un materiale voluminoso e trasformarlo fino ad ottenere oggetti nanometrici, rimuovendo il materiale in eccesso. Tali processi si basano sull'utilizzo di tecniche litografiche per la definizione della geometria delle strutture da realizzare, in combinazione con processi di attacco chimico e fisico per la rimozione del materiale. Tipicamente si utilizza litografia a fascio elettronico, insieme a processi di incisione e rimozione (etching) mediante l'azione di un plasma reattivo e/o di soluzioni chimiche in fase liquida. Altre tecniche prevedono l'impiego di un fascio ionico o laser pur mantenendo lo stesso controllo selettivo delle forme e dimensioni del materiale nanostrutturato. Queste procedure presentano però

lo svantaggio di incorporare impurezze nel materiale di crescita o creare imperfezioni strutturali nella nanostruttura, e pertanto sono scarsamente utilizzate rispetto all'altro tipo di approccio. Gli approcci *bottom-up* invertono il modello alla base dei processi *top-down*. In questo caso invece di cominciare da un materiale bulk da assottigliare fino alla scala nanometrica, atomi e molecole vengono assemblati (*self-assembly*) su larga scala fino a creare nanostrutture. Tali approcci sono il nuovo paradigma nel campo della nanofabbricazione, poiché permettono di creare materiali nuovi, di dimensioni non raggiungibili con le metodologie *top-down* ed impiegando set-up di sintesi relativamente economici. Inoltre è possibile produrre materiali nanocristallini di purezza elevata o variarne intenzionalmente il drogaggio. Varie tipologie di nanostrutture sono state sintetizzate utilizzando approcci *bottom-up*. Ognuna di queste presenta proprietà chimico-fisiche, composizione e morfologie uniche. Tali approcci presuppongono una piena comprensione delle forze a corto raggio che si instaurano tra atomi e molecole, quali forze di Van-der-Waals, forze elettrostatiche e tutta una varietà di forze interatomiche ed intermolecolari. La sintesi di sistemi nanostrutturati di Qdots mediante *self-assembly* di atomi o molecole ha reso i processi *top-down* virtualmente obsoleti. Un consistente numero di diverse tecniche di *self-assembly* è stato impiegato finora per la sintesi di Qdots e questi possono essere ampiamente suddivisi in metodi chimici in fase liquida e metodi in fase vapore. Processi sol-gel, microemulsioni, decomposizione in soluzione ad alta temperatura e metodi elettrochimici sono tutti metodi che rientrano generalmente nella prima categoria. *Self-assembly* di nanostrutture mediante epitassia da fasci molecolari (MBE) anche nota come deposizione di film sottili o tecniche di sputtering rientrano invece tra quelli in fase vapore.

Processi chimici nei quali sostanze o soluzioni liquide sono mescolate e le conseguenti reazioni producono materiali solidi insolubili in un range dimensionale tra 1 e 100 nm, sono detti processi colloidali. Il materiale formatosi tenderebbe a precipitare se non fosse stabilizzato in soluzione da molecole surfattanti. Ad una miscela di sostanze solubili e insolubili è dato, appunto, il nome di colloide.

Metodi chimici in fase liquida I Qdots colloidal (CQdots) sono una classe di solidi cristallini di semiconduttore di dimensioni nanometriche rivestiti da molecole surfattanti che contribuiscono a mantenerli sospesi in soluzione. Essi rappresentano uno degli esempi più efficaci di come i materiali possono comportarsi alla nanoscala e per queste ragioni sono considerati come una piattaforma promettente su cui investire per lo sviluppo di numerose classi di dispositivi optoelettronici quali celle fotovoltaiche, fotodetectors o dispositivi emettitori di luce [5]-[7]. Mediante metodi chimici in fase liquida, è possibile crescere cristalli di semiconduttore in un normale pallone di reazione. Quando le dimensioni di tali cristalli sono dell'ordine delle decine di nm, la loro energia di bandgap diventa dipendente dalla dimensione, in tal modo permettendo una modulazione delle loro proprietà chimico-fisiche quali ad esempio assorbimento o emissione ottica semplicemente variando una o più dimensioni. Louis Brus è stato il primo a mostrare che quando il raggio del cristallo diventa inferiore al raggio dell'eccitone di Bohr del materiale bulk, l'energia di confinamento dell'eccitone modifica l'energia del bandgap [8]. L'incredibile potere di modulare le proprietà ottiche (e non solo) di un materiale semplicemente cambiandone le dimensioni in aggiunta ai vantaggi legati ai metodi di sintesi in fase liquida, ovvero la processabilità delle nanoparticelle risultanti, sono alla base delle enormi potenzialità dei CQdots nel settore della tecnologia dell'illuminazione a stato solido, per i display o per applicazioni in campo biologico. CQdots di semiconduttore possono essere preparati in mezzi acquosi usando agenti stabilizzanti, con procedure simili a quelle usate per la sintesi di nanoparticelle di oro. Questi approcci, tuttavia, danno nanocristalli molto difettosi e con dimensioni non omogenee tra loro. Più recentemente, sono state riportate procedure per la sintesi di CQdots di alta qualità, uniformi in dimensione e con strutture cristalline prive di difetti, preparati principalmente mediante approcci ad alta temperatura che prevedono l'utilizzo di surfattanti specifici (il cosiddetto metodo *hot-injection*). I processi in fase liquida generalmente si basano sulla precipitazione, favorita da una serie di parametri chimico-fisici, del nanomateriale solido in soluzione. Il processo di precipitazione

del solido principalmente coinvolge: una rapida fase di nucleazione del materiale solido seguito da una lenta fase di crescita [9]. A sua volta la nucleazione può essere classificata in i) omogenea, ii) eterogenea e iii) nucleazione secondaria. Considerata la forte dipendenza delle proprietà chimico-fisiche dagli aspetti morfologici, un fattore critico nella sintesi di CQdots risiede nella necessità di governare appieno la dimensione ed eventualmente la forma dei cristalli. Questo richiede la comprensione e il controllo dei processi chiave di nucleazione e crescita. La nucleazione omogenea si realizza quando, senza l'intervento di una interfaccia liquido-solido pre-esistente (i.e. cristalli pre-esistenti in soluzione), atomi o molecole di soluto si combinano (self-assembly) per formare piccoli semi (nuclei) fino a raggiungere una dimensione critica. Sono necessarie a tale scopo condizioni termodinamiche favorevoli. Una soluzione sovrassatura di specie reattive è un sistema altamente instabile, poiché esso rappresenta una condizione in cui le specie chimiche presenti non sono più stabili nella fase liquida a quelle alte concentrazioni, per cui precipitano per formare una nuova fase. Alla variazione di energia libera ΔG associata alla formazione dei nuclei contribuiscono due termini. Un primo termine è dato dall'energia libera che il sistema guadagna dalla formazione di una nuova fase, che dipende dal suo volume (dove le specie chimiche sono più stabili che in soluzione). Un secondo termine è rappresentato invece dall'energia libera che il sistema deve spendere per creare una nuova interfaccia (tra la nuova fase e quella pre-esistente in soluzione). Quest'ultimo è perciò associato ad una energia interfacciale e dipenderà pertanto da un fattore di superficie. Quindi, in un modello semplificato di nucleo sferico, ΔG può essere espresso nel modo seguente:

$$\Delta G = \Delta G_V + \Delta G_S \quad (4)$$

in una forma più esplicita:

$$\Delta G = \frac{4}{3}\pi r^3 \Delta g_V + 4\pi r^2 \gamma \quad (5)$$

dove r rappresenta il raggio dei nuclei, Δg_V l'energia libera di formazione di una nuova fase per unità di volume e γ la tensione interfacciale per unità di area. La formazione della nuova fase sarà un evento spontaneo se il termine di

volume è maggiore in valore assoluto al termine di superficie. In altre parole, quando il ΔG (equazione 4) è negativo, la nucleazione avverrà spontaneamente (seguendo la convenzione del segno di ΔG per un processo spontaneo). In Figura 4 si riporta la dipendenza dal raggio di entrambe i termini e del ΔG totale. E' possibile identificare un punto in cui il ΔG è massimo. Tale valore può essere considerato come la barriera di energia che il sistema deve superare per poter dar avvio alla nucleazione, mentre il raggio corrispondente a questo valore è indicato come raggio critico. Questa dimensione può essere determinata derivando (rispetto al raggio) l'equazione 5 ed eguagliandola a zero:

$$\Delta G^* = \frac{16\pi\gamma}{3\Delta G} \text{ ed } r^* = -\frac{2\gamma}{\Delta G_V} \quad (6)$$

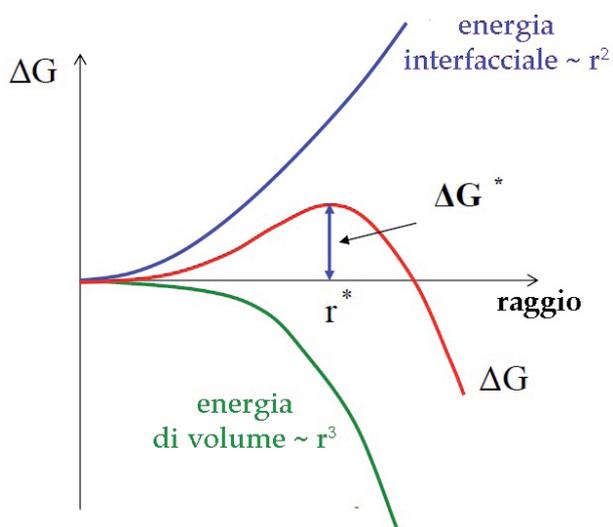


Figura 4: *Differenti contributi all'energia libera totale ΔG .*

La barriera energetica ΔG^* rappresenta l'energia di attivazione per la nucleazione e può essere compresa nel modo seguente: un nucleo che è appena più grande del raggio critico r^* , crescendo ulteriormente diminuirà l'energia libera totale della sua formazione. Questo si traduce nel concetto che qualsiasi nucleo che avrà dimensioni oltre il raggio critico, avrà buone probabilità di crescere sempre di più durante la reazione. Al contrario, un nucleo che ha dimensioni inferiori al raggio critico, si ri-solubilizzerà poiché in tal modo diminuirà l'energia libera di formazione. Il raggio critico dipende inoltre da una serie di parametri coinvolti nella sintesi dei nanocristalli, i più importanti dei quali sono la concentra-

zione delle specie reattive in soluzione e la temperatura [10]. La nucleazione eterogenea, d'altro canto, prende luogo sulla superficie di una fase condensata preesistente. E' intuitivo che tale processo sia energeticamente favorito rispetto alla nucleazione omogenea, poiché in generale il sistema non dovrà spendere energia per formare una nuova fase solida dal nulla. Particelle potranno nucleare su un supporto solido già esistente in soluzione quali semi cristallini di un altro materiale, impurezze solide, difetti reticolari, dislocazioni, bordi di grano, o molto spesso, sulle pareti stesse del contenitore. Il motivo è da associare al fatto che la creazione del nucleo distrugge parte del difetto preesistente, rilasciando così la sua energia ed abbassando quindi la barriera di attivazione. In termini chimici il processo trova una equivalenza nella catalisi eterogenea, nella quale la presenza del catalizzatore agisce nel senso di abbassare l'energia di attivazione. La nucleazione eterogenea è anche alla base della epitassia. La barriera di attivazione in tal caso dipenderà dalla energia di attivazione omogenea riscalata per un fattore di forma secondo la relazione:

$$\Delta G_{etero}^* = S(\theta) \cdot \Delta G_{omo}^* \quad (7)$$

Il fattore di forma $S(\theta)$ dipenderà dall'angolo di contatto (interfaccia) tra le due fasi solide, ovvero tra quella originaria e quella che si sta formando su di essa. Minore è l'angolo di contatto (cioè grande bagnabilità della seconda fase sulla prima), più piccolo sarà il valore del fattore di forma, e maggiore sarà invece il guadagno in energia di attivazione che si otterrà per la nucleazione eterogenea [11],[12]. La nucleazione eterogenea è alla base della sintesi di eterostrutture, ovvero nanocristalli costituiti da due o più domini nanodimensionati uniti almeno su una faccetta cristallina (vedi Figura 5).

La sintesi in fase liquida di tali sistemi multi-componente prevede un meccanismo a due fasi, ovvero la nucleazione omogenea e la crescita di un primo materiale che successivamente fungerà da seme per la nucleazione eterogenea, e quindi per la crescita, di un secondo materiale. Indipendentemente dalla tecnica di sintesi, tale procedura può essere realizzata in un unico processo sintetico o in due processi separati. Il secondo

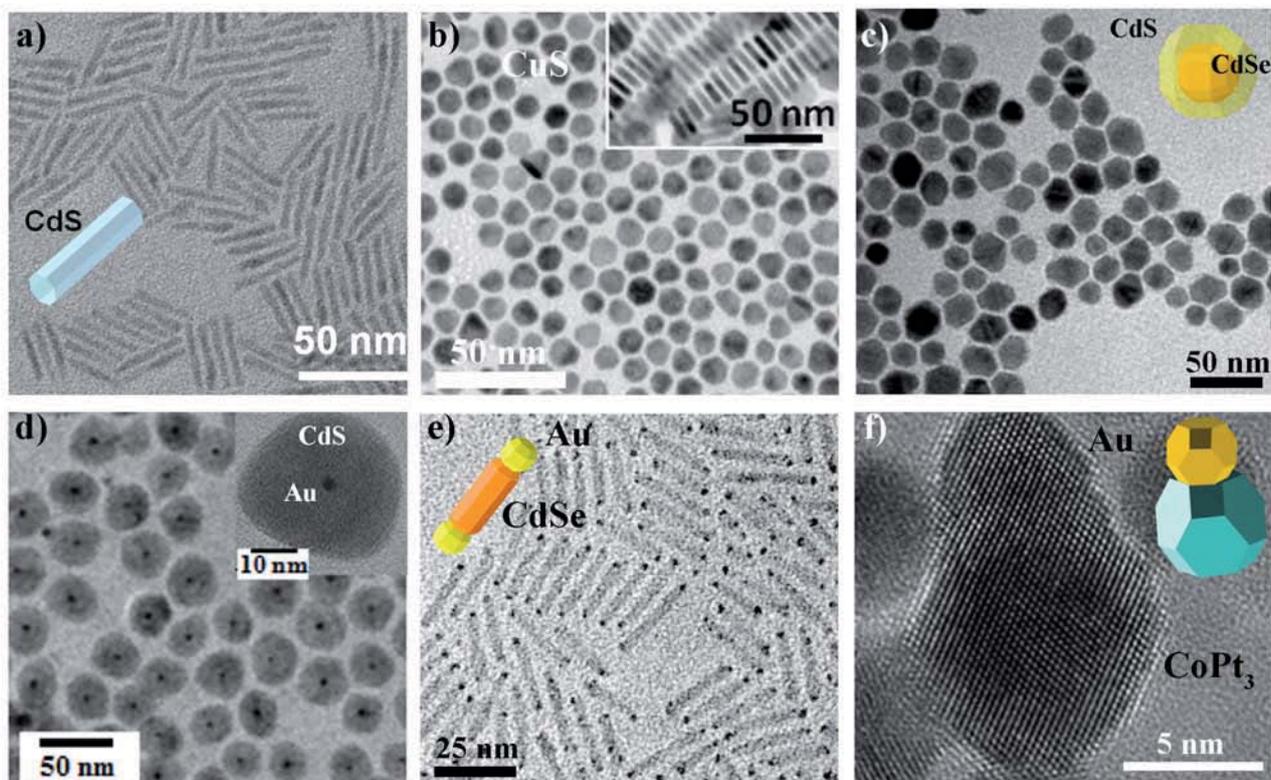


Figura 5: Immagini di nanocristalli a singolo materiale ed eterostrutture, ottenute mediante analisi in microscopia elettronica a trasmissione. Oltre a Qdots sferici core/shell (c), il panel mostra nanoarchitetture di forme differenti. Immagini tratte rispettivamente dai seguenti riferimenti: (a) [15] (b) [16] (c) [17] (d) [18] (e) [15] (f) [19].

materiale potrà ricoprire in modo omogeneo il primo formando una shell uniforme, nel caso di una completa bagnabilità (significativamente influenzata dai surfattanti superficiali) tra i due sistemi, oppure sviluppando regioni separate su una o più faccette del seme (vedi Figura 6).

Si potranno realizzare quindi eterostrutture core-shell (nel primo caso) o sistemi multi-decorati che espongono una o più segregazione di fasi (nel secondo). La nucleazione è seguita dalla fase di crescita, ovvero le specie non reagite rimaste in soluzione si aggiungono in maniera selettiva ai nuclei precedentemente formati [13]. Allo scopo di compensare gli elevati rapporti superficie-volume esistenti nei CQdots, molecole surfattanti come rivestimento delle superfici dei nanocristalli sono utilizzate nella maggior parte degli approcci sintetici. Questi ligandi sono costituiti da molecole tensioattive con una lunga catena idrocarburica apolare (C12 ÷ C18) ed una testa idrofilica polare. In genere i surfattanti più utilizzati sono: acidi carbossilici, ammine, fosfine e tioli. Questi sono in grado di modulare la reattività dei vari precursori molecolari, influenzando in maniera significativa gli

step di nucleazione e crescita [14]. In aggiunta la fase cristallina, la morfologia e le proprietà ottiche dei CQdots sono condizionate dalla scelta dei surfattanti di sintesi. La scelta è guidata e determinata dalla chimica di coordinazione dettata dalla superficie dei nanocristalli. Molte strategie a basso costo e di alta qualità produttiva sono possibili per la sintesi di nanoparticelle di forma e dimensione controllata. Allo scopo di evitare nucleazioni secondarie, mentre è in corso la crescita del nanomateriale, il concetto chiave alla base dei diversi meccanismi di sintesi è la netta separazione tra gli step di nucleazione e quello di crescita. Questa è pienamente realizzata nella maggior parte dei processi di sintesi attualmente impiegati quali: decomposizione termica realizzata mediante aggiunta del precursore già a caldo (*hot-injection*) o per riscaldamento da temperatura ambiente (*heating-up*), metodi sol-gel, micelle, sintesi idrotermica, etc. Il metodo *hot-injection* prevede l'iniezione ad alte temperature (200-350°C), dei precursori di reazione, organici od inorganici, nella miscela di reazione contenente un solvente organico ed uno o più surfattanti. Secondo tale approccio, la temperatura

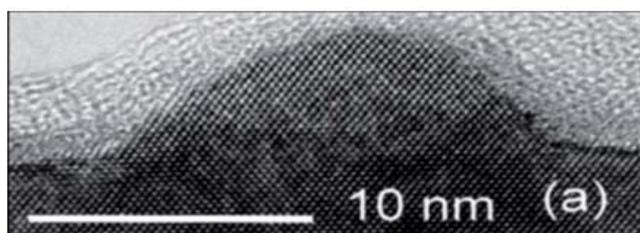
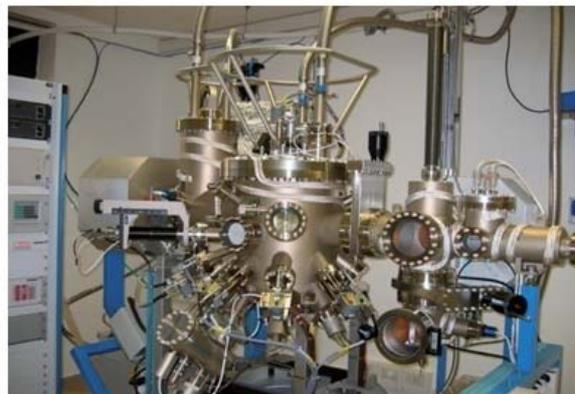
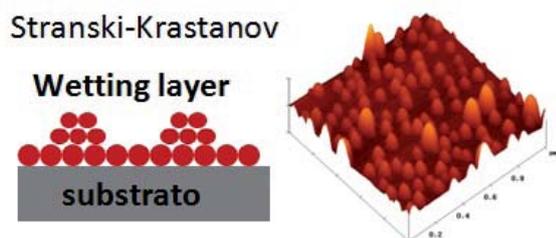
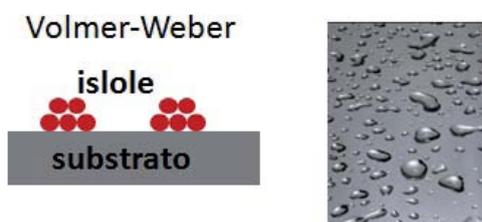


Figura 6: Possibili deposizioni eterogenee di un secondo materiale (in rosso) su un materiale preesistente (substrato, in grigio). Il modello Frank van der Merwe prevede una crescita per strati atomici successivi che si manifestano in terrazzamenti sulla superficie. Il modello Volmer-Weber, caratterizzato dalla non bagnabilità del substrato da parte dello strato depositato, evolve a isole, come delle gocce d'acqua su una superficie. L'ultimo regime è quello Stranski-Krastanov, dove lo strato dapprima bagna il substrato crescendo in modo planare e poi evolve in una nanostruttura tridimensionale a causa dell'elevata energia interna accumulata. In alto a destra si può vedere il sistema di Epitassia da Fascio Molecolare dei laboratori CNR-Nanotec, mentre in basso a destra un'immagine di un quantum dot di InAs in matrice di GaAs acquisita in sezione con un microscopio elettronico a trasmissione ad alta risoluzione.

d'iniezione è un fattore critico, poiché regola la decomposizione dei precursori molecolari. Fra i semiconduttori più comuni e studiati con tale metodo ci sono quelli del tipo II-VI (e.g., ZnS, ZnSe, CdS, CdSe, CdTe e HgS) e quelli del tipo III-V (e.g., GaAs, InP e InAs). Processi *sol-gel* sono anch'essi stati ampiamente usati per diversi anni per la sintesi di CQdots [20],[21]. Un sol è una dispersione colloidale liquida nella quale le particelle disperse hanno le dimensioni tra 1 e 100 nm, mentre un gel è una struttura formata da una rete continua di particelle legate tra loro. In un tipico esperimento, il precursore metallico solubilizzato in un mezzo acido o basico idrolizza e condensa per formare il sol, questo è seguito dalla sua polimerizzazione per formare un network (gel). Tale approccio non permette un grande controllo nella distribuzione delle dimensioni e della qualità cristallina delle nanoparticelle sintetizzate. Anche i metodi *idrotermici*, basati sulla cristallizzazione di sali inorganici da una soluzio-

ne acquosa a pressione e temperatura controllata, hanno avuto un ruolo importante nella produzione di CQdots. Variazioni di pressione, temperatura e tempi di reazione influenzano forma e dimensioni delle nanoparticelle. La solubilità di tali composti inorganici tipicamente diminuisce riducendo la temperatura e/o la pressione [22].

Metodi in fase vapore Tali metodi consistono principalmente nell'utilizzo di tecniche epitassiali alla base dei processi di realizzazione di tutti i dispositivi a semiconduttore, che nel caso dei Qdots vengono opportunamente manipolate verso una crescita auto-organizzata. Fin dagli inizi del secolo scorso, erano stati osservati processi di crescita di film sottili mal riusciti in quanto caratterizzati da ondulazioni ed elevata rugosità superficiale. Solo negli anni '90 si comprese che tali morfologie erano il mezzo per ridurre lo stress accumulato nella deposizione del film e, se opportunamente controllate, pote-

vano portare alla realizzazione di array uniformi di nanostrutture otticamente attive. La crescita auto-organizzata può essere causata da differenze nelle cinetiche di trasporto e di incorporazione per specie differenti, da effetti di segregazione indotti da strain o, come accade per la crescita in regime Stranski-Krastanov (SK), da transizioni morfologiche indotte direttamente dallo strain, cioè dalla deformazione elastica del reticolo cristallino. In generale, la formazione di interfacce inorganiche multilayer realizzate mediante tecniche di crescita in fase vapore o di eterostrutture prodotte con metodi chimici in fase liquida può essere interpretata termodinamicamente. Quando un secondo materiale (qui indicato con "2") deve crescere su un primo materiale preesistente (qui indicato con "1"), il segno del ΔG_S , ovvero la variazione dell'energia libera superficiale, determinerà la modalità di crescita:

$$\Delta G_S = \gamma_1 - \gamma_2 + \gamma_{1,2} \quad (8)$$

dove γ_1 e γ_2 rappresentano le energie superficiali associate coi rispettivi materiali, mentre $\gamma_{1,2}$ rappresenta l'energia interfacciale tra i due. Se il materiale 2 espone più basse energie superficiali ($\gamma_2 < \gamma_1$) e/o possiede un mismatch reticolare con 1 alquanto basso ($\gamma_{1,2}$ piccolo), allora la deposizione sarà uniforme ed omogenea ($\Delta G > 0$, regime Frank-van der Merwe, Figura 6a). Al contrario se il materiale 2 espone energie superficiali più grandi ($\gamma_2 > \gamma_1$) e/o possiede parametri reticolari ben diversi da 1 ($\gamma_{1,2}$ grande), allora questi crescerà sul materiale 1 formando domini isolati in modo da minimizzare l'energia interfacciale ($\Delta G_S < 0$, regime Volmer-Weber, Figura 6b). Una ulteriore possibilità può riguardare la crescita a due stadi in regime Stranski-Krastanov (Figura 6c). In questo caso il materiale 2 inizialmente cresce rivestendo omogeneamente 1 ($\Delta G_S > 0$), ma aumentando lo spessore si intensifica lo strain interfacciale per cui, raggiunta una soglia critica di spessore, il materiale 2 segrega cominciando a formare piccole isole ($\Delta G_S < 0$) [23]. La crescita Stranski-Krastanov è tipica di sistemi come InGaAs/GaAs o InSb/GaSb, entrambi caratterizzati da un mismatch reticolare superiore al 6%. La tecnica di epitassia da fasci molecolari (Molecular Beam Epitaxy-MBE) rappresenta uno degli approcci maggiormente

in uso per la deposizione di composti o leghe nanostrutturati in condizioni di ultra alto vuoto ($\sim 10^{-10}$ Torr). Un fascio di atomi o molecole è originato per evaporazione. Il fascio può essere formato da solidi (i.e. Ga ed As elementare sono impiegati per la crescita di Qdots di GaAs) od una combinazione di solidi e gas (i.e. AsH_3 , PH_3 o metallorganici quali tri-metil- o tri-etil-gallio). La tecnica MBE (Figura 6) è stata utilizzata per autoassemblare Qdots di semiconduttore III-V [24], e II-VI [25], sfruttando mismatch elevati (6-7%) in modo da favorire un regime Stranski-Krastanov.

4 Applicazioni

Display a Qdots

Fino a pochissimi anni fa, la maggior parte dei monitor per computer e TV utilizzavano display a cristalli liquidi (LCD) caratterizzati da elevati consumi di potenza elettrica. Un'alternativa a questa tecnologia sono stati i display a diodi ad emissione di luce basati su semiconduttori organici (OLED - Organic Light Emitting Diode), molto più luminosi ed efficienti dal punto di vista energetico. I costi elevati e i limitati tempi di vita di questi dispositivi ne restringevano però l'utilizzo a sistemi di piccole dimensioni. In questo campo applicativo l'utilizzo dei Qdots presentava potenzialità rilevanti come un consumo di potenza fino a cinque volte inferiore rispetto ai sistemi LCD, luminosità e tempi di vita superiori a quelli degli schermi OLED, e costi di produzione decisamente inferiori. Nel 2011, i ricercatori della Samsung Electronics hanno dimostrato il primo display full-color basato su tecnologia a Qdot [26]. In questo prototipo, una soluzione di Qdots veniva distribuita su un substrato di silicio. I Qdots usati corrispondevano a diverse combinazioni core/shell/shell di CdSe/CdS/ZnS, ingegnerizzate per ottenere i tre colori primari, rosso, verde e blu. Il solvente veniva fatto evaporare e successivamente, uno stampo in plastica modellato in modo da esibire sulla superficie un array a strisce, veniva compresso sul substrato. Di seguito, lo stampo veniva staccato dal substrato, rimuovendo con sé i Qdots trasferiti sulle strisce, per poi farlo aderire su un substrato desiderato, in vetro o plastica. In tal modo risultava possibile trasferire strisce di Qdo-

ts sul substrato di destinazione. In un display a colori, ogni pixel contiene dei subpixel dei colori fondamentali, rosso, verde e blu. Questi colori sono combinati in intensità variabili per produrre milioni di colori. Usando ripetutamente la tecnica di stampaggio proposta, si può creare un pattern ripetuto di strisce rosse, verdi e blu. Le strisce venivano trasferite direttamente su un array di transistor a film sottili di ossido amorfo di Hf, In o Zn, che consentono di avere correnti più elevate e stabili dei convenzionali transistor in silicio amorfo. Il display risultante mostrava così subpixel larghi $50\ \mu\text{m}$ e lunghi $100\ \mu\text{m}$, abbastanza piccoli da poter essere impiegati negli schermi dei cellulari. Dopo qualche anno, nel 2013, i primi display combinati con LCD e Qdots, in cui il ruolo di questi ultimi è di migliorare la resa cromatica del dispositivo agendo essenzialmente come dei filtri, sono stati commercializzati da Sony e, ad oggi, diverse sono le compagnie che ne hanno avviato la produzione. Le attuali stime di mercato per dispositivi e componenti a Qdots per i prossimi anni sono altissime, con previsione per la domanda di questi nuovi materiali a livelli di diverse tonnellate. Questa tipologia di applicazione è quella per cui si prevede il principale mercato, anche se ancora molto lavoro deve essere fatto rispetto alla resa dei processi e alla durata dei sistemi realizzati per sostituire completamente le tecnologie esistenti.



Figura 7: Prototipo di display a colori basato sulla combinazione di Qdots di diversi colori e transistor a film sottili di ossido di Hf, In e Zn.

Laser a Qdot per telecomunicazioni

I Qdots stanno anche emergendo come materiale promettente per altri tipi di dispositivi nel campo delle telecomunicazioni. Una delle più recenti e importanti innovazioni scientifiche e tecnologiche che hanno interessato il campo della fotonica è stato il laser a Qdot, dove le nanostrutture di materiale semiconduttore sono usate come mezzo attivo di guadagno nella tipica giunzione tra semiconduttori p ed n che costituisce un laser. Fin dalla sua prima teorizzazione da parte di Arakawa e Sakaki, nel 1982 [1], questo tipo di dispositivo ha infatti stimolato un ampio filone di ricerca, volto, da un lato, alla comprensione e determinazione delle proprietà elettroniche di queste strutture, e, dall'altro, alla definizione e ottimizzazione dei processi tecnologici necessari per la loro fabbricazione. Effettivamente, i numerosi vantaggi introdotti dall'utilizzo dei Qdots come mezzo attivo per un laser a semiconduttore rispetto alle più tradizionali quantum well, sono controbilanciati dalla maggiore complessità dei fenomeni fisici che si verificano nelle nanostrutture e dalle oggettive difficoltà tecnologiche che ne hanno a lungo impedito la realizzazione pratica. Le proprietà dei livelli energetici per elettroni e lacune nei Qdots descritte nei paragrafi iniziali di questi lavoro si traducono, se trasferite a un dispositivo laser, in una serie di prestazioni migliorate rispetto a un dispositivo classico bidimensionale, a quantum well. L'effetto più importante si può osservare dallo spettro di guadagno di questo materiale in funzione dell'energia, che risulta essere molto alto (fino a $1.5 \cdot 10^6\ \text{cm}^{-1}$), simmetrico ed estremamente stretto ($0.15\ \text{meV}$ per un singolo dot). Questo, a sua volta, si traduce in un aumento del guadagno differenziale e della temperatura caratteristica, in una riduzione della densità di corrente di soglia (fino a pochi A/cm^2) e nella soppressione dell' indesiderato effetto chirp, che porta alla modulazione della lunghezza d'onda di lasing per effetto della corrente iniettata. La possibilità di ingegnerizzare la lunghezza d'onda di emissione del dispositivo tramite una opportuna progettazione delle dimensioni, della forma e della composizione dei suddetti dot ha permesso inoltre di accedere a range spettrali che non erano accessibili con altri sistemi. Proprio in questo contesto, è diventata

sempre più stringente l'esigenza di sorgenti laser per sistemi telecom a 1.3 μm o 1.55 μm , dove la trasmissione dei dati in fibra ottica è al massimo della sua efficienza. I Qdots di InGaAs in matrice di GaAs hanno permesso di accedere a questa lunghezza d'onda, in precedenza limitata dalla scarsità di materiali semiconduttori aventi banda proibita in questa regione spettrale e accordabili reticolarmente a un substrato di utilizzo comune in optoelettronica, come GaAs o InP. A fronte delle migliori prestazioni del dispositivo e della maggiore libertà di progettazione, il laser a Qdots ha presentato delle problematiche tecnologiche non indifferenti, insite soprattutto nella realizzazione di un array di nanostrutture con un elevato grado di uniformità nella forma, dimensione e composizione, e con densità elevate. Le eventuali disuniformità tra i dot di uno stesso piano causano un allargamento dello spettro di luminescenza del materiale e quindi del suo spettro di guadagno. Inoltre, i Qdots rappresentano sempre un ricoprimento parziale della superficie su cui vengono realizzati, per cui il loro alto valore di guadagno di materiale deve essere corretto da un opportuno coefficiente di ricoprimento per avere il corrispondente guadagno modale. In numeri, considerati un array di dot con densità dell'ordine di $10^{10} \text{ dot}/\text{cm}^2$ di dimensioni medie 7 nm, se partiamo da un guadagno di saturazione del materiale di 10^5 cm^{-1} arriviamo a 10 cm^{-1} come guadagno modale. Da qui l'importanza di controllare la densità delle nanostrutture e portarla a valori più alti possibile, insieme ad un elevato grado di uniformità, sempre mantenendo elevati standard qualitativi del materiale. A questo scopo sono risultate particolarmente efficienti le tecniche di crescita epitassiale, nella forma dell'epitassia da fascio molecolare MBE e della deposizione chimica da fase vapore metallorganica (Metal organic chemical vapor deposition-MOCVD).

Qdots in bioimaging applications: FRET

Una intensa attività di ricerca si è di recente focalizzata verso l'utilizzo delle singolari proprietà ottiche tipiche dei Qdots per usi nel campo dell'imaging biologico [27],[28]. Finora, il bioimaging ha concentrato la maggior parte dei suoi sforzi sui tradizionali coloranti organici, anche se tutta

una serie di punti deboli sono associati al loro impiego. Ad esempio è ben noto: i) il fenomeno dell'auto-fluorescenza, nella regione del visibile, della cellula, come conseguenza degli effetti di assorbimento e di scattering che hanno l'emoglobina e l'acqua, la quale tende a mascherare il segnale di diversi coloranti organici; ii) i coloranti organici sono piuttosto instabili sotto fotirraggiamento con conseguenti brevi tempi di osservazione; iii) presentano una stretta finestra di eccitazione, che rende difficoltosa l'eccitazione contemporanea di diversi coloranti; iv) sono fortemente sensibili alle condizioni esterne, particolarmente alle variazioni di pH; infine v) hanno uno spettro di emissione piuttosto slargato. Quest'ultima caratteristica crea sovrapposizioni tra diversi canali di fotorivelazione rendendo così difficile la quantificazione di ognuno di essi. D'altro canto i Qdots rivestono un grande interesse in campo bio, per una serie di ragioni quali: i) coefficienti di estinzione estremamente alti; ii) elevata efficienza di luminescenza; iii) scarsa tendenza al photo-bleaching; iv) emissione ed assorbimento modulabili con le dimensioni; v) larghissimi spettri di assorbimento contrapposti a stretti picchi di emissione (20-30 nm); vi) possibilità d'impiego di diverse famiglie di Qdots contemporaneamente senza impedimenti di reciproca interferenza; vii) presentano una tossicità inferiore a quella di un classico colorante organico; viii) possono essere funzionalizzati con differenti specie bio-attive e bio-compatibili. Infine l'utilizzo di Qdots che emettono nella regione del vicino-IR (NIR), ix) permette di evitare la sovrapposizione della loro luminescenza con l'autofluorescenza della cellula. Tra i materiali nanostrutturati, l'emissione di colori basata sui Qdots e modulabile con le dimensioni, rappresenta l'opzione che gode delle maggiori potenzialità per lo sviluppo di approcci di bio-etichettatura multicolore. Ed in particolare, Qdots colloidali di CdSe sono i materiali maggiormente investigati per studi di imaging e diagnosi in vivo ed in vitro, in alternativa ai coloranti organici. Come già sopra descritto, i vari approcci sintetici in fase liquida fanno sì che tali specie mantengano un layer di surfattanti adsorbito sulla superficie che ne favorisce la processabilità in soluzione. I ligandi usati in fase di sintesi possono quindi essere scambiati successivamente per cambiarne le

proprietà di solubilità, influenzare e controllare approcci diversi di deposizione, oppure accelerare l'interazione tra i vari CQdots in modo da favorire la formazione di film omogenei. Inoltre, a causa del core inorganico di semiconduttore, i CQdots sono estremamente più resistenti alla degradazione causata da photo-bleaching rispetto ad un qualsiasi colorante. Questa stabilità può essere ulteriormente intensificata, rivestendo i core con una shell di materiale inorganico di più grande band gap. Diverse sono le tecniche di bioimaging, ognuna delle quali sfrutta una delle diverse proprietà dei Qdots: *i*) stretta e modulabile fluorescenza che va dall'UV al NIR; *ii*) trasferimento di energia per risonanza (Förster Resonance Energy Transfer-FRET), che consiste in un trasferimento di energia tra fluorofori; *iii*) spettroscopia Raman amplificata da superfici; *iv*) radio opacità, proprietà paramagnetiche e/o utilizzo come agenti di contrasto. Tra queste vogliamo soffermarci sull'impiego di CQdots in combinazione con il trasferimento di energia per risonanza (FRET) come sistema nanosensore. Il FRET consiste nel trasferimento di energia elettronica da un cromoforo donore ad uno accettore [29]. Tale trasferimento si realizza tra interazioni intermolecolari dipolo-dipolo su distanze tra 1 e 10 nm rendendo tale tecnica un mezzo ideale per determinare ad esempio distanze tra biomolecole, o per rivelare la presenza e l'interazione tra due fluorofori, o come segnale di conferma che un farmaco ha raggiunto il suo target. CQdots sono solitamente usati come sorgenti di elettroni (donori) in interazioni FRET, poiché variazioni della intensità di emissione possono essere facilmente quantificate in presenza di molecole che tendono ad accettare elettroni (accettrici). In base alla teoria di Förster, la velocità di trasferimento di energia dipende dalla sovrapposizione spettrale tra l'emissione del donore e l'assorbimento dell'accettore e dalla disposizione spaziale donore-accettore. *Mattoussi* e collaboratori hanno investigato il trasferimento di energia tra CQdots donori e molecole organiche accettrici [30]. Essi riportano un complesso costituito da una proteina capace di legare il maltosio (Maltose Binding Protein-MBP) covalentemente legata ad un fluoroforo organico (cianina, accettore). Tale complesso viene immobilizzato sulla superficie di CQdots core/shell di CdSe/ZnS solubili

in acqua. L'immobilizzazione è guidata sia da interazioni elettrostatiche tra MBP e la superficie dei nanocristalli, sia dalla coordinazione con i cationi metallici superficiali dei CQdots. Si è notato che, aumentando la frazione di maltosio in soluzione, l'emissione del colorante incrementava, mentre quella del CQdots diminuiva. In aggiunta, l'intensità di emissione da parte della proteina MBP etichettata con fluoroforo risultava dipendente dal colore emesso dal CQdot, ovvero dalla dimensione di questi e dalla sovrapposizione spettrale. *Mattoussi* e collaboratori hanno anche studiato CQdots come accettori in un processo FRET. Analogamente, *Zhang* e collaboratori utilizzarono singoli CQdots core-shell di CdSe-ZnS coniugati a sonde di DNA per investigare parti di DNA implicate in malattie genetiche [31]. Essi misero a punto un sistema a tre elementi (vedi Figura 8) basato rispettivamente su: 1) un frammento di DNA a singola catena etichettato con una molecola di biotina (sonda di cattura), capace di legare 2) un CQdot funzionalizzato in superficie con streptavidina. La streptavidina è una proteina tetramerica (60KDa) con un'altissima affinità per la biotina ($K_a=10^{15} M^{-1}$); la loro interazione biologica è il più forte legame di tipo non covalente che si conosca. 3) Un terzo elemento del sistema è costituito da un altro frammento di DNA a singola catena etichettato con un fluoroforo (sonda reporter). In presenza di un DNA target, le due sonde gli si legano in maniera complementare in tal modo portando in prossimità il CQd donore ed il fluoroforo accettore. In questo caso il CQdot trasferisce non-radiativamente energia all'accettore e questi emette un fotone: l'emissione conferma la presenza del DNA target. In assenza dell'accettore, l'efficienza FRET diventa zero, poiché tutta la fluorescenza sarebbe ascrivibile unicamente al donore.

5 Conclusioni

Negli ultimi anni, la ricerca orientata alla crescita di Qdots epitassiali e colloidali ed allo studio delle loro proprietà chimico-fisiche, ha fatto enormi passi in avanti, evolvendo da semplici studi pionieristici ad impieghi estremamente innovativi in campo opto-elettronico o all'ottenimento di accresciute performance in campo dispositivo. Lo sviluppo di metodi di sintesi ormai

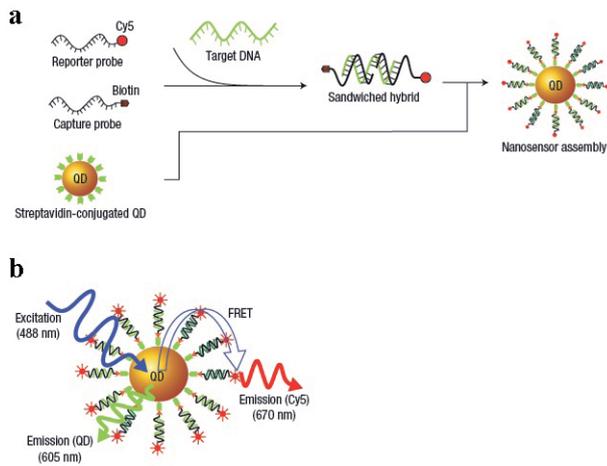


Figura 8: (a) Rappresentazione schematica di un nanosensore in presenza di un DNA target, basato su meccanismo FRET. (b) Luminescenza del fluoroforo accettore mediante meccanismo FRET, in seguito ad eccitazione del Qdot donore. Immagine adattata da [31].

ben consolidati ha permesso di poter disporre di una vasta gamma di proprietà derivanti dalla ingegnerizzazione del band-gap di questi materiali, ottenuta spesso tramite semplici procedure quali ad esempio approcci chimici in fase liquida. La forza di tali concetti è stata ben espressa e provata in applicazioni optoelettroniche che stanno rivoluzionando il nostro modo di vivere e di comunicare, quali dispositivi elettroluminescenti per illuminazione e display, laser o dispositivi fotovoltaici, o biosensori e strumenti per bio-imaging. Tuttavia, nonostante la mole di conoscenze sviluppate, diversi sono i meccanismi fondamentali la cui comprensione risulta piuttosto incompleta. Questo si traduce ad esempio, in differenze nelle prestazioni di un dispositivo pur utilizzando materiali apparentemente simili. Non di rado tali fattori limitanti derivano da fenomeni di “intrappolamento” dei portatori di carica (elettroni e lacune), a cui contribuiscono difetti strutturali del materiale, ma principalmente gli atomi presenti in superficie non opportunamente “passivati”, ovvero non impegnati in un legame chimico con un altro materiale o con una molecola surfattante. In questi casi orbitali vuoti o cariche in eccesso (a seconda che si tratti di un catione od anione) sono offerte dalle smisurate superfici che quindi interagiranno con le cariche foto/elettricamente generate. Nuove architetture basate su sistemi multi-materiale e l’evoluzione di alcune tra le tecniche di crescita descritte

stanno contribuendo in maniera significativa al miglioramento delle interfacce risultanti ed al raggiungimento di un accurato controllo della cristallinità del materiale. C’è ancora tanto spazio, quindi, nell’infinitamente piccolo universo dei quantum dot rivelato dalle nanotecnologie, per migliorarne l’efficienza, sviluppare tecniche di fabbricazione applicabili su scale sempre più elevate, utilizzare nuovi materiali e nuove configurazioni, comprenderne i tempi di vita e l’interazione con l’organismo umano, e rispondere a tutte le domande che la comunità scientifica e tecnologica ha ancora su questa tematica così giovane eppure così tanto versatile della ricerca.



- [1] Y. ARAKAWA AND H. SAKAKI: “Multidimensional quantum well laser and temperature dependence of its threshold current”, *Appl. Phys. Lett.* **40** (1982) 939.
- [2] A.L. EFROS AND A.L. EFROS: “Interband absorption of light in a semiconductor sphere”, *Sov. Phys. Semicond.* **16** (1982) 772.
- [3] M. NIRMAL, B.O. DABBOUSI, M. G. BAWENDI, J.J. MACKLIN, J.K. TRAUTMAN, T.D. HARRIS AND L.E. BRUS: “Fluorescence intermittency in single cadmium selenide nanocrystals”, *Mater. Chem. Phys. Nature* **383** (1996) 802.
- [4] C. GALLAND, Y. GHOSH, A. STEINBRÜCK, M. SYKORA, J.A. HOLLINGSWORTH, V.I. KLIMOV AND H. HTOON: “Two types of luminescence blinking revealed by spectroelectrochemistry of single quantum dots”, *Nature* **479** (2011) 203.
- [5] G. KONSTANTATOS, I. HOWARD, A. FISCHER, S. HOOGLAND, J. CLIFFORD, E. KLEM, L. LEVINA AND E.H. SARGENT: “Ultra-sensitive solution-cast quantum dot photodetectors”, *Nature* **442** (2006) 180.
- [6] S. KEULEYAN, E. LHUILLIER, V. BRAJUSKOVIC AND P. GUYOT-SIONNEST: “Mid-infrared HgTe colloidal quantum dot photodetectors”, *Nat. Photon.* **5** (2011) 489.
- [7] L. QIAN, Y. ZHENG, J. XUE AND P.H. HOLLOWAY: “Stable and efficient quantum-dot light-emitting diodes based on solution-processed multilayer structures”, *Nat. Photon.* **5** (2011) 543.
- [8] L.E. BRUS: “Electron-electron and electron-hole interactions in small semiconductor crystallites: The size dependence of the lowest excited electronic state”, *J. Chem. Phys.* **80** (1984) 4403.
- [9] N.T.K. THANH, N. MACLEAN AND S. MAHIDDINE: “Mechanisms of Nucleation and Growth of Nanoparticles in Solution”, *Chem. Rev.* **114** (2014) 7610.
- [10] C. BURDA, X. CHEN, R. NARAYANAN AND M.A. EL-SAYED: “Chemistry and Properties of Nanocrystals of Different Shapes”, *Chem. Rev.* **105** (2005) 1025.

- [11] N. TAUB, O. KRICHEVSKI AND G. MARKOVICH: "Growth of Gold Nanorods on Surfaces", *J. Phys. Chem. B* **107** (2003) 11579.
- [12] H.-Y. TUAN, D.C. LEE, T. HANRATH AND B.A. KORGEL: "Germanium Nanowire Synthesis: An Example of Solid-Phase Seeded Growth with Nickel Nanocrystals", *Chem. Mater.* **17** (2005) 5705.
- [13] B.L. CUSHING, V.L. KOLESNICHENKO AND C.J. O'CONNOR: "Recent Advances in the Liquid-Phase Syntheses of Inorganic Nanoparticles", *Chem. Rev.* **104** (2004) 3893.
- [14] R. GARCÍA-RODRÍGUEZ, M.P. HENDRICKS, B.M. COSSAIRT, H. LIU AND J.S. OWEN: "Conversion Reactions of Cadmium Chalcogenide Nanocrystal Precursors", *Chem. Mater.* **25** (2013) 1233.
- [15] L. CARBONE, A. JAKAB, Y. KHALAVKA AND C. SÖNNICHSEN: "Light-Controlled One-Sided Growth of Large Plasmonic Gold Domains on Quantum Rods Observed on the Single Particle Level", *Nano Letters* **9** (2009) 3710.
- [16] Y. XIE, L. CARBONE, C. NOBILE, V. GRILLO, S. D'AGOSTINO, F. DELLA SALA, C. GIANNINI, D. ALTAMURA, C. OELSNER, C. KRYSCHI AND P.D. COZZOLI: "Metallic-like Stoichiometric Copper Sulfide Nanocrystals: Phase- and Shape-Selective Synthesis, Near-Infrared Surface Plasmon Resonance Properties, and Their Modeling", *ACS Nano* **7** (2013) 7352.
- [17] B.N. PAL, Y. GHOSH, S. BROVELLI, R. LAOCHAROENSUK, V.I. KLIMOV, J.A. HOLLINGSWORTH AND H. HTOON: "Giant' CdSe/CdS Core/Shell Nanocrystal Quantum Dots As Efficient Electroluminescent Materials: Strong Influence of Shell Thickness on Light-Emitting Diode Performance", *Nano Letters* **12** (2012) 331.
- [18] J. ZHANG, Y. TANG AND K. LEE, M. OUYANG: "Nonepitaxial Growth of Hybrid Core-Shell Nanostructures with Large Lattice Mismatches", *Science* **327** (2010) 1634.
- [19] T. PELLEGRINO, A. FIORE, E. CARLINO, C. GIANNINI, P.D. COZZOLI, G. CICCARELLA, M. RESPAUD, L. PALMIROTTA, R. CINGOLANI AND L. MANNA: "Heterodimers Based on CoPt₃-Au Nanocrystals with Tunable Domain Size", *J. Am. Chem. Soc.* **128** (2006) 6690.
- [20] L. SPANHEL AND M.A. ANDERSON: "Semiconductor clusters in the sol-gel process: quantized aggregation, gelation, and crystal growth in concentrated zinc oxide colloids", *J. Am. Chem. Soc.* **113** (1991) 2826.
- [21] J. BANG, H. YANG AND P.H. HOLLOWAY: "Enhanced and stable green emission of ZnO nanoparticles by surface segregation of Mg", *Nanotechnology* **17** (2006) 973.
- [22] H. YANG, W. YINA, H. ZHAO, R. YANG AND Y. SONG: "A complexant-assisted hydrothermal procedure for growing well-dispersed InP nanocrystals", *J. Phys. Chem. Solids* **69** (2008) 1017.
- [23] L. CARBONE AND P.D. COZZOLI: "Colloidal heterostructured nanocrystals: Synthesis and growth mechanisms", *Nano Today* **5** (2010) 449.
- [24] D. LEONARD, M. KRISHNAMURTHY, C.M. REAVES, S.P. DENBAARS AND P.M. PETROFF: "Direct formation of quantum-sized dots from uniform coherent islands of InGaAs on GaAs surfaces", *Appl. Phys. Lett.* **63** (1993) 3203.
- [25] S.H. XIN, P.D. WANG, A. YIN, C. KIM, M. DOBROWOLSKA, J.L. MERZ AND J.K. FURDYNA: "Formation of self-assembling CdSe quantum dots on ZnSe by molecular beam epitaxy", *Appl. Phys. Lett.* **69** (1996) 3884.
- [26] T.-H. KIM, K.-S. CHO, E.K. LEE, S.J. LEE, J. CHAE, J.W. KIM, D.H. KIM, J.-Y. KWON, G. AMARATUNGA, S.Y. LEE, B.L. CHOI, Y. KUK, J.M. KIM AND K. KIM: "Full-colour quantum dot displays fabricated by transfer printing", *Nat. Photon.* **5** (2011) 176.
- [27] J.K. JAISWAL, E.R. GOLDMAN, H. MATTOUSSI AND S.M. SIMON: "Use of quantum dots for live cell imaging", *Nat. Meth.* **1** (2004) 73.
- [28] W.J. PARAK, T. PELLEGRINO AND C. PLANK: "Labelling of cells with quantum dots", *Nanotechnology* **16** (2005) R9.
- [29] R.M. CLEGG: "Fluorescence resonance energy transfer", *Curr. Opin. Biotech.* **6** (1995) 103.
- [30] A.R. CLAPP, I.L. MEDINTZ, J.M. MAURO, B.R. FISHER, M.G. BAWENDI AND H. MATTOUSSI: "Fluorescence Resonance Energy Transfer Between Quantum Dot Donors and Dye-Labeled Protein Acceptors", *J. Am. Chem. Soc.* **126** (2004) 301.
- [31] C.-Y. ZHANG, H.-C. YEH, M.T. KUROKI AND T.-H. WANG: "Single-quantum-dot-based DNA nanosensor", *Nat. Mater.* **4** (2005) 826.



Vittorianna Tasco: ricercatrice presso l'Istituto CNR-NANOTEC e si occupa di studiare i processi di crescita epitassiale nei materiali semiconduttori nanostrutturati in vista del loro utilizzo per lo sviluppo di dispositivi innovativi con applicazioni nel campo della fotonica e dell'elettronica.

Luigi Carbone: ricercatore presso l'Istituto CNR-NANOTEC ed i suoi argomenti di ricerca vertono sulla sintesi in fase liquida e rispettiva caratterizzazione delle proprietà morfologico-strutturali ed opto-elettroniche, di nanocristalli di calcogenuri metallici e di nuovi materiali non-metallici capaci di esibire risonanze plasmoniche da plasmoni localizzati di superficie.

Organic Light-Emitting Diodes: la nuova era della luce

Fabrizio Mariano

CNR-Nanotec, Lecce

Introduzione

Negli ultimi 30 anni la fotonica ha conosciuto una vera e propria rivoluzione tecnologica a seguito dell'intensa attività di ricerca che ha visto protagonista una nuova classe di dispositivi emettitori di luce che sfruttano le caratteristiche di elettroluminescenza di una vasta gamma di composti organici e che, soprattutto, offrono prospettive di sviluppo immediato in innumerevoli applicazioni, a tal punto da renderli estremamente competitivi rispetto ai dispositivi realizzati mediante semiconduttori inorganici.

I diodi organici emettitori di luce (OLED) sono diventati oggi una realtà in molte applicazioni pratiche. Basti pensare al fatto che diverse grandi compagnie industriali impegnate nella produzione di componenti per l'optoelettronica, come quelle che realizzano display per applicazioni video, telefonia mobile o altri prodotti di largo consumo, negli ultimi anni hanno investito grosse risorse nel campo dei dispositivi organici elettroluminescenti.

Il vasto campo applicativo di questi dispositivi può essere sommariamente suddiviso in tre principali categorie:

1) Illuminazione (figura 1, in basso), in cui l'intera superficie del diodo emette luce alla stessa

intensità. Utilizzando dei composti che emettono luce con colori complementari è possibile avere emissione bianca. Gli OLED bianchi sono quindi adatti ad essere sfruttati nel settore dell'illuminotecnica e, rispetto alle tecnologie standard, garantiscono ottima resa visiva, producono luce diffusa da una larga area emissiva, hanno un colore adeguato a fornire le migliori condizioni di illuminazione, sono a basso impatto ambientale e possono essere fabbricati su supporti flessibili.

2) Indicatori alfanumerici, caratterizzati da qualche centinaio di pixel.

3) Display (figura 1, in alto), con un numero di pixel tale da garantire la risoluzione richiesta nelle diverse applicazioni video (dai 180000 pixel del piccolo schermo da 2,5 pollici di una comune videocamera digitale, ai 1920x1080 pixel di un display ad alta definizione). Tali schermi non hanno bisogno di retroilluminazione; rispetto alla tecnologia dei cristalli liquidi pertanto, un display OLED può essere più sottile e leggero, è più efficiente ed è caratterizzato da un miglior contrasto.

Tra le ragioni che stanno alla base della diffusione degli OLED nel mercato dell'optoelettronica vanno sicuramente sottolineati i numerosi vantaggi che essi offrono rispetto a dispositivi analoghi realizzati con tecnologie tradizionali a



Figura 1: In alto, display OLED. In basso, prototipo di OLED bianco flessibile per illuminazione.

base di semiconduttori inorganici. Tra quelli più evidenti meritano di essere citati:

- distribuzione quasi lambertiana dell'emissione che rende possibili ampi angoli visivi;
- disponibilità di tutti i colori dello spettro del visibile;
- tensioni di funzionamento relativamente basse a temperatura ambiente;
- brevi tempi di risposta;
- costi di produzione ridotti.

Il processo più rilevante, tra quelli che concorrono nel definire le prestazioni di un OLED, è senza dubbio l'elettroluminescenza. Con questo termine si intende l'emissione di luce dal materiale organico in conseguenza della ricombinazione dei portatori di carica iniettati dagli elettrodi.

L'elettroluminescenza in materiali organici semiconduttori venne realizzata per la prima volta negli anni '60. Helfrich [1] risolse il problema dell'iniezione degli elettroni nei materiali organici (fino ad allora era stato possibile solo l'iniezione di lacune da una soluzione di iodio-potassio) utilizzando una soluzione di ioni negativi di antracene. Questa fu la prima prova sperimentale che l'elettroluminescenza ha bisogno di un'iniezio-

ne di cariche bipolari all'interno del dispositivo, della cattura di cariche di segno opposto in una zona di ricombinazione e del decadimento radiativo delle coppie eccitate elettrone-lacuna così create. Tuttavia, l'efficienza di ricombinazione era piuttosto bassa.

Bisognò aspettare la fine degli anni '80 prima che una ricerca più intensa fosse stimolata dal lavoro di Tang del gruppo Kodak [2]. Il loro dispositivo elettroluminescente (avente uno spessore complessivo di 135 nm) è ancora il prototipo degli OLED di oggi. Esso consisteva di uno strato di ossido di indio-stagno (ITO, un semiconduttore inorganico trasparente altamente drogato che consente l'iniezione delle lacune), una diamina aromatica (TPD) come strato di trasporto delle lacune, uno strato di emissione di Alq3 (aluminium-tris-(8-hydroxy-quinolate)) ed una lega di magnesio-argento (Mg:Ag) come contatto per l'iniezione di elettroni. Tutti gli strati furono sublimati in un sistema a vuoto ed erano amorfi.

Il LED del 1987 necessitava di 5.5 V per raggiungere una luminanza di 100 cd/m² (la luminosità tipica dei display) e aveva una efficienza di corrente di circa 2 cd/A. L'emissione verde aveva origine dallo strato di Alq3. Questa svolta nell'efficienza e nella tensione operativa del dispositivo diede il via ad un'intensa ricerca sui semiconduttori organici e sul loro uso per la costruzione di OLED.

Nel 1990, Burroughes [3] realizzò un OLED ad alta efficienza basato su polimeri (PLED). Il materiale emettitore del loro prototipo era il PPV (poly(p-phenylene-vinylene), che, come l'Alq3, emette nella regione della lunghezza d'onda del verde.

Per i PLED, il numero di strati è, di solito, limitato a due, dal momento che il solvente di ogni nuovo livello dissolverebbe, lo strato polimerico sottostante. Al contrario, i LED organici costruiti utilizzando small molecules possono essere costituiti da tanti livelli quanti ne sono necessari per ottenere le prestazioni migliori. Quindi, l'ottimizzazione dei LED può essere ottenuta per mezzo dell'aggiunta di nuovi strati molecolari con le caratteristiche più appropriate.

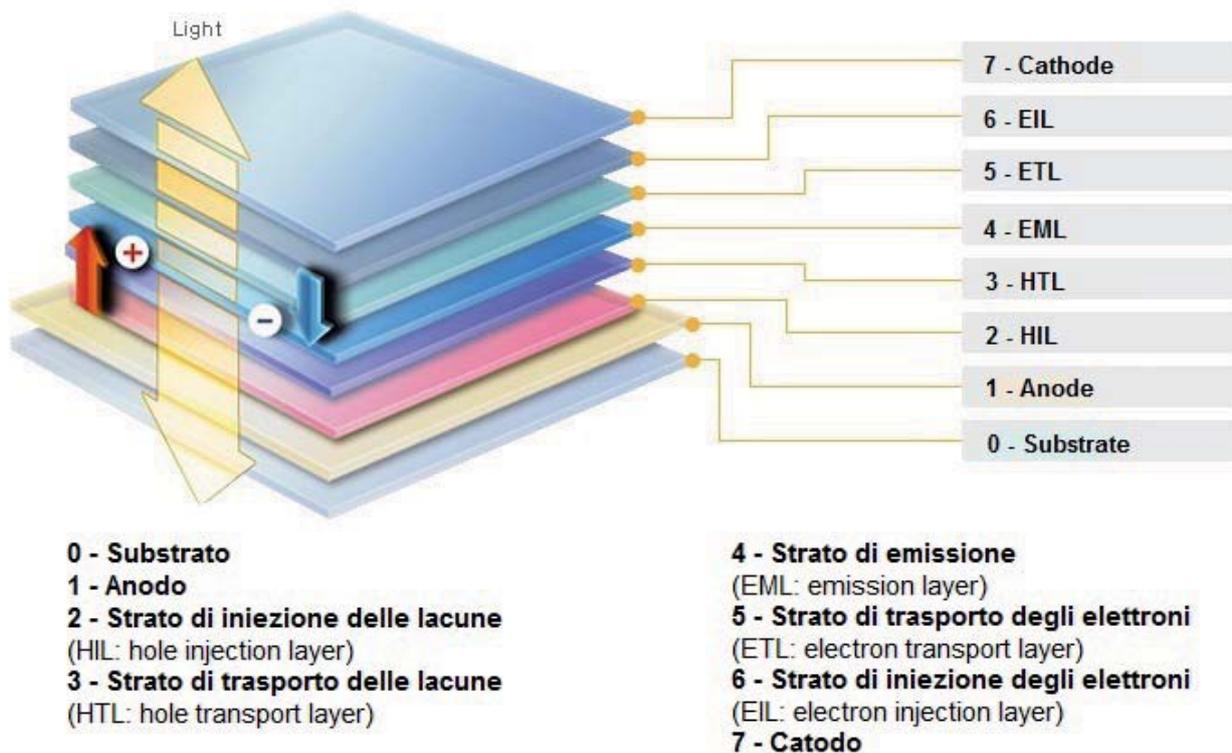


Figura 2: Schema di principio di un OLED e funzioni svolte dagli strati che lo compongono.

Struttura di un OLED

Un OLED consiste di una sequenza di strati di semiconduttori organici, spesso poche decine di nanometri, compresa tra due elettrodi (figura 2). Con il termine organico si intende un materiale principalmente costituito da atomi di carbonio. Questi materiali si comportano da semiconduttori se il livello energetico contenente gli elettroni di valenza (Highest Occupied Molecular Orbital, HOMO) e il primo livello successivo disponibile (Lowest Unoccupied Molecular Orbital, LUMO) sono distanziati, in termini di energia, di pochi elettronvolts (eV), permettendo facilmente agli elettroni di passare allo stato eccitato. Nella figura seguente viene illustrata la schematizzazione generale di un OLED e la funzione svolta da ciascuno strato che lo compone.

Quando viene applicata una tensione positiva, gli elettroni sono iniettati dal catodo mentre le lacune (ovvero le vacanze di elettroni) dall'anodo. I portatori di carica fluiscono attraverso gli strati di trasporto fino ad arrivare allo strato di emissione dove formano delle quasi-particelle chiamate eccitoni, in cui la lacuna è nel livello di valenza mentre l'elettrone è in quello eccitato. Gli eccitoni hanno una certa probabilità di

decadere radiativamente; in tal caso, le transizioni dell'elettrone dallo stato eccitato a quello fondamentale producono fotoni, con lunghezza d'onda proporzionale al salto energetico, che attraversano l'anodo trasparente.

Al fine di raggiungere un'alta efficienza (da intendere come numero di fotoni emessi per elettroni iniettati), gli strati devono soddisfare certi requisiti (figura 3). L'HIL deve facilitare il passaggio delle cariche positive dall'anodo allo strato di trasporto delle lacune HTL. L'HTL dovrebbe avere un'alta mobilità delle lacune e dovrebbe impedire agli elettroni provenienti dal catodo di raggiungere l'anodo. Le posizioni energetiche dell'HOMO e del LUMO dello strato di emissione EML deve rendere possibile l'iniezione delle lacune e degli elettroni dagli strati vicini. In più, l'efficienza di fotoluminescenza (ovvero il rapporto tra gli eccitoni che decadono radiativamente e quelli creati) del materiale di questo strato deve essere alta e lo spettro di assorbimento deve essere spostato rispetto a quello di emissione (deve avere, cioè, un ampio Stokes-shift; in questo modo si evita il riassorbimento della luce emessa). Le caratteristiche dello strato di trasporto degli elettroni ETL e dello strato di iniezione EIL

degli stessi devono essere complementari a quelle dell'HIL e dell'HTL (ovvero adeguati livelli di LUMO, alta mobilità elettronica). Il catodo, poi, dovrebbe essere un metallo con una bassa funzione lavoro (definita come l'energia minima necessaria per portare un elettrone delle sue molecole a distanza infinita), come il magnesio o il calcio. Infine, almeno un elettrodo deve essere trasparente alla luce visibile. Per questo motivo nella maggior parte dei casi si utilizza l'ossido di indio-stagno (ITO: Indium-Tin Oxide) come anodo trasparente.

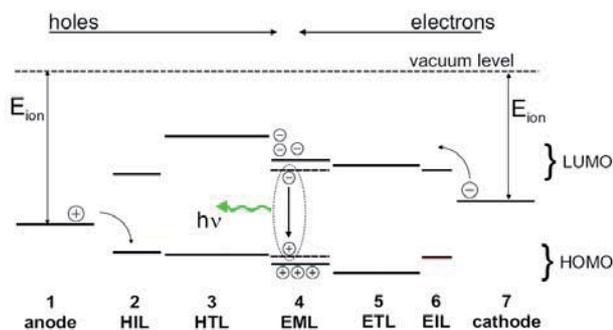


Figura 3: Visualizzazione schematica di una situazione energeticamente favorevole in un OLED.

Quasi tutti gli OLED che vengono realizzati, tuttavia, sono costituiti da meno strati di quelli proposti in figura 2. Ciò è dovuto al fatto che grazie alle caratteristiche di alcuni materiali organici è possibile combinare più proprietà, fra quelle menzionate prima, in un unico strato.

Il comportamento optoelettronico di un OLED è determinato da una complessa concomitanza di diversi effetti:

1) l'iniezione di trasportatori di carica dall'anodo e dal catodo negli strati organici vicini;

2) la redistribuzione del campo elettrico negli strati organici (che può causare l'iniezione di carica agli elettrodi) quando si applica la tensione di alimentazione;

3) l'insorgenza di spazi di carica sulle interfacce interne;

4) la dipendenza dal campo elettrico e dalla temperatura della mobilità dei trasportatori di carica;

5) l'attraversamento, da parte delle cariche, delle barriere energetiche interne, favorito dal campo elettrico presente;

6) la creazione di eccitoni sulle interfacce interne;

7) la creazione di eccitoni nel bulk;

8) il decadimento radiativo o non-radiativo degli eccitoni.

Di seguito vengono trattati i principali fenomeni che intervengono nella determinazione dell'efficienza di elettroluminescenza di un OLED, il parametro che riassume le prestazioni di un dispositivo.

Efficienza di elettroluminescenza

La determinazione dell'efficienza quantica in un materiale organico e in un OLED è una delle questioni chiave per la ricerca e lo sviluppo di dispositivi ad alte prestazioni. Ad oggi, l'efficienza quantica ottenuta nella maggior parte dei diodi elettroluminescenti non supera pochi punti percentuali [4].

L'efficienza degli OLED può essere indicata tramite l'efficienza quantica, l'efficienza di potenza e l'efficienza luminosa.

L'efficienza di potenza, in lumen/Watt (lm/W), indica il flusso luminoso emesso per unità di potenza elettrica fornita. L'efficienza luminosa, in candele/Ampere (cd/A, dove $cd=lm/steradiante$), misura l'intensità luminosa emessa per unità di corrente fornita.

L'efficienza quantica η_q di un dispositivo si compone di due parti, interna ed esterna:

Efficienza quantica interna: η_{int} o IQE , è il numero di fotoni generati all'interno del dispositivo per numero di coppie elettrone-lacuna iniettate. Una buona parte dei fotoni generati rimane intrappolata e assorbita all'interno del dispositivo (vedere il paragrafo "Dissipazioni ottiche negli OLED").

Efficienza quantica esterna: η_{ext} o EQE , è il numero di fotoni emessi dal dispositivo per numero di coppie elettrone-lacuna iniettate.

Efficienza quantica interna

Per OLED che utilizzano emettitori fluorescenti,

$$\eta_{int} = IQE = \gamma \eta_s \phi_f$$

dove γ indica la frazione di cariche iniettate che producono eccitoni, η_s la frazione di eccitoni che

danno origine a singoletti e φ_f è l'efficienza quantica di fluorescenza, ovvero la frazione di energia rilasciata sotto forma di luce dal materiale fluorescente. Sulla base della statistica degli spin, η_s può avere un valore massimo del 25%. In altre parole, solo il 25% degli eccitoni sono dei singoletti in grado di cedere energia mediante emissione di fotoni (vedere paragrafo "Ricombinazione dei portatori di carica e meccanismi di emissione: il *triplet harvesting*").

L'efficienza quantica di fluorescenza φ_f può avvicinarsi al valore unitario, ma solo in soluzioni diluite.

Il problema generale è quello di mantenere alta la φ_f in stato solido. Pochi materiali utilizzati negli OLED raggiungono una φ_f superiore al 50%. Anche in questo caso, una opportuna ingegnerizzazione delle molecole organiche può generare materiali per dispositivi OLED fluorescenti con IQE vicino al 25%. Ad esempio, uno studio mostra che un materiale per il trasporto di elettroni, che è anche un buon emettitore blu-verde, presenta una resa quantica assoluta del 97% [?].

Bassi valori di φ_f sono anche causati da effetti fotonici. Ad esempio, la vicinanza ad un elettrodo metallico riflettente può causare sia il trasferimento di energia dall'eccitone ai plasmoni di superficie (vedere il paragrafo "Dissipazioni ottiche negli OLED"), sia l'annichilazione dei fotoni, riducendo il tasso di emissione radiativa [6]. La distanza ottimale tra lo strato emissivo e il catodo è dell'ordine di 50 nm.

Fattori come questo devono essere considerati nella progettazione dei dispositivi OLED.

Per OLED con emettitori fosforescenti, si considera η_{ph} , ovvero la frazione di eccitoni che danno origine a tripletti, che non dovrebbe essere superiore al 75%. Tuttavia, come illustrato nel paragrafo "Ricombinazione dei portatori di carica e meccanismi di emissione: il *triplet harvesting*", utilizzando dei materiali emettitori fosforescenti, entrambi gli stati di singoletto e tripletto possono essere raccolti per produrre fotoni. Pertanto, almeno in quel caso, η_{ph} può essere prossimo all'unità.

Tra i fenomeni fondamentali che possono anche portare ad una ridotta efficienza quantica interna ci sono:

1) il *quenching* tra eccitoni [7]:

2) la dissociazione dell'eccitone [8]:

Alcuni di questi meccanismi di annichilazione degli eccitoni sono comuni a entrambi i sistemi fluorescenti e fosforescenti.

Efficienza quantica esterna

Nonostante tutti i progressi nel massimizzare l'efficienza quantica interna degli OLED, solo il 20% circa di tutti i fotoni generati dal materiale attivo riesce a fuoriuscire dal dispositivo. In un semplice dispositivo planare, oltre l'80% della luce può essere perso a causa dell'assorbimento interno e degli effetti di guida d'onda (figura 4). La maggior parte dei fotoni vengono riflessi dalle interfacce e vengono alla fine riassorbiti.

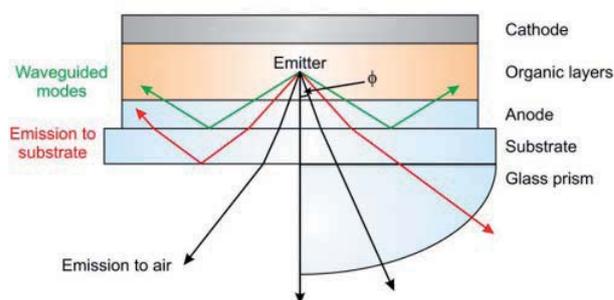


Figura 4: Visualizzazione delle riflessioni interne ad un OLED causate dalla presenza di strati ad elevato indice di rifrazione (ITO usato come anodo, vetro...). La figura illustra come, utilizzando dei sistemi esterni al dispositivo, si possa raccogliere una parte della luce che altrimenti sarebbe riflessa all'interno.

La riflessione interna dei fotoni, causata dalla presenza di strati ad alto indice di rifrazione, è la causa principale della scarsa efficienza energetica degli OLED.

L'efficienza quantica esterna η_{ext} (o EQE) è legata all'efficienza quantica interna η_{int} (o IQE) dalla seguente relazione:

$$\eta_{ext} = R_e \eta_{int}$$

dove R_e (o η_{out}) è l'efficienza di estrazione (o di *outcoupling*) che rappresenta il numero di fotoni emessi dall'OLED per numero di fotoni generati all'interno del dispositivo. La scarsa estrazione della luce è oggi il più importante fattore che limita l'efficienza quantica esterna degli OLED.

L'efficienza di *outcoupling* R_e è legata all'indice di rifrazione dello strato emissivo (n) dalla

seguente relazione:

$$R_e = \frac{1}{2n^2}$$

La maggior parte dei materiali usati negli OLED hanno un indice di rifrazione pari a circa 1,7, che è all'origine del suddetto limite percentuale di estrazione della luce. Sono stati proposti diversi metodi per migliorare l'efficienza di estrazione degli OLED:

- *Texturing* superficiale - mediante riflessioni multiple, dà ai fotoni la possibilità di trovare il cono di fuga. Con questa tecnica è stato finora raggiunto un fattore 2 di miglioramento della R_e [9].
- Uso di substrati a basso indice di rifrazione - E' stato riportato un miglioramento di R_e di un fattore 2,5 [10].

Altre tecniche sono state proposte e sono in fase di sperimentazione. Queste includono l'aggiunta, sulla superficie del dispositivo, di una matrice di microlenti (in linea di principio, si tratta di una *texturing* superficiale controllata) con la quale è stato dimostrato un miglioramento di un fattore 1,5 [11].

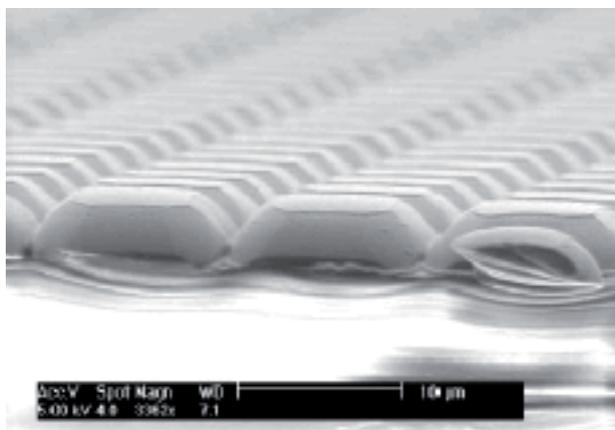


Figura 5: Matrice di microlenti.

La figura 5 illustra la forma della matrice di microlenti depositata sulla superficie di un substrato di vetro che ha portato a circa il 50% di miglioramento dell'efficienza di *outcoupling*.

Un'altra tecnica utilizza uno strato ordinato di microsferi di silice [12].

Oggi si ritiene che un perfezionamento sistematico dei metodi attualmente conosciuti e l'esplorazione di nuove strategie porterà ad un miglioramento della η_{ext} , che può raggiungere il

60-80% nei prossimi anni. Ciò rappresenta una opportunità importante per aumentare l'efficienza di emissione tramite l'ingegnerizzazione del dispositivo.

Dissipazioni ottiche negli OLED

Vediamo quali possono essere i fenomeni di accoppiamento ottico, radiativi e non, del dipolo elettrone-lacuna. Esso può accoppiarsi o radiativamente o non radiativamente a seconda delle condizioni che trova intorno a sé. Nel caso di accoppiamento radiativo si verifica quanto descritto sopra con l'emissione da parte del dipolo di un fotone che possiede un'energia pari alla differenza tra gli stati energetici in cui si trovava il dipolo prima e dopo l'emissione, per esempio nel caso in cui la luce riesca fuoriuscire dal dispositivo oppure se i fotoni vengono intrappolati nelle guide d'onda (figura 6). Per quanto riguarda invece il decadimento non radiativo possiamo trovarci di fronte al fenomeno di accoppiamento con i modi di risonanza plasmonica. Descriviamo brevemente i fenomeni che possono avvenire se un materiale emettitore viene posto in presenza di interfacce:

- Accoppiamento con fenomeni di guida d'onda: la guida d'onda è un fenomeno per cui l'onda elettromagnetica che sta attraversando un volume di materiale con indice di rifrazione n_1 impatta su una superficie di separazione con un altro mezzo di indice di rifrazione n_2 e viene riflessa per *total internal reflection*, rimanendo intrappolata se ha una determinata direzione di propagazione. L'effetto è un fenomeno di ottica che può essere spiegato, per quanto concerne il nostro studio, tramite il riferimento alle fibre ottiche. Quando un'onda elettromagnetica passa da un mezzo a un altro subisce una deviazione in funzione degli indici di rifrazione (n) dei due mezzi attraversati dall'onda, secondo la legge di Snell:

$$n_1 \sin \theta_1 = n_2 \sin \theta_2$$

Qualora l'angolo di incidenza θ dell'onda elettromagnetica superi un certo valore critico:

$$\theta_{crit} = \arcsin \frac{n_2}{n_1}$$

Ci troviamo di fronte a una riflessione interna. Si deve inoltre tener presente che i valori di n_1 e di n_2 variano in funzione della lunghezza d'onda impattante.

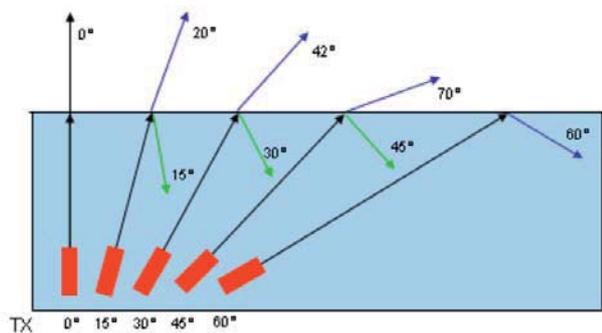


Figura 6: Schema sul meccanismo alla base del fenomeno della guida d'onda.

- Accoppiamento con modi “leaky” (letteralmente “che perdono”): questi modi, derivati dallo stesso principio fisico alla base del fenomeno di guida d'onda, si accoppiano con la radiazione emessa non permettendone la fuoriuscita diretta dal dispositivo. A differenza della tipologia di accoppiamento citata precedentemente, questa non intrappola totalmente la radiazione emessa ma permette ad una parte di essa di essere convogliata all'esterno. Ciò significa che tali modi non sono del tutto dannosi per l'efficienza del dispositivo e danno anche una misura indiretta della parte di radiazione che riesce a fuoriuscire da esso.
- Accoppiamento con modi plasmonici di superficie: per quanto riguarda tale meccanismo, occorre tener presente che qualsiasi materiale che possiede un indice di rifrazione n avente una parte reale negativa e una piccola parte immaginaria positiva (come ad esempio i metalli in un dispositivo a microcavità) permette la formazione di modi di risonanza plasmonici (*Surface Plasmon Resonance*, SPR). I plasmoni non sono altro che delle oscillazioni quantizzate degli elettroni presenti nel livello energetico di conduzione della superficie del metallo causate da una radiazione elettromagnetica incidente. Il modo di risonanza plasmonica si verifica quando l'angolo di incidenza della radiazione elettromagnetica sulla superficie e la

sua lunghezza d'onda risultano “accoppiati” in modo tale da determinare l'oscillazione degli elettroni appartenenti alla superficie metallica. Le oscillazioni plasmoniche generano radiazioni che si propagano debolmente al di fuori della superficie metallica, ma generalmente non sono nel campo del visibile, quindi vengono detti accoppiamenti non radiativi.

Lo studio dei plasmoni abbraccia notevoli rami di ricerca e di applicazione, dall'ingegneria dei dispositivi alla biologia passando per la fabbricazione di dispositivi tramite litografia. Questo settore ha subito recentemente grandi sviluppi grazie della possibilità di manipolare la struttura dei metalli fino alla scala nanometrica.

Possiamo classificare i modi plasmonici come propaganti o localizzati sulla superficie (figura 7). Nel caso del modo plasmonico propagante, la propagazione avviene lungo gli assi x e y sulla superficie di interfaccia del mezzo metallico, per una distanza che varia tra i pochi e le centinaia di micron. Il plasmon poi decade lungo l'asse z . Le tecniche di analisi dell'interazione delle onde elettromagnetiche confinate sulla superficie metallica e gli strati molecolari del dispositivo possono essere di tre tipi:

1. ad angolo predeterminato, dove misuriamo la riflettività in funzione della lunghezza d'onda incidente sul campione;
2. a lunghezza d'onda predeterminata, dove misuriamo la riflettività in funzione dell'angolo di incidenza sul campione;
3. mappatura ad angolo e lunghezza d'onda variabile, dove l'analisi di riflettività viene effettuata attraverso l'utilizzo di un doppio fascio di luce (uno ad angolo, l'altro a lunghezza d'onda costante).

Per quanto riguarda invece i plasmoni localizzati, la radiazione interagisce con particelle più piccole della lunghezza d'onda incidente; i plasmoni oscillano localmente attorno alla particella con una determinata frequenza di risonanza, dipendente dalla geometria della particella oltre che dal materiale e dal mezzo circostante. L'interazione con l'ambiente circostante genera variazioni di lunghezza d'onda di risonanza. Grazie

all'analisi delle dissipazioni ottiche, è possibile effettuare uno studio sulle variazioni di efficienza degli OLED al variare dell'ambiente ottico in cui le molecole emissive vengono poste. Infatti, considerando che il rapporto di smorzamento fra la *rate* di decadimento spontaneo dei dipoli nello spazio libero (b_0) rispetto alla *rate* di decadimento modificata dalla densità dei modi fotonici (b) è funzione dell'integrale sul vettore d'onda nel piano u della potenza dissipata, vale la seguente relazione:

$$\frac{b}{b_0} = (1 - q) + q \int_{u=0}^{u=\infty} I_{iso}(u) du$$

dove I_{iso} è la funzione della potenza dissipata derivata dalla scomposizione del campo riflesso, in funzione di u , per dipoli orientati isotropicamente; q è l'efficienza quantica intrinseca dell'emettitore e la differenza $(1 - q)$ rappresenta la potenza persa dall'emettitore per processi intrinseci non radiativi. È possibile scomporre I_{iso} nei vari contributi di dissipazione della potenza emessa, evidenziando nell'equazione precedente i termini radiativi e non radiativi:

$$\frac{b}{b_0} = (1 - q) + qC_G + qFC_R + q(1 - F)C_R$$

dove F è la frazione radiativa definita come il rapporto fra luce effettivamente fuoriuscita dal dispositivo e potenza totale accoppiata con i modi radiativi, qC_G è la porzione di luce persa nei modi guidati, qC_R è quella accoppiata ai modi radiativi di cui qFC_R viene effettivamente estratta dal dispositivo, mentre $q(1 - F)C_R$ viene assorbita dagli strati della struttura. Possono dunque essere definiti due parametri che mettono in relazione diretta l'analisi delle dissipazioni con la potenza emessa e con l'efficienza:

$$B = qFC_R$$

$$\eta = \frac{qFC_R}{(1 - q) + qC_G + qFC_R + q(1 - F)C_R}$$

dove la radianza B è la potenza integrata totale che viene estratta dal dispositivo mentre η è l'efficienza radiativa, ossia la frazione dell'emissione spontanea che viene effettivamente emessa dalla struttura ottica.

Abbiamo visto, quindi, che i parametri e i fenomeni che intervengono nella determinazione

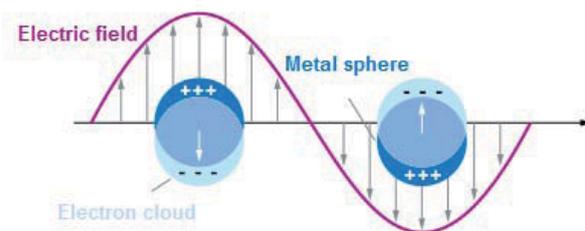
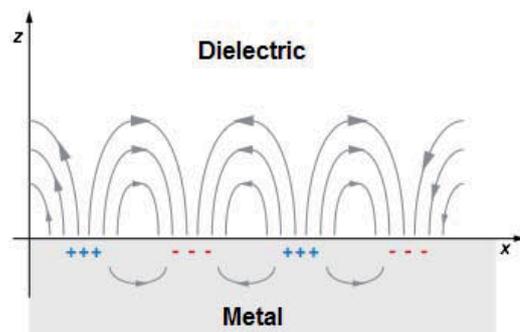


Figura 7: In alto, plasmoni propaganti; in basso, plasmoni localizzati.

delle prestazioni di un dispositivo, in termini di efficienza quantica interna η_{int} , sono di diversa natura: γ , legato alle cariche elettriche iniettate all'interno del dispositivo, $\eta_{s/ph}$, correlato alla statistica di formazione degli eccitoni di singoletto (s) e di tripletto (ph), φ_f , vincolato ai processi di emissione. La comprensione e il controllo di questi fenomeni sono di fondamentale importanza per il raggiungimento di una maggiore efficienza.

Di seguito illustreremo alcune tecnologie mediante le quali è possibile modificare i singoli parametri di cui abbiamo discusso; introdurremo, pertanto, il *doping* elettrico che ha una diretta influenza su γ , il *triplet harvesting* che consente di intervenire sulla $\eta_{s/ph}$, scegliendo opportunamente le molecole dello strato emissivo, e la tecnologia delle microcavità mediante la quale si può agire su φ_f .

Il drogaggio elettrico degli strati di trasporto organici

In un OLED, il processo di trasporto delle cariche elettriche dagli elettrodi allo strato attivo è cruciale per il buon funzionamento del dispositivo. Per rendere questo trasporto il più efficiente possibile (cioè con basse perdite resistive), è utile

avere degli strati di trasporto altamente conduttivi. Inoltre, le cariche devono essere efficacemente iniettate dai contatti inorganici per cui una iniezione efficiente richiede basse barriere energetiche, una regione di carica spaziale sottile o entrambe. È importante sottolineare come negli anni passati, la svolta della tecnologia del silicio è avvenuta nel momento in cui il tipo di conduzione non era più determinato da impurezze ma ha cominciato ad essere controllato tramite un drogaggio (o *doping*) volontario. I semiconduttori organici sono generalmente depositati come film nominalmente non drogati; tuttavia, il *doping* controllato e stabile di tali film si è imposto negli ultimi anni, come una tecnica efficace per l'incremento delle efficienze dei dispositivi a base organica poiché si riesce a ridurre le perdite ohmiche e a facilitare l'iniezione dai contatti.

Una tecnica adoperata con successo è quella del *doping* per co-evaporazione di una matrice con un drogante molecolare [13]. Il vantaggio di tale approccio rispetto all'uso di altre sostanze, come i metalli e i gas, è l'elevata stabilità che può essere raggiunta con i droganti molecolari. Sia per il drogaggio di tipo *p* [14] che per quello di tipo *n* [15], la conducibilità può essere incrementata di molti ordini di grandezza, ben al di sopra della conducibilità intrinseca dei materiali.

L'introduzione di strati di trasporto drogati porta diversi miglioramenti, in primo luogo la riduzione della resistenza ohmica degli strati stessi. In figura 8, a sinistra, è rappresentata schematicamente la tipica struttura a bande di un LED inorganico. In un tale dispositivo, lo strato emettitore (posto al centro, con *band-gap* inferiore) è inserito tra due strati di trasporto altamente drogati. Grazie alla elevata conducibilità di questi strati, le cadute di potenziale sono molto piccole dal momento che le bande sono praticamente piatte. La tensione di funzionamento del dispositivo è quindi molto prossima al *band-gap* dell'emettitore.

Nei LED organici con livelli di trasporto non drogati, la concentrazione dei portatori è bassa e si hanno perdite resistive. In tal caso, come mostrato nello schema a destra della figura 8, è necessario applicare un campo elettrico sufficientemente grande per guidare le cariche attraverso il dispositivo, tale che la tensione di funzionamento può essere notevolmente maggiore del

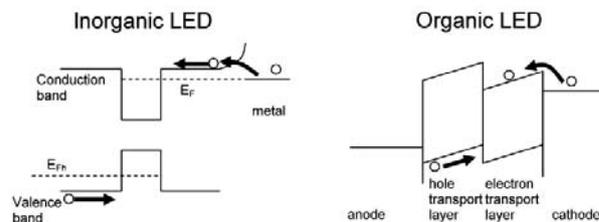


Figura 8: Schematizzazione della struttura a bande di un tipico LED inorganico (a sinistra), prossimo alla condizione di banda piatta durante il funzionamento (l'iniezione dai contatti si ha per effetto tunnel attraverso le sottili regioni di carica spaziale) e di un LED organico non drogato (a destra), in cui, a causa della bassa concentrazione di portatori, i livelli di energia non sono più piatti ma possiedono una barriera che i portatori di carica possono superare con l'applicazione di una opportuna tensione agli elettrodi.

band-gap del materiale emettitore.

È vero che una bassa caduta di potenziale può anche essere ottenuta utilizzando strati organici molto sottili, ma esistono almeno tre buone ragioni per usare strati drogati piuttosto spessi: 1) si riduce la probabilità di cortocircuiti, 2) si garantisce un maggiore controllo della deposizione e 3) si può facilmente regolare lo spessore della microcavità ottica nei dispositivi optoelettronici, al fine di ottimizzare la distribuzione dei campi al loro interno (vedere il paragrafo "Modulazione dello spettro di elettroluminescenza di OLED bianchi tramite multicavità organo-metalliche accoppiate").

Un altro aspetto molto importante è che gli strati di trasporto drogati danno origine a regioni di carica spaziale molto strette ai contatti, abbastanza sottili da consentire alle cariche di attraversarle per effetto tunnel, come schematicamente illustrato nella parte sinistra della figura 8. In questo modo, è possibile realizzare contatti ohmici anche in presenza di grandi barriere energetiche tra il livello di trasporto e il contatto. Ciò è molto vantaggioso in quanto, negli OLED con livelli di trasporto non drogati, l'ottimizzazione delle barriere richiede degli sforzi maggiori di ingegnerizzazione della intera struttura e la realizzazione di contatti con materiali aventi basse funzioni lavoro, quindi piuttosto instabili. Gli OLED con livelli di trasporto drogati sono quindi molto più flessibili in termini di scelta dei materiali utilizzabili.

I principi alla base del drogaggio dei semiconduttori organici sono simili a quelli dei materiali inorganici dal momento che vengono introdotte, all'interno del materiale trasportatore, delle cariche libere per mezzo di donatori o di accettori di elettroni. Nei materiali organici, si devono aggiungere molecole che o donino elettroni al più basso orbitale non occupato (LUMO) (drogaggio di tipo *n*) o rimuovano elettroni dal più alto orbitale non occupato (HOMO), al fine di generare lacune (drogaggio di tipo *p*) (figura 9).

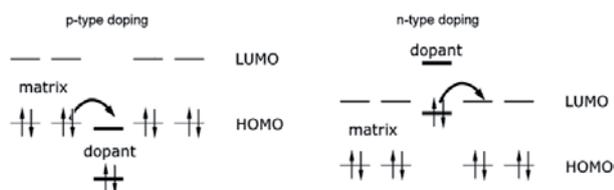


Figura 9: Schematizzazione del drogaggio molecolare di tipo *p* (a sinistra) o di tipo *n* (a destra). Nel drogaggio di tipo *p*, il drogante molecolare funge da accettore di elettroni; in quello di tipo *n*, da donatore. Al fine di avere un doping efficace, è necessario che ci sia una buona sovrapposizione dei livelli energetici della matrice e del drogante.

Ricombinazione dei portatori di carica e meccanismi di emissione: il triplet harvesting

Come detto in precedenza, con il termine elettroluminescenza si intende l'emissione di luce dal materiale organico in conseguenza della ricombinazione dei portatori di carica iniettati dagli elettrodi. I processi fisici alla base dell'elettroluminescenza sono strettamente correlati alle proprietà del materiale che costituisce lo strato emettitore, ed in particolare alla struttura elettronica. Caratteristica comune dei semiconduttori organici è la presenza di un *gap* energetico nella loro struttura a bande proprio come nel caso dei semiconduttori inorganici tradizionali. Ma assai diversi sono gli approcci teorici che spiegano i meccanismi da cui esso ha origine.

È utile discutere come può essere schematizzato il processo di ricombinazione elettrone-lacuna, la formazione di un eccitone neutro ed il popolamento di uno stato eccitato della molecola di emettitore. Si introducono ora i processi che avvengono all'interno dello strato emettitore di un

OLED. Questo strato, è costituito da una matrice organica, drogata con molecole di emettitore. Nel modello presentato si presume che la ricombinazione di elettroni e lacune avviene a livello delle molecole droganti.

La figura 10 visualizza un modello semplificato e schematico per descrivere il processo di formazione dell'eccitone [18]. La prima fase è caratterizzata dalla cattura di un portatore di carica. Tale processo induce (per un breve intervallo di tempo) la formazione di un complesso ossidato. Il processo di intrappolamento dei portatori di carica può indurre una riorganizzazione delle molecole del materiale emettitore. Tuttavia, questo effetto non è rappresentato nel modello della figura 10. Applicando un potenziale esterno, l'elettrone migra verso l'anodo attraversando il materiale usato come matrice. Di solito, questo processo di migrazione degli elettroni richiede una energia di attivazione termica dovuta alle disomogeneità legate al disordine spaziale ed energetico e agli effetti di riorganizzazione della matrice. I conseguenti *shift* energetici, risultano minori o dell'ordine dell'energia termica $k_B T$ in cui k_B è la costante di Boltzmann e T la temperatura assoluta. Per chiarezza, il diagramma in figura è semplificato e non mostra la distribuzione disomogenea dei livelli energetici delle molecole della matrice e dei loro *shift* energetici indotti dal potenziale esterno.

Quando l'elettrone è ancora lontano dalla lacuna intrappolata, migra indipendentemente da questa lacuna verso l'anodo. In tal caso, la lacuna e l'elettrone non sono legati. Tuttavia, quando l'elettrone migra in una regione con una specifica distanza R tra elettrone-lacuna, la lacuna attrarrà l'elettrone. A causa dell'attrazione di Coulomb, si genera un legame elettrone (*e*) - lacuna (*h*). L'energia di legame $\Delta E(e-h)$ è proporzionale a $1/\epsilon R$ e dipende dalla distanza R e dalla costante dielettrica ϵ del materiale che costituisce la matrice. Indotto da questa attrazione, si forma quindi l'eccitone.

Per poter comprendere i fenomeni successivi, è molto importante prendere in considerazione anche gli *spin* di elettrone e lacuna. Lo *spin* della lacuna è dato dallo *spin* dell'elettrone non appaiato nella molecola di emettitore. In accordo alla trattazione quanto-meccanica, i due *spin* possono essere accoppiati a quattro nuovi stati combinati:

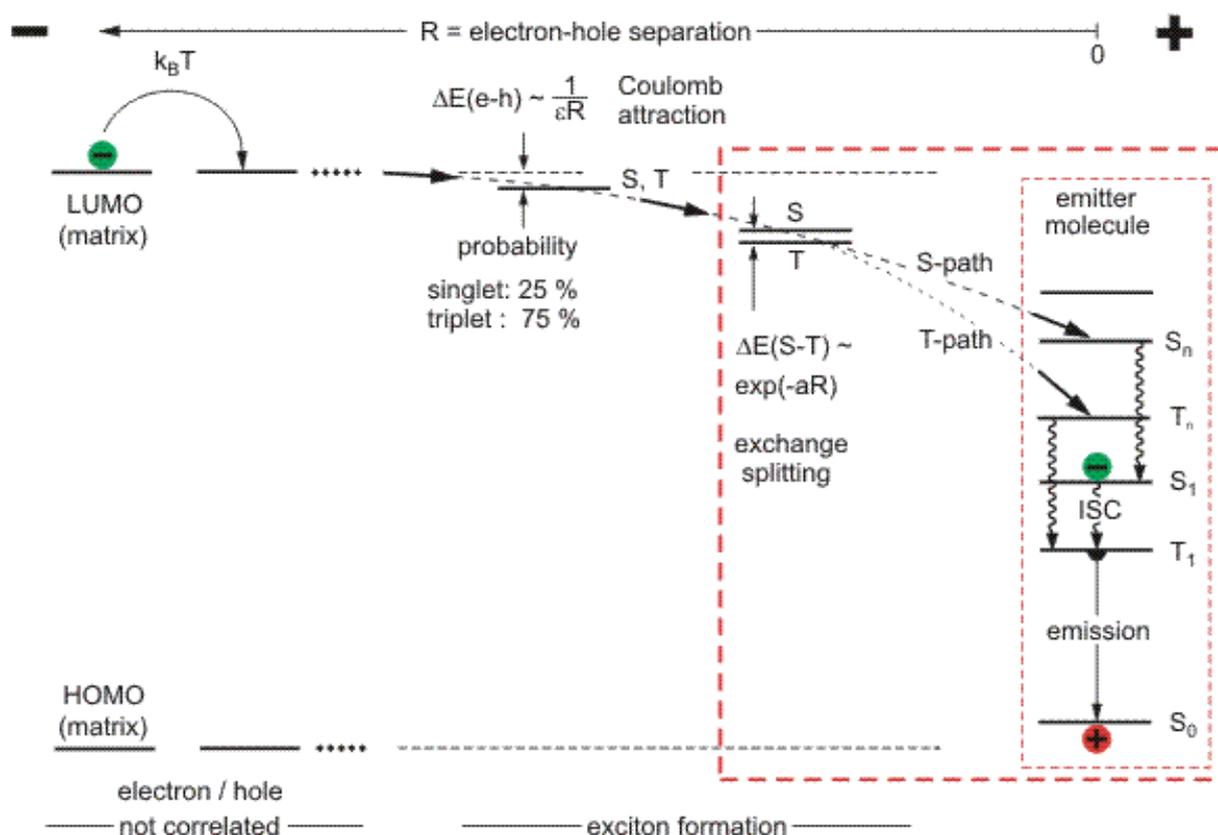


Figura 10: Dinamiche di formazione degli eccitoni. ISC: intersystem crossing.

uno stato di singoletto e uno stato di tripletto. Lo stato di tripletto si compone di tre sottostati. Questi sottostati differiscono tra loro principalmente per le orientazioni di *spin* relative (figura 11).

Lo *split* energetico tra gli stati di singoletto e di tripletto può essere ignorato al di sopra di grandi distanze elettrone-lacuna. Pertanto, lo stato eccitonico corrispondente è rappresentato in figura 10 (al centro) solo da un livello di energia. Dal punto di vista statistico, si formeranno i quattro sottostati di questo stato eccitonico (occupato) con pari probabilità. Di conseguenza, si ottiene un rapporto di popolazione di uno a tre per i sottostati di singoletto e di tripletto, rispettivamente.

Guidato dall'attrazione tra elettrone e lacuna, l'elettrone si muoverà ulteriormente nella matrice verso la lacuna intrappolata. Quando l'elettrone raggiunge le molecole della matrice prossime a quella dell'emettitore, avviene una sovrapposizione delle funzioni d'onda associate all'elettrone e alla lacuna. L'interazione a corto raggio risul-

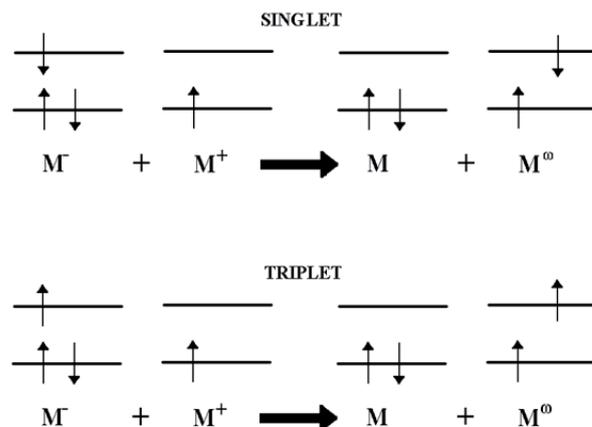


Figura 11: Schematizzazione dei livelli di spin negli stati di singoletto e tripletto. Quando uno dei due elettroni è eccitato a un livello energetico superiore, i loro spin potranno essere dello stesso segno oppure di segno opposto. Nel primo caso si parla di stati di tripletto, nel secondo caso di stati di singoletto.

tante porta ad uno *split* $\Delta E(S - T)$ tra l'energia dello stato di singoletto (S) e quello di tripletto (T). Questa energia è proporzionale approssimativamente a $\exp(-\alpha R)$, dove α è una costante che

dipende dalla matrice e dall'emettitore utilizzati.

Nella fase finale, l'elettrone salta direttamente, in tempi molto rapidi, alla molecola di emettitore che passa così al suo stato eccitato. Questo processo può coinvolgere sia stati di singoletto che di tripletto (*S-path*, *T-path*) a seconda dello *spin* iniziale della coppia elettrone-lacuna. Le costanti di tempo corrispondenti sono dell'ordine di un picosecondo. Il popolamento degli stati S_m e T_m , come mostrato in figura 10, è raffigurato solo come esempio. Successivamente, il sistema presenterà il tipico comportamento di una molecola di emettitore otticamente eccitata, con processi di rilassamento verso stati eccitati a più bassa energia ed emissione di radiazione. Per una discussione dettagliata delle proprietà fotofisiche di materiali metallo-organici con emissione da tripletto si rimanda a [16], [20] e [21]. Gli stati elettronici del materiale emettitore sono illustrati anch'essi in figura 10.

In un dato punto dello strato elettroluminescente, la probabilità che un elettrone collida con una lacuna è proporzionale al prodotto delle concentrazioni delle due specie in prossimità del punto considerato. Questo meccanismo è certamente reso più efficiente in dispositivi a multistrato dove il confinamento indotto dall'eterostruttura causa un rilevante incremento della densità di carica locale e quindi della probabilità di ricombinazione dei portatori. I materiali che costituiscono gli strati adiacenti sono scelti in modo che i livelli di LUMO (HOMO) siano sufficientemente differenti da creare un'effettiva barriera per gli elettroni (lacune). La densità di portatori viene pertanto ad essere incrementata nelle vicinanze dell'interfaccia e la probabilità di ricombinazione è aumentata in tale regione.

Come detto, la *rate* di ricombinazione è proporzionale al prodotto della densità di elettroni (n) e di lacune (p) secondo il modello di *Langevin*:

$$R = \gamma(np)$$

dove la costante di proporzionalità γ è data dalla relazione $\gamma = 4\pi e\mu_m/\varepsilon$, in cui μ_m è il massimo tra μ_n e μ_p , con cui sono indicate rispettivamente le mobilità degli elettroni e delle lacune, ed ε è la costante dielettrica del mezzo.

Integrando quest'espressione lungo lo spessore del dispositivo L si ricava il valore della densità

di corrente netta di ricombinazione J_r :

$$J_r = \int_0^L eRdx = J_n(L) - J_n(0)$$

dove $J_n(x)$ rappresenta la densità di corrente di elettroni in funzione della posizione x .

In realtà, però, quando si forma un eccitone, la probabilità di emissione di luce dipende dalla competizione tra il canale di decadimento radiativo e quello non radiativo. Se entrambi i canali coinvolgono una sola molecola con tempi di decadimento τ_r e τ_{nr} , rispettivamente nel caso radiativo e non radiativo, allora la *rate* di decadimento di una popolazione di N eccitoni è dato dalla relazione:

$$\frac{dN}{dt} = \frac{N}{\tau_r} + \frac{N}{\tau_{nr}}$$

Le proprietà di luminescenza del mezzo attivo sono determinate dai diversi meccanismi chimico-fisici che inducono atomi e molecole a portarsi in uno stato eccitato, ossia a trovarsi in configurazioni di energia maggiore rispetto a quella dello stato fondamentale. Il decadimento radiativo da uno stato eccitato a quello fondamentale dà luogo all'emissione di un fotone nel rispetto del principio di conservazione dell'energia. Va sottolineato, comunque, che non tutte le transizioni da uno stato eccitato ad uno a minor energia si traducono in emissione di luce. In accordo col principio di Pauli, gli elettroni che occupano un orbitale molecolare non possono avere gli stessi 4 numeri quantici, ossia ci saranno due elettroni con *spin* opposto su uno stesso orbitale. Le transizioni dagli stati di tripletto e singoletto a quello fondamentale sono governati dalla regola di selezione, che permette solo quelle tra stati aventi identica molteplicità di *spin*. Questo dovrebbe implicare che non saranno possibili transizioni tra gli stati di tripletto e quelli di singoletto. In realtà, tali transizioni avvengono, anche se con bassissima probabilità. Le differenti transizioni tra gli stati possono essere riassunte nel diagramma di *Jablonski* (figura 12), in cui possiamo distinguere 2 principali meccanismi di emissione:

Fluorescenza: in cui la transizione avviene dal più basso livello energetico corrispondente allo stato di singoletto S1 allo stato fundamen-

tale S₀. L'emissione è in questo caso spontanea ed il tempo di decadimento dipende dalla temperatura.

Fosforescenza: in cui la transizione avviene dal più basso livello energetico di tripletto T₁ allo stato fondamentale S₀. L'emissione è stimolata ed il tempo di decadimento è in genere molto maggiore rispetto a quello del processo di fluorescenza.

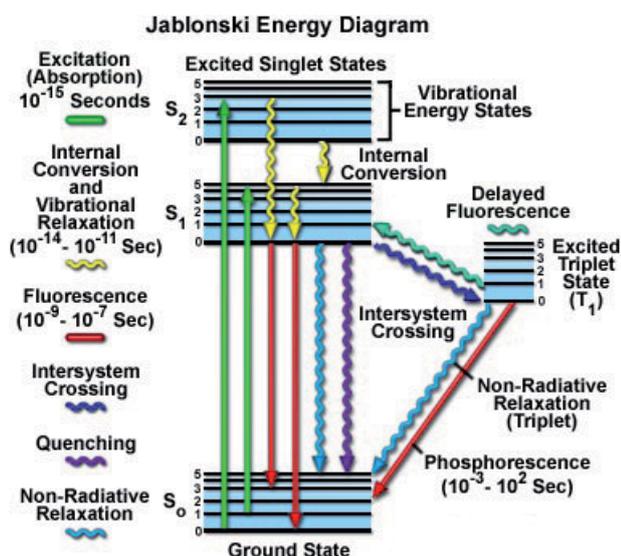


Figura 12: Diagramma di Jablonsky. In esso sono schematizzate le principali transizioni tra stati energetici nelle molecole organiche.

Come detto, però, oltre alle transizioni appena citate, caratterizzate entrambe da *decadimento radiativo*, possono verificarsi anche transizioni che non comportano l'emissione di un fotone ossia si può avere il *decadimento non radiativo* dell'eccitone. Definiamo, senza entrare nei dettagli, i principali canali di decadimento radiativo intramolecolari, riferendoci per semplicità al caso in cui l'eccitazione sia generata all'assorbimento di un fotone di opportuna lunghezza d'onda.

Nel caso in cui l'eccitone si formi in seguito all'assorbimento di un fotone, e se l'energia di eccitazione posseduta da quest'ultimo non coincide esattamente con la differenza energetica tra i due stati elettronici (stato fondamentale e stato eccitato), la molecola si porta su uno dei livelli rotovibrazionali dello stato elettronico eccitato. Dopo l'eccitazione la molecola rilassa sul primo livello vibrazionale dello stato elettronico eccitato, dissipando energia in modo non radiativo. Tale decadimento non radiativo è detto *internal conversion* (IC) (figura 12).

Un altro processo di diseccitazione non radiativa si può avere se il livello vibrazionale dello stato elettronico eccitato si sovrappone energeticamente con un livello vibrazionale di un altro stato elettronico, non accoppiato radiativamente con lo stato fondamentale. La molecola passa dal primo al secondo stato elettronico mediante *internal conversion* e poi ricombina non radiativamente con lo stato fondamentale sempre per *internal conversion*.

Triplet Harvesting

Scegliendo opportunamente il materiale emettitore drogante, è possibile incrementare l'efficienza di elettroluminescenza. A tal proposito, i composti organometallici hanno riscosso un interesse crescente negli ultimi anni poiché presentano caratteristiche interessanti per nuove applicazioni fotofisiche e fotochimiche. Il loro colore di emissione può variare in tutta la gamma del visibile, dal blu al rosso e anche all'infrarosso [16]. In questo paragrafo, si analizza l'emissione da tripletti di composti organici che contengono metalli di transizione, poiché sono stati utilizzati con successo nella fabbricazione di OLED ad alta efficienza [17]. Con l'uso di questi composti, è possibile avere, almeno in linea di principio, una efficienza di elettroluminescenza fino a quattro volte maggiore di quella che si può ottenere con la sola emissione da singoletto. Questa proprietà è legata ai meccanismi specifici di formazione dell'eccitone nella zona della ricombinazione elettrone-lacuna e ad un veloce ed efficiente *intersystem crossing* (ISC) dallo stato eccitato di singoletto allo stato di emissione di tripletto. Questo processo di accumulazione dell'energia di eccitazione nel più basso stato di tripletto eccitato è spesso chiamato *triplet harvesting*.

L'accoppiamento spin-orbita indotto dallo ione metallico posto al centro del complesso emettitore non altera fortemente il meccanismo di formazione dell'eccitone all'interno di una matrice organica, ma ha effetti notevoli sull'efficienza di elettroluminescenza in un dispositivo OLED. Per illustrare questa proprietà, metteremo a confronto l'efficienza raggiungibile con una molecola puramente organica con l'efficienza ottenibile con un complesso metallo-organico, se entrambe le molecole presentano uguali rendimenti quantici

di fotoluminescenza. Se si assume che il processo iniziale di formazione degli eccitoni si verifica statisticamente in funzione degli orientamenti di spin, si ottiene, come detto, il 25% di eccitoni di singoletto e il 75% di eccitoni di tripletto. Per le piccole molecole, questo risultato è ampiamente accettato [17]-[19].

Dopo la formazione dell'eccitone e il rilassamento dallo stato di singoletto o di tripletto, come discusso in precedenza, il più basso degli stati eccitati di singoletto o di tripletto è stato popolato. Ciò vale sia per i materiali emettitori puramente organici che per quelli organo-metallici.

La molecola organica può esibire fluorescenza, che si ha nella rapida transizione $S1 \rightarrow S0$, dato che di solito il tempo di decadimento $S1 \rightarrow T1$ è maggiore. D'altra parte, poiché la rate di transizione radiativa $T1 \rightarrow S0$ è anche piccola, la disattivazione dello stato $T1$ è normalmente non radiativo a temperatura ambiente. Pertanto, il 75% degli eccitoni, gli eccitoni di tripletto, si perdono. La loro energia viene dissipata sotto forma di calore. Le condizioni perché ciò avvenga sono più favorevoli per complessi contenenti metalli di transizione, in cui lo ione metallico comporta un notevole accoppiamento spin-orbita. Per questi complessi, l'*intersystem crossing* allo stato di tripletto $T1$ è solitamente molto efficiente e quindi un'emissione dal singoletto $S1$ non è osservabile. Inoltre, la rate di transizione radiativa $T1 \rightarrow S0$ può diventare sufficientemente alta da dare origine a fosforescenza anche a temperatura ambiente (per una discussione più dettagliata si veda [31]). Di conseguenza, tutti i quattro possibili orientamenti di spin degli eccitoni possono essere raccolti per popolare lo stato $T1$ più basso. E' quindi per tale processo di raccolta di emissione dai tripletti che si può ottenere, in linea di principio, una efficienza di elettroluminescenza fino a quattro volte più grande rispetto all'emissione da singoletto.

La tecnologia delle microcavità

I processi che coinvolgono l'interazione tra luce e materia sono di fondamentale importanza in numerosi ambiti. Un esempio importante è il trasferimento dell'energia di eccitazione da una molecola donatore eccitata ad una molecola accettore, attraverso l'interazione dipolo-dipolo (RDDI).

Oltre al suo ruolo chiave nella fotosintesi, [22] questo processo è di crescente importanza come mezzo per migliorare l'efficienza dei diodi emettitori di luce e laser basati su materiali organici [23].

Il controllo dell'emissione spontanea di luce può essere ottenuto mediante l'uso di strutture nelle quali la densità dei modi fotonici viene modificata, manipolando così i modi ottici per i quali l'emissione può aver luogo [24]. Si basano su questo concetto numerosi studi che riguardano, ad esempio, l'elettrodinamica quantistica in cavità [?]. E' noto che, ponendo un emettitore all'interno di una geometria confinata, non solo si altera il percorso della radiazione dell'emettitore ma si modifica anche la *rate* di decadimento della sua emissione spontanea (SpE).

La fluorescenza, come visto in precedenza, è un esempio di processo di emissione spontanea. In questo caso l'emettitore è un atomo, molecola o ione, e si considera che sia inizialmente in uno stato eccitato. L'emissione spontanea deriva dal decadimento dell'emettitore ad uno stadio inferiore di energia in seguito al quale viene generato un fotone che porta con sé l'energia dell'eccitazione.

L'emissione spontanea è l'archetipo di un processo meccanico quantistico; non possiamo dire nulla su quando si verificherà un certo evento di emissione spontanea; piuttosto, possiamo solo prevedere la probabilità di decadimento di un insieme di emettitori. È quindi sorprendente quando si scopre che il processo di emissione spontanea non è del tutto intrinseco ma è soggetto a influenze esterne.

La probabilità di emissione spontanea segue la *regola d'oro di Fermi* [?]:

$$\Gamma_{ij} \propto |M_{ij}|^2 \rho(\nu_{ij})$$

dove Γ_{ij} è la *rate* di transizione dallo stato eccitato i a quello a minore energia j ; M_{ij} , è un elemento di matrice che collega il livello eccitato a quello a più bassa energia ed è determinata dalle funzioni d'onda associate a tali livelli; $\rho(\nu_{ij})$ è la densità del campo ottico alla frequenza di transizione, di seguito denominata densità dei modi fotonici (PMD), e verrà discussa più avanti. Questo controllo della rate di decadimento attraverso la PMD è stato osservato (a radiofrequenze) da

Edward Mills Purcell nel 1946 [27].

Come può un'interfaccia modificare la *rate* di emissione spontanea? Si può intervenire sul termine M_{ij} nell'equazione precedente? In altre parole, può un'interfaccia modificare la funzione d'onda dell'emettitore? La funzione d'onda può essere perturbata, a condizione che l'emettitore e l'interfaccia siano abbastanza vicini. Tipicamente, perché ciò accada, la separazione tra l'emettitore e un atomo che costituisce l'interfaccia, deve essere quella caratteristica della funzione d'onda, di circa 1Å .

Anche se questo tipo di fenomeno è molto importante, ci interessa il secondo termine $\rho(\nu_{ij})$, la PMD alla frequenza di transizione. Attraverso questo termine l'interfaccia può modificare il processo di emissione spontanea; in tal caso entra in gioco la separazione caratteristica tra la molecola e l'interfaccia su cui è efficace, che deve essere dello stesso ordine di grandezza della lunghezza d'onda dell'emissione.

Ci sono due principali modi di vedere la PMD: in primo luogo, come fluttuazione del vuoto, che è un punto di vista essenzialmente quantomeccanico; in secondo luogo, come la capacità della struttura che circonda l'emettitore di "sostenere" il modo elettromagnetico relativo al fotone emesso, che è un punto di vista essenzialmente classico. Si tratta di interpretazioni complementari dal momento che forniscono le stesse risposte quantitative quando applicate allo stesso problema. In realtà, ciò non è così sorprendente, poiché l'interpretazione quantistica della luce si distingue realmente da quella classica solo nelle sue previsioni sul carattere statistico (cioè sul comportamento temporale) del campo ottico [28]. Qui tratteremo l'argomento principalmente dal punto di vista classico.

Si consideri un emettitore posto in un volume sferico chiuso la cui superficie interna è perfettamente riflettente. Se l'emettitore decade, si eccita il campo elettromagnetico all'interno del volume chiuso. Dal momento che questa energia non può essere irradiata al di fuori del volume, potrebbe causare successivamente una nuova eccitazione dell'emettitore (assumendo che quest'ultimo sia un semplice sistema a due livelli); dopodiché l'intero processo può ricominciare. Se la cavità non è un multiplo della lunghezza d'onda di emissione, allora nessun modo elet-

tromagnetico può essere sostenuto all'interno a causa dell'interferenza distruttiva delle riflessioni successive sulle superfici che definiscono la cavità. L'emettitore non è in grado di emettere e la struttura circostante non è in grado di supportare il modo. In questo caso, diciamo che la PMD alla frequenza di emissione è molto bassa e l'emissione spontanea è inibita.

Se rilassiamo la condizione al contorno e gli permettiamo di trasmettere una piccola frazione della radiazione emessa, allora vedremo che la cavità è popolata da un fotone e che l'energia oscilla tra il fotone e l'eccitone: queste sono le cosiddette oscillazioni di Rabi. Se lo scambio risonante di energia tra la cavità e l'emettitore è maggiore della probabilità di trasmissione, allora si avranno due diverse frequenze di risonanza consentite dalla cavità (*Rabi splitting*), a causa del forte accoppiamento tra la cavità e l'emettitore.

La cavità discussa sopra è un caso ideale ed è di scarso valore pratico. Vi sono, tuttavia, molte altre strutture che possono essere utilizzate per modificare la PMD attorno ad un emettitore. Numerosi studi sono stati condotti su strutture più complesse specificamente destinate a permettere il controllo della PMD, in particolare le strutture fotoniche [?] e la microcavità (figura 13) [30]. L'importanza di tali strutture, in particolare nel controllo dell'emissione spontanea nei LED e nei laser, è ben nota [31].

La modifica dell'emissione spontanea prodotta da un'interfaccia planare è dello stesso tipo di quella che si verifica nella cavità sferica ideale considerata in precedenza. Il campo riflesso dall'interfaccia interferisce con l'emettitore. Se il campo riflesso è in fase, la PMD intorno all'emettitore sarà alta e di conseguenza, l'emissione sarà incrementata. Se il campo riflesso è fuori fase, l'emissione sarà inibita (figura 14). In questo caso occorre considerare tutti i diversi modi per i quali l'emissione spontanea può aver luogo. La presenza dell'interfaccia fornisce percorsi aggiuntivi che coinvolgono la riflessione dalla superficie ed è l'inclusione di questi percorsi aggiuntivi che produce la modifica. Questo è forse l'esempio più semplice della branca della fisica chiamata elettrodinamica quantistica in cavità (QED).

Osservazioni sulla fluorescenza in presenza di una superficie piana hanno fornito la prima di-

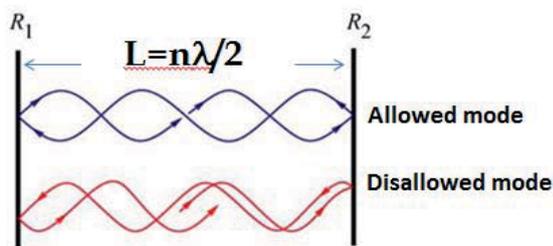
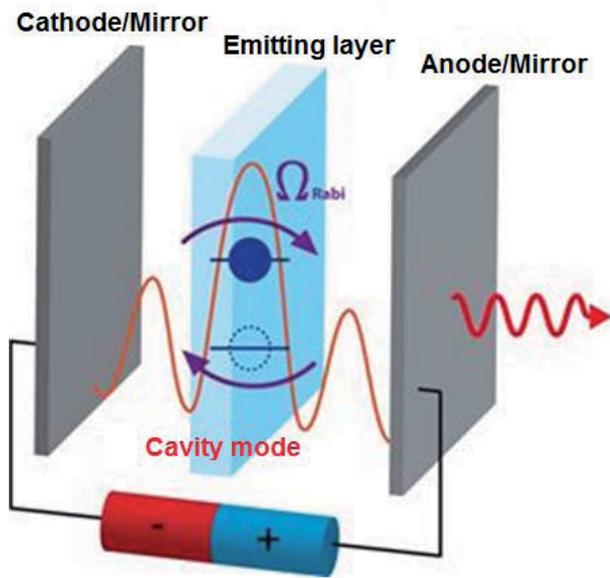


Figura 13: In alto, rappresentazione di una microcavità con specchi metallici, di cui uno semitrasparente. In basso: se la distanza tra le due superfici piane riflettenti è un multiplo della lunghezza d'onda di emissione, questa sarà supportata dalla cavità e si formerà un'onda stazionaria delimitata dagli specchi riflettenti.

mostrazione sperimentale che la *rate* di emissione spontanea potrebbe essere modificata cambiando la PMD locale. Fu Drexhage [32] [33] nel 1960 ad eseguire i primi esperimenti. Questa è stata dimostrata misurando il tempo di vita dei fotoni emessi da un singolo strato attivo in funzione della sua distanza da uno specchio metallico planare. Per separare il film emissivo dallo specchio sono stati usati degli strati di materiale dielettrico. Il tempo di vita dell'emissione spontanea, ovvero l'inverso della frequenza, ha evidenziato un comportamento oscillatorio in funzione della distanza emettitore-metallo ed è stata osservata sia l'inibizione che l'incremento dell'emissione. In seguito ad una riduzione della distanza al di sotto dei 50 nm, si è osservato il *quenching* della fluorescenza, che è stato associato con l'accoppia-

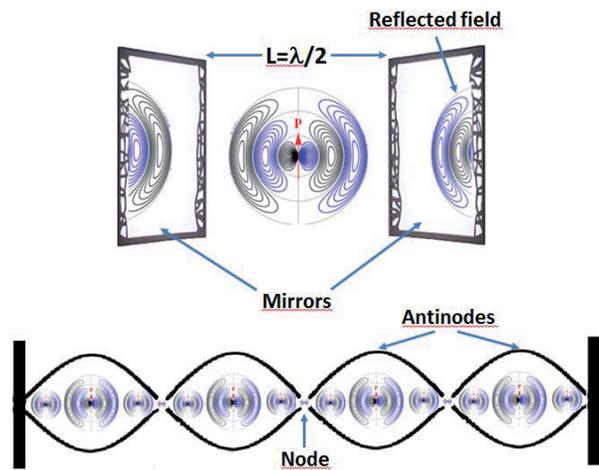


Figura 14: In alto: rappresentazione di un dipolo (nel nostro caso un eccitone che, decadendo, emette luce) posto all'interno di una cavità delimitata da interfacce planari riflettenti; il campo emesso dalla molecola sarà riflesso tra gli specchi. In basso: poichè il campo elettrico è nullo in corrispondenza dei nodi e massimo in corrispondenza degli antinodi, si può avere una amplificazione o uno smorzamento dell'emissione a seconda della posizione del dipolo all'interno della cavità.

mento ai modi di superficie non radiativi dello specchio d'argento.

L'incremento della rate di emissione dovuto all'introduzione di un modo di guida d'onda all'interno di una microcavità è stato anche studiato teoricamente e sono state ricavate delle equazioni piuttosto semplici [36]. La modifica della *rate* di emissione dipende dall'orientamento del momento di dipolo degli emettitori [34]. L'incremento maggiore è relativo all'introduzione del primo modo di guida d'onda all'interno della microcavità. Per una microcavità riempita di vuoto delimitata da specchi perfetti, il miglioramento è di un fattore 3 e si verifica quando il momento di dipolo è parallelo alle interfacce [35]. Le microcavità sono già state utilizzate sia per selezionare spettralmente le emissioni da un emettitore a larga banda [37] sia per aumentare l'intensità di emissione in una direzione specifica [38]. Entrambi questi effetti sono causati dalla modifica che induce una microcavità sulla densità dei modi fotonici.

Vediamo ora come è stato possibile intervenire sui parametri γ , $\eta_{s/ph}$ e φ_f in un caso concreto: la fabbricazione di un OLED ad emissione di luce bianca.

Modulazione dello spettro di elettroluminescenza di OLED bianchi tramite multicavità organo-metalliche accoppiate.

La tecnologia degli OLED bianchi (WOLED) attira notevole interesse per la potenziale applicazione nell'illuminazione di interni. Tuttavia, costituisce un obiettivo ambizioso il simultaneo raggiungimento di una elevata efficienza luminosa (LE) e di un alto indice di resa cromatica (CRI), compatibile con l'utilizzo di substrati eventualmente flessibili.

I dispositivi ad altissima efficienza mostrano valori solitamente bassi di CRI, non adatti per applicazioni di illuminazione, e viceversa. Recentemente, è stata ideata e realizzata una nuova architettura a multi-cavità per WOLED, con la quale è stato possibile controllare la distribuzione della densità dei modi fotonici (PMD) ed ottenere la simultanea ottimizzazione di tutte le caratteristiche optoelettroniche del dispositivo [39]. Infatti, senza l'utilizzo di ottiche di accoppiamento o di substrati ad alto indice di rifrazione, è stato raggiunto un CRI di 85 ed LE pari a 33 lm/W e 14 lm/W in dispositivi realizzati su vetro privo di ITO e su substrati flessibili, rispettivamente.

L'architettura del dispositivo è riportata in figura 15. Il dispositivo (in seguito indicato con l'acronimo MC-WOLED) è formato da tre cavità (C1, C2 e C3) composte da quattro specchi metallici di argento separati da film organici spessi alcune decine di nanometri. Gli strati metallici comuni, M1 ed M2, sono semitrasparenti in quanto hanno la funzione di accoppiare le cavità; il grado di accoppiamento è determinato dal loro spessore.

Il primo strato metallico AgB è anch'esso semitrasparente per consentire ai fotoni generati di fuoriuscire dal dispositivo, mentre l'ultimo, AgT, è completamente riflettente. La struttura è realizzata su un substrato trasparente (vetro o PET) con la seguente sequenza di strati: AgB / O1 / M1 / O2 / M2 / O3 / AgT. O1 è costituito da una sequenza di film organici, ed è detto *stack* attivo, poiché è quello che contiene i materiali emettitori; viene alimentato elettricamente tramite AgB ed M1, che quindi fungono anche da elettrodi del dispositivo. Al contrario, O2 e O3 sono de-

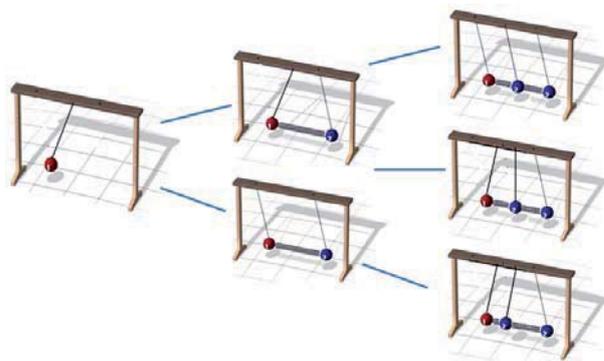


Figura 15: Il principio fisico alla base delle cavità accoppiate è analogo a quello su cui si basa un sistema costituito da due o più oscillatori armonici collegati tramite delle molle; nel caso delle cavità accoppiate, gli strati organici O1, O2 e O3 fungono da oscillatori, mentre gli strati metallici M1 ed M2 da molle. Il grado di accoppiamento delle cavità (associabile alla costante elastica delle molle) è determinato dallo spessore degli strati di metallo. In tal modo si verranno a formare due o più frequenze di oscillazione proprie del sistema.

gli *stack* organici passivi. L'uso di strati metallici sottili consente di avere valori di trasmittanza relativamente alti tali per cui le tre cavità sono in comunicazione l'una con l'altra. Si vengono così a generare tre modi elettromagnetici indipendenti, le cui lunghezze d'onda sono completamente modulabili semplicemente agendo sugli spessori degli strati metallici e organici (figura 16).

Il dispositivo è stato progettato per generare tre modi di risonanza corrispondenti a tre lunghezze d'onda: una rossa, una verde ed una blu (RGB); al fine di indurre gli effetti di amplificazione dell'emissione di ciascun colore, si è agito sugli spessori degli strati di trasporto delle cariche in maniera tale che i picchi dei modi di risonanza relativi allo *stack* O1 si trovassero in corrispondenza degli strati di emissione RGB. Infatti, come visto in precedenza, quando gli emettitori sono posizionati in corrispondenza degli antinodi dei campi interni alla cavità C1, si verifica una amplificazione della forza di oscillatore delle molecole [40] [41], con un conseguente aumento della luminosità del dispositivo. Questa configurazione determina un ampio spettro di emissione e un elevato effetto di risonanza su tutti i modi RGB.

Il miglioramento della luminanza è legato alla ridistribuzione della densità dei modi fotonici, indotta dagli effetti di cavità, attorno alle lun-

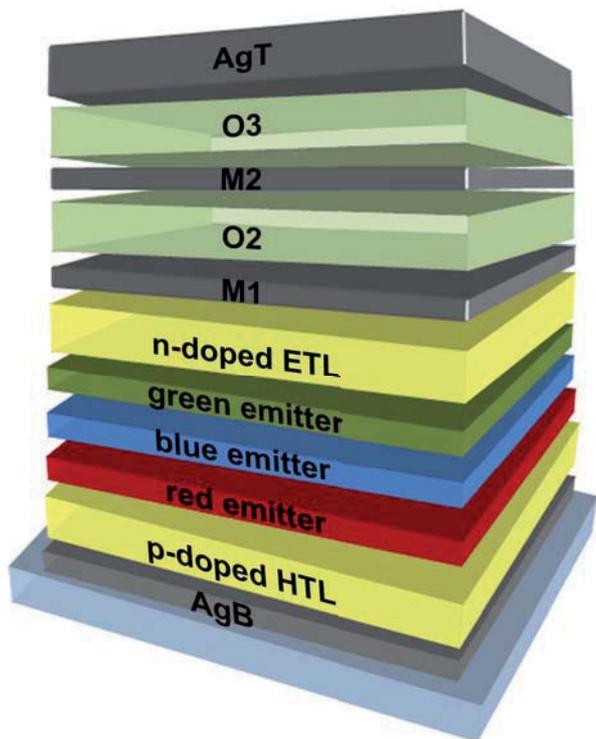


Figura 16: Schematizzazione del MC-WOLED. Lo strato emissivo è posto nella prima cavità, C1, ed è composto da tre film di materiali fosforescenti: rosso, verde e blu.

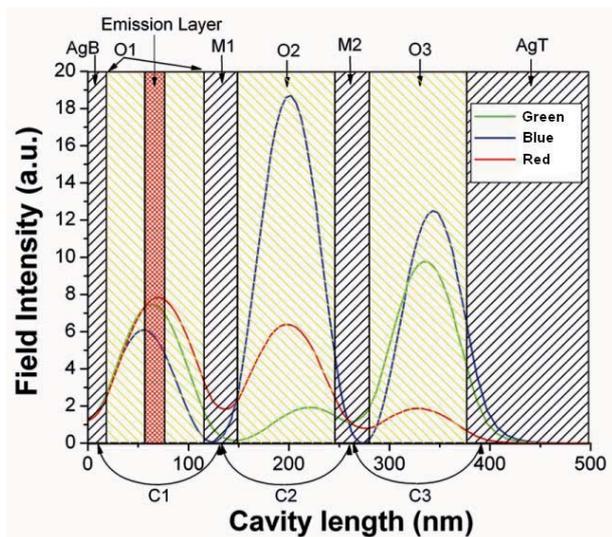


Figura 17: Distribuzione dei campi ottici risonanti, relativi alle lunghezze d'onda RGB, all'interno della migliore struttura MC-WOLED realizzata (CRI=85). La regione in arancione rappresenta lo strato attivo completo, dove sono collocate le molecole fosforescenti ad emissione rossa, verde e blu.

ghezze d'onda di emissione dei composti luminescenti, con un incremento della *rate* di decadimento radiativo degli eccitoni [40] [41]; abbiamo

quindi visto come, utilizzando gli effetti di cavità, si è potuto intervenire sull'efficienza quantica φ_f . La distribuzione dei campi del migliore MC-WOLED è illustrato in figura 17.

Inoltre, avendo utilizzato come emettitori dei composti fosforescenti a base di iridio, è stato sfruttato il meccanismo del *triplet harvesting* per la raccolta degli eccitoni di tripletto con un conseguente intervento sul parametro $\eta_{s/ph}$.

Come detto, il corretto posizionamento dei picchi dei modi di risonanza all'interno della cavità C1 è stato ottenuto agendo sugli spessori degli strati di trasporto delle cariche, reso possibile, senza effetti negativi sulle prestazioni del dispositivo, grazie all'utilizzo della tecnologia *p-i-n* che prevede il drogaggio elettrico di tali strati. Il drogaggio consente infatti di disaccoppiare le caratteristiche ottiche da quelle elettriche, rendendo più semplice l'ottimizzazione della struttura del dispositivo. I vantaggi del drogaggio non si limitano solo all'effetto suddetto, ma comprendono anche la forte riduzione delle perdite ohmiche all'interfaccia metallo/semiconduttore ed un aumento dell'iniezione delle cariche elettriche per effetto tunnel. Il drogaggio, in sostanza, ha permesso di modificare il parametro γ , che interviene nella determinazione dell'efficienza quantica interna, poiché consente di aumentare il numero di cariche che possono trasformarsi in eccitoni.

L'utilizzo di queste tecnologie ha permesso di raggiungere alti valori di CRI e di LE, simultaneamente. Si è così potuto ottenere un aumento del 35% della luminanza e un miglioramento del CRI da 79 a 85, rispetto a dispositivi bianchi tradizionali.

Conclusioni

Finora gli OLED hanno trovato applicazione principalmente nel settore dei display; la tecnologia sviluppata fino ad oggi ha permesso, infatti, di realizzare dispositivi ultrasottili con una vastissima gamma di colori. Gli OLED sono quindi riusciti ad entrare nel mercato e in pochi anni sostituiranno del tutto i tradizionali schermi LCD.

Oggi, la sfida che si sta affrontando è quella di rendere questa tecnologia adatta ad essere utilizzata in quei campi, quali *l'automotive* e *l'illu-*

minazione, che necessitano di emissioni di luce ad alta potenza; infatti per questo tipo di applicazioni, gli OLED presentano ancora dei limiti dal momento che è necessario combinare caratteristiche che sono spesso incompatibili per le tradizionali architetture OLED, ovvero: elevata luminosità, stabilità e purezza del colore, lunga durata e larga area emissiva.

In questo articolo sono stati illustrati i principali meccanismi che determinano il funzionamento di un OLED e alcune delle tecnologie utili per intervenire efficacemente sulle prestazioni dei dispositivi, al fine di aumentarne la durata e l'efficienza.

Questi approcci costituiscono un punto di partenza per lo sviluppo di OLED ad alta luminosità e di lunga durata, adatti a sostituire le tradizionali sorgenti luminose.

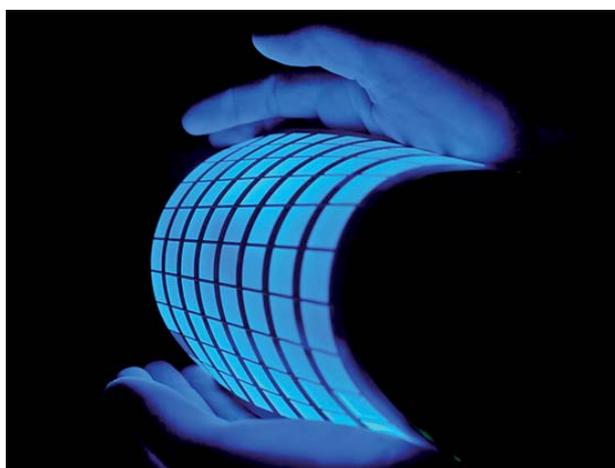


Figura 18: Matrice di OLED flessibili ad emissione blu



- [1] W. HELFRICH, W. G. SCHNEIDER: "Recombination Radiation in Anthracene Crystals", *Phys. Rev. Lett.* **14** (1965) 229.
- [2] C. W. TANG, S. A. VANSLYKE: "Organic electroluminescent diodes", *Appl. Phys. Lett.* **51** (1987) 913.
- [3] J. H. BURROUGHS, D. D. C. BRADLEY, A. R. BROWN, R. N. MARKS, K. MACKAY, R. H. FRIEND, P. L. BURNS AND A. B. HOLMES: "Light-emitting diodes based on conjugated polymers", *Nature* **347** (1990) 539.
- [4] H.S. NALWA, AND L.S. ROHWER: *Handbook of Luminescence, Display Materials, and Devices*. American Scientific Publishers, Los Angeles, California, USA, 2003 (
- [5] H. MURATA, Z.H. KAFABI, AND M. UCHIDA: "Efficient organic light-emitting diodes with undoped active layers based on silole derivatives", *Appl. Phys. Lett.* **80** (2002) 189.
- [6] H. BECKER, S.E. BURNS, AND R.H. FRIEND: "Effect of metal films on the photoluminescence and electroluminescence of conjugated polymers", *Phys. Rev. B - Condensed Matter* **56** (1997) 1893.
- [7] M.A. BALDO, C. ADACHI, AND S.R. FORREST: "Transient analysis of organic electrophosphorescence. II. Transient analysis of triplet-triplet annihilation", *Phys. Rev. B.-Condensed matter* **62** (2000) 10967.
- [8] J. SZMYTKOWSKI, W. STAMPOR, J. KALINOWSKI, AND Z.H. KAFABI: "Electric field-assisted dissociation of singlet excitons in tris-(8-hydroxyquinolato) aluminum (III)", *Appl. Phys. Lett.* **80** (2002) 1465.
- [9] I. SCHNITZER, AND E. YABLONOVICH: "30 light-emitting diodes", *Appl. Phys. Lett.* **63** (1993) 2174.
- [10] G. GU, D.Z. GARBUZOV, P.E. BURROWS, S. VENKATESH, S.R. FORREST, AND M.E. THOMPSON: "High-external-quantum-efficiency organic light-emitting devices", *Opt.Lett.* **22** (1997) 396.
- [11] C. F. MADIGAN, M. H. LU, AND J. C. STRURM: "Improvement of output coupling efficiency of organic light-emitting diodes by backside substrate modification", *Appl. Phys. Lett.* **76** (2000) 1650.
- [12] T. YAMASAKI, K. SUMIOKA, AND T. TSUTSUI: "Organic light-emitting device with an ordered monolayer of silica microspheres as a scattering medium", *Appl. Phys. Lett.* **76** (2000) 1243.
- [13] F. MARIANO, M. MAZZEO, Y. DUAN, G. BARBARELLA, L. FAVARETTO, S. CARALLO AND G. GIGLI: "Very low voltage and stable p-i-n Organic-Light-Emitting-Diodes using a linear S,S-dioxide oligothiophene as emitting layer", *Appl. Phys. Lett.* **94** (2009) 063510.
- [14] M. PFEIFFER, A. BEYER, T. FRITZ, AND K. LEO: "Controlled doping of phthalocyanine layers by cosublimation with acceptor molecules: A systematic Seebeck and conductivity study", *Appl. Phys. Lett.* **73** (1998) 3202.
- [15] A. NOLLAU, M. PFEIFFER, T. FRITZ, AND K. LEO: "Controlled n-type doping of a molecular organic semiconductor: Naphthalenetetracarboxylic dianhydride (NTCDA) doped with bis(ethylenedithio)-tetrathiafulvalene (BEDT-TTF)", *J. Appl. Phys.* **87** (2000) 4340.
- [16] H. YERSIN, D. DONGES: "Low-Lying Electronic States and Photophysical Properties of Organometallic Pd(II) and Pt(II) Compounds. Modern Research Trends Presented in Detailed Case Studies", *Top. Curr.Chem.* **81** (2001) 214.
- [17] Y.-Y. LIN, S.-C. CHAN, M.C.W. CHAN, Y.-J. HOU, N. ZHU, C.-M.CHE, Y. LIU; Y. WANG: "Structural, Photophysical, and Electrophosphorescent Properties of Platinum(II) Complexes Supported by Tetradentate N2O2 Chelates", *Chem. Eur. J.* **9** (2003) 1263.
- [18] H. YERSIN: "Triplet Emitters for OLED Applications. Mechanisms of Exciton Trapping and Control of Emission Properties", *Top. Curr. Chem.* **241** (2004) 1.
- [19] S. LAMANSKY, R.C. KWONG, M. NUGENT, P.I. DJUROVICH, M.E. THOMPSON: "Molecularly doped polymer light emitting diodes utilizing phosphorescent Pt (II) and Ir (III) dopants", *Org. Electron.* **2** (2001) 53.

- [20] W. FINKENZELLER, H. YERSIN: "Emission of Ir(ppy)₃. Temperature dependence, decay dynamics, and magnetic field properties", *Chem. Phys. Lett.* **377e** (2003) 299.
- [21] W. FINKENZELLER, P. STOESEL, H. YERSIN: "Emission and absorption of Ir(ppy)₂(CO)(Cl)-temperature dependence, phosphorescence decay dynamics, and assignment of excited states", *Chem. Phys. Lett.* **397** (2004) 289.
- [22] J.R. OPPENHEIMER: "Titolo", *Phys. Rev.* **60** (1941) 158.
- [23] M.A. BALDO, M.E. THOMPSON, AND S.R. FORREST: "High-efficiency fluorescent organic light-emitting devices using a phosphorescent sensitizer", *Nature* **403** (2000) 750.
- [24] K.H. DREXHAGE: *Progress in Optics*. E. Wolf Ed., North-Holland, Amsterdam, vol. XII, 1974 (
- [25] .chiave25P. BERMAN: *Cavity Quantum Electrodynamics, Advances in Atomic, Molecular and Optical Physics*. Academic Press, London, 1974 (
- [26] .chiave26E. FERMI: "Quantum theory of radiation", *Rev. mod. Phys.* **4** (1932) 87.
- [27] E.M. PURCELL: "Resonance absorption by nuclear magnetic moments in a solid", *Phys. Rev.* **69** (1946) 681.
- [28] P. MILONNI: *The Quantum Vacuum*. Academic Press, London, 186, 1994 (
- [29] .chiave29J.W. HAUS: "A brief review of theoretical results for photonic band structures", *J. Mod. Optics* **41** (1994) 195.
- [30] D. MESCHÉDE: "Radiating atoms in confined space: From spontaneous emission to micromasers", *Phys. Rep.* **211** (1992) 201.
- [31] R.E. SLUSHER AND W. WEISBUCH: "Optical microcavities in condensed matter systems", *Solid St. Commun.* **92** (1994) 149.
- [32] K.H. DREXHAGE, M. FLECK, F.P.SCHAFFER, AND W. SPERLING: "Beeinflussung der Fluoreszenz eines Europiumchelates durch einen Spiegel", *Ber. Bunsenges. Phys. Chem.* **20** (1966) 1179.
- [33] K.H. DREXHAGE, H. KUHN, AND F.P. SCHAFFER: "Variation of fluorescence decay time of a molecule in front of a mirror", *Ber. Bunsenges. Phys. Chem.* **72** (1968) 329.
- [34] R.R. CHANCE, A. PROCK, AND R. SILBEY: "Molecular fluorescence and energy transfer near interfaces", *Adv. Chem. Phys.* **37** (1978) 1.
- [35] S.M. DUTTRA AND P.L. KNIGHT: "Spontaneous emission in a planar Fabry-Pérot microcavity", *Phys. Rev.* **53** (1996) 3587.
- [36] G. BJORK, S. MACHIDA, Y. YAMAMOTO, AND K. IGETA: "Titolo", *Phys. Rev. A* **44** (1991) 669.
- [37] A. DODABALAPUR, L.J. ROTHBERG, R.H. JORDAN, T.M. MILLER, R.E. SLUSHER, AND J.M. PHILIPS: "Physics and applications of organic microcavity light emitting diodes", *Appl. Phys.* **80** (1996) 6954.
- [38] K.G. SULLIVAN AND D.G. HALL: "Enhancement and inhibition of electromagnetic radiation in plane-layered media. II. Enhanced fluorescence in optical waveguide sensors", *J. Opt. Soc. Am. B* **14** (1997) 1160.
- [39] M. MAZZEO, F. MARIANO, A. GENCO, S. CARALLO, AND G. GIGLI: "High Efficiency ITO-free flexible white organic light-emitting diodes based on multi-cavity technology", *Org. Electron.* **14** (2013) 2840.
- [40] W.L. BARNES: "Fluorescence near interfaces: the role of photonic mode density", *J. Mod. Opt.* **45** (1998) 661.
- [41] P.T. WORTHING, W.L. BARNES: "Spontaneous emission within metal-clad microcavities", *J. Opt. A: Pure Appl. Opt.* **1** (1999) 501.



Fabrizio Mariano: ha conseguito nel 2011 il titolo di Dottore di Ricerca in Scienze e Tecnologie Interdisciplinari presso l'Università del Salento. La sua attività di ricerca ha riguardato la realizzazione di OLED innovativi per applicazioni quali l'automotive e l'illuminotecnica.

Attualmente si occupa principalmente dello studio di LED ibridi e transistor elettroluminescenti per applicazioni optoelettroniche.

I Metamateriali: Invisibilità e Rifrazione Negativa in Mezzi Nanochirali

Marco Esposito

Francesco Todisco

*Dipartimento di Matematica & Fisica "Ennio De Giorgi" - Università del Salento,
CNR Nanotec
Dipartimento di Matematica & Fisica "Ennio De Giorgi" - Università del Salento,
CNR Nanotec*

I metamateriali sono, attualmente, una delle aree di ricerca più interessanti ed eccitanti nel settore delle nanotecnologie. Le proprietà elettromagnetiche che si possono ottenere con questi materiali sono inesistenti in natura, e per questo hanno attirato l'attenzione mondiale da innumerevoli punti di vista. Negli ultimi venti anni, alcuni dei fenomeni previsti dapprima solo teoricamente, come l'invisibilità, la rifrazione negativa e le superlenti, sono stati ampiamente dimostrati in diversi sistemi. Molte restano, tuttavia, le strade inesplorate, tutte ugualmente affascinanti e potenzialmente rivoluzionarie. Da questo punto di vista, dati i progressi delle tecnologie di fabbricazione, l'unico limite sembra essere la nostra capacità di immaginare nuove geometrie, nuovi fenomeni: nuovi

metamateriali.

Molte delle proprietà macroscopiche dei materiali sono spiegate dalla fisica dello stato solido attraverso una diretta relazione con le loro caratteristiche microscopiche. Le proprietà elettromagnetiche, ad esempio, dipendono fortemente da due fattori: le caratteristiche chimiche dei singoli atomi che compongono il materiale e il modo in cui questi sono ordinati nello spazio. Si tratta di due evidenze fondamentali che, insieme allo sviluppo, negli anni, di tecniche di nanofabbricazione sempre più elaborate, hanno aperto nuovi orizzonti nella ricerca di materiali dalle proprietà ottiche esotiche, difficili o addirittura impossibili da trovare in natura. I metamateriali, artificialmente prodotti in laboratorio, sono composti da celle elementari che prendono il posto delle molecole e degli atomi dei materiali tradizionali. In essi la geometria e l'ordine nello spazio riescono a modificare l'interazione con le onde elettromagnetiche, dalle microonde fino al visibile e l'infrarosso, fornendo loro caratteristiche uniche.

1 Dalle Equazioni di Maxwell ai Metamateriali

Le proprietà elettromagnetiche dei materiali sono descritte, attraverso le equazioni di Maxwell, da due quantità fondamentali: la permittività elettrica ϵ e la permeabilità magnetica μ . Si tratta di quantità specifiche dei materiali e sono, in generale, delle grandezze tensoriali. Dal punto di vista fisico, la permittività descrive in che modo un campo elettrico e il materiale si influenzano a vicenda, ovvero quanto il materiale si polarizza quando interagisce con un campo elettrico. Lo stesso vale per la permeabilità, dove il campo di interazione, però, è quello magnetico:

$$\mathbf{P} = \epsilon_0 (\epsilon_r - 1) \mathbf{E}$$

$$\mathbf{M} = (\mu_r - 1) \mathbf{H}$$

dove \mathbf{P} ed \mathbf{M} sono la polarizzabilità dielettrica e magnetica, rispettivamente, \mathbf{E} ed \mathbf{H} sono i campi elettrico e magnetico, ϵ_0 e μ_0 ed ϵ_r e μ_r sono la permittività e la permeabilità del vuoto e quelle relative del mezzo, rispettivamente. Semplificando rispetto alle costanti del vuoto, si definiscono le grandezze caratteristiche dei materiali $\epsilon = \epsilon_0 \epsilon_r$ e $\mu = \mu_0 \mu_r$.

Dal punto di vista macroscopico è possibile suddividere tutti i materiali esistenti in natura sulla base dei valori che assumono ϵ e μ , come riportato in Figura 1. Nella regione 1 si trovano i materiali con ϵ e μ entrambi positivi. Si tratta del caso più comune, e ne fanno parte molti materiali dielettrici. Nel secondo quadrante ci sono i materiali che possiedono permittività negativa. Questo accade nei metalli, nei semiconduttori drogati e nei materiali ferroelettrici, almeno in determinati intervalli di lunghezza d'onda, al di sotto della cosiddetta frequenza di plasma. La regione 4 comprende alcuni materiali a base di ferrite che offrono permeabilità negative, ma il cui comportamento magnetico decade oltre il range delle microonde. Il terzo quadrante, quello dei materiali che offrono simultaneamente permittività e permeabilità negative, è, al contrario degli altri, completamente vuoto: non esistono in natura sostanze con questo tipo di caratteristiche.

Forse proprio per questa loro curiosa inesistenza, i materiali del terzo quadrante hanno attira-

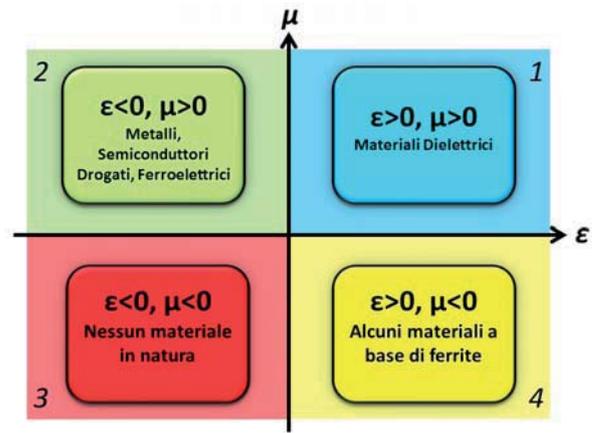


Figura 1: Spazio delle proprietà elettromagnetiche dei materiali, caratterizzate da permittività elettrica (asse x) e permeabilità magnetica (asse y).

to interesse, dapprima teorico, per predirne e valutarne le possibili proprietà elettromagnetiche. Nel 1968, infatti, fu il fisico teorico russo Victor Veselago ad analizzarne le possibili caratteristiche [1]. Consideriamo, ad esempio, cosa succede quando un'onda piana monocromatica si propaga in un mezzo del genere. In generale, le componenti elettrica e magnetica dell'onda si possono scrivere come $\mathbf{E}(\omega, \mathbf{k}) = \mathbf{E}_0 \cdot e^{i\mathbf{k} \cdot \mathbf{r} - i\omega t}$ e $\mathbf{H}(\omega, \mathbf{k}) = \mathbf{H}_0 \cdot e^{i\mathbf{k} \cdot \mathbf{r} - i\omega t}$, dove ω è la frequenza e \mathbf{k} il vettore d'onda. Le equazioni di Maxwell, in forma locale, sono

$$\begin{aligned} \nabla \cdot \mathbf{B} &= 0, & \nabla \cdot \mathbf{D} &= \rho \\ \nabla \times \mathbf{H} &= \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}, & \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \end{aligned}$$

dove $\mathbf{D} = \epsilon \mathbf{E} = \epsilon_0 \epsilon_r \mathbf{E}$ e $\mathbf{B} = \mu \mathbf{H} = \mu_0 \mu_r \mathbf{H}$ sono i campi di induzione elettrica e magnetica, rispettivamente. Se consideriamo il caso in cui non sono presenti né cariche libere (ρ) né correnti (\mathbf{J}), le equazioni si semplificano:

$$\nabla \times (\mathbf{E}_0 \cdot e^{i\mathbf{k} \cdot \mathbf{r} - i\omega t}) = -\frac{\partial}{\partial t} (\mu \cdot \mathbf{H}_0 \cdot e^{i\mathbf{k} \cdot \mathbf{r} - i\omega t}) \Rightarrow$$

$$\Rightarrow i\mathbf{k} \times \mathbf{E} = i\omega \mu \mathbf{H}$$

$$\nabla \times (\mathbf{H}_0 \cdot e^{i\mathbf{k} \cdot \mathbf{r} - i\omega t}) = \frac{\partial}{\partial t} (\epsilon \cdot \mathbf{E}_0 \cdot e^{i\mathbf{k} \cdot \mathbf{r} - i\omega t}) \Rightarrow$$

$$\Rightarrow i\mathbf{k} \times \mathbf{H} = -i\omega \epsilon \mathbf{E}$$

Si ottiene, quindi, il seguente sistema di equazioni

$$\begin{cases} \mathbf{k} \times \mathbf{E} = \mu\omega\mathbf{H} \\ \mathbf{k} \times \mathbf{H} = -\varepsilon\omega\mathbf{E} \end{cases}$$

Da queste equazioni si ricava che i vettori \mathbf{k} , \mathbf{E} ed \mathbf{H} formano una terna di vettori destrorsa quando un'onda piana propaga in un mezzo dielettrico con μ ed ε positive. Al contrario, se μ ed ε sono entrambe negative la tripletta è sinistrorsa e il vettore di Poynting, definito come $\mathbf{S} = \mathbf{E} \times \mathbf{H}$, è antiparallelo al vettore d'onda \mathbf{k} . Per la conservazione della causalità [2], inoltre, l'indice di rifrazione di questi materiali, definito come $n = \pm\sqrt{|\varepsilon_r||\mu_r|}$, assume segno negativo. A causa di queste ultime due proprietà, i materiali con ε e μ simultaneamente negativi, prendono il nome di materiali sinistrorsi o materiali a indice di rifrazione negativo.

Come descritto da Veselago nel suo lavoro, questi materiali avrebbero proprietà completamente fuori dalla *normalità* che conosciamo. Consideriamo, ad esempio, la rifrazione della luce in materiali ad indice di rifrazione positivo. Quando un'onda elettromagnetica si propaga da un mezzo con indice di rifrazione n_1 ad uno con indice di rifrazione n_2 , la sua direzione di propagazione θ_1 rispetto alla normale, viene deviata lungo la direzione θ_2 definita dalla legge di Snell

$$n_1 \sin\theta_1 = n_2 \sin\theta_2$$

Cosa succede se, invece, la luce si propaga passando da un mezzo ad indice di rifrazione positivo ad uno con indice di rifrazione negativo? In questo caso, in modo controintuitivo e assolutamente inesistente in natura, l'angolo di ingresso e quello di uscita devono avere seno di segno opposto per continuare a soddisfare la legge di Snell, ovvero la luce rifratta propagherà ad angoli negativi, dalla stessa parte della luce incidente, come indicato schematicamente in Figura 2. Questo è solo un esempio di quello che accade con materiali sinistrorsi. Anche l'effetto Doppler e l'effetto Cherenkov, ad esempio, sarebbero invertiti in questo tipo di materiali.

Dopo che Veselago descrisse teoricamente le proprietà di questi materiali nel 1968, la prima realizzazione sperimentale si ebbe nel 1996, quando il fisico inglese John Pendry realizzò un

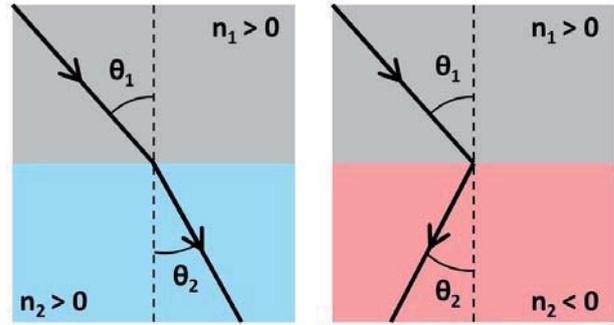


Figura 2: Rifrazione della luce in materiali ad indice di rifrazione positivo (a sinistra) e negativo (a destra).

sistema capace di bypassare i limiti imposti dalla natura, creando, così, il primo metamateriale. L'idea è, teoricamente, molto semplice. È, infatti, possibile ricreare in laboratorio materiali dalle proprietà ottiche ingegnerizzabili a piacimento, partendo da materiali comuni ma organizzandoli in singoli elementi (noti come meta-atomi) distribuiti periodicamente o in modo random, ed aventi dimensioni e spaziature molto più piccole della lunghezza d'onda della radiazione elettromagnetica con cui interagiscono. In questo modo, infatti, le proprietà microscopiche di ogni meta-atomo risultano essere "invisibili" alla radiazione elettromagnetica e la risposta del materiale viene influenzata dal comportamento collettivo di tutti i meta-atomi insieme. In altre parole, l'insieme disomogeneo di meta-atomi si può descrivere, dal punto di vista della risposta macroscopica, come un mezzo omogeneo caratterizzato da permittività elettrica e permeabilità magnetica effettive, $\varepsilon_{r,eff}$ e $\mu_{r,eff}$. Sulla base di queste osservazioni, Pendry ha dapprima dimostrato le proprietà di un array periodico di fili di rame che, con specifiche dimensioni, esibiva una risposta elettromagnetica con permittività elettrica negativa [3]. Successivamente lo stesso Pendry ha proposto array di split ring resonator (antenne a forma di anelli non completamente chiusi) che esibivano permeabilità magnetica negativa [4]. La combinazione delle due strutture portò, infine, alla realizzazione della struttura in Figura 3, che possiede simultaneamente $\mu_r < 0$ e $\varepsilon_r < 0$ nella regione spettrale delle microonde: il primo metamateriale.

Le Equazioni di Maxwell

Le equazioni di Maxwell, sviluppate dal matematico e fisico scozzese James Clerk Maxwell a metà del XIX secolo, sono un sistema di quattro equazioni differenziali (due vettoriali e due scalari) alle derivate parziali lineari. Questo sistema, insieme all'espressione della forza di Lorentz, permette di descrivere in maniera completa un sistema elettromagnetico, in quella che viene definita come elettrodinamica classica (distinta da quella, più complessa, quantistica che non trascura la natura quantizzata dei campi elettromagnetici). Nella loro forma globale, le equazioni di Maxwell possono essere scritte come

$$\begin{aligned} \oint_{\partial V} \mathbf{D} \cdot d\mathbf{S} - \iiint_V \rho dV &= 0, & \oint_{\partial S} \mathbf{E} \cdot d\mathbf{l} &= -\frac{\partial}{\partial t} \iint_S \mathbf{B} \cdot d\mathbf{S} \\ \oint_{\partial V} \mathbf{B} \cdot d\mathbf{S} &= 0, & \frac{\partial}{\partial t} \iint_S \mathbf{D} \cdot d\mathbf{S} - \oint_{\partial S} \mathbf{H} \cdot d\mathbf{l} + \iint_S \mathbf{J} \cdot d\mathbf{S} &= 0 \end{aligned}$$

Dove S è una superficie chiusa, ∂S il suo contorno, V un volume e ∂V la superficie che lo racchiude. Gli integrali su ∂S e V definiscono, quindi, il flusso delle grandezze integrate, l'integrale di linea su ∂S definisce una circuitazione, mentre l'integrale su V è un integrale di volume.

Pubblicate per la prima volta nel 1865 nel testo "A Dynamical Theory of the Electromagnetic Field" e riformulate nella forma a noi più nota nei primi anni del 1900 dal fisico britannico Oliver Heaviside, queste equazioni descrivono la dinamica evolutiva e i vincoli a cui è soggetto il campo elettromagnetico in relazione alle distribuzioni di carica e corrente elettrica da cui è generato. Di fatto, queste equazioni unificano i concetti già noti di elettricità e magnetismo e introducono il concetto, per l'epoca rivoluzionario, di campo elettromagnetico, ovvero dell'unica entità di cui il campo elettrico e quello magnetico sono semplicemente due manifestazioni differenti. L'importanza di questa rivoluzione scientifica, però, non si esaurisce nella semplice capacità sintetica di queste quattro equazioni. Al contrario, le equazioni di Maxwell fissano un punto di non ritorno per la fisica moderna e per il mondo scientifico in generale, aprendo la strada alla predizione, e alla successiva verifica sperimentale per mano di Heinrich Rudolf Hertz, dell'esistenza delle onde elettromagnetiche. Questa scoperta, a sua volta, ha dato il via a tutta una serie di invenzioni e scoperte che hanno cambiato profondamente la nostra società, dalla telegrafia senza fili fino alla teoria della relatività.

2 Dalle Superlenti all'Invisibilità: la Meta-Rivoluzione

Sebbene inizialmente fosse stata la loro assenza in natura a stimolare lo studio dei materiali a indice di rifrazione negativo, dopo la scoperta dei primi metamateriali l'attenzione del grande pubblico si è presto focalizzata sullo studio e sulla realizzazione di alcune delle loro potenzialità rivoluzionarie. In particolare, la possibilità di realizzare superlenti capaci di superare il limite di diffrazione della luce ha costituito sicuramente uno dei primi e più significativi esempi di sfida applicativa vinta dai metamateriali.

Nei sistemi ottici convenzionali non è possibile distinguere due punti se questi sono ad una distanza inferiore a $\lambda/2n$, dove n è l'indice di rifrazione del mezzo circostante. È questo il cosiddetto

limite di diffrazione. Questo limite intrinseco è dovuto al fatto che l'informazione riguardante strutture di dimensioni paragonabili alla lunghezza d'onda della luce, si propaga otticamente sotto forma di onde evanescenti, ovvero onde elettromagnetiche che decadono esponenzialmente man mano che si propagano. Di conseguenza, l'immagine di questi oggetti viene persa prima di poter essere raccolta e ricreata tramite una lente o un obiettivo. Fu ancora una volta John Pendry a proporre l'utilizzo di un metamateriale ad indice di rifrazione negativo come superlente (o lente perfetta), capace, cioè, di ricreare l'immagine di un oggetto senza perdere informazione sulla sua struttura, anche al di sotto del limite di diffrazione. Questa proprietà deriva dal fatto che, in un materiale ad indice di rifrazione negativo, le onde evanescenti non sono

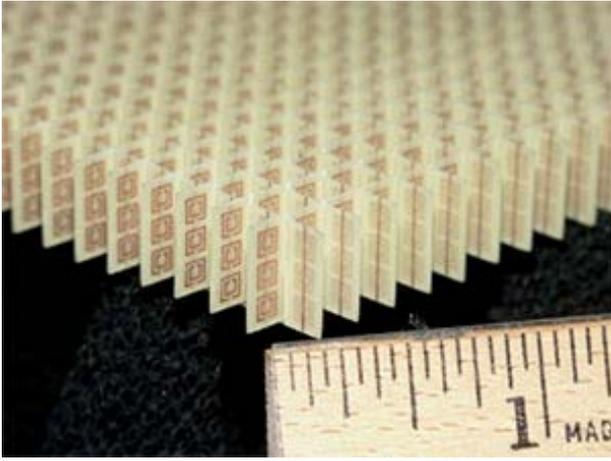


Figura 3: Il primo esempio di metamateriale proposto da John Pendry, costituito da un array periodico di split ring resonators e fili di rame. La dimensione della cella elementare è di 5 mm e il sistema lavora nella regione delle microonde. La frequenza delle microonde è dell'ordine di 10 GHz, che corrisponde ad una lunghezza d'onda nel vuoto di 3 cm: è, quindi, verificata la condizione di avere dimensioni delle celle elementari inferiori a quelle della lunghezza d'onda della luce.

più esponenzialmente decrescenti ma, al contrario, vengono amplificate. Allo stesso tempo, le onde propaganti vengono focalizzate automaticamente grazie all'indice di rifrazione negativo, come indicato schematicamente in Figura 4. Come risultato, una superlente riesce a focalizzare le onde propaganti e quelle evanescenti in un unico punto, superando il limite di diffrazione, come dimostrato sperimentalmente anche in regime ottico dai ricercatori dell'Università della California [5].

Un ulteriore campo di applicazione per i metamateriali, altamente affascinante sia intrinsecamente che dal punto di vista industriale, è la possibilità reale di rendere un oggetto invisibile alla radiazione elettromagnetica (il cosiddetto cloaking), proprio come il più famoso mantello di Harry Potter in Figura 5. Per comprendere questo incredibile fenomeno, partiamo da una domanda: come possiamo rendere invisibile un oggetto? Le risposte possono essere due: possiamo far passare la luce attraverso l'oggetto, oppure possiamo curvare la luce intorno all'oggetto. La prima è la proprietà dei comuni materiali trasparenti che però, come ben sappiamo, non possono rendere trasparenti oggetti terzi. Nel secondo caso, invece, i metamateriali possono

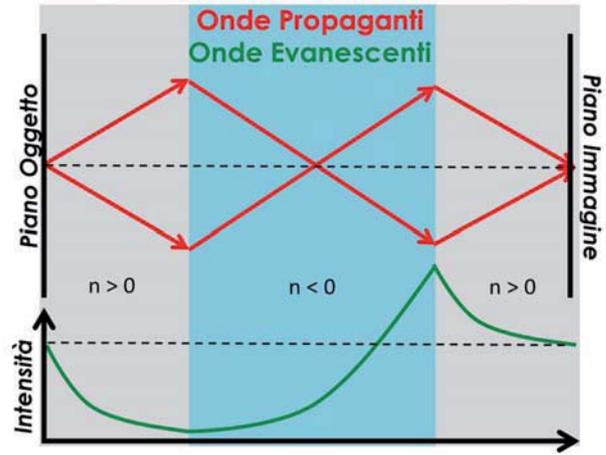


Figura 4: Schema di funzionamento di una superlente. Le onde propaganti vengono focalizzate dall'indice di rifrazione negativo, mentre quelle evanescenti vengono amplificate.

essere la chiave per l'invisibilità.



Figura 5: Harry Potter nel primo capitolo della serie cinematografica, insieme al più celebre esempio di "materiale" capace di rendere invisibili gli oggetti che avvolge.

Nell'ottica, il principio di Fermat afferma che la luce viaggia tra due punti seguendo il percorso che minimizza il cammino ottico. La stessa legge di Snell, così come la legge della riflessione, sono dirette conseguenze di questo principio. Consideriamo, quindi, il percorso della luce indicato in Figura 6a [6], in cui i raggi di luce girano attorno ad una regione vietata centrale di raggio R_1 , vengono schiacciati nella corona circolare di raggi R_1 ed R_2 e restano, infine, invariati altrove. Se riuscissimo a creare un materiale con un indice di rifrazione tale da ricreare questa distorsione spaziale, qualsiasi cosa all'interno della sfera centrale proibita risulterebbe totalmente invisibile, a causa dell'impossibilità di interagire con la luce circostante. Per fare questo è possibile sfrutta-

re la cosiddetta ottica trasformazionale [7] che permette di calcolare, data una certa distorsione dello spazio, la dispersione spaziale delle permittività elettrica e magnetica che la ricreano. Un così ampio e specifico controllo delle proprietà ottiche dei materiali si può ottenere con i metamateriali, il cui funzionamento come “mantelli dell’invisibilità” è stato provato dal solito John Pendry e da David Smith nel 2006 [8] a frequenze nel regime delle microonde, utilizzando dieci layer cilindrici di split ring resonators di forma variabile, come indicato in Figura 6b.

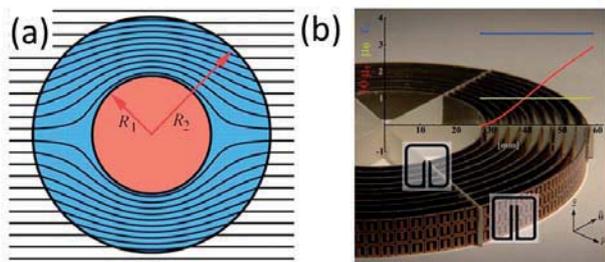


Figura 6: (a) Distorsione dello spazio per ottenere invisibilità degli oggetti: i percorsi di minimo cammino ottico della luce non passano attraverso il cerchio centrale di raggio R_1 , all’interno del quale ogni cosa sarà invisibile, ma vengono deformati e deviati nella corona circolare di raggi R_1 ed R_2 . (b) Prima dimostrazione di metamateriali per invisibilità nel regime delle microonde utilizzando dieci cilindri di split ring resonators.

3 La Terza Dimensione dei Metamateriali: la Chiralità

Una classe emergente di metamateriali è rappresentata dai metamateriali chirali. Lord Kelvin definì chirale un oggetto che non può essere sovrapposto alla sua immagine speculare, anche detta enantiomero. La chiralità è uno dei fenomeni più affascinanti che si verifica in natura e gioca un ruolo importante nel campo della chimica, della farmacologia e della biologia. In natura è presente una vasta gamma di sistemi dotati di chiralità, come gli amminoacidi, il DNA o le nostre stesse mani (Figura 7). Si tratta di una caratteristica fondamentale della materia vivente e un requisito talvolta necessario per la stessa vita: se un giorno l’ingegneria molecolare dovesse clonare i meccanismi autoassemblanti della

vita, certamente la chiralità svolgerebbe un ruolo cruciale.

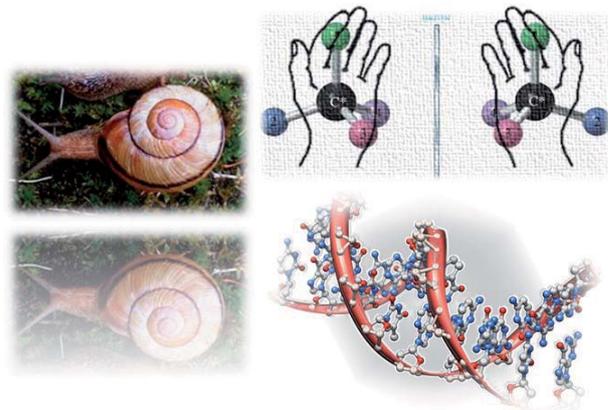


Figura 7: Le strutture chirali sono largamente diffuse in natura: il guscio delle lumache e la sua forma a spirale, le mani, la doppia elica del DNA.

Generalmente, i materiali chirali presentano una differente risposta ottica quando interagiscono con luce avente polarizzazione circolare destra e sinistra (RCP ed LCP rispettivamente, come indicato in Figura 7). Inoltre l’accoppiamento incrociato tra il campo elettrico e il campo magnetico all’interno di questi materiali, inteso come l’induzione di un campo magnetico attraverso il campo elettrico della luce incidente e viceversa, è identificato dal cosiddetto parametro di chiralità κ , determina una sensibile variazione degli indici di rifrazione relativi a queste due polarizzazioni. Nel 2004, John Pendry osservò come fosse possibile ottenere un indice di rifrazione negativo in mezzi chirali nanostrutturati [9], dimostrando, inoltre, come fosse più semplice entrare in questo regime rispetto ai metamateriali classici. Questi ultimi, infatti, devono essere ingegnerizzati in modo tale che sia la permittività elettrica che la permeabilità magnetica risultino simultaneamente negativi. Nei materiali chirali invece, quando il parametro di chiralità κ è sufficientemente grande, è possibile raggiungere rifrazione negativa per una delle due polarizzazioni circolari, senza necessariamente che ϵ e μ siano negativi. In questo caso si può parlare di metamateriali chirali.

Per capire come questo sia possibile, analizziamo in dettaglio cosa accade quando luce propagante nel vuoto interagisce con mezzi chi-

I Mezzi Bi-Isotropi

In generale, i mezzi chirali appartengono alla più ampia classe di mezzi bi-isotropi, per i quali le relazioni costitutive dei campi elettrico e magnetico si possono scrivere:

$$\begin{cases} \mathbf{D} = \varepsilon_0 \varepsilon_r \mathbf{E} + (\chi + i\kappa) \sqrt{\mu_0 \varepsilon_0} \mathbf{H} \\ \mathbf{B} = \mu_0 \mu_r \mathbf{H} + (\chi - i\kappa) \sqrt{\mu_0 \varepsilon_0} \mathbf{E} \end{cases}$$

I mezzi bi-isotropi, quindi, si differenziano dai mezzi isotropi standard per la presenza di due termini extra in queste relazioni. Il termine χ identifica la reciprocità del materiale ($\chi \neq 0$ indica un mezzo non reciproco, ovvero in cui i tensori ε e μ non sono simmetrici), mentre κ è il parametro della chiralità (materiali con $\kappa \neq 0$ sono detti chirali). In funzione dei parametri χ e κ , i materiali bi-isotropi possono essere identificati come

- 1) materiali bi-isotropi generali ($\chi \neq 0, \kappa \neq 0$)
- 2) materiali di Tellegen ($\chi \neq 0, \kappa = 0$)
- 3) materiali di Pasteur ($\chi = 0, \kappa \neq 0$)

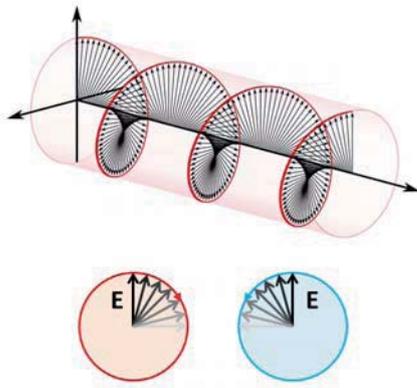


Figura 8: Luce con polarizzazione circolare destra e sinistra.

rali. Consideriamo dapprima un'onda piana propagante in un mezzo reciproco, isotropo e chirale. In queste condizioni, le equazioni di Maxwell possono essere combinate, portando all'equazione d'onda per il campo elettrico \mathbf{E} :

$$\mathbf{k} \times (\mathbf{k} \times \mathbf{E}) = -k_0^2 (\varepsilon \mu - \kappa^2) \mathbf{E} - 2i\kappa k_0 (\mathbf{k} \times \mathbf{E})$$

dove \mathbf{k} è il vettore d'onda nel mezzo, k_0 il vettore d'onda nel vuoto e κ la chiralità. Sostituendo la classica forma d'onda per il campo elettrico, e fissando la direzione di propagazione $\mathbf{k} = k\hat{z}$, si ottiene la soluzione per il vettore d'onda

$$k_{\pm} = k_0 (n \pm \kappa)$$

con n l'indice di rifrazione del mezzo senza chiralità e k_{\pm} i vettori d'onda per la polarizzazione circolare destra (+) e sinistra (-). Da questo risultato, si ricava l'espressione per il parametro di chiralità κ

$$\kappa = \frac{k_+ - k_-}{2k_0}$$

Scrivendo gli autovettori per il campo elettrico nella forma generica $E(r) = (E_{0x}\hat{x} + E_{0y}\hat{y}) e^{ikz}$, possiamo giungere alle relazioni

$$\frac{E_{0y}}{E_{0x}} = \frac{k_0^2 (n^2 - k^2) - k_{\pm}^2}{2ik_0 \kappa k_{\pm}} = \pm i$$

Introducendo, infine, gli indici di rifrazione n_{\pm} per le due polarizzazioni circolari tramite la relazione $k_{\pm} = n_{\pm} k_0$, si ottiene la relazione caratterizzante

$$n_{\pm} = n \pm \kappa$$

Di conseguenza la condizione necessaria per raggiungere la rifrazione negativa in mezzi chirali per una delle due polarizzazioni circolari, è che il parametro di chiralità sia sufficientemente grande rispetto all'indice di rifrazione del mezzo non chirale. In particolare, come già anticipato, non è richiesto che ε e/o μ assumano simultaneamente valori negativi.

L'esistenza di due indici di rifrazione distinti, determina l'esistenza di proprietà chiro-ottiche caratteristiche per questi materiali. Ad esempio,

considerando un'onda incidente polarizzata linearmente come la combinazione di due componenti con polarizzazione circolare destrorsa e sinistrorsa, il suo piano di polarizzazione risulterà ruotato quando attraversa un mezzo chirale, a causa dei due diversi indici di rifrazione associati a RCP ed LCP. Questo fenomeno è noto come rotazione ottica (Figura 9) e l'angolo di rotazione del piano di polarizzazione della luce incidente è espresso da $\theta = \arg(T_{++}) - \arg(T_{--})$, dove T_{++} e T_{--} sono i coefficienti complessi di trasmissione della luce RCP ed LCP rispettivamente. Poiché stiamo considerando un mezzo chirale, inoltre, le due polarizzazioni circolari incidenti vengono assorbite in maniera differente mentre propagano al suo interno. Questo fenomeno è noto come dicroismo circolare ed è definito dalla differenza normalizzata delle trasmissioni delle due polarizzazioni:

$$\eta = \frac{1}{2} \sin^{-1} \left(\frac{|T_{++}|^2 - |T_{--}|^2}{|T_{++}|^2 + |T_{--}|^2} \right)$$

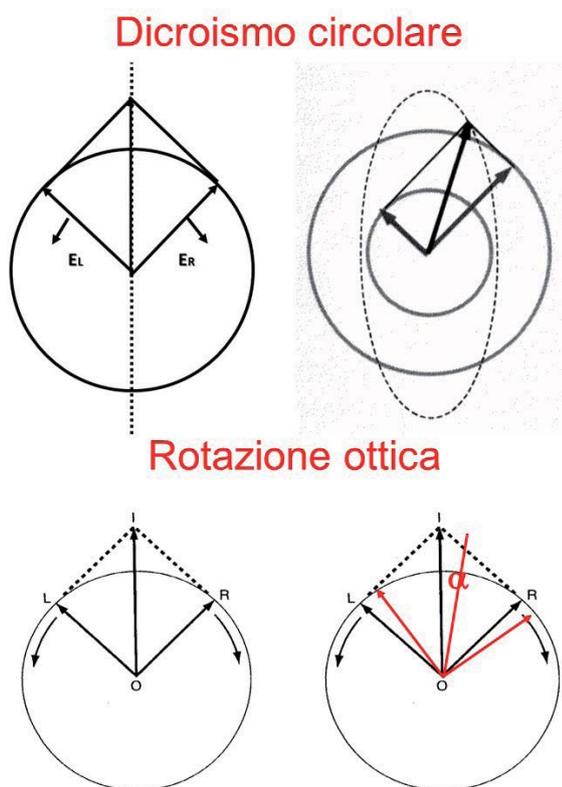


Figura 9: Rappresentazione grafica del dicroismo circolare e della rotazione ottica.

4 Metamateriali Chirali e Nanotecnologie

I sistemi chirali esistenti in natura presentano quasi sempre un momento di dipolo piuttosto piccolo, si accoppiano, quindi, debolmente con il campo elettromagnetico e le proprietà ottiche chirali ne risultano fortemente ridotte. Per superare queste difficoltà naturali, limitazioni, si può ricorrere ai più avanzati metodi di nanofabbricazione per ridurre le dimensioni dei materiali bulk fino alla nanoscala, dove l'interazione tra nanostrutture e luce, intensificata dall'elevato rapporto superficie-volume e influenzata dalla riduzione del numero di gradi di libertà, dà origine ad esclusive ed affascinanti proprietà ottiche. Un esempio del genere è costituito dai plasmoni localizzati, oscillazioni del gas di elettroni semi-liberi dei materiali metallici che, quando il metallo stesso viene nanostrutturato a dimensioni dell'ordine di decine o poche centinaia di nanometri, generano un intenso momento di dipolo e determinano un forte accoppiamento con il campo elettromagnetico.

Queste strutture plasmoniche sono state ingegnerizzate in particolari geometrie bidimensionali al fine di ottenere simultaneamente proprietà chirali, elevato momento dipolare e un'elevata efficienza di accoppiamento con la luce incidente [9, 10]. Sono stati realizzati in questo modo, per la prima volta, metamateriali plasmonici chirali 2D. D'altra parte, però, la propagazione del campo elettromagnetico avviene con una torsione dei vettori di campo lungo la direzione di propagazione. Ne consegue che, per massimizzare l'interazione con la luce, i metamateriali chirali dovrebbero estrudere nella direzione di propagazione della radiazione elettromagnetica. Per questo motivo la ricerca si è concentrata, inizialmente, su strutture denominate "quasi-3D", cioè composte da una ripetizione periodica della cella fondamentale del metamateriale in piani perpendicolari alla direzione di propagazione, separati e ruotati tra loro per ottenere forte interazione chirale con la luce [11, 12, 13, 14, 15, 16, 17, 18].

Con l'avanzamento di tecniche di nanofabbricazione sempre più complesse, la massima espressione della chiralità ha, infine, preso forma in strutture monolitiche intrinsecamente tridimensionali e chirali, come l'elica, che rappre-

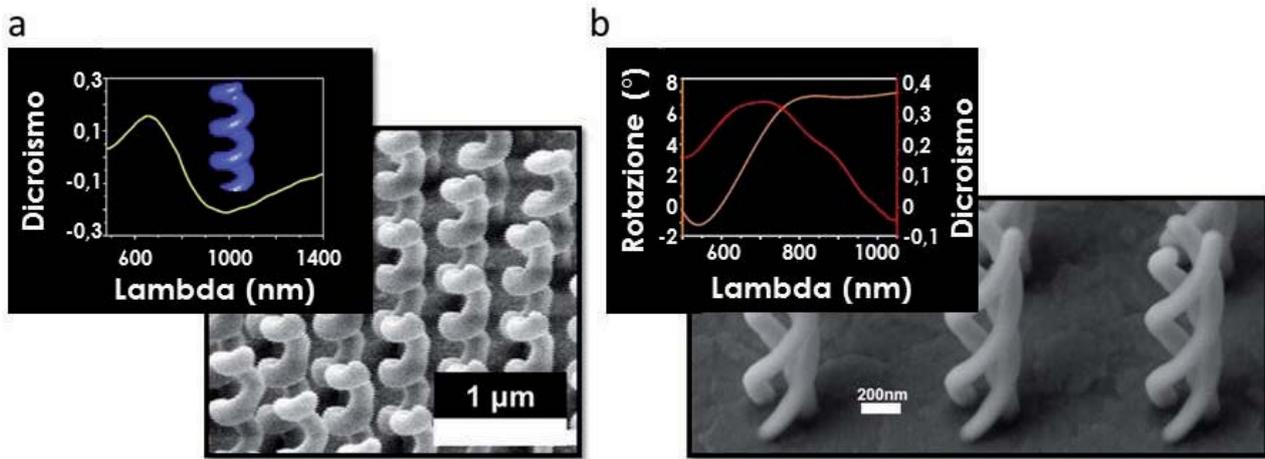


Figura 10: (a) Dicroismo circolare per eliche di platino a singolo filo con 3 giri. (b) Dicroismo circolare e rotazione ottica di eliche triple di platino.

senta anche la migliore struttura geometrica candidata a manipolare lo stato di polarizzazione della luce. Nel 2009 un gruppo di ricercatori tedeschi ha, per la prima volta, realizzato e dimostrato le potenzialità di strutture elicoidali tridimensionali di oro operanti nel regime dell'infrarosso come polarizzatori circolari [19]. Per spostarne il range di lavoro verso lo spettro visibile, strutture simili sono state studiate e scalate a dimensioni circa dieci volte più piccole attraverso complesse tecniche di nanofabbricazione basate su deposizione indotta da fascio ionico (FIBID) o elettronico (FEBID) focalizzato [20, 21, 22]. Queste tecnologie sfruttano un precursore metallo-organico che viene dissociato da un fascio elettronico o ionico che rende volatile la parte organica mentre le particelle metalliche vengono ionizzate e accelerate verso il substrato. Attraverso il controllo dei parametri del fascio incidente (corrente e tensione di accelerazione) e di quelli di pattern e crescita (pressione locale del precursore, temperatura, ecc.) è possibile realizzare, con una precisa localizzazione spaziale, complesse strutture tridimensionali, di qualsiasi forma e dimensione. Queste strutture hanno evidenziato una forte attività ottica nella regione del visibile e del vicino infrarosso, esibendo un elevato dicroismo circolare in un'ampia banda di lavoro così come una forte rotazione ottica rispetto alle strutture planari e quasi 3D (Figura 10a).

Uno dei principali fattori limitanti delle strutture elicoidali è, però, costituito dalla perdita di

simmetria rotazionale dovuta alla punta dell'elica stessa che, insieme al suo asse, introduce birifrangenza lineare e alta sensibilità della rotazione ottica all'orientazione del campione (anisotropia in piano). Questo inconveniente può essere cruciale nell'aprire campi applicativi per questo tipo di dispositivi. Recentemente il nostro gruppo di ricerca ha dimostrato che la simmetria rotazionale può essere efficacemente ripristinata attraverso la disposizione di multieliche concentriche [23]. Eliche triple funzionanti nella regione del visibile sono state realizzate per la prima volta attraverso la combinazione del concetto di tomografia con la tecnologia FIBID (Figura 10b), ottenendo qualcosa di simile ad una stampante 3D sensibile al nanometro [24]. La chiralità intrinseca e la recuperata simmetria rotazionale, insieme alle forti interazioni plasmoniche tra le tre nanoeliche attorcigliate, hanno permesso di migliorare praticamente tutte le figure di merito di questo tipo di oggetti. Elevatissimi valori di dicroismo circolare sono stati ottenuti in una banda dicroica che copre quasi l'intera regione dello spettro visibile e arriva nel vicino infrarosso. La trasmissione delle polarizzazioni circolari ha raggiunto un elevato grado di "purezza", ovvero sono ridotte al minimo le contaminazioni indotte dalla polarizzazione opposta. La forte simmetria rotazionale, inoltre, ha determinato sia lo sviluppo di una forte rotazione ottica (Figura 10b) in ampia banda di funzionamento che l'azzeramento dell'anisotropia in piano, aprendo la strada, ad esempio, verso efficaci polarizzatori

circolari ad ampio spettro.

5 Il Futuro dei Metamateriali

La ricerca nel campo dei metamateriali, dopo appena venti anni, si può dire essere ancora molto giovane. Diversi sono gli ambiti indagati dai ricercatori di tutto il mondo: nuovi materiali o combinazioni di questi, nuove geometrie, nuove ed elaborate tecniche di fabbricazione. Proprio questo aspetto è, allo stato attuale, il punto sensibile per questo tipo di tecnologia, dato che la sua reale applicabilità nella vita quotidiana passerà necessariamente attraverso la facilità, la ripetibilità e la scalabilità dei processi di fabbricazione. Diverse sono le tecniche finora utilizzate, dagli approcci top-down come la litografia da fascio elettronico e ionico e la scrittura via laser, fino a quelli bottom-up, come il self-assembly su substrati nanostrutturati. Si tratta di processi più o meno complessi, le cui possibilità di allargamento a scala industriale verranno certamente indagate sulla base delle applicazioni richieste.

Dal punto di vista dei fenomeni fisici indagabili nei metamateriali, molto interessanti sembrano essere le caratteristiche previste per sistemi accoppiati ad elementi non lineari, che vanno dall'amplificazione degli effetti non lineari dovuti alla concentrazione del campo elettromagnetico, fino a fenomeni esotici che comprendono transizioni isteretiche e propagazione di solitoni [25]. Applicativamente parlando, invece, sono state dimostrate potenzialità sia nel campo dei dispositivi di memoria che nella realizzazione di metamateriali dinamici, il cui comportamento cambia a seconda di stimoli esterni.

Appena agli albori, ma non meno interessante, sembra essere, infine, il campo dei metamateriali quantistici. In principio, infatti, il concetto di metamateriale può essere applicato a qualsiasi tipo di onda e su qualsiasi tipo di scala. Non fa eccezione, quindi, la materia e la descrizione ondulatoria delle particelle, come elettroni e neutroni. Ricercatori del Regno Unito hanno dimostrato, teoricamente, che in un layer di grafene si può osservare rifrazione negativa e focalizzazione degli elettroni, mentre un gruppo dell'università della California ha recentemente progettato un mantello dell'invisibilità per onde di materia, mediante un reticolo tridimensionale di fasci laser.

6 Conclusioni

L'universo dei metamateriali è tutt'oggi in continuo divenire. Le prospettive, le applicazioni, gli esperimenti si arricchiscono costantemente sulla base delle geometrie proposte dai fisici teorici e di quelle investigate dagli sperimentali. La capacità di plasmare e modificare le proprietà di interazione dei materiali con la radiazione elettromagnetica, costituiranno la base per lo sviluppo di tecnologie completamente nuove che si apprestano a rivoluzionare la società in cui viviamo. In particolare, i metamateriali chirali sono ottimi candidati per applicazioni di biofotonica, sensoristica e come componenti passivi e attivi nella progettazione di circuiti nanofotonici. Da questo punto di vista lo studio di fenomeni non lineari, l'accoppiamento con elementi attivi (emettitori) e lo studio di forme, dimensioni e composizione delle nanoeliche è di fondamentale importanza per permettere alla ricerca in questo campo di fare un passo in avanti verso una tecnologia matura e disponibile.

Dati i progressi delle tecnologie di fabbricazione e caratterizzazione, l'unico limite sembra essere la nostra capacità di immaginare nuove geometrie, nuovi fenomeni, nuovi metamateriali.

Benvenuti nella meta-rivoluzione.



- [1] V. G. VESELAGO: "The Electrodynamics of Substances with Simultaneously Negative Values of ϵ and μ ", *Soviet Physics Uspekhi* **4** (1968) 509–514.
- [2] V. V. Varadan, R. Ro "Analyticity, causality, energy conservation and the sign of the imaginary part of the permittivity and permeability" *IEEE Xplore Abstract Record*:C//Users//User//AppData//Roaming//Zotero//Zotero//Profiles 0wn608a7.default//zotero//storage//ZARPH4HQ//login.html:text/html 2006499–502
- [3] J. B. PENDRY, A. J. HOLDEN, W. J. STEWART, I. YOUNGS: "Extremely Low Frequency Plasmons in Metallic Mesostructures", *Phys. Rev. Lett.* **25** (1996) 4773–4776.
- [4] J. B. PENDRY, A. J. HOLDEN, D. J. ROBBINS, W. J. STEWART: "Magnetism from conductors and enhanced nonlinear phenomena", *IEEE Transactions on Microwave Theory and Techniques* **11** (1999) 2075–2084.
- [5] N. FANG, H. LEE, C. SUN, X. ZHANG: "Sub-Diffraction-Limited Optical Imaging with a Silver Superlens", *Science* **312** (2005) 534–537.
- [6] J. B. PENDRY, D. SCHURIG, D. R. SMITH: "Controlling Electromagnetic Fields", *Science* **5781** (2006) 1780–1782.

- [7] Y. LIU, X. ZHANG: "Metamaterials: a new frontier of science and technology", *Chem. Soc. Rev.* **2011** (5) 2494–2507.
- [8] D. SCHURIG *et al.*: "Metamaterial Electromagnetic Cloak at Microwave Frequencies", *Science* **314** (2006) 977–980.
- [9] J. B. PENDRY: "A Chiral Route to Negative Refraction", *Science* **306** (2004) 1353–1355.
- [10] M. KUWATA-GONOKAMI *et al.*: "Giant Optical Activity in Quasi-Two-Dimensional Planar Nanostructures", *Phys. Rev. Lett.* **95** (2005) 227401.
- [11] F. EFTEKHARI, T. J. DAVIS: "Strong chiral optical response from planar arrays of subwavelength metallic structures supporting surface plasmon resonances", *Phys. Rev. B* **86** (2012) 075428.
- [12] X. YIN, M. SCHÄFERLING, B. METZGER, H. GIESSEN: "Interpreting Chiral Nanophotonic Spectra: The Plasmonic Born–Kuhn Model", *Nano Lett.* **13** (2013) 6238–6243.
- [13] E. PLUM, V. A. FEDOTOV, A. S. SCHWANECKE, N. I. ZHELUDEV, Y. CHEN: "Giant optical gyrotropy due to electromagnetic coupling", *Applied Physics Letters* **90** (2007) 223113.
- [14] N. LIU, H. LIU, S. ZHU, H. GIESSEN: "Stereometamaterials", *Nat Photon* **3** (2009) 157–162.
- [15] C. HELGERT *et al.*: "Chiral Metamaterial Composed of Three-Dimensional Plasmonic Nanostructures", *Nano Lett.* **11** (2011) 4400–4404.
- [16] A. CHRISTOFI, N. STEFANOY, G. GANTZOUNIS, N. PAPANIKOLAOU: "Giant Optical Activity of Helical Architectures of Plasmonic Nanorods", *J. Phys. Chem. C* **116** (2012) 16674–16679.
- [17] M. DECKER *et al.*: "Strong optical activity from twisted-cross photonic metamaterials", *Optics Letters* **34** (2009) 2501.
- [18] J. ZHAO *et al.*: "Localized surface plasmon resonance biosensors", *Nanomedicine* **1** (2006) 219–228.
- [19] J. GANSEL *et al.*: "Gold Helix Photonic Metamaterial as Broadband Circular Polarizer", *Science* **235** (2009) 1513–1515.
- [20] M. ESPOSITO *et al.*: "Three Dimensional Chiral Metamaterial Nanospirals in the Visible Range by Vertically Compensated Focused Ion Beam Induced-Deposition", *Advanced Optical Materials* **2** (2014) 154–161.
- [21] M. ESPOSITO *et al.*: "Nanoscale 3D Chiral Plasmonic Helices with Circular Dichroism at Visible Frequencies", *ACS Photonics* **2** (2015) 105–114.
- [22] M. ESPOSITO *et al.*: "Tailoring chiro-optical effects by helical nanowire arrangement", *Nanoscale* **7** (2015) 18081–18088.
- [23] J. KASCHKE, J. K. GANSEL, M. WEGENER: "On metamaterial circular polarizers based on metal N-helices", *Optics Express* **20** (2012) 26012.
- [24] M. ESPOSITO *et al.*: "Triple-helical nanowires by tomographic rotatory growth for chiral photonics", *Nat. Commun.* **6** (2015) 6484.
- [25] A. A. ZHAROV, I. SHADRIVOV, Y. S. KIVSHAR: "Nonlinear properties of left-handed metamaterials", *Physical Review Letters* **91** (2003) 037401.

Francesco Todisco: Svolge attività di ricerca sperimentale nel gruppo Advanced Photonics del CNR Nanotec di Lecce, dove si occupa dello studio ottico nel campo della plasmonica e dei metamateriali.

Marco Esposito: Svolge attività di ricerca sperimentale nella facility di Nanofabbricazione del CNR Nanotec di Lecce, dove si occupa dello sviluppo e studio di materiali nanostrutturati, nanostrutture plasmoniche e metamateriali chirali.

Strong Coupling: l'entanglement tra luce e materia

Armando Genco

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

Le interazioni luce-materia sono alla base di un numero elevatissimo di eventi della nostra realtà quotidiana. Basti pensare che nell'occhio umano le cellule foto-ricettrici interagiscono continuamente con la radiazione elettro-magnetica tramite processi di assorbimento. Il progresso scientifico e tecnologico ha permesso di far interagire radiazione e materia anche in modi più inusuali, per esempio intrappolando la luce: posizionando due specchi ad alta riflettività l'uno di fronte a l'altro e distanziandoli di poche centinaia di nanometri otteniamo una microcavità in cui le onde elettromagnetiche risuonano e rimangono confinate a causa di fenomeni di interferenza. Tra le tante applicazioni, questo effetto può essere sfruttato per far interagire fortemente la luce con atomi e molecole che possono assorbire e rimettere fotoni tramite processi di transizione degli elettroni tra il livello energetico fondamentale e quelli eccitati. Se le energie in gioco per queste transizioni sono vicine

a quelle dei fotoni intrappolati nelle microcavità i due elementi si possono unire a formare nuovi stati ibridi detti polaritoni (Figura 1). In questo regime di interazione di luce e materia si aprono nuove ed entusiasmanti frontiere per lo studio dei fenomeni di elettrodinamica quantistica e nello sviluppo di dispositivi per la computazione quantistica.

Microcavità

Per poter illustrare teoria ed applicazioni dei polaritoni è utile descrivere brevemente i principi alla base del funzionamento dei dispositivi di confinamento ottico, che permettono ai fotoni di rimanere localizzati per un tempo sufficientemente lungo da entrare in regime di accoppiamento forte con la materia. Come le onde acustiche prodotte da una corda di chitarra risuonano nella cassa armonica, le onde elettromagnetiche subiscono fenomeni di risonanza nelle microcavità ottiche. In questi sistemi la risonanza è causata dall'interferenza costruttiva della radiazione che subisce riflessioni multiple sugli specchi del risonatore (Figura 2), distanziati di un multiplo della lunghezza d'onda della luce da confinare:

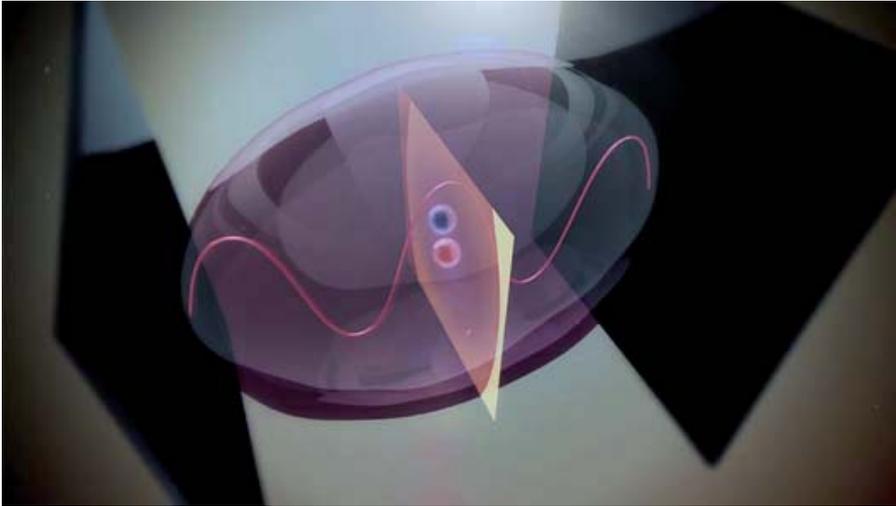


Figura 1: *Illustrazione concettuale di un sistema microcavità-transizione elettronica in accoppiamento forte: la microcavità è formata da due specchi contrapposti che confinano la radiazione elettromagnetica con cui si accoppiano le transizioni di eccitazione ed emissione di atomi, molecole o cristalli semiconduttori.*

nel caso della radiazione visibile siamo nell'ordine delle centinaia di nanometri da cui il prefisso micro. Gli specchi tipicamente utilizzati si dividono in due tipologie a seconda del tipo riflessione: nella prima viene sfruttata la riflessione in corrispondenza di una singola interfaccia ottica, per esempio una superficie metallica o un'interfaccia tra due materiali dielettrici. Il secondo approccio sfrutta le riflessioni parziali in strutture con numerose interfacce ottiche composte da materiali dielettrici alternati: gli spessori dei diversi materiali seguono una condizione periodica di tipo Bragg in modo che il cammino ottico (cioè il prodotto nd dell'indice di rifrazione n e lo spessore d dello strato) sia un quarto della lunghezza d'onda λ . Le porzioni di radiazione riflessa su ogni interfaccia sono in fase tra di loro e interferiscono costruttivamente amplificandosi e ottenendo così valori molto alti di riflettività. Se la periodicità viene realizzata lungo una sola dimensione utilizzando strati alternati con diverso indice di rifrazione si parla di specchi DBR (Distributed Bragg Reflector), mentre se è presente in due o tre dimensioni la struttura prende il nome di cristallo fotonico (Photonic Crystal). Gli specchi DBR saranno approfonditi nel prossimo paragrafo in cui illustreremo come dal punto di vista classico possiamo descrivere una microcavità planare in cui il confinamento avviene in una sola dimensione.

Visione classica: cavità di Fabry-Perot

Il sistema di confinamento ottico che permette la maggiore comprensione dei fenomeni in gioco è la microcavità planare costituita da due strati

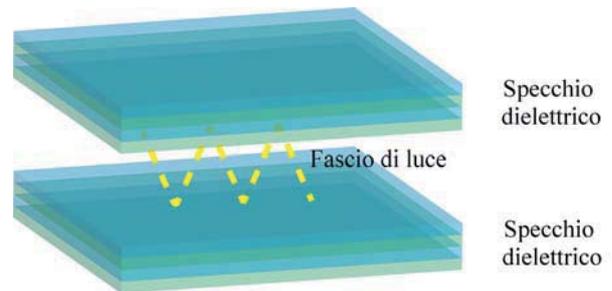


Figura 2: *Schema concettuale di una microcavità planare con specchi dielettrici: le onde elettromagnetiche rimangono confinate per effetto di interferenze costruttive all'interno della struttura ottica.*

di specchi ideali posti ad una distanza L (cavità di Fabry-Perot). Considerando degli specchi ideali si ha che il campo elettrico è nullo in corrispondenza delle due interfacce tra gli specchi e il mezzo che riempie la cavità (Figura 1): questa assunzione equivale ad imporre la condizione di stazionarietà per un'onda elettromagnetica nella direzione perpendicolare al piano degli specchi (direzione z). Le onde stazionarie possibili saranno quelle per cui lo sfasamento dell'onda in un percorso completo tra i due specchi sia un multiplo intero di 2π da cui si ottiene che:

$$k_z = \frac{m\pi}{nL} = m\omega_0, \quad (1)$$

dove k_z è la componente del vettore d'onda lungo z , ω_0 è frequenza di risonanza della luce, n è l'indice di rifrazione del mezzo della cavità, L è la lunghezza della cavità, m è un numero intero che definisce l'ordine di risonanza. La presenza del fattore m fa sì che si creino nello stesso risonatore

diversi ordini di risonanza a frequenze multiple di quella corrispondente al primo ordine. Un altro effetto importante del confinamento è che in corrispondenza della frequenza di risonanza la trasmissione della luce attraverso la cavità è permessa mentre in un singolo specchio sarebbe stata riflessa: la sola possibilità di creazione di onde stazionarie tra gli specchi per interferenza costruttiva rende possibile alla luce di penetrare attraverso essi. Questa proprietà viene illustrata nel grafico in figura 3 che mostra uno spettro di trasmissione/riflessione di una microcavità in funzione della lunghezza d'onda. È utile ricordare che solo lungo z sussiste la condizione di stazionarietà: in generale infatti la frequenza di risonanza di una cavità di Fabry-Perot dipende da q , la componente del vettore d'onda sul piano x, y :

$$\omega_c(q) = \frac{c}{\sqrt{\varepsilon_b}} \sqrt{k_z^2 + q^2} = \frac{c}{\sqrt{\varepsilon_b}} \sqrt{\frac{(m\pi)^2}{L^2} + q^2}, \quad (2)$$

dove c è la velocità della luce ed ε_b è la costante dielettrica del materiale tra gli specchi. Vedremo in seguito come questa dipendenza angolare della frequenza di risonanza sarà determinante per alcune proprietà dei polaritoni. Esiste un parametro che fornisce una misura di quanto la luce viene confinata nelle strutture: il fattore di qualità (Q-factor). Esso è definito come il rapporto tra la frequenza di risonanza della cavità, ω_0 e la larghezza in frequenza a metà altezza del picco di risonanza (detta larghezza di banda del modo di cavità γ_c), definita anche come $\delta\lambda$ in un grafico dell'intensità di trasmissione o riflessione in funzione della lunghezza d'onda come quello in figura 3. Intuitivamente il fattore di qualità equivale al numero di rimbalzi che compie un fotone tra gli specchi prima di fuoriuscire dalla cavità.

A sua volta la larghezza di banda del modo di cavità è legata alla "velocità" con cui la radiazione all'interno della cavità viene dissipata, a causa di assorbimento, diffusione, o fuoriuscita di radiazione da specchi imperfetti. Il suo inverso $\tau = 1/\gamma_c$ è definito come il tempo medio in cui il fotone rimane intrappolato.

Negli studi sui polaritoni si utilizzano princi-

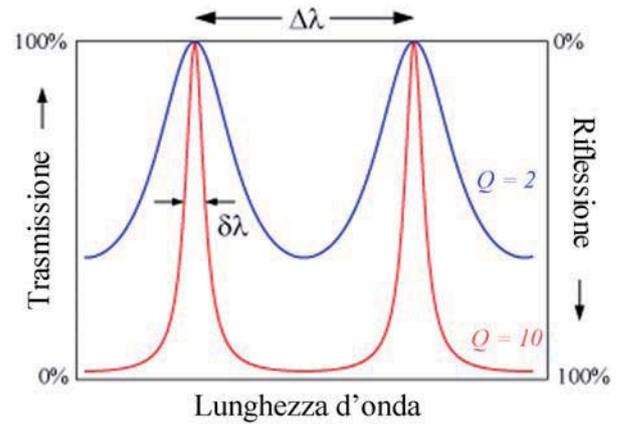


Figura 3: Distribuzione spettrale dei modi di risonanza di una microcavità. In ascisse troviamo la lunghezza d'onda, sull'asse verticale sinistro viene indicata l'intensità di trasmissione e sull'asse verticale destro l'intensità di riflessione che ha un andamento inverso alla trasmissione. La curva blu rappresenta una cavità con fattore di qualità pari a 2, mentre la curva rossa si riferisce a $Q=10$. Sul grafico viene indicata anche la larghezza di banda del modo di cavità ($\delta\lambda$) e il range spettrale libero ($\Delta\lambda$).

palmente le microcavità planari poiché, oltre a poter essere fabbricate più facilmente, presentano modi di risonanza che sono descrivibili da modelli teorici semplici ma al tempo stesso vicini alle condizioni ideali. Queste strutture sono quindi ottime per verificare la validità della teoria nella descrizione dei dati osservati. Descriviamo in maniera più approfondita gli specchi dielettrici dato che grazie ad essi si ottengono i fattori di qualità migliori per questa tipologia di cavità. La riflettività di uno specchio DBR dipende dal numero di coppie dei due materiali utilizzati e dal rapporto tra i loro indici di rifrazione.

$$R = 1 - 4 \frac{n_{ext}}{n_{cav}} \left(\frac{n_L}{n_H} \right)^{2N}, \quad (3)$$

dove N è il numero di coppie, n_{ext} è l'indice di rifrazione del mezzo esterno alla cavità, n_{cav} è l'indice di rifrazione del mezzo interno, n_L e n_H sono gli indici rispettivamente del materiale a basso e ad alto indice di rifrazione.

Dall'interferenza delle onde elettromagnetiche si crea una banda di lunghezze d'onda per cui la radiazione viene riflessa completamente (stop-band). La larghezza della stop-band è proporzionale alla differenza tra gli indici di rifrazione degli strati dielettrici.

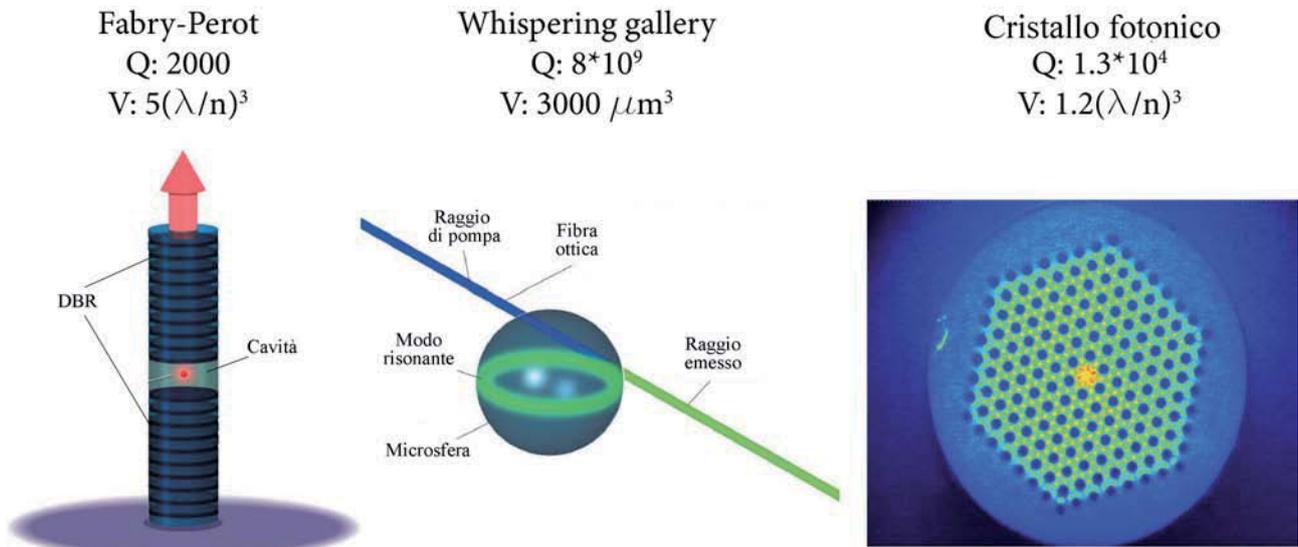


Figura 4: Esempi di strutture di confinamento ottico con i relativi fattori di qualità indicati dalla lettera Q e volumi modali (volume occupato dall'onda elettromagnetica confinata) definiti in μ^3 o in termini della dimensione del difetto ottico definita come lunghezza d'onda di risonanza fratto indice di rifrazione del mezzo in corrispondenza del difetto.

La presenza di un difetto lungo la periodicità degli strati, che può essere costituito da uno strato con spessore o materiale diverso da quelli del DBR, genera automaticamente una microcavità la cui frequenza di risonanza dipende dalle dimensioni e dal tipo di difetto: con la risonanza si crea l'apertura di una finestra di trasmissione nella stop-band degli specchi. Nella figura 5 viene mostrato lo spettro di intensità di luce riflessa dal DBR in cui è presente un difetto: si nota chiaramente la stop-band nella regione ad alta riflettività da circa 790 a 890 nm all'interno della quale si crea il modo di risonanza della cavità a circa 840 nm. Grazie alla alta riflettività dei DBR, il fattore di qualità di queste microcavità raggiunge valori elevati $Q \approx 2000$.

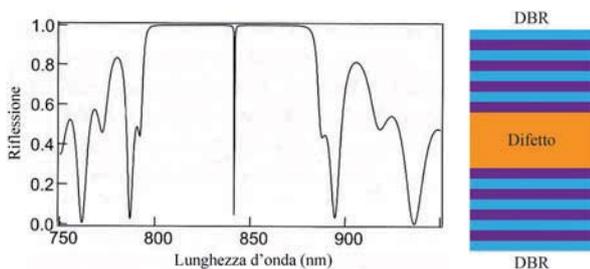


Figura 5: Intensità della luce riflessa (riflettanza) in funzione della lunghezza d'onda per una cavità costituita da un difetto nella periodicità di uno specchio DBR schematizzato nell'immagine accanto al grafico.

Estendendo la condizione di periodicità di

Bragg su più di una dimensione si ottengono le strutture ad alta riflettività 2D e 3D dei cristalli fotonici. Anche in questi sistemi una microcavità può essere fabbricata inserendo un difetto nel reticolo ottico. In figura 4 ad esempio viene mostrato un cristallo fotonico 2D in cui la periodicità ottica si genera scavando dei buchi con un pattern esagonale su un supporto dielettrico ad alto indice di rifrazione. Il contrasto di costanti dielettriche è dato dunque dal materiale del supporto e dall'aria. Si nota che il difetto nella periodicità ottica si ottiene lasciando al centro del cristallo fotonico una zona non forata in cui si creano le condizioni di risonanza e confinamento del campo elettromagnetico. Un altro modo per confinare la luce si ottiene fabbricando delle "whispering galleries" che sono strutture circolari di materiale dielettrico: essi sfruttano la riflessione totale interna della radiazione all'interfaccia tra aria e il corpo convesso (come una sfera o un disco) ad alto indice di rifrazione per fare risuonare la luce lungo la circonferenza esterna. Nella figura 4 vengono mostrati i valori dei fattori di qualità Q tipici delle strutture schematizzate dai disegni. Le cavità di Fabry-Perot sono caratterizzate da valori di Q più bassi rispetto ad altre tipologie di risonatori ottici poiché il confinamento avviene principalmente in una sola direzione.

Abbiamo trattato i principi di funzionamento di una microcavità e abbiamo considerato le di-

verse tipologie di risonatori ottici che possiamo utilizzare sperimentalmente approfondendo la descrizione delle microcavità planari di Fabry-Perot. A questo punto vedremo come è possibile trattare questi sistemi anche con un approccio puramente quantistico ottenendo una trattazione sorprendentemente semplice ed intuitiva della fisica dei fotoni confinati nelle microcavità.

Fotoni in scatola: visione quantistica

Il modello quantistico che descrive una microcavità sfrutta l'approccio della seconda quantizzazione del campo elettromagnetico.[2] Per la trattazione seguente è utile richiamare il concetto di oscillatore armonico quantistico: in questo sistema, tipicamente associato ad una particella confinata in un potenziale parabolico (se consideriamo ad esempio una molecola diatomica sarà dato dalla forza di repulsione tra due nuclei), i valori di energia che essa può assumere sono quantizzati come schematizzato in figura 6. C'è quindi una forte analogia con le onde stazionarie che si creano nelle microcavità solo in corrispondenza delle frequenze di risonanza. L'Hamiltoniana di un oscillatore armonico quantistico è in generale descritta dagli operatori posizione e quantità di moto ma si può semplificare introducendo degli operatori che chiameremo di creazione a^\dagger e distruzione a .

$$H_{osc} = \hbar\omega_0(a^\dagger a + \frac{1}{2}), \quad (4)$$

dove ω_0 è la frequenza di risonanza naturale dell'oscillatore. Gli autovalori dell'Hamiltoniana dell'oscillatore armonico saranno i livelli di energia discreti riportati in figura 6. Gli operatori di creazione e distruzione hanno un significato sottile poiché agiscono sullo stato del sistema rispettivamente incrementando o diminuendo di una unità il numero di quanti di oscillazione. L'operatore $a^\dagger a$ è chiamato operatore numero ed il suo valore di aspettazione è un numero intero e positivo e rappresenta proprio il numero di oscillazioni quantizzate presenti nel sistema.

Nel caso della quantizzazione della luce i quanti di oscillazione sono rappresentati dai fotoni. L'autostato $|N\rangle$ associato all'operatore numero è definito stato di Fock del fotone. Partendo dalle equazioni di Maxwell per la radiazione elettromagnetica che si propaga nello spazio libe-

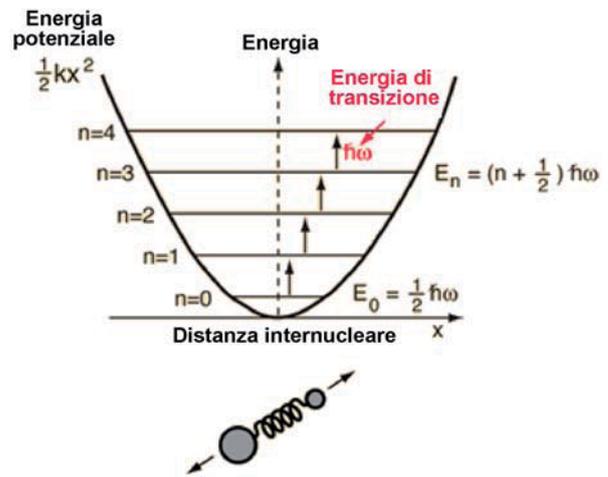


Figura 6: Esempio di un oscillatore armonico quantistico: una molecola diatomica vibra come un sistema di due masse accoppiate da una molla ed è caratterizzata da un'energia potenziale che dipende dal quadrato della distanza dalla posizione di equilibrio ($x=0$ nel grafico). Grazie al confinamento quantico i livelli energetici risultanti sono discreti e quantizzati.

ro ed effettuando una trasformazione di Fourier nello spazio reciproco si può scomporre il campo elettromagnetico come sommatoria di onde piane ciascuna con una specifica frequenza. In analogia con gli oscillatori armonici quantistici a ciascuna frequenza viene associato un numero intero di fotoni e i relativi operatori di creazione e distruzione. Nel caso della radiazione elettromagnetica quindi tali operatori rappresentano l'incremento o la diminuzione del numero di fotoni, ma non saranno funzione dell'operatore posizione e quantità di moto bensì del campo elettrico e del potenziale vettore. Seguendo questo approccio l'Hamiltoniana della luce che si propaga nello spazio libero ($H_{FreeSpace}$) può essere espressa come:

$$H_{FreeSpace} = \sum_{k,j} \hbar\omega_k a_{kj}^\dagger a_{kj}, \quad (5)$$

dove k è l'indice relativo alla singola onda piana con frequenza ω_k e j è un indice relativo alla polarizzazione della radiazione. Per la trattazione seguente conviene scrivere anche il potenziale vettore in termini di operatori di creazione e distruzione:

$$A_k(r) = \sum_{k,j} \lambda_{kj} (a_{kj} e^{ikr} + a_{kj}^\dagger e^{-ikr}), \quad (6)$$

$$\lambda_{kj} = \sqrt{\frac{\hbar}{2\epsilon_0 V \omega_k}} e_j, \quad (7)$$

dove k è il vettore d'onda, r è il vettore posizione, e_j è il versore di polarizzazione, ϵ_0 è la costante dielettrica nel vuoto e V è il volume modale (volume occupato dall'onda elettromagnetica confinata).

A questo punto si possono fare alcune considerazioni sui sistemi che interessano il nostro studio e compiere delle approssimazioni. Anzitutto in una microcavità il campo non risuona a tutte le frequenze ma esistono onde stazionarie solo alla frequenza di risonanza ω_{cav} , per cui la sommatoria dell'Hamiltoniana sarà costituita da un singolo termine o al massimo da quelli degli ordini di risonanza superiori. Inoltre considerando che le interazioni della luce avvengono in sistemi atomici o molecolari e che tipicamente in questi esperimenti si considerano frequenze nell'intervallo della luce visibile si ha che il prodotto kr è molto piccolo, quindi vale l'approssimazione di dipolo elettrico e la dipendenza del potenziale vettore dalla posizione può essere trascurata. Infine per semplicità consideriamo una singola polarizzazione. Con queste approssimazioni si ottengono l'Hamiltoniana della cavità H_{cav} e il potenziale vettore A_{cav} :

$$H_{cav} = \hbar\omega_{cav} a^\dagger a, \quad (8)$$

$$A_{cav}(r) = A_{cav}(0) = \sqrt{\frac{\hbar}{2\epsilon_0 V \omega_{cav}}} e_1 (a + a^\dagger). \quad (9)$$

Ottenute queste semplici equazioni che descrivono i fotoni confinati in una microcavità, passiamo alla descrizione delle loro interazioni con la materia.

Strong Coupling in microcavità

Le tipologie di interazione tra materia e fotoni in cavità possono essere suddivise in accoppiamenti deboli e forti. In accoppiamento debole la probabilità di una transizione elettronica spontanea aumenta se all'energia di transizione corri-

sponde una elevata densità di stati fotonici (in poche parole più fotoni ci sono e più l'atomo o la molecola ha probabilità di emettere un altro fotone). Nel regime di accoppiamento forte (Strong Coupling) invece il fotone e l'eccitone formano un sistema unico caratterizzato da due nuovi stati a due energie diverse da quella degli stati di partenza: a questi stati si possono associare delle quasi-particelle chiamate polaritoni che possiedono proprietà ibride rispetto a quelle degli elementi non accoppiati. Questo regime si può ottenere in qualsiasi sistema associabile a due oscillatori armonici accoppiati e quindi ha un perfetto analogo nella meccanica classica: consideriamo ad esempio due pendoli identici accoppiati da una molla (figura 7). Se la costante elastica della molla è sufficientemente grande, il doppio pendolo potrà oscillare a due frequenze corrispondenti all'oscillazione in fase e fuori fase dei singoli pendoli diverse dalla frequenza di oscillazione naturale dei due pendoli non accoppiati. La differenza tra le due nuove frequenze dipenderà dalla costante elastica della molla e quindi dal grado di accoppiamento degli oscillatori armonici.

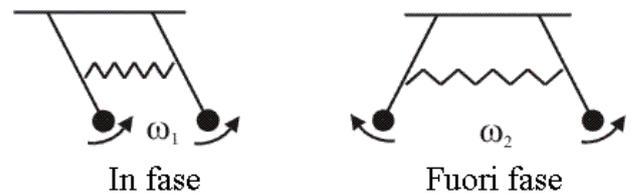


Figura 7: Oscillazioni naturali di due pendoli accoppiati in fase e fuori fase. Le frequenze delle oscillazioni sono diverse da quella del singolo pendolo e la differenza tra le due è proporzionale al grado di accoppiamento.

La condizione necessaria per entrare nel regime di Strong Coupling è inserire in una struttura di confinamento ottico un materiale in cui possono avvenire transizioni elettroniche (vedi schema in Figura 1). Consideriamo quindi un materiale in cui gli elettroni di valenza compiono transizioni dallo stato fondamentale a quello eccitato e formano con le lacune nello stato fondamentale (vacanze di elettroni con carica opposta) delle quasi-particelle chiamate eccitoni. Se l'energia della transizione eccitonica è uguale o molto vicina a quella dei fotoni confinati nel modo di risonanza della cavità, viene raggiunto l'accop-

piamento forte luce-materia quando eccitoni e fotoni si scambiano energia più rapidamente rispetto ad eventuali processi di decadimento (dephasing) dell'eccitone e di fuoriuscita del fotone dalla cavità. Questo scambio energetico è associato a un ciclo di emissione e riassorbimento del fotone chiamato ciclo di Rabi.

Il regime di strong coupling è stato raggiunto in tantissimi sistemi molto diversi tra di loro, ma quelli che hanno mostrato i fenomeni più interessanti sono i gas di atomi in cavità ultra-fredde e le microcavità di semiconduttori in stato solido. Nel prosieguo vedremo come descrivere un sistema in regime di strong coupling sia dal punto di vista ideale, con una trattazione puramente quantistica che ben si presta per analizzare esperimenti di strong coupling di singoli atomi, sia approfondendo la trattazione semiclassica con un approccio utile per analizzare le microcavità in stato solido. La prima parte si chiuderà con la descrizione degli esperimenti di *entanglement* a singolo polaritone, nella seconda presenteremo esperimenti di condensazione di Bose-Einstein in dispositivi a stato solido e a temperatura ambiente.

Il modello quantistico: sistemi a due livelli accoppiati

L'approccio quantistico offre un modello dell'accoppiamento luce-materia molto intuitivo e utile per avere una visione generale di qualsiasi sistema che presenta delle transizioni energetiche. In questa descrizione la radiazione confinata in microcavità sarà considerata come descritto nella sezione precedente, ossia come fotoni associati ad onde stazionarie alla frequenza di risonanza della cavità, mentre l'eccitone sarà considerato come un sistema atomico in cui un elettrone può compiere transizioni tra due livelli energetici.[2] Quest'ultima approssimazione è valida, oltre che per il campo della fisica atomica, anche in tutti i casi dove c'è confinamento elettronico, per esempio in strati sottili di molecole dove ci sono poche interazioni inter-molecolari e gli elettroni sono confinati o in strutture inorganiche di confinamento elettronico chiamate Quantum Wells.

Descriviamo dunque lo stato dei due sistemi non accoppiati schematizzati in figura 8: l'atomo

può essere descritto da un sistema a due livelli, lo stato fondamentale $|g\rangle$ e quello eccitato $|e\rangle$ che rappresentano la base per descriverne lo stato. I fotoni confinati invece saranno descritti da un sistema che ha per base gli stati di Fock del fotone alla frequenza di risonanza della cavità. Per semplificare la trattazione consideriamo solo gli stati di Fock $|0\rangle$ e $|1\rangle$ che rappresentano rispettivamente la situazione in cui non ci sono fotoni in cavità oppure ce n'è uno.

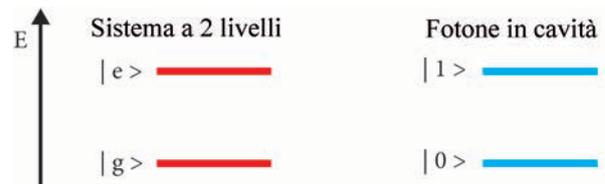


Figura 8: Stati energetici dell'elettrone in un sistema a due livelli e stati di Fock del fotone in cavità.

L'Hamiltoniana del sistema *fotone confinato e atomo* sarà descritta dall'espressione seguente:

$$H = \frac{1}{2m} (p - eA)^2 + V(r) + H_{cav}. \quad (10)$$

Il primo termine della sommatoria rappresenta l'energia cinetica di un elettrone immerso in un campo elettromagnetico dove p è l'operatore quantità di moto dell'elettrone, e la carica dell'elettrone, m la sua massa, A l'operatore potenziale vettore del campo elettromagnetico. Il secondo termine rappresenta il potenziale atomico, mentre il terzo è l'Hamiltoniana del campo elettromagnetico quantizzato in cavità (vedi equazione 8). Sviluppando questa espressione ricaviamo:

$$H = H_{cav} + H_e + H_{int} \quad (11)$$

con

$$H_e = \frac{p^2}{2m} + V(r), H_{int} = \frac{e^2 A^2}{2m} - \frac{epA}{m}. \quad (12)$$

Abbiamo isolato i termini relativi all'energia del solo elettrone (H_e) da quelli di interazione radiazione-materia (H_{int}). Come abbiamo visto sfruttando l'approccio della seconda quantizzazione possiamo riscrivere i termini dell'Hamiltoniana riferiti al campo elettromagnetico introducendo gli operatori di creazione e distruzione a^\dagger e a per i fotoni. Si può dimostrare che anche

il sistema atomico a due livelli può essere considerato come un oscillatore armonico quantistico analogamente al caso della radiazione elettromagnetica. Infatti le transizioni elettroniche tra i due livelli energetici del sistema atomico possono essere approssimate come un dipolo elettrico oscillante con frequenza di oscillazione pari a ω_{exc} . Introduciamo quindi gli operatori b^\dagger e b di creazione e distruzione di eccitazioni elettroniche. Se applichiamo l'operatore b^\dagger allo stato $|g\rangle$ avremo lo stato $|e\rangle$, ossia l'elettrone nello stato fondamentale ha compiuto una transizione nello stato eccitato. Applicando invece b a $|e\rangle$ otterremo lo stato $|g\rangle$, ossia il sistema elettronico nello stato fondamentale. Dunque l'Hamiltoniana elettronica si potrà scrivere come:

$$H_e = \hbar\omega_{exc}b^\dagger b. \quad (13)$$

È utile anche definire l'operatore quantità di moto in termini di questi nuovi operatori:

$$p = -\frac{m\omega_{exc}d}{e}(b + b^\dagger), \quad (14)$$

dove d è il momento di dipolo elettrico. Ricordiamo che il momento di dipolo elettrico è una misura della distanza tra due particelle cariche opposte che nell'eccitone sono identificate dall'elettrone e dalla lacuna. Sostituendo agli operatori p ed A nelle espressioni citate in precedenza, l'Hamiltoniana dell'intero sistema fotone-eccitone verrà quindi scritta in modo molto semplice:

$$H = \hbar\omega_{cav}a^\dagger a + \hbar\omega_{exc}b^\dagger b + \hbar\Omega(b + b^\dagger)(a + a^\dagger) + \frac{\hbar^2\Omega^2}{\hbar\omega_{cav}}(a + a^\dagger)(a + a^\dagger), \quad (15)$$

dove Ω è chiamata frequenza di Rabi. Nella condizione di risonanza per cui $\omega_{exc} = \omega_{cav} = \omega$ e considerando i dipoli elettrici allineati con la polarizzazione della radiazione, essa è uguale a:

$$\Omega = d\sqrt{\frac{\hbar\omega}{2\epsilon_0 V}}. \quad (16)$$

dove d è il momento di dipolo elettrico e V è il volume modale. Questa grandezza rappresenta la frequenza di interazione tra fotone ed eccitone e può essere vista come una misura dell'accoppiamento luce-materia. A questo punto possia-

mo fare una ulteriore approssimazione chiamata Rotating Wave Approximation (RWA). Consideriamo innanzitutto trascurabile il termine in cui compare il rapporto $\frac{\Omega^2}{\omega}$ poiché nei sistemi fisici studiati sperimentalmente la frequenza di Rabi è tipicamente molto più bassa della frequenza di risonanza. Inoltre sviluppando il termine in cui compare ω si possono fare altre considerazioni:

$$H_{int} = \hbar\Omega(ab + ab^\dagger + a^\dagger b + a^\dagger b^\dagger). \quad (17)$$

Possiamo suddividere i quattro termini dell'Hamiltoniana di interazione nei termini detti risonanti che moltiplicano ab^\dagger e $a^\dagger b$ e in quelli antirisonanti in cui compaiono $a^\dagger b^\dagger$ e ab . I termini risonanti si chiamano così perché considerano le transizioni tra fotone ed elettroni in risonanza e conservano l'energia del sistema dato che creano una eccitazione e contemporaneamente ne distruggono un'altra. Essi rappresentano transizioni reali come ad esempio nell'emissione spontanea si crea un fotone nella cavità e un eccitone decade. In figura 9 viene mostrato come agiscono i termini risonanti di interazione. I termini antirisonanti invece creano o distruggono contemporaneamente due eccitazioni, non seguono la conservazione dell'energia e quindi sono associati a transizioni virtuali che hanno una probabilità molto bassa di avvenire. Nella RWA i termini antirisonanti vengono trascurati proprio perché descrivono transizioni poco probabili.

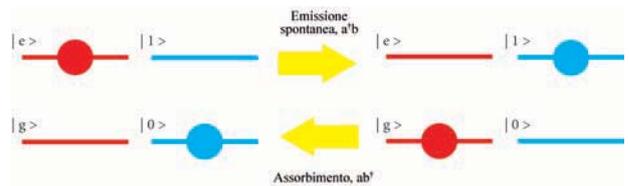


Figura 9: Visione quantistica dei processi di emissione spontanea e di assorbimento tra fotone ed elettrone. Nella situazione di sinistra l'atomo è eccitato e non ci sono fotone in cavità, mentre in quella di destra l'atomo è nello stato fondamentale ed è presente un fotone: le transizioni tra le due configurazioni vengono descritte dai termini risonanti di interazione ab^\dagger e $a^\dagger b$.

Con queste approssimazioni possiamo scrivere l'Hamiltoniana del sistema nella forma seguente:

$$H = \hbar\omega_{cav}a^\dagger a + \hbar\omega_{exc}b^\dagger b + \hbar\Omega(a^\dagger b + ab^\dagger). \quad (18)$$

In questa forma essa viene chiamata Hamiltoniana di Jaynes-Cummings in onore dei due scienziati che per primi hanno ideato il modello. [3] A causa della presenza dei termini di interazione luce-materia l'Hamiltoniana di Jaynes-Cummings non ha come autostati gli stati del fotone e dell'eccitone tal quali, ma viene diagonalizzata utilizzando una combinazione lineare delle basi $|g, 1\rangle$ e $|e, 0\rangle$ di un sottospazio dello spazio prodotto tra gli spazi di Hilbert dei due sistemi di partenza:

$$\begin{aligned} |UP\rangle &= \alpha |g, 1\rangle - \beta |e, 0\rangle, \\ |LP\rangle &= \alpha' |g, 1\rangle + \beta' |e, 0\rangle. \end{aligned} \quad (19)$$

I nuovi stati misti risultanti vengono chiamati stati polaritonici (Upper Polariton e Lower Polariton) e gli autovalori di energia considerando $\omega_{exc} = \omega_{cav} = \omega$ risultano:

$$E_{UP,LP} = \hbar(\omega \pm \Omega). \quad (20)$$

Grazie all'interazione fra i due sistemi dagli stati degeneri in energia di fotone ed eccitone si creano i due stati polaritonici che avranno una differenza di energia pari a $2\hbar\Omega$. Essendo metà fotone e metà eccitoni, i polaritoni possiedono caratteristiche ibride che li rendono interessanti per tanti studi e applicazioni: possiedono la capacità di interazione tipica delle particelle cariche e dotate di massa ma con una lunghezza di coerenza molto più elevata tipica del fotone. Inoltre nei polaritoni fotone ed eccitone sono *entangled* poiché questi stati non si possono scrivere come semplice prodotto delle basi dei sistemi separati ossia non sono fattorizzabili.

La trattazione quantistica presentata ci fornisce un risultato esatto delle possibili interazioni luce-materia in maniera semplice ed elegante, ma è valida in un sistema ideale e non tiene conto delle dissipazioni che determinano perdite di fase e coerenza. Per tener conto di questi effetti ci viene in aiuto la trattazione semiclassica che vedremo successivamente. Nonostante le difficoltà sono stati effettuati degli esperimenti in cui il sistema è stato progettato per lavorare in con-

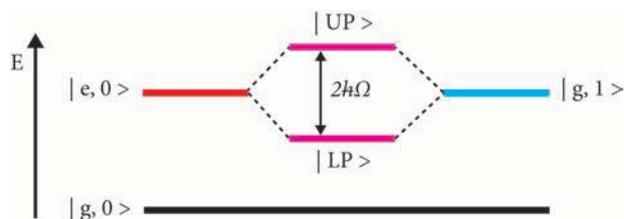


Figura 10: Formazione degli stati polaritonici Upper Polariton (UP) e Lower Polariton (LP) a partire dagli stati eccitone e fotone in cavità. La differenza in energia dei nuovi stati $2\hbar\Omega$ è direttamente proporzionale al grado di accoppiamento del sistema.

dizioni estreme in modo da rendere le perdite trascurabili: in questo apparato è stato possibile fare esperimenti che sono valsi il premio Nobel per aver studiato fenomeni di *entanglement* e decoerenza quantistica con singoli polaritoni.

Atomi e cavità entangled

Data la natura "Entangled" di fotoni ed eccitoni in regime di SC è possibile sfruttare le proprietà dei polaritoni per studiare questo fenomeno e per testare la decoerenza quantistica, ossia come e perchè nei sistemi macroscopici non vediamo i fenomeni tipici della fisica quantistica. Inoltre l'*entanglement* è una risorsa di fondamentale importanza per la computazione quantistica dove i bit classici sono sostituiti dai qubit, degli oggetti quantistici in sovrapposizione di stato che possono assumere valori continui tra 0 e 1. Se i qubit non fossero in stato *entangled* un computer quantistico sarebbe facilmente simulato da un computer classico. Ciascun qubit infatti può essere simulato da $2n$ bit con una accuratezza pari a n . Per simulare m qubit non *entangled* con la stessa accuratezza ci vorrebbero $2mn$ bit, mentre se i qubit fossero *entangled* il numero di bit equivalenti sarebbe $2^m n$! L'*entanglement* dunque rappresenta il vero valore aggiunto nella capacità di calcolo di un computer quantistico. Lo studioso che ha dato il contributo più significativo nell'*entanglement* tra fotoni ed elettroni è il premio Nobel Serge Haroche.[5] Nei suoi esperimenti atomi di Rydberg vengono inviati in una microcavità con altissimo fattore di qualità costituita da specchi superconduttori (Figura 11).

Descriviamo l'apparato sperimentale utilizzato: gli atomi di Rydberg sono degli atomi con

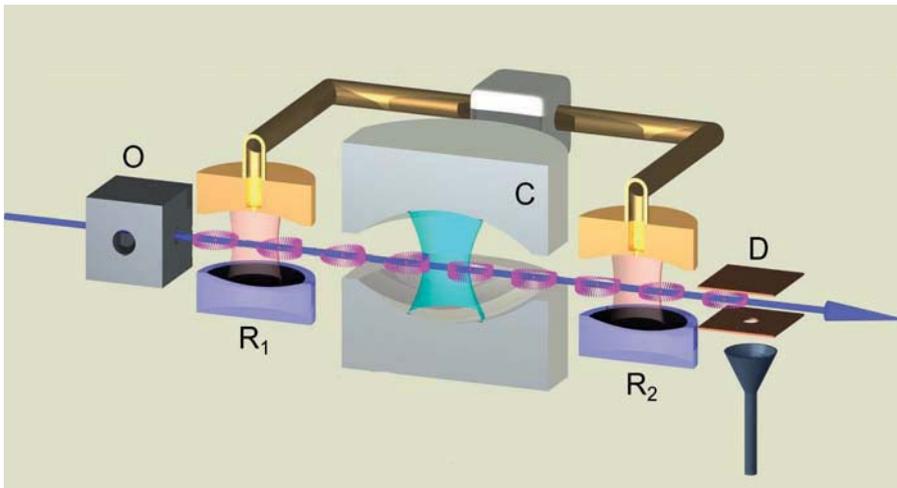


Figura 11: Schema dell'apparato sperimentale per testare l'accoppiamento atomi di Rydberg-cavità. Gli atomi provenienti da un raggio atomico vengono selezionati in base alla loro velocità da un pompaggio laser prima di essere preparati in stati di Rydberg (O). Essi attraversano la cavità C e vengono misurati in uscita da un detector D che applica un campo ionizzante. Eventualmente lo stato degli atomi può essere preparato tramite delle cavità ausiliarie R_1 e R_2 prima e dopo la cavità C per effettuare esperimenti di interferometria con la tecnica di Ramsey.

alto numero atomico in cui gli elettroni di valenza sono molto distanti dal nucleo e sono quindi schermati dal campo elettrico prodotto da esso. Tali elettroni vengono eccitati in modo che la configurazione elettronica sia simile all'atomo di idrogeno e sono quindi ottimi sistemi per studiare le proprietà quantistiche dei polaritoni poiché si comportano essenzialmente come un sistema a due livelli. La cavità è fabbricata in modo da confinare fotoni nell'intervallo delle microonde (frequenza di 51.1 GHz e lunghezza d'onda di 6 mm). Il lungo tempo di permanenza dei fotoni (circa 1 ms) dato dall'elevato fattore di qualità permette di far interagire la cavità con un atomo di Rydberg per volta se inviato all'interno di essa con l'opportuna velocità. Il modo di risonanza della cavità è uguale o vicino alla transizione tra due livelli energetici atomici $|e\rangle$ e $|g\rangle$. Per minimizzare le perturbazioni causate dalla radiazione termica l'apparato viene raffreddato fino a temperature bassissime (circa 1 K). Nella parte finale si inseriscono dei detector D che misurano lo stato dell'atomo dopo il passaggio dalla microcavità applicando un campo ionizzante che fornisce un segnale di tipo on/off se l'energia di ionizzazione è sufficiente o meno affinché l'elettrone venga strappato dall'atomo. Nel paragrafo precedente abbiamo descritto la frequenza di Rabi Ω come frequenza caratteristica del ciclo di scambio di energia tra fotoni in cavità ed atomi. La probabilità di trovare l'atomo nello stato $|e\rangle$ o $|g\rangle$ è infatti una funzione periodica caratterizzata dalla frequenza Ω . Nel setup di Haroche questa frequenza è elevata (circa 50

KHz) grazie all'alto grado di accoppiamento tra gli atomi di Rydberg e le microonde ed è quindi possibile misurare direttamente le oscillazioni tra i due stati dell'atomo. Inviando gli atomi uno alla volta nella cavità e cambiando il tempo di interazione luce-materia è possibile ricostruire sperimentalmente le oscillazioni di Rabi in funzione del tempo come mostrato in figura 12. Il tempo di interazione viene modificato a comando applicando un campo elettrico tra gli specchi della cavità che, grazie all'effetto Stark, modifica i livelli energetici degli atomi mandando il sistema fuori risonanza. In questo modo si può "congelare" l'oscillazione nella condizione voluta. Per esempio, come mostrato in figura 12, è possibile fermare l'interazione nella fase $\pi/2$ del ciclo di Rabi realizzando una sovrapposizione

$$|e, 0\rangle + |g, 1\rangle. \quad (21)$$

Questo stato *entangled* atomo-fotone sopravvive anche dopo che l'atomo ha lasciato la cavità. Si può anche interrompere il ciclo di Rabi dopo una fase π ottenendo l'effetto di invertire lo stato atomico e quello della cavità. Infine facendo compiere un ciclo di 2π si può cambiare il segno dello stato atomo-cavità (analogamente ad una rotazione di 2π di un sistema a spin 1/2). Tutte queste operazioni permettono di ottenere delle porte logiche quantistiche. Infatti il sistema a due livelli (negli stati $|e\rangle$ e $|g\rangle$) e il fotone in cavità (negli stati $|0\rangle$ e $|1\rangle$) si possono considerare come qubit che interagiscono grazie allo Strong Coupling.

Le potenzialità dello Strong Coupling per stu-

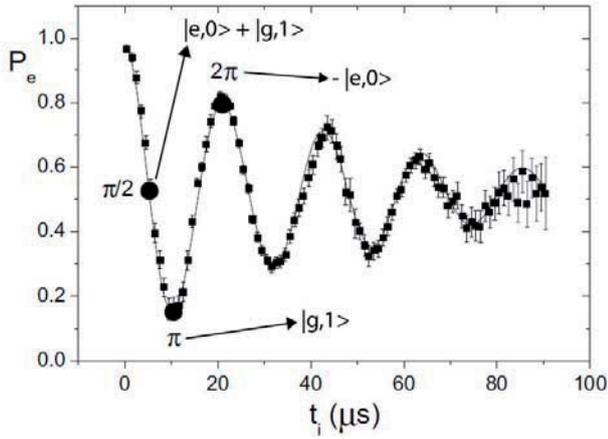


Figura 12: Oscillazioni di Rabi misurate sperimentalmente per atomi di Rydberg inviati nella cavità nello stato iniziale $|e\rangle$. Ripetendo l'esperimento per diversi tempi di interazione atomo-cavità si può ricostruire la probabilità P_e di trovare l'atomo in stato $|e\rangle$ dopo essere uscito dalla cavità. A seconda del momento in cui cessa l'interazione il sistema può essere "congelato" in qualsiasi stato.

diare i sistemi nei regimi quantistici non si esauriscono qui. Per esempio l'oscillazione di Rabi quantistica fornisce un metodo efficace per preparare due atomi in stato *entangled*. Dopo che un atomo è stato inviato nella cavità ed è stato congelato dopo una fase $\pi/2$, si invia un secondo atomo inizialmente in stato $|g\rangle$ che entra nel ciclo di rabi per una fase π , riassorbendo il fotone che era già *entangled* con il primo atomo. Alla fine del processo la cavità rimane senza fotoni e i due atomi diventano *entangled* nello stato

$$|e, g\rangle + |g, e\rangle \quad (22)$$

Il gruppo di Haroche ha compiuto numerosi test su questo stato dimostrandone l'effettiva proprietà di *entanglement* e studiando le proprietà di correlazione. Lo stesso apparato sperimentale può essere sfruttato per tantissimi studi e applicazioni, per esempio per testare la decoerenza quantistica in sistemi in sovrapposizione di stati tipo "gatto di Schrödinger". Utilizzando una cavità ausiliaria prima della cavità principale C si possono preparare gli atomi che entrano nell'apparato in stato $|e\rangle$ in sovrapposizione di stati

$$|g\rangle + |e\rangle \quad (23)$$

tramite un impulso laser che faccia compie-

re una evoluzione di $\pi/2$ nel ciclo di Rabi. In questa cavità ausiliaria a basso fattore di qualità gli atomi sono in regime di strong coupling, ma il campo elettromagnetico all'interno di essa si comporta classicamente, quindi non produce stati *entangled*. Supponiamo che gli atomi in sovrapposizione di stati dopo l'impulso della cavità R_1 non interagiscano con la cavità centrale ad alto Q ma passino direttamente attraverso una terza cavità R_2 sempre caratterizzata da un basso fattore di qualità. Anche in questo caso verrà inviato un impulso di $\pi/2$ alla stessa frequenza di R_1 che ha la funzione di applicare una seconda trasformazione allo stato degli atomi: se gli impulsi in R_1 e in R_2 sono in fase allora nel caso ideale ci sarà il 100% di probabilità che il detector finale misuri l'atomo nello stato $|g\rangle$. Viceversa se c'è uno sfasamento questa probabilità sarà in generale una funzione sinusoidale della differenza di fase ϕ tra gli impulsi e lo stato atomico:

$$P_e = 1 - P_g = (1 - \cos\phi)/2 \quad (24)$$

che a sua volta è definita dalla relazione seguente:

$$\phi = (\omega_r - \omega_{eg})T_r \quad (25)$$

dove ω_r è la frequenza degli impulsi nelle cavità ausiliarie, ω_{eg} è la frequenza della transizione atomica e T_r è il tempo impiegato dagli atomi per passare dalla prima alla seconda cavità. Variando dunque la differenza tra la frequenza di transizione e quella degli impulsi si ottiene il grafico mostrato in figura in cui compaiono le frange di interferenza. Questo apparato prende il nome di interferometro di Ramsey (schematizzato in figura 13) e può essere sfruttato per misurare la decoerenza di un sistema quantistico. La presenza di interferenza infatti è direttamente legata alla coerenza degli atomi tra le due cavità R_1 e R_2 : qualsiasi perturbazione che intacca la coerenza degli atomi in sovrapposizione di stato dopo R_1 , per esempio la presenza di fotoni nella cavità C, avrà l'effetto di distruggere l'interferenza e diminuire il contrasto tra le frange. Haroche ha sfruttato proprio questo effetto per misurare i tempi di decoerenza di fotoni confinati nella cavità centrale.[6]

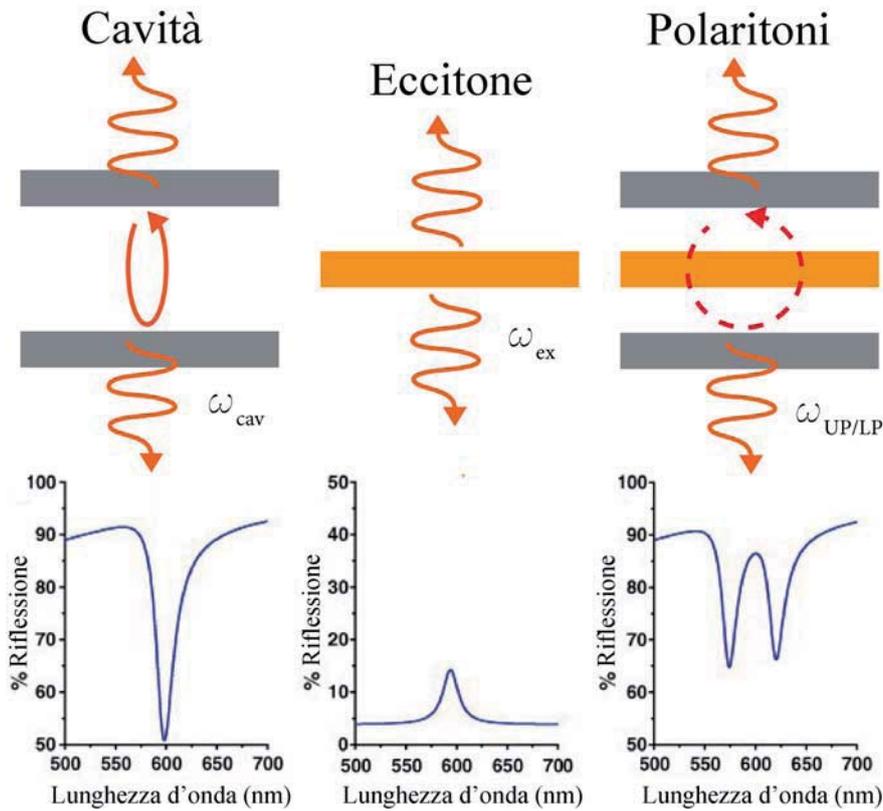


Figura 14: Spettri di riflettanza sperimentali (1-T) dei sistemi cavità e materiale che assorbe e riemette fotoni prima e dopo l'istaurarsi del regime di Strong Coupling. Lo spettro del materiale identifica la transizione eccitonica. Gli spettri dei sistemi non accoppiati sono riconducibili a funzioni di oscillatori armonici, da cui è possibile ricavare la frequenza complessa di oscillazione naturale la cui parte immaginaria è una misura del tempo di dephasing di eccitoni e fotoni in cavità.

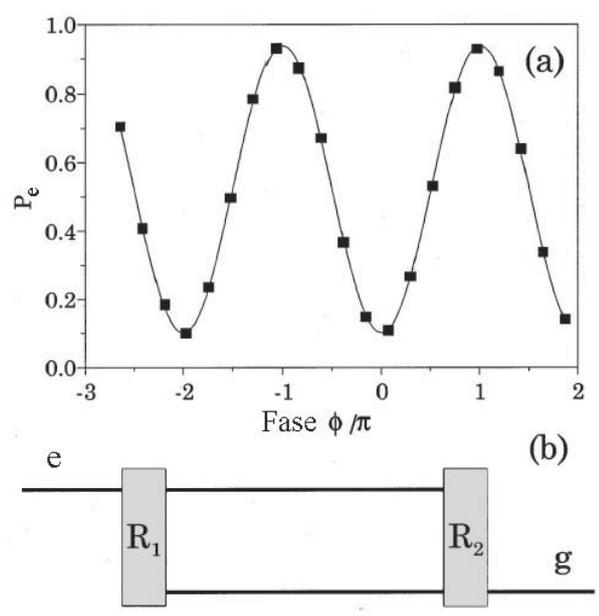


Figura 13: Frange di interferenza di Ramsey: a) grafico per la transizione $e \rightarrow g$ della probabilità P_e di misurare l'atomo nello stato $|e\rangle$ in funzione della differenza di fase ϕ tra i due impulsi $\pi/2$ in R_1 e R_2 espressa in unità di π . I punti sono misure sperimentali mentre la linea è un fit sinusoidale. La probabilità non arriva al 100% a causa di perturbazioni e perdite nell'apparato. b) Schema che mostra i due cammini seguiti dagli stati di un atomo nell'interferometro di Ramsey.

Strong Coupling in stato solido

I fenomeni dello Strong Coupling sono stati studiati e sfruttati anche in dispositivi a stato solido utilizzando dei materiali semiconduttori. Rispetto alla trattazione fatta nel caso ideale con cui si possono correttamente modellizzare sistemi atomici in cavità ultra-fredde, in questo caso occorre considerare le perdite delle cavità e dei materiali emissivi usando la trattazione semiclassica. Consideriamo anche in questo caso le transizioni eccitoniche e i modi fotonici della cavità come degli oscillatori armonici ma di tipo classico. Le frequenze di transizione eccitonica e di risonanza della cavità sono delle grandezze complesse: le parti immaginarie saranno rispettivamente relative alla velocità di decadimento dell'eccitone, che corrisponde approssimativamente alla larghezza a metà altezza del picco di assorbimento, e alla velocità di dissipazione del fotone nella cavità, che corrisponde alla larghezza di banda dello spettro di riflettanza (vedi Fig. 14). Dall'equazione che descrive il moto di un oscillatore armonico smorzato e forzato possiamo ricavare la frequenza complessa di risonanza la cui parte reale corrisponde alla frequenza di picco di risonanza, mentre la parte immaginaria si riferisce al termine di smorzamento (quindi

di dissipazione). Applicando una trasformata di Fourier si ottiene uno spettro di intensità in funzione della frequenza che riconduce il coefficiente di smorzamento alla larghezza a metà altezza del picco spettrale dell'oscillatore.

Conosciute le frequenze dei sistemi non accoppiati, per risalire alle frequenze degli stati polaritonici possiamo scrivere l'equazione di Schrödinger indipendente dal tempo come:

$$\begin{bmatrix} \omega_{exc} - i\gamma_{exc} & V \\ V & \omega_{cav} - i\gamma_{cav} \end{bmatrix} \begin{pmatrix} |Exc\rangle \\ |Cav\rangle \end{pmatrix} = \omega \begin{pmatrix} |Exc\rangle \\ |Cav\rangle \end{pmatrix}, \quad (26)$$

dove V rappresenta il termine di interazione. Diagonalizzando la matrice si ottengono gli autovalori:

$$\omega_{UP/LP} = \frac{\omega_{exc} + \omega_{cav} - i(\gamma_{exc} + \gamma_{cav})}{2} \pm \sqrt{V^2 + \frac{1}{4}[\omega_{exc} - \omega_{cav} - i(\gamma_{exc} - \gamma_{cav})]^2}. \quad (27)$$

Quando ω_{exc} e ω_{cav} coincidono (condizione di anticrossing), possiamo ricavare la frequenza di Rabi Ω sapendo che la differenza di energia tra i due stati polaritonici (Rabi splitting) è uguale a $2\hbar\Omega$:

$$\Omega = \sqrt{V^2 - \frac{1}{4}(\gamma_{cav} - \gamma_{ex})^2}. \quad (28)$$

Dall'equazione precedente risulta che la condizione per cui la radice sia reale è che il suo argomento deve essere positivo ossia che $|V| > |\gamma_c - \gamma_{ex}|$, quindi il termine di interazione deve essere maggiore della differenza tra i termini di dissipazione del fotone e dell'eccitone. Confrontando questo risultato con quello ottenuto da una trattazione puramente classica si può identificare il termine di interazione V con i parametri sperimentali del sistema cavità-semiconduttore:

$$V^2 = \frac{4AdQ\gamma_{ex}\gamma_{cav}}{\pi}, \quad (29)$$

dove A è il coefficiente di assorbimento del materiale, d è lo spessore di tale strato semiconduttore e Q è il fattore di qualità della cavità.[4]

Da questa relazione ricaviamo una dipendenza diretta del termine di accoppiamento con i parametri sperimentali della microcavità: per aumentare l'accoppiamento è necessario usare dei materiali con alto coefficiente di assorbimento e delle cavità con alto fattore di qualità.

Nella condizione di anticrossing ($\omega_{exc} = \omega_{cav}$) i coefficienti α e β di eccitone e fotone (Eq. 19) sono uguali per entrambi gli stati, quindi i polaritoni avranno uguale carattere fotonico ed eccitonico, ma questa condizione non è costante per tutti gli angoli. Come abbiamo visto in precedenza infatti la frequenza di risonanza di una microcavità dipende dall'angolo, mentre in film amorfi di semiconduttori la frequenza di assorbimento ha una dispersione piatta. Questo implica che in una microcavità ci sarà solo un ristretto intervallo di angoli per cui le energie di fotone ed eccitone coincidono; per tutti gli altri angoli i coefficienti α e β saranno in generale diversi e i polaritoni avranno più carattere fotonico o eccitonico. Questo si riflette anche sull'andamento dell'energia dei polaritoni in funzione dell'angolo che sarà più piatto quando il carattere è più eccitonico, mentre nel caso opposto seguirà più l'andamento della cavità (Figura 15). Come vedremo nella sezione dedicata ai condensati di Bose-Einstein, è utile associare ai polaritoni una massa efficace: essa dipende dai coefficienti di mixing, ma in genere sarà dominata da quella del fotone essendo molto più piccola di quella dell'eccitone (il fotone in cavità ha una massa efficace non nulla proprio a causa del confinamento).[7]

Per la scelta del materiale su cui creare gli eccitoni si possono considerare semiconduttori organici o inorganici: i semiconduttori organici presentano eccitoni di tipo Frenkel, che sono localizzati, hanno alte energie di legame ed elevati coefficienti di assorbimento, proprietà vantaggiose per avere Strong Coupling a temperatura ambiente. Il fenomeno di elettroluminescenza da polaritoni a temperatura ambiente è stato osservato per la prima volta usando una struttura a base di semiconduttori organici contenente un sottile strato di aggregati di molecole organiche racchiuso tra due strati di materiali per l'iniezione di buche e ed elettroni e due specchi metallici che fungevano anche da elettrodi.[8] Un'altra caratteristica interessante dei semiconduttori organici è che grazie alla forte energia

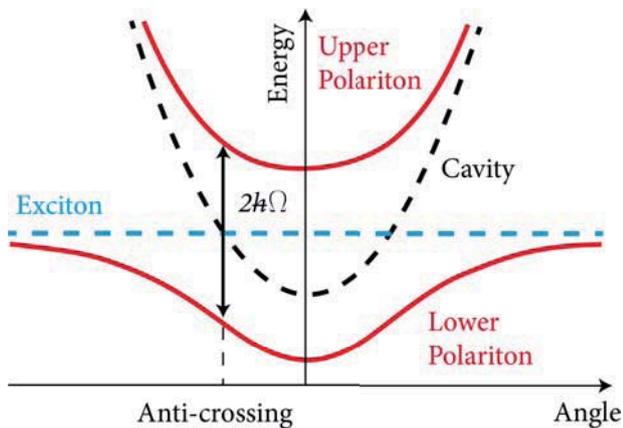


Figura 15: Grafico dell'energia associata agli stati polaritonici in funzione dell'angolo. Il punto in cui l'energia dell'eccitone eguaglia quella della cavità prende il nome di anticrossing. L'andamento delle dispersioni polaritoniche è più simile a quello dell'eccitone o della cavità non accoppiati se lo stato polaritonico è più eccitonico o fotonico.

di legame che caratterizza gli eccitoni Frenkel, si possono raggiungere Rabi Splitting molto elevati: questo implica la comparsa di un regime di accoppiamento spinto chiamato Ultra Strong Coupling (USC). Tale regime viene raggiunto quando il rapporto tra il Rabi Splitting e l'energia della transizione del materiale, chiamato grado di accoppiamento g , diventa considerevole (convenzionalmente si considera come soglia il valore 0.2). In questo regime il sistema non può più essere descritto dal modello quantistico approssimato di Jaynes-Cummings che sfrutta la Rotating Wave Approximation ma si devono tenere in conto anche i termini anti-risonanti che generano nuovi fenomeni interessanti legati al vuoto quantistico che però non approfondiremo in questo articolo.[9]

I semiconduttori inorganici presentano eccitoni di tipo Wannier-Mott che sono più delocalizzati nel reticolo cristallino ed hanno energie di legame più basse. Questa caratteristica rende gli eccitoni molto meno stabili al punto che anche a temperatura ambiente elettrone e lacuna si possono slegare perdendo la possibilità di accoppiarsi con i fotoni in cavità, quindi è necessario fare esperimenti a basse temperature. D'altro canto la delocalizzazione eccitonica è molto utile per ottenere facilmente condensati di Bose-Einstein di polaritoni come vedremo nel prossimo paragrafo.

Condensati di Bose-Einstein di polaritoni

Oltre alle sorprendenti proprietà descritte, i polaritoni sono particelle bosoniche e si possono quindi utilizzare per ottenere condensati di Bose-Einstein (BEC) in sistemi in stato solido: in particolare i sistemi in Strong Coupling mostrano delle caratteristiche vantaggiose per il verificarsi di questo fenomeno. Per descrivere i BEC è utile fare una digressione sui sistemi ideali di gas atomici ultra freddi. Quando la temperatura di un gas ideale di bosoni si abbassa fino a una certa temperatura critica avviene una transizione di fase dello stato quantico del sistema: la condensazione di Bose-Einstein. L'origine di questo fenomeno deriva dalla natura ondulatoria e dalla indistinguibilità intrinseca dei bosoni: essi possono occupare lo stesso stato quantico e possono essere descritti come dei pacchetti d'onda che hanno un'estensione pari alla lunghezza d'onda di De Broglie

$$\lambda_{dB} = \frac{h}{\sqrt{2\pi m k_B T}}, \quad (30)$$

dove h è la costante di Planck, m è la massa della particella, k_B è la costante di Boltzmann e T è la temperatura del sistema. La lunghezza d'onda di De Broglie può essere intesa come l'incertezza spaziale sull'estensione delle particelle che per il principio di indeterminazione di Heisenberg è inversamente proporzionale alla quantità di moto associata ad esse. Se la temperatura del sistema scende, la lunghezza d'onda verrà incrementata, fino a che non diventa comparabile con la distanza media fra le particelle causando una sovrapposizione dei pacchetti d'onda. Anche la quantità di moto e l'energia dei bosoni diminuiscono con la temperatura fino a che non avviene una transizione ad una certa T critica per cui le particelle iniziano a collapsare nello stesso stato quantico fondamentale a più bassa energia. A differenza dei fermioni infatti, non esiste un limite per il numero di bosoni che possono occupare lo stesso stato quantico.

Il condensato di Bose-Einstein si forma quando la temperatura raggiunge lo zero assoluto e tutte le particelle sono nello stesso stato. È interessante far notare che se le particelle non interagiscono ciascuna di esse avrà una funzione d'onda ugua-

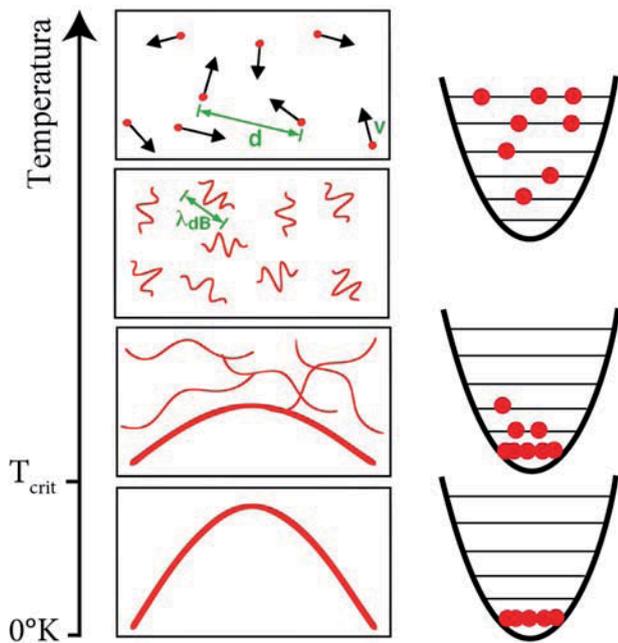


Figura 16: Schema degli stadi per la transizione verso il condensato di Bose-Einstein di un gas atomico in funzione della temperatura. Raggiunta la temperatura critica la lunghezza d'onda di De Broglie è tale da permettere la sovrapposizione delle funzioni d'onda e le particelle iniziano a collassare in un unico stato quantico.

le alle altre ma rimarrà comunque indipendente. Se sono presenti interazioni invece, le particelle collassano in un oggetto macroscopico caratterizzato da un unico stato quantistico con una specifica fase. Un condensato di Bose-Einstein interagente è quindi un oggetto coerente la cui funzione d'onda quindi può essere descritta dalla seguente espressione:

$$\Psi(r) = \sqrt{N}e^{i\varphi(r)}, \quad (31)$$

dove $\varphi(r)$ è la fase comune e N è il numero totale di particelle condensate. [10] L'interazione tra le particelle quindi è responsabile della coerenza del condensato e in aggiunta permette di ottenerlo a temperature maggiori dello zero assoluto.

Veniamo dunque al caso dei polaritoni per cui la condensazione di Bose-Einstein ha dei vantaggi tecnologici notevoli dovuti soprattutto alla possibilità di avere questo fenomeno a temperature meno estreme. I polaritoni infatti sono caratterizzati da una massa efficace e da un tempo di vita essenzialmente dominati dalla componente fotonica; quindi a parità di temperatura, la

lunghezza d'onda di De Broglie associata a un polaritone è molto più grande di quella atomica o eccitonica (nell'ordine dei 100 micron) e questo permette l'interazione delle particelle in fase di condensazione ad una temperatura più alta rispetto ad altri sistemi. Inoltre la componente eccitonica dei polaritoni è responsabile di forti interazioni coulombiane tra le particelle. Queste interazioni polaritone-polaritone facilitano il collasso delle particelle nello stato a minima energia, che nella dispersione tipica di un sistema in Strong Coupling corrisponde allo stato lower polariton ad angolo zero (vedi Fig. 15). I condensati di polaritoni sono interessanti tecnologicamente anche perchè producono emissione laser: dato che le particelle condensate sono nello stesso stato quando decadono sotto forma di fotoni avranno la stessa lunghezza d'onda e la stessa fase, come nell'emissione laser, ma senza il requisito dell'inversione di popolazione e quindi a soglie di potenza più basse.

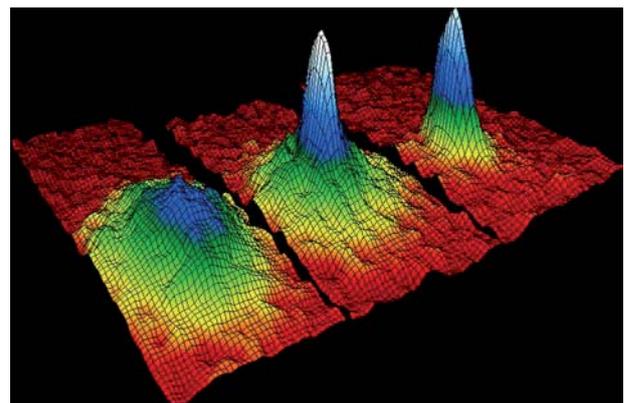


Figura 17: Tre grafici di distribuzione delle velocità di un gas di atomi di rubidio; in quello intermedio e nell'ultimo a destra si forma e si stabilizza un condensato di Bose-Einstein.

Dalla prima osservazione di fenomeni polaritonici in dispositivi a semiconduttore avvenuta nel 1992 [11] sono state compiute numerose ricerche utilizzando materiali organici ed inorganici in microcavità verso sempre nuovi ed interessanti studi ed applicazioni. Sono stati osservati negli anni fenomeni come il lasing polaritonico, condensati di Bose-Einstein di polaritoni a temperatura ambiente,[12] [14] e come citato in precedenza sono stati fabbricati dispositivi polaritonici elettroluminescenti.[13]

Conclusioni

Grazie al confinamento delle onde elettromagnetiche le interazioni luce-materia si possono spingere al punto in cui fotoni ed eccitoni arrivano a fondersi. Le particelle ibride che si creano, i polaritoni, hanno caratteristiche interessantissime sia per studi di base che per la fabbricazione di dispositivi fotonici innovativi. Studiare i condensati di Bose-Einstein in sistemi in stato solido, facili da fabbricare ed analizzare, può aprire le porte a nuovi e sorprendenti esperimenti per esplorare i confini tra mondo quantistico e classico. Il passo tecnologico successivo è raggiungere questo stato tramite controllo elettrico in un dispositivo integrato. Questa tematica è di interesse cruciale per i futuri computer quantistici che, basandosi sui sorprendenti principi della fisica quantistica, rivoluzioneranno il modo di fare computazione.



- [1] K.J. VAHALA: "Optical microcavities", *Nature* **424** (2003) 839-846.
- [2] A. KAVOKIN, J.J. BAUMBERG, G. MALPUECH, F. LAUSSY: *Microcavities*. Oxford University Press, Oxford (2011).
- [3] E.T. JAYNES, F.W. CUMMINGS: "Comparison of quantum and semiclassical radiation theories with application to the beam maser", *Proc. IEEE* **51(1)** (1963) 89-109.
- [4] R.J. HOLMES, S.R. FORREST: "Strong exciton-photon coupling in organic materials", *Organic Electronics* **8** (2007) 77-93.
- [5] S HAROCHE: "Nobel Lecture: Controlling photons in a box and exploring the quantum to classical boundary", *Reviews of Modern Physics* **85(3)** (2013) 1083-1102.
- [6] M. BRUNE, E. HAGLEY, J. DREYER, X. MAI[^]TRE, A. MAALI, C. WUNDERLICH, J. M. RAIMOND, S. HAROCHE: "Observing the Progressive Decoherence of the "Meter" in a Quantum Measurement", *Physical Review Letters* **77(24)** (1996) 4887-4890.
- [7] P. MICHETTI AND G. C. LA ROCCA: "Polariton states in disordered organic microcavities", *Physical Review B* **71** (2005) 115320.
- [8] M. MAZZEO, A. GENCO, S. GAMBINO, D. BALLARINI, F. MANGIONE, O. DI STEFANO, S. PATANÈ, S. SAVASTA, D. SANVITTO, G. GIGLI: "Ultrastrong light-matter coupling in electrically doped microcavity organic light emitting diodes", *Applied Physics Letters* (2014).
- [9] C. CIUTI, G. BASTARD AND I. CARUSOTTO: "Quantum vacuum properties of the intersubband cavity polariton field", *Physical Review B* **72** (2005) 115303.
- [10] W. KETTERLE, D.S. DURFEE, D.M. STAMPER-KURN: "Making, probing and understanding Bose-Einstein condensates", (1999) <http://arxiv.org/abs/cond-mat/9904034>

- [11] WEISBUCH, NISHIOKA, ISHIKAWA, ARAKAWA: "Observation of the coupled exciton-photon mode splitting in a semiconductor quantum microcavity", *Physical Review Letters* (1992).
- [12] J. KASPRZAK, M. RICHARD, S. KUNDERMANN, A. BAAS, P. JEAMBRUN, J. M. J. KEELING, F. M. MARCHETTI ET AL.: "Bose-Einstein condensation of exciton polaritons", *Nature* **7110** (2006) 409-414.
- [13] S. I. TSINTZOS, N. T. PELEKANOS, G. KONSTANTINIDIS, Z. HATZOPOULOS, P. G. SAVVIDIS: "A GaAs polariton light-emitting diode operating near room temperature", *Nature* **7193** (2008) 372-375.
- [14] K. S. DASKALAKIS, S. A. MAIER, R. MURRAY, S. KÉNA-COHN: "Nonlinear interactions in an organic polariton condensate", *Nature Materials* (2014).



Armando Genco: Dottorando in "Fisica e Nanoscienze" presso il Dipartimento di Matematica e Fisica "Ennio De Giorgi" di Lecce. Lavora in collaborazione con il CNR-NANOTEC di Lecce nel campo della fabbricazione di dispositivi organici elettro-luminescenti per la nano-fotonica quantistica.

Neutrini e Cosmologia

Elles sont engendrées par un raisonnement mathématique, établies sur des calculs, réductibles à des schémas d'une grande sécheresse. Mais dans ces cadres sévères, une sorte de fièvre presse et multiplie les figures; un étrange génie de complication enchevêtre, replie, décompose et recompose leur labyrinthe.

H. Focillon, *Vie des Formes*

Gianpiero Mangano

Istituto Nazionale di Fisica Nucleare, Sezione di Napoli, Italy

I neutrini sono le più elusive particelle elementari note. La loro esistenza fu ipotizzata da Wolfgang Pauli per salvare la conservazione dell'energia e la statistica degli spin nei decadimenti β . La loro rivelazione e le loro proprietà hanno sfidato i fisici per decenni. Mostrano una fenomenologia incredibilmente ricca e lasciano tracce in molti ambiti, dalle centrali nucleari all'interno del pianeta terra, dai processi di fusione nucleare all'interno delle stelle, all'esplosione delle supernovae. In questo breve articolo descriverò come i neutrini influenzino molti stadi dell'evoluzione dell'universo, il cui studio è normalmente chiamato *Cosmologia*.

Neutrini

I neutrini apparvero sulla scena nel 1930, quando Wolfgang Pauli, per salvare la conservazione dell'energia e la relazione tra gli spin e la statistica nei decadimenti nucleari β , propose che, insieme con un elettrone, fosse emessa una particella neutra di spin 1/2, che, a quei tempi, Pauli

chiamò *neutrone*.

Pauli confessò più tardi al suo collega, l'astronomo Walter Baade, che (citato in [1]):

“Oggi ho fatto qualcosa di terribile, qualcosa che nessun fisico teorico dovrebbe mai fare. Ho suggerito qualche cosa che non potrà mai essere verificato sperimentalmente.

Questa è, forse, la sola volta che Pauli sbagliò. Trenta anni dopo, nel 1956, i neutrini furono scoperti da Reines e Cowan.

Oggi sappiamo molto dei neutrini, ma loro custodiscono ancora qualche segreto. Poiché le loro proprietà sono state raccontate in dettagli nel precedente numero di *Ithaca*, in particolare si vedano [3, 4, 5], riassumerò lo stato dell'arte e introdurrò il minimo numero di concetti necessari che userò in seguito.

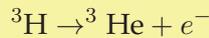
Cosa sappiamo

i) Sapori e interazioni dei neutrini.

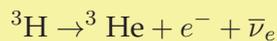
Ci sono tre specie di neutrini (o *sapori*), ν_e, ν_μ e ν_τ , e tre corrispondenti antiparticelle (antineutrini) $\bar{\nu}_e, \bar{\nu}_\mu$ e $\bar{\nu}_\tau$, che appaiono sempre in processi di interazione accompagnate dai loro partner carichi e^\pm, μ^\pm and τ^\pm , rispettivamente. I neutrini interagiscono tra loro e

L'idea di Pauli

Nel decadimento β di nuclei stabili come il trizio ${}^3\text{H}$, l'energia cinetica dell'elettrone E_e sarebbe costante se ci fosse una sola particella emessa



e data dal Q – valore della reazione, $Q = M({}^3\text{H}) - M({}^3\text{He}) - M(e^-)$, dove M è la massa della particella corrispondente. Questo non è quello che si riscontra negli esperimenti, che invece mostra che l'energia dell'elettrone copre una distribuzione continua di energia da un valore minimo fino al massimo valore Q . La conservazione sarebbe violata, a meno che una particella neutra venga emessa, in modo che la carica elettrica sia conservata: un elettrone e un antineutrino



L' (anti)neutrino dovrebbe avere spin semi-intero per soddisfare la relazione tra spin e statistica (ancora una idea fondamentale di Pauli: il principio di esclusione.) Dall'analisi del decadimento β del nucleo ${}^3\text{H}$ sappiamo che m_ν non può essere più grande di 2 eV, un limite ottenuto osservando il massimo di energia dell'elettrone $E_e \leq Q - m_\nu$. L'esperimento KATRIN [2] dovrebbe abbassare questo limite superiore di un ordine di grandezza, o misurare la massa del neutrino, se questa è maggiore di 0.2 eV.

con le altre particelle solo mediante le interazioni debole e gravitazionale, diversamente dai leptoni carichi, sensibili anche all'interazione elettromagnetica, e ai quark che sono sensibili a tutte le interazioni fondamentali, inclusa l'interazione forte.

Quando le energie e le quantità di moto sono sufficientemente piccoli (nel seguito noi ci occuperemo con energie della scala del MeV= 10^6 eV) i processi possono essere descritti dal modello di *Fermi-Gamow-Teller*, la prima descrizione coerente delle interazio-

ni deboli proposta da Fermi nel 1934. Nel regime di bassa energia l'intensità dell'interazione è definita dalla costante di Fermi $G_F \simeq 1.166 \cdot 10^{-5} \text{ GeV}^{-2}$ ($\text{GeV} = 10^9 \text{ eV}$). Dagli anni settanta dell'ultimo secolo, la teoria di Fermi è stata meravigliosamente integrata nel Modello Standard delle interazioni elettrodeboli di Glashow, Weinberg e Salam. Come per le forze elettromagnetiche, dove particelle elettromagneticamente cariche interagiscono tra loro scambiandosi fotoni, i processi deboli sono mediati dallo scambio di *bosoni intermedi* W^\pm and Z^0 . Il corto raggio e la debole intensità di questi processi è alla fine dovuto alla grande massa di queste particelle, 80 GeV e 91 GeV per W^\pm e Z^0 , rispettivamente.

Le interazioni gravitazionali sono tipicamente trascurabili in esperimenti in laboratorio. In questo caso, i neutrini diffondono, appaiono o sono assorbiti solo per interazione debole. In ambienti molto densi e grandi, come l'universo considerato nella sua globalità, la gravità è piuttosto importante e, come vedremo, il contributo dei neutrini al campo gravitazionale globale non è trascurabile.

ii) Masse dei neutrini e oscillazioni di sapore.

Sebbene Pauli menzionasse esplicitamente nella sua lettera alle "signore e signori radioattivi" nel 1930 che i neutrini ... *differiscono dai quanti di luce nel fatto che non viaggiano alla velocità della luce. La massa del neutrone (come abbiamo detto Pauli aveva chiamato neutrone la nuova particella) dovrebbe essere dello stesso ordine di grandezza della massa dell'elettrone e in ogni caso non più grande di 0.01 della massa del protone*, fino alla grande rivoluzione dopo la scoperta delle oscillazioni del neutrino, si ipotizzava che le masse dei neutrini fossero nulle.

Il quadro teorico appoggiava questa idea poiché nella versione più semplice del Modello Standard delle interazioni elettrodeboli non c'è alcuna necessità di introdurre una componente destrorsa del neutrino, o una componente sinistrorsa dell'antineutrino. In questo caso il neutrino è rappresentato da uno spinore di Weyl privo di massa (si veda [4]). Tuttavia, nelle ultime tre decadi, i

Distillato di Cosmologia I: evidenze

Il modello del Big Bang caldo è, attualmente, la nostra descrizione dell'universo. In effetti, uno dei risultati principali dell'applicazione della teoria della Relatività Generale di Einstein a questo grande sistema è che noi possiamo pensare in termini di un universo in evoluzione, qualche cosa che nel linguaggio di Newton sarebbe difficile da implementare. Ci sono tre osservazioni principali in favore della teoria del Big Bang.

- 1) **La legge di Hubble.** Il 1929 è l'anno in cui la legge dell'allontanamento delle galassie fu formulata da Edwin Hubble. Entità come le galassie osservate nello spazio profondo a distanze sufficientemente grandi presentano spostamenti Doppler delle linee spettrali dovuti alla loro velocità relativa rispetto alla terra. Questa velocità v di una galassia che si allontana dalla terra è approssimativamente proporzionale alla distanza d dalla terra, almeno per le galassie lontane fino a poche centinaia di megaparsec, si veda la figura 1, $v = H_0 d$, dove $H_0 \sim 70 \text{ Km/s Mpc}^{-1}$ è la costante di Hubble. In realtà la legge di Hubble fu derivata teoricamente alcuni anni prima da Georges Lemaître dalla Relatività Generale di Einstein come una caratteristica tipica dell'universo in espansione.
- 2) **La radiazione cosmica di fondo (CMB).** A meno di piccole fluttuazioni dell'ampiezza dell'ordine di 10^{-5} , la terra riceve una radiazione isotropa di fotoni con una proprietà rilevante: la distribuzione delle frequenze (ν) descrive perfettamente l'emissione di un corpo nero! Rivelata da Penzias e Wilson nel 1964, fu subito interpretata come eco dei primi stadi dell'espansione dell'universo come già ipotizzato da George Gamow negli anni 40.

Una distribuzione di corpo nero è l'analogo della distribuzione di Maxwell-Boltzmann della velocità di equilibrio dei fotoni

$$\text{flusso}(\nu) = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/kT} - 1} \text{W/m}^2/\text{sr/Hz}$$

Oggi i fotoni mantengono questa perfetta distribuzione di equilibrio perché erano in equilibrio nel passato, dato che il presente l'universo è, in media, troppo diluito e freddo, e l'interazione elettromagnetica non è sufficiente ad ottenere e mantenere l'equilibrio. Questo fu raggiunto nel passato quando l'universo era molto più denso e caldo dell'attuale ed i fotoni interagivano molto efficacemente con la materia.

La temperatura odierna della CMB è $T = 2.752 \pm 0.002^\circ\text{K}$. Nell'espressione di sopra k , h e c sono, rispettivamente, le costanti di Boltzmann, di Planck e la velocità della luce nel vuoto.

risultati sperimentali sul flusso di neutrini dal nostro sole e su quello dovuto ai raggi cosmici che colpiscono i nuclei nell'atmosfera della terra hanno chiaramente mostrato che i neutrini sono soggetti a oscillazione di sapore durante la loro propagazione [5]. Questa è una chiara evidenza che ci sono tre stati di neutrini con masse diverse, e tre *autostati di sapore* ν_e , ν_μ and ν_τ che sono

combinazioni lineari di questi autostati di massa. Questo effetto è ben descritto dalla meccanica quantistica. Sin dall'ipotesi di de Broglie nel 1924, è noto che le particelle hanno un comportamento tipico delle onde, in particolare possono interferire e, in analogia con le onde elettromagnetiche, anche le onde di *materia* soddisfano il principio della sovrapposizione lineare durante la loro

3) **La nucleosintesi del Big Bang.** Si ritiene che tutti gli elementi che si trovano in natura sulla terra e nei pianeti attorno siano stati prodotti nelle stelle, durante la loro vita, oppure, per stelle molto massive, durante la loro catastrofica fine come Supernovae. Tuttavia, la quantità di ^4He , un prodotto tipico dei processi di fusione nucleare che hanno luogo nelle stelle, prevista dai nostri modelli, è molto più piccola di quella osservata. Qual'è la causa di questo eccesso? La risposta è ancora radicata nell'idea che nel passato l'universo fosse molto più caldo e denso, e, in un'epoca particolare, funzionasse come un enorme reattore nucleare a fusione. Il prodotto principale di questa fase è il nucleo di ^4He , ma furono prodotti anche, ^3He e ^7Li . Questa è la nucleosintesi del Big Bang (BBN), un tipico fenomeno di non equilibrio: più l'universo si espande più si riduce la probabilità che avvengano processi nucleari, e, ad un certo momento l'abbondanza delle specie nucleari si congela. C'è un modo semplice per comprendere questo fenomeno: se i processi di fusione non si fossero arrestati tutte le specie nucleari sarebbero evolute in ^{56}Fe , il nucleo con la massima energia di legame per nucleone. La BBN implica un complicato insieme di reazioni nucleari, rappresentate nella Figura 3.

propagazione.

In realtà le oscillazioni sondano le differenze tra i quadrati delle masse dei tre tipi di neutrini m_i , $i = 1, 2, 3$, che oggi sono note con l'accuratezza di poche parti per cento

$$\sqrt{m_2^2 - m_1^2} = 0.009 \text{ eV}$$

$$\sqrt{m_3^2 - m_2^2} = 0.05 \text{ eV}$$

Questo significa che almeno due neutrini hanno massa diversa da zero, mentre lo stato più leggero potrebbe essere privo di massa. Infatti, quello che le oscillazioni non possono indicarci è il valore assoluto della scala di massa dei neutrini m_0 .

Ciò che non conosciamo

i) Scala delle masse dei neutrini.

Il limite più stringente della scala della massa dei neutrini, cioè il valore della massa m_0 del più leggero dei neutrini, viene dall'analisi dello spettro degli elettroni emessi nel decadimento del nucleo ^3H . Dalla conservazione dell'energia, la massima energia cinetica dell'elettrone emesso è data dal Q valore del processo meno la massa del neutrino più leggero. Il limite attuale è di circa 2 eV. Come vedremo, informazioni indirette

su m_0 vengono dalla cosmologia. Il valore di m_0 definisce la scala nell'evoluzione dell'universo quando i neutrini diventano particelle non-relativistiche, ovvero quando la loro tipica energia cinetica diventa più piccola della loro massa. Quando avviene questo fenomeno, i neutrini cominciano ad essere sensibili alla buche di potenziale dei campi gravitazionali che circondano le strutture inhomogenee in formazione (galassie, cluster), e contribuiscono alla loro formazione.

Il fatto che, in ogni caso, la massa del neutrino elettronico non possa eccedere pochi eV, cioè che sia vari ordini di grandezza inferiore della massa del più leggero leptone carico (l'elettrone ha una massa di 0.511 MeV) ha sconcertato la comunità dei fisici per decenni, e sta ancora ispirando una delle principali linee di ricerca sia dal punto di vista teorico che sperimentale. Al momento, forse, la più elegante soluzione teorica per capire perché i neutrini siano così speciali è il meccanismo *see=saw*. Dato che sono particelle neutre, i neutrini sono speciali nel senso che, per loro, particelle e antiparticelle potrebbero coincidere. Per gli elettroni questo sarebbe impossibile dato che l'anti-elettrone (positrone) ha carica elettrica opposta, e questo avviene anche per gli altri leptoni carichi

e per i quark! Questo apre la possibilità al fatto che i neutrini possano essere particelle di Majorana e che i valori delle loro masse, insieme con il celebrato meccanismo di Higgs, siano legate ad effetti fisici che hanno luogo ad una nuova grande scala energetica.

ii) Dirac o Majorana?

Nel 1937 Ettore Majorana [6] trovò che ipotizzare l'esistenza di particelle che coincidono con le loro antiparticelle è perfettamente consistente con la teoria speciale della relatività di Einstein. È noto che queste particelle possono essere descritte da *spinori di Majorana*, mentre le particelle che sono diverse dalle loro antiparticelle sono collegate agli *spinori di Dirac*. Tutte le particelle elettricamente cariche sono necessariamente particelle di Dirac, ma i neutrini in questa situazione sono speciali. Parlano dialetto siciliano (Majorana era di Catania) o il perfetto inglese di Cambridge (Dirac)? Non conosciamo ancora la risposta, ma abbiamo uno strumento sperimentale per affrontare il problema: il doppio decadimento beta senza neutrini. Trovare l'evidenza sperimentale per questo processo estremamente raro sarebbe un chiaro segnale che i neutrini sono particelle di Majorana. Per una rassegna in questo volume si veda [7]. Vorrei mettere in evidenza che se si trovasse che il neutrino è di tipo Majorana questo non sarebbe solo l'indicazione di una soluzione del perché i neutrini sono così leggeri, per il già menzionato meccanismo see-saw. Sarebbe, forse, la sola evidenza, al momento, per un nuovo fenomeno che non può essere descritto in termini di ciò che comprendiamo delle interazioni fondamentali, quello che nel gergo si definisce come *Fisica oltre il modello Standard*.

iii) Ci sono altre specie di neutrini?

Il numero magico tre per le specie di neutrini è doppiamente radicato. Da una parte si prevede che sia tre per la chiara evidenza che i neutrini sono accoppiati con i loro partner leptoni carichi, e che ci sono tre specie di leptoni carichi. D'altra parte, qualsiasi nuova specie di neutrini che interagisse mediante l'interazione debole sareb-

be stata trovata nell'analisi del decadimento dello Z^0 nell'esperimento LEP al CERN. Questo non è quello che dicono i dati: il risultato di LEP è che il numero di neutrini *attivi*, cioè che interagiscono debolmente, è $2.98 \pm 0.07(stat) \pm 0.07(syst)$ (dove *stat* è l'incertezza statistica e *syst* quella sistematica). Teoria ed esperimento combaciano perfettamente! In ogni caso la fantasia dei teorici va sempre oltre. Non solo! In effetti, c'è un gruppo di dati sperimentali anomali, vedi [8], che sembrano suggerire che ci possano essere altre particelle neutre, che interagiscono solo gravitazionalmente (si usa normalmente la parola neutrini *sterili* che indica che non sono sensibili ad alcuna interazione nota tranne quella gravitazionale, come abbiamo detto.) Devo dire che la situazione degli esperimenti di laboratorio è, al momento, piuttosto confusa. Vedremo che qualche informazione indipendente può provenire dalla cosmologia.

Cosmologia

La cosmologia è lo studio quantitativo delle proprietà e dell'evoluzione dell'universo come entità unica. La scoperta da parte di Hubble nel 1929 della relazione tra distanza e spostamento verso il rosso degli spettri atomici e molecolari provenienti da galassie lontane fu il primo chiaro indizio in favore dell'ipotesi di un universo in espansione, che può essere ben descritta in termini della soluzione di Friedmann e Lemaître delle equazioni di Einstein. Non fosse per i risultati ottenuti da Einstein, non avremmo una nuova prospettiva per la descrizione della gravità e della cosmologia.

Alla base del modello di Friedmann e Lemaître c'è l'osservazione empirica che, su grande scala, l'universo è notevolmente omogeneo (le osservazioni sono indipendenti dalla posizione dell'osservatore) ed isotropo (le osservazioni sono le stesse in ogni direzione spaziale). Questo fatto sperimentale è normalmente chiamato *Principio Cosmologico*. Ipotizzando che il nostro punto di osservazione non sia privilegiato, si è naturalmente condotti alla conclusione che tutte le osservazioni fatte da punti differenti nell'universo debbano sembrare all'incirca le medesime

e siano indipendenti dalla direzione di osservazione. Omogeneità e isotropia selezionano una forma unica della metrica dello spazio-tempo, l'ingrediente principale della teoria di Einstein. I modelli cosmologici possono quindi essere studiati dopo aver specificato il contenuto di materia che agisce come sorgente di curvatura. I risultati possono essere confrontati con dati astrofisici che negli ultimi decenni hanno raggiunto una notevole precisione.

In realtà, il Principio Cosmologico funziona solo su scale maggiori di 100 Mpc. Si tratta comunque è un'ipotesi molto forte. Molte osservazioni, come la distribuzione nel cielo della radiazione cosmica di fondo (CMB), identificata da Penzias e Wilson nel 1964, mostrano disomogeneità molto piccole, così che possono essere trattate come perturbazioni di un modello di riferimento che è omogeneo ed isotropo.

L'idea di un universo in espansione conduce alla non ovvia conclusione che tutta la materia fosse caratterizzata nel passato da densità e temperature maggiori di quelle attuali, fino ad una singolarità iniziale dove ci si aspetta che gli aspetti quantistici della gravità fossero importanti, al punto tale che per una descrizione di questa fase si è forzati ad abbandonare la teoria classica di Einstein. Al di là di questo poco compreso momento iniziale, il modello di *Big Bang* caldo fa delle previsioni per molte quantità osservabili che possono essere verificate sperimentalmente. Questo modello è autosconsistente ove si tenga conto del fatto che nei primi tempi l'espansione fosse accelerata durante una fase nota come. Questa fase è oggi un pilastro del modello cosmologico standard e prevede precisi segnali che influenzano l'evoluzione successiva, come il fatto che le disomogeneità si sviluppano da qualche seme iniziale prodotto durante questa fase dell'evoluzione, o che i fotoni della CMB venendo da direzioni molto diverse condividono la stessa distribuzione in energia, a meno di piccolissime fluttuazioni dell'ordine di 10^{-5} , come rilevato per primo dal satellite COBE.

La validità del modello del Big Bang caldo è profondamente collegata all'osservazione che i fotoni della CMB sono caratterizzati da una distribuzione di corpo nero notevolmente accurata, oggi con una temperatura $T = 2.725 \pm 0.002$ K. Questo è un chiaro segnale che i fotoni erano,

anticamente, in equilibrio termodinamico con altre particelle (elettroni, nuclei ecc.) e quindi, che l'universo fosse un plasma di nuclei, leptoni carichi, fotoni, neutrini ecc. In altre parole, andando indietro nel tempo ed immaginando di prendere istantanee delle condizioni dell'universo, lo troveremmo molto più caldo e denso di quanto sia oggi. Lo strumento naturale per descrivere le proprietà di questo plasma è il meccanismo dell'equilibrio statistico. Finché le interazioni tra le particelle sono abbastanza forti da garantire condizioni di equilibrio, come succede per esempio agli atomi o alle molecole di un gas contenuto in una scatola, è possibile introdurre il concetto di temperatura e descrivere l'evoluzione temporale di tutte le specie usando il meccanismo dell'equilibrio statistico, o la sua controparte macroscopica, l'equilibrio termodinamico.

Tuttavia, il fatto empirico che i sistemi evolvono sempre spontaneamente verso le configurazioni di equilibrio è spiegabile se il sistema è imperturbato. Nell'universo, se il tasso di espansione è troppo veloce, le particelle non riescono, in certe epoche, a raggiungere l'equilibrio. Questa osservazione è cruciale per poter spiegare la produzione di nuclei leggeri durante la nucleosintesi primordiale e l'abbondanza primordiale di barioni (protoni e nuclei leggeri quali ^3He and ^4He) o la densità primordiale di densità di materia oscura. Per descrivere tutti questi fenomeni, è necessario abbandonare l'equilibrio termodinamico e usare la teoria cinetica, uno strumento matematico che descrive come un sistema fatto di particelle evolva nel tempo, ed eventualmente tenda a raggiungere una configurazione di equilibrio. Fu sviluppata da Ludwig Boltzmann in un suo articolo famoso nel 1872. Il suo approccio consiste nel guardare l'evoluzione temporale della funzione di distribuzione delle particelle come soluzione di un'equazione integro-differenziale, conosciuta oggi come equazione di Boltzmann.

Perché dovremmo usare la teoria cinetica per descrivere (almeno alcune) caratteristiche dell'evoluzione dell'universo? La ragione è che, come abbiamo detto, durante l'espansione, l'equilibrio non è garantito. Nell'universo in espansione ci sono due effetti in competizione nell'evoluzione verso l'equilibrio: espansione e processi di interazioni tra particelle. Il tasso di espansione è inserito nel parametro di Hubble H , mentre

l'equilibrio è stabilizzato dai processi di interazione come dalla diffusione, che ridistribuisce gli impulsi delle particelle, e da quei processi in cui il numero di particelle di una data specie non è conservato, che conduce verso l'equilibrio tra specie differenti. Poiché l'universo è in espansione, l'equilibrio è mantenuto se il tasso di interazioni microscopiche è maggiore del tasso di espansione. In questo caso l'evoluzione del sistema è *quasi statica*, e può essere considerata come una sequenza di stati in equilibrio. La condizione di equilibrio è quindi $\Gamma \gg H$, dove Γ è il tasso di interazione tipica (il numero di processi di interazione per unità di tempo).

In realtà le epoche più interessanti nella storia dell'universo sono quelle in cui l'equilibrio non è raggiunto! L'evidenza di una grande quantità di *materia oscura*, l'assenza di antimateria nell'universo osservabile, la formazione di nuclei pochi secondi dopo il big bang, la stessa CMB: tutti questi fenomeni sono chiare indicazioni di diverse fasi prive di equilibrio. Per ultimo il fondo cosmico di neutrini (CNB), l'analogo per i neutrini della traccia fossile di fotoni delle prime fasi dell'espansione, porta un chiaro segnale di un fenomeno di non equilibrio: quando l'età dell'universo era di appena pochi secondi, le interazioni deboli divennero troppo lente per mantenere i neutrini in equilibrio termico con fotoni e coppie di elettroni e positroni.

Un'introduzione molto incompleta, breve, ma più quantitativa è presentata negli inserti *distillati di Cosmologia* I e II. Nella sezione seguente, utilizzando la teoria di Einstein, descriverò la meccanica statistica dell'equilibrio, la teoria cinetica ed il Modello Standard delle interazioni elettro-deboli e come ci si aspetta che i neutrini riempiano l'universo durante la sua evoluzione. Tuttavia, ed in maniera più interessante, confrontando le osservazioni con la teoria possiamo guadagnare informazioni ulteriori su proprietà sconosciute.

Prima di discutere questo punto, farò una piccola digressione e riassumerò l'attuale comprensione delle diverse fasi dell'evoluzione dell'universo. Ogni fase prevede un diverso tasso di espansione e può essere limitata dai dati osservativi. Per esempio, un'epoca dominata dalla materia, cioè quando le particelle non relativistiche sono la sorgente principale del campo gravitazio-

nale, sarebbe un disastro durante la nucleosintesi del Big Bang (BBN).

i) Inflazione.

Le equazioni di Einstein prevedono che *particelle*, sia relativistiche come i fotoni, o non relativistiche come la sconosciuta materia oscura o i barioni, rappresentino la sola sorgente di gravità. L'espansione dell'universo sarà sempre rallentata, non diversamente dalla mela di Newton, che può viaggiare fino a distanze infinite, ma sempre sentendo l'attrazione della terra. Innocente come può apparire, questa osservazione solleva problemi concettuali su tutta la storia dell'universo. Nei primi anni '80 dell'ultimo secolo si trovò che una fase iniziale *accelerata* può risolvere questi problemi, al prezzo di ipotizzare che vi sia una sorgente di gravità che accelera lo sviluppo: una qualche entità la cui densità di energia non cambi nel tempo, una *costante cosmologica* o qualcosa di molto simile. Fu scoperto, in seguito, che questo scenario, chiamato inflazione, è anche capace di fare semplici previsioni su come inhomogeneità casuali iniziali si possano evolvere in tutte le strutture che noi vediamo nel cielo (galassie, cluster, ecc).

ii) Dominio della radiazione.

Per un periodo di tempo relativamente lungo rispetto ai tempi tipici delle interazioni tra particelle tutto era radiazione, un fluido fatto di particelle con una grande energia cinetica rispetto alle loro masse. C'è una sonda notevolmente robusta che ci dice che questo è quello che accadde quando l'età dell'universo era compresa tra pochi secondi e pochi minuti: la nucleosintesi del Big Bang. La formazione di nuclei leggeri non sarebbe consistente con i dati osservati se a quel tempo l'universo si fosse espanso con un tasso anche di poco diverso (basta qualche percento) rispetto a quello determinato dalla radiazione. La sola domanda ancora rimasta senza risposta è quanta radiazione fosse presente a quel tempo. Discuteremo quel punto successivamente, poiché è importante capire quanti neutrini ci fossero nell'universo in quella epoca.

iii) Dominio della materia.

La radiazione si muove troppo velocemente per permettere formazione di strutture. Galassie e cluster si sarebbero formate con un tasso molto più lento di quello osservato, se particelle relativistiche fossero responsabili del bilancio energetico dell'universo, perché non sono facilmente catturate in buche di potenziali gravitazionali. Sono troppo veloci, e l'effetto netto è quello di sopprimere la crescita di strutture locali. Evidenze a scale di lunghezza diverse (le curve di rotazione delle galassie, il deficit di massa nei cluster di galassie se consideriamo solo la materia luminosa, le proprietà della CMB, ecc.) tutto conduce all'idea che ci sono particelle pesanti che sono state non relativistiche negli ultimi miliardi di anni. Insieme ai barioni, queste particelle di *materia oscura* rappresentano il contributo principale alla parte destra dell'equazione di Einstein (si veda distillato di Cosmologia II) fino ad epoche molto recenti, ed essendo molto lente, sono le principali responsabili della crescita di inhomogeneità, conducendo ai complicati disegni che noi vediamo nel cielo su scale sufficientemente piccole (galassie, cluster, filamenti e grande vuoti). La ricerca della natura della materia oscura è ancora aperta, ma tutti i dati sperimentali indicano la sua esistenza.

iv) La costante gravitazionale vince oggi.

Questo è un argomento per un altro articolo. I dati astrofisici, soprattutto l'osservazione delle velocità di Supernovae molto distanti di tipo Ia, ci dicono che, oggi e in epoche recenti, l'espansione dell'universo è accelerata. Poiché non è rilevante per l'argomento della nostra analisi, non descriverò oltre questo interessante argomento, ma menzionerò come la scienza sia ancora in debito con la filosofia.

Neutrini e cosmologia

Un universo pieno di neutrini

La descrizione standard dei primi istanti dell'universo è quella di un plasma denso e caldo co-

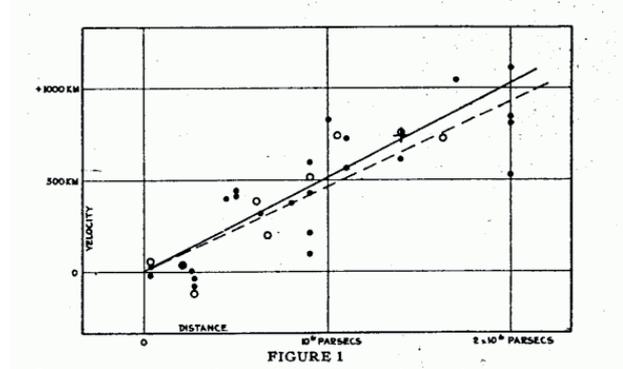


Figura 1: La figura originale della velocità di recessione in funzione della distanza osservata nell'articolo di Hubble del 1929.

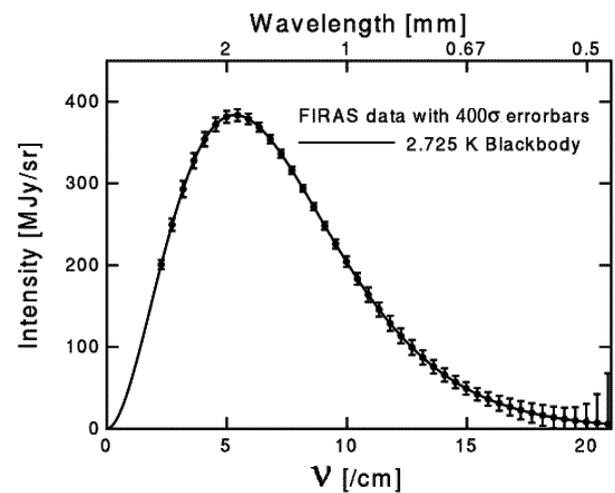


Figura 2: La distribuzione di frequenza della CMB osservata dall'apertura FIRAS (punti) confrontata con una distribuzione di corpo nero (linea continua).

stituito da tutte le particelle note (e presumibilmente anche di altre particelle non ancora note!) che a quel tempo erano tutte relativistiche. Le loro interazioni continue presto portarono ad uno stato di equilibrio termodinamico. A causa dell'espansione, come avviene in un gas che si espande in una scatola termodinamicamente isolata (espansione adiabatica), la loro temperatura T , la densità numerica, quelle di massa e di energia diminuivano. In particolare, la legge di conservazione dell'energia e pressione presentata nel riquadro *Distillato di cosmologia II: strumenti*, prevede che la temperatura diminuisca in proporzione inversa rispetto al fattore di scala, $T \times a(t) = \text{costante}$. Questo è noto come *spostamento verso il rosso*.

Lo spostamento verso il rosso è una quantità

Distillato di cosmologia II: strumenti

Nonostante i grovigli matematici della relatività generale, la teoria dell'interazione gravitazionale di Einstein è meravigliosamente semplice. Per una data distribuzione di masse, corpi, fluidi ecc., che producono un campo gravitazionale, la quantità fondamentale è rappresentata da una particolare struttura dello spazio-tempo che decide quale sia il percorso più breve tra due punti. Le equazioni di base, analoghe alla legge di gravitazione universale di Newton, sono dette equazioni di Einstein e sono espresse come:

$$\text{gravità} = 8 \pi G_N (\text{densità di energia} / \text{pressione}).$$

La parte sinistra (gravità) è l'incognita, ed il suo valore è definito una volta che la parte destra venga specificata, cioè quando siano date le proprietà del campo gravitazionale. Queste ultime sono determinate dalla densità di energia ρ e dalla pressione P (quest'ultima non è sorgente di gravità nella teoria di Newton) o, più in generale, attraverso il tensore *stress-energia*. Naturalmente G_N è la costante di gravitazione di Newton. Tutte le *accettabili* sorgenti di gravità devono obbedire all'analogo della conservazione di energia della meccanica classica che, nelle parole di Einstein, è la *conservazione covariante* del tensore stress-energia.

Una particella di prova messa in un campo gravitazionale, che non sia sottoposta ad altre forze, si muove seguendo il percorso più breve, un moto *geodesico* che dipende dalla metrica, definita dalla soluzione dell'equazione di Einstein, ottenuta quando le proprietà dei corpi celesti e le loro distribuzioni sono assegnate. Fondamentalmente è lo stesso modo di ragionare della costruzione di Newton: sorgente \rightarrow gravità \rightarrow moto di un corpo di prova, ma presentato in un linguaggio diverso, con la, per niente ovvia, nuova caratteristica legata al fatto che esista una velocità massima per la propagazione dei segnali (la velocità della luce).

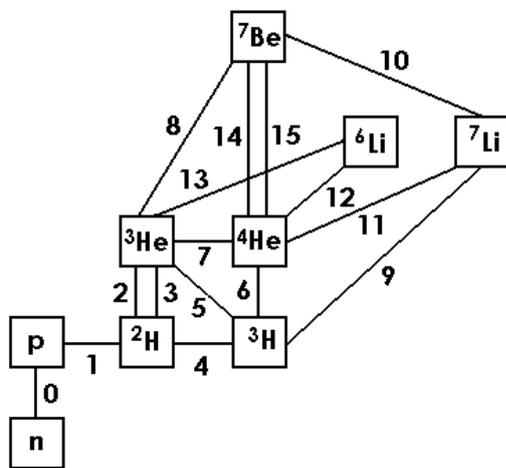


Figura 3: Un diagramma che illustra le principali reazioni nucleari durante la nucleosintesi del Big Bang.

più fondamentale della diminuzione della temperatura, cioè dell'energia cinetica media delle specie relativistiche in equilibrio termico. Questo perché lo spostamento verso il rosso dello spettro di emissione sussiste anche per particelle

che non siano in equilibrio termico. In un universo in espansione, infatti, il momento lineare di una particella diminuisce sempre come $a(t)^{-1}$ sino a quando è in caduta libera, cioè segue il percorso più piccolo nella metrica di Friedmann, Robertson, Walker and Lemaître. È da notare che, poiché $a(t)^3$ è il modo con il quale scala il volume nell'universo in espansione, abbiamo $T \sim V(t)^{\gamma-1}$, e l'indice adiabatico è $\gamma = 2/3$, da confrontare con il risultato standard per un gas monoatomico, $\gamma = 5/3$.

Prima o poi, le particelle diventeranno non relativistiche, a meno che non siano prive di massa, questo poiché la loro energia cinetica diminuisce con l'espansione. Quando il valore della loro energia cinetica diventa inferiore a quello della loro massa a riposo, la loro annichilazione diventa più conveniente per minimizzare l'energia. Consideriamo, ad esempio, il muone. All'equilibrio l'annichilazione di una coppia di muoni $\mu^+ + \mu^- \rightarrow \gamma + \gamma$ è compensata dal processo inverso $\gamma + \gamma \rightarrow \mu^+ + \mu^-$, che stabilisce l'equilibrio cinetico tra muoni e fotoni. Ad un certo punto, tuttavia, i fotoni non hanno energia suf-

Friedmann, Roberston, Walker e Lemaître (FRWL), furono i primi ad applicare la teoria di Einstein alla dinamica dell'Universo, sfruttando, in modo semplice, una singola ipotesi: che su una scala sufficientemente grande, il nostro universo sia abbastanza omogeneo e isotropo. Questa idea, nota come *Principio Cosmologico* è sufficiente per definire la forma della metrica a meno di una funzione sconosciuta, il fattore di scala $a(t)$ che dipende solo dal tempo e fissa, ad ogni tempo, la distanza tra due osservatori: se $a(t)$ è una funzione crescente del tempo, la loro distanza cresce nel tempo: questo è un universo in espansione.

Nel modello FRWL la teoria di Einstein si riduce a due semplici relazioni: la legge di Friedmann

$$H^2 = \left(\frac{1}{a(t)} \frac{da(t)}{dt} \right)^2 = \frac{8\pi G_N}{3} \rho - \frac{k}{a^2}$$

che afferma che il tasso di espansione (H) dell'Universo dipende dalla densità di energia della materia, e la legge di conservazione dell'energia e pressione

$$\frac{d\rho}{dt} + 3H(\rho + P) = 0$$

Quella che Hubble scoprì essere la costante che lega la distanza di una data sorgente astrofisica con la sua velocità di regressione è, nella cosmologia di FRWL, il valore attuale della costante H . Un'ultima osservazione circa il parametro k , noto come curvatura dello spazio. È un parametro libero che può essere positivo, negativo oppure nullo. Il suo valore decide l'eventuale destino di un dato Universo. Per $k \leq 0$ l'espansione continuerà per sempre, poiché la parte destra della prima equazione, e quindi la velocità di espansione, è sempre positiva. Per $k > 0$ c'è un valore di a per cui l'espansione si arresta, quando la parte destra è nulla, e l'universo inizia una nuova fase di collasso. Grazie agli esperimenti degli ultimi dieci anni, in particolare WMAP e Planck [9], sappiamo che il nostro universo è molto vicino ad un universo spazialmente piatto, cioè $k = 0$.

ficiente per produrre coppie di muoni, e solo il primo processo diventa possibile. Tutti i muoni si trasformano allora in fotoni, e così per elettroni, neutrini ecc. con processi simili. Ad esempio, oggi, non rimane alcun muone primordiale nell'universo.

Per una ragione che apparirà immediatamente chiara, la fase in cui l'universo aveva temperatura dell'ordine del MeV è particolarmente rilevante. Un'istantanea del contenuto di particelle relativistiche a quell'età mostra che rimangono solo coppie di elettroni e positroni, fotoni e neutrini, le specie che sono ancora relativistiche. Sono presenti anche protoni e neutroni sebbene siano molto massivi, in una piccolissima frazione rispetto all'abbondanza dei fotoni. Il rapporto tra densità di fotoni e barioni può essere misurato da due osservabili cosmologiche, BBN e CMB, e si trova che è $n_b/n_\gamma \sim 10^{-9}$. La ragione per la qua-

le i barioni stanno ancora popolando l'universo è legata al fatto che la conservazione del numero barionico, una simmetria fondamentale delle interazioni a scale di bassa energia, li protegge da una totale annichilazione in specie più leggere, (come invece avviene nel caso dei muoni!). Ritorno su questo punto. All'epoca in cui $T \geq$ pochi MeV, la radiazione domina grandemente il bilancio energetico totale. La densità di energia dei fotoni all'equilibrio può essere semplicemente stimata, e assume l'espressione standard della distribuzione di Stefan-Boltzmann (la densità di energia è proporzionale a T^4)

$$\rho_\gamma = 2 \times \frac{\pi^2 k^4}{30 \hbar^3 c^3} T^4$$

dove $\hbar = h/2\pi$ è la costante di Plank ridotta. Elettroni e positroni e neutrini contribuiscono

per simili quantità

$$\rho_{e^\pm} = 2 \times 2 \times \frac{7}{8} \frac{\pi^2 k^4}{30 \hbar^3 c^3} T^4$$

$$\rho_\nu = 3 \times 2 \times \frac{7}{8} \frac{\pi^2 k^4}{30 \hbar^3 c^3} T^4$$

Il fattore $7/8$ è dovuto alle diverse proprietà statistiche di elettroni e neutrini rispetto ai fotoni. Essendo i primi dei fermioni soddisfano il principio di esclusione di Pauli, mentre non lo fanno i fotoni. I fattori numerici sono dovuti alle possibili orientazioni dello spin (2), la somma sugli stati di particella e antiparticella (2) e, per i neutrini, il numero di sapori (3). Oggi è una tradizione ben radicata ridefinire la densità di energia dei neutrini nel parametro *numero effettivo di neutrini* N_{eff} , in modo che la densità di energia delle specie totalmente relativistiche ρ_R quando e^\pm sono ancora particelle relativistiche (vedi dopo) diventa

$$\rho_R = \rho_{e^\pm} + \rho_\gamma \left(1 + \frac{7}{8} N_{\text{eff}} \right)$$

Questa notazione può sembrare un po' barocca, poiché dalle formule presentate sopra sembra ovvio che $N_{\text{eff}} = 3$. Tuttavia questo succede se: i) c'è solo una specie di neutrini; ii) sono in completo equilibrio con i fotoni (cioè sono alla stessa temperatura); iii) non ci sono strutture esotiche nella loro distribuzione in funzione del momento lineare (dato da $h\nu/c$ in accordo con la relazione di de Broglie). Questo è dato dall'analogia con la funzione di corpo nero per i fotoni, con un cruciale segno più nel denominatore che indica che sono particelle fermioniche, cioè soddisfano il principio di esclusione di Pauli (confronta con la distribuzione di fotoni nel riquadro *Distillato di cosmologia I: evidenze*).

$$\text{flux}_{\text{neutrino}}(\nu) = 3 \times \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/kT} + 1}$$

in unità di $\text{W m}^{-2} \text{sr}^{-1} \text{Hz}^{-1}$.

In altre parole, N_{eff} è un modo semplice per parametrizzare quello che non conosciamo del comportamento dei neutrini nell'era del MeV. Mostro come il valore di N_{eff} possa essere fortemente vincolato dai dati sperimentali.

Abbiamo già detto che i neutrini sono tenuti in equilibrio con altre specie attraverso interazio-

ni deboli. Alla temperatura dell'ordine di pochi MeV il loro tasso tipico scende sotto il parametro di espansione di Hubble, e i neutrini si *disaccoppiano*. Da questo momento in poi, i loro effetti sulla materia dell'universo si manifesteranno solo attraverso la gravità. Non essendo più soggetti a processi di diffusione, essi si propagano liberamente e la loro distribuzione è congelata, tranne che per lo spostamento verso il rosso della loro temperatura. Questo significa che l'espressione del loro flusso presentata sopra è quella che noi misuriamo oggi, ammesso che si scali la temperatura al valore attuale.

Poco dopo il disaccoppiamento dei neutrini, gli elettroni e i positroni cominciano a sentirsi intollerabilmente pesanti. A $T \sim 0.511$ MeV, si annichilano in coppie di fotoni in processi $e^+e^- \rightarrow \gamma\gamma$, *ma non in neutrini*¹, poiché il processo analogo $e^+e^- \rightarrow \nu\bar{\nu}$ è un processo debole che in questa epoca è troppo lento per avvenire. Questo significa che tutta l'energia accumulata in elettroni e positroni scalda i fotoni ma non i neutrini. Dopo questo istante la temperatura del neutrino T_ν è quindi inferiore alla temperatura del fotone T_γ per un fattore che può essere calcolato usando la conservazione dell'entropia durante l'espansione, $T_\nu = (4/11)^{1/3} T_\gamma$. Dato che oggi conosciamo abbastanza accuratamente T_γ , possiamo inferire la temperatura dei neutrini nell'universo attuale

$$T_\nu(\text{oggi}) = (4/11)^{1/3} 2.752 \text{ K} = 1.964 \text{ K}$$

che è estremamente bassa! Conseguentemente, il numero di neutrini e antineutrini per unità di volume è dell'ordine di 56 cm^{-3} per sapore, inferiore a quello dei fotoni della CMB (411 cm^{-3}), ma notevolmente grande (rappresenta il maggior flusso di origine astrofisica, più abbondante, ad esempio, di quello dei neutrini provenienti dal sole!). Se il modello di Big Bang caldo è corretto, siamo circondati da un intenso flusso di neutrini fossili che, in principio, come i fotoni della CMB trovati da Penzias e Wilson, potrebbe essere rivelato da qualche *antenna di neutrini*. Sfortunatamente questo è molto difficile perché i neutrini del fondo cosmico (CNB) sono molto freddi (pochi Kelvin) e, quindi, trasportano poca

¹C'è un piccolo rilascio di energia in neutrini dell'ordine del percento.

energia cinetica da trasferire in processi di interazione con il rivelatore. Inoltre, interagiscono solo mediante l'interazione debole, quindi il tasso di interazione è estremamente basso. In ogni caso, una prova diretta dell'esistenza del CNB sarebbe di importanza straordinaria, essendo, come la CMB, la legge di Hubble e la nucleosintesi primordiale una previsione chiave del modello del Big Bang. Molti ricercatori studiano oggi possibili tecniche di rilevazione di tali neutrini. Vi sono idee interessanti che si spera possano essere implementate sperimentalmente in un futuro non troppo lontano.

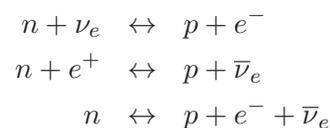
I neutrini e la sintesi degli elementi primordiali

C'è un'altra chiara evidenza che i neutrini hanno popolato l'universo a partire da pochi secondi dopo il big bang: la nucleosintesi dal Big Bang (BBN). La BBN si è attivata in un breve periodo della storia dell'universo durante la quale nuclei leggeri, soprattutto ^4He , ^2H , ^3He e, in misura minore, ^7Be e ^7Li furono prodotti da reazioni di fusione, come avviene oggi, in genere mediante diversi meccanismi nucleari, all'interno delle stelle. Ad energie sufficientemente alte, maggiori delle tipiche energie di legame, la cui scala è dell'ordine di 1 - 10 MeV per nucleone, ogni specie nucleare prodotta come reazione di fusione potrebbe essere fissionata dal fondo di fotoni ad alta temperatura o da altri processi di fusione. Il fatto è che quando la temperatura del plasma nell'universo, quindi quella dei fotoni, è più grande di alcuni MeV, tutti i barioni sono nella forma di protoni o neutroni liberi, e non si ha la formazione di nuclei.

Da quanto ho detto prima (le particelle pesanti tendono a scomparire dal bagno termico per poter minimizzare l'energia totale), il lettore può essere sorpreso dal fatto che queste particelle pesanti (i nucleoni hanno energie dell'ordine del GeV) siano ancora in giro a queste temperature così basse. La ragione per la quale questo è possibile è legata al fatto che l'universo non è simmetrico nella composizione di materia-antimateria, in particolare contiene un maggior numero di barioni (protoni, neutroni) che di anti-barioni (antiprotoni, antineutroni). Infatti, protoni e nuclei leggeri, insieme con gli elettroni, costituiscono i

mattoni fondamentali per la costruzione di pianeti, stelle, galassie, ecc. ma non ci sono tracce di antiprotoni, anti-nuclei o positroni, se non quelli prodotti da interazioni ad alta energia di raggi cosmici o da sorgenti astrofisiche di particelle di alta energia. Questa osservazione fu considerata all'inizio come un'ipotesi *ad hoc* sulle condizioni iniziali dell'universo nel *big bang*. Fu chiarito in seguito, soprattutto grazie al fisico russo Andrej Sacharov, che una soluzione più elegante consisteva nell'ipotizzare che, in una qualche epoca molto antica, fosse attiva una interazione che violava la conservazione del numero barionico. Quando la temperatura scese con l'espansione, queste interazioni diventarono troppo deboli (il loro tasso era sceso sotto il tasso di espansione generato dal tasso di espansione di Hubble). Tutti gli altri processi, quelli che noi studiamo in esperimenti di laboratorio (elettromagnetici, deboli, forti e gravitazionali) non possono cambiare il numero iniziale di barioni (cioè il numero di barioni meno il numero di anti-barioni) perché tutte conservano il numero barionico, così come le interazioni elettromagnetiche conservano la carica. Per un valore della temperatura dell'ordine della massa del nucleone, i nucleoni e gli anti-nucleoni diventano particelle non relativistiche e la loro rapida annichilazione in particelle più leggere condurrebbe ad un universo senza barioni se non fosse per il piccolo eccesso di barioni sugli anti-barioni. Questa frazione di barioni sopravvive e rappresenta l'attuale contenuto globale di barioni dell'universo.

Dopo questa digressione, ritorniamo alla BBN. I neutroni ed i protoni sono tenuti in equilibrio dai processi deboli



Poiché i neutroni sono leggermente più pesanti dei protoni, a basse temperature, inferiori alla differenza tra le masse del neutrone e del protone $\Delta m_N = 1.29 \text{ MeV}$, queste reazioni avvengono preferibilmente nella direzione destra (sempre per minimizzare l'energia, poiché i protoni pesano leggermente meno) e tutti i neutroni scompaiono. Questo significa un universo con solo atomi di idrogeno. Il punto è che poco prima, in ma-

niera analoga a quanto è avvenuto per i processi deboli che mantengono i neutrini in equilibrio, anche questi processi diventano troppo lenti e il rapporto tra protoni e neutroni si *congela* ad un valore fissato, che può essere calcolato esplicitamente. È curioso che *proprio per i valori esatti* della costante di Fermi G_F , che regola il tasso di processi deboli, della costante di Newton G_N , che appare nella legge di Friedmann (si veda il riquadro *Distillato di cosmologia II: strumenti*), e di Δm_N , i neutroni sopravvivono, i nuclei si formano e quindi si evolvono strutture complesse, inclusi gli esseri umani. Proprio dopo questo congelamento dei processi deboli, cominciano a formarsi nuclei di deuterio. C'è un piccolo guadagno di energia in questo, dato che l'energia di legame nucleare è di 2.23 MeV. Le condizioni di temperatura e di densità in questa epoca permettono una serie di complicate reazioni nucleari che sono sufficientemente efficienti da produrre anche nuclei più pesanti fino al ${}^7\text{Be}$ e ${}^7\text{Li}$ (vedi la Figura 3). Ad un certo punto, tuttavia, la catena nucleare si ferma perché l'espansione dell'universo domina e, ad esempio, non si producono *metalli* (nel gergo astrofisico i metalli sono C, N, O, ogni elemento non prodotto nella BBN). L'evoluzione con la temperatura dei nuclei leggeri durante il BBN è mostrata nella Figura 4. Il parametro chiave per fissare il valore finale del loro numero è il già menzionato ammontare di barioni per fotone, chiamato η_b . È un parametro libero, legato all'asimmetria iniziale tra materia e anti-materia. In effetti, la BBN fu, storicamente, la prima maniera di definire il suo valore. La stima presente è, come detto prima, $\eta_b \sim 10^{-9}$. Per questo valore, le previsioni teoriche per l'abbondanza dei nuclei leggeri di origine primordiale sono in ottimo accordo con le osservazioni astrofisiche, in particolare le abbondanze di deuterio e ${}^4\text{He}$. Per l'abbondanza del ${}^3\text{He}$ al momento abbiamo solo un limite superiore. Lo stato del ${}^7\text{Li}$ è leggermente più complicato, e non è chiaro se le osservazioni stiano veramente misurando il suo valore *primordiale*.

Qual è il ruolo dei neutrini nella BBN? È doppio.

- i) I neutrini elettronici sono direttamente coinvolti nei processi deboli che mantengono neutroni e protoni in equilibrio. Se non ci fossero stati ν_e e $\bar{\nu}_e$ nell'universo, o se il loro

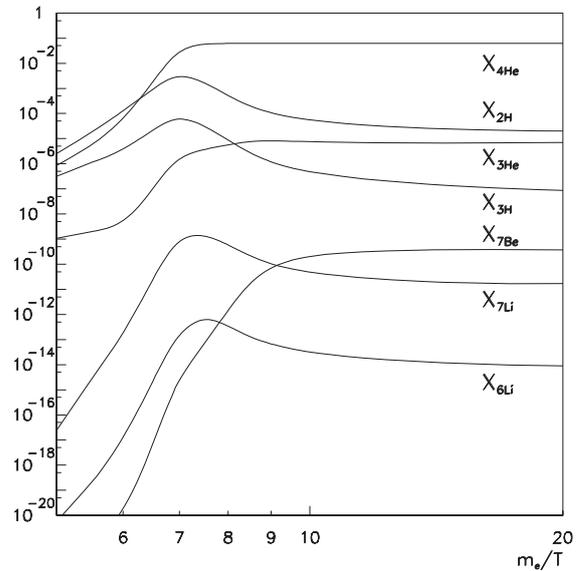


Figura 4: L'evoluzione delle abbondanze nucleari prodotte nella BBN in funzione di m_e/T , dove m_e è la massa dell'elettrone. Le quantità X_i sono le frazioni del numero di nuclidi i normalizzate al numero totale di protoni e neutroni nell'universo.

numero fosse stato troppo grande o troppo piccolo rispetto a quello discusso nella sezione precedente, le previsioni della BBN sarebbero state completamente errate.

- ii) Il tasso di espansione dell'universo, oltre al valore di G_N , dipende in modo cruciale dalla densità di energia ρ durante la BBN. Il primo risultato è che durante la BBN l'universo *dovrebbe essere dominato dalla radiazione*. Se la materia oscura, o anche peggio, una costante cosmologica, avessero determinato l'espansione, le abbondanze nucleari sarebbero completamente diverse da quelle che si osservano.

Noi vediamo che la densità di energia delle specie relativistiche può essere descritta da un solo parametro, N_{eff} . Se il suo valore è quello atteso (3), lo scenario della BBN funziona abbastanza bene. Questo significa che la CNB dovrebbe essere molto vicina alle nostre attese. Troppi neutrini, o troppo pochi, o qualche esotica caratteristica nel loro flusso, avrebbero rovinato il buon accordo tra teoria e dati sperimentali. Questi ultimi al momento escludono che N_{eff} possa valere 4 o 2, per esempio.

Quest'ultimo punto ci porta ad un'altra osservazione importante. Al di là dei tre neutrini standard, tutte le altre specie di particelle, ammesse che siano relativistiche, contribuiscono a N_{eff} . Questo perché la gravità (la legge di Friedmann) è cieca alle proprietà delle particelle ma è sensibile solo alla loro energia (e pressione). Abbiamo menzionato in precedenza che ci sono risultati di laboratorio che suggeriscono che ci potrebbero essere delle specie sterili di neutrini. Se fossero state prodotte nell'universo primordiale, e questo dovrebbe essere il caso se i risultati sperimentali fossero spiegati in termini di queste specie, il valore di N_{eff} dovrebbe essere maggiore di tre. Questo è sfavorito dai risultati della BBN. Vedremo come il limite su questi stati sterili sia anche maggiore se si usano i risultati della CMB.

L'impronta dei neutrini sulla radiazione cosmica di fondo (CMB) e sulle strutture a grande scala.

La CMB si è formata quando l'età dell'universo era di circa 380000 anni. In questa epoca, la temperatura dei fotoni era di una frazione di eV, e gli elettroni e i protoni (e i nuclei di ^4He , la specie più abbondante dopo i protoni) si ricombinarono per formare i primi atomi. Poco dopo, la diffusione di fotoni per il tramite di processi Thomson sui pochi elettroni rimasti liberi diventarono molto rari. Quindi l'universo diventò *trasparente* alla luce. I fotoni *si spostarono verso il rosso*, cioè perdettero parte della loro energia a causa dell'espansione dell'universo ed, eventualmente, oggi raggiungono i nostri rivelatori. È molto simile a quanto accade alla luce emessa dal Sole. I fotoni prodotti all'interno del sole diffondono molte volte, e sono continuamente assorbiti ed emessi. Solo alla superficie la densità degli ioni ed elettroni è sufficientemente bassa, e i fotoni sono quindi liberi di viaggiare fino alla terra. La superficie nel cielo, dalla quale riceviamo i fotoni della CMB è *l'ultima superficie di diffusione*, uno dei più antichi strumenti di osservazione che abbiamo per investigare l'evoluzione dell'universo, l'equivalente della superficie del Sole.

La radiazione della CMB è notevolmente isotropa. La luce sembra avere le stesse proprietà indipendentemente dalla direzione di arrivo e possiede una temperatura media comune T . In

particolare, la distribuzione di fotoni, come abbiamo già menzionato, è quella di un perfetto corpo nero. Ci sono, tuttavia, piccole fluttuazioni nella temperatura se si guarda ad angoli diversi nel cielo. Questo non è solo inatteso, ma è piuttosto una benedizione, poiché ci offre la possibilità di comprendere la ragione per cui l'universo non è *perfettamente omogeneo e isotropo*.

L'intendimento comune è ancora radicato nel paradigma dell'inflazione. Durante la fase iniziale dell'espansione, minuscole perturbazioni del potenziale gravitazionale su piccole distanze furono allungate a scale cosmologiche, e sono i semi primordiali di tutte le disomogeneità che noi osserviamo oggi, dalle anisotropie della CMB fino alle strutture delle galassie, dei cluster, ecc.

Così come è privo di significato chiedersi perché la Terra si trovi esattamente a quella distanza dal sole che permise alla vita di svilupparsi, è ugualmente privo di interesse cercare di capire perché osservando il cielo della CMB, i fotoni provenienti da una particolare direzione siano poco più caldi, o freddi, rispetto al valor medio della temperatura della CMB. Invece, è interessante studiare le proprietà statistiche delle fluttuazioni. Se avessimo molti universi da osservare, potremmo vedere che condividono caratteristiche comuni, anche se non identiche. La stella più vicina alla Terra non sarebbe α -Centauri, ma la *probabilità* che ci sia una stella vicina alla distanza di pochi anni luce sarebbe analoga a quella che si ha nel nostro universo osservabile. In altre parole, tutti i diversi universi, ammesso che siano stati generati con le stesse condizioni iniziali, sarebbero molto simili tra loro, analizzandoli in termini di probabilità di avere un certo numero di galassie, o di cluster, in un dato volume. Cosa è veramente significativo per comprendere le proprietà delle disomogeneità, che sono variabili stocastiche? Questo ci indica a quale classe di universi noi apparteniamo! Queste perturbazioni stocastiche hanno distribuzioni gaussiane? Qual è l'ampiezza del loro spettro quando ci muoviamo da piccole a grandi scale?

Le osservazioni ci dicono che tutto è stato originato da piccole disomogeneità gaussiane. Queste perturbazioni iniziali sono soggette a diversi destini, a seconda della loro lunghezza d'onda. Poiché l'universo si espande, e la velocità della luce è finita, c'è una distanza massima che

la luce può percorrere. Questa distanza è normalmente chiamata *orizzonte delle particelle*. Le perturbazioni di una data ampiezza, possono essere cambiate da meccanismi dinamici (diffusioni, smorzamenti, amplificazioni di instabilità gravitazionali) solo su scale di distanze inferiori a quella dell'orizzonte delle particelle. Tutti questi regimi sono chiaramente visibili nelle anisotropie della CMB. Questo perché la causalità implica che non ci siano meccanismi fisici che possono trasferire informazioni più rapidamente della velocità della luce.

L'osservabile fondamentale della CMB è la funzione di correlazione temperatura-fluttuazione. Per un dato punto nel cielo, identificato da due angoli θ e ϕ (per esempio ascensione retta e declinazione), si misura il valore attuale della temperatura della CMB $T(\theta, \phi)$, e definendo

$$\delta T(\theta, \phi) = \frac{T(\theta, \phi) - T}{T},$$

la funzione di correlazione

$$\overline{\delta T(\theta, \phi) \delta T(\theta', \phi')}$$

rappresenta l'eccesso di probabilità, confrontato con una distribuzione casuale di temperature, di trovare simili valori di T ad una data distanza $(\theta - \theta', \phi - \phi')$.

La barra in alto indica la media. Questa quantità è comunemente espressa in termini di coefficienti C_l che, tecnicamente, sono ottenuti facendo uno sviluppo in termini di polinomi di Legendre (si veda il riquadro *I coefficienti C_l e lo spettro di potenza $P(k)$*). La distribuzione sperimentale è mostrata nella figura 5 insieme ad una curva di fit. Ci sono tre caratteristiche che possono essere apprezzate da questa figura.

- i) Nel regime di piccoli l (grandi distanze angolari), lo spettro della CMB è piuttosto piatto. È ciò che rimane delle fluttuazioni primordiali prodotte dall'inflazione.
- ii) Nella regione intermedia, c'è una serie di picchi e valli. Queste sono dovute alla dinamica durante l'ultima fase della diffusione. Il plasma di fotoni, elettroni e barioni oscilla sotto l'effetto di due forze in competizione: la pressione dei fotoni da una parte e la gravità dei barioni ed elettroni dall'altra.

- iii) A grandi valori di l , i valori dei coefficienti C_l scendono esponenzialmente. I fotoni diffondono a piccole distanze per moto Browniano. Questo conduce ad uno smorzamento delle inhomogeneità e dei valori di C_l .

Ci sono due modi attraverso i quali i neutrini possono influenzare questo percorso: attraverso le loro proprietà di fondo cosmico e attraverso le loro perturbazioni. Con le proprietà di fondo cosmico intendo le loro proprietà omogenee, come la densità media, o proprietà fisiche, come la loro massa. Ma i neutrini, come tutte le specie, presentano alcune disomogeneità che forniscono un contributo sia alla crescita di strutture materiali o al loro smorzamento. Qui sotto c'è un riassunto dei loro effetti sullo spettro di potenza della CMB.

- i) Durante la formazione della CMB (quando la temperatura dei fotoni è una frazione di eV) i neutrini contribuiscono alla densità di energia da radiazione (la N_{eff} che abbiamo introdotto prima), a meno che non abbiano una massa maggiore della scala dell'eV. Se il valore di N_{eff} aumentasse (o si abbassasse), questo cambierebbe il rapporto tra l'ammontare della radiazione e della materia nell'universo, e quindi, produrrebbe il cosiddetto effetto *Integrated Sachs-Wolfe* (ISW). Durante il loro viaggio dall'ultima superficie di diffusione alla terra, i fotoni sono soggetti a buche di potenziale e picchi prodotti da (crescenti) disomogeneità. Come abbiamo detto, il modo in cui il potenziale gravitazionale cambia nel tempo dipende dall'espansione. In un universo dominato dalla radiazione o dalla costante gravitazionale, buche o picchi localizzati decadono nel tempo. Quando un fotone cade in una buca, accelera e trasforma la sua energia gravitazionale in energia cinetica (temperatura). L'opposto avviene quando risale dalla buca. Per la conservazione dell'energia, l'energia cinetica finale del fotone rimarrebbe la stessa se il potenziale gravitazionale non cambiasse nel tempo, ed i fotoni della CMB mantenessero la loro distribuzione di temperatura (ovviamente spostata verso il rosso a causa dell'espansione). Quindi lo spettro finale della CMB

sarebbe modificato grazie all'effetto ISW dall'epoca dominata dalla radiazione a quella dominata dalla materia (Early ISW) o da quest'ultima a quella dominata dalla.

ii) Sulle piccole scale (nella Figura 5, il range $l > 1000$) la diffusione dei fotoni cancella le fluttuazioni di temperatura. La tipica scala di lunghezza per questo meccanismo è $\sqrt{t_u}$, dove t_u è l'età dell'universo all'epoca della separazione radiazione materia. La radice quadrata indica il comportamento tipico del moto Browniano: collisioni casuali portano ad uno spostamento delle particelle che, in media, è nullo ma la cui varianza cresce con il tempo, ma più lentamente del moto della particella (lo spostamento sarebbe proporzionale al tempo in questo caso). D'altra parte, il primo picco nello spettro di C_l , chiaramente visibile attorno a $l = 200$, corrisponde alla più grande scala casualmente collegata alla ricombinazione, l'oscillazione finale del plasma di fotoni ed elettroni prima del disaccoppiamento. Questa scala è dell'ordine di grandezza dell'orizzonte delle particelle all'epoca del disaccoppiamento ed è una funzione lineare del tempo t_u . Confrontando la posizione del picco nel disegno in funzione di l con il range di l dello smorzamento possiamo inferire qualche informazione su t_u , e quindi sulla velocità dell'espansione dell'universo all'epoca della formazione della CMB. Poiché il valore di H dipende dall'ammontare totale della densità di energia di radiazione (e di materia), questo si modifica in un limite su N_{eff} . L'esperimento Planck usa questo fenomeno per restringere fortemente la stima del numero effettivo di neutrini $N_{\text{eff}} = 3.04 \pm 0.18$ [9]. Non c'è spazio per una quarta specie di neutrini, a meno che la sua densità nelle prime ere dell'universo sia molto inferiore a quella dei neutrini attivi.

iii) La massa del neutrino ha un doppio effetto sulla CMB, se non è troppo più piccola della temperatura dell'ultima superficie di diffusione che è dell'ordine di qualche frazione di eV. Da una parte fissa il tempo nella storia dell'espansione quando i neutrini diventano non relativistici, quindi quando non sono

più *radiazione* ma diventano *materia*. È notevole il fatto che i neutrini abbiano una massa piccola, e che siano le sole particelle per le quali possiamo vedere questa transizione usando le osservazioni! Se fossero diventati non relativistici durante la formazione della CMB, questo avrebbe cambiato il già menzionato effetto ISW. Ad un certo punto l'universo sente più materia di quella attesa se i neutrini fossero particelle prive di massa, e la disomogenità comincia a crescere prima, come in un universo dominato dalla materia. Il secondo effetto della massa del neutrino è l'effetto *lente* sulla CMB. I fotoni emessi dall'ultima superficie di diffusione viaggiano verso la Terra e lungo il loro cammino incontrano galassie, cluster ecc. Il loro effetto è la famosa previsione della teoria di Einstein: il *lensing* cioè la curvatura delle loro traiettorie, e un cambiamento della loro energia nel campo gravitazionale dei corpi massivi. L'effetto *di lente* è naturalmente legato all'ammontare di materia lungo la linea di vista, ed il suo peso relativo rispetto alla radiazione. Un neutrino massivo contribuirebbe all'effetto lente in maniera alquanto differente rispetto ad uno senza massa. L'aumento della massa del neutrino sopprime le aggregazioni su scale inferiori alla dimensione dell'orizzonte delle particelle al tempo della transizione non relativistica (prima di questa era, in effetti, i neutrini sono radiazione), e quindi l'effetto lente è inferiore a queste scale. L'effetto lente sulla CMB è stato misurato dall'esperimento Planck che ha trovato uno stretto vincolo sulle somme delle masse dei neutrini, in altre parole, sulla scala di massa assoluta m_0 che abbiamo menzionato prima, dell'ordine di 0.2 eV. Notiamo che questo valore è la sensibilità a cui mira l'esperimento Katrin.

iv) Non diversamente dai fotoni, anche i neutrini hanno piccole perturbazioni, che rappresentano il primo segno dello stadio inflazionario. Queste perturbazioni evolverebbero in maniera differente se i neutrini fossero particelle prive di massa, se si fossero mosse liberamente durante l'epoca della CMB, cioè se non fossero state soggette ad alcun proces-

so di interazione, diffusione, annichilazione ecc., o se fossero massive. I recenti risultati dell'esperimento Planck ci dicono che le nostre previsioni sono abbastanza consistenti con i dati. I neutrini sono come Pauli li immaginava: interagiscono debolmente, sono particelle leggere, si muovono alla velocità della luce attraverso la CMB.

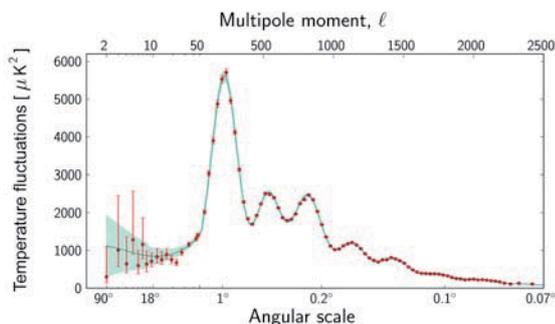


Figura 5: Lo spettro di potenza della CMB (C_l) in funzione della scala angolare (scala inferiore) e del momento di multipolo l (scala superiore).

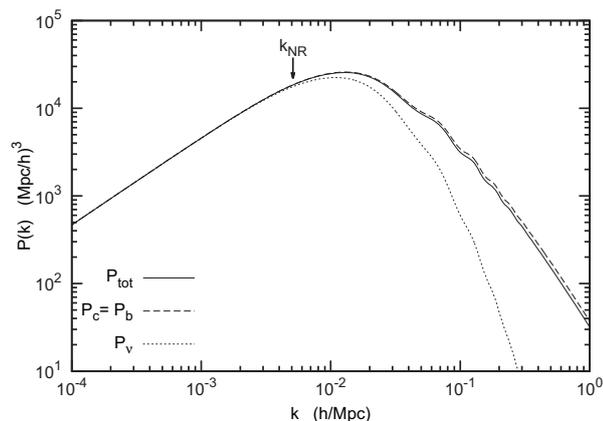


Figura 6: Lo spettro di potenza della materia $P(k)$ in funzione dell'inverso della distanza $k = 2\pi/d$, misurato in unità di h / Mpc^{-1} . L'incertezza sul valore attuale della costante di Hubble è inserita nel parametro h , definito come $H_0 = 100 h \text{ Km/s Mpc}^{-1}$. La linea continua è il valore totale di $P(k)$, la linea tratteggiata è il contributo della materia oscura (c =cold dark matter) e dei barioni (b). Infine, la linea punteggiata è lo spettro di perturbazione per tre neutrini con massa di 0.3 eV ognuno. Il numero d'onda k_{NR} è l'inverso della scala di lunghezza alla quale i neutrini diventano particelle non relativistiche.

Il valore della massa dei neutrini può anche essere vincolato dalla quantità osservata di galassie

e cluster dell'universo. Il meccanismo fondamentale è uguale quello che abbiamo descritto nel caso della CMB. La transizione del fondo di neutrini da un fluido relativistico a non relativistico cambia il modo in cui loro influiscono sulla formazione delle strutture. L'equivalente del C_l per le disomogeneità della materia è il cosiddetto spettro di potenza, $P(k)$. È la funzione di correlazione a due punti per galassie, cluster ecc. a distanze spaziali d dell'ordine di $k = 2\pi/d$, un altro modo di studiare: i) i valori iniziali delle disomogeneità prodotte durante l'epoca iniziale dell'inflazione, e ii) la dinamica di queste perturbazioni casuali sotto l'effetto della gravità. La definizione di $P(k)$ è richiamata nel riquadro II C_l e lo spettro di potenza $P(k)$, e il suo comportamento in funzione di k , espressa in Mpc^{-1} , è mostrata dalla linea continua nella Figura 6. Questo comportamento è ancora consistente con la nostra comprensione attuale su come le strutture si formano nell'universo. Minuscole fluttuazioni sono state amplificate da instabilità gravitazionali in presenza di particelle sufficientemente pesanti (la materia oscura), che attivano il collasso.

Il ruolo dei neutrini è legato alla loro massa anche in questo caso. Se fossero privi di massa, avrebbero soppresso la crescita delle strutture indipendentemente dalla scala $1/k$ alla quale noi li osserviamo. Tuttavia, se la loro massa non è trascurabile, c'è una scala peculiare $1/k_{NR}$ che distingue due regimi differenti. Ho già menzionato che si possono sviluppare strutture per mezzo dell'instabilità gravitazionale su scale che sono inferiori all'orizzonte delle particelle. Questo significa che le strutture a piccola scala si sviluppano prima di quelle a grande scala, semplicemente perché l'orizzonte delle particelle cresce con il tempo, poiché è la distanza che ha percorso il fotone. Nelle prime epoche i neutrini sono relativistici, e il loro ruolo è quello di smorzare parzialmente la formazione delle strutture. Hanno una grande velocità, quella della luce, e sono liberi di muoversi nell'universo rendendo omogeneo il potenziale gravitazionale. Tutte le scale che sono abbastanza piccole, in particolare più piccole dell'orizzonte delle particelle in questa epoca, soffrono di questo effetto e crescono meno di quanto accadrebbe se i neutrini non fossero presenti nell'universo. D'altra parte, le disomogeneità a grande scala, che iniziano a col-

Il C_l e lo spettro di potenza $P(k)$

L'osservabile principale della mappa della CMB è la funzione di correlazione temperatura-temperatura

$$\overline{\delta T(\theta, \phi) \delta T(\theta', \phi')}$$

(si veda il testo per la definizione), dove la media è fatta sulla probabilità di distribuzione della variabile stocastica $T(\theta, \phi)$. Mediare significa che se avessimo molti universi a nostra disposizione potremmo fare molti esperimenti e dedurre le proprietà di $T(\theta, \phi)$. Questo è impossibile. Tuttavia, se consideriamo correlazioni a piccole scale angolari, possiamo mediare su molti settori del cielo che evolvono abbastanza indipendentemente. L'universo non è omogeneo sulle piccole scale, ma è isotropo. In questo modo la funzione di correlazione può dipendere solo dall'angolo relativo tra i punti di osservazione ϑ . Il suo valore per un certo ϑ è inserito nel parametro C_l

$$C_l \sim \overline{\delta T \delta T}(\vartheta) \Big|_{\vartheta=\pi/l}$$

In una formulazione matematica più formale, i C_l sono le trasformate di Legendre della funzione di correlazione a due punti.

Un approccio simile può essere utilizzato per definire la funzione di correlazione galassia-galassia (cluster-cluster, ecc.). In questo caso, diversamente dalla CMB, che viene da una sfera ad una data distanza da noi, l'ultima superficie di diffusione, le galassie sono distribuite nello spazio a tre dimensioni. Se $\overline{\delta(\mathbf{x})\delta(\mathbf{x}')}$ è la funzione di correlazione della densità di materia (si veda il testo), essa può dipendere solo dal modulo $|\mathbf{x} - \mathbf{x}'|$ (isotropia). La sua *trasformata di Fourier*, lo spettro di potenza, è analogo a quello di C_l per le fluttuazioni di temperatura della CMB

$$k^3 P_k \sim \overline{\delta \delta}(|\mathbf{x} - \mathbf{x}'|) \Big|_{|\mathbf{x} - \mathbf{x}'|=2\pi/k}$$

lassare dopo che il neutrino diventò non relativistico, non sono influenzate dal loro scorrere liberamente. In aggiunta a questo, anche i neutrini cominciano a sentire i pozzi di potenziale gravitazionale, poiché la loro velocità può essere più piccola della velocità di fuga da, diciamo, un cluster di galassie in formazione, e cominciano anch'essi a collassare. Questo causa un ulteriore incremento nella formazione delle strutture. Da analisi delle galassie, quali il Sloan Digital Sky Survey [10], possiamo quindi inferire un limite superiore della massa del neutrino. Come per il caso della CMB, quello che possiamo misurare è la scala della massa del neutrino, più precisamente la somma delle masse delle specie dei neutrini. Il limite è dello stesso ordine di grandezza di quello suggerito dalla CMB: una frazione di eV.

Conclusioni

La cosmologia dei neutrini è più di quello che ho descritto in questo breve articolo. È un campo di ricerca molto attivo, il cui scopo, come ho cercato di illustrare, è quello di vincolare le proprietà dei neutrini usando la cosmologia. Alcuni aspetti sono stati discussi qui. Molti altri, come le asimmetrie neutrino-antineutrino, le interazioni esotiche dei neutrini, le proprietà elettromagnetiche dei neutrini, il loro ruolo nel produrre un'asimmetria barione-antibarione, meriterebbero un altro articolo. Occasionalmente, la cosmologia fornisce indizi prima delle misure di laboratorio, come nel caso del numero di neutrini attivi (N_{eff}), che fu già trovato essere tre dalla BBN, prima della evidenza diretta agli esperimenti di LEP.

Forse, la prospettiva di una rivelazione dei neutrini fossili, un'altra eco del Big Bang, è al momento solo un sogno, sognato da molti fisici teorici e sperimentali. Usando le parole di Pauli,

pensando a loro

... è una cosa terribile ..., qualcosa che nessun fisico teorico dovrebbe mai fare, ... qualcosa che non può essere verificata sperimentalmente ...

Se Pauli si era sbagliato sulla rivelazione dei neutrini, possiamo sperare che anche i neutrini che riempiono l'universo possano essere, forse fortunatamente, scoperti in un futuro non troppo lontano.



- [1] F. HOYLE: "Concluding Remarks", *Proceedings of the Royal Society A* **301** (1967) 171.
- [2] <https://www.katrin.kit.edu>
- [3] G. CO': "Neutrinos and weak interaction", *Ithaca VI* (2016) 7.
- [4] E. LISI: "Neutrinos: messengers of new physics", *Ithaca VI* (2016) 117.
- [5] D. MONTANINO: "Neutrino oscillations", *Ithaca VI* (2016) 19.
- [6] E. MAJORANA: "Teoria simmetrica dell'elettrone e del positrone", *Il Nuovo Cimento* **14** (1937) 171.
- [7] F. VISSANI: "Maoirana's conundrum", *Ithaca VI* (2016) 47.
- [8] P. BERNARDINI: "The frontier of sterile neutrinos", *Ithaca VI* (2016) 109.
- [9] <http://www.cosmos.esa.int/web/planck>
- [10] <http://www.sdss.org>



Gianpiero Mangano: è ricercatore in fisica teorica dell'Istituto Nazionale di Fisica Nucleare, Sezione di Napoli, Italia. Il suoi principali interessi scientifici sono la fisica dei neutrini, cosmologia e idee bizzarre circa la natura dello spazio-tempo a piccole distanze. È co-autore del libro *Neutrino Cosmology* (Cambridge University Press, 2013), un divertissement a quattro mani scritto con Julien Lesgourgues, Gennaro Miele and Sergio Pastor.

Ascoltando l'Universo di onde gravitazionali: implicazioni e prospettive dopo la scoperta

Imagination is more important than knowledge. Knowledge is limited to all we now know and understand, while imagination embraces the entire world, and all there ever will be to know and understand.

Albert Einstein

Paola Leaci Dipartimento di Fisica - Università "La Sapienza di Roma e INFN Sezione di Roma

La data dello scorso 14 Settembre 2015 ha ufficialmente segnato l'inizio di una nuova era, quella dell'*Astronomia Gravitazionale*. In questo articolo, dopo una breve introduzione sul concetto di onde gravitazionali, strumentazione utilizzata per la loro rivelazione e sorgenti astrofisiche, saranno illustrati i dettagli della prima osservazione diretta dell'esistenza di onde gravitazionali, evidenziando le ricadute in am-

bito scientifico e tecnologico derivanti da questa sensazionale scoperta.

Cosa sono le onde gravitazionali?

Le onde gravitazionali sono increspature dello spazio-tempo, predette dalla Teoria della Relatività Generale di Einstein nel 1916, che si propagano alla velocità della luce [1]. Esse sono prodotte da accelerazioni di massa a simmetria non assiale, in maniera analoga a cariche elettriche che, in moto accelerato, emettono onde

elettromagnetiche che viaggiano a velocità della luce. Le due tipologie di onde, entrambe trasversali, sono comunque piuttosto differenti. Le onde gravitazionali sono caratterizzate da due stati di polarizzazione, denotati come “+” e “×”, che differiscono tra loro per una rotazione di 45 gradi attorno all’asse di propagazione, mentre le onde elettromagnetiche differiscono tra loro per una rotazione di 90 gradi [2].

Quando un’onda gravitazionale attraversa la materia, la distanza spazio-temporale tra i suoi costituenti viene modificata di una quantità estremamente piccola. L’effetto di un’onda gravitazionale che attraversa un anello di particelle libere è schematizzato in Fig. 1 notevolmente amplificato per essere più chiaro. L’anello di particelle diventa una delle ellissi e ritorna ad essere un cerchio durante la prima metà del ciclo, diventa quindi l’altra ellisse e ritorna ad essere un cerchio durante la successiva metà.

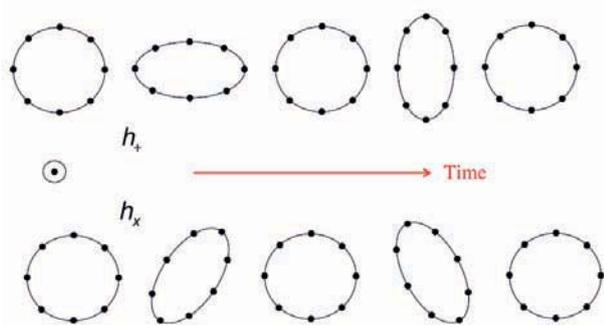


Figura 1: Risposta di un anello di particelle libere (punti in figura), in funzione del tempo, ad onde gravitazionali polarizzate “+” (h_+) e “×” (h_x), che si propagano perpendicolarmente al piano della figura.

A differenza delle onde elettromagnetiche, le onde gravitazionali interagiscono in maniera trascurabile con la materia, viaggiando quasi indisturbate dalle zone più remote del cosmo a noi, trasportando quindi tutta quell’informazione che le onde elettromagnetiche non trasportano.

La quantità di energia emessa sotto forma di onde gravitazionali da un qualsiasi sistema meccanico costruito dall’uomo è estremamente piccola, pertanto molto difficile da misurare. Per questa ragione si preferisce osservare la radiazione gravitazionale emessa da eventi catastrofici e da sorgenti astrofisiche.

Rivelatori di onde gravitazionali

Il pioniere della rivelazione di onde gravitazionali è stato Joseph Weber negli anni 1960 [3], che ha sviluppato il primo rivelatore a barra risonante e più tardi ha investigato la più promettente interferometria laser. I primi rivelatori interferometrici sono stati suggeriti negli anni 1960 [4] e 1970 [5] e, grazie a studi riguardanti la caratterizzazione del rumore, successivamente migliorati [6, 7], fino a diventare strumenti a larga banda estremamente sensibili in grado di rivelare onde gravitazionali [8, 9, 10, 11].

Un rivelatore interferometrico di onde gravitazionali è costituito da due bracci lunghi da 3 a 4 km, posti a 90 gradi l’uno dall’altro e alla cui estremità sono posizionate delle masse libere, gli specchi, su cui viene riflessa la radiazione laser che percorre entrambi i bracci. Un’onda gravitazionale che attraversa un interferometro tenderà ad allungare leggermente un braccio e ad accorciare un altro, producendo uno sfasamento tra i due fasci di luce laser, che è direttamente proporzionale all’intensità dell’onda gravitazionale in transito. L’effetto di un’onda gravitazionale è quello di cambiare la distanza fra le estremità dei bracci dell’interferometro di una quantità ΔL . Le tecniche attuali permettono di identificare ΔL dell’ordine di 10^{-18} m, ossia distanze cento milioni di volte più piccole delle dimensioni di un atomo. L’intensità delle onde gravitazionali produce fenomeni proporzionali al rapporto $\Delta L/L$, risulta pertanto chiaro che bracci più lunghi consentono di rivelare segnali più deboli. Per rivelatori come LIGO (Laser Interferometer Gravitational wave Observatory), con bracci lunghi $L = 4$ km, questo significa essere in grado di rivelare segnali con un’ampiezza (*strain*) pari a $h \sim \Delta L/L \sim 10^{-22}$. Per maggiori dettagli consultare per esempio [12].

I rivelatori interferometrici di onde gravitazionali attualmente esistenti sono:

- i rivelatori americani LIGO, situati uno ad Hanford (Washington) e l’altro a Livingston (Louisiana), entrambi (ribadiamo) con una lunghezza di bracci pari a 4 km;
- il rivelatore europeo Virgo, situato a Cascina (Italia), avente una lunghezza di bracci di 3 km;

- il rivelatore anglo-tedesco GEO, situato nelle vicinanze di Hannover (Germania), con bracci di lunghezza pari a 600 m.

Dopo il 2019 si prevede l'entrata in funzione del rivelatore giapponese KAGRA e successivamente al 2022 di quello americano-indiano LIGO-India. Pertanto, in un futuro molto vicino, questa rete mondiale di rivelatori, schematizzata in Fig. 2, garantirà un'ottima copertura del cielo. Questi rivelatori rappresentano gli interferometri di seconda generazione. Ad essi seguiranno rivelatori di terza generazione, sotterranei, criogenici e con lunghezza di bracci di decine di km, come i progetti futuri ed ambiziosi Einstein Telescope [13], LIGO Voyager e LIGO Cosmic Explorer [14].

Tutti i rivelatori descritti in questa sezione sono rivelatori terrestri caratterizzati da una risposta a larga banda, in grado di coprire l'intervallo di frequenze da ~ 10 Hz a ~ 2 kHz. Del tutto complementare è il futuro rivelatore spaziale LISA, che mira a rivelare onde gravitazionali a bassissima frequenza, da 0.1 mHz a 1 Hz, dato che non risente dei disturbi ambientali di origine terrestre di cui sono affetti i rivelatori in Fig. 2.

Sorgenti di onde gravitazionali e tecniche di analisi dati

Le sorgenti di onde gravitazionali si possono raggruppare in sorgenti impulsive, che generano segnali transienti di onde gravitazionali e sorgenti che danno luogo a segnali continui di onde gravitazionali.

Sorgenti impulsive

- Una delle principali sorgenti di onde gravitazionali è la *coalescenza di sistemi compatti*, formati da due stelle di neutroni, due buchi neri oppure una stella di neutroni ed un buco nero. Le fasi della coalescenza includono: (i) lo spiraleggiamento, in cui i due corpi orbitano uno attorno all'altro perdendo energia e momento angolare attraverso emissione di onde gravitazionali; i due oggetti pertanto si avvicinano, le loro velocità aumentano, come anche la frequenza ed ampiezza delle onde gravitazionali emesse

finché i due oggetti si fondono durante la fase di (ii) *merger*, a cui segue l'ultima fase di (iii) *ringdown*, in cui si viene a formare l'oggetto compatto finale che emette onde gravitazionali ad una frequenza costante ed ampiezza che si smorza esponenzialmente.

Il metodo che si utilizza per rivelare onde gravitazionali emesse da questo tipo di sorgenti è basato sul cosiddetto filtro adattato (*matched filtering*), che consiste nel correlare i dati (raccolti dai rivelatori) con varie forme d'onda note (dette *templates*) ognuna delle quali descrive le caratteristiche del segnale mediante valori diversi di parametri non noti, quali le masse e gli *spins* degli oggetti compatti [15]. Chiaramente, maggiore è lo spazio dei possibili valori di masse e *spins* che si vogliono considerare e tanto più onerosa sarà la ricerca dal punto di vista computazionale. Questa tecnica è stata utilizzata per identificare il primo segnale gravitazionale, indicato con GW150914 (dato che si è verificato il 14 Settembre 2015), in maniera diretta [16].

- Altra sorgente di segnali transienti di onde gravitazionali è quella prodotta dal collasso asimmetrico della parte centrale di stelle massive durante eventi di supernovae, fenomeno durante il quale sono prodotti *bursts* di onde gravitazionali.

La tecnica che si utilizza per rivelare questo tipo di segnale si basa sulla ricerca di un eccesso di potenza nello spazio tempo-frequenza utilizzando una decomposizione *wavelet* [17]. Ricerche per questo tipo di segnale si svolgono analizzando anche la controparte elettromagnetica utilizzando, ad esempio, eventi di tipo *gamma-ray bursts* identificati da vari telescopi [18]. Anche la ricerca basata sulla decomposizione *wavelet* è stata utilizzata per identificare il primo segnale gravitazionale in maniera diretta [16].

Sorgenti continue

- Si prevede inoltre l'esistenza di un fondo stocastico di onde gravitazionali, sia di natura astrofisica, per esempio dovuto alla sovrapposizione di molti segnali transienti, sia di natura cosmologica, per esempio dovuto alla sovrapposizione di molti segnali continui.

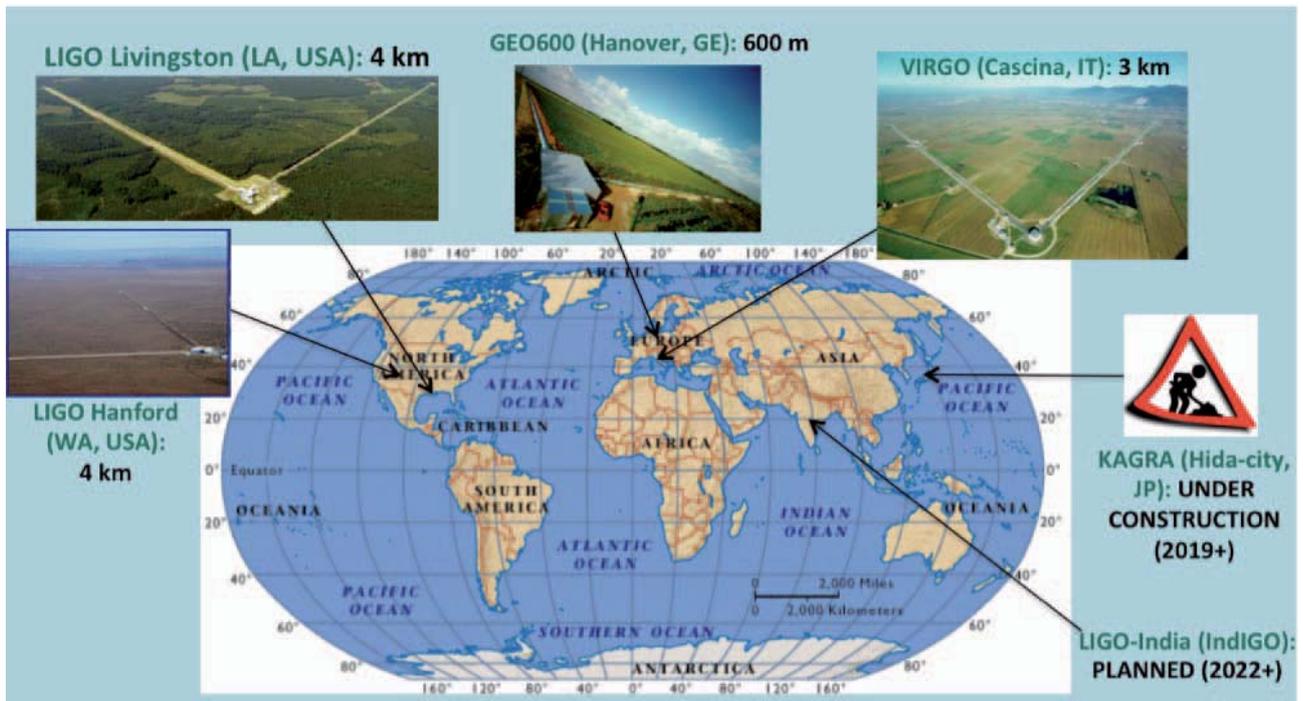


Figura 2: Rete di rivelatori interferometrici di onde gravitazionali.

posizione incoerente di segnali emessi da una vasta popolazione di sorgenti distanti, che di natura cosmologica. Quest'ultimo si pensa sia stato emesso nei primissimi istanti di vita dell'Universo e, se rivelato, ci fornirà preziose informazioni su ciò che è avvenuto all'era di Planck, ossia 10^{-43} s dopo il Big Bang. La metodologia di ricerca utilizzata per distinguere tale debole segnale dal rumore strumentale si basa sulla correlazione tra i dati raccolti almeno da due rivelatori. La rivelazione dell'evento GW150914 ha permesso di stimare un fondo stocastico di onde gravitazionali da binarie di buchi neri che risulta più alto delle predizioni precedenti, pertanto di possibile rivelazione con la rete di rivelatori avanzati LIGO e Virgo operanti a sensibilità ottimali [19].

- Una classe di sorgenti di onde gravitazionali molto promettente è rappresentata da stelle di neutroni (caratterizzate da una simmetria non assiale) rapidamente rotanti nella nostra Galassia. Si noti che un oggetto compatto a simmetria sferica non emette onde gravitazionali perchè la propria distribuzione di massa non ha una componente quadrupolare variabile nel tempo. La deviazione dalla simmetria sferica può per

esempio essere causata da stress magnetici interni per stelle di neutroni rapidamente rotanti o da accrescimento di materia su stelle di neutroni facenti parte di stelle binarie a raggi X di piccola massa. Sebbene questa tipologia di segnali non sia stata finora identificata, sono stati posti stringenti limiti superiori sull'intensità di onde gravitazionali emesse da stelle di neutroni isolate (come ad esempio [20, 21]) e sono stati raggiunti importanti risultati nel caso di onde gravitazionali emesse da stelle di neutroni facenti parte di sistemi binari (si vedano ad esempio le referenze [22, 23]).

La prima rivelazione diretta di onde gravitazionali

In seguito ad un importante aggiornamento, e conseguente miglioramento in sensibilità [12], i rivelatori LIGO sono tornati ad essere operativi con il *run* scientifico O1 durante il quale gli interferometri hanno raccolto dati tra Settembre 2015 e Gennaio 2016¹. Il 14 Settembre 2015 i due rivelatori LIGO hanno osservato per la prima volta un segnale di onde gravitazionali transiente pro-

¹Il rivelatore Virgo era in fase di aggiornamento in questo periodo, pertanto non utilizzabile per la presa dati.

dotto dalla fusione di due buchi neri, processo questo mai osservato finora, con masse pari a 36 e 29 volte quella del Sole e distanti più di un miliardo di anni luce da noi. In [16] (e nei riferimenti bibliografici in esso contenuti) sono stati riportati i dettagli di questa duplice scoperta, che riassumiamo nel seguito. Il segnale è caratterizzato da un aumento in ampiezza e frequenza, da 35 a 150 Hz in otto oscillazioni, per poi raggiungere il picco massimo, con l'ampiezza del segnale che raggiunge il valore massimo di 10^{-21} , come si può notare in Fig. 3. Le serie temporali illustrate sono state filtrate per rimuovere vari disturbi strumentali. Si possono notare le ampiezze (*strain*) osservate da Livingston (L1, prima colonna di destra), dove l'evento è giunto prima e da Hanford (H1, prima colonna di sinistra), 7 ms più tardi. Per una migliore visualizzazione, i dati di H1 sono stati spostati temporalmente di una tale quantità ed invertiti (in maniera tale da considerare le diverse orientazioni dei rivelatori) [16]. La seconda riga di Fig. 3 mostra le forme d'onda ottenute utilizzando la relatività numerica² per un sistema con parametri consistenti con quelli dell'evento GW150914 (linee solide) e le forme d'onda ricostruite con due metodi basati sul filtro adattato (utilizzando 150 000 *templates*) e la decomposizione *wavelet* (aree ombreggiate). Si può notare l'accordo tra le forme d'onda, che si traduce in una conferma dell'esattezza della teoria della Relatività Generale di Einstein [24]. Nella rappresentazione tempo-frequenza dell'ultima riga di Fig. 3 si nota che la frequenza del segnale aumenta nel tempo, indicando che il sistema osservato è costituito da due oggetti che ruotano uno intorno all'altro, avvicinandosi fino a fondersi in un unico oggetto. L'informazione su quanto massivi siano gli oggetti in questione viene fornita dal repentino aumento della frequenza, mentre il valore massimo di frequenza raggiunto ci dice quanto vicini (~ 240 km) fossero i due oggetti prima del contatto. La ragione che porta a dedurre che si tratti di due buchi neri è dovuta al fatto che, per raggiungere una frequenza orbitale di 75 Hz (metà della frequenza dell'onda gravitazionale), i due oggetti devono essere molto vicini e molto compatti. Gli unici

²Si utilizza la locuzione relatività numerica per indicare una soluzione numerica delle equazioni di Einstein per una situazione di accoppiamento forte della gravità.

oggetti che conosciamo essere tanto massivi e compatti sono i buchi neri (per maggiori dettagli consultare [16]). Inoltre, l'intensità del segnale gravitazionale che ha raggiunto la Terra fornisce informazioni sulla loro distanza e quindi su quanto tempo fa l'evento si è verificato, ossia poco meno di un miliardo e mezzo di anni fa, quando sul nostro pianeta iniziavano a comparire le prime cellule evolute in grado di utilizzare l'ossigeno.

Il violento processo di fusione osservato è caratterizzato da un rilascio di energia pari a tre volte la massa del Sole, ma risulta invisibile. Le onde gravitazionali, che ribadiamo essere in grado di attraversare indisturbate profondi strati di materia, risultano quindi l'unico messaggero in grado di fornire informazioni su ciò che è veramente accaduto, essendo assenti le emissioni di segnali elettromagnetici.

Precisiamo che entrambi i rivelatori LIGO operavano in condizioni stabili all'epoca dell'evento, consentendoci di escludere, dopo studi dettagliati, che l'evento in questione potesse esser dovuto a un disturbo strumentale. Inoltre è stata vagliata anche la possibilità che potesse trattarsi di una *blind injection*, ossia un segnale artificiale introdotto nel rivelatore a livello *hardware* per controllare l'abilità dei vari algoritmi di ricerca di identificare un segnale di questo tipo. Questa eventualità è stata banalmente esclusa dato che i rivelatori non operavano in tale modalità al tempo dell'evento.

Onde gravitazionali tra scienza e progresso tecnologico

Dal punto di vista scientifico, la scoperta delle onde gravitazionali, oltre che convalidare la teoria della Relatività Generale di Einstein, ci permette di comprendere quello che accade a densità sopranucleari nelle fusioni di stelle di neutroni, di conoscere il grado di asimmetria tipico della stessa stella di neutroni, di conoscere la densità di stelle di neutroni nell'universo e l'evoluzione dei sistemi binari oggetto di studio. Tutte queste informazioni saranno possibili dopo aver rivelato più segnali di onde gravitazionali, sia transienti che continui. La scoperta di un fondo stocastico cosmologico di onde gravitazionali ci

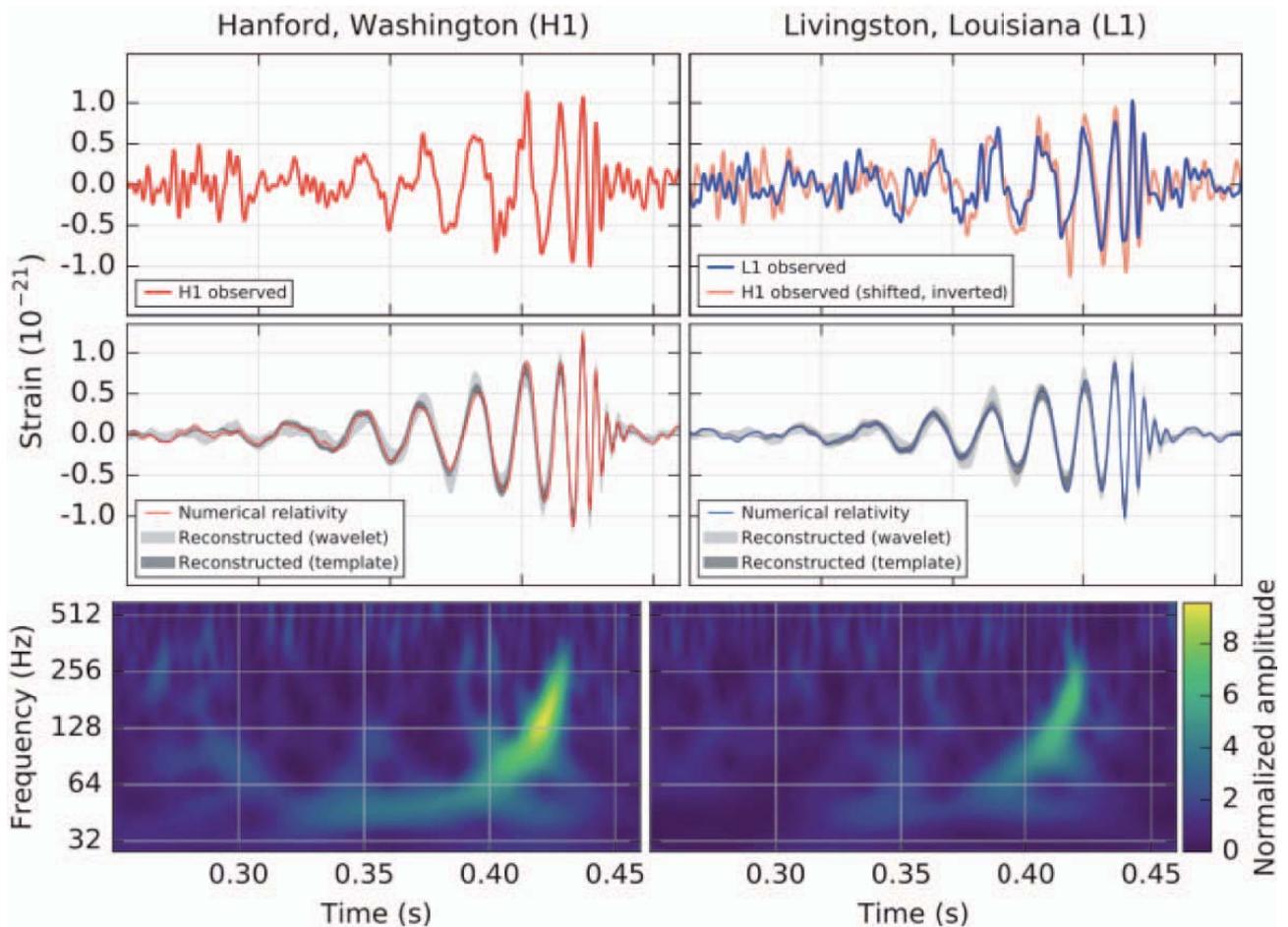


Figura 3: Evento GW150914 osservato dai rivelatori LIGO Hanford e Livingston [16]. Per maggiori dettagli fare riferimento al testo.

consentirebbe poi di guardare indietro nel tempo, di avere quindi accesso ai primissimi istanti di vita dell'Universo (ossia a circa 13.8 miliardi di anni fa), di conoscerne i dettagli e di osservare quest'ultimo diversamente da come abbiamo fatto finora.

Oltre alle tante implicazioni di carattere scientifico, ce ne sono varie anche in ambito tecnologico e pratico. La Relatività Generale di Einstein, che predice l'esistenza di onde gravitazionali, ha fornito per esempio un contributo essenziale alla messa a punto e al grado di precisione ottenibile dalle attuali reti satellitari GPS. Ci sono inoltre molte analogie tra alcune delle tecniche utilizzate per l'identificazione del segnale gravitazionale continuo e quelle utilizzate nel campo dell'elaborazione digitale delle immagini, che trovano impieghi rilevanti anche nel campo della videosorveglianza. Inoltre, il processo di estrazione del segnale gravitazionale è nella maggior parte dei casi molto complicato dal punto di vista com-

putazionale. Esso richiede familiarità con vari software e tecniche per creare algoritmi efficienti in grado di funzionare su processori paralleli. Alcuni di questi programmi e tecniche sono largamente utilizzati non solo nel mondo accademico, ma anche in molti settori dell'industria. Un esempio immediato riguarda la tecnologia superiore basata sulle unità di elaborazione grafica (note come GPU), che è ampiamente utilizzata in campi disparati come l'elaborazione di immagini scientifiche, l'esplorazione petrolifera e perfino in ambito finanziario. Vi sono inoltre sistemi di isolamento sismico e accelerometri ultrasensibili utilizzati per il funzionamento dei rivelatori interferometrici che possono trovare applicazioni anche in settori industriali per quanto riguarda i problemi connessi ad oscillazioni di ponti e grattacieli, o perfino in geofisica. Questi sono solo alcuni esempi per evidenziare la notevole ricaduta tecnologica e non solo scientifica.

Conclusioni e prospettive future

La comunità scientifica delle onde gravitazionali è da sempre consapevole della necessità di raccogliere dati con più rivelatori installati in siti lontani tra loro. Questo consente non solo di abbattere i disturbi ambientali mediante tecniche di coincidenza, ma di ottenere notevoli miglioramenti nella stima dei parametri e nella precisione dell'identificazione della posizione celeste di una certa sorgente astrofisica. Questo sarà fattibile in maniera ancora più efficace nei prossimi anni, quando la rete di rivelatori di onde gravitazionali diventerà più grande e ancora più sensibile. Ulteriori vantaggi saranno ottenuti dall'instaurare collaborazioni sempre più strette con i colleghi astronomi ed effettuare ricerche di onde gravitazionali che consentono di identificare una controparte elettromagnetica nei processi astrofisici osservati. Questo ci consentirà infatti di conoscere la fisica degli oggetti compatti e delle esplosioni violente.

Si ribadisce che la scoperta scientifica annunciata costituisce una importantissima conferma della teoria della Relatività Generale di Einstein. Essa apre un nuovo capitolo della Fisica delle Interazioni Fondamentali e inaugura l'Astronomia Gravitazionale, una nuova finestra di osservazione dell'Universo, che può fornire informazioni non altrimenti accessibili (neanche ai rivelatori elettromagnetici). Tale scoperta ricompensa gli sforzi ininterrotti di generazioni di fisici e il sostegno delle agenzie finanziatrici, che hanno avuto il coraggio di investire a lungo termine in un'impresa tanto accattivante, ma non di ovvia riuscita. La tenacia di noi fisici è stata certamente fomentata dalla brillante prova indiretta sull'esistenza di onde gravitazionali, che risale al 1974. In tale anno, infatti, Hulse and Taylor scoprirono la prima radio pulsar³ in un sistema binario, il cui periodo orbitale decresceva in perfetto accordo con quanto predetto dalla teoria della Relatività Generale riguardo all'emissione di onde gravitazionali [25].

Con l'osservazione diretta del primo segnale gravitazionale transiente, lo scorso 14 Settembre

³Una pulsar è una stella di neutroni rotante, caratterizzata da valori molto elevati di campo magnetico ed emettente fasci di radiazione elettromagnetica, che si possono osservare a Terra come impulsi emessi ad intervalli di tempo molto regolari.

2015 si è aperta una finestra del tutto nuova sull'Universo. Rimangono tuttavia altre classi di segnali gravitazionali da rivelare e senza dubbio il miglioramento in sensibilità dei nostri rivelatori ed algoritmi di ricerca porterà grandi risultati. Durante questo percorso non sarà sorprendente scoprire eventualmente oggetti inattesi, dal momento che abbiamo appena iniziato ad esplorare uno spazio di parametri a cui non abbiamo mai avuto accesso.

Nota

La trattazione qui presentata è una rassegna sommaria che illustra in modo divulgativo lo stato dell'arte della ricerca di onde gravitazionali e dell'astronomia che si può iniziare a fare. Pur non avendo la pretesa di essere completa, si auspica di dare una chiara panoramica dell'argomento, che permetta ai lettori più interessati di approfondire vari aspetti, sia nei riferimenti bibliografici di seguito elencati, sia consultando la vasta letteratura disponibile sul Web.



- [1] A. EINSTEIN: "Die Grundlage der allgemeinen Relativitätstheorie", *Annalen der Physik* **49** (1916) 769.
- [2] C. MISNER, K. THORNE, J. WHEELER: *Gravitation*. W. H. Freeman, S. Francisco USA (1973).
- [3] J. WEBER: "Detection and Generation of Gravitational Waves", *Phys. Rev.* **117** (1960) 306.
- [4] E. GERTSENSHTEIN AND V. I. PUSTOVOIT: "On the Detection of Low-Frequency Gravitational Waves", *Sov. Phys. JETP* **16** (1963) 433.
- [5] G. E. MOSS, L. R. MILLER, AND R. L. FORWARD: "Photon-Noise-Limited Laser Transducer for Gravitational Antenna", *Appl. Opt.* **10** (1971) 2495.
- [6] R. Weiss, MIT Report No. 105 (1972) : <https://dcc.ligo.org/public/0038/P720002/001/P720002-00.pdf>
- [7] R. W. P. Drever in "Gravitational Radiation", p. 321, N. Deruelle and T. Piran eds., North-Holland, Amsterdam (1983).
- [8] R. W. P. Drever, F. J. Raab, K. S. Thorne, R. Vogt, and R. Weis, Laser Interferometer Gravitational-wave Observatory (LIGO) Technical Report (1989) : <https://dcc.ligo.org/public/0065/M890001/003/M890001-03%20edited.pdf>
- [9] A. ABRAMOVICI ET AL.: "LIGO: The Laser Interferometer Gravitational-Wave Observatory", *Science* **256** (1992) 325.

- [10] A. Brilliet, A. Giazotto *et al.*, Virgo Project Technical Report No. VIR-0517A-15 (1989): <https://tds.ego-gw.it/itf/tds/index.php?callContent=2&callCode=12122>
- [11] J. Hough *et al.*, Proposal for a joint German-British interferometric gravitational wave detector, MPQ Technical Report 147 No. GWD/137/JH (1989) : <http://eprints.gla.ac.uk/114852/>
- [12] B. P. ABBOTT ..., P. LEACI *et al.* (THE LIGO SCIENTIFIC COLLABORATION AND THE VIRGO COLLABORATION): “GW150914: The Advanced LIGO Detectors in the Era of First Discoveries”, *Phys. Rev. Lett.* **116** (2016) 131103.
- [13] Einstein Telescope Science Team, Einstein gravitational wave Telescope conceptual design study, ET Document: ET-0106C-10, 2011: <http://www.et-gw.eu/>
- [14] LIGO Scientific Collaboration, Instrument Science White Paper, LIGO-T1400316-v4, 2010: <https://dcc.ligo.org/public/0113/T1400316/004/T1400316-v5.pdf>
- [15] B. P. Abbott ..., P. Leaci *et al.* (The LIGO Scientific Collaboration and the Virgo Collaboration): <http://arxiv.org/abs/1602.03840>
- [16] B. P. ABBOTT ..., P. LEACI *et al.* (THE LIGO SCIENTIFIC COLLABORATION AND THE VIRGO COLLABORATION): “Observation of Gravitational Waves from a Binary Black Hole”, *Phys. Rev. Lett.* **116** (2016) 061102.
- [17] B. P. Abbott ..., P. Leaci *et al.* (The LIGO Scientific Collaboration and the Virgo Collaboration): <http://arxiv.org/abs/1602.03843>
- [18] J. AASI ..., P. LEACI *et al.* (THE LIGO SCIENTIFIC COLLABORATION AND THE VIRGO COLLABORATION): “Method and results of a search for gravitational waves associated with gamma-ray burst using the GEO 600, LIGO, and Virgo”, *Phys. Rev. D* **89** (2014) 122004.
- [19] B. P. ABBOTT ..., P. LEACI *et al.* (THE LIGO SCIENTIFIC COLLABORATION AND THE VIRGO COLLABORATION): “Implications for the stochastic gravitational wave background from binary black holes”, *Phys. Rev. Lett.* **116** (2016) 131102.
- [20] J. AASI ..., P. LEACI *et al.* (THE LIGO SCIENTIFIC COLLABORATION AND THE VIRGO COLLABORATION): “Einstein@Home all-sky search for periodic gravitational waves in LIGO S5 data”, *Phys. Rev. D* **87** (2013) 042001.
- [21] J. AASI ..., P. LEACI *et al.* (THE LIGO SCIENTIFIC COLLABORATION AND THE VIRGO COLLABORATION): “First low frequency all-sky search for continuous gravitational wave signals”, *Phys. Rev. D* **93** (2016) 042007.
- [22] J. AASI ..., P. LEACI *et al.* (THE LIGO SCIENTIFIC COLLABORATION AND THE VIRGO COLLABORATION): “First all-sky search for continuous gravitational waves from unknown sources in binary systems”, *Phys. Rev. D* **90** (2014) 062010.
- [23] P. LEACI AND R. PRIX: “Direct searches for continuous gravitational waves from binary systems: Parameter-space metrics and optimal Scorpius X-1 sensitivity”, *Phys. Rev. D* **91** (2015) 102003.
- [24] B. P. Abbott ..., P. Leaci *et al.* (The LIGO Scientific Collaboration and the Virgo Collaboration): <http://arxiv.org/abs/1602.03841>
- [25] R. A. HULSE AND J. H. TAYLOR: “Discovery of a pulsar in a binary system”, *Astrophys. J.* **195** (1975) L51.



Paola Leaci: è una ricercatrice di Fisica presso l'Università “La Sapienza di Roma. Si occupa di Fisica delle Onde Gravitazionali, Analisi dati di Onde Gravitazionali, Astrofisica Relativistica e Cosmologia.

Numero VII Anno 2016

Viaggio nella Scienza

Ithaca

Luce

