

Componenti per ottica integrata

Alberto Tibaldi

16 maggio 2011

Indice

1	Introduzione	4
1.1	Introduzione all'ottica	4
1.2	Sviluppo storico dei fenomeni ottici	13
1.2.1	Propagazione della luce	13
1.2.2	Teoria elettromagnetica	18
1.2.3	Interazione luce-materia e ottica moderna	19
1.3	Cenni sui reticoli di diffrazione	20
2	Ottica dei mezzi stratificati	25
2.1	Equazioni di Maxwell nel dominio del tempo e della frequenza	25
2.2	Mezzi materiali	28
2.2.1	Materiale dielettrico isotropo	29
2.2.2	Mezzi materiali anisotropi	34
2.2.3	Polarizzazione	35
2.3	Onde piane	36
2.3.1	Osservazioni preliminari	36
2.3.2	Derivazione delle onde piane	39
2.4	Mezzi non omogenei: discontinuità piana	47
2.4.1	Metodo delle linee di trasmissione	47
2.4.2	Coefficienti di Fresnel - onde TE	54
2.4.3	Coefficienti di Fresnel - onde TM	73
2.4.4	Dielettrici con perdite	79
2.4.5	Rombo di Fresnel - Fasci gaussiani	80
2.5	Struttura a tre mezzi materiali	83
2.5.1	Casistica 1: $n_1 = n_3 < n_2$	88
2.5.2	Casistica 2: $n_1 = n_3 > n_2$	91
2.5.3	Strati antiriflesso	94
3	Strutture periodiche	100
3.1	Introduzione agli specchi - Riflettore di Bragg	100
3.2	Introduzione al formalismo delle onde di Bloch	104

3.2.1	Cenni sull'interferometro di Fabry-Perot	107
3.3	Concetti fondamentali di Algebra Lineare	109
3.3.1	Usi di basi non ortogonali	114
3.3.2	Funzioni di una matrice	117
3.3.3	Problema agli autovalori generalizzato - generalizzazioni varie	119
3.3.4	Applicazione ai sistemi lineari	120
3.3.5	Esempi di applicazione del cambio di base	122
3.4	Specchi di Bragg	125
3.4.1	Curve di dispersione per le onde di Bloch	134
3.4.2	Cenni al comportamento delle curve di dispersione per $\vartheta_i \neq 0$	136
3.4.3	Coefficienti di riflessione della struttura	138
3.4.4	Confronto con la fisica dei solidi: energy gap	141
3.4.5	Osservazioni conclusive	143
3.5	Note conclusive sui riflettori di Bragg	144
3.5.1	Riflettori di Bragg accordabili	145
3.5.2	Birifrangenza di forma	145
3.5.3	Caso della singola cella come degenerazione caso generale	147
3.5.4	Confronto tra reticoli di Bragg e reticoli di diffrazione .	148
3.5.5	Analisi per piccole riflessioni	150
4	Interferometri di Fabry-Perot	152
4.1	Introduzione e concetti preliminari	152
4.2	Osservazioni e calcoli sull'interferometro di Fabry-Perot	154
4.2.1	Banda dell'interferometro	156
4.2.2	Analisi in transitorio della cavità con condizioni iniziali	161
4.3	Progetto di un interferometro di Fabry-Perot con specchi di Bragg	165
4.3.1	Analisi	170
5	Guide d'onda dielettriche	173
5.1	Introduzione	173
5.1.1	Principali tipi di guide d'onda ottiche	175
5.2	Analisi di una struttura accademica	176
5.2.1	Analisi modale della struttura	177
5.2.2	Analisi di una guida a facce piane parallele	179
5.2.3	Verifica della natura modale delle espressioni ricavate .	187
5.3	Studio della struttura <i>asymmetric slab</i>	192
5.3.1	Note aggiuntive	201
5.4	Numero di modi presenti in una guida dielettrica	203

5.4.1	Modi guidati	206
5.4.2	Modi non guidati	207
5.4.3	Note conclusive sulle guide d'onda dielettriche	208
5.4.4	Strutture planari	217
5.4.5	Metodo dell'indice di rifrazione efficace	217
5.5	Teoria dei modi accoppiati	219
5.5.1	Cenni al formalismo di Marcuvitz e Schwinger	220
5.5.2	Studio dell'accoppiamento modale su guide ottiche	224

Capitolo 1

Introduzione

1.1 Introduzione all'ottica

In questo capitolo si vuole introdurre un insieme di concetti preliminari rispetto al resto della trattazione. Parlando di ottica, l'entità con cui si avrà maggiormente a che fare è senza dubbio la luce; la luce è tuttavia studiabile con due tipi di approcci, di interpretazioni: un approccio ondulatorio o un approccio corpuscolare, particellare; in questa introduzione si parlerà per sommi capi di entrambi gli approcci.

Per quanto riguarda uno studio macroscopico, ondulatorio della luce, l'approccio più interessante è senza dubbio quello "classico", basato sulle equazioni di Maxwell: la luce, a tutti gli effetti, è un campo elettromagnetico, dunque è possibile utilizzare le nozioni di Campi Elettromagnetici al fine di studiare le fenomenologie in questione. Tra i vari metodi con cui si può avere a che fare, uno dei più interessanti è l'ottica geometrica: si tratta di un'approssimazione molto utilizzata, al fine di studiare la propagazione della luce nello spazio o in mezzi guidanti. Come ogni modello, come ogni approssimazione, l'ottica geometrica vale solo in un range limitato (per quanto vasto) di applicazioni: per capire il concetto, si può per esempio pensare agli sviluppi di Taylor:

$$\sqrt{1+x} \sim 1 + \frac{x^2}{2}$$

l'approssimazione, se $x = 10^{-6}$, è assolutamente eccellente; se $x = 0,1$, è pessima, dal momento che volendo conservare dei decimali, essi sarebbero insignificanti. L'ottica geometrica presenta limiti nelle cosiddette *caustiche*. Si immagini per esempio di avere a che fare con una lente convergente:

Se un fascio di luce va sulla lente, i raggi tendono a convergere sul fuoco (il termine *fuoco* deriva dal fatto che, nell'antichità, l'uomo si è reso conto che

mettendo della carta sul fuoco della lente l'energia tendeva a concentrarsi, portando pure alla combustione l'oggetto); questa cosa si può vedere anche elettromagneticamente, pensando che l'incremento di temperatura sia legato a un incremento di energia, dunque di campo elettromagnetico; il modello dell'ottica geometrica tuttavia propone risultati sbagliati, in questo contesto: calcolando il campo con l'ottica geometrica, si trova un campo infinito: sicuramente sarà grande, ma infinito sicuramente non può essere: questo accade, dal momento che il modello dell'ottica geometrica, nel fuoco, è fuori dal proprio campo di validità.

Un altro modello spesso utilizzato in elettromagnetismo è quello della **carica puntiforme**: volendo modellare la presenza di una carica nello spazio, a grande distanza da essa, si può dire che il campo elettromagnetico da essa generato sia generato da una equivalente sorgente localizzata in un punto nello spazio; se tuttavia ci si avvicina a questa carica, studiare il campo elettromagnetico utilizzando l'ipotesi di campo generato da sorgente puntiforme vacilla: avvicinandosi, si ottengono risultati irragionevoli.

Un'altra descrizione della luce "popolare", è quella particellare: la luce vista come un insieme di fotoni. Questa descrizione, come d'altra parte può esserlo anche quella ondulatoria, è assolutamente ingannevole, tant'è che Lamb, un noto fisico, diceva che ci "vorrebbe la patente" per utilizzare la parola "fotone": queste descrizioni corpuscolari infatti sono molto ingannevoli, come detto, nel senso che serve un certo livello di astrazione. Un tipo di fenomenologia che potrebbe essere spiegata per esempio con il comportamento corpuscolare della luce è la formazione di fasci, visibili a occhio nudo, in presenza di fumo: nei film, per esempio, per mostrare dei "raggi LASER" si introduce molto fumo in un ambiente, e si invia un LASER in esso: il LASER interagisce con le particelle di fumo, generando scattering, rendendo dunque visibile a occhio nudo il fascio. In una stanza priva di fumo, ciò non è possibile, a causa di un particolare tipo di rumore: la luce presente nella stanza. Come detto, il fascio è visibile solo in presenza di scattering con delle altre particelle; in una stanza sicuramente delle particelle sono presenti, essendovi polvere; la luce che illumina una stanza, tuttavia, come vedremo tra breve, si comporta sostanzialmente come un rumore, rendendo sostanzialmente insignificante il fenomeno di scattering con i vari pulviscoli presenti nell'aria. La questione è sostanzialmente legata al **rapporto interferenza su segnale**: la luce che genera interferenza rispetto al segnale scatterato è molto maggiore di esso, rendendolo non apprezzabile.

Si può cercare di capire quanti fotoni ci sono nell'ambiente: in un LASER, tendenzialmente, i fotoni sono tutti in un fascio, e tendono a rimanere confinati in esso; in realtà si ha una probabilità non nulla ma molto ridotta, del fatto che alcuni fotoni si disperdano, andando in un'altra parte della stanza.

Si può dire quanti fotoni ci sono, dal momento che un fotone trasporta una certa quantità di energia; si può dunque sapere, in un modo di propagazione, quanti fotoni ci siano; il modo di propagazione è tuttavia un concetto intrinsecamente esteso nello spazio: si sa che all'interno del modo c'è una certa quantità di energia, dunque un certo numero di fotoni, ma non si può dire in quale punto dello spazio siano localizzati: per questo motivo, parlare di fotoni come di "palline", è sostanzialmente fuorviante. Lo studio della propagazione dei fotoni (della luce) nello spazio tuttavia è semplice da studiare con un approccio basato sulle equazioni di Maxwell, dunque sulla teoria ondulatoria, mentre è estremamente complicato partendo dal formalismo particellare; quest'ultimo torna invece particolarmente utile quando si deve studiare l'interazione dei fotoni con la materia; la trattazione sarà sostanzialmente incentrata sullo studio della luce come campo elettromagnetico.

Si parla, in ambito accademico o in ogni ambito, di "fotonica": dal momento che l'elettronica, nome attribuito alla scienza che studia il comportamento degli elettroni, è stato di estremo successo, è stato dato il nome "fotonica" a quella branca dell'ingegneria che studia la luce; nell'ambito del Politecnico di Torino, solitamente si lega il nome "fotonica" alla propagazione della luce, mentre il termine "optoelettronica" è più legato alla generazione/rilevazione di luce; questa distinzione in realtà non vale in ambiti più generali (un libro di fotonica potrebbe contenere entrambi gli argomenti senza introdurre particolari distinzioni: sia teoria ondulatoria/propagativa sia studio dell'interazione con la materia).

L'ottica si può dividere in due grandi branche:

- ottica classica: quella "storica", che si studia fin da Euclide, per quanto gli studi più avanzati in tal senso provengono probabilmente dal 1600 in poi, da personaggi come Galilei, Snell, Hamilton, Rayleigh; il 1600 è un periodo importante, poichè è il periodo in cui nasce un vero e proprio pensiero scientifico;
- ottica moderna: quella di cui si stava facendo cenno poco fa; parte sostanzialmente dal 1960, periodo in cui è stato introdotto il primo LASER.

Ci si potrebbe chiedere quale sia la vera differenza tra ottica classica e ottica moderna; questa, sostanzialmente, risiede nelle **sorgenti**.

Nell'ottica classica, si ha sostanzialmente a che fare con sorgenti termiche, le quali sono sorgenti a densità spettrale di potenza sostanzialmente piatta. Si dice che la luce si comporti come un *rumore bianco*, dal momento che il "bianco" è la sensazione visiva che fornisce un rumore di questo tipo, in

ambito ottico. Le sorgenti termiche sono a banda estremamente larga: dai 400 ai 700 THz circa (dal rosso al violetto). Sorgenti come i LED o i LASER hanno bande estremamente più strette.

Si noti che, spesso, per tradizione, il parametro che si utilizza per descrivere un sistema a microonde è la lunghezza d'onda λ , definita come rapporto tra la velocità di fase v_f e la frequenza f con cui si ha a che fare (se ne parlerà tra breve):

$$\lambda = \frac{v_f}{f}$$

Questa è una tradizione, che nasce sostanzialmente dalle misure: un tempo, per “misurare la frequenza della luce”, si utilizzavano degli interferometri, e ciò che si misurava con gli interferometri era la distanza tra due specchi: si parla dunque di misure di distanza, di spazio, non di frequenza (le quali sono sostanzialmente nate quando sono nati i contatori di eventi). In realtà la frequenza f è un parametro molto più interessante, dal momento che, se un sistema è LTI (Lineare Tempo Invariante), essa è invariante in ogni punto del sistema: avendo a che fare con un sistema che lavora a una certa frequenza, questa è sicuramente uguale in ogni punto del sistema; la stessa cosa non si può dire di λ , dal momento che, esplicitandone la formula, si ha:

$$\lambda = \frac{c}{n(f)f}$$

dove c è ovviamente la velocità della luce. $n(f)$ è il coefficiente di rifrazione, ed è una funzione della frequenza, dunque, a seconda del punto del sistema che si considera, λ sarà diversa. Si parla di sistemi WDM (Wavelength Division Multiplexed) quando si parla di sistemi a più portanti; in questo caso, in realtà, si ha a che fare con frequenze equispaziate, più che con λ equispaziate. Si noti che:

$$\Delta\lambda = -\frac{c}{f^2}\Delta f$$

dunque di fatto avere una Δf costante è diverso che avere una $\Delta\lambda$ costante.

I LASER sono sorgenti che invece hanno una banda estremamente stretta, confrontabile sostanzialmente con quella di un oscillatore a microonde.

Un LED è costituito da materiale attivo: un sistema alimentato e che dunque produce luce; anche un LASER in realtà è costituito da una parte attiva, ma contiene anche un risonatore (come d'altra parte tutti gli oscillatori, che sono costituiti da un materiale attivo e da un dispositivo che determini la frequenza). In un LASER ciò che determina la frequenza è,

per esempio, una cavità risonante: una cavità risonante è un risonatore per frequenze altissime, in quanto a parametri distribuiti. Dal momento che la larghezza di banda è tanto più piccola quanto più grande il risonatore, è possibile realizzare, regolando le dimensioni, dei risonatori con ottimo fattore di qualità.

Le tecniche apprese ai corsi di Campi Elettromagnetici vanno benissimo per studiare la luce emessa da LASER, ma vanno malissimo per studiare l'ottica classica, dunque sorgenti termiche: per studiare uno spettro elettromagnetico a banda larga come quello, sono più idonei approcci come quelli noti dai corsi di Comunicazioni Elettriche, basati sull'uso di funzioni di mutua correlazione e di autocorrelazione.

Il fatto che l'approccio elettromagnetico sia ottimo per studiare la luce LASER è sostanzialmente insito nella seguente affermazione: il LASER produce una luce **coerente**. Cosa significa coerente? Beh, per capirlo, si potrebbe usare la seguente equazione empirica:

$$\text{luce} + \text{luce} = \text{buio}$$

Cosa significa ciò? Beh, se si considera una stanza, e la si illumina con 1, 2, N lampadine, aumentando il numero di lampadine la luce continua ad aumentare, diventa sempre più uniforme. Se invece si usano due LASER, a certe condizioni, si può ottenere il buio: anzichè sommarsi in fase, le luci, essendo coerenti, possono sommarsi in controfase, ottenendo un fenomeno di interferenza distruttiva, dunque annullandosi. Non essendo la luce termica coerente, invece, questo non è sostanzialmente possibile.

I fenomeni di interferenza, per essere compresi, possono essere analizzati a partire da fenomeni più semplici da vedere; il più classico dei fenomeni è quello delle sorgenti di onde nell'acqua: avendo un punto che viene "alzato" e "abbassato" nell'acqua, si finisce per avere delle onde; se i punti sono 2, quando le onde si incontrano (quando le superfici a fase costante si incontrano), si può avere sollevamento ulteriore o abbassamento ulteriore dell'acqua. L'interferenza si vede studiando le superfici a differenza di fase costante rispetto ai punti di sorgente, e si ottengono profili come iperboloidi. Se la differenza di fase è 0 o è un multiplo di 2π , le due onde raggiungono il valore massimo contemporaneo, e quindi raddoppiano l'ampiezza della singola onda nel punto di incontro; se la differenza di fase è π , 3π , 5π , si ha controfase, dunque l'acqua è sostanzialmente piatta; questa stessa cosa si applica anche nell'ambito delle onde elettromagnetiche.

Studiando il libro di fisica nel quale si può studiare questo fenomeno elementare, si può osservare quale sia l'ipotesi fondamentale: il fatto che le perturbazioni centrali, le sorgenti, abbiano oscillazioni del tipo $\cos(\omega_0 t)$: il

fatto che la sorgente sia **monocromatica**; questo è, di fatto, l'unico caso nel quale sia possibile avere interferenza.

Un ulteriore esempio, più noto a chi ha studiato Campi Elettromagnetici, è quello del diagramma d'onda stazionario per una linea di trasmissione: volendo studiare il modulo della tensione $|V(z)|$ al variare della sezione z considerata sulla linea, si hanno massimi e minimi: questo accade dal momento che la tensione totale è costituita dalla somma di un modo progressivo e di un modo regressivo nella linea, che si sommano con una certa differenza di fase; quando sono in fase, dunque quando la differenza di fase è nulla, si ottiene un massimo; quando si sommano in controfase, dunque quando la differenza di fase è π , si ha il minimo; l'ipotesi che sta dietro a tutto ciò, tuttavia, è il fatto che si abbia a che fare con un generatore **sinusoidale**.

Se il generatore non fosse sinusoidale ma anche solo a banda stretta, si avrebbe un fenomeno diverso:

Quando si hanno piccole distanze dal carico, tutto resta sostanzialmente uguale; quello che si ha, tuttavia, avendo una banda $\Delta\omega$ non nulla, è avere tante λ diverse, una per ciascuna frequenza; per ciascuna λ ; di fatto, si ha uno zero; man mano che ci si allontana dal carico, i minimi tendono a riempirsi e, sommando i risultati, si ottiene una sorta di media, che va "adattare" il ROS, rendendolo circa costante.

In realtà vi sono casi intermedi: come vi è il caso di luce totalmente coerente, quando la sorgente è monocromatica, esiste anche il caso di luce totalmente incoerente, quando sostanzialmente è rumore (come nel caso delle sorgenti termiche).

I fenomeni di interferenza non riguardano solamente il trasmettitore, la sorgente, ma anche il ricevitore; se si considera infatti un LASER, incidente in modo quasi radente su un foglio di carta, si possono vedere tanti puntini, tante "macchioline" nere; questo deriva sostanzialmente dalla **rugosità** del materiale di incidenza: di fatto, se essendo la carta composta da fibre, si ha che il LASER interagisce con fasi diverse nei vari punti, dunque il ricevitore (l'occhio umano) riceve punti di luce e punti neri, ossia punti in cui si ha un'interferenza distruttiva; una luce del genere non può andar bene per leggere, dal momento che in certi punti va bene, in altri "toglie luce" alla carta. Uno specchio è una superficie idealmente con errori superficiali inferiori a $\lambda/20$ (o anche meno per specchi migliori): ciò che fa non è tanto la conducibilità o il n del materiale, dunque, quanto la sua rugosità; minore è la rugosità del materiale, maggiore sarà la similitudine tra il comportamento del materiale a quello di un riflettore che segue idealmente le leggi di Snell: non si ha rugosità, non si ha scattering. Le lampade alogene sono un esempio applicativo di ciò: una lampada alogena illumina il soffitto, il quale è una

superficie ruvida, dunque che scatterà, illuminando l'ambiente (andando i raggi diffratti in ogni direzione).

Si è parlato di coerenza; esistono due tipi di coerenza, per quanto riguarda la luce: coerenza temporale e coerenza spaziale.

Ha senso definire la prima delle due, la coerenza temporale, mediante la banda temporale, la quale a sua volta è definita a partire dalla teoria di Fourier:

$$v(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} V(\omega) e^{j\omega t} d\omega$$

si parla di ω e t come di due variabili “coniugate”: sono due domini, nei quali tuttavia si ha una certa relazione; di queste relazioni, fondamentale è il principio di indeterminazione: più una funzione è “larga” in un dominio, più “stretta” sarà nel dominio reciproco.

Cosa si può dire riguardo la coerenza spaziale? Una delle caratteristiche della luce LASER è quella di essere molto direttiva: il “punto” finale, che si vede, è molto piccolo. Fare ciò con una lampadina, con una torcia elettrica, non è possibile: si riescono a fare cerchi di una certa dimensione, non puntini: il fascio non si riesce a concentrare, dal momento che si disperde.

Al fine di comprendere il motivo per cui questo fascio si disperde, proviamo a introdurre un'analogia, basata sull'antenna a tromba: a sinistra dell'antenna a tromba, per alimentarla, si usa una guida d'onda, la quale è collegata a un generatore sinusoidale, che ha una certa ω_0 ; la guida, per ipotesi, è monomodale, ossia il campo sull'apertura è un modo TE_{10} : osservando il grafico solitamente presentato nei libri di Campi Elettromagnetici, si può vedere la classica rappresentazione del TE_{10} mediante un coseno. Prima di tutto si deve ricordare il fatto che questo è un disegno nel dominio della **frequenza**, ma che è ricollegabile facilmente al dominio del tempo. In ogni punto il campo elettromagnetico oscilla, con frequenza ω_0 . La cosa fondamentale è che in tutti i punti, l'oscillazione è simultanea: il ritmo di oscillazione è sempre lo stesso. In modo più formale, su tutta la faccia della tromba, c'è una ben precisa relazione di fase, cosa che deriva dalla presenza di un ben precisa pulsazione ω_0 dell'eccitazione.

Consideriamo un caso opposto: una lampadina. In una lampadina, dispositivo costituito da un filamento in cui si fa scorrere della corrente, gli atomi saltano su livelli superiori, che sono instabili; essendo instabili, si ha un decadimento degli elettroni, emettendo una luce con frequenza ν :

$$\nu = \frac{E_2 - E_1}{2\pi\hbar}$$

dove gli E_i sono livelli energetici e \hbar è la costante di Planck. Il punto essenziale è che, nella lampadina, ciascun atomo si comporta in modo indipendente dagli altri atomi (a parte in quelli molto prossimi): non si ha nessun rapporto di simultaneità nelle oscillazioni, dunque nessuna relazione di fase tra i vari elementi, tra i vari atomi: non si ha una struttura che oscilla all'unisono. Si hanno piuttosto delle piccole "celle" che internamente si comportano così, ma le varie celle sono tra loro con fase arbitraria, regolata da un processo casuale.

Per le antenne, come noto, per determinare il campo irradiato dalla bocca, dall'apertura, si può fare la trasformata di Fourier del campo all'apertura; ciò si potrebbe applicare anche alla lampadina, ma solo su ciascuna "cella", la quale sarà estremamente piccola; essendovi il principio di indeterminazione, essendo le celle molto piccole, allora la luce emanata sarà molto ampia, dunque il fascio sarà larghissimo, non collimato. In un LASER la cosiddetta "apertura" è sostanzialmente tale da avere coerenza su tutta la faccia, su tutto il "buco" dell'uscita: si ha la relazione di fase precisa. Il fascio LASER è tanto più direttivo quanto maggiore l'estensione del fascio iniziale.

Per capire a quanto si può arrivare, si consideri il seguente esempio: gli astronauti, nel periodo delle missioni lunari, lasciarono un catarifrangente sulla luna: esso è costituito non da uno specchio (dal momento che, per Snell, rifletterebbe l'onda secondo un angolo opposto rispetto alla normale della superficie), ma da elementi retrorifrangenti costituiti da strutture a L (vertici di cubi). Al fine di poter avere del segnale utile "indietro", è necessario avere un eco da questo catarifrangente lunare con un rapporto segnale-rumore molto elevato. Ciò che si può fare per rendere il fascio più collimato, più direttivo, è aumentare la sezione trasversale; questo si può fare mettendo il LASER nell'oculare di un telescopio, per esempio mediante un telescopio Cassegrain: ciò allarga il fascio, rendendolo più collimato in uscita per la trasformata di Fourier. Con questi metodi, si ottengono fasci sulla Luna larghi 20 chilometri, su una distanza di 400000.

Esistono diverse applicazioni: telecomunicazioni in fibra ottica, ma anche signal processing: un esempio di signal processing è la realizzazione della trasformata di Fourier di una certa distribuzione di campo.

Se si considera un sistema di questo genere, con una lente biconvessa, ponendo una *slide* (un'immagine, una fotografia per esempio), illuminata con una luce coerente, la distribuzione del campo nel piano focale posteriore non c'entra nulla visivamente, ma in realtà è la trasformata doppia di Fourier della slide. Mettendo dunque la lente "al contrario", invece che formarsi l'immagine all'infinito (come suggerirebbe l'ottica geometrica), si ottiene ciò.

Altri campi di applicazione riguardano l'industria meccanica: sempre più spesso si utilizzano LASER per saldare o tagliare metalli, plastiche, materiali

vari; quelli più diffusi, specialmente in questo ambito, sono i LASER ad anidride carbonica, che raggiungono potenze nell'ordine dei kW. Si stanno sempre più diffondendo, inoltre, i LASER a fibra ottica.

Altra applicazione. quella probabilmente più interessante, è quella delle comunicazioni ottiche: le fibre ottiche utilizzate come mezzi di trasmissione. Fondamentale è parlare di frequenza (o lunghezza d'onda che sia), dal momento che il comportamento dell'attenuazione per unità di lunghezza della fibra ottica non è indipendente da λ (si ricordi sempre che per λ si intende λ_0): per lunghezze d'onda nell'ordine dei 1550 nm, si ha un'attenuazione di 0,2 dB/km: si ha dunque a che fare con fibre ottiche estremamente trasparenti. Un vetro per finestre (che comunque è una struttura come si può immaginare e vedere abbastanza trasparente per le onde elettromagnetiche nel campo dell'ottico) è cinque ordini di grandezza più attenuante. Il fatto che l'attenuazione sia così alta in un vetro comune è dovuto alle impurità presenti nel vetro: eliminando le impurità, si è scoperto che è possibile ottenere valori così incredibili di attenuazione.

L'altro aspetto importante che riguarda le comunicazioni ottiche, aspetto in questo caso sfavorevole, è la dispersione: queste fibre ottiche sono dispersive, ossia hanno un comportamento che varia con la frequenza; introducendo dunque un insieme di bit, quello che si ottiene è un aumento della durata di ciascun bit, rischiando che, quindi, si abbia interferenza intersimbolica. Per evitare l'interferenza intersimbolica sarebbe necessario "distanziare temporalmente" i vari bit, ma ciò non ha senso, dal momento che ridurrebbe il bitrate, quindi le prestazioni della struttura guidante.

Ci sono sostanzialmente tre cause che possono portare alla dispersività della struttura guidante:

- la multimodalità: se sono presenti più modi di propagazione all'interno della guida, ciascuno di essi ha diverse caratteristiche propagative, diverse v_f o v_g ;
- la dispersione della guida d'onda;
- la dispersione del materiale: il materiale infatti ha un indice di rifrazione n che è funzione della frequenza f ; al variare della frequenza, dunque, variano anche velocità di fase e velocità di gruppo.

In una fibra d'ottica monomodale, il problema della dispersione dovuto alla presenza di più modi ovviamente non è presente; ciò che si ha ancora, tuttavia, sono i secondi due fenomeni, dei quali la dispersione causata dalle caratteristiche del materiale è di sicuro prevalente. Ciò che si può tuttavia

fare è studiare l'andamento di questa dispersione, definendo un “zero dispersion point”: si può vedere che, localmente, a $\lambda = 1300$ nm circa, la velocità di gruppo non cambia, o comunque varia in modo trascurabile; essendo tuttavia le due λ ‘ottime’ in valori diversi, servirebbe un compromesso; sono state inventate, per soddisfare questa problematica, le fibre ottiche a dispersione spostata, in modo che la dispersione sia minima a una λ prossima ai già citati “classici” 1550 nm. Si noti che 1550 nm è ottima anche per un altro motivo: a queste frequenze è possibile avere amplificatori ottici, ossia sistemi in grado di amplificare il segnale ottico senza che vi sia necessità di effettuare una conversione in elettronica per dover fare condizionamento.

1.2 Sviluppo storico dei fenomeni ottici

Si vuole, in questa sezione, introdurre uno sviluppo storico di varie branche della scienza coinvolte nell’ottica, proponendo nel mentre alcune riflessioni riguardo alcuni concetti fondamentali.

1.2.1 Propagazione della luce

- I primi sviluppi dell’ottica possono riguardare gli “antichi”: Empedocle ed Euclide, che propongono studi sulla propagazione della luce in linea retta, e alcuni racconti sugli specchi ustori (per quanto questi siano poco credibili, dal momento che servirebbero specchi con fuochi distanti centinaia di metri, dunque enormi lenti sostanzialmente piatte).
- Nel 1600 nasce il pensiero scientifico, dunque si inizia a chiedersi come funziona la natura, invece che “imporre” il proprio pensiero, “dicendo come funziona”. Si iniziano a formulare idee riguardo la pressione della luce: la luce, quando incide su una superficie, vi imprime una certa forza, una certa pressione; questa cosa è stata proposta in tempi più moderni per muovere ipotetiche astronavi mediante le vele solari: vele enormi che, a partire dalla forza impressa dal Sole, siano in grado di muovere le astronavi.
- Fermat introduce il proprio principio, secondo il quale la luce segue il percorso che richiede il minore tempo. In tal direzione, ha senso dire due parole sul metodo variazionale: un approccio alternativo alla soluzione di sistemi di equazioni differenziali, è basato sulla definizione di configurazioni nelle quali si trova un sistema, per esempio una configurazione A e una configurazione B ; il metodo variazionale è basato sul

minimizzare una certa quantità, che influisce la “traiettoria” di passaggio da una configurazione del sistema a un’altra. Scrivendo in questo caso l’equazione dell’ottica geometrica, minimizzando, si ottiene questa formulazione.

- Si parla di anelli di Newton: appoggiando una lente su una superficie piana, la distanza tra la lente e la superficie piana è diversa nei vari punti: al centro è nulla, aumentando la distanza aumenta. I raggi luminosi dunque si muovono facendo percorsi diversi, si incontrano strati di aria di spessore diverso, e a seconda del fatto che l’interferenza sia costruttiva o distruttiva, si vedono anelli luminosi o scuri. Hooke ha scoperto questi anelli, Newton li ha studiati.

- Hooke è anche tra i primi a osservare e studiare gli effetti della diffrazione; spendiamo, a questo punto, alcune parole sulla diffrazione.

Cos’è la diffrazione? La diffrazione è sostanzialmente la “realtà”, “meno l’ottica geometrica”. Il modello sul quale si basa di solito il nostro studio dell’ottica è proprio l’ottica geometrica, la quale, come già detto, essendo un modello, ha un campo di validità limitato. Si immagini una situazione di questo genere:

Secondo il modello dell’ottica geometrica, inviando dei raggi (supponendo che sia tutto omogeneo, quindi raggi “dritti”), si dovrebbe avere, sullo schermo dietro la fessura, un’illuminazione “a porta”: i raggi vanno dritti e illuminano, mentre tutto il resto dovrebbe essere una zona d’ombra. Ciò che capita veramente, però, è qualcosa di diverso: si hanno, anche a una certa distanza dalla proiezione della fessura, contributi di illuminazione non nulli: sebbene i raggi, le linee di flusso dell’energia, siano teoricamente (per l’ottica geometrica) solo in quella regione, si ha una apertura del fascio, del fronte d’onda: ciò non è assolutamente prevedibile dall’ottica geometrica. Il risultato finale è una sorta di sinc (seno cardinale). Questa cosa è particolarmente evidente quando il buco è piccolo: essendo le λ che andiamo a considerare dell’ordine del mezzo micron, se il buco è largo qualche lunghezza d’onda, la cosa è particolarmente visibile.

Un fenomeno del genere si può vedere anche a livello molto più macroscopico (ai tempi di Euclide di sicuro non era facile realizzare buchi grandi quanto la lunghezza d’onda), per esempio andando al mare: se si va in Liguria, per esempio, si hanno delle barriere di scogli, con delle aperture di una ventina di metri; questa apertura è dell’ordine della λ delle onde marittime (essendo λ la distanza tra due creste dell’onda,

dunque una ventina di metri circa anche in questo caso); applicando il modello dell'ottica geometrica (cosa fattibile, dal momento che comunque si ha a che fare con dei fenomeni ondulatori di un qualche tipo), si dovrebbe avere un'onda d'acqua non nulla solo in prossimità della proiezione dell'apertura sulla spiaggia, ma quello che si osserva in pratica è un allargamento del fronte d'onda il quale è, a tutti gli effetti, un fenomeno di diffrazione dell'onda marina.

Osservazione finale per questo punto: quando si ha un puntatore LASER, quella che si propaga nello spazio è un'onda piana. La cosa sembra assurda: sullo schermo vediamo un puntino, e un'onda piana è invece una configurazione di campo con superficie di fase infinitamente estesa trasversalmente; se il puntino è esteso su un centimetro, qualche millimetro, sembra insensato che esso sia "infinitamente esteso". Quando si parla di Campi Elettromagnetici, tuttavia, le dimensioni non vanno rapportate a ciò che umanamente si pensa grande o piccolo, bensì alla λ considerata: se in questo caso $\lambda \sim 0,5 \mu\text{m}$, qualche centimetro significa avere decine di migliaia di lunghezze d'onda: la larghezza del fascio, dunque, tendenzialmente, è infinita (estremamente estesa). Quando si calcolano i coefficienti di riflessione a un'interfaccia, per avere un'onda luminosa piana, si può dunque utilizzare un LASER, che produce sostanzialmente un'onda piana. Per un esperimento di questo tipo, bisogna tenere conto di varie ipotesi: quando si fa un conto analitico, bisogna conoscere l'indice di rifrazione della superficie di discontinuità (e deve essere ben noto), ma non solo: un LASER non è esattamente un'onda piana, bensì un *fascetto* di onde piane: ciò che si può fare è considerare la presenza di un impulso gaussiano con una portante che ne segue l'inviluppo. Si può¹ decomporre l'impulso nel dominio della frequenza, utilizzando la teoria della trasformata di Fourier: le equazioni delle linee valgono per segnali sinusoidali, monocromatici. La trasformata di Fourier di questo segnale è dato da un contributo intorno a una certa portante ω_0 , dunque si hanno sinusoidi attorno a ciò. Il fascio gaussiano del LASER ha dunque non solo un'onda piana, ma anche alcune "onde piane" vicine. Per ciascuna onda piana, ciascuna frequenza, si calcola il coefficiente di riflessione, si somma tutto e antitrasforma.

- Nel 1666 Newton, personaggio oramai estremamente importante per il suo tempo, ipotizza che la luce abbia un comportamento particellare, ossia che sia composta da particelle; questo è, per il tempo, un abbaglio

¹Orta R. - Teoria delle linee di trasmissione - Cap. 8

enorme, ma di fatto poco contestabile, dal momento che il personaggio ai tempi era troppo importante per essere contestato. Si ricorda che, fino a qua, la luce è considerata come un'entità a velocità infinita.

- A partire dallo studio delle orbite dei satelliti di Giove, osservando le eclissi (nate dal fatto che, per il loro moto, i satelliti finivano nascosti dal loro pianeta), si notano irregolarità, ritardi, e questo derivava dal fatto che la luce impiegava un certo tempo per arrivare da Giove alla Terra; si riesce a effettuare una prima misura della velocità della luce, dunque, sfruttando le nozioni sulle distanze interplanetarie.
- Nel 1678 Huygens nota proprietà di polarizzazione della luce: la luce polarizzata lungo direzioni perpendicolari non interagisce: non si può avere interferenza nè costruttiva nè distruttiva. Questo fenomeno viene scoperto studiando i cristalli, come il quarzo: il quarzo è un materiale con una struttura cristallina, dunque con un reticolo regolare: si hanno velocità di propagazione diverse a seconda della direzione in cui l'onda piana si propaga, e diverse polarizzazioni: effetti di anisotropia. La manifestazione più evidente dell'anisotropia è la birifrangenza: se si prende un cristallo trasparente (per esempio lo "spato d'Islanda"), si hanno due raggi identici, traslati: due fenomeni di rifrazione con un'unica onda incidente.
- Sempre a Huygens si deve il cosiddetto principio di Huygens: ogni punto di un fronte d'onda è sorgente di onde microscopiche che per effetto dell'interferenza danno luogo a un fronte d'onda dopo; questo principio è collegato al cosiddetto **teorema di equivalenza**.
- Studio della aberrazione da stelle fisse: le stelle fisse sono stelle vere e proprie, non pianeti (pianeta, dal greco, "errante", mobili rispetto alle stelle fisse); l'aberrazione è una anomalia: questo fenomeno si osserva anche, per esempio, quando si corre con un ombrello: bisogna piegare l'ombrello, quando si corre, perchè si ha una velocità che si compone con quella di caduta verticale della pioggia, dunque quando si ha una composizione di velocità, rispetto al sistema di riferimento in cui noi siamo origine, si ha che la pioggia scende con un certo angolo. La Terra si muove con 30 km/s, dunque le stelle sembrano descrivere una piccola ellisse attorno alla loro posizione media del cielo: un'immagine dell'orbita terrestre. Bradley, studiando questo effetto, riesce a proporre una stima della velocità della luce.

- Eulero, nel 1746, spinge per una teoria ondulatoria della luce: essendo un personaggio piuttosto importante per l'epoca, la sua opinione è importante.
- Young parla di "iridescenza": si tratta di ciò che capita con i film sottili, che generano interferenza: questo fenomeno per esempio si può vedere quando si ha della benzina su una pozzanghera: a seconda della posizione dell'osservatore si vede un diverso colore, e ciò è causato dalla presenza di questi film sottili. A seconda dello strato e dell'angolo di osservazione, si vedono colori diversi, e questo per l'interferenza.
- Nel 1818, Fresnel distrugge la teoria particellare, proponendo tra le altre teorie l'esperimento della macchia illuminata; si inizia a parlare di diffrazione, che poi porta alla GTD (Geometric Theory of Diffraction): di tutti i raggi, ve ne è una serie che colpisce esattamente il bordo; questi raggi, colpendo il bordo, cambiano direzione; questa cosa è evidente, osservando una tazzina di caffè illuminata dall'alto: se si guarda, intorno al bordo, si vede un anello luminoso; questo è dovuto al fatto che il cerchio è una caustica per il modello di ottica geometrica, generato dal fatto che i raggi luminosi illuminano il bordo, vengono diffratti, e il campo aumenta in questo anello. Fenomeni di questo tipo si vedono per esempio anche nelle antenne a paraboloide, quando si parla di *back lobes*: questi si osservano, a causa della presenza di parte del campo di illuminazione che interagisce con i bordi del paraboloide.

Per capire cosa sia la luce, un'idea interessante da riprendere è quella della pressione: come detto, la luce esercita una forza, dunque una pressione, sulla superficie sulla quale incide; il fatto che si abbia una pressione, comporta la presenza di onde di taglio (onde trasversali) e onde di pressione (onde longitudinali). La polarizzazione è una quantità che dipende dalle direzioni trasversali a quella di propagazione; la luce è una pressione, dunque si può pensare a un fluido: in un fluido si hanno solo onde di pressione, onde longitudinali, ma non di taglio, trasversali. Se si batte un pugno su un tavolo nella direzione del pugno si manda un'onda di pressione, ma nel mentre si hanno anche onde di taglio. Se non si ha interferenza, significa che le onde interferenti non sono quelle longitudinali, ma quelle trasversali: quelle che danno la proprietà di polarizzazione. L'interferenza o meno delle onde dunque non dipende dalle caratteristiche longitudinali dell'onda, ma da quelle trasversali.

In un fluido, le onde sono solo longitudinali: sembrerebbe dunque che l'etere, definito precedentemente, sia un solido, poichè solo nei solidi si possono avere onde trasversali, onde che permetterebbero l'iden-

tificazione di una proprietà di polarizzazione; ciò ha creato ulteriore scompiglio.

- Sono stati notati, da Fresnel e Cauchy, indicazioni sull'origine della dispersione, basandosi sulla struttura molecolare della materia (idea estremamente innovativa per l'epoca): in effetti n dipende proprio dalla struttura atomica, dal fatto che si ha una dinamica nella materia, presenza di comportamenti risonanti.
- Nel 1870 Tyndall introduce la prima struttura ottica guidante: un serbatoio con un buco, illuminato: la luce era guidata dentro al getto d'acqua uscente dal buco.

1.2.2 Teoria elettromagnetica

Parallelamente alle scoperte nel campo ottico, ne sono state fatte anche nell'ambito elettromagnetico, come già detto molto imparentato con l'ottica.

- In principio non si poteva parlare di elettromagnetismo, non avendo ancora scoperto l'accoppiamento tra i due tipi di fenomeni. Una prima osservazione che si vuole fare riguarda il campo magnetico: le prime osservazioni fatte, nella storia, riguardavano infatti fenomeni legati al magnetismo: il fatto che un oggetto di ferro possa essere attirato da della magnetite. Il motivo per cui la calamita attira un pezzo di ferro, il punto essenziale, è il fatto che il campo magnetico generato dalla calamita è non uniforme: se fosse uniforme, infatti, si avrebbe solo un orientamento del pezzo di ferro, senza però spostarsi, anche perchè non si avrebbe una direzione di spostamento, essendo il campo ovunque uguale; nella calamita il campo è più forte che da altre parti. Un fenomeno duale è uguale per l'ambra, o per la plastica, che attira, se caricata elettrostaticamente, dei pezzettini di carta, esercitando una forza.
- Faraday inizia i propri esperimenti, esperimenti che avrebbero permesso a Maxwell di teorizzare.
- Nel 1873 Maxwell propone la propria sintesi dei fenomeni elettrici e magnetici, introducendo il noto termine di accoppiamento, noto come "corrente di spostamento":

$$\frac{d\mathcal{D}}{dt}$$

questo termine, a frequenze sufficientemente elevate, è non nullo, quindi accoppia le equazioni di Maxwell tra loro, rendendo campo elettrico e campo magnetico due “facce della stessa medaglia”. Si noti che nel testo di Maxwell non si parla di vettori: i vettori sarebbero stati inventati solo dopo da Gibbs e Heaviside, con grandi contestazioni dal gruppo di inventori del precedente calcolo tensoriale (Levi-Civita, Ricci). Nella teoria circuitale, Heaviside si è anche reso famoso grazie al metodo operativo per lo studio dei circuiti in AC, basato sulla trasformata di Laplace.

- Mediante misure di capacità (condensatori) si riesce a determinare una misura di ε_0 , mediante misure di induttanza (bobine) si riescono a ottenere misure di μ_0 ; si trova:

$$v = \frac{1}{\sqrt{\varepsilon_0 \mu_0}}$$

che, guarda caso, è simile alle misure fatte precedentemente: la luce, dunque, è effettivamente riconducibile a fenomenologie di tipo elettromagnetico.

- Tra le varie problematiche lasciate in sospeso, c’era il problema dell’etere: pare che l’etere fosse un mezzo dalle proprietà assolutamente contraddittorie, dal momento che doveva essere immobile, e solido, ma la Terra si muove, dunque la cosa non sembrerebbe avere senso; data inoltre una $v = c$, le proprietà di elasticità/rigidità sarebbero dovute essere estreme, molto più di quelle dell’acciaio.
- La soluzione dell’etere viene proposta da Einstein: l’etere non esiste (come evidenziato da Michelson e Morley qualche anno prima). A questo punto rimane un problema aperto: le onde elettromagnetiche sono sì onde, dunque vibrazioni, oscillazioni, ma non di materia: cos’è che oscilla, se esse si possono propagare anche nel vuoto, dunque in assenza di energia? In effetti, è difficile trovare risposta a una domanda del genere.

1.2.3 Interazione luce-materia e ottica moderna

Si vogliono introdurre rapidamente, essendo meno importanti o già citati, alcuni concetti conclusivi.

- Un fenomeno sicuramente importante per quanto riguarda l’interazione tra luce e materia, è quello scoperto da Fraunhofer: la presenza di righe

nere nello spettro della luce del sole; queste righe nere sono provocate dall'assorbimento della luce da parte di particelle localizzate negli strati alti dell'atmosfera, portando dunque a non vederne più il contributo a terra. Questo è stato poi utilizzato per identificare i materiali: a seconda della posizione spettrale delle righe di assorbimento, si può identificare l'elemento che ha interagito con l'onda, assorbendola. Ciò dipende dal fatto che una certa lunghezza d'onda permette di far fare una precisa transizione a un elemento.

- Per quanto riguarda l'ottica moderna, una prima scoperta riguarda la scoperta dei modi di propagazione in una fibra ottica.
- Miller effettua il primo passo nell'ottica geometrica, proponendo circuiti simili sostanzialmente a quelli in "microstriscia", però a frequenze ottiche: si tratta di lastre di dielettrico in cui sono realizzate guide d'onda.
- Si parla di DFB LASER: LASER a Distributed FeedBack, ossia si iniziano a usare risonatori distribuiti (come in cavità).

1.3 Cenni sui reticoli di diffrazione

Uno degli argomenti che verranno introdotti nella trattazione è quello dei reticoli di diffrazione: essi sono utilizzati molto spesso, anche in applicazioni assolutamente banali, come la carta da regalo, o i riflettori dei CD. I reticoli di diffrazione sono utilizzati per la realizzazione di risonatori distribuiti, DFB.

Un reticolo di diffrazione può essere costruito mediante una struttura di questo tipo:

Un CD può servire come specchio: tenendolo in una certa posizione, si può vedere della luce riflessa. Questo sfrutta un meccanismo di "ordine 0": si vede la lampada, riflessa di un angolo ϑ_i , con quindi angolo di incidenza e riflessione uguale.

Esistono altri ordini di diffrazione? Ordine di diffrazione -1: la luce arriva con un certo angolo, e la luce torna indietro con lo stesso angolo di incidenza, parallelo. Mettendo il CD in una certa posizione, si vede effettivamente il CD illuminato, senza distinguere la forma della lampada, ma vedendo una riga colorata: questa deriva dal meccanismo di diffrazione di ordine -1. Si può far vedere che la direzione di orientamento della freccia dipende dalla lunghezza d'onda: a seconda della posizione, essendoci dipendenza dalla lunghezza d'onda, si può vedere, con vari valori di inclinazione a partire dal primo, con vari colori; dopo una certa inclinazione, tuttavia, si rivede lo stesso

colore già visto: cambiando l'inclinazione, prima si sfrutta l'ordine -1 per tutti i vari angoli, per tutte le lunghezze d'onda; aumentando oltre una certa inclinazione, si rivedono le stesse frequenze di prima, con un ordine superiore, -2. L'angolo di propagazione delle varie componenti di colore dipende dalla lunghezza d'onda del colore stesso: questo permette di effettuare un'analisi spettrale! Di scomporre la luce in diverse componenti spettrali, facendo una sorta di trasformata di Fourier al variare dei vari angoli (questo è ciò che si usa nei telescopi moderni).

Si ha:

$$\underline{k}_{\text{incidente}} = k_x^i \hat{x} + k_z^i \hat{z}$$

dove l'indice "i" sta per "incidente": questo è il \underline{k} dell'onda incidente sul reticolo di diffrazione. I vari ordini, etichettati da un certo intero n , hanno un \underline{k} di questo tipo:

$$\underline{k}^n = k_x^n \hat{x} + k_z^n \hat{z}$$

questa è, per ciascun valore di n , l'espressione del vettore d'onda diffratto. In questo caso:

$$k_x^n = k_x^i + n \frac{2\pi}{d}$$

dove d è il periodo di disposizione delle varie "fosse" del reticolo. Se ho un ordine negativo, addirittura "tolgo" al k_x^i un valore, facendo tornare indietro l'onda. A forza di aumentare, i k_z diventano immaginari, ottenendo onde coincidenti con i modi sotto taglio (onde evanescenti).

I reticoli di diffrazione possono anche essere di più dimensioni, come il puntatore "a scacchiera" del LASER, dove vi sarebbe un reticolo di diffrazione bidimensionale, per l'appunto a scacchiera. Si hanno, in questi casi, due periodicità: d_x e d_y . Il motivo per cui si hanno varie onde riflesse deriva dal fatto che si ha una struttura periodica. Tornando per un attimo al caso monodimensionale, si consideri il semplice reticolo, in x : se il reticolo di diffrazione è una struttura periodica, con una periodicità in x di d , allora è possibile espandere la funzione del reticolo $f(x)$ in serie di Fourier:

$$f(x) = \sum_{n=0}^{\infty} c_n e^{-jn \frac{2\pi}{d} x}$$

Volendo fare lo spettro, la trasformata di Fourier di questa funzione periodica (spaziale, ma esattamente identica), si ottiene un pettine di δ , la cui ampiezza segue l'involuppo dettato dalla trasformata di Fourier della singola

cella del reticolo. Essendo la cella del reticolo un “rettangolo”, si avrà a che fare con la trasformata di una porta, dunque di una sinc.

In questo caso, si hanno le seguenti caratteristiche, per questo spettro:

- la distanza tra due δ , tra due impulsi, è:

$$\Delta f = \frac{2\pi}{d}$$

- la posizione dello zero dell’involuppo, è legato alla trasformata della singola cella, alla sinc, supposta spessa d_1 ; si ha dunque:

$$f_{\text{zero}} = \frac{2\pi}{d_1}$$

essendo poi d_1 più piccolo di d , gli inversi sono nella relazione opposta: ci sono diversi impulsi, prima dello zero!

La posizione $n\frac{2\pi}{d}$ è proprio la distanza delle due δ , di due impulsi, la base per il calcolo degli ordini, è la posizione della n -esima δ rispetto alla fondamentale. Nel caso si ha $k_x = 0$, tutto a posto; nel caso $k_x \neq 0$, è come se tutta la struttura venisse traslata di un certo parametro.

I vari punti, le varie δ , corrispondono a delle direzioni di propagazione: se la struttura è bidimensionale, invece di avere un rastrello si ha un “letto da fachiro”, tante δ su una superficie bidimensionale. Ciascun \underline{k}^n è una direzione, diversa, che andrà sul muro a identificare dei punti: ciascun “punto” è un’onda piana, nata dal reticolo di diffrazione.

Si ha un involuppo: questo mostra che le onde, allontanandosi dalla centrale, vanno “a morire”: allontanandosi dalla δ centrale la sinc si attenua, quindi la luce proiettata sul muro si attenua con essa. Si noti che:

- il fattore di attenuazione delle onde, allontanandosi dalla centrale, dipende esclusivamente dallo scavo: è lo scavo che detta l’involuppo delle ampiezze delle δ ;
- la geometria del fascio, dunque la posizione dei puntini proiettati, dipende dalla periodicità del reticolo.

Questa cosa potrebbe ricordare una schiera di antenne, dove l’elemento singolo è il pozzetto, il fattore di schiera è regolato dalla periodicità.

Nota: se ruotiamo il LASER, ruotiamo anche il fascio sul muro. Banale? No, assolutamente: questa è la dimostrazione sperimentale di una proprietà

delle trasformate di Fourier bidimensionali: la distribuzione di campo a grande distanza, come nel caso delle antenne, è la trasformata di Fourier della distribuzione di campo sull'antenna, sul LASER (approssimazione di far field sull'integrale di irradiazione); ruotando dunque la distribuzione di campo sull'antenna, ruota anche la trasformata di Fourier. Questa cosa non è banale, dal momento che non è detto che l'operatore di rotazione (operatore lineare) commuti con l'operatore trasformata di Fourier (anch'esso operatore lineare): ciò che ci aiuta in questo caso è il fatto che il nucleo della trasformata di Fourier in questo caso sia un $\underline{k} \cdot \underline{r}$, e questo termine è invariante per rotazione: ruotando il sistema di riferimento, il prodotto scalare non cambia, essendo il prodotto scalare invariante per rotazione.

Si faccia ora un altro esperimento: si consideri un cubo di materiale trasparente, e si metta per esempio un dito davanti a esso: il dito, illuminato dalla luce esterna, sarà sicuramente visibile attraverso il vetro, essendo esso trasparente. Si può tuttavia anche osservare il dito attraverso una delle facce laterali del vetro: quello che accade è avere una doppia superficie di discontinuità, ma quella che ci interessa è sostanzialmente la seconda: se la luce, entrata attraverso la prima superficie, vuole essere osservata al di sopra di un certo angolo, detto "angolo limite", il campo dall'altra parte è un'onda evanescente. Se si schiaccia il dito, nel punto a maggiore pressione si avrà visione dell'onda, ma allontanandosi, nei punti "meno schiacciati", c'è un'ombra: questo deriva dal fatto che il dito in certi punti è più lontano e, essendovi un modo evanescente, la costante di decadimento del campo evanescente (il quale si attenua esponenzialmente) è tale da far andare "in ombra" parte del dito. Ciò si può provare anche con un bicchiere, se le superfici esterne sono asciutte!

L'effetto della riflessione totale, e dell'effetto del dito che si vede e non si vede attraverso il cubo/bicchiere, è una manifestazione dell'effetto tunnel: pensando all'elettrone come a un oggetto con proprietà ondulatorie (secondo il dualismo di De Broglie), è evidente il fatto che si possa avere un effetto di questo tipo. Questo effetto è molto usato per esempio in questo ambito:

Se si ha una coppia di prismi separati in aria, se il sistema è progettato in modo che si abbia riflessione totale, si ha un'onda riflessa e nient'altro; nel caso ci sia però l'altro prisma, ciò che rimane del modo evanescente entra nel secondo prisma e lì si può propagare: questo è un divisore di potenza ottico. A seconda della distanza tra i due prismi, si può capire quanto accoppiamento si ha. L'accoppiamento non può essere troppo sbilanciato, a causa della rugosità dell'oggetto: la rugosità di fatto introduce un fondo di radiazione, trasmettendo in tante direzioni.

La riflessione totale, come vedremo, è il meccanismo che permette il guidaggio della fibra ottica: se si riflette al di sopra dell'angolo limite, i raggi

non escono, bensì rimangono confinati nella struttura guidante. Si scopre che nella fibra ci sono dei modi di propagazione: configurazioni di campo che si mantengono invariate mentre si propagano nella direzione longitudinale. Una sorgente eccita i modi di una struttura, in misura diversa.

Si può dire che i modi siano una “tastiera”, e la sorgente sia il “pianista”: il pianista suona, ed eccita alcuni modi. Misurando il campo in un punto, avrò la somma di tutti i modi eccitati. Questa è una rappresentazione modale, ed è diversa da quella raggistica; si può tuttavia dimostrare che ogni modo è un insieme di raggi, e ogni raggio si può descrivere con un insieme di modi, e tutto ciò si può rappresentare mediante una trasformata: la “formula di Poisson”: la formula di Poisson è quella che converte l’interpretazione raggistica nell’interpretazione modale.

Capitolo 2

Ottica dei mezzi stratificati

2.1 Equazioni di Maxwell nel dominio del tempo e della frequenza

Si vuole a questo punto introdurre la base dei discorsi che verranno affrontati nella trattazione, partendo “da zero”: dalle equazioni di Maxwell. Per equazioni di Maxwell si intende, nel modo più proprio, la coppia (dunque due) di equazioni differenziali *del rotore*:

$$\begin{cases} \nabla \times \underline{\mathcal{E}}(\underline{r}, t) = -\frac{\partial}{\partial t} \underline{\mathcal{B}}(\underline{r}, t) \\ \nabla \times \underline{\mathcal{H}}(\underline{r}, t) = \frac{\partial}{\partial t} \underline{\mathcal{D}}(\underline{r}, t) + \underline{\mathcal{J}}_e(\underline{r}, t) \end{cases}$$

queste sono le **due** equazioni di Maxwell dal momento che, quelle comunemente dette “equazioni della divergenza”, sono una conseguenza di queste e della legge di continuità. È dunque ragionevole richiedere che valga anche la seguente equazione, detta **legge di continuità**:

$$\nabla \cdot \underline{\mathcal{J}}_e + \frac{\partial \rho}{\partial t} = 0$$

infatti, la carica non “evapora” nè proviene dal nulla: tutta la carica che sta in un posto, finisce in un altro posto.

Si noti l’uso dei caratteri corsivi: i caratteri corsivi, nella notazione utilizzata nella trattazione, indicano una dipendenza delle funzioni dal tempo.

Spesso, i campi di interesse sono armonici, dunque dipendenti armonicamente dal tempo; in altre parole, sinusoidali. Ha dunque senso definire i fasori come:

$$\underline{\mathcal{E}}(\underline{r}, t) = \operatorname{Re} \{ \underline{E}(\underline{r}) e^{j\omega_0 t} \}$$

$\underline{E}(\underline{r})$ è il fasore, ed è un vettore complesso, le cui componenti sono numeri complessi; questo evidenzia una dipendenza armonica, sinusoidale, dal tempo.

Si noti che in diversi testi (per esempio di fisici) si preferisce usare, come nucleo per il passaggio da fasore a grandezza nel tempo, $e^{-j\omega_0 t}$; questo comporterà delle differenze che verranno discusse in seguito.

Come noto, è possibile convertire nel dominio dei fasori le suddette equazioni, ottenendo le equazioni di Maxwell nel dominio della frequenza, a partire dalla seguente proprietà:

$$\frac{\partial}{\partial t} \iff j\omega_0$$

quindi

$$\begin{cases} \nabla \times \underline{E}(\underline{r}) = -j\omega_0 \underline{B}(\underline{r}) \\ \nabla \times \underline{H}(\underline{r}) = j\omega_0 \underline{D}(\underline{r}) + \underline{J}_e(\underline{r}) \end{cases}$$

In questo modo, la dipendenza dal tempo è sparita, ottenendo espressioni molto più semplici da maneggiare. Questo presuppone tuttavia una cosa particolarmente scomoda: la presenza di segnali puramente sinusoidali. Nel caso si abbia a che fare con segnali più complicati, come un impulso, questa trattazione non basta. Ciò che si può fare, in questo caso, è proporre un metodo diverso per la conversione in frequenza, considerando $\underline{\mathcal{E}}(\underline{r}, t)$ come antitrasformata di Fourier di una certa funzione:

$$\underline{\mathcal{E}}(\underline{r}, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \underline{E}(\underline{r}, \omega) e^{j\omega t} d\omega$$

In questo caso, si può applicare l'operatore "derivata" al campo, e ottenere ciò:

$$\frac{\partial}{\partial t} \underline{\mathcal{E}}(\underline{r}, t) = \frac{\partial}{\partial t} \frac{1}{2\pi} \int_{-\infty}^{+\infty} \underline{E}(\underline{r}, \omega) e^{j\omega t} d\omega$$

in questo caso, l'operatore di derivazione opera solo sull'esponenziale, dunque si ottiene qualcosa di molto simile a prima:

$$\frac{\partial}{\partial t} \iff j\omega$$

Si noti una differenza visivamente sottile: prima si aveva una ω_0 , ossia un valore ben definito di frequenza: la frequenza di oscillazione del campo elettromagnetico; ora, invece, ω è un numero qualsiasi. Utilizzando questa

seconda rappresentazione (basata sulla trasformata di Fourier invece che sui fasori), appaiono le stesse equazioni, con solo ω invece di ω_0 .

Le due rappresentazioni sono in realtà imparentate tra loro: l'integrale infatti è una somma nel continuo, una somma infinita, che può essere approssimata introducendo una discretizzazione dell'asse della frequenza:

si può dire che:

$$\underline{\mathcal{E}}(r, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \underline{E}(r, \omega) e^{j\omega t} d\omega \sim \sum_i \frac{1}{2\pi} \underline{E}(r, \omega_i) \Delta\omega e^{j\omega_i t}$$

Si noti che qua non si ha a che fare con la parte reale, ma si ha questa espressione; quella, deriva semplicemente dal fatto che prima si consideravano assieme, inglobate, frequenze positive e negative, ordinando i contributi positivi e quelli negativi in maniera tale da ottenere un risultato complessivamente reale.

Si consideri un segnale di questo tipo:

questo significa sostanzialmente avere una luce nulla prima e dopo un certo punto (impulso), con un certo colore, dettato dalla frequenza della cosinusoide. Volendo calcolare $E(\omega)$ per questo, si ha a che fare con la somma di due sinc traslate in frequenza su $\pm\omega_0$; questo oltretutto, tenendo conto che, con l'interpretazione appena introdotta, è stato introdotto un campionamento della frequenza. Cosa introduce il “ \sim ”, invece della semplice eguaglianza? Beh, sostanzialmente due motivazioni:

- la \sum_i è finita: si avrà a un certo punto un troncamento, che porta a un fenomeno di Gibbs: quello che si ha è un overshoot, la cui altezza è costante, e il cui spessore dipende dal numero di armoniche considerate; maggiore è il numero di armoniche mantenute, minore è lo “spessore” dell'overshoot, a parità di altezza;
- discretizzando, si hanno tanti *alias*; se $\Delta\omega \rightarrow 0$, le copie (nel dominio del tempo), distanziate tra loro $\frac{2\pi}{\Delta\omega}$, sono molto lontane, idealmente all'infinito; questa cosa può essere importante quando si devono fare calcoli numerici: se $\Delta\omega$ è per esempio troppo grande, o se si han finestre eccessivamente larghe, si finisce per rischiare di ottenere risultati numerici errati.

Per quale motivo utilizziamo le sinusoidi come segnali, come “base” della nostra trattazione? La motivazione è abbastanza semplice, concettualmente: le sinusoidi o, meglio, gli esponenziali complessi, sono funzioni interessanti, dal momento che si comportano “bene” rispetto alle derivate. Cosa vuol dire questo? Vediamo:

$$\frac{\partial}{\partial t} e^{j\omega t} = j\omega e^{j\omega t}$$

gli esponenziali sono autofunzioni dell'operatore di derivazione. Vediamo cosa significa: dall'algebra lineare, si sa che

$$\underline{\underline{A}}u = \lambda u$$

u è l'autovettore: si tratta semplicemente del vettore tale per cui, se gli si applica l'operatore $\underline{\underline{A}}$, rappresentato mediante la sua matrice, produce in uscita un altro vettore parallelo a esso, quindi con la stessa direzione, al più con norma diversa (λ). Estendendo alle funzioni e agli operatori differenziali, se si ha un operatore che non modifica la funzione, allora questa è autofunzione del medesimo: la funzione finale è proporzionale a quella di partenza, ossia uguale a meno di un certo fattore di proporzionalità. Questo è il motivo per cui la rappresentazione spettrale è utile: l'operatore di derivazione è sparito. Questo, in realtà, vale solo se si ha a che fare con equazioni differenziali a coefficienti costanti.

Nota finale: sotto certi punti di vista, fasori e trasformate di Fourier sono leggermente diversi: un fasore passa da un $\underline{\mathcal{E}}$ a un \underline{E} , dunque si ha:

$$[\underline{E}]_{\text{fasori}} = \frac{V}{m}$$

per quanto riguarda la trasformata di Fourier, in realtà si passa da $\underline{\mathcal{E}}$ a $\underline{E}\Delta\omega$, dunque in realtà si ha qualcosa di leggermente diverso:

$$[\underline{E}]_{\text{Fourier}} = \frac{V}{m \text{ Hz}}$$

in questo caso dunque si ha a che fare con una densità spettrale di campo, non esattamente con un campo, almeno sotto il punto di vista dimensionale.

2.2 Mezzi materiali

Le equazioni di Maxwell sono due equazioni, in quattro incognite: campo elettrico $\underline{\mathcal{E}}$, campo magnetico $\underline{\mathcal{H}}$, induzione elettrica (anche detta spostamento elettrico o spostamento dielettrico) $\underline{\mathcal{D}}$, induzione magnetica $\underline{\mathcal{B}}$. In realtà, le induzioni sono legate ai rispettivi campi: le induzioni, come il nome stesso suggerisce, sono ciò che si ha dall'interazione di un campo con un qualche mezzo materiale. Sotto il nostro punto di vista, dunque, le grandezze fondamentali saranno i campi, $\underline{\mathcal{E}}$ e $\underline{\mathcal{H}}$.

Si può utilizzare un approccio da “teoria dei sistemi”, e avere qualcosa di questo genere:

Il mezzo materiale è come una “scatola”, un “sistema”, nel quale si introduce una certa quantità, e che ne produce un'altra in uscita. A seconda del tipo di mezzo materiale, si avrà dunque una certa relazione tra campi e induzioni.

Il caso più semplice e di base rispetto alle altre situazioni, è quello di propagazione nel vuoto:

$$\underline{\mathcal{D}}(\underline{r}, t) = \varepsilon_0 \underline{\mathcal{E}}(\underline{r}, t)$$

$$\underline{\mathcal{B}}(\underline{r}, t) = \mu_0 \underline{\mathcal{H}}(\underline{r}, t)$$

dove

$$\mu_0 = 4\pi \times 10^7 \text{ H/m}$$

questa costante è stata “decisa a tavolino”.

$$\varepsilon_0 = \frac{1}{\mu_0 c}$$

dove c è un osservabile, è misurabile, e vale circa 3×10^8 F/m. Si può dunque dedurre che:

$$\varepsilon_0 \sim 8,854 \times 10^{-12} \text{ F/m}$$

2.2.1 Materiale dielettrico isotropo

Un primo step, una prima complicazione, è quella di introdurre un mezzo senza struttura cristallina, dunque irregolare, come per esempio un vetro; il fatto di avere una completa irregolarità, una completa assenza di un reticolo, permette di non avere comportamenti particolari in direzioni diverse: l'irregolarità porta a non avere “direzioni privilegiate”. Questo tipo di comportamento va studiato (e si spiegherà tra breve perchè) nel dominio della frequenza, ottenendo:

$$\underline{D}(\underline{r}, \omega) = \varepsilon(\omega) \underline{E}(\underline{r}, \omega)$$

$$\underline{B}(\underline{r}, \omega) = \varepsilon(\omega) \underline{H}(\underline{r}, \omega)$$

Si osservi che l'induzione in un punto dipende esclusivamente dal campo elettrico nel medesimo punto; questa è un'osservazione assolutamente non

banale, dal momento che, nelle equazioni di Maxwell, si può vedere che \underline{B} dipende da $\nabla \times \underline{E}$: dalle derivate spaziali di \underline{E} . Parlare di derivate, significa automaticamente parlare di intorni, dunque non esclusivamente del punto desiderato; questo si può vedere ricordando che, per una generica funzione, una delle interpretazioni che si può dare all'operatore di derivazione è quello di limite del rapporto incrementale:

$$\frac{\partial f}{\partial x} = \lim_{h \rightarrow 0} \frac{f(x + \frac{h}{2}) - f(x - \frac{h}{2})}{h}$$

come si vede, il valore dipende dai valori del campo nell'interno.

Si noti che si ha dipendenza dalla frequenza: ciò complica, in effetti, l'espressione nel dominio del tempo.

Questo approccio sistemistico che si sta utilizzando per presentare i concetti farebbe venire in mente la teoria dei sistemi LTI: Lineari Tempo Invarianti. Ci chiediamo a questo punto, se sia possibile utilizzare i concetti relativi ai sistemi LTI per studiare il nostro sistema.

Prima di tutto, il nostro sistema, ossia il dielettrico, è lineare? Assolutamente no: in natura di lineare non esiste nulla! D'altra parte però è anche noto che, se si eccita un sistema lineare con un livello basso dell'ingresso, è possibile effettuare una linearizzazione, ossia considerare senza sbagliare molto il sistema come lineare. Questa cosa non è scontata: se si considera un flusso di dati molto elevato, per esempio 10 Gbit/s, perchè il segnale che si trasporta sia facilmente rilevabile (per garantire un SNR sufficientemente elevato per i rivelatori), si deve avere una certa energia; se la velocità di trasmissione è grande come in questo caso, si ha un'elevata energia al secondo, dunque un'elevata potenza; questa potenza, elevata, è in uno spazio molto ridotto: qualche micron quadro: si ha un'elevatissima densità di potenza, dunque campo elevato, dunque non è detto che ε sia effettivamente costante: potrebbe cambiare di valore, ottenendo un sistema che a tutti gli effetti non è lineare.

A parte questo aspetto, il sistema è TI, ossia Time Invariant? Beh, tendenzialmente si hanno invecchiamenti del materiale, però essi non sono importanti: ε generalmente non è una funzione del tempo (in senso di invecchiamento), per quanto sono possibili situazioni in cui, per esempio eccitando con un'onda elastica un vetro, si possano ottenere situazioni in cui questa affermazione non è vera.

Supponendo che il sistema sia dunque approssimabile a LTI, è noto che un metodo, nel dominio della frequenza, per rappresentare un sistema, sia quello di usare la sua funzione di trasferimento, definita come il rapporto tra la grandezza in uscita e quella in ingresso al sistema, nel dominio della frequenza: in questo caso, la funzione di trasferimento è $\varepsilon(\omega)$. Una seconda

grandezza che comunemente si utilizza è l'antitrasformata della funzione di trasferimento, detta "risposta all'impulso"; in questo caso, essa è indicata con $g_e(t)$, ed è:

$$g_e(t) = \mathcal{F}^{-1} \{ \varepsilon(\omega) \}$$

questa è la **funzione di Green**. Questo si ha dal momento che l'equazione differenziale che modella il sistema ha a secondo membro una delta di Dirac, quindi il risultato è proprio una funzione di Green. Si ha, nel dominio del tempo,

$$\underline{\mathcal{D}}(\underline{r}, t) = \int_{-\infty}^{+\infty} g_e(t - t') \underline{\mathcal{E}}(\underline{r}, t') dt'$$

questo porta a capire che lo spostamento a un istante dipende da tutta la storia precedente, proprio come la tensione $v(t)$ in una capacità. Nel caso di una resistenza, invece, si ha:

$$g_R(t) = R\delta(t)$$

per questo motivo, si ha che la funzione di Green della resistenza è una δ nel tempo. Il fatto di avere, in questo caso, una dipendenza da ω , è detto **dispersione**.

Una nota aggiuntiva sul modello che stiamo utilizzando: tutte le ipotesi sono fondate sul fatto che il materiale sia **omogeneo**, ossia composto da una sorta di *materiale continuo*: che non vi siano asperità. Le "asperità" derivano dalle discontinuità interne al materiale, dunque dalla struttura osservata a livello atomico; questo fatto è ragionevole dal momento che le λ con cui si ha a che fare sono dell'ordine, al minimo, dei 300 nm, dunque molto, molto maggiori della distanza tra due atomi, intorno ai 0,1 nm. Questa cosa non sarebbe ragionevole per esempio se usassimo i raggi X: essendo le lunghezze d'onda molto più basse, si avrebbe a che fare con fenomeni di scattering multiplo non trascurabili (cosa peraltro utilizzata per lo studio della struttura della materia), perdendo di fatto la possibilità di "mediare" le caratteristiche del materiale, dunque la possibilità stessa di definire una ε .

ε complesse: perdite

Si usa normalizzare la ε rispetto a quella del vuoto, in modo da isolare il comportamento "relativo" del mezzo:

$$\varepsilon = \varepsilon_0 \varepsilon_r(\omega)$$

$$\mu = \mu_0 \mu_r(\omega)$$

dove però, di solito, nei nostri casi, $\mu_r(\omega) \sim 1$. Per caratterizzare i materiali ottici, si introduce l'indice di rifrazione n come:

$$n \triangleq \frac{c}{v_f} = \frac{\sqrt{\varepsilon \mu}}{\sqrt{\varepsilon_0 \mu_0}} \sim \sqrt{\varepsilon_r}$$

A questo punto però ci chiediamo: ε_r è un numero reale o è un numero complesso? Beh, dato il nostro approccio sistemistico, dal momento che le funzioni di trasferimento sono in genere funzioni complesse, non c'è ragione di affermare che ε sia un numero reale. Il fatto che ε abbia componenti complesse è usualmente (non sempre) legato alla presenza di fattori di conducibilità nel dielettrico, ossia termini che introducono una dissipazione nell'onda. Per modellare ciò, si introduca un termine di corrente di conduzione, \underline{J}_c , nelle equazioni di Maxwell:

$$\nabla \times \underline{H} = j\omega\varepsilon \underline{E} + \underline{J}_e + \underline{J}_c$$

dove

$$\underline{J}_c = \gamma \underline{E}$$

questa, semplicemente, è la legge di Ohm in forma microscopica. Quello che usualmente si fa è considerare insieme i due termini in \underline{E} , ottenendo:

$$j\omega\varepsilon \underline{E} + \gamma \underline{E} = j\omega \left(\varepsilon - j\frac{\gamma}{\omega} \right) \triangleq j\omega \tilde{\varepsilon}$$

dove dunque $\tilde{\varepsilon}$ è un numero complesso; indicando ε il termine complessivo, si usa di solito la notazione:

$$\varepsilon = \varepsilon' - j\varepsilon''$$

Si ha dunque una parte reale, che è il termine puramente propagativo, e una parte immaginaria, che è quella che introduce la dissipazione di energia. Una nota su questa espressione: il fatto di aver usato il segno “-” deriva, facendo i conti, dall'aver usato, come termine di antitrasformazione dei fasori, $e^{+j\omega t}$; comunemente, per esempio nelle comunità di fisici, si trova lo stesso, con segno - a esponente; ciò porta ad annullare il segno -, ottenendo un +.

Si noti che:

- se $\varepsilon'' > 0$, si ha a che fare con un materiale passivo, ossia con un materiale che dissipa potenza;

- se $\varepsilon'' < 0$ si ha a che fare con un materiale attivo, ossia con un materiale che aumenta la potenza, che la amplifica.

Di solito, per i buoni dielettrici, si ha a che fare con una ε debolmente immaginaria, dunque si definisce il parametro “tangente di perdita” come:

$$\tan \delta \triangleq \frac{\varepsilon''}{\varepsilon'}$$

questa, di solito, è dell'ordine di 10^{-2} .

Un caso particolare sono i metalli: a microonde essi hanno una conducibilità elevatissima, portandoli a modellarli come conduttori perfetti, ma a $0,6 \mu\text{m}$, per esempio, possono avere

$$\varepsilon_{r,Ag} = -14,06 - j0,45$$

ε' è negativo dal momento che si ha una risposta opposta rispetto al segnale trasmesso. A queste frequenze ottiche, il metallo si comporta come un gas di elettroni, dunque studiandolo come un plasma, usando il modello di Drude, si può trovare un'espressione di questo genere.

Molto spesso, tra le grandezze indicate nei libri di fisica, si trova un “indice di rifrazione complesso”, come:

$$\tilde{n} = n - jk$$

questo, per l'ingegneria elettromagnetica, è molto poco indicato, dal momento che K , parametro legato alla dissipazione, è anche il modo in cui si indica normalmente la costante di propagazione, il numero d'onda.

Relazioni di Kramers-Kroenig

Un aspetto interessante e molto poco intuitivo, è il fatto che la parte reale e la parte immaginaria della costante dielettrica sono, di fatto, collegate tra loro. Si consideri di avere una ε tale per cui:

$$\varepsilon(\omega) = \varepsilon'(\omega) - j\varepsilon''(\omega)$$

Valgono, dunque, le seguenti relazioni:

$$\varepsilon'(\omega) - \varepsilon_0 = \frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{\varepsilon''(\alpha)}{\alpha - \omega} d\alpha$$

$$\varepsilon''(\omega) = \frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{\varepsilon'(\alpha) - \varepsilon_0}{\alpha - \omega} d\alpha$$

L'unica ipotesi sotto cui queste relazioni sono valide, è la **causalità** del sistema: il fatto che \underline{D} sia una “conseguenza”, \underline{E} una “causa”; ciò è assolutamente ragionevole, dal momento che è ovvio che ciò che viene indotto nel mezzo materiale sia una conseguenza del campo elettrico/magnetico in esso inviato. Queste relazioni sono trasformate integrali: di fatto, sono le trasformate di Hilbert della funzione.

P è la “parte principale secondo Cauchy”: si tratta di un metodo per il calcolo dell'integrale indefinito, anche in presenza di poli. Dato l'intervallo di integrazione α , ciò che si fa è dividere l'intervallo di integrazione in due parti:

si fa in modo da integrare da $-\infty$ fino a “appena prima il polo”, a $\alpha = \omega$, quindi si riparte da “appena dopo”, e si integra fino a $+\infty$. La parte principale secondo Cauchy si definisce quindi come:

$$P \int_{-\infty}^{+\infty} \frac{f(x)}{x} dx = \lim_{\delta \rightarrow 0} \left\{ \int_{-\infty}^{-\delta} \frac{f(x)}{x} dx + \int_{\delta}^{+\infty} \frac{f(x)}{x} dx \right\}$$

in questo modo, si evita di integrare al polo.

Queste formule, apparentemente teoriche, hanno anche un'importante valenza sperimentale: misurando in qualche modo o parte reale o parte immaginaria, se la grandezza non misurata è difficile da ottenere sperimentalmente, è possibile in questo modo ottenere l'altra mediante il calcolo dell'integrale. Questo può essere anche usato per legare, in ambito circuitale, R a X , oppure G a B .

2.2.2 Mezzi materiali anisotropi

L'ultimo caso interessante da analizzare, è quello dei dielettrici con struttura regolare, ossia in cui si hanno strutture ordinate. In questo caso, l'ipotesi di linearità è sempre valida, tuttavia non è detto che i vettori \underline{D} e \underline{E} siano paralleli.

Per vedere ciò, è semplicemente sufficiente generalizzare ε in $\underline{\underline{\varepsilon}}$:

$$\underline{D} = \underline{\underline{\varepsilon}} \underline{E}$$

dove $\underline{\underline{\varepsilon}}$ è una diadica, rappresentabile mediante una matrice 3×3 , essendo lo spazio nel quale lavora tridimensionale:

$$\begin{bmatrix} D_x \\ D_y \\ D_z \end{bmatrix} = \begin{bmatrix} \varepsilon_{xx} & \varepsilon_{xy} & \varepsilon_{xz} \\ \varepsilon_{yx} & \varepsilon_{yy} & \varepsilon_{yz} \\ \varepsilon_{zx} & \varepsilon_{zy} & \varepsilon_{zz} \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix}$$

La matrice associata a $\underline{\underline{\varepsilon}}$ è simmetrica e diagonalizzabile; il fatto che sia diagonalizzabile è interessante, dal momento che significa che esiste un riferimento ortogonale che, a partire da quello (x, y, z) di partenza, arrivando a (X, Y, Z) , permette di avere $\underline{\underline{\varepsilon}}$ diagonale. Questo particolare sistema, è quello degli **assi principali del cristallo**. Si hanno poi materiali per cui solo due dei numeri sono diversi tra loro, e altre cose particolari.

Nella pratica, utilizzare il riferimento proprio, ossia quello dove gli assi principali del cristallo coincidono con gli assi del sistema di riferimento, è impossibile: non è infatti detto che, montando il cristallo, il campo possa agire secondo le componenti principali; ciò ci dice comunque che esistono direzioni privilegiate, lungo le quali la velocità di fase v_f non sarà la stessa; allo stesso modo, anche la polarizzazione del campo potrebbe variare nelle varie direzioni principali.

2.2.3 Polarizzazione

Si vuole a questo punto riprendere il concetto di polarizzazione. Si ha:

$$\underline{\mathcal{E}}(t) = \Re \{ \underline{E} e^{j\omega_0 t} \}$$

quello che si fa usualmente nel calcolo è separare parte reale e parte immaginaria del fasore, ottenendo:

$$\underline{\mathcal{E}}(t) = \underline{E}' \cos \omega_0 t - \underline{E}'' \sin \omega_0 t$$

dove:

$$\underline{E} = \underline{E}' + j\underline{E}''$$

Il piano di polarizzazione è il piano su cui giacciono i vettori \underline{E}' e \underline{E}'' .

Al fine di comprendere la polarizzazione, si disegna un parallelogramma che ha come semimediane \underline{E}' e \underline{E}'' , quindi si disegna l'ellisse inscritto nel parallelogramma; questo ellisse sarà tangente nei punti medi del medesimo. A seconda degli istanti di tempo, si vede che prevale uno o l'altro vettore, quindi si può identificare la posizione dell'ellisse, e determinare il verso della rotazione della polarizzazione.

Esistono particolari tipi di polarizzazione:

- se $\underline{E}' \times \underline{E}'' = 0$, ossia se i due vettori sono paralleli (o se uno dei due è nullo), la polarizzazione è lineare: i due vettori “collassano” in uno solo, o sono comunque paralleli, e quindi il campo elettrico oscilla solo su una linea; si usa indicare la polarizzazione “verticale” come quella per cui il **campo elettrico** è per l'appunto verticale;

- se $|\underline{E}'| = |\underline{E}''|$, e $\underline{E}' \cdot \underline{E}'' = 0$ (ossia se i due vettori sono ortogonali e uguali in modulo), la polarizzazione è circolare;
- se non valgono le precedenti condizioni, la polarizzazione è ellittica.

2.3 Onde piane

2.3.1 Osservazioni preliminari

Si vuole a questo punto proporre un metodo di soluzione per le equazioni di Maxwell, basato su un certo insieme di ipotesi preliminari; l'obiettivo di questa sezione, nella fattispecie, è quello di introdurre i concetti fondamentali riguardanti le onde, e nel dettaglio un particolare tipo di onde: le onde piane.

Le equazioni di Maxwell, come ben noto ormai, sono le seguenti:

$$\begin{cases} \nabla \times \underline{E} = -j\omega\mu\underline{H} \\ \nabla \times \underline{H} = j\omega\varepsilon\underline{E} \end{cases}$$

Si ha a che fare con due ipotesi fondamentali:

- il primo, si può notare dalla forma delle equazioni: non si ha a che fare con sorgenti; quando si vuole determinare una soluzione di tipo “onda” (sia essa piana, sferica, cilindrica o di qualche altro tipo), non si devono considerare le sorgenti nel problema (la cosa verrà ancora ripresa e discussa);
- si assumono le seguenti ipotesi sul dielettrico: dielettrico omogeneo e illimitato.

Precedentemente, è stato introdotto un punto di vista “sistemistico”, fondato sulla Teoria dei Sistemi, al fine di introdurre alcuni concetti fondamentali; qua si potrebbe riprendere questa interpretazione, per quanto riguarda il primo punto: il campo, a tutti gli effetti, dovrebbe derivare da un certo insieme di sorgenti che, oscillando, lo producono: si dovrebbe partire da sorgenti, ingressi del “sistema”, ottenendo campi elettrici e magnetici (si ricordi infatti che, a meno che $\omega = 0$, i fenomeni elettrici e magnetici sono tra loro accoppiati, come si può anche notare dalle equazioni di Maxwell). Verrebbe da dire che, senza sorgenti, non potrebbero neanche esistere campi, ma in realtà non è così: potremmo avere anche delle risposte che non derivano dalle sorgenti; questa cosa può essere interpretata a partire da un concetto matematico.

Quando risolviamo un sistema lineare del tipo

$$\underline{\underline{A}} \underline{x} = \underline{n}$$

la soluzione di questo sistema, normalmente, esiste solamente quando $\det \{\underline{\underline{A}}\} \neq 0$; nel caso il determinante della matrice sia nullo, o esistono infinite soluzioni, o la soluzione non esiste proprio, ottenendo un sistema impossibile. Il ruolo di \underline{n} , in questo contesto, è quello di fare “da sorgente”: rappresenta un po’ il ruolo delle \underline{J}_e nelle equazioni di Maxwell. Nel caso dunque il determinante della matrice sia nullo e \underline{n} appartenga allo spazio delle colonne della matrice $\underline{\underline{A}}^1$, la soluzione esisterà, e non sarà unica; altrimenti, non esiste.

Esiste tuttavia un caso di sistema sicuramente non impossibile, in ogni situazione: il sistema omogeneo, ossia il sistema in forma

$$\underline{\underline{A}} \underline{x} = 0$$

nella nostra interpretazione, \underline{x} sono i campi, $\underline{\underline{A}}$ sono i vari operatori differenziali in gioco che vengono applicati sui campi. In questo caso, si noti che $\underline{x} = 0$ è sicuramente soluzione, quindi esiste; in questo caso, tuttavia, se il determinante di $\underline{\underline{A}}$ è nullo, allora quella “banale” non è l’unica soluzione: il fatto che il determinante sia nullo implica che lo spazio nullo, il *kernel*, ossia lo spazio dei vettori che, se vi si applica l’operatore rappresentato dalla matrice, vengono portati nell’origine, sono tutti soluzione, dal momento che, andando nell’origine, rispettano l’eguaglianza appena proposta. Dire che c’è uno spazio nullo significa dunque che ci sono dei vettori \underline{x} non nulli che rispettano l’equazione.

Un esempio visivo di vedere ciò è: dato un operatore, una matrice, che descrive l’operazione lineare di proiezione di un vettore su un certo piano, per esempio xy , tutti i vettori perpendicolari al piano, che sono infiniti, vengono proiettati nell’origine: lo spazio nullo ha dimensione 1, ossia è costituito dall’insieme dei vettori perpendicolari al piano.

Questa idea, applicata sulle matrici, dunque su “operatori” che operano su uno spazio di dimensione finita, può essere applicata anche su cose matematicamente più complicate, come un’equazione differenziale; quando si risolve un’equazione del tipo:

$$\frac{dy}{dt} = Ay$$

in questo caso, si può dire che gli operatori di derivazione si comportino *come una matrice con determinante nullo*: questo non è sicuramente molto

¹per il teorema di Rouchè-Capelli

formale da dirsi ma, dal momento che lo spazio delle soluzioni ha dimensione 1 (un esponenziale complesso), non è neanche troppo errata. La molteplicità delle soluzioni di un'equazione differenziale, dunque, nasce sostanzialmente da un concetto di questo tipo.

Discussi questi aspetti preliminari, è possibile procedere alla soluzione delle equazioni di Maxwell. Prima di tutto, consideriamo un'osservazione sul dielettrico: essendo esso omogeneo e illimitato, l'equazione è di fatto a coefficienti costanti, dal momento che μ e ε non subiscono variazioni nello spazio. ω è inoltre considerata fissa.

Al fine di risolvere l'equazione, l'idea è ipotizzare che la soluzione sia quella esponenziale (cosa ragionevole dal momento che gli operatori differenziali in gioco, i vari rotori, sono sostanzialmente combinazioni lineari delle derivate, nelle varie direzioni); per risolvere il sistema, tuttavia, è ancora necessario definire una cosa: il set di variabili indipendenti da utilizzare, ossia il tipo di sistema di riferimento. Una scelta è quella delle coordinate cartesiane, (x, y, z) ; un'altra, quella delle coordinate sferiche (r, ϑ, φ) ; un'altra ancora, quella delle coordinate cilindriche (ϱ, φ, z) : questi sono tutti differenti modi per rappresentare la posizione di un punto. La cosa interessante, tuttavia, è il fatto che, **a seconda del sistema di rappresentazione scelto, si ottiene una diversa famiglia di soluzioni**: se per esempio si usa un sistema di tipo cartesiano, la soluzione che si otterrà sarà una famiglia di **onde piane** (e vedremo in seguito cosa significa ciò); con un sistema sferico si otterranno onde sferiche, e così via. La notazione vettoriale ci svincola dal riferimento: a seconda delle coordinate utilizzate, si avranno comunque sempre soluzioni delle equazioni di Maxwell.

Si noti che si sta per utilizzare questo formalismo per risolvere un problema di fatto astratto: un problema in assenza di sorgenti; volendo risolvere un problema concreto, sembrerebbe che ciò che stiamo facendo non abbia senso; in realtà però, quello che stiamo facendo è perfettamente motivato, dal momento che noi in realtà stiamo cercando soluzioni in questa forma al fine di determinare delle "soluzioni base" per la soluzione del problema reale, quello **del mondo vero**, che tiene conto anche delle sorgenti; di fatto stiamo suddividendo un problema in vari step, e questo è solamente uno step intermedio; lo step successivo a questo potrebbe essere quello di prendere le soluzioni da noi trovate, la *base* di onde piane, per la rappresentazione delle sorgenti: data una sorgente, ci si può porre la domanda "quali onde piane eccita?". Questa soluzione, secondo onde piane, è comoda in un certo insieme di casi, ossia quello di sorgenti a simmetria planare; nel caso si avesse a che fare con simmetria rispetto a un certo asse, un asse di invarianza, dunque una simmetria cilindrica, le onde cilindriche sarebbero più indicate; nel caso invece del campo di un'antenna, campo generato da un punto, le

onde sferiche sono il caso più indicato: nel caso delle onde sferiche infatti il campo parte per l'appunto da un punto e si propaga secondo superfici a fase costante (approfondiremo il concetto) sferiche.

Nota conclusiva: tutto quello che stiamo per fare parte dalla soluzione di un sistema omogeneo con “rango inferiore al rango massimo”; questo presuppone di lavorare con infinite soluzioni, ma in realtà non è così: una volta che si è specificato tutto ciò che si deve specificare, ossia le varie condizioni iniziali o al contorno, di tutte le infinite soluzioni, se ne “seleziona” una proprio mediante queste condizioni.

2.3.2 Derivazione delle onde piane

Date le osservazioni precedenti, supponendo di utilizzare un sistema di riferimento piano, soluzioni di tipo esponenziale, si assume che:

$$\underline{E}(\underline{r}) = \underline{E}_0 e^{-jk_x x} e^{-jk_y y} e^{-jk_z z}$$

In questo, considerando che:

$$\underline{r} = x\hat{x} + y\hat{y} + z\hat{z}$$

$$\underline{k} = k_x\hat{x} + k_y\hat{y} + k_z\hat{z}$$

si ha

$$\underline{E}(\underline{r}) = \underline{E}_0 e^{-j\mathbf{k}\cdot\mathbf{r}}$$

$$\underline{H}(\underline{r}) = \underline{H}_0 e^{-j\mathbf{k}\cdot\mathbf{r}}$$

A questo punto, vogliamo procedere; sappiamo che la soluzione ha questa forma. Dal momento che si ha a che fare con operatori differenziali, come ∇ nella fattispecie, sarebbe interessante capire come esso operi sui campi. Si può vedere che:

$$\nabla e^{-j\mathbf{k}\cdot\mathbf{r}} = -j\mathbf{k} e^{-j\mathbf{k}\cdot\mathbf{r}}$$

infatti, ∇ rappresenta una combinazione lineare delle derivate, applicate sui campi; applicandolo nelle varie modalità (gradiente, rotore, divergenza), si può vedere che:

$$\nabla[] \longleftrightarrow -j\mathbf{k}$$

ovviamente, ciò vale **solo** quando si **lavora con degli esponenziali** (questa si può vedere come un'estensione, nel dominio della trasformata tripla di Fourier, della proprietà di trasformazione secondo Fourier della derivata di una funzione). Procediamo dunque nei conti, applicando questo tipo di proprietà:

$$-j\underline{k} \times \underline{E}_0 e^{-j\underline{k} \cdot \underline{r}} = -j\omega\mu \underline{H}_0 e^{-j\underline{k} \cdot \underline{r}}$$

$$-j\underline{k} \times \underline{H}_0 e^{-j\underline{k} \cdot \underline{r}} = j\omega\varepsilon \underline{E}_0 e^{-j\underline{k} \cdot \underline{r}}$$

a questo punto, si può vedere che, semplificando i vari esponenziali, che sono diversi da zero, sparisce la dipendenza da \underline{r} ; questo conferma la nostra ipotesi iniziale, riguardante il fatto che i coefficienti sono effettivamente costanti: se avessimo equazioni funzione di \underline{r} , non sapremmo in effetti che pesci pigliare, anche dal momento che non sarebbe più detto che le soluzioni siano esponenziali. Ordinando i termini, dividendo per i vari $-j$, si ottiene il seguente sistema di equazioni:

$$\begin{cases} \underline{k} \times \underline{E}_0 = \omega\mu \underline{H}_0 \\ \underline{k} \times \underline{H}_0 = -\omega\varepsilon \underline{E}_0 \end{cases}$$

Da questo set di equazioni si possono determinare le prime proprietà interessanti: nella prima, \underline{H}_0 deriva dal prodotto esterno di \underline{k} e \underline{E}_0 ; nella seconda, \underline{E}_0 dal prodotto esterno di \underline{E}_0 e \underline{k} ; questo ci dice che sia \underline{E}_0 sia \underline{H}_0 son esterni a \underline{k} , e che \underline{E}_0 e \underline{H}_0 sono **perpendicolari tra loro**. In altre parole, questo ci dice che \underline{k} , \underline{E}_0 e \underline{H}_0 sono una terna di vettori ortogonali.

A questo punto, si ricavi \underline{H}_0 dalla prima delle due equazioni, per sostituirlo nell'altra:

$$\underline{H}_0 = \frac{\underline{k} \times \underline{E}_0}{\omega\mu}$$

Si noti che questa relazione è molto importante, ed è detta **relazione di impedenza**, come si vedrà in seguito; si ha, proseguendo:

$$\frac{1}{\omega\mu} \underline{k} \times (\underline{k} \times \underline{E}_0) + \omega\varepsilon \underline{E}_0 = 0$$

A questo punto si moltiplica tutto per $\omega\mu$, in modo da eliminare la frazione, e contemporaneamente si semplifica il doppio prodotto, utilizzando la proprietà vettoriale nota:

$$-(\underline{k} \cdot \underline{k}) \underline{E}_0 + (\underline{k} \cdot \underline{E}_0) \underline{k} + \omega^2 \varepsilon \mu \underline{E}_0 = 0$$

ma, essendo \underline{E}_0 e \underline{k} ortogonali come noto dalle relazioni a sistema, si può dire che valga la seguente relazione:

$$(\underline{k} \cdot \underline{k}) \underline{E}_0 - \omega^2 \varepsilon \mu \underline{E}_0 = 0$$

Volendo soddisfare questa equazione, vi sono sostanzialmente due strade: $\underline{E}_0 = 0$, ossia la soluzione banale, la soluzione non interessante, o $\underline{E}_0 \neq 0$: in questo caso, si può raccogliere

$$(\underline{k} \cdot \underline{k} - \omega^2 \varepsilon \mu) \underline{E}_0 = 0$$

quindi

$$\underline{k} \cdot \underline{k} - \omega^2 \varepsilon \mu = 0$$

questa equazione è detta **relazione di dispersione**: essa infatti soddisfa la tipica espressione delle relazioni di dispersione, $f(\omega, k) = 0$. Per ora non abbiamo informazioni sul modulo di \underline{E}_0 , sulla sua ampiezza: questo sarà un obiettivo futuro. \underline{H}_0 è legato a \underline{E}_0 , come noto, dalla relazione di impedenza. Il significato della relazione di dispersione è il seguente:

$$\underline{k} \cdot \underline{k} = |\underline{k}|^2 \triangleq k^2$$

intendendo dunque con k^2 il modulo quadro². Volendo scrivere \underline{k} come:

$$\underline{k} = k \hat{s}$$

dove \hat{s} ha caratteristiche non ancora definite. La relazione di dispersione ha il compito di fissare il **modulo** del vettore \underline{k} , ma non la sua direzione. Per ora, le possibili soluzioni che stiamo cercando sono infinite: abbiamo infiniti possibili \hat{s} che soddisfano le condizioni finora trovate, per le onde piane.

È giunto il momento di parlare del nome “onde piane”: come mai queste onde sono dette “piane”? Beh, se si sostituisce l’espressione di \underline{k} dentro all’espressione di partenza, si ha che un’onda piana diretta lungo \hat{s} ha espressione (la particolare, l’unica onda piana che si propaga lungo quella direzione, \hat{s}):

$$\underline{E}|_{\hat{s}} = \underline{E}_0 e^{-jk\hat{s} \cdot \underline{r}}$$

Ciò che permette di caratterizzare un’onda, sono le superfici a fase costante: i luoghi dei punti nello spazio per cui la fase, ossia l’argomento dell’esponenziale, è costante. Parlare di superfici a fase costante è importante dal momento che le onde sono un qualcosa di difficile da descrivere: un’onda

²in realtà, più che di modulo, sarebbe corretto parlare di *pseudomodulo*, che coincide con il modulo se tutto è reale

è una sorta di nuvola, di nebbia, e vedere da “dentro” capire che forma ha una nuvola è molto difficile; la superficie a fase costante permette di riconoscere l’onda: i “fronti d’onda” (sinonimo di superfici a fase costante). Data superficie di fase ϕ ,

$$\phi = k\hat{s} \cdot \underline{r}$$

il luogo dei punti per cui ciò è costante è il piano: il piano perpendicolare a \hat{s} (dal momento che \underline{r} è alla prima potenza); essendo la superficie a fase costante un piano, si ha un’onda piana; se essa fosse un cilindro o una sfera, sarebbe rispettivamente un’onda cilindrica o sferica.

Come si può dunque rappresentare un’onda piana? In letteratura, la cosa più comune è disegnare il versore \hat{s} , quindi una famiglia di fronti d’onda, ossia di piani a esso normali.

Si tenga ben presente, di ciò, che essi sono i fasori:

$$\underline{\mathcal{E}}(\underline{r}, t) = \Re \{ \underline{E}_0 e^{j\omega_0 t} e^{-jk\hat{s} \cdot \underline{r}} \}$$

ipotizzando che \underline{E}_0 sia nella consueta forma

$$\underline{E}_0 = \underline{E}'_0 + j\underline{E}''_0$$

si ha

$$\underline{\mathcal{E}}(\underline{r}, t) = \underline{E}'_0 e^{-jk\hat{s} \cdot \underline{r}} \cos(\omega_0 t) - \underline{E}''_0 e^{-jk\hat{s} \cdot \underline{r}} \sin(\omega_0 t)$$

I punti sulla stessa superficie a fase costante sono caratterizzati dal fatto che $k\hat{s} \cdot \underline{r}$ hanno un valore specifico: la grandezza deve avere la stessa quantità; se questa quantità è tale per cui, per un certo istante di tempo t il coseno è massimo, si considerano piani per cui il campo è grande. Al mare, le creste sono l’insieme dei punti più alti della superficie dell’onda marina; le valli, quelle per cui i punti sono bassi. Dal momento che \underline{E} è perpendicolare a \underline{k} , l’ellisse di polarizzazione appartiene proprio ai piani disegnati, i piani delle superfici a fase costante.

Si può, infine, scrivere la relazione di impedenza nella sua forma più nota:

$$\underline{H}_0 = \sqrt{\frac{\varepsilon}{\mu}} \hat{s} \times \underline{E}_0$$

$$\underline{E}_0 = \sqrt{\frac{\mu}{\varepsilon}} \underline{H}_0 \times \hat{s}$$

dove \hat{s} , ossia il versore di propagazione dell’onda, è “nome e cognome” dell’onda piana: è ciò che la specifica univocamente.

Note sulla propagazione delle onde (in particolare piane)

Si vuole a questo punto introdurre una serie di concetti riguardanti la propagazione delle onde. Come noto, nelle linee di trasmissione (o nelle guide d'onda, analizzate con il formalismo delle linee modali), la propagazione avviene lungo \hat{z} ; si ha, considerando di avere su questa linea un modo solo progressivo,

$$V(z) = V_0^+ e^{-jkz}$$

volendo antitrasformare nel tempo, si ottiene una forma del tipo:

$$v^+(z, t) = \cos(\omega_0 t - kz)$$

in questo caso, è evidente la dipendenza della fase dal tempo e dallo spazio (in questo caso, il problema è unidimensionale). Il fatto che l'onda sia progressiva è evidente per il fatto che, all'aumentare di z , l'onda è progressiva. Lo studio del fatto che l'onda sia progressiva dipende sostanzialmente dalla velocità di fase: la velocità di fase v_f è definita come la velocità che l'**osservatore** deve avere, in modo tale che la fase in quel campo sia costante³. In questo caso la fase dipende sia dal tempo, sia dallo spazio: è dunque una legge che lega tempo e spazio. Data ϕ la fase, chiedere che ϕ sia costante significa chiedere che il suo differenziale sia nullo, in modo che ci si trovi su di un punto stazionario per la fase: $d\phi = 0$. Calcolando il differenziale:

$$d\phi = \omega_0 dt - k_0 dz$$

dunque:

$$\frac{dz}{dt} = \frac{\omega_0}{k} \triangleq v_f$$

nel caso nostro, dato k ,

$$v_f = \frac{\omega_0}{k} = \frac{1}{\sqrt{\epsilon\mu}}$$

Questo termine è maggiore di zero, dunque la velocità di fase è positiva, rispetto all'aumentare delle z , quindi l'onda è progressiva: z cresce al passare del tempo, quindi l'onda è progressiva.

Confrontando ciò con il caso tridimensionale (dell'onda piana), si vede che l'onda avanza lungo \hat{s} .

³definizione alternativa, come si può vedere da Collins - Foundations of ... , è "la velocità di propagazione della portante, supposta come attiva da $t = -\infty$

Parliamo ancora della velocità di fase: se nella struttura guidante unidimensionale prima introdotta la velocità di fase è definita in maniera univoca, nel caso tridimensionale ciò non è detto:

Si immagini di avere un'onda piana che si propaga lungo un certo \hat{s} , e si immagini che l'osservatore “corra” dietro l'onda piana, con una certa velocità, in maniera da osservare la fase costante: questo significa, considerando l'intersezione della direzione di moto dell'osservatore con i piani a fase costante, che, se dopo un certo tempo Δt si è passati da un piano a fase costante a un altro, l'osservatore, per muoversi con l'onda, per osservare la fase costante, dovrà essersi mosso da un'intersezione a un'altra, da un “punto” a un altro”. Come si può vedere dalla costruzione geometrica, tuttavia, l'osservatore dovrà percorrere una strada maggiore rispetto a quella dell'onda: se la velocità con cui questi piani si spostano nella direzione \hat{s} è c , l'osservatore, per stare allineato, dovrà andare ancora più veloce, dal momento che si ha qualcosa di equivalente a un trapezio, con il suo “lato lungo”: $\frac{1}{\sqrt{\epsilon\mu}}$ è il valore **minimo** della velocità dell'osservatore, al fine di mantenersi “agganciati” alle intersezioni dei piani! Il caso limite è poi quello per cui l'osservatore si muove in una direzione parallela ai piani, ossia normale a \hat{s} : in questo caso, infatti, la velocità di fase, al fine di rimanere “agganciati”, dovrebbe essere infinita!

Un fenomeno del genere si può osservare, parlando di fenomeni noti, in una guida d'onda: come noto infatti, in una guida d'onda, si può pensare al campo elettromagnetico che si propaga come a una serie di “onde piane” che rimbalzano da una parete a un'altra.

Questo è il motivo per cui la velocità di fase di un modo è più grande della velocità della luce. La velocità di fase, riferita sempre rispetto all'asse z (la “direzione per cui può correre il nostro osservatore”), non sarà parallela quindi ai \hat{s} delle onde piane all'interno della struttura guidante, quindi sicuramente maggiore. Il caso “estremo” è poi la condizione di cut-off: al taglio, le onde si propagano in modo sempre più verticale, dunque i piani saranno “sempre più paralleli” a \hat{z} , avendo un osservatore che ha bisogno di una velocità infinita per vedere costante la fase.

Il discorso è invece diverso per quanto concerne la velocità di gruppo, v_g : la velocità di gruppo è perfettamente definita (c'è un solo valore). Un modo per descrivere, identificare tutte le varie onde piane, mediante una costruzione geometrica, è basata sullo spazio \underline{k} :

In questo caso, nel sistema (x, y, z) , non si riportano le dimensioni spaziali, ma le dimensioni reciproche: k_x, k_y, k_z , che dimensionalmente sono:

$$[k_i] = \text{m}^{-1}$$

Un generico punto di questo spazio è uno dei vettori \underline{k} che appaiono nell'espressione dell'onda piana. Se, per una certa frequenza fissata, i vettori \underline{k} possono essere qualsiasi, a patto che rispettino l'equazione di dispersione, si ha che si definisce una sfera, in questo spazio: l'immagine geometrica del fatto che il modulo è fisso, ma la direzione non ancora definita. Un punto di questa sfera è di fatto un'onda piana: è il vettore \underline{k} a esso corrispondente, è la direzione dell'onda piana a esso associata. In realtà, di onde piane se ne hanno 2, per ciascun punto: noi sappiamo che \underline{E}_0 deve essere perpendicolare a \underline{k} , ma non si hanno informazioni di come stia, nel piano tangente: nel piano tangente, infatti, vi sono due direzioni linearmente indipendenti, che si possono scegliere "arbitrariamente", al punto in cui siamo arrivati: sono due direzioni di polarizzazione (se la polarizzazione è lineare; nel caso non lo sia, si può proiettare sugli elementi di una base di polarizzazioni lineari).

A ogni punto della sfera sono dunque collegati due vettori, ma (per motivi per ora non esplicitati), è bene che a una superficie sia associato un solo vettore: si deve "contare due volte" la superficie, considerando come la presenza di due superfici sferiche coincidenti. Questo è utile dal momento che, se il dielettrico è un cristallo anisotropo, facendo gli stessi conti visti, invece di avere una "sfera doppia" si hanno due superfici diverse, distinte, per ciascuna delle quali si ha una sola polarizzazione; il caso dei mezzi isotropi è "degenere", dal momento che le superfici diventano sferiche, e se ne hanno "due coincidenti", che differiscono solo per le possibili polarizzazioni⁴.

La velocità di gruppo, nel caso 1-dimensionale, è definita come:

$$v_g \triangleq \frac{d\omega}{dk}$$

in questo caso, però, non abbiamo un solo k , ma \underline{k} : k_x, k_y, k_z : la velocità di gruppo sarà un vettore, con tre componenti:

$$v_{g,x} = \frac{\partial\omega}{\partial k_x} \quad v_{g,y} = \frac{\partial\omega}{\partial k_y} \quad v_{g,z} = \frac{\partial\omega}{\partial k_z}$$

dunque, v_g sarà in realtà un vettore: \underline{v}_g .

Un'osservazione: tornando al nostro spazio \underline{k} , la superficie sferica con cui si ha a che fare, è anche pensabile come la superficie a ω costante: considerando ω come una funzione di k , dal momento che, invertendo l'equazione di dispersione, si ha;

⁴Questa cosa si può ricondurre, mediante un'analogia, al caso di soluzione di equazioni algebriche con $\Delta = 0$: si hanno "due soluzioni coincidenti", un po' come in questo caso; tutto ciò deriva sostanzialmente dall'equazione di dispersione per le onde piane: è essa a fissare il modulo: $\underline{k} \cdot \underline{k} = \omega\sqrt{\varepsilon\mu}$, e solo se il mezzo è isotropo, si hanno due sfere uguali

$$\omega = k \frac{1}{\sqrt{\varepsilon\mu}}$$

allora la sfera è effettivamente il luogo dei punti per cui, se k è costante, allora anche ω è costante; la velocità di gruppo, tuttavia, è:

$$\underline{v}_g = \frac{\partial\omega}{\partial k_x} \hat{x} + \frac{\partial\omega}{\partial k_y} \hat{y} + \frac{\partial\omega}{\partial k_z} \hat{z}$$

ossia, è sostanzialmente il gradiente della superficie a ω costante! Come è noto, il gradiente è sempre perpendicolare alla superficie di livello, dal momento che il gradiente indica la direzione di massima pendenza, di massima variazione: in altre parole, è diretta verso \hat{s} (il quale è normale alla superficie di livello). In altre parole, si può dimostrare, calcolando le derivate, che:

$$\underline{v}_g = \frac{1}{\varepsilon\mu} \hat{s}$$

Il discorso della velocità di gruppo è anche legato alla potenza trasportata dall'onda: come noto, la potenza si calcola mediante il flusso del vettore di Poynting, ossia con il flusso del vettore \underline{S} così rappresentabile:

$$\underline{S} = \underline{E} \times \underline{H}^*$$

Dal momento che tuttavia il dielettrico è privo di perdite, gli esponenziali si cancellano, si ha che i campi sostanzialmente sono reali:

$$\underline{S} = \underline{E}_0 \times \underline{H}_0$$

Come noto dalla relazione di impedenza, però:

$$\underline{H}_0 = \sqrt{\frac{\mu}{\varepsilon}} \underline{k} \times \underline{E}_0$$

si ottiene dunque:

$$|\underline{S}| = \frac{|\underline{E}_0|^2}{\sqrt{\frac{\mu}{\varepsilon}}} \hat{s}$$

dove

$$[\underline{S}] = \frac{\text{W}}{\text{m}^2}$$

Anche ciò permette di dire che l'onda si muove lungo \hat{s} : la potenza, l'energia, viene trasportata lungo \hat{s} . Il campo elettromagnetico immagazzina

energia, e, spostandosi il campo, anche l'energia si sposta: ciò porta ad avere un flusso di energia che dunque si sposta, con una certa velocità. Si può dimostrare, facendo il conto tra il flusso di potenza e l'energia immagazzinata per unità di volume, che la velocità dell'energia coincide con la velocità di gruppo. Ciò rafforza ulteriormente l'interpretazione del fatto che l'onda piana si sposta lungo \hat{s} .

Per ogni punto, c'è anche il punto diametralmente opposto: quello che descrive l'onda che si muove lungo $-\hat{s}$.

2.4 Mezzi non omogenei: discontinuità piana

Si vuole a questo punto fare un passo avanti: le onde piane sono soluzione delle equazioni di Maxwell in caso di mezzi omogenei; una prima cosa, semplice, che si può fare, è introdurre una discontinuità piana, che introduce una non omogeneità nel mezzo: si avrebbero, in questo caso, due materiali diversi, con una discontinuità relativamente "semplice" da affrontare (come vedremo). Quando si ha a che fare con un problema di questo genere, si finisce per ottenere una situazione di questo tipo:

Lavorando sulle equazioni di Maxwell (senza dunque aggiungere nulla dall'esterno), è possibile trovare delle "condizioni di interfaccia", ossia delle condizioni che garantiscano la regolarità del campo elettromagnetico; queste condizioni sono di fatto condizioni al contorno, che introducono richieste sui campi elettrici e magnetici **tangenziali all'interfaccia**, ossia trasversali alla direzione di propagazione dell'onda: $\underline{E}_{\text{trasverso}}$, $\underline{H}_{\text{trasverso}}$. Essendoci due condizioni vettoriali, queste daranno luogo a 4 condizioni: 2 sul campo elettrico, 2 sul campo magnetico (essendo sui trasversi, vi sono solo due direzioni indipendenti trasversali a quella di propagazione).

Come noto dalla Fisica delle onde, dato un campo incidente sotto forma di onda piana, si generano un'onda riflessa e una trasmessa, per cui valgono le leggi di Snell, leggi in grado di valutare gli angoli di riflessione/rifrazione (trasmissione); ciò che le leggi di Snell non quantificano, tuttavia, sono le ampiezze dei campo elettrico e magnetico riflessi e trasmessi; questo sarà l'obiettivo finale del metodo che verrà ora introdotto.

2.4.1 Metodo delle linee di trasmissione

Si propone la seguente idea: la struttura prima presentata può essere pensata come una guida d'onda, con sezione infinita (per esempio quadrata); per questa, è possibile effettuare uno studio dei modi della struttura, applicando

il formalismo modale, ossia il metodo comunemente utilizzato sulle guide d'onda.

Si introduce a questo punto un'ipotesi semplificativa: da qui in poi si supporrà che le onde piane considerate stiano sempre nel piano xz , ossia si supporrà che $k_y = 0$:

Se dunque la componente del k incidente lungo y è nulla, si può vedere che il campo elettrico non dipende da y ; questo deriva banalmente dal fatto che:

$$e^{-jk_y y} = e^{-j0} = 1$$

questo, ovviamente, comporta anche il fatto che:

$$\frac{\partial E}{\partial y} = 0$$

essendoci poi la relazione di impedenza, questo porta anche a dire che

$$\frac{\partial H}{\partial y} = 0$$

Questo vale per campo incidente, riflesso e trasmesso: se il campo incidente non dipende da y , infatti, solo una dipendenza dall'interfaccia da y porterebbe ad avere una variazione, una derivata non nulla rispetto a y ; essendo tuttavia l'interfaccia planare, nella fattispecie un piano normale a \hat{z} (e quindi il piano xy), essa non subisce variazioni al variare di y . Ciò che si è dunque voluto motivare a parole, è semplicemente il fatto che, per tutti i casi che considereremo,

$$\frac{\partial}{\partial y} = 0$$

Questa cosa ha validità abbastanza generale: se si ha un'onda piana che incide "di sbieco", è semplicemente possibile ruotare il sistema di riferimento; questa cosa ovviamente però non funziona, nel caso in cui di onde piane "di sbieco" se ne hanno due, con direzioni di propagazione non parallele: in questo caso si potrebbe "aggiustare" un'onda piana, ma certamente non l'altra.

Per procedere, si scrivano a questo punto le equazioni di Maxwell, per componenti (ossia, un po' come "le aveva scritte Maxwell" nel 1873):

- la prima equazione si può scrivere mediante il seguente sistema:

$$\left\{ \begin{array}{l} \frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} = -j\omega\mu H_x \\ \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} = -j\omega\mu H_y \\ \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = -j\omega\mu H_z \end{array} \right.$$

- la seconda equazione si può scrivere mediante il seguente sistema:

$$\left\{ \begin{array}{l} \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} = j\omega\varepsilon E_x \\ \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} = j\omega\varepsilon E_y \\ \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} = j\omega\varepsilon E_z \end{array} \right.$$

Questi sono i set di equazioni complete, senza aver introdotto la nostra semplificazione; una volta introdotta, si ottiene ciò:

$$\left\{ \begin{array}{l} \frac{\partial E_y}{\partial z} = j\omega\mu H_x \quad \Leftarrow \\ \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} = -j\omega\mu H_y \\ \frac{\partial E_y}{\partial x} = -j\omega\mu H_z \quad \Leftarrow \end{array} \right.$$

$$\left\{ \begin{array}{l} -\frac{\partial H_y}{\partial z} = j\omega\varepsilon E_x \\ \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} = j\omega\varepsilon E_y \quad \Leftarrow \\ \frac{\partial H_y}{\partial x} = j\omega\varepsilon E_z \end{array} \right.$$

In questi, sono stati identificate, mediante delle frecce, equazioni “dello stesso gruppo”: le equazioni scritte infatti non sono tutte interdipendenti tra loro, ma solo “a gruppi”. Sono interdipendenti tra loro quelle indicate “con la freccia”, e quelle “senza la freccia”. Si hanno, in sostanza due sottogruppi, ai quali, come stiamo per osservare, appartengono campi elettromagnetici dotati di particolari proprietà.

- il sottogruppo “con freccia”, contiene le seguenti componenti di campo:

$$\begin{bmatrix} 0 \\ E_y \\ 0 \end{bmatrix}, \quad \begin{bmatrix} H_x \\ 0 \\ H_z \end{bmatrix}$$

questa configurazione è detta TE: Trasverso Elettrico. Questo nome deriva dal fatto che $E_z = 0$, ossia dal fatto che non si hanno componenti del campo elettrico lungo l’asse z . Non siamo interessati in realtà all’asse x , dal momento che il fatto di non avere componenti lungo z è una proprietà generale, mentre quella di non averne lungo x è una proprietà che discende dall’aver imposto $k_y = 0$, dunque non è particolarmente interessante ai fini della trattazione.

- l'altro sottogruppo ha le seguenti configurazioni di campo:

$$\begin{bmatrix} E_x \\ 0 \\ E_y \end{bmatrix}, \quad \begin{bmatrix} 0 \\ H_y \\ 0 \end{bmatrix}$$

sarà scontato il fatto che questa configurazione è detta TM, Trasverso Magnetico, per le stesse motivazioni prima proposte per quanto riguarda il TE, applicate sul campo magnetico.

Quelli appena introdotti sono due insiemi di campi. Come si può vedere, le equazioni di Maxwell presentano derivate lungo x e lungo z ; l'interfaccia, tuttavia, è coincidente con l'asse xz , dunque il sistema è invariante per traslazione rispetto a x , non lo è rispetto a z (ovvio: se cambio la posizione di x , non cambia nulla, dal momento che se io "sposto" l'interfaccia lungo x , essendo essa infinita, il sistema non subisce alcune mutazione; se la "sposto" lungo z invece sposto il punto di interazione dell'onda, modificando drasticamente il sistema). Data questa osservazione di invarianza, è dunque utile introdurre, per quanto riguarda i campi espressi lungo x , una rappresentazione di tipo spettrale, ossia basata sull'antitrasformazione secondo Fourier:

$$E_y(x, z) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \tilde{E}_y(\xi, z) e^{-j\xi x} d\xi$$

dove ξ è la variabile spettrale rispetto a x .

Proviamo a motivare queste ultime frasi: il metodo operativo per la soluzione dei circuiti RLC, è basato sulla seguente assunzione: considerando per esempio di dover trasformare (secondo Fourier o Laplace che sia) la derivata associata a una capacità, C , si ha che ciò ha senso dal momento che C è costante per tutti i valori di frequenza, ossia è costante nel tempo; in questo caso la trasformata di Fourier in questione è spaziale, ma, dal momento che la grandezza interessante è costante rispetto a x , ha senso effettuare la trasformazione; questa, oltretutto, comporta che:

$$\frac{\partial}{\partial x} e^{-j\xi x} \longleftrightarrow -j\xi$$

e ciò ovviamente semplifica molto le espressioni.

Applicazione alle onde piane TE

Come noto, le onde piane TE sono descritte mediante il seguente sistema di equazioni:

$$\begin{cases} \frac{\partial E_y}{\partial z} = j\omega\mu H_x \\ \frac{\partial E_y}{\partial z} = -j\omega\mu H_z \\ \frac{\partial \tilde{H}_x}{\partial z} - \frac{\partial H_z}{\partial x} = j\omega\varepsilon E_y \end{cases}$$

Trasformiamo secondo Fourier spaziale la variabile x ; tutte le componenti di campo saranno conseguentemente trasformate, ottenendo l'eliminazione delle derivate secondo x (non le altre, ovviamente, dal momento che la trasformazione non coinvolge z per i motivi già detti):

$$\begin{cases} -\frac{\partial \tilde{E}_y}{\partial z} = -j\omega\mu \tilde{H}_x \\ -j\xi \tilde{E}_y = -j\omega\mu \tilde{H}_z \\ \frac{\partial \tilde{H}_x}{\partial z} + j\xi \tilde{H}_z = j\omega\varepsilon \tilde{E}_y \end{cases}$$

Dalla seconda delle tre equazioni si può ricavare l'espressione della componente longitudinale del campo magnetico:

$$\tilde{H}_z = \frac{\xi}{\omega\mu} \tilde{E}_y$$

sostituendo dunque questa nella terza equazione, si ottiene:

$$\frac{\partial \tilde{H}_x}{\partial z} + j\xi \frac{\xi}{\omega\mu} \tilde{E}_y = j\omega\varepsilon \tilde{E}_y$$

A questo punto, posso prendere questa equazione, raccogliere al termine destro un $\frac{1}{\omega\mu}$, e notare che si ottiene in risultante k^2 , essendo

$$k^2 = \omega^2\varepsilon\mu$$

Considerando il fatto che le derivate parziali possono anche essere scritte (questione di gusto) come derivate totali, dal momento che ξ è sostanzialmente un parametro, una costante; si ha dunque che questa equazione è scrivibile come:

$$\frac{d\tilde{H}_x}{dz} = j\frac{1}{\omega\mu} (k^2 - \xi^2) \tilde{E}_y$$

questa è una equazione; l'altra, semplicemente, è:

$$-\frac{d\tilde{E}_y}{dz} = -j\omega\mu \tilde{H}_x$$

mettiamole a sistema:

$$\begin{cases} -\frac{d\tilde{E}_y}{dz} = -j\omega\mu\tilde{H}_x \\ \frac{d\tilde{H}_x}{dz} = j\frac{1}{\omega\mu}(k^2 - \xi^2)\tilde{E}_y \end{cases}$$

Questo è un sistema di equazioni differenziali, che può ricordare qualcos'altro. Nella fattispecie, si effettuino le seguenti “sostituzioni”, al fine di identificare di cosa si intende parlare:

$$V(\xi, z) = \tilde{E}_y(\xi, z)$$

$$I(\xi, z) = -\tilde{H}_x(\xi, z)$$

$$k_z^2 = k^2 - \xi^2$$

$$Z_\infty^{\text{TE}} = \frac{\omega\mu}{k_z}$$

sostituendo, si trovano le seguenti equazioni:

$$\begin{cases} -\frac{dV}{dz} = jk_z Z_\infty^{\text{TE}} I \\ -\frac{dI}{dz} = jk_z Y_\infty^{\text{TE}} V \end{cases}$$

queste sono semplicemente le equazioni delle linee di trasmissione, ossia le equazioni delle linee modali equivalenti del nostro problema. Questo è un sistema di equazioni differenziali, dunque andrebbe risolto, ma dalla teoria di Campi Elettromagnetici noi sappiamo che:

$$V(z) = V_0^+ e^{-jk_z z} + V_0^- e^{+jk_z z}$$

$$I(z) = V_0^+ Y_\infty^{\text{TE}} e^{-jk_z z} - V_0^- Y_\infty^{\text{TE}} e^{+jk_z z}$$

Tutto ciò è noto dalla teoria della propagazione guidata. Come si può vedere dalle “sostituzioni”, ξ sarebbe semplicemente qualcosa di analogo alla costante di propagazione critica, k_c : la costante di propagazione longitudinale infatti è

$$k_z = \sqrt{k^2 - k_c^2}$$

ξ può essere un numero qualsiasi: esso è stato introdotto come variabile di integrazione in un integrale da $-\infty$ a $+\infty$; questo significa che qualsiasi valore è accettato. In una guida vera e propria, invece, possiamo scegliere qualsiasi valore di m e n intero per la definizione del modo, e di conseguenza si ha

un certo k_c ; ogni coppia (m, n) , individua univocamente un modo. Possiamo dire, in questo caso, che ogni valore di ξ sia la “etichetta” di un particolare modo, un TE_ξ (si ha un pedice solo, dal momento che il problema è con $k_y = 0$); nel caso della guida, i campi variano sia con x , sia con y , quindi gli indici modali sono due. V è quindi la tensione modale, I la corrente modale; si noti inoltre che la tensione modale, per come è stata introdotta, è la trasformata di Fourier di E_y , dunque della componente trasversale del campo elettrico, mentre la corrente modale è la trasformata di Fourier della componente H_x : corrente e tensione modali sono imparentati con i campi trasversali rispetto alla direzione di propagazione. Cerchiamo di capire tutto ciò con un esempio: immaginiamo di avere un campo $E_y(x)$ a “porta”:

un campo non nullo solo su una finestra. In questo caso, date le formule precedentemente proposte, si ha:

$$V(\xi, 0) = \int_{-\infty}^{+\infty} E_y(x) e^{+j\xi x} dx$$

dato un campo “a porta”, si ha:

$$V(\xi, 0) = LE_0 \frac{\sin\left(\xi \frac{L}{2}\right)}{\xi \frac{L}{2}}$$

questa espressione è una sinc; si può vedere che la posizione dello zero è:

$$\frac{\xi L}{2} = \pi \implies \xi = \frac{2\pi}{L}$$

Questo è l’andamento della tensione modale relativa a ciascun modo: essendo essa in funzione di ξ , ossia dell’indice modale, questo è semplicemente l’andamento della tensione modale, data ipotesi di campo E_y a porta, per ciascun modo: il contributo di ciascun modo nel campo. Si hanno infiniti modi; questo non è una novità, dal momento che anche nella guida era così, tuttavia qua l’infinità ha la potenza del continuo: dal momento che la distanza tra i vari modi tende ad annullarsi, si ha questo tipo di andamento: un ξ che varia con la potenza del continuo: i coefficienti con cui posso sommare le tante sinusoidi, i tanti modi.

Applicazione alle onde piane TM - Cenni

Per il caso TM, sostanzialmente si può dire che:

$$\begin{cases} V(\xi, z) = \tilde{E}_x(\xi, z) \\ I(\xi, z) = \tilde{H}_y(\xi, z) \end{cases}$$

poi

$$Z_{\infty}^{\text{TM}} = \frac{k_z}{\omega \varepsilon}$$

$$k_z = \sqrt{k^2 - \xi^2}$$

Per ogni ξ , quindi, si ha a che fare con un modo TE e un modo TM: con due polarizzazioni diverse possibili per l'onda; in questo caso non si hanno esclusioni, come per le guide (dove per esempio in certi casi i modi non possono avere un indice nullo o cose simili).

2.4.2 Coefficienti di Fresnel - onde TE

Precedentemente, è stata proposta la rappresentazione della tensione modale come secondo una somma di sinusoidi, di esponenziali complessi; quello che si può fare, tuttavia, è considerare una sola di queste “sinusoidi”, e chiederci come essa si propaghi, ossia come essa “cambi con z ”; si consideri dunque un ben preciso valore di ξ , si calcoli il k_z , quindi si può trovare, supponendo di avere una componente di campo puramente progressiva:

$$E_y^+(x, z) = V_0^+ e^{-jk_z z}$$

ma sappiamo che:

$$E_y(x, z) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} V(\xi, z) e^{-j\xi x} d\xi$$

di questo integrale, prendiamo un singolo termine, una singola ξ : siamo interessati solo a una certa ξ . Ciò presupporrebbe, per essere formali, di avere solo una $\delta(\xi)$. Proviamo a vedere la cosa in pratica:

Si supponga di avere a che fare con un'onda TE; in questo caso, dunque, si ha una situazione di questo genere:

$$\underline{E}_{\text{inc}}(x, z) = E_y \hat{y} = E_0 e^{-jk_0 n_1 \hat{s} \cdot x} \hat{y}$$

infatti, nel caso di onda TE, rispetto al sistema di riferimento stabilito, il campo elettrico è diretto sicuramente lungo \hat{y} (come d'altro canto dimostrato precedentemente). Il versore \hat{s} sarà inclinato di un certo angolo sul piano xz ; dato ϑ_1 l'angolo di incidenza sul piano rispetto alla sua normale, si ha:

$$\hat{s} = \sin \vartheta_1 \hat{x} + \cos \vartheta_1 \hat{z}$$

calcolando dunque il prodotto scalare, si finisce per ottenere:

$$\underline{E}_{\text{inc}}(x, z) = E_y \hat{y} = E_0 e^{-jk_0 n_1 \sin \vartheta_1 x} e^{-jk_0 n_1 \cos \vartheta_1 z} \hat{y}$$

Si consideri ora la rappresentazione spettrale della tensione, in questo caso coincidente con la tensione progressiva, per la sezione $z = 0$:

$$\begin{aligned} V^+(\xi, 0) &= \int_{-\infty}^{+\infty} E_y(x, 0) e^{+j\xi x} dx = \int_{-\infty}^{+\infty} E_0 e^{-jk_0 n_1 \sin \vartheta_1 x} e^{j\xi x} dx = \\ &= 2\pi E_0 \delta(\xi - k_0 n_1 \sin \vartheta_1) \end{aligned}$$

questo permette di “selezionare”⁵ un preciso valore di ξ .

Questa, è un’onda piana: essa è caratterizzata dal solito fattore di propagazione: un esponenziale che dipende da x e z , con un certo ξ selezionato.

Facciamo un discorso geometrico:

Si può vedere che un certo valore di ξ indica una particolare direzione dell’onda piana; se si ha un’espressione del tipo

$$V_0^+ e^{-jk_z z} e^{-j\xi x}$$

guardando questo esponenziale, se \underline{k} è composto da k_z e ξ , essendo $k_y = 0$, per forza si ha che ξ è la componente lungo l’asse x del vettore \underline{k} ! Questo è il significato fisico/geometrico del nostro ξ : esso è la componente trasversale del vettore d’onda; al contrario, k_z è la componente lungo z del vettore d’onda.

Essendo il modo TE, la tensione modale è proporzionale alla trasformata di E_y : come detto, il campo elettrico va proprio lungo \hat{y} . La direzione del campo magnetico può essere determinata, applicando la regola della mano destra alla relazione di impedenza, trovando che si ha ciò:

questo, dal momento che:

$$\underline{H} \propto \hat{s} \times \underline{E}$$

Alternativa, è dire che, se si considera E_y positivo (uscendo dal foglio), la corrente modale è l’ammettenza modale moltiplicata per la tensione, ma H_x ha un segno $-$: la componente di H_x deve essere dunque negativa, sempre dall’espressione della tensione modale, quindi questo è il risultato finale.

Riassumendo: considerare un particolare valore di ξ coincide con il considerare una specifica onda piana. Spesso, ciò che si dà, per indicare un’onda piana, è un certo angolo di incidenza, ϑ_1 ; questo permette di capire, dal disegno (o dalle espressioni precedentemente ricavate), che:

⁵il fatto che infatti $V^+(\xi, 0)$ sia non nulla per $\xi = k_0 n_1 \sin \vartheta_1$, infatti, significa che si sta considerando un solo modo, e che non ha senso calcolare circuiti per altri ξ

$$\xi = k \sin \vartheta_i = k_0 n_1 \sin \vartheta_i$$

k_z si può trovare o con la formula delle guide o, equivalentemente, con il teorema di Pitagora, ricavando banalmente (come da prima)

$$k_z = k \cos \vartheta_i = k_0 n_1 \cos \vartheta_i$$

Impedenza del mezzo e impedenza modale

Precedentemente, si era parlato dell'impedenza intrinseca del mezzo, che si può definire come rapporto tra i moduli del campo elettrico e del campo magnetico:

$$Z_0 = \sqrt{\frac{\mu}{\varepsilon}} = \frac{|\underline{E}|}{|\underline{H}|}$$

L'impedenza modale, invece, è un qualcosa di diverso:

$$Z_{\infty}^{\text{TE}} = \frac{V^+}{I^+} = \frac{E_y}{H_x}$$

Le due sono diverse: di fatto, il campo E_y coincide con il campo totale, dal momento che il campo elettrico ha solo componente lungo \hat{y} , ma la stessa cosa non si può dire per il campo magnetico, che ha componente sia lungo \hat{x} , sia lungo \hat{z} ! In questo caso, dunque, si ha:

$$\frac{E_y}{H_x} = \frac{|\underline{E}|}{|\underline{H}| \cos \vartheta_i}$$

manipolando l'espressione:

$$= \sqrt{\frac{\mu}{\varepsilon}} \frac{1}{\cos \vartheta_i}$$

questo significa che:

$$Z_{\infty}^{\text{TE}} = \frac{Z_0}{\cos \vartheta_i}$$

Questo risultato, geometrico, deve essere ottenuto anche tramite il formalismo elettrico: l'impedenza modale deve infatti essere il rapporto tra $\omega\mu$ e k_z , ma ciò dà lo stesso risultato.

A questo punto si può affrontare il problema specifico: studiare l'ampiezza dell'onda trasmessa e di quella riflessa di una certa onda incidente, con un certo angolo. Il metodo si mette in pratica come per le guide; come prima cosa

si disegna il circuito modale equivalente, quindi si hanno una linea a sinistra con una certa impedenza caratteristica e costante di propagazione, una linea a destra con altre impedenze caratteristiche e costanti di propagazione, quindi una discontinuità in un punto A.

Si noti che, una volta disegnata in questo modo questa struttura, ci siamo già “sbilanciati molto”, senza tenere conto di un effetto molto importante: di fatto, il fatto che si disegnino due linee modali relative allo stesso modo, implica che, dato un certo ϑ_1 in ingresso, si ha un solo modo che si propaga dopo la discontinuità; in altre parole, ciò che è stato implicitamente detto con questo disegno è il fatto che, in una situazione come quella appena descritta, non si eccitano altri modi.

Questa affermazione è molto forte, dal momento che di solito non è verificata: se si considera un’iride in guida metallica, un “diaframma”, e vi si manda un modo TE_{10} sopra, non è assolutamente detto che nessun altro modo si eccita: dal momento che si devono soddisfare le condizioni al contorno, e che il campo elettrico del TE_{10} è indipendente da y , si deve garantire che il campo elettrico sul metallo sia nullo, e diverso da 0 nel buco, solo con un TE_{10} : il fatto di avere questa discontinuità porta a richiedere la presenza di altri elementi della base modale, nella fattispecie dei modi superiori, almeno in prossimità della discontinuità, al fine di poter soddisfare le condizioni al contorno. Quello che si ha in pratica è una situazione di questo genere:

Si eccita un TE_{10} riflesso, un TE_{20} riflesso e uno trasmesso, TE_{30} , TE_{40} e così via riflessi e trasmessi: si ha un fenomeno di accoppiamento modale. La “matrice scattering” dice sostanzialmente quanto di un modo venga eccitato dal modo di partenza. I vari modi, più salgono di ordine, più sono evanescenti; si eccitano infiniti modi, diversi a seconda del tipo di discontinuità. Accoppiamento modale significa: mandando un modo in un punto, in questo punto si eccitano molti modi.

Il fatto di unire, senza avere questo “blocco di accoppiamento”, due linee modali con lo stesso modo, presuppone il fatto che semplicemente si abbia, dopo la discontinuità, lo stesso modo, nel nostro esempio TE, con lo stesso indice modale. Noi dell’onda conosciamo l’angolo ϑ_1 , ma si tenga ben presente il fatto che non è ϑ_1 l’indice modale: è ξ l’indice modale. I due angoli di incidenza e di trasmissione, sono diversi, ma i modi sono gli stessi: gli ξ devono essere uguali, dunque, indipendentemente dagli angoli, lo devono essere anche i modi.

Ciò che ci garantisce il fatto che le due linee modali possano essere collegate “a cuor leggero”, sono le condizioni all’interfaccia, sui campi tangenziali, dunque, per come è disposta l’interfaccia, trasversali a z : essi sono E_y e H_x . Dire che sono continui, significa che devono esserlo per ogni valore di x .

A sinistra, i campi dipendono da x come $e^{-j\xi_1 x}$; a destra, come un $e^{-j\xi_2 x}$; è dunque evidente che, se i due fattori devono essere identici, $\xi_1 = \xi_2 = \xi$; perchè dunque la condizione al contorno sia verificata, si deve avere:

$$\xi_{\text{sinistra}} = \xi_{\text{destra}} \implies k_0 n_1 \sin \vartheta_i = k_0 n_2 \sin \vartheta_t$$

questa, banalmente, è la legge di Snell: il fatto che essa sia verificata, ci tranquillizza sul fatto che probabilmente tutto ciò che è stato fatto finora è ragionevolmente giusto.

A questo punto, possiamo **risolvere il problema**, e questo significa calcolare le ampiezze del campo riflesso e del campo trasmesso; questo, in sostanza, coincide con il calcolo del coefficiente di riflessione all'interfaccia, Γ_{A^-} ; dal momento che conosciamo tuttavia la teoria delle linee di trasmissione, è banale capire che:

$$\Gamma_{A^-} = \frac{Z_{\infty 2} - Z_{\infty 1}}{Z_{\infty 2} + Z_{\infty 1}}$$

a questo punto, sostituendo le espressioni per le onde TE:

$$= \frac{\frac{\omega\mu}{k_{z2}} - \frac{\omega\mu}{k_{z1}}}{\frac{\omega\mu}{k_{z2}} + \frac{\omega\mu}{k_{z1}}} = \frac{k_{z1} - k_{z2}}{k_{z1} + k_{z2}} =$$

esprimendo tutto come

$$k_{z2} = k_0 n_2 \cos \vartheta_t$$

è scomodo, perchè si dovrebbe calcolare l'angolo; l'idea, è quella di usare direttamente la legge di Snell, e vedere, per il teorema di Pitagora, che

$$k_{z2} = k_0 \sqrt{n_2^2 - n_1^2 \sin^2 \vartheta_i}$$

quindi:

$$= \frac{n_1 \cos \vartheta_i - \sqrt{n_2^2 - n_1^2 \sin^2 \vartheta_i}}{n_1 \cos \vartheta_i + \sqrt{n_2^2 - n_1^2 \sin^2 \vartheta_i}}$$

Si vogliono proporre a questo punto alcune note, riguardo tutto ciò che è stato fatto.

La prima osservazione che si può fare, è che, nel modello da noi introdotto, non si ha a che fare con un generatore: conosciamo infatti solo il campo incidente sulla superficie; nella fattispecie, noi possiamo supporre di conoscere la tensione progressiva sulla discontinuità, $V_{A^-}^+$, quantità **associata a un modo**; ciò che identifica il modo, come già detto, è un preciso valore di ξ .

Per calcolare Γ_{A^-} , utilizzando il formalismo delle linee modali equivalenti, è necessario utilizzare le ben note formule:

$$\Gamma_{A^-} = \frac{\zeta_{A^-} - 1}{\zeta_{A^-} + 1}$$

dove

$$\zeta_{A^-} = \frac{Z_A}{Z_{\infty 1}}$$

e

$$Z_A = Z_{\infty 2}$$

Si noti che si sta parlando di Z_A , senza specificare che si consideri la parte sinistra o destra della sezione; questo fatto è dovuto al fatto che Z_A è un'impedenza, non un'impedenza normalizzata, quindi è indipendente dalla zona in cui si trova. Un aspetto più interessante invece è quello che riguarda un altro tipo di notazione: sarebbe interessante usare, come notazione, \vec{Z}_A : la freccia “verso destra” sta a indicare il fatto che questa impedenza è quella che si vede “guardando verso destra”. La necessità di definire questo dettaglio non dipende, si noti, dal fatto che si ha un'onda incidente verso destra, bensì dalla convenzione di segno della corrente: guardare a “destra” significa sostanzialmente guardare verso dove si ha il verso della corrente positiva: questa è la convenzione standard per la corrente, per misurare un'impedenza positiva.

Si noti che il fatto che l'onda sia “da destra a sinistra”, in un secondo problema, si può analizzare in maniera “furba”: ciò che si può tuttavia fare, al fine di studiare questo problema, è considerare lo stesso problema di prima, scambiando i valori di n_1 e n_2 : in questo modo il problema è a tutti gli effetti uguale, e non è necessario raddoppiare le formule.

Si era visto che:

$$\Gamma_{A^-} = \frac{n_1 \cos \vartheta_i - \sqrt{n_2^2 - n_1^2 \sin^2 \vartheta_i}}{n_1 \cos \vartheta_i + \sqrt{n_2^2 - n_1^2 \sin^2 \vartheta_i}}$$

Queste radici quadrate, per questioni computazionali, sono da tenere d'occhio; prima di tutto, si può notare che esse hanno un significato preciso: esse sono legate a k_{z2} , ossia a k_z nel mezzo 2, dal momento che:

$$k_{z2} = k_0 \sqrt{n_2^2 - n_1^2 \sin^2 \vartheta_i}$$

questa radice, quindi, è proporzionale alla costante di propagazione nel mezzo 2.

Casistica 1: $n_1 < n_2$

Data questa espressione, studiamola, al fine di capire nelle varie circostanze che valori può assumere; sostanzialmente, dobbiamo scomporre in casistiche, a seconda dei vari parametri. È uso utilizzare, come prima casistica, quella per cui

$$n_1 < n_2$$

ossia, la casistica per cui si ha a sinistra un mezzo meno denso di quello a destra.

Prima di tutto, si può osservare (non solo in questa casistica in realtà), che l'espressione non dipende dalla frequenza; questo, dal momento che non si ha k_0 tra i parametri. Questa è un'ipotesi ragionevole, che però non è sicuramente corretta, dal momento che gli indici di rifrazione sono funzioni della frequenza a causa della dispersività del materiale; possiamo tuttavia ignorare questo fatto; come unica variabile, dunque, abbiamo ϑ_i , l'angolo di incidenza.

Il fatto di non avere un indice di rifrazione che dipende dalla frequenza è estremamente interessante, anche perchè deriva da una proprietà generale: se si pensa a quale sia l'unico componente circuitale la cui impedenza non dipende dalla frequenza, quello è la resistenza; se in un circuito c'è un condensatore, o un induttore, o comunque una combinazione di reattanze e resistenze, si ha dipendenza dalla frequenza; questa cosa dipende dal fatto che, se sono presenti oggetti che possono immagazzinare energia, allora si ha dipendenza delle caratteristiche della struttura dalla frequenza: dispersione. La struttura che stiamo analizzando non è dispersiva, dal momento che, se n_1 e n_2 non dipendono dalla frequenza, l'energia non si potrebbe immagazzinare da nessuna parte; i campi infatti immagazzinerebbero energia, ma, dal momento che viaggiano, non c'è nessuna "scatola" che possa mantenere l'energia in un certo posto: le energie riflessa e trasmessa *scappano all'infinito*. Come si vedrà, se ci sono due pareti, l'energia tende a rimanere dentro un volume limitato.

Abbiamo detto che ϑ_i è la variabile del problema; studiamo a questo punto varie situazioni.

Per $\vartheta_i = 0$, si ha che i coseni vanno a 1, i seni a 0, quindi l'espressione del Γ diventa (dove per Γ si intende ovviamente l'unico che ci interessa, Γ_{A-}):

$$\Gamma = \frac{n_1 - n_2}{n_1 + n_2}$$

per ipotesi, tuttavia, si ha che $n_1 < n_2$; per questo motivo, dunque, sicuramente $\Gamma < 0$. Per andare avanti con i ragionamenti, al variare di ϑ_i , si può notare che:

$$n_1 \cos \vartheta_i = \frac{k_{z1}}{k_0} = \sqrt{n_1^2 - n_2^2 \sin^2 \vartheta_i}$$

Questa notazione è molto utile dal momento che la radice quadrata, al proprio interno, ha il \sin^2 , proprio come l'altra; ciò che cambia è n_1^2 in una, n_2^2 nell'altra; questo significa che la prima radice sarà sempre più piccola della seconda e che quindi, per ogni valore di ϑ_i , si avrà un Γ negativo, minore di zero.

Infine, si può osservare che il valore massimo del Γ , in modulo, sia per $\vartheta_i \rightarrow \frac{\pi}{2}$: questo è il valore limite; a $\pi/2$ non si può arrivare, dal momento che, se l'angolo fosse esattamente $\pi/2$, l'onda sarebbe parallela all'interfaccia, radente a essa. Il risultato sarebbe $\Gamma = -1$.

Disegnando sulla carta di Smith questa cosa, si può vedere che dunque tutti i punti sono reali (entrambe le radici infatti hanno radicandi maggiori di zero), e negativi: sulla carta di Smith, che rappresenta il piano complesso dei Γ , tutti i valori staranno dunque sull'asse orizzontale (asse reale), a sinistra.

Un altro modo di studiare $\Gamma(\vartheta_i)$ è quello di disegnarlo su un normale diagramma cartesiano:

A questo punto, per concludere, è interessante disegnare il diagramma di tensione e di corrente di questa linea, ossia i diagrammi di onda stazionaria, per un certo angolo ϑ_i **fissato**. Come si può ricordare dalla teoria precedentemente studiata, si ha che $|V| \propto |E_y|$, dal momento che l'unica componente non nulla del campo elettrico è proprio quella lungo \hat{y} ; per quanto riguarda la corrente, invece, si ricordi che la componente che conta, per quanto riguarda la determinazione di I , è la componente trasversale del campo: H_x .

La carta di Smith, come si può ricordare dai corsi di Campi Elettromagnetici, è la rappresentazione in coordinate polari del coefficiente di riflessione di tensione Γ ; per fare il diagramma di tensione, dunque, si deve avere che:

$$V(z) = V^+(z) + V^-(z)$$

dove

$$V^+(z) = V_0^+ e^{-jk_z z}$$

e

$$V^-(z) = \Gamma(z)V^+(z)$$

La carta di Smith è utile per capire “dove ci si trova”: si finisce infatti per avere un’espressione del tipo:

$$V(z) = V_0^+ [1 + \Gamma(z)]$$

Questo significa che, studiando il vettore $1 + \Gamma(z)$ sulla carta di Smith, si può determinare l’intensità della tensione totale, rispetto alla tensione progressiva V_0^+ , che può per esempio essere un coefficiente di normalizzazione rispetto a cui si disegna il diagramma d’onda stazionaria. Si noti inoltre, volendo fare tutto “a mano”, che $1 + \Gamma(z)$ è anche il “coefficiente di trasmissione” sulla carta di Smith (spesso indicato come $T(z)$), e ciò è interessante dal momento che spesso è indicato sia in modulo sia in fase sulle scale del diagramma.

Si noti che la dipendenza da z , all’esponenziale, è da k_z , **non da k** : perchè la fase dell’esponenziale abbia fase che vari di 2π , si deve avere che $z = \lambda_g$: si deve introdurre un concetto di lunghezza d’onda guidata. La costante di propagazione della linea modale è k_{z1} , dal momento che sulla linea si risolve solo la questione relativa alla propagazione sulla direzione z , e quindi la λ in realtà è una λ_g . Per definizione, infatti, la lunghezza d’onda è l’incremento che si deve dare a z per fare in modo che una funzione riacquisti lo stesso valore: un periodo spaziale. Si ha dunque:

$$e^{-jk_z\lambda_g} = 1 \implies k_z\lambda_g = 2\pi \implies \lambda_g = \frac{2\pi}{k_z}$$

Tutto questo ragionamento, la “formula” per $V(z)$, vale sicuramente a sinistra della discontinuità di impedenza A; cosa si può tuttavia dire, a destra? Vale ancora tutto? No: a destra si ha sostanzialmente una linea di lunghezza infinita, dunque non si ha nessuno “specchio” che possa far tornare indietro del segnale: c’è solo l’onda progressiva, dunque non si hanno fenomeni di interferenza tra modi progressivi e regressivi, quindi l’ampiezza dell’onda è costante. Si noti infine che sulla discontinuità, si ha un minimo: considerando infatti come “origine” dell’asse z la discontinuità, cosa sensata dal momento che il Γ è noto proprio in quella posizione, si ha che lì

$$V(A) = V_0^+ [1 + \Gamma_{A-}]$$

Il diagramma di onda stazionario si disegna calcolando mediante il solito procedimento noto dai corsi di Campi, introducendo l’impedenza normalizzata rispetto all’impedenza caratteristica della linea equivalente sulla carta di Smith, e valutandone il valore in modulo e fase; dalle considerazioni finora fatte, tuttavia, abbiamo la certezza del segno di Γ_{A-} , dal momento che sappiamo che esso è negativo; questo significa, per il diagramma di tensione, che

ci si trova in un punto di minimo, dal momento che il vettore “coefficiente di trasmissione” è minimo sui punti dell’asse reale negativi.

Casistica 2: $n_1 > n_2$

A questo punto, cambiamo l’ipotesi, considerando $n_1 > n_2$; questo porterà a conseguenze abbastanza particolari, delle quali sarà necessario tenere conto anche in un contesto computazionale (ossia anche al momento di lavorare su un software per il calcolo).

La principale conseguenza è sostanzialmente sulla radice quadrata, la quale ha un radicando di cui prima non ci si preoccupava: prima si aveva $n_2 > n_1$, e n_1 era oltretutto moltiplicato per un seno, cosa che poteva solo ridurlo ulteriormente, garantendo la positività del radicando. Ora invece n_2 è più piccolo, ma, a seconda di ϑ_i , si potranno avere valori positivi o negativi del radicando.

Ripetiamo le osservazioni di prima, al variare di ϑ_i (ripetendo dunque il “primo esperimento” della precedente sottosezione):

- per $\vartheta_i = 0$, l’espressione è quella che si era ottenuta prima:

$$\Gamma_{A^-} = \frac{n_1 - n_2}{n_1 + n_2}$$

si ha solo una sostanziale differenza: in questo caso, infatti, si ha che $\Gamma > 0$. Questo sarà generalmente vero, almeno fino a certi valori di angolo;

- al crescere di ϑ_i , inizialmente, Γ cresce, tendendo ad arrivare a +1;
- si osservi la tanto temuta radice quadrata:

$$\sqrt{n_2^2 - n_1^2 \sin^2 \vartheta_i}$$

il suo argomento continua a diminuire, facendo tendere Γ a +1; a un certo punto, tuttavia, all’aumentare di ϑ_i , si ha che

$$n_2^2 = n_1^2 \sin^2 \vartheta_i$$

considerando tutte le radici positive (discorso che si dovrà discutere tra breve), si arriva a una condizione limite per cui

$$\sin \vartheta_i = \frac{n_2}{n_1}$$

invertendo, si trova il cosiddetto **angolo critico**:

$$\vartheta_c = \arcsin\left(\frac{n_2}{n_1}\right)$$

questo angolo viene chiamato “critico” dal momento che rappresenta il “punto di separazione” tra due situazioni ben diverse: quella per cui Γ è un numero reale, e quella per cui Γ è un numero complesso, con fase diversa da 0 o π .

- se $\vartheta_i > \vartheta_c$, si ha:

$$\Gamma = \frac{n_1 \cos \vartheta_i + j\sqrt{n_1^2 \sin^2 \vartheta_i - n_2^2}}{n_1 \cos \vartheta_i - j\sqrt{n_1^2 \sin^2 \vartheta_i - n_2^2}}$$

come mai abbiamo scritto questa cosa? Cerchiamo di interpretarla: prima di tutto, abbiamo “scambiato i due termini sotto radice”, “estraendo”, “raccogliendo” un termine $\sqrt{-1}$, che fa j ; in realtà, però, è stato anche fatto un cambio di segno ulteriore; l’espressione iniziale, infatti, era:

$$\Gamma_{A^-} = \frac{n_1 \cos \vartheta_i - \sqrt{n_2^2 - n_1^2 \sin^2 \vartheta_i}}{n_1 \cos \vartheta_i + \sqrt{n_2^2 - n_1^2 \sin^2 \vartheta_i}}$$

questa cosa è giusta, ma va spiegata con calma.

Come mai, nel punto conclusivo, è stato fatto questo trucco? La risposta è abbastanza semplice: si ricordi che la tensione progressiva (è sufficiente essa, dal momento che per la regressiva il discorso risulta corretto una volta messo a posto per la progressiva), si ha a che fare con:

$$V(z) = V_0^+ e^{-jk_z z}$$

questa cosa deriva dalla notazione fasoriale che abbiamo scelto, ossia dal fatto di utilizzare $e^{j\omega t}$ al momento di effettuare l’antitrasformata del fasore. k_z , nei casi analizzati prima d’ora, era sempre un numero reale, dunque il termine $V(z)$ aveva un esponenziale che portava solo a una variazione di fase. Nel caso in cui tuttavia k_z sia un numero complesso, si ha un termine esponenziale con argomento reale, ossia un termine che porta ad avere una variazione nell’ampiezza.

A questo punto si ha un problema di cui bisogna tenere conto: quando si calcola la radice quadrata di qualcosa (come in questo caso), si può per esempio avere:

$$\sqrt{-4} = \pm j2$$

questo è un problema ed è un problema importante, dal momento che di solito si usa sempre la radice con parte reale positiva, ma in questo caso non è questa “abitudine” la priorità: ora la priorità è avere un modello matematico che funzioni, e un modello funziona ogni qual volta esso non diverge. Se si avesse un termine esponenziale per cui, al crescere della distanza z dal punto di partenza, cresce anche la tensione, si avrebbe qualcosa di fisicamente insensato: il tradizionale caso in cui si crea energia dal nulla. Bisogna dunque fare in modo che, al crescere di z , la tensione venga smorzata esponenzialmente e, come si può vedere, questo si ha solo se la parte immaginaria ottenuta dalle radici è **negativa**; dato infatti un numero immaginario del tipo

$$z = -j|z|$$

dove ovviamente $|z|$ è reale positivo, si ha:

$$e^{-j(-j|z|)} = e^{-|z|}$$

che è effettivamente fisicamente possibile. Ha dunque senso utilizzare una scrittura del tipo “ $-j|z|$ ”, in modo da essere sicuri del segno della radice negativa; il fatto di aver scambiato questo segno in effetti garantisce l’assenza di indeterminazioni sulla scrittura: “funziona sempre”.

Prima di proseguire con l’analisi, a questo punto si può discutere un aspetto computazionale: la soluzione dell’equazione $x^2 = a$, dove a è un certo numero, non è una funzione, dal momento che ci sono più valori di uscita per un singolo ingresso (infatti, \sqrt{a} produce $\pm a$, e questo in una funzione non ha senso, dal momento che una funzione associa un ingresso a una singola uscita). Di solito, per questioni “tradizionali”, delle due *determinazioni*, si utilizza quella a parte reale positiva.

Siamo a questo punto interessati a come lavori MATLAB, dal momento che esso (e i software simili a esso) di fatto hanno dei criteri per la determinazione della fase delle radici. MATLAB calcola le radici basandosi sul fatto che, nel piano complesso, si ha una “linea di ramificazione” che parte da $-\infty$ e sta leggermente sotto l’asse reale:

questa, permette di determinare la fase di un numero z ; si ha, nella fattispecie, la funzione $\text{angle}(z)$. In MATLAB, si ha che:

$$-\pi < \text{angle}(z) \leq \pi$$

si noti che da una parte si ha un $<$, dall'altra un \leq : l'uguale è messo "davanti", in modo da eliminare quindi l'indeterminazione sulla fase del numero complesso. Questo discorso è collegato a quello delle radici, dal momento che MATLAB calcola le radici a partire dalla formula di De Moivre:

$$\sqrt{z} = \sqrt{|z|} e^{j\frac{\angle z}{2}}$$

MATLAB nella fattispecie calcola solo la prima. Studiamo a questo punto il comportamento, nel primo citato caso $\sqrt{-4}$; applicando la formula di De Moivre sulla prima radice,

$$\sqrt{-4} = \sqrt{2} e^{j\frac{\pi}{2}} = j2$$

la radice positiva deriva dal fatto che il \leq è sul secondo termine. Questa, purtroppo, è la determinazione "sbagliata": a noi, infatti, servirebbe quella di segno negativo. Questa cosa va tenuta in conto, al momento di programmare su software di questo tipo.

Torniamo alla formula di prima:

$$\Gamma = \frac{n_1 \cos \vartheta_i + j\sqrt{n_1^2 \sin^2 \vartheta_i - n_2^2}}{n_1 \cos \vartheta_i - j\sqrt{n_1^2 \sin^2 \vartheta_i - n_2^2}}$$

Come si può notare, il numeratore ha una parte reale e una parte immaginaria, e il denominatore è il complesso coniugato del numeratore; si può vedere, calcolando il modulo, che:

$$|\Gamma| = 1$$

$$\angle \Gamma = 2 \arctan \left(\frac{\sqrt{n_1^2 \sin^2 \vartheta_i - n_2^2}}{n_1 \cos \vartheta_i} \right)$$

Utilizzando queste osservazioni è possibile determinare il comportamento del coefficiente di riflessione per angoli superiori all'angolo critico; nella fattispecie, per $\vartheta \rightarrow \frac{\pi}{2}$, il coseno tende a 0, quindi l'argomento tende a $+\infty$, e l'arcotangente a π . Sulla carta di Smith, si ottiene qualcosa di questo genere:

Si parte da $\frac{n_1 - n_2}{n_1 + n_2}$ che è positivo, quindi facendo crescere ϑ_i si cresce fino a $+1$, quindi si inizia a ruotare fino ad arrivare a -1 . Questo risultato è

fondamentale, dal momento che esso semplicemente esprime il fenomeno della riflessione totale⁶.

Si parla di riflessione totale, ma per quale motivo? Di fatto, nel mezzo 2, il campo non è nullo, bensì è esponenzialmente decrescente! Di fatto dunque, il coefficiente di riflessione è unitario, ma si ha del campo anche nell'altro mezzo. La risposta alla domanda è nascosta in un concetto diverso: in realtà, come stiamo per dimostrare, la riflessione è totale sotto il punto di vista energetico, dal momento che **non si ha transito di potenza nel mezzo 2**. Dimostriamolo:

$$P_r = \frac{1}{2} \operatorname{Re} \{ V_{A^+}^+ I_{A^+}^{+*} \}$$

in realtà, questa sarebbe semplicemente la ben nota:

$$P(z) = \frac{1}{2} \operatorname{Re} \{ V(z) I^*(z) \}$$

dal momento tuttavia che “a destra” si ha sostanzialmente una condizione di adattamento, le tensioni e correnti totali coincidono con quelle progressive. Si sa, poi, che:

$$I_{A^+}^+ = Y_{\infty 2} V_{A^+}^+$$

Dove però l'impedenza/ammettenza caratteristica è quella modale, la cui espressione è ben nota:

$$Y_{\infty 2} = \frac{k_{z2}}{\omega \mu}$$

ma, d'altra parte, per le questioni di notazione/convenzione precedentemente discusse, si ha che:

$$k_{z2} = -j |k_{z2}|$$

ossia, esso è un numero puramente immaginario; si ha dunque:

$$P_r = |V_{A^+}^+| \operatorname{Re} \{ Y_{\infty 2} \} = 0$$

dal momento che l'impedenza caratteristica modale ha solamente una parte immaginaria pura. Questo dimostra il fatto che la P_2 è dunque 0: tutta la potenza P_1 incidente viene riflessa. Questa cosa può essere dimostrata anche utilizzando la formula usuale per la potenza:

⁶nota: si parla di riflessione totale dal momento che $|\Gamma| = 1$; non ha senso, per esempio nel caso di perdite, dire la stessa cosa: il fatto che il modulo del coefficiente di riflessione sia unitario è un fatto vincolante, per la definizione appena fornita

$$P_{A^-} = \frac{1}{2} \frac{|V_{A^-}^+|^2}{Z_{\infty 1}} [1 - |\Gamma_{A^-}|^2]$$

nella condizione in cui siamo, ossia la riflessione totale, il modulo del coefficiente di riflessione è infatti unitario, dunque la potenza sulla sezione A^- è sostanzialmente nulla, il che significa che tutta la potenza che vi arriva deve anche “tornare indietro”: la componente di potenza progressiva eguaglia la componente di potenza regressiva. Come dimostrato e accennato, dunque la riflessione è totale sotto il punto di vista energetico.

Una nota: come mai è stata usata la formula “di partenza”, quella complicata, e non semplicemente

$$P_2 = \frac{|V_{A^+}|^2}{2Z_{\infty 2}}$$

al fine della dimostrazione? Beh, per il semplice fatto che essa è sbagliata (anche dal momento che non fa 0): questa formula, vale esclusivamente per valori di Z_{∞} reali, ossia nel caso non ci siano perdite o nel caso ci siano solo piccole perdite; in altri casi, la formula non è valida; per questo motivo, al fine di effettuare una dimostrazione rigorosa, sarebbe necessario fare uso della formula di partenza, che invece non richiede osservazioni di questo genere.

Parliamo a questo punto del grafico della tensione, ossia del diagramma d’onda stazionaria; se l’angolo di incidenza è minore dell’angolo critico, si ha sostanzialmente un punto di massimo, invece che di minimo, all’interfaccia (dal momento che ora siamo “a destra” e non “a sinistra” dell’origine della carta di Smith, ma sempre sull’asse reale); si gira poi verso il generatore sulla carta di Smith. A “destra dell’interfaccia” si ha un’onda progressiva che si propaga. Nel caso invece $\vartheta_i > \vartheta_c$, si ha qualcosa di diverso: nel punto di partenza, siamo “prima del massimo”: ruotando sulla carta di Smith *toward generator*, si arriva verso il massimo, e poi verso il minimo, e così via; il minimo in questo caso è 0, dal momento che il modulo di Γ è unitario, dunque si ha una cancellazione completa della tensione. A destra, si ha il campo esponenzialmente decrescente: $\propto e^{-|k_{z2}|z}$.

Un’altra nota: dal punto di vista dei vettori d’onda, vale la solita legge di Snell:

$$n_1 \sin \vartheta_i = n_2 \sin \vartheta_t$$

Quello che si può fare, dunque, è considerare le due situazioni: se $n_1 < n_2$, dal punto di vista dei vettori d’onda, si ha ciò:

in questo caso, come si può vedere dall’equazione precedente, si ha evidentemente che $\vartheta_t < \vartheta_i$; questo significa che la lunghezza del vettore d’onda

\underline{k} a sinistra è $k_0 n_1$; quella del vettore a destra, è $k_0 n_2$; dal momento che i due ξ devono essere uguali, sicuramente si avrà k_{z2} più lungo di k_{z1} .

Più interessante è sicuramente il caso $n_2 < n_1$:

in questo caso, se faccio crescere ϑ_i , si arriva a un certo punto per cui $\vartheta_t = \pi/2$; se dunque si ha $\vartheta_i > \vartheta_c$, cosa succede? Beh, precedentemente, la teoria ci ha detto che:

$$k_{z2} \longrightarrow -j |k_{z2}|$$

L'angolo critico è l'angolo di incidenza tale per cui $\vartheta_t = 90^\circ$. Cosa capita, se $\vartheta_i > \vartheta_c$? Beh, se si volesse interpretare il disegno "meccanicamente", sembrerebbe che l'onda torna indietro; questo ovviamente non ha assolutamente senso, dunque non bisogna farsi ingannare dal disegno:

$$\underline{k}_2 = \xi \hat{x} - j |k_{z2}| \hat{z}$$

questo è il modello che deve guidarci. L'onda piana "a destra" non è definita a sinistra, ma solo a destra: non può tornare indietro.

Per quanto riguarda dunque l'onda "a destra", ha senso chiedersi quali siano i piani a fase costante, ma anche i piani ad ampiezza costante; l'onda piana, come si può vedere da \underline{k}_2 , ha un'espressione del tipo:

$$e^{-j\xi x} e^{-|k_{z2}|z}$$

disegnare infatti la componente reale e la componente immaginaria, per un vettore, non è possibile; questa è la soluzione grafica più semplice da realizzare.

Come si può vedere, la fase dipende solo da ξ , dunque i piani a fase costante sono i piani a x costante: fissato un certo valore di x , la fase è fissa, al variare delle altre variabili; nella fattispecie, dal momento che il nostro problema è bidimensionale, l'unica altra variabile presente è z . Cosa significa ciò? Significa che l'onda ha fase variabile lungo \hat{x} , ma **non** lungo \hat{z} .

Si noti che l'onda incidente è un'onda piana: non è detto che l'onda sia incidente solo nel punto di incidenza della "freccia", dal momento che le superfici di fase sono dei piani, e i piani sono superfici illimitate; non ha dunque senso dire che il campo esiste solo dove esiste la freccia; nella fattispecie, il campo esiste su tutta l'interfaccia, su ogni punto in cui il piano si interseca con l'interfaccia! Qualunque perpendicolare ai piani a fase costante si può usare come vettore \underline{k} . Si noti che ξ è un numero sempre positivo:

$$\xi = k_0 n_1 \sin \vartheta_i$$

e, facendo il conto, viene sempre positivo.

Per quanto riguarda i piani ad ampiezza costante, essi sono sostanzialmente i piani a z costante: fissato un z , per tutti gli x , l'ampiezza ha sempre gli stessi valori, dal momento che l'andamento dell'ampiezza è una coda esponenziale lungo l'asse z . Questa “coda esponenziale” è l'**onda evanescente**. Per l'onda evanescente si può definire una “costante di spazio”, in maniera analoga a quanto si fa nei circuiti RC :

$$d = \frac{1}{|k_{z2}|}$$

volendo, è possibile esplicitare meglio questa espressione:

$$d = \frac{1}{k_0 \sqrt{n_1^2 \sin^2 \vartheta_i - n_2^2}} = \frac{\lambda_0}{2\pi \sqrt{n_1^2 \sin^2 \vartheta_i - n_2^2}}$$

Questa formula ci permette di smentire un “luogo comune”: di solito si dice che la profondità a cui si spinge l'onda evanescente è pari a λ_0 circa; questo non è assolutamente vero, dal momento che si ha un'importante dipendenza da parte dell'angolo di incidenza dell'onda piana sulla superficie di discontinuità: se l'angolo di incidenza è appena superiore all'angolo critico, infatti, il denominatore potrebbe assumere valori come $\frac{1}{100}$ o cose del genere, ottenendo di fatto campi che decadono in molte lunghezze d'onda; per determinare d , dunque, bisogna fare i conti con ϑ_i .

Questa onda ha dunque i piani a modulo costante e a fase costante tra loro ortogonali; questo tipo di onda è detta “onda non omogenea”; le onde omogenee, invece, non hanno sostanzialmente neanche definiti i piani a modulo costante, dal momento che il modulo è costante ovunque.

Volendo approfondire il discorso delle onde omogenee, si può dire che un'onda è detta “omogenea” quando, dato

$$\underline{k} = \underline{k}' + j\underline{k}''$$

si ha

$$\underline{k}' // \underline{k}''$$

ossia, quando si può dire che la direzione di propagazione dell'onda è reale.

L'onda evanescente esiste, e può essere vista come una sorta di “barba del campo” sull'interfaccia: il campo infatti “sborda” dall'uscita, ma ciò può essere usato, in modo intenzionale o meno, per effettuare un accoppiamento. Quello che sappiamo, dal modello basato sulla teoria modale per la rappresentazione secondo linee equivalenti del sistema ottico in questione, è

il fatto che non si ha potenza transitante sulla “linea a destra”; questo tuttavia considera e risolve il problema solo e unicamente per quanto riguarda la direzione z : la linea di trasmissione infatti, noti i suoi parametri, risolve il problema solo lungo z . Nello spazio, tuttavia, il vettore di Poynting non è nullo, dal momento che si hanno parti reali del medesimo non nulle lungo \hat{x} , dunque sostanzialmente sull’interfaccia, e questo perchè ξ è positivo; si ha della potenza che “scivola” lungo la superficie, “aggrappata all’interfaccia”.

Un’altra nota: una prima osservazione sembrerebbe mostrare che, se $d \rightarrow 0$, si avrebbe un qualcosa di simile a una “discontinuità di potenza” all’interfaccia; la verità è che però, come si può osservare dalle formule complete, più d tende a ridursi, più il punto del diagramma d’onda stazionario tende a diventare un minimo: quando il coefficiente di riflessione tende a diventare -1 , da un lato d si riduce, ma dall’altro si ha una sorta di “transizione verso la condizione di corto circuito” (ossia, riducendo d , il valore della tensione all’interfaccia tende a diminuire, ad annullarsi).

In ambito di applicazione, conviene massimizzare o ridurre d ? Sostanzialmente, dipende dall’applicazione: l’onda evanescente si può considerare come le capacità parassite nell’elettronica: un termine “extra” che porta ad avere un campo elettrico di accoppiamento tra due termini. Ciò può essere voluto, o non voluto.

Rappresentazione grafica della legge di Snell

Si vuole a questo punto introdurre, per fissare meglio i concetti, una rappresentazione grafica alternativa della legge di Snell, che permette anche di avere un’idea sull’onda evanescente.

Precedentemente, si era parlato dello spazio \underline{k} , come di quello spazio per cui ciascun vettore identificava una particolare onda piana; ciò che si può fare, dunque, è, nello spazio xz che si sta utilizzando, disegnare una circonferenza (sarebbe una sfera ma, poichè $k_y = 0$, essa degenera in una circonferenza), dove ciascun vettore identifica una certa onda piana. Anche questa volta, è possibile considerare i due casi: n_1 maggiore o minore di n_2 .

Se $n_1 < n_2$, si ha qualcosa del genere:

Sulla sfera a sinistra si vede il \underline{k} dell’onda incidente, sull’altro quello dell’onda rifratta, dell’onda trasmessa; nel caso $n_1 < n_2$, a destra si ha una circonferenza più grande di quella che si aveva a sinistra. Ciò che si può fare è dunque, dato un valore di ξ , identificare ϑ_i , come angolo per cui ξ a partire dall’asse z interseca la circonferenza. Proiettando questo punto sull’altra circonferenza (in modo da preservare ξ), si può identificare immediatamente il \underline{k}_2 , ma anche il ϑ_t , e questo utilizzando banali considerazioni geometriche: questa è una rappresentazione grafica della legge di Snell.

Il secondo caso, quello per cui $n_1 > n_2$, è in effetti più interessante:

In questo caso, evidentemente, si vede che la circonferenza “destra” è più piccola. In questo caso, si hanno tre possibilità:

- la prima situazione è sostanzialmente duale a quella precedente: se si ha un angolo per cui ξ esiste sia su una circonferenza sia sull'altra, significa che siamo ad un angolo inferiore rispetto all'angolo critico: in questo caso, aldilà del fatto che $\vartheta_t > \vartheta_i$, tutto è perfettamente uguale a prima;
- quando si arriva in una situazione tale per cui ξ viene a coincidere con il raggio della circonferenza a destra, pari a $k_0 n_2$, siamo nella situazione “critica”, la situazione di transizione: in questo caso, si ha a che fare con $\vartheta_i = \vartheta_c$;
- se l'angolo è maggiore dell'angolo critico, l'intersezione non sembrerebbe più esserci.

Approfondiamo l'ultimo punto, usando un piccolo paragone algebrico: quando si ha a che fare con un'equazione di secondo grado, si sa che esiste un parametro, noto come determinante (o discriminante), Δ , che fornisce indicazioni di massima sulla soluzione; se $\Delta > 0$, si ha a che fare con due soluzioni reali distinte; se $\Delta = 0$, le due soluzioni degenerano in una singola soluzione; se $\Delta < 0$, si vuol dire che le soluzioni “non esistono”, che “non ci sono intersezioni con l'asse”. Questa cosa è assolutamente collegata con ciò che succede qua: le soluzioni, in realtà, esistono, ma non sono nel dominio reale, bensì in quello complesso. In questo caso, infatti, si può dire che:

$$k_{z2} = \sqrt{k_0^2 n_2^2 - \xi^2}$$

questo, se è immaginario, deve essere scritto come:

$$= -j\sqrt{\xi^2 - k_0^2 n_2^2}$$

Si può vedere che la radice è un'iperbole: si ha un andamento a iperbole.

Tutto ciò sembra molto complicato, ma in realtà c'è una cosa molto semplice, che ci viene ad aiutare: il k_{z2} o è reale, o è immaginario: non può essere complesso. Questa cosa è estremamente interessante, dal momento che ci permette di effettuare una costruzione geometrica “intuitiva”, o quantomeno visibile. Quello che si potrebbe fare è considerare la curva come variabile nel piano uscente dal foglio, perpendicolare al foglio, dal momento che è ragionevole pensare che l'asse immaginario sia rappresentabile come un

piano ortogonale a z . Per avere infine un'idea più visiva, si può prendere questa iperbole disegnata uscente dal foglio, e "ruotarla", sovrapponendo (in realtà non c'è sovrapposizione) l'asse immaginario a quello reale:

in questo modo si può vedere che effettivamente un'intersezione c'è, ma non è reale.

Volendo rappresentare il coefficiente di riflessione, un altro grafico è quello modulo-fase:

Il modulo sale fino all'angolo critico, e resta bloccato fino a $\pi/2$; per quanto riguarda la fase, invece, essa rimane bloccata a 0, fino a quando poi non inizia a crescere, fino a $\pi/2$ (come tra l'altro si può vedere banalmente, studiando la carta di Smith).

2.4.3 Coefficienti di Fresnel - onde TM

Per il caso TE, sono già state completate tutte le analisi; a questo punto si vogliono proporre e discutere i risultati concernenti il caso di onde TM, in modo da completare le nozioni teoriche su questo tipo di problema. Come si vedrà tra breve, l'espressione del coefficiente di riflessione nel caso TM, Γ^{TM} , sarà un poco diversa da quella del caso TE, cosa che comporterà implicazioni interessanti.

Si consideri dunque il solito problema, con però un'onda incidente puramente TM:

Esattamente come prima, si ha:

$$\xi = k_0 n_1 \sin \vartheta_i$$

e

$$k_z = \sqrt{\xi^2 - k^2}$$

si ha però una sostanziale differenza:

$$Z_\infty^{\text{TM}} = \frac{k_{z1}}{\omega \varepsilon_1}$$

Questo fa sì che, calcolando il Γ , si abbia:

$$\Gamma_{A^-} = \frac{n_1 \sqrt{n_2^2 - n_1^2 \sin^2 \vartheta_i} - n_2^2 \cos \vartheta_i}{n_1 \sqrt{n_2^2 - n_1^2 \sin^2 \vartheta_i} + n_2^2 \cos \vartheta_i}$$

Valutiamo a questo punto il comportamento di questa espressione al variare dell'angolo di incidenza ϑ_i :

- per $\vartheta_i = 0$, si ha, come peraltro anche nel caso TE:

$$\Gamma_{A^-} = \frac{n_1 - n_2}{n_1 + n_2}$$

nel caso si consideri la casistica $n_1 < n_2$, ossia quella per cui **non** si ha la riflessione totale, si vede che il coefficiente di riflessione è, per incidenza normale ($\vartheta_i = 0$), minore di zero: negativo;

- se $\vartheta_i \rightarrow \frac{\pi}{2}$, si ha che i coseni vanno a zero, quindi si ha che

$$\Gamma_{A^-} \rightarrow +1$$

A questo punto è necessario introdurre un certo numero di discorsi, al fine di chiarificare alcuni concetti. Nel caso $n_1 < n_2$, esattamente come prima, si ha che la radice non può avere radicando negativo, dunque sicuramente i vari numeri saranno tutti reali (dunque sull'asse reale della carta di Smith); d'altra parte, se si passa da numeri negativi a +1, ci sarà sicuramente un punto di attraversamento dello zero, ossia un punto per cui il coefficiente di riflessione è nullo; dal momento che, ora come ora, la nostra unica variabile indipendente è quella angolare, questo “punto” sarà un angolo di incidenza, un certo ϑ_i ; si chiama dunque “angolo di Brewster” quell'angolo per cui il coefficiente di riflessione si annulla; annullando il coefficiente di riflessione e facendo alcuni calcoli algebrici, si può dimostrare agevolmente che:

$$\vartheta_B = \arctan \left(\frac{n_2}{n_1} \right)$$

Cosa significa ciò? Come mai è possibile questa cosa, sebbene vi sia una discontinuità? Beh, si ricordi sempre che, all'origine di tutto, si ha il modello secondo le linee di trasmissione, ossia a partire dalle impedenze:

$$\Gamma_{A^-} = \frac{Z_{\infty 2} - Z_{\infty 1}}{Z_{\infty 2} + Z_{\infty 1}}$$

Il fatto che si abbia una discontinuità nei materiali ma comunque coefficiente di riflessione nullo significa che la discontinuità geometrica effettivamente c'è, ma anche che non c'è una vera e propria discontinuità elettrica: il mezzo è elettricamente omogeneo. Sotto il punto di vista di un'onda con polarizzazione TM, dunque, è come se non ci fosse discontinuità.

Discutiamo a questo punto un altro aspetto molto interessante: le due formule, per $\vartheta_i = 0$, sono identiche. Come mai? Si deve osservare che le due configurazioni, in questo caso, presentano una certa simmetria rispetto alla discontinuità:

I due campi sono **ruotati**, cosa che non coincide assolutamente con **sfasati**, di 90° . Il termine “sfasamento” è infatti collegato con il concetto di “fase”, il quale è collegato a quello che nella teoria dei circuiti è il fondamentale fasore, o “vettore rotante”: $e^{j\omega t}$. Questa cosa è importante, in teoria dei circuiti, dal momento che il piano complesso è una corrispondenza biunivoca con la retta reale (considerata, in questo momento, come un dominio di tipo temporale): il dominio reale di una dimensione temporale e quello dei fasori di una dimensione (il “piano complesso” in realtà è un dominio monodimensionale, per quanto rappresentabile mediante piano di Gauss); in questo “piano complesso” posso disegnare dei vettori, che al passare del tempo ruotano su questo piano, variando la propria **fase**.

Il termine “sfasamento” va sostanzialmente collegato solo al dominio del tempo, mentre la “rotazione” dei vettori nel paragone tra TE e TM è una rotazione nel dominio spaziale: per questo motivo il termine “sfasamento” è improprio: sfasamento indica un ritardo, almeno in ambito di teoria dei circuiti.

Si noti inoltre che il piano su cui giacciono i campi elettrico e magnetico è un piano “fisico”, un piano che come dimensioni ha dimensioni spaziali; il “piano di Gauss” che si utilizza per rappresentare nei fasori è un’astrazione geometrica che permette di semplificare la visualizzazione dei calcoli, dandone un significato: nell’ambito dei fasori, la rotazione nel piano complesso, la “variazione di fase”, è solo un metodo furbo di trattare i ritardi.

Il fatto che si abbia lo stesso coefficiente di riflessione in quella situazione, si può ricavare, almeno sotto un punto di vista qualitativo, mediante una semplice osservazione della giunzione: la giunzione infatti è isotropa, dunque non vi è una direzione preferenziale per l’orientamento dei campi; in altre parole, qualsiasi sia l’orientamento dei campi, la giunzione si comporta sempre alla stessa maniera (cosa che non capiterebbe, per esempio, in un cristallo, o in una lente polarizzata).

Come mai, nonostante ci sia questa simmetria per questo valore di angolo, dopo non vi sia più? Come mai in altre parole si ha la differenziazione tra i due coefficienti? La risposta qui è abbastanza semplice: se prima, per $\vartheta_i = 0$, c’era una simmetria, non si può dire altrettanto per il “dopo”: quello che si fa, di fatto, è, variando l’angolo, “rompere la simmetria”: le equazioni di Maxwell sono invarianti se si scambiano campo elettrico e magnetico, μ con $-\varepsilon$: solo in questo caso si ha a che fare con equazioni invarianti, quindi per questo si ha la differenziazione del comportamento. Se si considerassero due mezzi magnetici, dunque con diverse μ , ma la stessa ε , si avrebbe a che fare con un angolo di Brewster per la polarizzazione TE, ma non per la TM; se si avessero inoltre alcuni parametri ben precisi, sarebbe possibile avere entrambi gli angoli di Brewster in entrambe le polarizzazioni, o anche in nessuna delle

due.

Delle varie casistiche, consideriamo ora l'altra: quella per cui, nel caso TE (e pure qua, vedremo), si aveva riflessione totale: $n_1 > n_2$. In questo caso, di nuovo, si ha:

$$k_{z2} = -j |k_{z2}|$$

In questo caso, il coefficiente di riflessione, utilizzando questo metodo per la rappresentazione della costante di propagazione longitudinale, diventa:

$$\Gamma_{A^-} = \frac{-jn_1 \sqrt{n_1^2 \sin^2 \vartheta_i - n_2^2} - n_2 \cos \vartheta_i}{-jn_1 \sqrt{n_1^2 \sin^2 \vartheta_i - n_2^2} + n_2 \cos \vartheta_i}$$

Nel piano complesso, questo significa avere qualcosa di questo genere:

si ha che il numeratore ha parte reale negativa, parte immaginaria negativa; il denominatore ha ancora parte immaginaria negativa, ma ora parte reale positiva; i due numeri dunque hanno lo stesso modulo: si può dire che

$$|\Gamma_{A^-}| = 1$$

Cosa si può dire riguardo la fase? Beh, si può dimostrare, facendo l'arcotangente della parte reale e di quella immaginaria, che⁷:

$$\angle \Gamma_{A^-} = -\pi + 2 \arctan \left(\frac{n_1 \sqrt{n_1^2 \sin^2 \vartheta_i - n_2^2}}{n_2 \cos \vartheta_i} \right)$$

Questa cosa si può anche visualizzare sulla carta di Smith:

All'inizio si hanno dei Γ reali e positivi, quindi aumentando l'angolo si "va indietro", fino a raggiungere l'angolo critico; fino a questo punto, si ha fase o positiva o negativa, dunque non si hanno situazioni intermedie. La situazione "intermedia", per cui la fase varia, si ha per $\vartheta_i > \vartheta_c$, situazione in cui la fase per l'appunto "ruota", fino ad arrivare a 0: $\Gamma_{A^-} = +1$.

Volendo disegnare il diagramma d'onda stazionaria per questa situazione, si ha ciò:

Questo risultato si ottiene prendendo la carta di Smith con un certo angolo ϑ_i fissato ma tale da essere superiore all'angolo critico: girando T.G. (Toward Generator), verso quindi orario sulla carta di Smith, si vede che prima si raggiunge un punto di minimo, o addirittura in questo caso di 0 (essendo il modulo del Γ unitario), quindi un massimo, un altro zero e così via. Si vede che ora, a differenza del caso TE, all'interfaccia si ha un punto angoloso,

⁷in alcuni testi si ha un'espressione con la sola arcotangente; questa si ottiene definendo un coefficiente di riflessione "di corrente", definito dunque a partire dal campo magnetico

ossia un punto per cui si ha una discontinuità della derivata della funzione della tensione; per il resto, a destra, il solito esponenziale decrescente.

Nel caso TM si ha dunque qualcosa di più “spigoloso”; c’è modo di determinare ciò a priori? In realtà sì: ricordiamo che tutti i modelli che stiamo studiando partono sostanzialmente da modelli basati sulle linee di trasmissione; l’equazione delle linee di trasmissione ha un’espressione del tipo:

$$-\frac{dV}{dz} = jk_z Z_\infty^{\text{TE}} I = j\omega\mu I$$

volendo analizzare questa espressione, si vede che dipende da μ , la quale è costante ovunque; dal momento che inoltre la sezione A è una discontinuità di impedenza caratteristica, si ha la certezza che la corrente sia continua; dal momento che, per il caso TE:

$$Z = \frac{\omega\mu}{k_z}$$

la derivata della tensione è dunque una funzione continua, quindi si ha la certezza che essa sia derivabile, “liscia”. Non si può dire purtroppo altrettanto sulla derivata seconda, dal momento che vale la seguente equazione (di partenza):

$$\frac{d^2V}{dz^2} + k_z^2 V = 0$$

in questo caso, le k_z a sinistra e a destra sono diverse, dunque la derivata seconda sarà certamente discontinua. Dal momento che nel caso TM invece la dipendenza è da $\omega\varepsilon$, dal momento che il nostro problema è sostanzialmente basato su una discontinuità di costante dielettrica relativa, si ha per forza il fatto che la derivata della tensione sia discontinua (per quanto ora comunque la corrente non lo sia, per lo stesso ragionamento di prima).

Volendo analizzare i grafici complessivi del modulo di $\Gamma_{A^-}^{\text{TE, TM}}$, nei casi rispettivi con e senza riflessione totale, si ha qualcosa di questo genere:

Dal grafico relativo alla presenza di riflessione totale è evidente il fatto che nasce un certo angolo critico, ossia un certo angolo a partire del quale si ha che i coefficienti di riflessione diventano unitari. Questo si può pensare come un fenomeno per cui si “spinge” in su la curva dei Γ , dei coefficienti di riflessione, schiacciandola un po’ a fisarmonica.

Dai grafici, è evidente che $\Gamma^{\text{TE}} > \Gamma^{\text{TM}}$, sempre, dal momento che il secondo deve passare per zero e, non essendoci cambi bruschi di concavità, è questo l’andamento. Nel caso di polarizzazione TE si parla anche di “pola-

rizzazione perpendicolare”, o “S”⁸, e di polarizzazione “P” o “parallela” per la TM.

Si vuol dire che, con gli occhiali a lenti polarizzate, si ha una “vista migliore”; cosa significa ciò? Beh, il problema in questione, sostanzialmente, è lo studio del rapporto segnale su interferente di un sistema: il segnale è la parte utile delle onde che intendiamo vedere, l’interferente è la luce del sole, sia in porzione diretta, sia in porzione riflessa. Si immagini la seguente situazione:

Si immagini di essere, durante una giornata soleggiata, al lago a guardare le barche; la fonte di illuminazione dei vari punti del sistema è sostanzialmente il Sole; quando dal lago arrivano i raggi contenenti l’informazione sulle barche, si vede sia la luce del Sole diretta, sia la luce del Sole riflessa, sia la luce della barca: la luce riflessa dal lago (componente importante, dal momento che l’acqua riflette bene), è un interferente rispetto alla luce della barca.

Si vuol dire che la luce del Sole non sia polarizzata; in realtà questa è un’affermazione inesatta, dal momento che di fatto la luce che arriva dal Sole ha un certo campo elettrico che oscilla, e ha sia una componente orizzontale sia una componente verticale; la frase deriva in realtà dal fatto che questa somma, dal momento che la luce del Sole è sostanzialmente rumore, ha “pesi aleatori”: si tratta di una funzione caotica, per la quale però si ha una particolarità: ciascuna componente ha la stessa densità di potenza (la luce è “bianca”); non si può di sicuro dire che essa sia ellittica, lineare o circolare, ma di sicuro il campo ha delle componenti nelle varie direzioni, dal momento che il fatto di avere una certa direzione è una proprietà dei vettori. L’unica cosa che ha è il fatto che la polarizzazione verticale e quella orizzontale hanno la stessa densità di potenza per unità di superficie, con però una relazione di fase aleatoria, ottenendo una “somma caotica”. Avendo tuttavia una lente polarizzatrice, ossia una lente che di tutta la luce fa passare solo quella polarizzata lungo una certa direzione, la luminosità non cambia: questo a testimonianza del fatto che la densità di potenza sia costante, in tutte le possibili polarizzazioni: il rumore si somma in potenza, non in ampiezza.

Consideriamo lo step successivo: ciò che ci disturba molto è la luce che viene riflessa dal lago. La superficie del lago porta ad avere un’incidenza quasi radente, con un coefficiente di riflessione molto grande, ottenendo un forte contributo di riflessione, con però una dominanza della polarizzazione TE: questa, dal momento che il coefficiente di riflessione TE è più grande di quello TM, portando il contributo riflesso ad essere predominante. Se dunque si hanno occhiali che attenuano la componente TE della luce rispetto alla TM di per esempio 20 dB, il rapporto segnale su interferente migliora

⁸dal tedesco Senkrecht

considerevolmente. La luce riflessa, dunque, è tendenzialmente quasi solo polarizzata TE.

L'angolo di Brewster è anche detto “angolo polarizzante”: questo, deriva da un metodo di realizzazione di un polarizzatore basato sul principio della riflessione.

Il seguente metodo è ora descritto: dati n_1 e n_2 , quindi un'onda piana incidente con contributi sia TE sia TM, se si fa in modo da avere $\vartheta_i = \vartheta_B$, si avrà che l'onda trasmessa avrà sia componenti TE, sia TM, ma l'onda riflessa solo componenti TE: il coefficiente di riflessione relativo alle onde TM è infatti nullo, all'angolo di Brewster. Questo permette, sapendo che la luce TE oscilla solo perpendicolarmente (sotto il punto di vista del campo elettrico), di ottenere un polarizzatore.

2.4.4 Dielettrici con perdite

L'ultimo caso di interesse è quello dei dielettrici con perdite. Si consideri, per il secondo dielettrico, la seguente casistica:

$$n_2 = n_2' - jn_2'' \quad n_2'' > 0$$

in questo caso, il mezzo 2 dissipa energia. Questa cosa in realtà non introduce nessuna modifica alla nostra teoria: aldilà dei risultati finali, la teoria è esattamente la stessa. Si ha:

$$k_{z2} = k_0 \sqrt{(n_2' - jn_2'')^2 - n_1^2 \sin^2 \vartheta_i} = \beta - j\alpha$$

Volendo analizzare il comportamento dei k_{z2} sul piano complesso, si hanno i seguenti andamenti (ovviamente, nel caso di riflessione totale):

Se non si hanno perdite, come detto in precedenza, $k_{z2} = k_0 n_2$, per $\vartheta_i = 0$; al crescere dell'angolo di incidenza, il radicando diventa sempre più piccolo, fino ad arrivare a un punto per cui il coefficiente k_{z2} si annulla; aumentando ulteriormente, esso diventa puramente immaginario; a forza di aumentare, si arriva al seguente caso limite:

$$-jk_0 \sqrt{n_1^2 - n_2^2}$$

si ha dunque una transizione “brusca” da “reale puro” a “immaginario puro”. Nell'altro caso, quello in cui si ha un dielettrico con perdite, perde di senso parlare di “angolo critico”, dal momento che si ha sempre un k_z complesso: esso, per angoli di incidenza ridotti, sarà tendenzialmente reale (ossia con parte reale prevalente sulla parte immaginaria), per angoli maggiori sarà

tendenzialmente immaginario (dal momento che prevarrà la parte immaginaria); la zona di transizione sarà ovviamente meno brusca, e non si avrà un valore vero e proprio per cui $k_{z2} = 0$. Il valore di costante di propagazione verso cui si tende, comunque, è sostanzialmente lo stesso di prima.

Questa cosa si riflette anche sul diagramma d'onda stazionario:

In questo caso non si ha più a che fare con zeri, ma solo con minimi, dal momento che il coefficiente di riflessione in questo caso non è più unitario, quindi non annulla completamente la tensione.

2.4.5 Rombo di Fresnel - Fasci gaussiani

Si vuole a questo punto concludere la sezione proponendo un esempio applicativo della riflessione a singola interfaccia, richiamando alcune nozioni e introducendo alcune precisazioni su alcuni concetti.

Si consideri un dispositivo di questo tipo:

Si consideri un'onda piana che arriva sulla faccia, inclinata in modo tale da avere un angolo di incidenza superiore a quello critico; si ha riflessione totale, quindi il raggio riflesso subisce un'ulteriore riflessione, ancora una volta ad angolo superiore a quello critico (il sistema va ovviamente progettato in modo tale da avere queste riflessioni effettivamente totali), quindi il raggio esce fuori dall'altra faccia. Questo tipo di dispositivo funziona da **polarizzatore**: le due polarizzazioni infatti subiscono storie di fatto indipendenti tra loro. Un'onda piana non è in generale o TE o TM: essa infatti può avere polarizzazione circolare, ellittica, o anche lineare, ma non è detto che anche nel caso lineare il vettore di campo elettrico sia puramente TE o puramente TM; quello che tuttavia si può fare, grazie alla linearità delle equazioni che studiamo, è scomporre questi vettori in componenti per l'appunto TE e TM, applicando dunque su ciascuna componente, la quale avrà in generale un certo valore di modulo e un certo valore di fase, le nozioni precedentemente presentate; quello che vale in generale è tuttavia il fatto che i coefficienti di riflessione, per le due componenti, saranno diversi. La riflessione cambia le caratteristiche di polarizzazione del campo riflesso, e ciò permette di realizzare, con questo dispositivo, un polarizzatore, dal momento che il dispositivo permette di discriminare i diversi contributi (TE e TM) dell'onda su di esso incidente. Essendoci inoltre due riflessioni, l'effetto sulla fase raddoppia, quindi quello che si può fare è cambiare la polarizzazione dell'onda incidente, per esempio partendo da un'onda lineare in ingresso e "uscendo" con un'onda a polarizzazione circolare.

Si noti una cosa molto importante: stiamo applicando il modello precedentemente elaborato, su una situazione che sembrerebbe, a prima vista, essere fuori dal suo range di validità: il dispositivo ha infatti una superficie

finita, mentre le onde piane hanno come superfici a fase costante per l'ap-punto dei piani, di dimensioni idealmente infiniti. Qui, invece, si incide da un lato su una superficie finita, dall'altro con onde non esattamente piane: come già detto infatti l'onda "piana" deve essere ricavata a partire da una qualche sorgente, per esempio un LASER; un LASER tuttavia non produce un'onda esattamente piana, dal momento che produce un "fascetto" di onde piane, nella fattispecie un fascio gaussiano (come si può sapere dai corsi di Optoelettronica).

Come si era detto precedentemente, la tensione è rappresentabile, spet-tralmente, come (considerando per esempio un'onda TE):

$$V^{\text{TE}}(\xi, z) = \int_{-\infty}^{+\infty} E_y(x, z) e^{+j\xi x} dx$$

e con la sua trasformazione inversa:

$$E_y(x, z) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} V(\xi, z) e^{-j\xi x} d\xi$$

Di tutte le possibili onde piane, ciò che è stato finora fatto, nella pre-sentazione del modello, è stato fissare un certo ξ , dunque "selezionare" una singola onda piana, e considerare solo questa; si aveva, come noto:

$$\xi = k_0 n_1 \sin \vartheta_i$$

Se come campo incidente si ha tuttavia un fascio gaussiano, non si avrà più un solo ξ , o una " $\delta(\xi)$ " come accennato precedentemente, bensì un insieme di ξ . Fascio gaussiano significa sostanzialmente che il campo sul piano ha un andamento gaussiano (si pensi all'analogia con le antenne, e il campo sull'apertura):

$$E_y(x, 0) = e^{-\frac{x^2}{2w_0}}$$

ossia, una gaussiana con varianza w_0 (w sta per *width*).

Si supponga, per semplicità, che il fascio sia a incidenza normale sul piano di discontinuità: i piani a fase costante dei fascetti sono tendenzialmente orientati normalmente al piano, dal momento che la funzione del campo è reale, dunque non interviene nella variazione di fase: questo significa che, lungo x , non si hanno variazioni di fase; in altre parole, i piani a fase costante sono tutti quelli normali a \hat{z} : i piani $z = \text{costante}$.

Volendo calcolare, a partire dal campo, la tensione, è sufficiente ricordare le proprietà della trasformata di Fourier, e vedere che:

$$V(\xi, 0) = E_0 L \sqrt{2\pi} e^{-\frac{\xi^2}{2} w_0^2}$$

ossia, si ha ancora una volta una gaussiana, in cui però la varianza è w_0^{-1} : il reciproco della varianza di prima. Questa osservazione deriva semplicemente dal “caso limite” del principio di indeterminazione: la gaussiana è infatti il tipo di funzione per cui si ha il minimo del principio di indeterminazione, ossia il caso in cui si ha che il prodotto delle larghezze nei due domini è uguale a una costante, e non “minore”.

Si supponga ora di avere un esempio del tipo: se $w_0 = 1$ mm, ricordando che le lunghezze d'onda λ_0 nell'ottica sono dell'ordine di grandezza del micron, si ha che $w_0 = 1000\lambda_0$: quindi 1 mm è effettivamente tantissimo, ossia assimilabile a infinito! Questo, dal momento che, come sempre in Elettromagnetismo, le dimensioni spaziali devono essere rapportate alla variazione di fase che si ha quando si percorre una certa distanza, ossia a k_0 ; ricordando che, ovviamente,

$$k = \frac{2\pi}{\lambda_0} n_1$$

con lunghezze d'onda di quel genere possiamo “stare tranquilli”.

Cosa significa ciò? Beh, semplicemente, che, nel dominio reciproco, k_0 è “molto più avanti”: dal momento che si ha una gaussiana, dunque, di “onde piane” se ne han tante, dal momento che tanti sono i valori di ξ di cui si dispone; tuttavia, la variazione dei valori di ξ è estremamente piccola se rapportata al k_0 . Per ogni ξ , come ben noto, si ha a che fare con un certo angolo, dunque con un certo valore di Γ , di coefficiente di riflessione; se tuttavia, sebbene la variazione di ξ sia non nulla, essa è totalmente trascurabile, quindi si può considerare il coefficiente di riflessione dell'onda sostanzialmente costante, dal momento che tutto il fascio occuperà, a stima, qualche decimo di grado (non è detto che sia così ma è per dare un ordine di grandezza): $\Gamma(\vartheta_i)$ è dunque circa costante.

In altre parole: stiamo lavorando con fasci gaussiani, ma, di fatto, essi sono assimilabili, per i nostri calcoli, a onde piane: il dispositivo si può dunque progettare a partire da questa osservazione. Questo è un esempio che permette di capire quale sia la potenza dello strumento analitico rappresentato dalle onde piane: il suo campo di applicazione è enorme, anche in problemi che sembrano molto complicati rispetto alla semplicità del modello che vi si vorrebbe applicare (e che in effetti, come appena visto, si può applicare in molte situazioni).

2.5 Struttura a tre mezzi materiali

A partire dalle nozioni precedentemente acquisite, si vuole a questo punto studiare una struttura più complicata della precedente, ma anche spesso utilizzabile in situazioni di vario tipo, come base per il progetto di dispositivi ottici. Si avrà dunque a che fare, ora, con una struttura di questo genere:

Si hanno tre mezzi dielettrici, con quindi due discontinuità. I dati per il problema sono k_{zi} , $Z_{\infty i}$, per le polarizzazioni TE e TM (di fatto, l'unica differenza tra le due sono i valori numerici delle Z_{∞}).

Un obiettivo del nostro studio potrebbe essere quello di determinare il coefficiente di riflessione e quello di trasmissione, quindi i vari grafici dei diagrammi d'onda. Prima di tutto, come si trova Γ_{A-} , ossia il coefficiente di riflessione visto da prima della prima discontinuità? Beh, sostanzialmente, esistono alcuni metodi:

- utilizzare un approccio “diretto”, “classico”, basato sull'analisi delle linee di trasmissione: noti i vari parametri delle linee di trasmissione, fare il calcolo e trovare dunque il coefficiente di riflessione;
- descrivere questa struttura come la cascata di due strutture, considerando un approccio di tipo matriciale, dove ciascuna matrice descrive una giunzione.

Il metodo classico presenta un sostanziale vantaggio sul secondo metodo, che sarà quello che andremo effettivamente a utilizzare nell'analisi: esso permette di determinare il comportamento “interno” del circuito: il metodo matriciale infatti considera il circuito come una “scatola nera”, guardata dall'esterno, della quale possiamo esclusivamente conoscere le uscite come funzione degli ingressi, senza interessarci di cosa capitano dentro; con il metodo classico, invece, è possibile determinare lo stato elettrico, dunque conoscere l'andamento del campo elettromagnetico anche internamente al sistema.

Come noto, la matrice scattering si definisce a partire dal seguente sistema di equazioni lineari:

$$\begin{cases} b_1 = S_{11}a_1 + S_{12}a_2 \\ b_2 = S_{21}a_1 + S_{22}a_2 \end{cases}$$

Dove le varie onde a_i sono le onde incidenti nel doppio bipolo, le b_i quelle scatterate da esso. Esse sono definite come:

$$a_1 = \frac{V_1^+}{\sqrt{Z_{r1}}}$$

$$b_1 = \frac{V_1^-}{\sqrt{Z_{r1}}}$$

Si tratta in sostanza di onde progressive e regressive normalizzate in potenza, in modo da far sostanzialmente scomparire la dipendenza dalle impedenze di riferimento (le impedenze rispetto cui si hanno i parametri scattering).

Come già detto, l'obiettivo è quello di studiare una struttura composta da due doppi bipoli in cascata. Esistono delle formule che permettono di trattare le due giunzioni come due "scatole", da mettere insieme; la condizione però per cui valgono, è quella di avere le impedenze di riferimento delle porte tra loro connesse uguale; questa è una condizione relativamente poco restrittiva, che permette tuttavia di avere formule semplificate rispetto a una casistica più generale. Per la struttura globale, si ha che:

$$S_{11} = S'_{11} + \frac{S'_{12}S'_{21}S''_{11}}{1 - S''_{11}S'_{22}}$$

Dove i coefficienti con un apice sono quelli relativi alla giunzione "sinistra". quelli a due apici alla giunzione "destra".

Per quanto semplificata, questa legge di connessione di strutture è ancora complicata. Quando si vogliono collegare tra loro strutture, tuttavia, la matrice \underline{S} non è l'unica, e neanche la più indicata: si può utilizzare la **matrice di trasmissione**. Essa sostanzialmente usa le stesse onde di potenza comunemente utilizzate per la matrice scattering, tuttavia "cambiandole di nome":

$$c_1^+ = a_1 \quad c_1^- = b_1 \quad c_2^+ = b_2 \quad c_2^- = a_2$$

A questo punto si devono scegliere delle variabili indipendenti e delle variabili dipendenti; una scelta (non è l'unica, come si vedrà), è:

$$\begin{bmatrix} c_1^+ \\ c_1^- \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} c_2^+ \\ c_2^- \end{bmatrix}$$

In questo tipo di matrice, in ogni configurazione, si "isolano", come variabili indipendenti, o le due rappresentanti lo stato elettrico all'ingresso, o le due rappresentanti lo stato elettrico all'uscita; nella scelta ora effettuata, si dice che "lo stato elettrico all'ingresso è uguale alla matrice di trasmissione, moltiplicata per lo stato elettrico all'uscita".

Questo tipo di rappresentazione è molto interessante dal momento che, dati i dispositivi in cascata, rappresentati mediante le relative matrici \underline{T}' e \underline{T}'' , la matrice equivalente di trasmissione sarà:

$$\underline{\underline{T}} = \underline{\underline{T'}} \underline{\underline{T''}}$$

dove per prodotto si intende il tradizionale prodotto matriciale riga per colonna.

Come mai è stata scelta questa definizione della matrice di trasmissione? Essa sembra “poco intuitiva” rispetto al nome, dal momento che, quando si parla di “trasmissione”, sembrerebbe più intuitivo avere, come variabili dipendenti, le variabili di stato dell’uscita; è stata tuttavia fatta questa scelta, dal momento che facendo così il prodotto delle matrici va fatto con le matrici dei dispositivi “da sinistra a destra”, ossia considerando il prodotto delle matrici da quella rappresentante il dispositivo più a sinistra verso quello più a destra, invece che al contrario, come sarebbe se si usasse la notazione appena presentata come “più intuitiva”; è sostanzialmente comunque solo questione di gusti. Considerando una di queste matrici per ogni interfaccia, è possibile calcolare con maggiore semplicità la matrice risultante da tutte queste.

Si ricordi di tenere bene a mente quali sono i nostri obiettivi: noi siamo interessati a Γ o al coefficiente di trasmissione, non agli elementi della matrice di trasmissione: ove possibile, è preferibile utilizzare la rappresentazione mediante parametri scattering, dal momento che essi sono più vicini alla fisica del dispositivo, e dal momento che essi sono misurabili (mediante un analizzatore di reti per esempio).

Esiste in realtà un terzo metodo, basato sulle “matrici catena” o “matrici $ABCD$ ” (o, anche in questo caso, dette “matrici di trasmissione”): si tratta di matrici simili a quelle di prima, in cui però invece che rappresentare lo stato elettrico mediante onde di potenza, lo si fa con tensioni totali, che generalmente in questo caso non sono osservabili e quindi poco interessanti.

Incominciamo a questo punto l’analisi della struttura, utilizzando un formalismo basato sulle matrici **scattering**. Si consideri $\underline{\underline{S'}}$ la matrice scattering relativa alla discontinuità tra A^- e A^+ : si può dimostrare agevolmente (analizzando semplicemente una discontinuità di impedenza) che:

$$\underline{\underline{S'}} = \begin{bmatrix} r_{12} & (1 + r_{12}) \sqrt{\frac{Z_{\infty 1}}{Z_{\infty 2}}} \\ (1 + r_{12}) \sqrt{\frac{Z_{\infty 1}}{Z_{\infty 2}}} & -r_{12} \end{bmatrix}$$

dove:

$$r_{12} = \frac{Z_{\infty 2} - Z_{\infty 1}}{Z_{\infty 2} + Z_{\infty 1}}$$

In questo caso, le impedenze di riferimento della struttura sono:

$$Z_{r1} = Z_{\infty1} \quad Z_{r2} = Z_{\infty2}$$

Ovviamente, le impedenze caratteristiche vanno calcolate, a seconda che si voglia aver a che fare con onde TE o TM, usando le formule precedentemente presentate.

r_{12} è un coefficiente di riflessione di Fresnel; esso ha sostanzialmente il seguente significato: data la struttura complessiva, si “estragga” la giunzione tra i due mezzi n_1 e n_2 , e la si consideri **isolata rispetto al resto del circuito**; in questo caso, se si fa ciò, è sostanzialmente come non avere la discontinuità nella sezione B, quindi il problema è identico a quello precedentemente analizzato, e possono essere utilizzati tutti i risultati precedentemente ricavati.

La stessa cosa può essere fatta per quanto riguarda la giunzione destra, ossia quella sulla sezione B; essa viene trattata nella stessa maniera della prima, ossia non considerando la giunzione in A, e dunque riconducendoci ancora una volta al solo problema di riflessione da singola discontinuità di mezzo dielettrico. In questo caso, si avrà qualcosa di molto simile a prima, dunque:

$$\underline{\underline{S''}} = \begin{bmatrix} r_{23} & (1 + r_{23}) \sqrt{\frac{Z_{\infty2}}{Z_{\infty3}}} \\ (1 + r_{23}) \sqrt{\frac{Z_{\infty2}}{Z_{\infty3}}} & -r_{23} \end{bmatrix}$$

$$Z_{r1} = Z_{\infty2} \quad Z_{r2} = Z_{\infty3}$$

Queste formule, di per sè, sono giuste. Non si sta tuttavia tenendo conto di un fatto: tra le due giunzioni, c'è una “linea”, ossia c'è una porzione di mezzo con coefficiente di rifrazione n_2 , di lunghezza d . Questa va tenuta in conto, considerando uno spostamento dei piani di riferimento su una delle due matrici. Scegliamo dunque di modificare $\underline{\underline{S''}}$, e, per “ricordare mnemonicamente”⁹ il procedimento, si ricordi che S''_{11} e S''_{22} sono coefficienti di riflessione, mentre S''_{12} e S''_{21} sono coefficienti di trasmissione; in altre parole, pensando alla matrice scattering come a una matrice in cui i vari coefficienti rappresentano dei collegamenti, si può intuitivamente dire che:

- S''_{11} collega la prima porta con la prima porta stessa, quindi si ha un doppio contributo di sfasamento ϕ : $e^{-j2\phi}$
- S''_{12} e S''_{21} collegano la prima e la seconda porta in un verso e in un altro; dal momento che solo la prima porta deve essere spostata di riferimento, si ha un contributo singolo di riferimento su ciascuna: $e^{-j\phi}$;

⁹per una dimostrazione più formale, Orta R. - Teoria delle linee di trasmissione

- S''_{22} collega la seconda porta con sè stessa; dal momento che la seconda porta non subisce spostamenti di riferimento, il coefficiente rimarrà uguale.

La vera matrice scattering da implementare sarà dunque la seguente:

$$\underline{\underline{S''}} = \begin{bmatrix} r_{23}e^{-j2\phi} & (1+r_{23})\sqrt{\frac{Z_{\infty 2}}{Z_{\infty 3}}}e^{-j\phi} \\ (1+r_{23})\sqrt{\frac{Z_{\infty 2}}{Z_{\infty 3}}}e^{-j\phi} & -r_{23} \end{bmatrix}$$

$$Z_{r1} = Z_{\infty 2} \quad Z_{r2} = Z_{\infty 3}$$

e dove

$$\phi = k_{z2}d = k_0d\sqrt{n_2^2 - n_1^2 \sin^2 \vartheta_i}$$

Elaborando dunque le espressioni, si può, facendo qualche conto, trovare che, per il blocco complessivo¹⁰ (i due blocchi in cascata):

$$S_{11} = \frac{r_{12} + r_{23}e^{-j2\phi}}{1 + r_{12}r_{23}e^{-j2\phi}}$$

$$S_{21} = \sqrt{\frac{Z_{\infty 1}}{Z_{\infty 3}}} \frac{(1+r_{12})(1+r_{23})e^{-j\phi}}{1 + r_{12}r_{23}e^{-j2\phi}}$$

Queste formule, si osservi, valgono tanto per le polarizzazioni TE quanto per quelle TM, a patto di avere le relative Z_{∞} .

Si osservi una cosa: queste due espressioni hanno lo stesso denominatore; le matrici scattering sono sostanzialmente delle matrici contenenti, volendo utilizzare un approccio *controllistico*, diverse funzioni di trasferimento dello stesso sistema. Un sistema, come noto, è caratterizzato dai poli della sua funzione di trasferimento, elementi che sono caratteristici del sistema: in ogni funzione di trasferimento del sistema, a partire da un certo ingresso e verso una certa uscita, si avranno sempre gli stessi poli.

Al fine di comprendere meglio il senso di questa affermazione, nonchè il funzionamento del sistema, verranno ora considerate due casistiche, con però un elemento in comune: $n_1 = n_3$ (ciò permette di avere delle espressioni semplici ma dalle quali comunque è possibile avere osservazioni molto interessanti).

¹⁰per dimostrare la formula di S_{11} conviene dimostrare che il determinante di $\underline{\underline{S'}}$ è unitario, facendo tutti i conti, e così l'espressione si semplifica notevolmente

2.5.1 Casistica 1: $n_1 = n_3 < n_2$

Questo è il caso tipico per cui, per esempio, i mezzi 1 e 3 sono aria, il mezzo 2 è una lastra di vetro: si tratta dunque di una lastra di vetro di spessore d immersa in aria. Come già visto:

$$\phi = k_0 d \sqrt{n_2^2 - n_1^2 \sin^2 \vartheta_i}$$

In questa condizione si può dimostrare che:

$$r_{23} = -r_{12}$$

In questo caso, dunque, si ha una semplificazione delle espressioni (che quindi perdono la loro generalità):

$$S_{11} = \frac{r_{12}(1 - e^{-j2\phi})}{1 - r_{12}^2 e^{-j2\phi}} \quad S_{21} = \frac{(1 - r_{12}^2 e^{-j\phi})}{1 - r_{12}^2 e^{-2j\phi}}$$

Il numeratore di S_{11} si annulla per:

$$1 - e^{-j2\phi} = 0 \implies \phi = n\pi$$

ϕ è un parametro che ha diverse variabili: k_0 (dunque la frequenza), ϑ_i , gli indici di rifrazione, lo spessore d ; un altro modo di dire ciò, è dire che, in questa condizione, data la lunghezza d'onda guidata:

$$\lambda_{g2} = \frac{2\pi}{k_{z2}}$$

che

$$d = n \frac{\lambda_{g2}}{2}$$

In altre parole, quando la lastra ha uno spessore pari a un multiplo di $\frac{\lambda_{g2}}{2}$, è come se la lastra non ci fosse! In altre parole, essendo $n_1 = n_3$, è come avere un mezzo omogeneo (sotto il punto di vista elettrico): come non avere discontinuità. Nel caso più generale per cui $n_1 \neq n_3$, se la medesima condizione è soddisfatta, il risultato è comunque simile: il mezzo 2 è elettricamente invisibile, ossia è come avere solamente una discontinuità tra i mezzi 1 e 3. Questo in realtà è fisicamente ovvio: abbiamo annullato il numeratore di S_{11} , che, come noto, è il coefficiente di riflessione all'ingresso della struttura.

Vi è, nel caso di polarizzazione TM, un altro caso per cui il coefficiente di riflessione della struttura è nullo: l'incidenza secondo angolo di Brewster. Questo fatto in realtà sembra molto poco intuitivo, ma quello che capita è che il raggio trasmesso dalla prima interfaccia incide sulla seconda interfaccia

ancora una volta con angolo di Brewster; **anche** in questo caso si ha a che fare con un coefficiente di riflessione all'ingresso nullo, ma questa casistica **non è di nostro interesse**, per un motivo che vedremo tra breve.

Consideriamo il caso per cui $S_{11} = 0$, senza andare a parlare dell'angolo di Brewster: abbiamo parlato di S_{11} , ma cosa si può dire di S_{21} ? Beh, se si fa variare la frequenza, il modulo del numeratore è costante, mentre quello del denominatore no, dal momento che il numeratore è dato da una parte reale interamente moltiplicata per un termine di fase, mentre il denominatore è dato dalla somma di una parte reale e di una immaginaria, e quella immaginaria è variabile con la frequenza. Si ha sostanzialmente qualcosa di questo genere:

Generalmente, $|r_{12}| < 1$, quindi il numeratore non si annullerà; il valore massimo del denominatore sarà dunque pari a $1 + r_{12}^2$, quello minimo a $1 - r_{12}^2$; nella seconda situazione, quella in cui il valore è minimo, si ha il massimo del coefficiente di trasmissione: $|S_{12}| = 1$ (dal momento che numeratore e denominatore sono uguali, con fase diversa). Si hanno i seguenti grafici:

I massimi di un grafico coincidono con i minimi dell'altro; questo è ragionevole dal momento che la potenza si conserva, dunque, della potenza incidente, parte viene riflessa, parte viene trasmessa; quando la trasmissione sarà massima, è ovvio che la riflessione sarà minima, e viceversa. Nel caso della riflessione, sono possibili come già detto situazioni in cui essa possa annullarsi, dunque i minimi del coefficiente $|S_{11}|$ saranno in realtà degli zeri, mentre quelli di $|S_{21}|$ saranno sempre e comunque dei minimi.

A questo punto, prendiamo una pausa, al fine di analizzare con critica tutto ciò che abbiamo appena studiato. Finora sono stati considerati modelli matematici, ma essi sono sensati? Hanno riscontro nella pratica?

Consideriamo, per rispondere a questa domanda, un caso "pratico": si consideri di avere una discontinuità, su cui si fa incidere un'onda; parte della potenza sarà riflessa, parte della potenza sarà trasmessa. Se si ha un forte salto di indice di rifrazione, quindi una forte riflessione, dentro alla seconda interfaccia passa poca potenza: il coefficiente di trasmissione, supponendo che ci sia una trasmissione alla prima interfaccia di potenza pari a $1/100$, sarà $T \sim 0,1$ (dal momento che il coefficiente di trasmissione è una grandezza lineare, la potenza è quadratica). Se nella seconda interfaccia si ha lo stesso T , si potrebbe avere una potenza pari a $1/10^{-4}$ di quella incisa; il modello, tuttavia, sembra suggerire che esistono particolari situazioni in cui **l'intera** potenza viene trasmessa dal primo al terzo mezzo; come mai ciò?

In realtà nella considerazione appena fatta non si tiene conto di ciò che capita tra le due interfacce, ossia nel mezzo 2:

Si immagini di incidere un'onda a sinistra della prima interfaccia: una piccola parte dell'onda passa, quindi cerca di andare nel terzo mezzo, ma non

riesce di nuovo: quello che capita, dunque, è che nel secondo mezzo **si ha una riflessione**, ossia si ha un'onda che man mano continua ad andare avanti e indietro, dal momento che i coefficienti di riflessione **anche all'interno** della “scatola”, tra le due interfacce, sono molto elevati. C'è da dire un'altra cosa: è vero che l'onda un poco “sgocciola” ogni volta che viene riflessa, ha ossia delle perdite, ma non stiamo analizzando una situazione in cui si manda un impulso e si stoppa il generatore: l'onda sta incidendo “CW”, Continuous Wave, ossia il generatore è “sempre attivo”: l'onda che mandiamo dentro la struttura non si ferma, dunque continua “ad accumularsi” al suo interno.

Se lo sfasamento ϕ è uguale a π , esiste un guadagno di anello (volendo sempre usare termini controllistici), legato alla propagazione in un verso, al Γ_B^{\rightarrow} (ossia al Γ nel punto B guardato da sinistra verso destra), all'altro sfasamento (l'onda dopo essere arrivata “torna indietro”), quindi al Γ_{A+}^{\leftarrow} , ossia al Γ guardato da destra verso sinistra:

$$G_{\text{anello}} = e^{-j\phi} \Gamma_B^{\rightarrow} e^{-j\phi} \Gamma_A^{\leftarrow}$$

Questo, si può però anche scrivere utilizzando i coefficienti delle matrici: “guardare” da una parte o dall'altra significa semplicemente considerare una o l'altra porta delle due matrici:

$$G_{\text{anello}} = S''_{11} S'_{22} e^{-j2\phi}$$

Come noto, il denominatore delle funzioni di trasferimento è sempre $1 - G_{\text{anello}}$: esattamente ciò che abbiamo trovato noi in altro modo. Si osservi che il guadagno di anello, se $\phi = \pi$, ha fase 2π più le fasi dei due coefficienti di riflessione; al nostro punto, tuttavia, essi sono reali, quindi o hanno fase nulla, o hanno fase pari a π : sommando le varie fasi, si ottiene sempre o fase pari a 2π , o pari a 0 , dunque equivalenti. Se il segnale fa un giro completo tra le due interfacce, si somma in fase con quello che arriva da sinistra, ottenendo quindi la condizione di “trasparenza” del mezzo 2: i vari contributi parziali fanno sì che si abbia un trasferimento completo di potenza verso sinistra. Tutto ciò, naturalmente, vale perchè stiamo parlando di sinusoidi, dunque perchè abbiamo una ben precisa ipotesi: quella di **monocromaticità**. Si noti che questo è un sistema che non può oscillare: dal momento che i coefficienti di riflessione sono minori di 1, allora anche il guadagno di anello sarà minore di 1, quindi non si hanno le condizioni per soddisfare il criterio di Bode.

Un commento aggiuntivo: questo tipo di risposta in frequenza potrebbe (almeno, in via di principio), essere utilizzata per la realizzazione di filtri; il valore minimo, infatti, è:

$$\frac{1 - r_{12}^2}{1 + r_{12}^2}$$

ma r_{12} si progetta quando si progettano i materiali: è sufficiente utilizzare il grafico dei Γ nelle polarizzazioni TE o TM e quindi scegliere cosa fare. Volendo per esempio un salto molto forte di indice di rifrazione n , si avrebbe un Γ molto elevato, quindi un $1 - |r_{12}|^2$ circa nullo:

In questo modo, il minimo di trasmissione è circa zero, quindi la risposta è sostanzialmente quella di un filtro passabanda. Questo in realtà è dunque un filtro passabanda, ma realizzato mediante un sistema **a parametri distribuiti**: raddoppiando la frequenza, si ha di nuovo la “banda passante”. Questo fatto è molto importante, ed è una regola generale: se un filtro a parametri concentrati (il classico filtro RC o simili) ha una sola banda passante, in un filtro a parametro concentrati esistono, almeno in linea di principio, infinite bande passanti. La banda a -3 dB dipenderà quindi da n_1/n_2 . Questo tipo di “lastra” potrebbe dunque essere usata come filtro passa banda, ma in realtà non si fa, dal momento che servirebbe un dielettrico trasparente e con n elevatissimo, cosa che si può fare ma è abbastanza complicata.

La zona “bassa” della risposta in frequenza è detta “banda attenuata”, andando a pensare ai filtri, ma in realtà ciò non è esatto (nè nella teoria generale dei filtri, nè in quella che stiamo analizzando noi nel dettaglio): un filtro infatti attenua non perchè dissipa potenza, bensì perchè **introduce un disadattamento**: non si ha a che fare con perdite.

2.5.2 Casistica 2: $n_1 = n_3 > n_2$

Si consideri a questo punto un'altra casistica, che permetterà di osservare altri fenomeni: $n_1 = n_3 > n_2$, ovviamente per cui $\vartheta_1 > \vartheta_c$ (infatti, se l'angolo è inferiore all'angolo critico, non si osserva nulla di interessante rispetto a quanto precedentemente descritto). In questo caso:

$$\phi = k_{z2}d = -j|k_{z2}|d$$

ossia, in questo caso si ha un k_{z2} puramente immaginario. Può capitare che:

$$e^{-j2\phi} = e^{-2|k_{z2}|d}$$

sia molto minore di 1, ossia sia un numero estremamente piccolo. Ricordando dunque l'espressione di S_{21} , si ha:

$$S_{21} = \frac{(1 - r_{12}^2)e^{-j\phi}}{1 - r_{12}^2e^{-j2\phi}}$$

L'esponenziale al denominatore è tale da ridurre di molto il contributo del secondo termine: possiamo dunque approssimare il denominatore a 1. L'espressione quindi sarà:

$$S_{21} \sim (1 - r_{12}^2)e^{-|k_{z2}|d}$$

mentre

$$S_{11} \sim 1$$

Abbiamo a questo punto un problema di interpretazione: se l'angolo di incidenza è superiore all'angolo critico, si è visto che si ha il fenomeno della riflessione totale: $|r_{12}| = 1$, quindi tutta la potenza è riflessa. Il campo "a destra della prima interfaccia" non è nullo, dal momento che si ha un'onda evanescente, ma d'altra parte, nel terzo mezzo, quello che si può vedere è che si ha di nuovo propagazione di potenza attiva! In altre parole, $|S_{21}| \neq 0$. Come è possibile, se, come abbiamo visto prima, non c'è transito di potenza attiva nel mezzo 2? Questo fenomeno è detto "FTR": Frustrated Total Reflection, ossia "riflessione totale frustrata". Supponendo per semplicità $n_3 = n_1$, come detto, si ha qualcosa di questo genere:

Nel mezzo 1 si hanno le solite oscillazioni, quindi dopo A l'attenuazione esponenziale, e dopo B il modulo rimane costante. Per avere una fenomenologia di questo genere, significa che da qualche parte la potenza attiva deve propagarsi, ma come è possibile, se la situazione è come quella di prima? La risposta è semplice: la situazione non è più uguale a prima, grazie alla presenza del mezzo 3: l'onda evanescente prima non trasportava infatti potenza attiva, dal momento che tensione e corrente erano in quadratura; ora, dal momento che si ha anche un contributo di onda evanescente **regressiva**, si perde questa condizione: la seconda discontinuità è dunque quella che ci fa ottenere un'onda evanescente che "torna indietro". In altre parole, per quanto non sia evidente, la parte centrale della curva prima mostrata, quella tra i due mezzi, non è in realtà un esponenziale, bensì la somma di due esponenziali; a testimonianza di ciò, c'è un piccolo aspetto analitico: la derivata sulla seconda discontinuità dielettrica, ossia in B, è costante, e si ha un punto a tangente orizzontale; questo, se l'esponenziale fosse uno solo, sarebbe privo di senso, dal momento che un esponenziale non può avere derivata nulla, in nessun punto.

Si può dimostrare che vale la seguente formula¹¹:

¹¹Orta R. - Teoria delle linee di trasmissione (edizione 1999), pag. 45 (sez. 3.4)

$$P = \frac{1}{|Z_\infty|} \text{Im} \{V^{+*}V^-\}$$

Si ha un effetto cooperativo dell'onda progressiva e di quella regressiva.

Il fatto di avere trasporto di potenza, come detto, si può vedere studiando $V(z)$ e $I(z)$, mediante il metodo dei vettori rotanti: la tensione progressiva generica è un certo vettore, la corrente un altro:

Prima di tutto si disegna una tensione progressiva; poi, come noto, la corrente progressiva a partire da questa, essendo noto che (per esempio per la polarizzazione TE, ma non cambia nulla per la TM):

$$I^+ = Y_\infty^{\text{TE}} V^+ = \frac{k_z}{\omega\mu} V^+ = -j \frac{|k_z|}{\omega\mu} V^+$$

la corrente progressiva, dunque, è in quadratura rispetto alla tensione progressiva. Lo stesso discorso si può fare con tensione e corrente regressive.

Quello che si vede, calcolando le risultanti dei vettori tensione progressiva e regressiva, corrente progressiva e regressiva, è che l'angolo tra tensione totale e corrente totale non è pari a $\pi/2$; anzi, quello che si può vedere, è che esso è sempre diverso da $\pi/2$, a meno che la V^- diventi troppo piccola; questo, in effetti, capita se d è elevato, ossia se si fa attenuare molto l'onda evanescente prima di incontrare il punto di riflessione. Se la sezione B è molto lontana, quindi, si ottiene ciò; fisicamente questo è ragionevole, dal momento che l'onda evanescente in B è quasi sparita, quindi non si può riflettere.

Questo effetto può essere sfruttato per la realizzazione di un particolare dispositivo ottico: il già citato **cube beam splitter**.

Dati due prismi, con un indice di rifrazione appropriato, si arriva sull'interfaccia con un certo ϑ_i , tale da essere maggiore dell'angolo critico; in questo modo si ha un'onda riflessa fuori. Se si progetta la distanza d tra i due prismi tale da permettere il transito di potenza, si ottiene un'onda trasmessa, oltre a quella riflessa, per la riflessione totale frustrata. A seconda di d e degli n_i , è possibile regolare l'ampiezza, ottenendo sostanzialmente un divisore di potenza ottico; se ovviamente d è molto grande, non si ha il fenomeno di FTR.

Si noti che la formula del S_{21} , come già accennato precedentemente, potrebbe suggerire risultati errati "nella pratica":

$$|S_{21}| = (1 - r_{12}^2) e^{-|r_{12}|d}$$

sembrerebbe, da questa espressione, che si possa attenuare a piacere, ossia scegliere un certo d , e attenuare anche di 100 dB, per esempio; questo purtroppo nella pratica non è vero, dal momento che non si può andare al

di sotto di un certo fattore di accoppiamento. Questo fatto è dovuto al fatto che l'interfaccia del prisma diventa importante, dal momento che se l'interfaccia è rugosa, ondulata (e per rugosa si può intendere anche semplicemente coperta di qualcosa di estremamente banale come una ditata), si ha dello scattering; questi fenomeni di scattering sono dovuti dunque alle non idealità della superficie, ma di fatto costituiscono un “rumore”, un “fondo”: sia esso - 60 dB o - 100 dB, non è possibile avere accoppiamento inferiore di questo fattore, volendo ovviamente ottenere qualcosa di sensato.

2.5.3 Strati antiriflesso

Si vuole a questo punto proporre una prima applicazione pratica dello studio precedentemente effettuato, in modo da vedere a cosa può servire. Precedentemente, è stato ipotizzato, dati due mezzi n_1 e n_3 , $n_1 = n_3$; questo non è assolutamente necessario, dal momento che i due mezzi possono in effetti essere tra loro diversi: in questo caso (supponendo per ora che non ci sia nulla in mezzo), ci sarà della riflessione; questa riflessione, in molte occasioni, può essere alquanto fastidiosa; ciò che si può fare, tuttavia, è introdurre tra i due mezzi uno **strato antiriflesso**, ossia uno strato con un particolare spessore e un particolare coefficiente di rifrazione, tale da avere $\Gamma_{A-} = 0$. La situazione è dunque la seguente:

Quello che si può fare, come si può osservare dal modello a linee di trasmissione, è ricondurci alle nozioni già note sull'adattatore a $\lambda/4$: se infatti si ha

$$AB = \lambda_{g2}/4$$

e

$$Z_{\infty 2} = \sqrt{Z_{\infty 1} Z_{\infty 3}}$$

si ha uno strato antiriflesso, ossia l'equivalente ottico dell'adattatore a quarto d'onda.

Al fine di dimensionare una struttura di questo genere, la prima idea potrebbe essere quella di determinare, a partire dalla richiesta su $Z_{\infty 2}$, il valore di n_2 : nelle espressioni delle impedenze modali, infatti, si trova un'equazione in funzione di n_2 . Prima di parlare di ciò, tuttavia, è opportuno evidenziare il fatto che, in questo sistema, vi sono delle arbitrarietà:

- qual è l'angolo di incidenza dell'onda (supposta piana) sulla struttura?
- quale polarizzazione stiamo utilizzando?

- a quale frequenza si sta lavorando?

Purtroppo, questa struttura presenta uno svantaggio enorme: essa funziona, ma solo fissati i tre parametri di cui si parlava. Al fine di comprendere il ragionamento dietro il progetto della struttura, tuttavia, si vogliono proporre alcuni calcoli esemplificativi. Si supponga di lavorare con un angolo ϑ_i **fisso e noto**, e con un'onda a **polarizzazione TE**. Detto ciò, si ha:

$$\frac{\omega\mu}{k_{z2}} = \sqrt{\frac{\omega\mu}{k_{z3}} \frac{\omega\mu}{k_{z1}}}$$

Questa si può riscrivere, semplificando tutto ed elevando al quadrato:

$$k_{z2}^2 = k_{z1} k_{z3}$$

a questo punto, si recuperino le espressioni dei vari k modali, e le si sostituiscano:

$$k_0^2 (n_2^2 - n_1^2 \sin^2 \vartheta_i) = k_0^2 \sqrt{(n_1^2 - n_1^2 \sin^2 \vartheta_i) (n_3^2 - n_1^2 \sin^2 \vartheta_i)}$$

Noto n_2 da questa equazione, essendo esso l'unica incognita di tutti i parametri contenuti, è noto che:

$$\lambda_{g2} = \frac{2\pi}{k_0 \sqrt{n_2^2 - n_1^2 \sin^2 \vartheta_i}}$$

ricordando poi che:

$$\lambda_0 = \frac{2\pi}{k_0}$$

si trova:

$$\lambda_{g2} = \frac{\lambda_0}{\sqrt{n_2^2 - n_1^2 \sin^2 \vartheta_i}}$$

Questo permette di determinare lo spessore effettivo dello strato, quindi, sapendo che deve essere pari a un quarto di quello appena analizzato, il dimensionamento è ultimato. Nel caso di polarizzazione TM, si ha qualcosa di un poco più complicato, a causa delle espressioni della $Z_{\infty 2}^{\text{TM}}$ (ma i conti sono sostanzialmente simili, se non per il fatto che si trova un'equazione di secondo grado in n_2 invece di una di primo grado).

Quella appena proposta è una possibile soluzione pratica di un problema. Quali sono i suoi limiti, i suoi problemi?

- L'angolo di incidenza spesso non è definito in maniera precisa: esso sta in un certo campo di variabilità; si immagini per esempio che si voglia progettare uno strato antiriflesso per un sistema ottico di una macchina fotografica: non è assolutamente detto che la luce del Sole incida con un angolo preciso e ben definito sull'obiettivo!
- La polarizzazione non è ben definita: quando si ha la luce dall'ambiente, la polarizzazione come già detto precedente è arbitraria.
- Non è detto di avere a che fare con una luce monocromatica: potrebbe interessare un certo range di colori di luce, quindi una certa banda; il progetto, invece, vale a una singola frequenza.

Quello che si può chiedere, in pratica, non è tanto un *funzionamento perfetto* per tutti i parametri, ma un *buon funzionamento* su un certo range di parametri; questo concetto si può sostanzialmente ricondurre al progetto di un filtro: un filtro si può progettare introducendo dei parametri iniziali, quindi eseguendo un programma, un *software*, in grado di ottimizzare i vari parametri liberi fino a ottenere un risultato soddisfacente.

La strategia di progetto è: non chiedere un risultato “ottimo” per un certo valore di frequenza, ma chiederlo per un valore di frequenza “mediano”, in modo tale da avere, a partire da questo valore di frequenza, una banda in cui il parametro che si intende ottimizzare (nel nostro caso, il Γ) abbia valori ragionevoli, rispetto alla specifica fornita.

Studiamo a questo punto il problema della variabilità in frequenza, considerando tuttavia fissati gli altri parametri: polarizzazione e angolo di incidenza. Si considera inoltre la frequenza f variabile in una certa banda ben definita:

$$f \in [f_{\min}, f_{\max}]$$

Volendo considerare un grafico dell'andamento di Γ funzione della frequenza f , si potrebbe avere qualcosa del genere:

Si cerca il valore f_0 mediano, dunque, usando la strategia di progetto prima proposta, $\Gamma(f_0) = 0$, e in un intorno il Γ cresce. Si deve quindi verificare il valore di Γ_{\max} accettabile, per vedere se, nella banda, il sistema soddisfa le specifiche. In realtà, purtroppo, con questo tipo di adattatore, a quarto d'onda, non è possibile fare di meglio. Si ha nella fattispecie la seguente situazione:

Nel caso $n_1 \sim n_3$, la curva è più “aperta”, dunque sarà tendenzialmente più semplice soddisfare le specifiche con essa; nel caso invece n_1 sia molto diverso da n_3 , l'andamento di $\Gamma(f)$ è molto più ripido; in altre parole, la

banda in cui l'adattatore funziona in maniera accettabile dipende molto dal salto di indice di rifrazione.

A questo punto, si ha un problema: non sono i progettisti quelli che possono decidere n_1 e n_3 , dal momento che spesso n_1 è l'aria (o comunque un mezzo molto poco denso), n_3 per esempio l'obiettivo della macchina fotografica; l'idea, dunque, è quella di "gradualizzare" il salto, introducendo più di una sezione:

Di solito si indica come prima linea una linea con caratteristiche derivanti dall'aria, o comunque da un mezzo esterno, come detto poco denso: n_0 ; i vari tratti di linea, tutti lunghi $\lambda_{g,i}/4$, hanno indice di rifrazione n_i , $i = 1, 2, \dots$; l'elemento finale, quello da "adattare", avrà indice di rifrazione n_s (s sta per substrato, dal momento che l'antiriflesso si ottiene mediante deposizione di vari strati di film sottili su questo substrato). L'obiettivo finale è quindi quello di ridurre il salto grosso in tanti salti piccoli, eliminando il problema della larghezza di banda.

Il fatto di introdurre diversi tratti di linea introduce molti parametri di progetto: a seconda di come si scelgono gli indici di rifrazione, si hanno risultati diversi in termini di curva di progetto; il fatto infatti di avere tutti questi parametri, si presta a utilizzare metodi "sistematici": metodi in grado di far realizzare, al coefficiente di riflessione, curve al variare della frequenza f che siano ben note, per esempio dalla teoria dei filtri. Si noti che quello che stiamo facendo presenta notevoli analogie con i filtri passa-banda, ma anche una sostanziale differenza: a noi, sostanzialmente interessa **solo la banda passante**, non la banda attenuata: fuori "banda" è probabile che non arrivino neanche segnali, quindi non si hanno problemi di questo tipo.

Dalla teoria dei filtri, due esempi di risposte molto famose sono le seguenti:

La prima è una risposta **alla Butterworth**, anche detta **a massima piattezza**: si tratta di una funzione nulla a f_0 e con un certo numero di derivate nulle; questa è la curva più piatta che si può ottenere. La seconda curva è detta **alla Chebyshev**, ed è estremamente interessante: essa permette di avere delle ondulazioni, ad ampiezza costante; questa di solito è più apprezzata, dal momento che presenta proprietà molto interessanti.

A questo punto, non si vuole presentare il procedimento esatto per progettare sistemi di questo tipo, ma solo alcune idee; progettare il sistema significa scegliere i coefficienti n_i ; vi sono diverse strade:

- scegliere una cosiddetta "sintesi diretta": trovare tutti gli n_i direttamente, ma ciò non è semplice;
- utilizzare una soluzione basata sulla "teoria delle piccole riflessioni".

Teoria delle piccole riflessioni

Senza entrare troppo nei dettagli, facendo riferimento al problema dei tre mezzi, si vuole a questo punto presentare l'idea dietro la teoria delle piccole riflessioni, dal momento che spesso la si incontra, studiando problemi di elettromagnetismo ingegneristico.

Precedentemente, è stata ricavata la seguente espressione:

$$S_{11} = \frac{r_{12} + r_{23}e^{-j2\phi}}{1 + r_{12}e^{-j2\phi}}$$

Non consideriamo la presenza del fenomeno delle riflessione totale: questo, dal momento che di solito si parte da un mezzo poco denso, verso uno più denso; gli r_{ij} , come noto, sono coefficienti di riflessione di Fresnel.

L'obiettivo della presente sottosezione è quello di ricavare una formula ragionevolmente accurata, a patto che essa lavori con coefficienti di Fresnel piccoli: questa è un'ipotesi ragionevole quando i salti tra i coefficienti di rifrazione n_i sono piccoli.

Studiando il denominatore, è possibile vedere che si ha il prodotto di due quantità "piccole": questa è quindi una "quantità piccola del secondo ordine", la quale è, ragionevolmente, trascurabile rispetto a 1; si può dunque dire che:

$$S_{11} = r_{12} + r_{23}e^{-j2\phi}$$

Questa formula, così presentata, è assolutamente iterabile; studiando il più semplice caso di linea di trasmissione con una discontinuità di impedenza, si ha qualcosa di questo genere:

L'obiettivo, in generale, è quello di ottenere una formula iterativa in grado di ottenere Γ_{A-} , come funzione di Γ_{A+} . La strada per fare ciò, come noto dai corsi di Circuiti a parametri distribuiti, è quella di passare per le impedenze, denormalizzare e rinormalizzare; lavorando analiticamente, tuttavia, è possibile trovare una formula di questo tipo:

$$\Gamma_{A-} = \frac{r_{12} + \Gamma_{A+}}{1 + r_{12}\Gamma_{A+}}$$

dove, ovviamente:

$$r_{12} = \frac{Z_{\infty 2} - Z_{\infty 1}}{Z_{\infty 2} + Z_{\infty 1}}$$

Si consideri a questo punto il solito esempio, a partire dai tre mezzi: due sezioni, A e B, in cui si hanno delle discontinuità. In questo caso, si ha che, nella formula:

$$\Gamma_{A+} \longleftrightarrow \Gamma_{B-} e^{-j2\phi}$$

e

$$\Gamma_{B-} = r_{23}$$

c'è dunque una perfetta corrispondenza con questa formula.

Questa formula, sviluppata secondo la teoria delle piccole riflessioni, è così approssimabile:

$$\Gamma_{A-} \sim r_{12} + \Gamma_{A+}$$

questa, ovviamente, è valida a patto che:

$$\begin{cases} r_{12} \ll 1 \\ |\Gamma_{A+}| \ll 1 \end{cases}$$

Questa formula, come detto, è applicabile iterativamente: dato un circuito con tanti pezzi di linee di trasmissione collegate assieme, con lunghezze generiche (per gli adattatori sono necessarie lunghezze ben definite, ma l'applicazione della teoria delle piccole riflessioni non dipende assolutamente dalla lunghezza delle linee, a meno che essa non sia collegata con i due parametri precedentemente introdotti). Si può vedere che la formula, approssimata (ovviamente anche quella più esatta, ma quella approssimata è molto più semplice), è applicabile iterativamente, ottenendo:

$$\Gamma_{A-} = r_{12} + r_{23}e^{-j2\phi} + r_{34}e^{-j4\phi} + \dots$$

dove, ogni volta, si aggiunge uno sfasamento, “venendo indietro”, di $e^{-j2\phi}$.

Si noti che i vari r_{ij} non sono noti, ma d'altra parte sono essi il risultato che si deve ottenere: questa approssimazione è stata applicata per avere un'espressione semplice, come d'altra parte questa è (si tratta solo di una somma pesata con degli esponenziali complessi), dei vari r_{ij} ; questa approssimazione è stata fatta dal momento che essa è importante ai fini di progettare il sistema, non di analizzarlo: usandola, è possibile quindi progettare gli strati antiriflesso (o altri sistemi) come appena visto.

Capitolo 3

Strutture periodiche

3.1 Introduzione agli specchi - Riflettore di Bragg

A questo punto si vuole analizzare un'altra struttura, per ora con i metodi precedentemente utilizzati, ma poi applicando un formalismo completamente diverso; per questo motivo, e per l'importanza di queste strutture (e di questi approcci), si è pensato che sia il caso di dedicare un nuovo capitolo a tutto ciò, sebbene la parte introduttiva sia sostanzialmente riconducibile a ciò che è già stato analizzato.

L'obiettivo di questa prima sezione è quello di realizzare, al contrario di prima, uno **specchio**: se prima l'obiettivo era quindi quello di eliminare le riflessioni, ora l'obiettivo è quello di ottenere una struttura molto riflettente.

Gli specchi esistono già, senza bisogno di usare le nozioni che stiamo per introdurre: utilizzando un film di argento depositato, quindi un vetro che da un lato protegge il film di argento, dandogli un supporto meccanico, e dall'altro permettendo di ottenere una rugosità superficiale molto ridotta (a seconda della qualità necessaria è possibile ottenere errori in frazioni di lunghezze d'onda: $\lambda/10$, $\lambda/20$, o altro a seconda della necessità). Come mai dunque si desidera parlare di specchi, quando già esiste questa eventualità? Beh, il motivo è semplice: gli specchi realizzati a partire da film depositati, film di materiali metallici, hanno il problema delle perdite: si ha una ε con parte reale, e parte immaginaria (per quanto quest'ultima sia una frazione della parte reale, generalmente). Il coefficiente di riflessione, dunque, non è esattamente 1, o quantomeno non è neanche possibile pensare di ottenerlo: qualsiasi sia il suo valore, in alcune situazioni potrebbe essere necessario fare di meglio.

Esposto il problema, esponiamo la soluzione: come è possibile ottenere uno specchio migliore di quello appena descritto? Ciò di cui disponiamo, ora, è un certo substrato n_s , sul quale vogliamo fare depositare degli strati di dielettrici, a coppie (ciascuna coppia è detta “cella”). Quanto sono lunghi i dielettrici nella cella? Beh, la soluzione, ancora una volta, è:

$$l_1 = \lambda_{g1}/4$$

$$l_2 = \lambda_{g2}/4$$

Come mai ancora una volta si ha a che fare con queste lunghezze elettriche? La risposta è abbastanza semplice: scegliendo gli spessori in questo modo, i campi riflessi si sommano tutti in fase: questa, quindi, è ancora una volta un’applicazione del principio di interferenza.

Si noti tuttavia un fatto: stiamo parlando immediatamente della lunghezza degli elementi della cella, ma non stiamo parlando di n_1 e n_2 ; la fondamentale differenza tra questo progetto e il precedente (quello dello strato antiriflesso) sta nel fatto che, prima, era fondamentale ricavare, a partire da un’equazione (quella ricavata a partire dalle linee modali imponendo la condizione di adattatore a $\lambda/4$), i valori di n_1 e n_2 ; questi, ora, sono “liberi”: di fatto, fornitici due materiali che possiamo depositare, senza avere bisogno di determinarli, è possibile realizzare uno specchio con un indice di rifrazione desiderato.

Utilizziamo il solito metodo di analisi:

Se i tratti di linea sono $\lambda_g/4$, non è nemmeno necessario ricorrere alla caratterizzazione secondo carta di Smith. Prima di tutto, si può determinare l’impedenza normalizzata alla sezione C^- :

$$\zeta_{C^-} = \frac{Z_{\infty,s}}{Z_{\infty 2}}$$

dal momento che il tratto di linea lungo un quarto della lunghezza d’onda, come noto dalla teoria delle linee di trasmissione, inverte l’impedenza normalizzata, si ha che:

$$\zeta_{B^+} = \frac{1}{\zeta_{C^-}} = \frac{Z_{\infty 2}}{Z_{\infty,s}}$$

quindi, si ha che:

$$\zeta_{B^-} = \zeta_{B^+} \frac{Z_{\infty 2}}{Z_{\infty 1}} = \frac{Z_{\infty 2}^2}{Z_{\infty,s} Z_{\infty 1}}$$

Andiamo avanti: per arrivare alla parte destra della sezione A, si incontra ancora una volta un invertitore di impedenza:

$$\zeta_{A^+} = \frac{1}{\zeta_{B^-}} = \frac{Z_{\infty,s} Z_{\infty}}{Z_{\infty 2}^2}$$

quindi,

$$\zeta_{A^-} = \frac{Z_{\infty,s} Z_{\infty}}{Z_{\infty 2}^2} \frac{Z_{\infty 1}}{Z_{\infty 0}} = \left(\frac{Z_{\infty s}}{Z_{\infty 0}} \right) \left(\frac{Z_{\infty 1}}{Z_{\infty 2}} \right)^2$$

Il primo fattore sarebbe l'unico presente, nel caso non ci fosse la cella; questa sarebbe, in questo caso, l'impedenza di ingresso. Il secondo fattore dipende invece dai due strati in mezzo; in questo caso, $Z_{\infty 1} \neq Z_{\infty 2}$, quindi il rapporto può essere maggiore o minore di 1.

Posizionando (come si può vedere osservando una carta di Smith) un valore di impedenza normalizzata maggiore di 1, esso andrà "a destra" della carta di Smith; posizionandone uno "minore di 1", esso andrà a sinistra (si ricordi che a sinistra c'è il corto circuito, a destra il circuito aperto, quindi è ragionevole il fatto che l'impedenza caratteristica si avvicini a una di queste condizioni).

Si immagini a questo punto di avere a che fare con $\frac{Z_{\infty,s}}{Z_{\infty 0}} < 1$: in questo caso, il "punto" va a sinistra, ottenendo un Γ reale ma minore di 0. In questa condizione, la strategia è quella di avere:

$$\frac{Z_{\infty 1}}{Z_{\infty 2}} < 1$$

in questo modo, essendo il punto già "a sinistra", esso viene spostato "ancora più a sinistra", ottenendo con maggiore facilità un coefficiente di riflessione elevato in modulo. Nella situazione duale, ossia se il "punto iniziale" è a destra della carta di Smith, si fa in modo da "andare ancora più a destra": sarebbe ovviamente possibile, in entrambe le condizioni, fare il contrario, ma ciò sarebbe sicuramente inutile e più dispendioso. Inutile, dal momento che, nel caso il rapporto tra le impedenze caratteristiche della cella risultasse "scomodo", è sufficiente scambiare l'ordine di deposizione dei materiali, invertendo di fatto il rapporto e ottenendo ciò che si desidera: basta scambiare i dielettrici!

A questo punto sorge un altro problema: potrebbe essere che un materiale abbia indice per esempio pari a 1,5, l'altro a 50: un materiale con questo secondo n è dunque difficile da realizzare (se non impossibile). Si è tuttavia detto, all'inizio della sezione, che gli n_i possono anche non essere a nostra scelta, bensì possono esserci forniti dall'esterno; come è possibile fare

ciò? Semplice: usando N celle! Se si hanno anche indici di rifrazione molto vicini, ma mettendo per esempio 40 celle, è possibile aumentare enormemente il coefficiente di riflessione della struttura. Si osservi che mettere tante celle è assolutamente ragionevole: le lunghezze d'onda in gioco (e quindi le lunghezze) sono molto ridotte, quindi lo spessore verrà di qualche micron, anche con molte celle in cascata.

Iterando il ragionamento, si ottiene ciò:

come si vede, dalla sezione C si vede Z_C , invece che $Z_{\infty,s}$; questa cosa tuttavia è interessante, dal momento che Z_C è quella di prima, quindi si può applicare la formula precedente, considerando come “primo termine” quello dipendente da Z_C . Fatto questo ragionamento, il risultato finale è:

$$\zeta_{A^-} = \frac{Z_{\infty,s}}{Z_{\infty 0}} \left(\frac{Z_{\infty 1}}{Z_{\infty 2}} \right)^{2N}$$

dove N è, come detto, il numero di celle.

Questo metodo, come si può osservare, è molto flessibile!

Questa struttura per la realizzazione di specchi è detta “riflettore di Bragg” o “specchio di Bragg”: la riflessione in questo caso è solo legata alle riflessioni causate dall'interfaccia dielettrica; questo permette sostanzialmente di ridurre le perdite, rispetto a quelle che si avevano quando si utilizzavano materiali metallici come base per la realizzazione dello specchio. Le uniche due posizioni che funzionano, per il progetto di questo sistema, sono il “punto più a sinistra” e il “punto più a destra” della carta di Smith; quale scegliere di questi, dipende sostanzialmente dal substrato e dalla polarizzazione.

Sostanzialmente, il progetto potrebbe essere di questo genere: a partire da un $\Gamma_{\min} = 0,99$ (per esempio, per dare un numero), si ricava la Z_{A^-} che si deve avere; questa deve essere ottenuta, a partire dai dielettrici a nostra disposizione, quindi si fanno i calcoli e si ricava un'equazione in cui N è l'unica incognita.

A questo punto ci poniamo un'altra domanda: cosa succede se la frequenza cambia, rispetto a quella di progetto? Quello che si può verificare (e che verificheremo) è sostanzialmente un comportamento di questo genere:

dove f_0 è la frequenza per cui la lunghezza dei tratti di dielettrico è $\lambda_g/4$, e in cui il Γ soddisfa bene la specifica.

Questo comportamento ricorda il comportamento di un filtro che ha riflessione (Γ) alla “frequenza di centro banda”: questo è sostanzialmente un filtro elimina banda, dal momento che alla banda “importante” esso riflette molto (essendo, sotto il punto di vista ottico, uno specchio).

Si noti che lo specchio appena presentato (il riflettore di Bragg) funziona sostanzialmente grazie a fenomeni di risonanza, dal momento che tutto è

basato sul fatto che si hanno segnali che si sommano in fase (ancora una volta, dunque, a partire dal principio dell'interferenza); a frequenze basse, il Γ è nullo, quindi quello che dovrebbe essere uno specchio in realtà è trasparente. Attorno a f_0 si ha una riflettività forte, ma non esattamente unitaria, come detto.

3.2 Introduzione al formalismo delle onde di Bloch

A questo punto ci poniamo una domanda, alla quale potremo per ora rispondere solo in maniera qualitativa: come mai le curve dell'andamento in frequenza hanno un andamento di questo genere?

Proviamo a proporre un'idea: la struttura appena presentata (il riflettore di Bragg) è stata analizzata dicendo che essa è composta da N celle: dato lo studio di una cella, abbiamo introdotto una generalizzazione.

Questo non è in realtà l'unico approccio che si possa utilizzare nello studio di un problema del genere, ossia nello studio di una **struttura periodica**: la struttura introdotta infatti presenta una periodicità importante, dal momento che le celle sono sempre uguali tra loro, e semplicemente ripetute. Quello che si può fare, tuttavia, è considerare una struttura periodica infinita, composta quindi da infinite celle tutte uguali tra loro, prendendone solo un pezzo: una **struttura periodica troncata**.

Un esempio sul quale si possono applicare questi due approcci diversi, è la linea di trasmissione: l'approccio che si utilizza, inizialmente, è quello di considerare un circuito di questo genere:

Questo disegno è sostanzialmente ottenuto “discretizzando” il comportamento di una linea di trasmissione, considerandone tanti modelli a parametri concentrati (una linea di trasmissione, come un coassiale, si modella meglio mediante un circuito a parametri distribuiti), dove ciascun parametro è discretizzato su una certa lunghezza Δz . A questo punto, si hanno due possibili approcci:

- applicare Kirchhoff, fare il limite per $\Delta z \rightarrow 0$, ottenendo le equazioni dei telegrafisti; le soluzioni delle equazioni dei telegrafisti sono un'onda progressiva e un'onda regressiva, ossia ciò che usiamo sempre come base per la rappresentazione dello stato elettrico in una struttura guidante distribuita;
- inserire “tante celle” in un simulatore (o fare il conto a mano, essendo una rete a scala), quindi determinare lo stato elettrico semplicemente mediante tensioni e correnti, come sempre fatto.

Il secondo metodo appena proposto, se ci si pensa, è esattamente coincidente con quanto fatto nell'approccio precedentemente proposto per la descrizione del riflettore di Bragg. Esso, a patto di rispettare alcune ipotesi, è un approccio sostanzialmente corretto: se Δz è piccolo rispetto a λ (10, 20, 100 volte), si ha la garanzia che i due approcci convergano allo stesso risultato, con un certo numero di cifre significative; l'approccio basato sui circuiti a parametri distribuiti, tuttavia, è un approccio "esatto".

C'è un altro aspetto interessante: se si osserva il modello da cui partono tutte le nostre osservazioni, esso è un modello circuitale, una normalissima rete a scala: se ci si pensa, immaginare che in quel circuito a parametri concentrati viaggino delle **onde**, è una cosa assolutamente non banale, per non dire impossibile: questo approccio ondulatorio di fatto sembra, fisicamente, estremamente scorrelato al nostro modello, per quanto la matematica (e l'osservazione sperimentale) suggeriscano che in realtà sono due visioni diverse dello stesso fenomeno, ovviamente a patto che siano rispettate le suddette ipotesi.

Con le onde progressiva e regressiva introdotte dal modello delle equazioni dei telegrafisti, si hanno sostanzialmente eccitazioni, del tipo e^{-jkz} , che sono "connaturate con una linea infinitamente lunga". Il termine "connaturate" significa sostanzialmente che "l'ambiente naturale" per cui queste onde progressive si propagano senza scatterare, senza subire fenomeni di interferenza, è la linea di lunghezza infinita: se infatti la linea non ha troncamenti o specchi, non c'è nulla che costringa l'onda a tornare indietro.

Ciò che si vuol fare, è descrivere la struttura periodica, ossia la struttura composta da un insieme di strati, mediante l'utilizzo di un formalismo ondulatorio, dove però le onde come ambiente naturale non avranno quello della struttura "libera", ma quello della **struttura periodica**: le "onde progressive" si propagheranno bene fintanto che resteranno confinate all'interno di una struttura che mantenga delle periodicità. Si noti che un'onda di quelle ricavate dalla soluzione delle equazioni dei telegrafisti, in un ambito come quello della struttura periodica che si vuole analizzare, non sarebbe assolutamente "nel suo ambiente naturale": ogni barriera porterebbe a fenomeni di scattering, ottenendo dunque riflessioni e trasmissioni multiple. Nell'approccio che utilizzeremo, invece, l'onda, che non sarà un'onda progressiva di quelle dei telegrafisti ma qualcosa di un poco diverso, si propagerà come in un "ambiente continuo", fintanto che rimarrà in una situazione ben precisa. Per queste onde che si propagano con naturalezza in una struttura periodica, dette **onde di Bloch**, si potrà definire una relazione di dispersione: una relazione del tipo $\underline{k}(\omega)$. Ciò ha per esempio conseguenze in ambito quantistico: le bande di energia in un reticolo cristallino sono infatti delle curve in cui sulle ascisse c'è proprio il vettore d'onda \underline{k} (o comunque il momento, la

quantità di moto), sulle ordinate l'energia.

La possibilità di introdurre delle onde che si possano propagare, permette quindi di applicare i concetti che ci sono già famigliari dalla teoria delle linee di trasmissione, usando tuttavia come onde propaganti queste onde di Bloch. Nella fattispecie, si avrà qualcosa di questo genere:

In questo caso, invece di avere le discontinuità causate dallo scattering multiplo, si avrà a che fare con solo due discontinuità: la prima, da “primo mezzo” a “struttura periodica”, la seconda da “struttura periodica” a “secondo mezzo”; il problema, in termini di linee di trasmissione, sarà ovviamente molto più semplice di quello basato sullo studio dello scattering multiplo. In altre parole, una struttura periodica sarà semplicemente assimilabile a un particolare singolo tratto di linea di trasmissione. In altre parole, le onde di Bloch si propagano bene nella struttura, fino a quando essa mantiene la propria periodicità. Come si vedrà in seguito, inoltre, le onde di Bloch sono costituite da una “giusta mistura” di onda progressiva e onda regressiva, rendendone dunque la determinazione cosa relativamente semplice; dire che un'onda di Bloch “rappresenta bene” il comportamento di una struttura periodica significa proprio che i coefficienti “di peso” per le varie componenti di onda sono quelli giusti per avere la propagazione dell'onda di Bloch, nella struttura, senza avere riflessioni interne a essa.

Si noti inoltre che le onde di Bloch possono essere propagative o evanescenti: nella fattispecie, la zona in cui si ha la “banda stoppata”, ossia quella per cui lo specchio si comporta effettivamente da specchio, è quella per cui si ha la FTR, sulle onde di Bloch: quella per cui le onde di Bloch sono proprio evanescenti.

A patto di credere a tutto ciò (e questi discorsi verranno motivati con la matematica), la curva di risposta diventa più comprensibile: si era infatti visto analizzando la lastra, che si ha un comportamento di questo genere:

n_0 e n_s nel nuovo caso non dipendono dalla frequenza, mentre nel blocco in mezzo sì (essendo una struttura sostanzialmente risonante); i “picchi” però compaiono comunque, essendo il comportamento formalmente simile o comunque riconducibile a quello di prima. Nella fattispecie, gli zeri di riflessione sono quelli per cui lo sfasamento delle onde di Bloch è multiplo di π (mentre ora, questa condizione riguardava lo sfasamento delle onde dei telegrafisti).

Giustificare ora la “banda piatta” attorno a f_0 è invece più difficile; quello che si può tuttavia dire è che essa sia sostanzialmente, in ambito ottico, l'equivalente del E_g nei semiconduttori, nella fisica dello stato solido: è un “band gap” di frequenza, ossia la regione per cui le onde di Bloch, come già accennato, sono evanescenti.

3.2.1 Cenni sull'interferometro di Fabry-Perot

L'interferometro di Fabry-Perot è una struttura di questo tipo:

Dati due specchi semitrasparenti, con un coefficiente di riflessione dunque elevato ma anche uno di trasmissione non nullo, posti tra loro a una certa distanza d , quando si manda un'onda piana da sinistra, si ha un contributo di onda riflessa e uno di onda trasmessa, e ciò sostanzialmente dipende dal rapporto tra d e λ . Questo fenomeno è ancora una volta basato sul concetto di interferenza: ad alcune frequenze si ha una forte riflessione, ad altre una forte trasmissione; utilizzando questo approccio, si fa qualcosa di analogo alla struttura di Bragg, per gli specchi.

In realtà, questo sistema può essere analizzato come una cavità risonante: il concetto di “risonanza” infatti è basato sull'idea che le caratteristiche ingresso-uscita di un dispositivo variano molto rapidamente con la frequenza, a causa della presenza di un picco.

I risonatori si possono fare in diversi modi: a frequenze basse, si fanno a parametri concentrati; a frequenze più alte, in altra maniera; a frequenze altissime, mediante una scatola (magari di metallo ma non è troppo detto), chiusa. Questo approccio sembrerebbe andare bene per il range di frequenze nell'ambito dell'ottico, ma in realtà non si può fare: le frequenze di risonanza, per strutture di dimensione maggiori a 1, tendono ad affollarsi. Come mai diciamo ciò? Beh, l'idea è abbastanza semplice: se come struttura di risonanza si ha una semplice linea monodimensionale chiusa su di un carico, le sue frequenze di risonanza sono:

$$f_n = n \frac{c}{2d}$$

Queste, evidentemente, sono tutte equidistanti tra loro. La stessa cosa purtroppo non vale per quanto riguarda strutture a più dimensioni, come una guida d'onda:

in una guida d'onda, si ha che:

$$k_z = \sqrt{\left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2}$$

Vicino all'origine, i punti sono “distanti” tra loro; allontanandosi dall'origine, tuttavia, la lunghezza del vettore (Pitagora) che collega due punti è sempre più simile; detto in altre parole, i vettori che collegano due punti, due “modi” adiacenti saranno molto più simili tra loro, rispetto a due vettori che rappresentano due modi vicini “all'origine”. In altre parole, “allontanandoci dall'origine”, dunque per frequenze molto elevate, come sono quelle ottiche, si ha a che fare con un affollamento di modi, con un affollamento di frequenze

di risonanza; una cavità di 1 centimetro, quindi, ai THz, non si può usare, dal momento che si avrebbero troppe frequenze di risonanza, quindi si avrebbe una cavità con una relazione non armonica di multipli di frequenza.

Un trucco che si utilizza in ottica è quello di considerare una cavità in cui si abbiano “solo le basi”, senza avere “le pareti”: le **cavità aperte**. Se si fa così, si evita di introdurre tutte le risonanze “trasversali”, ottenendo solo le “risonanze longitudinali”; questo è lo “spirito” della cavità di Fabry-Perot: l’idea è quello di “sfoltire le risonanze”, tenendo solo quelle longitudinali.

Un risonatore Fabry-Perot si usa soprattutto nei LASER a gas, con una piccola variante: quello che si fa, di solito, è avere sì degli specchi, ma “curvi”: si scopre che i modi di risonanza non sono onde piane, ma fasci gaussiani, ossia fasci di radiazione che assomigliano a quelli che escono dal puntatore LASER. I bordi degli specchi sono quindi poco illuminati, l’energia del materiale attivo finisce tutta sugli specchi, quindi il Q della cavità è elevato. Uno degli specchi non rifletterà 1, ma meno, quindi parte della luce uscirà fuori: se la luce esce, questa sarà quella del LASER.

Qui tutto è in fase: si ha un unico modo di risonanza della cavità che esce fuori dallo specchio, quindi la luce del LASER è una luce **coerente**.

I due ingredienti fondamentali per avere un LASER sono:

- un materiale attivo, che amplifica alla frequenza di interesse (e a molte altre frequenze); essendo attivo, significa che gli deve essere fornita dall’esterno; questo può essere, per esempio, un LED;
- il risonatore, che è quello che, di tutte le frequenze prodotte, amplificate dal materiale attivo, seleziona solo quella che interessa: è la parte che “fissa la frequenza” del sistema.

In altre parole, il LED emette, e la “riga” è quella del risonatore.

I LASER a semiconduttore in realtà non usano esattamente questa strategia, ma qualcosa di simile a ciò:

Si hanno due specchi, di cui uno dei due deve essere più sottile, più trasparente dell’altro, in modo che lasci passare dell’onda. All’interno, quindi, si ha uno spazio $d = \lambda/2$, in modo da avere una “cavità risonante”. Alla fine, volendo, questa è ancora una struttura Fabry-Perot, tenendo conto di come sono fatti gli specchi (con dei riflettori di Bragg). In realtà, questi LASER sono detti DBR: Distributed Bragg Reflectors.

Lo spazio in mezzo, il d , è detto “difetto”, dal momento che è spaziato $\lambda/2$ invece di $\lambda/4$; questo tipo di linguaggio nasce dal contesto del “PBG”: Photonic Band Gap: dal momento che le strutture a semiconduttore manifestano dei gap di energia, Yablonovitch inventò questo nome, per analogia con

gli E_g della fisica dello stato solido. La struttura sembrerebbe “difettosa”, per questo $\lambda/2$, ma in realtà ciò, sotto un punto di vista elettromagnetico, torna tutto: ciò permette di avere, alla fine dei conti, una vera e propria cavità risonante; in altre parole, un Fabry-Perot, ottenuto però con questo tipo di strutture.

3.3 Concetti fondamentali di Algebra Lineare

Al fine di capire da dove derivi la curva raffigurante l’andamento di S_{11} al variare della frequenza f , è necessario determinare particolari configurazioni dello stato elettrico, tali per cui esse si propagano senza scattering. Al fine di introdurle, tuttavia, è necessario riprendere alcuni concetti fondamentali legati ai problemi agli autovalori, in modo da poterli poi applicare nel formalismo matematico in questione.

Una generica matrice 2×2 ha una forma del tipo:

$$\underline{\underline{A}} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

Si definisce un problema algebrico agli autovalori come:

$$\underline{\underline{A}}\underline{u} = \lambda\underline{u}$$

dove \underline{u} è detto “autovettore”, λ “autovalore” (si noti che λ non è una lunghezza d’onda: si è usata questa notazione dal momento che, nell’ambito dell’algebra lineare, è la più usuale). Da un punto di vista di principio, al fine di calcolare gli autovettori, è possibile portare tutto al membro sinistro, ottenendo:

$$(\underline{\underline{A}} - \lambda\underline{\underline{I}})\underline{u} = 0$$

Questo è un sistema lineare omogeneo (dal momento che non si ha un termine noto); \underline{u} è una soluzione del sistema, ma ovviamente non è interessante essendo la soluzione banale; le soluzioni interessanti si hanno solamente se il rango della matrice associata al sistema lineare è minore del massimo, dunque se:

$$\det \{ \underline{\underline{A}} - \lambda\underline{\underline{I}} \} = 0$$

solo in questo caso, la soluzione è non unica. Quello che si può fare, dunque, è **richiedere** che questa condizione sia verificata, ottenendo un’equazione che permette di determinare λ . Si può dimostrare, come noto dall’algebra

lineare, che, scrivendo per esteso l'equazione imponente la condizione appena presentata, che:

$$\lambda^2 - \text{Tr} \{ \underline{\underline{A}} \} + \det \{ \underline{\underline{A}} \} = 0$$

questa è l'equazione che permette di determinare gli autovalori della matrice; Tr è detta “traccia” della matrice, ed è la somma degli elementi della diagonale principale; il determinante è cosa nota:

$$\det \{ \underline{\underline{A}} \} = A_{11}A_{22} - A_{21}A_{12}$$

Si vuole portare all'attenzione del lettore un concetto molto importante: la traccia e il determinante della matrice hanno un significato **invariante**: quando si ha a che fare con una certa matrice, essa si può interpretare come una matrice associata a un certo operatore lineare \mathcal{A} , che agisce su uno spazio vettoriale (stiamo considerando tutto di dimensione rigorosamente finita, dunque la matrice potrebbe per esempio essere 2×2 come quella prima riportata); se si introduce una base con la quale generare questo spazio vettoriale, con la quale generarlo, per esempio $(\underline{e}_1, \underline{e}_2)$, si ha una corrispondenza biunivoca tra l'operatore \mathcal{A} e la matrice con la quale lo si rappresenta, al momento di applicarlo sullo spazio rappresentato con la base, $\underline{\underline{A}}$. Ovviamente, se si ha a che fare con basi diverse, \mathcal{A} può essere applicato a basi diverse: lo stesso operatore può essere descritto, a seconda della base, con diverse matrici.

Questo concetto, che sembra molto astratto, ha in realtà una validità molto generale: un doppio bipolo, un generico circuito dunque, può essere rappresentato con la matrice catena (ABCD), o con la matrice di trasmissione, $\underline{\underline{T}}$: in entrambi i casi, si rappresenta un “operatore di trasmissione”, ossia si rappresenta lo stato elettrico in uscita da un certo blocco, noto lo stato elettrico in ingresso a esso. Gli autovalori di questa matrice, essendo una proprietà non solo di una matrice, ma proprio dell'operatore, saranno uguali in tutti e due i casi: lo stato elettrico è rappresentato mediante basi diverse (tensioni e correnti piuttosto che onde di potenza), ma, essendo l'operatore lo stesso, le matrici saranno diverse, gli autovalori uguali. Si può poi dimostrare che:

$$\text{Tr} \{ \underline{\underline{A}} \} = \sum_i \lambda_i$$

$$\det \{ \underline{\underline{A}} \} = \prod_i \lambda_i$$

Data una matrice 2×2 , è possibile ricavare i due autovalori (che, solitamente, saranno 2, come si vedrà); noti gli autovalori, è poi possibile calcola-

re gli autovettori, risolvendo il sistema lineare imponendo la presenza degli autovalori nel sistema lineare.

Al fine di fissare i concetti, si consideri un esempio specifico:

$$\underline{A} = \begin{bmatrix} 0 & a \\ b & 0 \end{bmatrix}$$

dove

$$a, b \in \mathbb{R}, > 0$$

in questo caso, la traccia è evidentemente nulla, dal momento che gli elementi sulla diagonale sono nulli; si ha quindi che gli autovalori saranno la soluzione di questa equazione:

$$\lambda^2 - ab = 0$$

quindi

$$\lambda_1 = \sqrt{ab} \quad \lambda_2 = -\sqrt{ab}$$

Si vogliono a questo punto calcolare gli autovettori relativi all'autovalore λ_1 :

$$\begin{bmatrix} -\lambda_1 & a \\ b & -\lambda_1 \end{bmatrix} \begin{bmatrix} u_1^{(1)} \\ u_2^{(1)} \end{bmatrix} = 0$$

l'apice indica il fatto che si ha a che fare con autovettori relativi all'autovalore λ_1 . Riscrivendo il sistema, si ottiene:

$$\begin{cases} -\lambda_1 u_1^{(1)} + a u_2^{(1)} = 0 \\ b u_1^{(1)} - \lambda_1 u_2^{(1)} = 0 \end{cases}$$

Si può a questo punto ricavare $u_2^{(1)}$ dalla prima equazione:

$$u_2^{(1)} = \frac{\lambda_1}{a} u_1^{(1)} = \sqrt{\frac{b}{a}} u_1^{(1)}$$

in alternativa, è possibile ricavare $u_2^{(1)}$ dalla seconda equazione:

$$u_2^{(1)} = \frac{b}{\lambda_1} u_1^{(1)}$$

Cosa abbiamo scoperto? La seconda equazione non dà informazioni ulteriori: risolvendo una a caso delle due equazioni, si ottengono già tutte le soluzioni possibili; provando a risolvere entrambe le equazioni, si otterrebbe una

forma indeterminata, come $0 = 0$: una cosa che non ci interessa. In realtà ciò non è assolutamente strano, dal momento che queste due equazioni non sono linearmente indipendenti, e ciò non è una scoperta particolarmente grandiosa: dal momento che λ_i è quel valore di λ tale per cui si annulla il determinante, è ragionevole immaginare che le due equazioni, alla fine, siano tra loro indipendenti. Volendo esprimere quindi la coppia autovalore-autovettore, si ha:

$$\lambda_1, \underline{u}^{(1)}$$

dove

$$\underline{u}^{(1)} = \begin{bmatrix} u_1^{(1)} \\ u_1^{(1)} \sqrt{\frac{b}{a}} \end{bmatrix}$$

Una nota: volendo calcolare mediante MATLAB autovalori e autovettori, esiste una specifica funzione:

$$[V, D] = \text{eig}(A);$$

D è una matrice diagonale che contiene, sulla diagonale, gli autovalori; V è una matrice piena, in cui ciascuna colonna rappresenta un autovettore relativo all'autovalore sulla stessa riga/colonna. Gli autovettori sono vettori, ma di fatto essi, per essere scritti in forma numerica, devono essere normalizzati in qualche modo; un modo è quello di scegliere una componente in maniera arbitraria, quindi l'altra in funzione della prima; un altro modo, invece, è quello di normalizzare secondo una certa norma il vettore:

$$|\underline{u}^{(1)}|^2 = 1$$

questa è nella fattispecie quella fatta da MATLAB.

Per quanto riguarda λ_2 , facendo gli stessi conti, si vede che:

$$\underline{u}^{(2)} = \begin{bmatrix} u_1^{(2)} \\ -\sqrt{\frac{b}{a}} u_1^{(2)} \end{bmatrix}$$

A questo punto, al fine di comprendere tutto ciò che è stato spiegato, possiamo porci una domanda: qual è il significato geometrico del problema agli autovalori? Prima di tutto, ha senso porsi la domanda a partire da una particolare interpretazione dello spazio vettoriale: introducendo come base dello spazio vettoriale su cui la matrice relativa alla base opera come

(x, y) (il piano cartesiano), è possibile disegnare, sul piano xy , il vettore, conoscendone le componenti (ovviamente la cosa è arbitraria, nel senso che possiamo scegliere noi la base, ma tutto, a partire dalla matrice, deve essere noto coerentemente con la base cartesiana).

Dato dunque \underline{u} vettore del piano, se vi si applica sopra la matrice \underline{A} , si otterrà, generalmente, un altro vettore, completamente scorrelato dal precedente: diverso modulo, diversa direzione. Ci potremmo porre la domanda: esistono particolari vettori \underline{u} per cui $\underline{A}\underline{u}$ sia parallelo a \underline{u} , ossia sostanzialmente identico a \underline{u} se non nel modulo e nel verso? La risposta a questo problema è il problema agli autovalori: si moltiplica un vettore per un numero! Infatti, fare ciò, non modifica la direzione del vettore. Si parla degli autovettori come di “vettori invarianti rispetto alla matrice”: per essi, infatti, la direzione è mantenuta costante, identica.

Scegliamo a caso come prima componente 1, la seconda di conseguenza; si avrà qualcosa di questo genere:

I due autovettori sono linearmente indipendenti (il fatto è garantito dal momento che due vettori sono linearmente indipendenti se, nello spazio, essi non sono paralleli); sono ortogonali? Vediamo: se sono ortogonali, il loro prodotto scalare dovrebbe essere nullo!

$$\underline{u}^{(1)} \cdot \underline{u}^{(2)} = u_1^{(1)} u_1^{(2)} \left(1 - \frac{b}{a} \right)$$

se $b \neq a$, questo è nullo. Se invece $b = a$, il prodotto scalare è nullo, dunque i vettori ortogonali; questo in realtà è un caso particolare di una situazione molto più generale: dall'algebra lineare è infatti noti che se la matrice è reale simmetrica, i suoi autovettori sono ortogonali.

Si sappia che gli autovalori per una matrice sono sempre N , ma gli autovettori possono non essere nello stesso numero, in casi tuttavia piuttosto complicati (quando ci sono autovalori coincidenti, potrebbero infatti mancare alcuni autovettori, ad alcune considerazioni che tuttavia noi non considereremo).

Il fatto che gli autovettori siano ortogonali non è particolarmente interessante: interessante è, piuttosto, il semplice fatto che essi siano non paralleli, dunque linearmente indipendenti; questo è interessante dal momento che, come base, è possibile utilizzare proprio gli autovettori! Ciò permetterebbe di ottenere risultati interessanti: $\underline{u}^{(1)}$, $\underline{u}^{(2)}$ costituiscono infatti una base collegata in modo inscindibile da \underline{A} , dal momento che, se essi costituiscono la base della matrice, \underline{A} assume una forma particolare. \underline{A} , come già detto, è un'immagine dell'operatore astratto in una certa base; questa base è molto interessante, dal momento che, in quella base, la matrice assume una forma

molto particolare: si può dimostrare che l'immagine dell'operatore nella base dei suoi autovettori sia semplicemente:

$$\mathcal{A} \longleftrightarrow \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

ossia, è una matrice diagonale. Quando si ha a che fare con una matrice diagonale, il suo modo di funzionare è molto semplice: le matrici diagonali, infatti, sono la cosa che “più assomiglia ai numeri”: per esempio, il prodotto di due matrici diagonali commuta; oppure, si può vedere che, una cosa del tipo

$$\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \lambda_1 x \\ \lambda_2 y \end{bmatrix}$$

Cosa significa ciò? L'uscita dipende solo dall'operando: questa è una matrice che non “scambia le carte in tavola”: il risultato per ciascuna componente dipende solo dalla componente stessa! In altre parole, se una matrice piena descrive un accoppiamento tra le variabili x e y , una matrice diagonale no: ciascuna componente è indipendente dalle altre.

3.3.1 Uso di basi non ortogonali

Vediamo come tutto ciò debba essere applicato: si è detto che gli autovettori sono tra loro indipendenti, ma non ortogonali; volendo dunque avere il vantaggio di gestire matrici diagonali, sarà necessario esprimere i componenti dei vari vettori come scomposti in una base non ortogonale. Si consideri per esempio il vettore \underline{v} :

$$\underline{v} = c_1 \underline{u}^{(1)} + c_2 \underline{u}^{(2)}$$

c_1 e c_2 sono incognite: noi stiamo tentando di rappresentare \underline{v} nella base $(\underline{u}^{(1)}, \underline{u}^{(2)})$, e questo problema è quello di determinare i pesi che ha ciascun vettore di base nella rappresentazione del finale; fare i conti richiederebbe l'applicazione della regola del parallelogramma, e, con un po' di goniometria, il calcolo delle componenti; per il ragionamento la geometria è dunque semplice, ma per le applicazioni calcolative, una volta fissati i concetti mediante i disegni, conviene utilizzare un approccio matriciale, come si sta per fare.

Riscriviamo, per ora, l'equazione nella base (x, y) : dal momento che è in questa base che sono note le componenti dei vettori di base $(\underline{u}^{(1)}, \underline{u}^{(2)})$, essendo la matrice \underline{A} riferita al piano cartesiano, si ha:

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = c_1 \begin{bmatrix} u_1^{(1)} \\ u_2^{(1)} \end{bmatrix} + c_2 \begin{bmatrix} u_1^{(2)} \\ u_2^{(2)} \end{bmatrix} = \begin{bmatrix} u_1^{(1)} & u_1^{(2)} \\ u_2^{(1)} & u_2^{(2)} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

l'ultimo passaggio semplicemente ordina in modo diverso le equazioni; si può dire che:

$$\begin{bmatrix} u_1^{(1)} & u_1^{(2)} \\ u_2^{(1)} & u_2^{(2)} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \underline{\underline{M}} \underline{c}$$

si osservi tuttavia la forma di $\underline{\underline{M}}$: essa è, sostanzialmente, la matrice degli autovettori, ossia la matrice tale per cui ciascuna colonna è uno degli autovettori relativo a uno degli autovalori. Si ha dunque che:

$$\underline{v} = \underline{\underline{M}} \underline{c}$$

questo significa che, dal momento che \underline{c} è l'incognita, che posso risolvere il sistema lineare, trovando:

$$\underline{c} = \underline{\underline{M}}^{-1} \underline{v}$$

In riferimento a quanto si usa di solito sui semiconduttori, spiegheremo a questo punto un concetto utile a partire dall'idea di **reticolo reciproco**: come noto dalla teoria della fisica dello stato solido, un cristallo ha un reticolo diretto, rappresentabile mediante due vettori \underline{d}_1 e \underline{d}_2 (bidimensionalmente); il reticolo reciproco è costituito da vettori **biortogonali** a questi:

Questo significa, dati i vettori $\underline{d}_1, \underline{d}_2$ del reticolo diretto, che i vettori \underline{k}_1 e \underline{k}_2 del reticolo reciproco sono ortogonali a essi, e quindi:

$$\underline{k}_2 \cdot \underline{d}_1 = 0$$

$$\underline{k}_1 \cdot \underline{d}_2 = 0$$

e poi, si impone che:

$$\underline{k}_1 \cdot \underline{d}_1 = 1$$

$$\underline{k}_2 \cdot \underline{d}_2 = 1$$

ossia, si impone quanto vale il prodotto scalare tra i 2 vettori. In questo caso, si è usata una normalizzazione a 1, anche se di solito si usa 2π . Queste quattro relazioni si possono riassumere come:

$$\underline{k}_i \cdot \underline{d}_j = \delta_{ij}$$

Se è valida questa relazione (che riassume le quattro precedenti), si dice che i vettori \underline{k}_i sono biortogonali a \underline{k}_j .

A cosa serve la base reciproca? Serve esattamente a fare ciò che è stato appena fatto con il formalismo matriciale: a trovare i componenti di un vettore rispetto a una base non ortonormale.

Dato

$$\underline{v} = c_1 \underline{d}_1 + c_2 \underline{d}_2$$

dove $(\underline{d}_1, \underline{d}_2)$ è una base non ortogonale, come si fa? Beh, se la base fosse ortonormale, si avrebbe semplicemente da applicare il procedimento di Gram-Schmidt: proiettare su ciascuna componente!

$$c_1 = \underline{d}_1 \cdot \underline{v} = c_1 \underline{d}_1 \cdot \underline{d}_1 + c_2 \underline{d}_2 \cdot \underline{d}_1$$

in questo caso, però, il secondo termine non è nullo, dunque l'equazione non si semplifica: se infatti gli elementi sono non ortogonali, la proiezione di un vettore sull'altro è non nulla!

Usando invece i vettori del dominio reciproco, si ha:

$$\underline{k}_1 \cdot \underline{v} = c_1 \underline{k}_1 \cdot \underline{d}_1 + c_2 \underline{k}_2 \cdot \underline{d}_1 = c_1$$

facendo la normalizzazione come fatto, si trova proprio questo risultato. Si può dunque dire che

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} \underline{k}^{(1)} \cdot \underline{v} \\ \underline{k}^{(2)} \cdot \underline{v} \end{bmatrix}$$

Recuperiamo, a questo punto, l'equazione di prima:

$$\underline{c} = \underline{\underline{M}}^{-1} \underline{v}$$

questa, come si può vedere, risolve lo stesso problema! Riordinando ancora una volta l'equazione matriciale, si ottiene:

$$\begin{bmatrix} \underline{k}^{(1)} \cdot \underline{v} \\ \underline{k}^{(2)} \cdot \underline{v} \end{bmatrix} = \begin{bmatrix} k_1^{(1)} & k_2^{(1)} \\ k_1^{(2)} & k_2^{(2)} \end{bmatrix}$$

questo ordine deriva dal fatto che si scrivono i $\underline{k}^{(i)}$ come vettori riga, al fine di completare il calcolo $\underline{k}^{(i)} \cdot \underline{v}$ in termini di prodotto riga-colonna (osservando la matrice e svolgendo il conto si vede che è giusto). Questa matrice e $\underline{\underline{M}}$ sono imparentate!

Questo ci spiega anche come costruire la base reciproca: questa matrice è \underline{M}^{-1} . Il procedimento dunque per trovare i vettori della base reciproca, numericamente, è prendere i componenti dei vettori di base per colonne, farne la matrice \underline{M} , farne l'inversa, e leggerla per righe: la i -esima riga è il i -esimo elemento della base reciproca (in altre parole, per esempio, $\underline{k}^{(1)}$ è la prima riga della matrice risultante, \underline{M}^{-1}).

Dovessimo poi continuare a parlare della teoria dei reticoli reciproci in fisica dello stato solido, finiremmo a parlare del fatto che in realtà le varie componenti del reticolo reciproco sono rappresentabili come coefficienti della serie di Fourier del reticolo periodico, ma non si vuole arrivare a ciò.

3.3.2 Funzioni di una matrice

Si è detto che le matrici diagonali sono quanto di più vicino a un numero, in termini di come vi si può operare; ci si vuol chiedere, a questo punto, se sia possibile definire, a partire da queste matrici "speciali", la funzione di una matrice, come $e^{\underline{A}}$, o una generica $f(\underline{A})$.

Volendo organizzare una funzione in uno sviluppo in serie, si sa che, a una dimensione:

$$f(x) = f(0) + \frac{df}{dx}x + \frac{1}{2!} \frac{d^2f}{dx^2}x^2 + \dots$$

Sapendo calcolare per esempio la potenza di una variabile scalare, ci chiediamo se sia fattibile farlo per matrici. Ovviamente sì, e basta utilizzare i prodotti riga per colonna, ma esiste un'altra via. Al contrario, l'esponenziale di matrice è più complicato da vedersi, ma con la visione appena introdotta, basata sulle serie di potenze, si può vedere che:

$$e^{\underline{A}} = \underline{I} + \underline{A} + \frac{1}{2} \underline{A}^2 + \dots$$

Questa è una possibile interpretazione, ma in realtà si può fare di meglio: si può usare la teoria delle matrici diagonali.

Si era visto che un problema agli autovalori è espresso come:

$$\underline{A}\underline{u} = \lambda\underline{u}$$

e che vi sono sostanzialmente due soluzioni accettabili:

$$\underline{A}\underline{u}^{(1)} = \lambda_1\underline{u}^{(1)}$$

$$\underline{A}\underline{u}^{(2)} = \lambda_1\underline{u}^{(2)}$$

Volendo scrivere queste due equazioni come una sola, si può scrivere una matrice \underline{M} come matrice tale per cui $\underline{u}^{(1)}$ e $\underline{u}^{(2)}$ siano le sue colonne; dunque, si ha:

$$\underline{A}\underline{M} = \underline{M}\underline{\lambda}$$

Si noti che l'ordine di moltiplicazione riga per colonne al membro destro è importante e non è univoco: questo è l'unico metodo giusto dal momento che una matrice diagonale che moltiplica da destra una matrice piena, fa in modo da moltiplicare la prima colonna della matrice piena per il primo numero (il primo autovalore), la seconda colonna per il secondo numero, e così via; dal momento che il significato fisico di \underline{M} è sulla colonna, noi vogliamo che le singole colonne, ossia i singoli autovettori (per come si è appena detto che si organizza la matrice \underline{M}), siano moltiplicati per i propri autovalori, ottenendo dunque esattamente questo tipo di scrittura. Ciò si può riscrivere come:

$$\underline{A} = \underline{M}\underline{\lambda}\underline{M}^{-1}$$

Si provi a questo punto a elevare al quadrato questa matrice:

$$\begin{aligned} \underline{A}^2 &= (\underline{M}\underline{\lambda}\underline{M}^{-1}) (\underline{M}\underline{\lambda}\underline{M}^{-1}) = \\ &= \underline{M}\underline{\lambda}^2\underline{M}^{-1} \end{aligned}$$

se poi al posto di \underline{A}^2 vi fosse \underline{A}^N , si potrebbe vedere che vale esattamente la stessa cosa. Si noti, inoltre, che $\underline{\lambda}^N$ coincide con l'avere N prodotti di matrici diagonali, dunque semplicemente gli elementi di questa matrice saranno le potenze dei singoli elementi!

Facendo un salto, giustificabile con il ragionamento delle serie di Taylor prima espresse, si può giustificare (cosa per esempio fatta su una matrice 2×2), per una generica funzione f sviluppabile in serie di Taylor su un certo raggio di convergenza:

$$f(\underline{A}) = \underline{M}f(\underline{\lambda})\underline{M}^{-1} = \underline{M} \begin{bmatrix} f(\lambda_1) & 0 \\ 0 & f(\lambda_2) \end{bmatrix} \underline{M}^{-1}$$

ciò che deve stare nel cerchio di convergenza della funzione, nella fattispecie, sono $\lambda_{1,2}$, ossia i due autovalori.

Questa cosa ci servirà dal momento che, nota la matrice di trasmissione \underline{T} di una cella, potremo fare con facilità \underline{T}^N , ma, usando questo formalismo, vedremo che le onde di Bloch sono collegate a tutto ciò che è stato detto finora.

3.3.3 Problema agli autovalori generalizzato - generalizzazioni varie

Il problema agli autovalori nella forma:

$$\underline{A}u = \lambda u$$

è detto **problema standard** agli autovalori; l'aggettivo deriva dal fatto che esiste anche il **problema generalizzato** agli autovalori, che ha forma:

$$\underline{A}u = \lambda \underline{B}u$$

Il problema standard è dunque semplicemente il problema generalizzato, in cui

$$\underline{B} = \underline{I}$$

In realtà i due problemi, a certe condizioni, sono uno riconducibile all'altro: se infatti si moltiplica a sinistra per \underline{B}^{-1} , si ottiene qualcosa del tipo:

$$\underline{B}^{-1}\underline{A}u = \lambda u$$

Questo problema sembrerebbe avere la stessa struttura del precedente; in realtà le cose non sono così semplici, dal momento che non è affatto scontato (e anzi spesso non accade) che \underline{B} sia una matrice invertibile. Esistono inoltre ulteriori generalizzazioni su questo problema; esso infatti può essere scritto in forma:

$$(\lambda \underline{B} - \underline{A})u = 0$$

La quantità tra parentesi è un polinomio di grado 1 in λ , che ha come coefficiente delle matrici; nulla vieta ovviamente di avere problemi di secondo o terzo grado:

$$(\lambda^2 \underline{C} + \lambda \underline{B} - \underline{A})u = 0$$

Anche in questo caso esistono possibilità per MATLAB di risolvere problemi di questo tipo; in realtà, si può anche dimostrare che un problema agli autovalori polinomiale (come questo) può essere ricondotto a un problema di primo grado (ossia con λ solo nella prima potenza), ma di ordine doppio; nel caso di problemi "difficili", questa non è per forza una soluzione valida: non è infatti detto che le matrici in gioco siano piccole, quindi, aumentando l'ordine, il costo computazionale di ciascuna operazione potrebbe aumentare enormemente.

3.3.4 Applicazione ai sistemi lineari

A questo punto vogliamo proporre un'interpretazione di questi metodi su un problema ben noto: la soluzione di sistemi lineari. Un sistema, come problema, ha quello di trovare \underline{x} , in un'equazione del tipo

$$\underline{A}\underline{x} = \underline{n}$$

Si sa, tuttavia, che si ha la presente possibile rappresentazione, per la matrice \underline{A} :

$$\underline{A} = \underline{M}\underline{\lambda}\underline{M}^{-1}$$

quindi, sostituendo:

$$\underline{M}\underline{\lambda}\underline{M}^{-1}\underline{x} = \underline{n}$$

premultiplico per \underline{M}^{-1} entrambi i membri, ottenendo:

$$\underline{\lambda}\underline{M}^{-1}\underline{x} = \underline{M}^{-1}\underline{n}$$

a questo punto, due definizioni:

$$\underline{x}' = \underline{M}^{-1}\underline{x}$$

$$\underline{n}' = \underline{M}^{-1}\underline{n}$$

in questo caso, il problema è ricondotto a:

$$\underline{\lambda}\underline{x}' = \underline{n}'$$

Cosa abbiamo di bello? Beh, si ha ancora una volta qualcosa nella forma di un sistema lineare, dove però la matrice dei coefficienti è una matrice diagonale: in questo caso, data la proprietà precedentemente vista (il fatto che, nel caso di matrici diagonali applicate a vettori, il risultato per ciascuna componente dipende solo dalla componente stessa), non si ha neanche un sistema: solo un insieme di equazioni tra loro indipendenti. La soluzione del sistema è banale:

$$\underline{x}' = \underline{\lambda}^{-1}\underline{n}'$$

ma l'operazione di inversione di una matrice diagonale è altrettanto banale: è sufficiente calcolare il reciproco di ciascun elemento. Per completare quindi il sistema, bisogna ricordare, invertendo la precedente relazione:

$$\underline{x} = \underline{M}\underline{x}'$$

continuando a svolgere i conti:

$$\underline{M}\underline{x}' = \underline{M}\underline{\lambda}^{-1}\underline{n}' = \underline{M}\underline{\lambda}^{-1}\underline{M}^{-1}\underline{n} = \underline{A}^{-1}\underline{n}$$

È ragionevole tutto ciò? Beh, basti ricordare che, per calcolare l'inversa di una matrice data dal prodotto di più matrici, si deve calcolare l'inversa di ciascuna matrice, quindi scambiare le posizioni dei termini:

$$\underline{A}^{-1} = (\underline{M}\underline{\lambda}\underline{M}^{-1})^{-1} = (\underline{M}^{-1}\underline{\lambda}^{-1}\underline{M})^T = \underline{M}\underline{\lambda}^{-1}\underline{M}^{-1}$$

proprio come atteso.

Questa è una proposta di procedimento, ma qual è il suo significato? Precedentemente, si è visto che, dato un vettore \underline{v} , rappresentato in una certa base, esso ha componenti:

$$\underline{v} = c_1\underline{u}^{(1)} + c_2\underline{u}^{(2)}$$

al fine di trovare c_1 e c_2 , ossia le componenti del vettore rispetto alla base, si era detto che si poteva scrivere l'equazione come:

$$\underline{v} = \underline{M}\underline{c} \implies \underline{c} = \underline{M}^{-1}\underline{v}$$

In altre parole, il fatto di premoltiplicare \underline{M}^{-1} a un certo vettore, ricordando che \underline{M} è la matrice costituita dagli autovettori presi come colonne, permette di rappresentare un vettore \underline{v} rispetto alla base degli autovettori, come visto precedentemente. Considerando dunque questo specifico caso, $\underline{Ax} = \underline{n}$, si ha che \underline{n} è un vettore riferito alla “base di partenza”, la base rispetto cui siamo comodi a rappresentare il problema; dal momento che si ha

$$\underline{n}' = \underline{M}^{-1}\underline{n}$$

quello che si fa è prendere il sistema e riscriverlo secondo un'altra base: la base degli autovettori. In altre parole, quello che si fa è passare, dalla base di partenza, a una base particolare, quella degli autovettori; di tutte le basi possibili, quella degli autovettori è la base migliore, dal momento che, se si scegliesse una qualsiasi altra base, al posto di $\underline{\lambda}$ si avrebbe una certa \underline{A}' , matrice non diagonale ma **piena**. Giunti dunque nella base degli autovettori, si effettuano tutte le operazioni che si devono effettuare (si risolve il sistema), quindi si torna a \underline{x} premoltiplicando per \underline{M} , in maniera da prendere il risultato ed esprimerlo nella “base naturale”. Si può identificare quindi uno schema.

1. Si parte da un sistema, espresso secondo una certa *base naturale* (che, per esempio, potrebbe essere il piano cartesiano, (x, y)).
2. A partire dal sistema, dove il problema è formulato, si passa alla *base degli autovettori*, base nella quale le operazioni necessarie per la risoluzione del sistema sono molto più semplici da applicare.
3. Si trova, nella base degli autovettori, la soluzione del sistema.
4. Presa la soluzione, si torna indietro alla base naturale.

A questo punto una domanda un poco scorrelata dal discorso, ma utile: si fa effettivamente così, per risolvere i sistemi lineari? In realtà no: calcolare gli autovalori e gli autovettori è cosa lunga e dispendiosa, anche per un software: finché si ha a che fare con sistemi semplici, di ordine ridotto, va benissimo, ma dal momento che si ha a che fare con qualcosa di più complicato, di ordine maggiore, è una scelta assolutamente inapplicabile. Esistono in realtà dei metodi che permettono di calcolare direttamente \underline{x} , senza invertire le matrici; in effetti, l'unico caso in cui si calcola \underline{A}^{-1} (essendo l'inversione un'operazione estremamente dispendiosa), è quello per cui nel sistema si abbia a che fare con moltissimi termini noti diversi: tanti \underline{n}_i rispetto a cui risolvere il sistema; in questo caso, di \underline{A}^{-1} se ne ha una sola, e invece che applicare i metodi diretti si può calcolare una volta per tutte l'inversa, e moltiplicarla.

Se per qualche motivo già si conoscessero gli autovalori e gli autovettori, ed essi siano ortogonali (caso abbastanza comune), è possibile costruire \underline{M} e farne l'inversa (che, nel caso di vettori ortogonali, si ridurrebbe alla trasposta).

3.3.5 Esempi di applicazione del cambio di base

Verranno a questo punto presentati, in modo relativamente informale, alcuni esempi in cui si applica il ragionamento appena proposto, in modo da poter fissare i concetti prima di passare all'applicazione di reale interesse: quella sugli specchi di Bragg.

Risoluzione di un problema ai valori iniziali

Si consideri il seguente problema ai valori iniziali:

$$\begin{cases} \frac{dx}{dt} = \underline{A}x \\ x(t)|_{t=0} = x_0 \end{cases}$$

Dalla teoria, è noto che la soluzione di questa equazione è:

$$\underline{x}(t) = e^{\underline{A}t} \underline{x}_0$$

Come si calcola quindi l'esponenziale di matrice? Beh, esso è una funzione di matrice; supponendo che \underline{A} sia una matrice 2×2 , si ha che:

$$\begin{aligned} e^{\underline{A}t} &= \underline{M} e^{\underline{\lambda}t} \underline{M}^{-1} = \\ &= \underline{M} \begin{bmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{bmatrix} \underline{M}^{-1} \end{aligned}$$

Quindi, invece di fare chissà quali stranezze, semplicemente si applica ai singoli autovalori la funzione esponenziale; riscrivendo tutto:

$$\underline{x}(t) = \underline{M} \begin{bmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{bmatrix} \underline{M}^{-1} \underline{x}_0$$

Possiamo introdurre delle interpretazioni: si ha un $\underline{M}^{-1} \underline{x}_0$, ossia si trasforma la condizione iniziale, secondo la base degli autovettori; a questo punto, quindi, il problema nella base degli autovettori è:

$$\frac{dx'}{dt} = \underline{\lambda} x'$$

Questa è l'immagine del problema iniziale nella base degli autovettori; esso, semplicemente, diventa:

$$\frac{d}{dt} \begin{bmatrix} x'_1 \\ x'_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 x'_1 & 0 \\ 0 & \lambda_2 x'_2 \end{bmatrix}$$

questo, come prima, non è più un sistema: si tratta solamente di due equazioni differenziali indipendenti, la cui soluzione è immediata; moltiplicando quindi per \underline{M} , si torna nello spazio naturale, nella base naturale.

Si noti che è importante una condizione, che non è stata detta esplicitamente: \underline{A} deve essere indipendente dal tempo. Se così non fosse, autovalori e autovettori sarebbero dipendenti dal tempo. In realtà, si hanno due sottocasi, di cui uno più fortunato:

- il caso fortunato è quello di autovettori costanti rispetto al tempo, autovalori variabili: in questo caso, la tecnica appena presentata funzionerebbe ancora, dal momento che la base degli autovettori sarebbe ancora la stessa, dunque l'equazione sarebbe a coefficienti variabili, ma molto più semplice da risolvere;

- nel caso in cui anche gli autovettori siano variabili rispetto al tempo, il metodo non sarebbe più applicabile in nessuna forma: non si avrebbe più una base costante rispetto al tempo.

Sistemi LTI: Lineari Tempo Invarianti

Dato un sistema LTI in cui si introduce un segnale $x(t)$, in uscita si avrà un $y(t)$, che a prima vista non avrà rapporti con il segnale di partenza: le forme del segnale in ingresso e di quello in uscita sono diverse, tranne in un caso, ossia per una particolare categoria di segnali: le armoniche, gli esponenziali di forma $e^{j\omega t}$. Nel caso delle armoniche, variano solo modulo e fase, ma non tipo di segnale, che si mantiene costante e con la stessa frequenza.

Il procedimento per determinare l'uscita del sistema è quello di passare da $x(t)$ a $X(\omega)$, mediante la trasformata di Fourier; si ottiene quindi $Y(\omega)$ moltiplicando $X(\omega)$ per $H(\omega)$, ossia per la funzione di trasferimento del sistema; la risposta nel tempo del sistema, ossia l'espressione dell'uscita, si ottiene effettuando una trasformazione inversa, ossia la trasformata inversa di Fourier. Il procedimento, dunque, è molto simile a quello effettuato in precedenza.

Concettualmente, la trasformata di Fourier è un qualcosa di molto simile rispetto a prima; si ha che:

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} X(\omega) e^{j\omega t} d\omega$$

Questa, poi, si può tradurre in una sommatoria (l'integrale di fatto è semplicemente una sommatoria nel continuo):

$$\Rightarrow \sim \sum_i X(\omega_i) \frac{\Delta\omega}{2\pi} e^{j\omega_i t}$$

Questo, di fatto, assomiglia molto a quanto visto prima:

$$\underline{v} = c_1 \underline{u}^{(1)} + c_2 \underline{u}^{(2)}$$

In realtà dunque le due cose sono molto simili, solo che quella appena proposta è formata da infiniti termini. Gli esponenziali sono gli argomenti della somma, dal momento che essi sono gli *autovettori del sistema LTI*, essendo essi gli elementi che *passano attraverso esso* rimanendo **paralleli a sè stessi**, ossia per cui si mantiene una certa porzione di funzione (si parla infatti anche di **autofunzioni**, quando si parla di autovettori in spazi infinito-dimensionali come gli spazi di Hilbert), più un certo coefficiente moltiplicativo (che sarà l'autovalore).

Guide d'onda: formalismo modale

Un altro esempio, forse un poco più complicato dei precedenti ma non per questo inappropriato, è quello del formalismo modale che si applica sulle guide d'onda: in laboratorio, a Campi Elettromagnetici, spesso si studia una guida d'onda alimentata da un coassiale, che irradia un campo dentro la guida; di questo campo poi se ne propaga solo una certa porzione.

Un problema di questo genere, da studiare, è estremamente complicato.

Cerchiamo di inquadrare meglio il problema: la nostra base naturale è lo spazio cartesiano, dunque (x, y, z) ; l'obiettivo è calcolare il campo ad una certa distanza, in un punto (2), a partire da ciò che si ha delle sorgenti. Le sorgenti, ossia il coassiale, irradiano un certo campo; quello che si può fare è considerare i campi \underline{E} e \underline{H} in prossimità dell'antenna, in una certa sezione ben definita, (1); quello che si può fare, a partire da ciò, è trovare delle particolari funzioni per cui le funzioni rimangano le stesse, a meno di un certo coefficiente moltiplicativo; queste funzioni sono quelle che a Campi Elettromagnetici vengono chiamate $\underline{e}_n(\underline{\varrho})$ e $\underline{h}_n(\underline{\varrho})$; note su $z = 0$, per tutta la teoria studiata a Campi, si sa che, nella sezione z , esse diverranno:

$$\begin{cases} \underline{e}_n(\underline{\varrho}) \implies \underline{e}_n(\underline{\varrho})e^{-jk_z n z} \\ \underline{h}_n(\underline{\varrho}) \implies \underline{h}_n(\underline{\varrho})e^{-jk_z n z} \end{cases}$$

Queste due funzioni \underline{e} e \underline{h} sono funzioni "particolari", dal momento che sono le funzioni che, nelle varie sezioni, sono uguali a sè stesse, a meno di un certo termine di fase, un certo **numero**. Queste funzioni vengono dette **autofunzioni**, relative ai vari modi. $e^{-jk_z n z}$ è dunque l'autovalore relativo all'autovettore $\underline{e}_n(\underline{\varrho})$.

Quello appena presentato è il dominio delle linee di trasmissione equivalenti; quello che si deve fare per risolvere questo tipo di problema, quindi, è **decomporre la sorgente nei modi**, ossia prendere la sorgente \underline{J} e trasportarla nella base delle autofunzioni, quindi risolvere il problema, e tornare indietro al dominio naturale.

3.4 Specchi di Bragg

A questo punto, supponendo fissati i concetti fondamentali, si vuole proporre una trattazione, basata per l'appunto su tutte queste idee, degli specchi di Bragg. Come già anticipato in precedenza, gli specchi di Bragg sono strutture composte da strati alternati con diversi coefficienti di rifrazione, $n_1 - n_2$. Al fine di definire correttamente il problema, abbiamo una struttura per ora semplificata:

Supponiamo che la prima cella sia quella che parte con n_2 , larga uno spazio d_2 , seguita quindi da una cella n_1 larga d_1 . La semplificazione della struttura sta nel fatto che a sinistra e a destra della prima cella si ha n_1 , invece di avere uno strato di aria (a sinistra), o di dielettrico denso (a destra) come usualmente capita con strutture di questo tipo. Essendo il nostro obiettivo rappresentare ciascuna di queste celle mediante il formalismo dei doppi bipoli, è necessario definire dei piani di riferimento per la struttura. Si scelgono, come piani di riferimento, quelli dentro il dielettrico n_1 esterno alle celle: come appena detto, infatti, il dielettrico n_1 parte da $-\infty$, ci sono le varie celle, l'ultima cella parte da n_2 , prende un tratto d_1 del dielettrico finale, il quale proseguirà fino a $z \rightarrow +\infty$. I piani di riferimento sono posizionati *appena prima il primo n_2* (per la “porta 1”) e *appena prima della fine di d_1* , nel dielettrico finale; si noti che il piano di riferimento terminale rispetterà **non** l'interfaccia, bensì **la periodicità**: queste definizioni sono basate sulla periodicità, non (solo) sulla geometria del sistema. In altre parole, la porta 1 sarà in A^- , la porta 2 in C^- . Volendo dunque analizzare ciascuna cella, si avranno le varie linee di trasmissione, si ha una situazione del genere:

Per ciascuna i -esima porta, come noto dalla teoria delle linee di trasmissione, è necessario definire un'impedenza di riferimento, $Z_{r,i}$; questa, al fine di effettuare la scelta “più semplice” (che potrebbe per esempio essere quella di garantire la continuità di impedenza caratteristica con la cella precedente, essendo comunque il formalismo che useremo basato sulle onde di potenza), sarà, per entrambe le porte, $Z_{\infty 1}$, ossia l'impedenza di riferimento relativa a n_1 (essendo entrambi i piani di riferimento nel dielettrico n_1).

A questo punto, quale rappresentazione possiamo utilizzare per ciascuna cella? Beh, la più idonea, come si può immaginare, è quella della matrice di trasmissione, \underline{T} : essendo questa struttura una struttura periodica, dunque una struttura composta da N celle tutte uguali tra loro messe in cascata, è di sicuro più semplice usare questo tipo di rappresentazione rispetto a una rappresentazione scattering, per quanto essa sia comunque più vicina alla fisica del sistema. L'impedenza di riferimento per la matrice sarà, dunque, $Z_{\infty 1}$ per entrambe le porte. Come noto, quindi:

$$\begin{bmatrix} c_1^+ \\ c_1^- \end{bmatrix} = \underline{T} \begin{bmatrix} c_2^+ \\ c_2^- \end{bmatrix}$$

dove

$$c_1^+ = \frac{V_{A^-}^+}{Z_{\infty 1}} \quad c_1^- = \frac{V_{A^-}^-}{Z_{\infty 1}} \quad c_2^+ = \frac{V_{C^-}^+}{Z_{\infty 1}} \quad c_2^- = \frac{V_{A^-}^-}{Z_{\infty 1}}$$

A questo punto, un'osservazione finale: essendovi N celle, la matrice “completa”, ossia la matrice del sistema risultante dalla cascata delle varie celle, sarà:

$$\underline{\underline{T}}_{\text{completa}} = \underline{\underline{T}}^N$$

Tuttavia, è stato visto che ciò può anche essere scritto, passando alla base degli autovettori al fine di rendere il problema più semplice, come:

$$\underline{\underline{T}}_{\text{completa}} = \underline{\underline{M}}\underline{\underline{\lambda}}\underline{\underline{M}}^{-1}$$

Al fine dunque di passare alla base degli autovettori, riducendo il problema “difficile” a un problema molto semplice, è necessario trovare gli autovettori della matrice di trasmissione $\underline{\underline{T}}$ e i suoi autovalori.

Come al solito, si partirà dagli autovalori, risolvendo l'equazione caratteristica per la matrice 2×2 (si ricordi che infatti si ha a che fare con dei doppi bipoli, quindi la matrice sarà evidentemente di questa forma):

$$\lambda^2 - \text{Tr} \{ \underline{\underline{T}} \} + \det \{ \underline{\underline{T}} \} = 0$$

A questo punto, ci viene in soccorso una proprietà molto importante: si può dimostrare che $\underline{\underline{T}}$ ha determinante unitario, pari a 1; ciò deriva dal fatto che la struttura è **reciproca**. Infatti, in quel caso, nella matrice $\underline{\underline{S}}$ si può dimostrare che $S_{12} = S_{21}$, e imponendo ciò nella conversione da matrice scattering a matrice di trasmissione, si può vedere che il vincolo è questo.

Usando dunque la formula tradizionale per la soluzione di equazioni del secondo ordine, si può ricavare banalmente che:

$$\lambda = \frac{1}{2} \text{Tr} \{ \underline{\underline{T}} \} \pm \sqrt{\left(\frac{\text{Tr} \{ \underline{\underline{T}} \}}{2} \right)^2 - 1}$$

Che dà luogo, come atteso, a **due autovalori**.

A questo punto, si ricordi per un momento qual è l'obiettivo preposto: l'operazione che si vuole studiare, mediante questo formalismo, è la propagazione attraverso una struttura (nella fattispecie periodica); dal momento che la base degli autovettori è la base per cui si hanno le *cose semplici*, il nostro obiettivo è avere la propagazione più semplice possibile: passando da un punto a un altro della struttura periodica, intesa come insieme di celle in cascata tra loro, si vuole avere la sola moltiplicazione per un esponenziale (in modo da ricondurci alla forma e^{-jkz}). C'è dunque modo di ricondurre la scrittura appena proposta in modo che ricordi un esponenziale? Prima di tutto, si raccolga un -1 all'interno della radice, ottenendo:

$$\lambda = \frac{1}{2} \text{Tr} \{ \underline{\underline{T}} \} \pm j \sqrt{\left(1 - \frac{\text{Tr} \{ \underline{\underline{T}} \}}{2} \right)^2}$$

A questo punto, abbiamo due espressioni; se diciamo che:

$$\frac{1}{2} \text{Tr} \triangleq \cos \phi$$

ossia, se imponiamo che questo termine sia uguale al coseno di un certo angolo ϕ , essendo l'altro termine la radice di 1, cui si sottrae lo stesso numero al quadrato, essa è semplicemente l'applicazione della formula fondamentale della goniometria:

$$\pm j \sqrt{\left(1 - \frac{\text{Tr} \{ \underline{\underline{T}} \}}{2} \right)^2} = \pm j \sqrt{1 - \cos^2 \phi} = \pm j \sin \phi$$

Questo significa che si ha:

$$\lambda = \cos \phi \pm j \sin \phi = e^{\pm j \phi}$$

dove l'ultimo passaggio è semplicemente una banale applicazione della formula di Eulero. Invertendo l'espressione del coseno, è possibile anche trovare un'espressione di ϕ :

$$\phi = \arccos \left(\frac{T_{11} + T_{22}}{2} \right)$$

essendo la traccia della matrice $\underline{\underline{T}}$ la somma degli elementi sulla diagonale principale, dunque la somma degli elementi che hanno pedici uguali tra loro. Infine:

$$\lambda = e^{\pm j \phi}$$

sono i due autovalori del nostro problema, dato ϕ calcolato con la formula precedente.

A questo punto abbiamo risolto il problema del calcolo degli autovalori; quali sono, quindi, gli autovettori? Beh, noti gli autovalori, si devono sostituire dentro al problema standard

$$\underline{\underline{T}} \underline{u} = \lambda \underline{u}$$

Cos'è \underline{u} ? Al fine di capirlo, è necessario riprendere la definizione di matrice di trasmissione, dunque il linguaggio della teoria dei circuiti: i vettori \underline{u} conterranno delle onde di potenza, ossia

$$\underline{T} \begin{bmatrix} c^+ \\ c^- \end{bmatrix} = \lambda \begin{bmatrix} c^+ \\ c^- \end{bmatrix} \quad \lambda = e^{\pm j\phi}$$

A questo punto, vedendo che negli autovettori \underline{u} appaiono le onde progressive e regressive, possiamo proporre un'interpretazione di tutto ciò: gli autovettori \underline{u} sono quelle combinazioni di onda progressiva e regressiva che, operate con \underline{T} (matrice costituente l'immagine dell'operatore nella base naturale delle onde progressive e regressive, descrive la "trasmissione" dello stato elettrico attraverso il blocco), siano "parallele" in ingresso e in uscita dal blocco: mantengono la propria forma.

La base naturale, costituita da (c^+, c^-) , non va bene: se si usassero le onde di potenza nel senso tradizionale, infatti, si avrebbero fenomeni di scattering multiplo ai vari punti dell'interfaccia; questo, complicherebbe enormemente i conti. La "base furba", quella degli autovettori, è quella che passa attraverso il blocco, subendo solo la moltiplicazione per un numero, per l'autovalore (il quale è stato ricondotto a uno sfasamento): questa è a tutti gli effetti la definizione di autovettore, applicata nella nostra situazione. ϕ è lo **sfasamento per cella**: i radianti di sfasamento che si hanno dall'ingresso all'uscita della cella.

A questo punto, si vuole fare una serie di osservazioni su quanto finora detto: si ha a che fare con un modello per cui, di fatto, non si hanno "attenuazioni per disadattamento" delle onde di potenza da una cella a un'altra: ciò sembrerebbe strano. Non è in realtà la prima volta che si ha a che fare con interpretazioni "particolari" e apparentemente poco intuitive come questa: si pensi per esempio al modello discretizzato di un cavo coassiale, ossia di una linea di trasmissione, rappresentabile mediante una rete a scala

Nel caso delle onde di potenza, il modulo della tensione rimane senza dubbio costante in ogni punto della linea, ma ciò non è assolutamente banale, considerando questo tipo di modello: dal momento che sull'induttanza per esempio vi è una corrente, su di essa vi sarà una qualche caduta di tensione, eppure ciò non è evidente dal modello a parametri distribuiti, il quale, a certe condizioni, fornirà in realtà lo stesso risultato di questo modello: è necessario prestare attenzione alle interpretazioni "a occhio" con i vari modelli, dal momento che si rischia di fare confusione o di proporre risultati sbagliati ancora prima di aver fatto i conti. Cosa simile capita con le onde appena introdotte: dal modello sembrerebbero essere evidenti fenomeni di scattering, che in realtà non ci sono, utilizzando le suddette onde al posto di quelle di potenza.

Nel caso dell'equivalenza linea-rete a scala, tutto ciò accade, ed è verificato, solo se la corrente che passa nel circuito, I , è:

$$I = \frac{V}{Z_\infty}$$

in altre parole, solo a queste condizioni, questo V - I è un autovettore: trovare l'autovettore è trovare il rapporto tra V e I , ossia l'impedenza caratteristica. L'impedenza caratteristica del cavo coassiale è la maniera per specificare gli autovettori.

Torniamo al discorso sulla nostra cella: quanto è lunga la cella in questione? Beh, banalmente, si può dire che sia lunga un certo d , dove

$$d = d_1 + d_2$$

si noti che qua non si sta per ora mettendo in relazione ciò con la lunghezza d'onda, dal momento che stiamo studiando una struttura del tutto generica. A questo punto, verrebbe la tentazione di calcolare il seguente rapporto:

$$\frac{\phi}{d} = k_B$$

È stata immediatamente data la lettera k per identificare questo parametro: k_B farebbe pensare alla costante di Boltzmann, ma in realtà questa osservazione è totalmente fuori luogo: la lettera k identifica il fatto che questo parametro, dimensionalmente, è radianti su metri, dunque è a tutti gli effetti una costante di propagazione. B dunque non sta per Boltzmann, ma per Bloch: questa è una sorta di costante di propagazione relativa alle onde di Bloch. Un altro nome importante nell'ambito è quello di Floquet, il matematico che ha studiato le equazioni differenziali a coefficienti periodici; Bloch sarebbe il ricercatore che ha generalizzato il lavoro di Floquet, alle equazioni di Schroedinger, alle derivate parziali.

Gli autovettori che abbiamo trovato sono le onde di Bloch, e hanno quella costante di propagazione: sono onde, come nel caso del coassiale, che si propagano senza scattering nella struttura, ma invece di propagarsi in un coassiale, si propagano nella struttura periodica. Bisogna prendere tutto ciò con le molle, compresa la definizione di k_B : le onde di Bloch hanno una caratteristica molto diversa dalle onde tradizionali, dal momento che quelle tradizionali sono definite in punti vicinissimi, con continuità lungo l'asse z ; nel caso delle onde di Bloch, si ha qualcosa di profondamente diverso: si ha una cascata di scatole nere, di doppi bipoli: le onde di Bloch sono definite solo **ai terminali**. Dire che lo sfasamento per cella ϕ è di 30° , per esempio, non permette di dire che “a metà cella” si abbia $\phi_{1/2} = 15^\circ$: sapere quanto vale la costante di propagazione di Bloch non è legato alla propagazione su un tratto di linea equivalente interno alla scatola nera, bensì solo alla struttura

periodica. k_B è una specie di costante di costante di propagazione “media”, ma non calcolabile immediatamente a partire dalle caratteristiche della linea (se non da T_{11} e T_{22}). L’**unica** maniera è passare da ϕ , quindi dalle onde di Bloch, le quali sono una combinazione delle onde di potenza.

Queste onde di Bloch sono dunque onde “discrete”, che sono quindi caratterizzate da dei valori solo ai confini della cella: si tratta, in sostanza, come di avere una successione di doppi bipoli, “accessibili” solo dai loro terminali: non si hanno informazioni di cosa stia dentro. Si può pensare, quindi, che se le onde di potenza sono definibili in ogni punto, queste onde di Bloch siano un po’ una “versione campionata” delle onde di potenza, nello spazio: solo in alcune posizioni spaziali è possibile ottenerne il valore.

Si ricorda a questo punto che:

$$\phi = \arccos\left(\frac{T_{11} + T_{22}}{2}\right)$$

Lo sfasamento progressivo dipende da n_1 e n_2 ; essi, tuttavia, sono funzioni della frequenza, dunque di ξ , il quale è calcolabile come:

$$\xi = k_0 n_1 \sin \vartheta_1$$

Oltre che dei parametri “fissi” ϑ_1 , n_1 , n_2 , questi coefficienti T_{11} e T_{22} presentano queste dipendenze:

$$T_{11} = T_{11}(\omega, \xi)$$

$$T_{22} = T_{22}(\omega, \xi)$$

Gli autovettori del sistema, come già detto, sono quelli che soddisfano la solita equazione agli autovettori:

$$\underline{T} \begin{bmatrix} c^+ \\ c^- \end{bmatrix} = \lambda \begin{bmatrix} c^+ \\ c^- \end{bmatrix}$$

dove la prima è la componente dell’ampiezza dell’onda viaggiante “verso destra”, ossia nel verso delle z crescenti (c^+), mentre la seconda è la componente delle ampiezze viaggianti “verso sinistra”, ossia nel verso delle z decrescenti; nell’ambito delle onde di potenza, si sarebbe detto che queste, semplicemente, sarebbero rispettivamente le onde progressiva (c^+) e regressiva (c^-).

Come mai non sono stati usati i termini “progressiva” e “regressiva” ? Si sa che:

$$\frac{c^-}{c^+} = \Gamma$$

Γ dovrebbe essere un parametro di caratterizzazione della linea di trasmissione, o no? La risposta è tendenzialmente no: Γ è un parametro che può dare informazioni riguardo al comportamento della linea, considerando tuttavia di utilizzare, come base per la rappresentazione dello stato elettrico, quella delle onde di potenza. In questo caso, supponendo che il rapporto delle componenti sia quello delle componenti delle onde di Bloch, Γ non è il coefficiente di riflessione nel senso finora considerato, bensì è un Γ_B : un “Gamma di Bloch”. Questo è sostanzialmente un parametro che permette di ricordare qual è la composizione dell’autovettore: le onde di Bloch, infatti, sono date da una mistura delle onde progressive e regressive, ma dunque quello che ci serve, per caratterizzare le onde di Bloch, non è tanto il valore assoluto dei coefficienti, quanto il rapporto dei loro pesi: proprio questo Γ_B ! La norma di un autovettore infatti di solito non è definita, dal momento che è scelta arbitrariamente: ciò che è imposto “dal problema” è solo questo rapporto delle componenti, che è ciò che fornisce una direzione particolare all’autovettore. Quando $\Gamma_B = 1$, significa che i coefficienti di combinazione lineare, di “mistura” delle onde progressive e regressive, sono uguali: le onde progressiva e regressiva sono presenti nella stessa quantità, e con lo stesso segno.

Si noti che si ha, come autovalori, $e^{\pm j\phi}$: si ha un doppio segno, per lo sfasamento, che sembrerebbe essere ciò che discrimina onda di Bloch progressiva e onda di Bloch regressiva; questo in effetti è vero, ma non nel senso che ci si potrebbe attendere: non è infatti detto che $e^{-j\phi}$ sia l’onda di Bloch progressiva, a priori, semplicemente guardando il segno. Guardare il segno non è sufficiente, dal momento che gli sfasamenti ϕ sono definiti solo per passi discreti: quando si dispone solo di un valore puntuale, senza sapere cosa capitino nell’intorno, dal momento che la fase non è definita in modo univoco (uno sfasamento di 90° in anticipo o di 270° in ritardo, sono assolutamente coincidenti, nel caso non si sappia quale sia, “nel continuo” l’informazione sugli sfasamenti), è necessario ricorrere ad altri trucchi, per capire quale sia, dei due contributi di sfasamento, quello relativo all’onda progressiva.

Ciò che si deve fare, per determinare l’onda progressiva, è calcolare la potenza: se si trova che l’autovalore relativo a $+\phi$ porta ad avere potenza positiva verso destra, ossia al crescere delle z , allora esso sarà quello relativo all’onda progressiva; l’altro, sarà quello relativo all’onda regressiva: è il segno della potenza dunque che permette di determinare se un’onda sia progressiva o regressiva. Ciò non era necessario nelle onde di potenza, dal momento che

lo sfasamento era osservabile per ogni z , quindi ulteriori misure non erano di fatto importanti.

Come si può procedere? Si supponga di avere a che fare con la seguente situazione:

$$\underline{T} \begin{bmatrix} c_1^+ \\ c_1^- \end{bmatrix} = \lambda_1 \begin{bmatrix} c^+ \\ c^- \end{bmatrix}$$

Cosa significa tutto ciò? c_1^\pm rappresentano le porzioni di onde di potenza progressiva (+) e regressiva (-) presenti nelle onde di Bloch; questi sono gli autovettori di \underline{T} relativi all'autovalore λ_1 (potrebbe, per esempio, essere $e^{-j\phi}$); se si calcola dunque:

$$P_1 = \frac{1}{2} |c_1^+|^2 - \frac{1}{2} |c_1^-|^2$$

questa è la potenza, definita positiva se viaggiante da sinistra verso destra; nel caso P_1 abbia valore positivo, allora l'autovettore relativo all'autovalore λ_1 è effettivamente l'onda di Bloch progressiva; altrimenti, è quella regressiva. Si noti inoltre che la struttura è simmetrica per riflessione; per questo motivo, si ha che l'onda di Bloch progressiva e quella regressiva hanno potenza uguale in modulo, opposta in segno (la somma delle due è nulla); conoscendo dunque il rapporto (il Γ_B , volendolo chiamare così) per l'onda di Bloch progressiva, è noto che, date strutture di questo tipo, il Γ_B dell'altra onda sarà il reciproco di questo: si scambiano semplicemente gli indici dell'autovettore.

Il criterio della potenza è sostanzialmente collegato con ciò che si può ottenere studiando il contributo delle onde in termini di componenti progressive e regressive: se è più grande c^+ , allora si avrà che l'onda di Bloch si sposta sostanzialmente verso destra; nell'altro caso, se ossia c^- è prevalente, si avrà il contrario; questo coincide con il dire che Γ_B può essere maggiore di 1: la componente c^- , può prevalere sulla c^+ . Questo è ragionevole, dal momento che non si parla più di onde di potenza, ma di onde di Bloch: essendo la base sostanzialmente diversa, si può pensare che i risultati precedenti possano non valere. Questi risultati, nel dettaglio, sono basati sulla seguente ipotesi: l'onda progressiva, in un modo o nell'altro, è un termine ondulatorio che consideriamo sempre prodotto da un generatore impostato dall'utente; se tutti i blocchi sono passivi, si può dimostrare che $|\Gamma| \leq 1$, quindi l'onda regressiva non potrà mai essere, a questa ipotesi, maggiore di quella regressiva; nel caso delle onde di Bloch, tuttavia, conosciamo meno cose, quindi, essendo un'onda di Bloch un'entità che non deve soddisfare il vincolo di cui si è appena fatto cenno, bensì una semplice soluzione omogenea di un problema elettromagnetico, non sappiamo cosa dia vita all'onda con precisione, quindi il discorso salta.

Come già visto per le onde di potenza, anche per le onde di Bloch potrebbe capitare che esse non trasportino potenza attiva: possono esistere, infatti, **onde di Bloch evanescenti**. In questo caso, capire quale sia l'onda progressiva, è molto più semplice: nel caso delle onde di Bloch evanescenti, si può dire con certezza che l'onda progressiva sia quella decrescente nel verso crescente delle z : al solito, non è ammesso il fatto di avere onde che divergano; per quanto dunque i punti siano sempre "campionati", in questo caso è possibile discriminare un comportamento, e, valutando le ampiezze nei singoli punti, è sufficiente vedere quale andamento garantisce un'onda che si attenua al crescere di z .

Prima di procedere con altre osservazioni, una osservazione finale: questa tecnica, questa teoria, vale sostanzialmente per strutture semplici come quelle analizzate, ma anche più complicate: avendo una struttura per esempio con una cella composta da 8 dielettrici, l'unica cosa che cambia è la matrice di trasmissione \underline{T} ; a partire da essa, il resto della teoria sarà esattamente identica.

3.4.1 Curve di dispersione per le onde di Bloch

Si consideri a questo punto un particolare problema: quello di polarizzazione TE, con $\vartheta_i = 0$ (ossia, incidenza normale). Dai risultati¹ noti, considerando il caso semplificato, si ottiene, ricordando che²:

$$\phi = \arccos\left(\frac{T_{11} + T_{22}}{2}\right) \implies \cos \phi = \frac{1}{2}(T_{11} + T_{22}) \sin(k_{z1}d_1) \sin(k_{z2}d_2)$$

$$\cos \phi = \cos(k_1d_1) \cos(k_2d_2) - \frac{1}{2}\left(\frac{n_1}{n_2} - \frac{n_2}{n_1}\right)$$

Dunque, chiamato *RHS* il membro destro, si ha:

$$\cos \phi = RHS$$

Si può dimostrare che si ha a che fare con un grafico di questo tipo:

Se infatti $f \rightarrow 0$, si ha che i coseni tendono a 1, i seni tendono a 0; al crescere di f , i coseni diminuiscono, i seni aumentano, e, dal momento che

$$\frac{1}{2}\left(\frac{n_1}{n_2} + \frac{n_2}{n_1}\right) > 1$$

¹vedi Yeh P. - Optical Waves in Layered Media - p. 122 e p. 125

²dove, sul libro prima citato, $d_1 = a$; $d_2 = b$; $T_{11} = A$; $T_{22} = D$; $\phi = K\Lambda$: $K = k_B$, $\Lambda = d = d_1 + d_2$

(dal momento che la somma di un numero più il suo inverso è sempre maggiore di 2, se ciascun numero è maggiore o uguale a 1), come il primo termine diminuisce, il secondo aumenta, e così via; quando il secondo termine finisce per prevalere sul primo, si ha che le ondulazioni arrivano sotto il -1 . Detto ciò, il termine, per periodicità, tende a ricrescere, cresce e arriva fin sopra a $+1$, e così via continua.

Questa curva rappresenta il **coseno** di ϕ ; quello che si dovrebbe fare, a questo punto, è determinare in qualche modo l'arcocoseno; ciò si può fare graficamente, usando il seguente stratagemma:

Un metodo grafico per fare l'arcocoseno è disegnare di fianco una circonferenza goniometrica, quindi impostare un riferimento per l'angolo. Come noto dalla goniometria, l'angolo per cui il coseno è unitario è 0; si può dunque dire che la situazione $f = 0$ (o $\omega = 0$), per cui si ha $\cos \phi = 1$, sia proprio quelli per cui $\phi = 0$: gli angoli dunque, per questo motivo, andranno presi rispetto alla verticale, essendo la verticale la situazione per cui il coseno è unitario. Andando avanti, la curva prosegue, e si possono trovare vari valori di ϕ . Prima di proseguire, una nota: come visto, esistono valori di ω per cui $|\cos \phi| > 1$: questo porta ad avere una situazione sostanzialmente simile a quella che si aveva nella legge di Snell: le soluzioni esistono, e si possono vedere come “usciti dal foglio”. essendo esse su un piano normale a esso, piano rappresentante la frequenza e la parte immaginaria di $\phi(\omega)$. Si userà questo stratagemma: anche in questo caso si avranno delle “iperboli”, che “vanno a braccetto” con la circonferenza.

A noi interessa, sostanzialmente, vedere come lo sfasamento progressivo ϕ vari al variare della frequenza f , o comunque della pulsazione ω ; si ha ciò:

Questo disegno è tradizionalmente fatto con ω sulle ordinate e ϕ sulle ascisse, per quanto, si vuole ricordare, ω sia la variabile indipendente.

Dire che $\cos \phi$ non possa essere maggiore di 1 è un errore concettuale: ϕ , in effetti, è **la variabile dipendente, non la variabile indipendente**: se ϕ fosse la variabile indipendente, in effetti non sarebbe possibile far assumere al coseno un valore maggiore di 1; questa è tuttavia una situazione diversa, in cui ϕ si deve “adattare” a ciò che capita, quindi è possibile anche questa situazione, tenendo conto delle “iperboli” per la presenza delle soluzioni di soluzioni in campo complesso. A parte questo commento preliminare, come si vede, partendo da $\phi = 0$, esso aumenta, fino ad arrivare alla pulsazione ω_1 , per cui si passa dall'avere ϕ reale e pari a π a ϕ immaginario; questi punti di transizione sono detti **band edge**. Da qui, come si vede nel disegno, si ha un andamento della funzione che, per correttezza, dovrebbe essere disegnato, tridimensionalmente, uscente dal foglio, in modo da indicare il fatto che è sull'asse immaginario; si immagini di ruotare l'asse fino a farlo coincidere con quello reale: tanto il numero può essere o solo reale o solo complesso, quindi

la cosa non dà luogo a sovrapposizioni. Ci sarà una certa frequenza per cui questa parte immaginaria sarà massima, quindi decrescerà fino a ω_2 , dove si annullerà; ω_2 è un altro band edge, quindi all'aumentare di ω aumenterà ancora ϕ , fino ad arrivare a 2π a ω_3 , quindi si avrà una parte immaginaria e così via.

La zona tra $\omega = 0$ a $\omega = \omega_1$, quella da ω_2 a ω_3 e simili, dove l'onda di Bloch è sopra taglio, ossia viaggia, è detta **banda passante**. Le zone in cui $\omega \in [\omega_1, \omega_2]$ e così via, dove l'onda di Bloch non è viaggiante, sono dette "bande attenuate", ossia bande dove ϕ è immaginario, e l'onda di Bloch è sotto taglio; in quest'ultima situazione, l'esponenziale assume un certo valore minore di 1, tale per cui si ha un'attenuazione del modulo dell'onda.

Questa curva rappresenta l'andamento di k , espresso in termini di ϕ (ossia dello sfasamento), funzione di ω ; per questo motivo, questa è una curva di dispersione, come visto precedentemente.

Un'osservazione finale: si ha un'onda sotto taglio, ma non si ha dissipazione: come mai le onde, dunque, si attenuano, facendo attenuare la potenza? Beh, si ricordi che le onde di Bloch sono costituite da combinazioni lineari di onde di potenza, quindi si possono pensare a questi fenomeni di attenuazione come derivanti dagli effetti di riflessioni multiple che si hanno tra le varie pareti: si può sostanzialmente pensare a un'ennesima applicazione del principio di interferenza sulle onde di potenza, che si combinano, in certe condizioni di frequenza, in modo da far ridurre l'ampiezza complessiva.

3.4.2 Cenni al comportamento delle curve di dispersione per $\vartheta_i \neq 0$

Vogliamo a questo punto capire cosa capiti, almeno a livello qualitativo, se $\vartheta_i \neq 0$; in realtà, come variabile, si userà ξ , che è notoriamente funzione di ϑ_i , al posto dell'angolo; in questo caso, si può vedere che si ha, per la polarizzazione TE, un grafico di questo tipo:

Le regioni "annerite" sono bande passanti, mentre le "fascette bianche" sono bande attenuate. Questo grafico³ è sostanzialmente legato al precedente: per $\xi = 0$ (condizione per cui $\vartheta_i = 0$), infatti, si vede la proiezione delle varie fasce sull'asse ω . Questo disegno potrebbe essere "montato" su un disegno per cui si ha un asse uscente dal foglio, che sarebbe l'asse $\phi(\omega)$; il grafico è semplicemente la proiezione di tutti i punti $\phi(\omega)$ su questo piano: le uniche zone in cui non si ha nulla, sono quelle di banda attenuata, ossia le zone per cui la parte reale di ϕ è nulla. Se $\xi \neq 0$, si hanno situazioni nuove rispetto a quelle note: quelle che prima erano ω_1 e ω_2 , i "band edge",

³ ξ in questo caso è anche chiamato β

al variare dell'angolo si spostano "più in su": questo indica il fatto che il comportamento è dispersivo in frequenza. La cosa è evidente dal momento che questo grafico non va letto tagliando i vari valori con rette orizzontali o verticali, ma con funzioni $\beta(\omega)$; esse hanno una forma del tipo:

$$\beta = \frac{\omega}{c} n_1 \sin \vartheta_i$$

tutti i parametri tranne ω sono fissi, quindi questa è una funzione in una sola variabile; essa è una **retta**, con coefficiente angolare dipendente da ϑ_i mediante un seno; questo significa che, variando l'angolo, è possibile che la retta (non parallela agli assi quindi) intersechi buona parte della regione attenuata del grafico, aumentando di fatto l'ampiezza della stop band del grafico.

Per frequenze molto basse, inoltre, si ha solo una banda attenuata, contrariamente a ciò che si aveva per $\xi = 0$.

Questo era il caso TE; nel caso TM, si ha qualcosa di molto simile, ma con una interessante differenza:

I confini delle bande attenuate, in questo caso, si intrecciano: si può dunque avere punti per cui la banda attenuata è nulla, per certi valori di ξ , dunque di ϑ_i . Man mano che si considerano bande attenuate sempre più elevate, si hanno bande con punti di intreccio diversi, sempre più alti; questi punti, inoltre, si adagiano perfettamente su una retta tratteggiata. Questa retta, si può vedere, è scrivibile come:

$$\xi = \frac{\omega}{c} n_1 \sin \vartheta_B$$

dove, in questo caso, ϑ_B è l'angolo di Brewster: l'equazione è quella di una retta nel piano (ϑ, ω) ; all'angolo di Brewster, inoltre, quello che capita è il fatto che l'impedenza caratteristica della linea di ingresso era uguale a quella della linea dopo, ottenendo adattamento: coefficiente di riflessione nullo. In questo caso è la stessa cosa, dove invece di avere in mezzo le onde di potenza, si hanno semplicemente quelle di Bloch: è tutto uguale a prima, in cui però le onde di Bloch coincidono con quelle di potenza: non c'è scattering, e il mezzo è **eletttricamente omogeneo**. Dal momento che le onde di Bloch coincidono con le onde di potenza, e che esse non subiscono riflessione, non può nemmeno esserci la banda attenuata: l'ampiezza della banda attenuata che tende a 0 è dunque legata al fatto che per quella frequenza il mezzo è omogeneo.

Cenno all'approccio alternativo: teorema di Floquet

Un approccio alternativo per il calcolo di tutto ciò che è stato proposto è basato sul teorema di Floquet, ossia il già citato teorema che permette di risolvere equazioni differenziali a coefficienti variabili ma con periodicità.

Il teorema di Floquet afferma che la soluzione di un'equazione differenziale a coefficienti costanti è:

$$E(z) = e^{-jk_B z} E_p(z)$$

dove

$$E_p(z + d) = E_p(z)$$

ossia, dove questa è una funzione periodica nel dominio z di periodo spaziale $d = d_1 + d_2$.

Questo approccio è alternativo a quello precedentemente utilizzato, basato sui doppi bipoli. Si ha dunque una funzione periodica, per un esponenziale dipendente dal k di Bloch, per z : in questo caso, dunque, si ha a che fare con una z che può variare con continuità, non avendo quindi più onde che sono definite solo sui morsetti delle “scatole nere”, ma non al loro interno. In questo caso le celle sono “perfettamente accessibili”.

3.4.3 Coefficienti di riflessione della struttura

Qual è il collegamento tra i grafici prima mostrati, e quelli mostrati a priori dell'introduzione del formalismo matematico, riguardanti il coefficiente di riflessione della struttura? Si era visto che:

Volendo per esempio progettare la struttura con un $\Gamma_{A-} = 0,99$, si può fare ciò. Imponendo (non è strettamente obbligatorio, ma è la scelta ottima)

$$d_1 = \frac{\lambda_{g1}}{4}$$

$$d_2 = \frac{\lambda_{g2}}{4}$$

si sceglie un certo N , numero di celle, in modo da avere, con i dielettrici dati, il coefficiente di riflessione massimo desiderato.

Come è collegato tutto ciò alla teoria prima vista? Al fine di utilizzare un riferimento diretto alla prima curva, si immagini che $\vartheta_i = 0$ (poi, ovviamente, con un calcolatore si possono fare i conti esatti): la ω_0 è la pulsazione per cui si è “in mezzo alla prima banda attenuata”: essa è dunque, la zona in cui la banda attenuata presenta la massima parte immaginaria per ϕ . In

questa situazione, dunque, l'onda incidente sulla struttura periodica eccita onde di Bloch evanescenti in essa, essendo ϕ immaginario. Dal momento che le discontinuità sulla regione periodica sono 2 (infatti, prima bisogna accedere alla struttura da sinistra, ma poi, essendo la struttura periodica finita, vi sarà, a “destra” di essa, un'ulteriore zona in cui si avrà il vuoto, dunque una discontinuità per le onde di Bloch), si hanno due onde di Bloch evanescenti nella struttura: quella progressiva e quella regressiva; in questo caso, quindi, si ha, per le onde di Bloch, la FTR (riflessione totale frustrata), e parte della potenza passa dall'altra parte dello specchio.

Al crescere della frequenza, quindi, si esce dalla banda attenuata: per certe frequenze, addirittura, arriva ad avere $\Gamma_{A-} = 0$. Cosa significa tutto ciò? Beh, come detto, nella struttura quelle che si hanno sono onde di Bloch; queste situazioni per cui il coefficiente di riflessione si annulla, sono quelle per cui

$$\phi_{\text{totale}} = n\pi$$

ossia, per cui lo sfasamento **complessivo** introdotto dalla struttura periodica è multiplo di π . Lo sfasamento da considerare è quello quindi sulle N celle, e sarà:

$$\phi_{\text{totale}} = N\phi$$

quindi, quando $N\phi = n\pi$, si ha $\Gamma = 0$.

Alla frequenza $3f_0$, quindi, si avrà una successiva zona di riflessione; come mai a $3f_0$ e non a $2f_0$? La risposta è abbastanza semplice: se le strutture sono state progettate come si deve, ossia con $d_1 = \frac{\lambda_{g1}}{4}$, d_2 idem, si ha la cosiddetta **struttura $\lambda/4$ stack**. In questo caso, per $f = 2f_0$, si ha che si raddoppia la frequenza, e gli strati, la cui lunghezza è fissa, sono tali da avere lunghezze elettriche pari a $\lambda_g/2$. Come noto da quanto studiato precedentemente, gli strati a mezz'onda sono invisibili, trasparenti, quindi di sicuro non possono riflettere, dal momento che il mezzo elettrico appare come elettricamente omogeneo. Si noti che questa cosa non è per niente banale: il fatto che raddoppiando la frequenza (e dunque dimezzando λ) si abbia un dimezzamento anche delle lunghezze d'onda guidata, non è assolutamente banale. Questo deriva non da proprietà intrinseche della struttura, bensì dalla rappresentazione utilizzata per rappresentare il campo incidente: quella basata su ϑ_i ; infatti:

$$k_z = \frac{\omega}{c} n_1 \cos \vartheta_i$$

se si usasse un'espressione (peraltro più formale, per quanto tradizionalmente non utilizzata) del tipo:

$$k_z = \sqrt{k_0^2 n_1^2 - \xi^2}$$

in questo caso, raddoppiando la frequenza non si raddoppierebbe la λ_g ; dunque, questa, è tutta questione di **rappresentazione**.

Si noti inoltre che il legame tra band edge e stop band non è evidente: studiando con le espressioni esatte l'andamento della riflettività, si vede che ci sono situazioni in cui esistono punti nei quali, sebbene la parte immaginaria sia nulla, si ha una riflessione elevatissima; questo, dipende dal fatto che in realtà la riflessione è dettata da due cause:

- da un lato, la presenza di onde di Bloch evanescenti, per le quali dunque si ha un fenomeno di riflessione totale all'ingresso della struttura;
- se tuttavia la struttura ha costanti di propagazione con parte immaginaria nulla, il punto precedente salta; quello che si ha, quindi, è un fenomeno diverso: un **disadattamento di impedenza**; questo è il secondo meccanismo che partecipa all'aumento della riflettività della struttura.

La situazione, come si era già discusso precedentemente (considerando uno strato centrale invece di uno specchio di Bragg), è simile a quella per cui si ha incidenza da angolo di Brewster, ma con una fondamentale differenza: se l'angolo di Brewster funziona per ogni frequenza, in questo caso il fenomeno è puramente legato a fenomeni di risonanza; fenomenologie diverse, ma il succo non cambia: la banda attenuata continua a non esserci, a $2f_0$.

Al fine di mostrare il comportamento della struttura in una casistica diversa, si analizzino le seguenti figure:

In questo caso, $a = b$; questo significa che, per nessuna frequenza, si hanno spessori entrambi pari a $\lambda_g/4$; il discorso su $2f_0$ appena visto dunque non si può applicare, e anche a $2f_0$ vi sarà una banda attenuata. I grafici usano un angolo ϑ per la descrizione della figura, definito come:

$$\vartheta = \frac{c}{\Lambda}$$

dove $\Lambda = d$, come già detto.

Per quanto riguarda il grafico relativo al modo TE, si può vedere che, per $\vartheta_i = 0$, le ondulazioni quasi non si vedono, ma si vedono solo i picchi di banda attenuata; al crescere di ϑ_i , quindi, i picchi si spostano un poco, dal momento che, come visto precedentemente, le bande attenuate (dal grafico precedente) si spostano a frequenze un poco più elevate. Inoltre, dal momento che Γ in modulo cresce con ϑ_i come visto precedentemente, si ha che i picchi

cregono: ciò dipende tuttavia esclusivamente dai coefficienti di riflessione di Fresnel. Aumentando l'angolo, le oscillazioni diventano sempre più presenti ed evidenti, e la banda passante diventa sempre più larga: β (o ξ) infatti è proporzionale a ω , quindi il grafico si taglia non verticalmente, ma con una retta con una certa pendenza: tagliando di sbieco, l'intervallo in frequenza, a certe frequenze, in cui la retta intreccia la banda passante è molto largo, da ciò l'intersezione si va a riportare in questo grafico, ottenendo quindi un allargamento della banda.

Per quanto riguarda il grafico TM, non c'è nulla da aggiungere, se non il fatto che, per $\vartheta_i = 45^\circ$, si ha un annullamento dei picchi: ciò dipende dal fatto che, per i mezzi, questo angolo è molto prossimo all'angolo di Brewster, ottenendo quindi un annullamento della banda attenuata, quindi della riflessione.

3.4.4 Confronto con la fisica dei solidi: energy gap

Quando è stata disegnata la curva di dispersione delle onde di Bloch, in realtà sono stati omessi alcuni dettagli molto importanti: si riparta dal grafico iniziale, per la determinazione di $\phi(\omega)$:

In realtà, quando si studiano le intersezioni, non ve ne è una sola come indicato precedentemente, bensì due per ogni caso! Per ogni situazione, dunque, vi sono due angoli. Oltre alla curva iniziale, vi sarà la simmetrica rispetto all'origine, e qui non c'è nulla di strano: le onde di Bloch come autovalori hanno $e^{\pm j\phi}$, quindi, per $\phi < 0$, si hanno comunque le curve. Le cose non finiscono tuttavia qui: una volta raggiunto il primo band edge e superato il secondo, dunque appena usciti dalla banda attenuata, non c'è soltanto un punto di risalita, ma due: vi sono, anche in questo caso, due intersezioni, una per la quale ϕ cresce, una per la quale ϕ decresce. Parallelamente alla curva che parte dal band edge di uscita, dunque, si ha una curva che sale in maniera duale! Le due curve cresceranno una fino a $\phi = 2\pi$, una fino a $\phi = 0$: ai bordi dunque tutto viene bene, dal momento che i due angoli sono coincidenti (dire 0 e 2π è la stessa cosa).

C'è ancora un aspetto molto negativo tuttavia: quando si considera l'arcocoseno di un certo angolo, si ha:

$$\arccos(x) = \vartheta + n2\pi$$

ossia, ogni volta che si calcola l'arcocoseno di un numero, si ha un angolo (o due che siano), più tutti i multipli di 2π : è questa la vera soluzione dell'equazione! Questo significa che la curva va traslata a ogni livello, in verticale e orizzontale.

La conclusione è: quando si disegnano le curve di dispersione nel piano (ϕ, ω) , vengono fuori cose di questo tipo:

Si tratta di una curva piuttosto complicata, con però una cosa positiva: essa presenta delle periodicità! Questa deriva dall'ambiguità che si ha nella definizione degli angoli, intesi come soluzione delle equazioni basate sull'arcocoseno.

Tutto ciò significa sostanzialmente che, considerando una sola porzione del grafico, si riesce a conoscere tutto l'andamento del grafico, dal momento che esso si ripete. La zona tra $-\pi$ e $+\pi$ è detta **prima zona di Brillouin**: tutto ciò che è contenuto in essa, replicato sul piano, dà la figura intera.

La "prima zona di Brillouin" deriva da un linguaggio spesso utilizzato nell'ambito della meccanica quantistica, scienza in cui si studia, tra le varie cose, l'andamento degli elettroni in un potenziale cristallino, potenziale periodico. Questi concetti, dunque, sono collegati a quelli della meccanica quantistica. Spesso, nei libri, si trova un disegno di questo tipo:

Si ha una parabola, che rappresenta l'equazione

$$E = \frac{p^2}{2m}$$

che è una delle equazioni fondamentali per la meccanica. Questa rappresenta la relazione di dispersione per un elettrone libero in una zona a potenziale costante: si tratta della formula dell'energia cinetica, ossia della legge di dispersione, di un elettrone libero in un mezzo omogeneo, in un potenziale dunque costante: la relazione di dispersione della particella libera. In meccanica quantistica si ha poi che:

$$p = \hbar k$$

e qua si vede il legame tra k e p .

Il fatto che ci sia un potenziale cristallino, invece, che si somma a quello della particella libera, crea delle "riflessioni": in questo caso, infatti, al potenziale della particella libero, va aggiunto quello imposto da un reticolo periodico di atomi.

La sostanza è che, anche qua, si creano degli intervalli, detti **band gap**: gap di energia, detti anche "intervalli di energia proibita". Questo termine, come abbiamo visto nei nostri esempi (di ottica/elettromagnetismo). In analogia coi nostri problemi, sembrerebbe che quelle frequenze, dunque, siano proibite, ma è veramente così? Beh, in realtà no: noi non abbiamo escluso nessun valore di frequenza nel nostro caso, tant'è che abbiamo disegnato anche una rappresentazione della parte immaginaria della curva di dispersione. Come mai in meccanica quantistica dunque si parla di "energie proibite", ma noi non parliamo di "frequenze proibite"?

La questione deriva sostanzialmente dalla diversa impostazione del problema: in meccanica quantistica non esistono particelle con energie che appartengono a questo intervallo, ma la cosa ha una spiegazione molto precisa: nell'ambito della meccanica quantistica, quello che si fa normalmente è calcolare le energie del sistema (che sarebbero per noi le frequenze); queste energie, in meccanica quantistica, hanno il significato delle "nostre" frequenze di risonanza; nei problemi di ottica, invece, il punto di vista è completamente diverso, dal momento che si ha un generatore equivalente, che specifica la frequenza nel circuito: la frequenza è dunque un "dato specificato", non un qualcosa da ricavare a partire da condizioni al contorno. Per noi il generatore **forza** delle frequenze; le frequenze di risonanza, tuttavia, sono indipendenti dal generatore. Si pensi a un circuito RLC o a una guida d'onda: le frequenze di risonanza sono nel primo caso quelle che annullano/rendono infinita l'impedenza, nel caso della guida d'onda quelle configurazioni che garantiscono un andamento della tensione che rispetti delle condizioni al contorno. I problemi della meccanica quantistica sono confrontabili con la ricerca delle frequenze dei modi in una guida d'onda: si impostano condizioni al contorno, che permettono di trovare degli andamenti del campo. A ogni ξ che si ricava dalle condizioni al contorno corrisponde una ben precisa frequenza: k , quindi ω/c , è una frequenza di risonanza, che deriva dall'aver fissato ξ , fissato per l'appunto dalle condizioni al contorno.

Dal momento che, nella meccanica quantistica, le condizioni al contorno periodiche (di Born) che fissano i valori energetici sono tali da imporre il fatto che p abbia valori del tipo:

$$p = \frac{m\pi}{L}$$

(condizione che peraltro ricorda molto quella delle guide); ciò implica che i valori di p , quindi di k , sono reali, quindi le energie equivalenti non potranno essere nei "band gap", come conseguenza dei possibili valori di p . È come se noi avessimo fissato i valori di ϕ ai soli reali e fossimo andati a vedere quali sono le ω che si possono di conseguenza avere.

3.4.5 Osservazioni conclusive

Come si fa, a questo punto, a disegnare la curva dei coefficienti di riflessione? Come detto:

$$\underline{\underline{T}}_{\text{completa}} = \underline{\underline{T}}^N = \underline{\underline{M}}\lambda\underline{\underline{M}}^{-1}$$

Si supponga di aver calcolato \underline{M} e la matrice degli autovalori (quest'ultima è stata effettivamente calcolata, quella degli autovettori no); facendo i conti, si scopre che:

$$\underline{T}_{\text{completa}} = \begin{bmatrix} T_{11}U_{N-1} - U_{N-2} & T_{12}U_{N-1} \\ CU_{N-1} & DU_{N-1} - U_{N-2} \end{bmatrix}$$

e

$$U_N = \frac{\sin [(N+1)\phi]}{\sin \phi}$$

Questo è il risultato finale necessario per il calcolo della matrice di trasmissione della struttura; volendo riavvicinarci alla fisica, da questa sarà necessario determinare la matrice scattering complessiva, utilizzando le formule di conversione. Si può vedere che:

$$r_N = \Gamma_{A^-} = \frac{CU_{N-1}}{AU_{N-1} - U_{N-2}}$$

questa formula è sostanzialmente coincidente con la formula per ricavare S_{11} di una matrice a partire dagli elementi della matrice di trasmissione⁴. Discorso simile per S_{21} , che sarà il coefficiente di trasmissione.

Tutte queste formule permettono di trovare la matrice scattering riferita a $Z_{\infty 1}$ a entrambe le porte; al fine di determinare la matrice scattering con le impedenze $Z_{\infty 0}$ e $Z_{\infty, s}$, è "semplicemente" (concettualmente semplice, ma richiede un po' di algebra) necessario risolvere un circuito di questo tipo:

Al fine di determinare i parametri scattering riferiti a un'altra impedenza di riferimento, bisogna considerare il circuito a parametri distribuiti con la matrice scattering, chiusa non sull'impedenza di riferimento nativa ($Z_{\infty 1}$), ma su quella che si desidera; caricando la porta 2 con $Z_{\infty, s}$ e calcolando S_{11} e S_{21} si risolve metà del problema, e idem chiudendo la 1 con $Z_{\infty 0}$ e facendo gli stessi conti.

3.5 Note conclusive sui riflettori di Bragg

In questa sezione si vuole presentare un certo numero di commenti aggiuntivi e conclusivi, rispetto a quanto già detto per quello che concerne gli specchi di Bragg.

⁴Vedi Orta R. - Teoria delle linee di trasmissione - cap. 7

3.5.1 Riflettori di Bragg accordabili

Un primo cenno si può fare riguardo i riflettori di Bragg accordabili: come già detto, i riflettori di Bragg sono sostanzialmente sistemi funzionanti sulla base di effetti di risonanza; se fosse possibile “accordare” questi strumenti, si potrebbe aumentare la banda degli specchi. Un’idea, sostanzialmente, è quella di modificare le dimensioni: dal momento che gli specchi funzionano dal momento che la lunghezza degli strati è pari a $\lambda_g/4$, quello che si può fare è variare le dimensioni fisiche del sistema, per esempio “schiacciando” il dielettrico; in questo modo, è possibile introdurre delle deformazioni (assolutamente piccole, si parla di qualche centinaio di nanometri), utilizzando metodi puramente ottici.

3.5.2 Birifrangenza di forma

Il discorso della birifrangenza è piuttosto complicato, quindi lo si può affrontare solo in maniera qualitativa, dal momento che esso presuppone una conoscenza del comportamento dei materiali anisotropi. Per ora, quello che noi faremo è analizzare un particolare aspetto della cosa.

Precedentemente è stata analizzata una struttura di questo tipo:

Si è detto che un’onda piana incidente sulla struttura eccita le onde di Bloch, le quali sono caratterizzate da ϕ , anche detto “sfasamento per cella”; ϕ è una funzione della frequenza. Si è inoltre definita una costante di propagazione di Bloch, k_B , come:

$$k_B = \frac{\phi}{d_1 + d_2}$$

A questo punto ci poniamo una domanda: cosa capita, se $\omega \rightarrow 0$? Beh, prima di tutto, è importante aprire una parentesi sul significato dell’ipotesi appena proposta: come sempre, non è che si pretenda che $\omega = 0$, o che comunque abbia valori particolarmente bassi: nella fattispecie, per dire che una pulsazione tende a 0, bisogna capire “rispetto a cosa” deve essere asintoticamente nulla: qual è il termine di confronto. Dire che la pulsazione è bassa deve essere relazionato allo sfasamento per blocco, ϕ : la frequenza deve “far variare poco” la fase. Questo coincide con il dire che:

$$\phi \ll 1$$

dove per “1” ovviamente si intende 1 radiante, dal momento che ϕ è una misura di angolo, di fase. In alternativa, quello che si può fare per quantificare la variazione di fase è definire una lunghezza d’onda di Bloch a partire dalla k_B , come:

$$\lambda_B = \frac{2\pi}{k_B}$$

in questo modo, se gli spessori degli strati, d , sono $d \ll \lambda_B$, significa sostanzialmente avere a che fare con una condizione equivalente alla suddetta $\omega \rightarrow 0$: la dimensione della cella deve essere molto piccola rispetto a λ_B .

Quello che si vede è che l'onda, in queste condizioni, “quasi non si accorge” delle variazioni di indice di rifrazione, per esempio da n_1 a n_2 , da n_2 a n_1 e così via; ciò che si ottiene, in pratica, è una sorta di media per cui il mezzo si compatta, come se fosse omogeneo, con un certo indice di rifrazione equivalente; al fine di fissare il concetto di $n_{\text{equivalente}}$ si può pensare alla microstriscia, struttura in cui si usa il concetto di n_{eff} al fine di avere gli strumenti per il calcolo della velocità di fase, quindi per trattare la microstriscia come una vera e propria linea di trasmissione.

Ciò che capita tuttavia in questo caso, ossia la trattazione di una struttura a indici omogenei, è sostanzialmente dipendente dalla polarizzazione: nel caso di polarizzazione TE o TM, infatti, la struttura si comporterà in maniera diversa; per studiare formalmente la cosa si può vedere che il mezzo è omogeneo, ma anisotropo, e che in sostanza il comportamento equivalente è quello di una struttura a due indici di rifrazione. Il comportamento che si potrebbe ottenere potrebbe anche essere estremamente variabile con la frequenza: un indice potrebbe essere n_1 , proprio di un metallo, e un altro un n_2 , proprio di una struttura trasparente: si mostra che, a seconda della polarizzazione, si avranno dunque due diverse onde rifratte. Il nome di questo fenomeno è “birifrangenza di forma”: “bi-rifrangenza” dal momento che si ha uno sdoppiamento dell'indice di rifrazione, “di forma” dal momento che questa dipende dalla periodicità presente nella forma del reticolo.

Questo concetto della “media” in realtà è molto più familiare di quanto non ci si possa aspettare: quando si dice che un mezzo ha un “vero” indice di rifrazione n_1 , in realtà, cosa si sta dicendo? Un mezzo, come noto dalla teoria della fisica dello stato solido, è composto da atomi non disposti da un reticolo (se si parla, ovviamente, di mezzo isotropo); quando si fa incidere un'onda elettromagnetica su un mezzo di questo tipo, il campo incide con gli elettroni più esterni, che iniziano a vibrare per seguirlo; vibrando, essi emettono onde secondarie, che vanno a interagire tra loro: si ha dunque in sostanza una molteplicità di onde scatterate in ogni direzione, che si combinano; ciò, tuttavia, si vede su distanze dell'ordine di 1 angstrom. La radiazione ottica ha delle λ dell'ordine di $1\mu\text{m}$, quindi molto maggiori della distanza interatomica: in una λ possono esserci anche 5000 passi reticolari, quindi per questo non dobbiamo particolarmente preoccuparci del fatto che n_1 sia una buona media o meno: n_1 è quindi, a tutti gli effetti, una “media”

del comportamento dei vari elettroni. Quando si dice dunque, per esempio, che $n_1 = 1, 5$, stiamo già implicitamente omogeneizzando un materiale!

In realtà, ora, dipende tutto dal materiale che si vuole utilizzare: se si utilizza il vetro, nessun problema, dal momento che esso è una struttura amorfa, dunque irregolare; volendo utilizzare del quarzo, il quale è un cristallo, esso avrebbe una struttura ordinata, dunque sarebbe pensabile come una sorta di struttura periodica, proprio come il reticolo di Bragg (per quanto la periodicità sarebbe in tre dimensioni); il periodo d sarebbe molto più piccolo della lunghezza d'onda, ma quindi si ricadrebbe nel caso di prima! In tal caso, l'omogeneizzazione sarebbe non corretta da applicare.

Un caso in cui, volendo far incidere un'onda piana, non si può più considerare l'ipotesi di omogeneizzazione prima introdotta, è quello dei raggi X: i raggi X infatti hanno una lunghezza d'onda confrontabile con le distanze atomiche, quindi ora il discorso di media salta: ora l'unico modo per studiare gli atomi è proprio quello di considerare ciascun atomo come un centro di scattering, e studiare quindi come si sommano in fase, controfase o in altri modi i vari contributi di riflessione. In questo caso, quando si scrive una cosa del tipo:

$$\underline{E}(r) = \underline{E}_0 e^{-jk \cdot r}$$

dove

$$\underline{k} = \hat{s} \omega \sqrt{\varepsilon \mu}$$

3.5.3 Caso della singola cella come degenerazione caso generale

Un'ulteriore interessante osservazione potrebbe riguardare un'interpretazione dell'uso delle onde di Bloch, per un caso molto particolare: quello di cella unica. Si supponga di avere qualcosa di questo tipo:

Questa struttura è stata sostanzialmente presentata ancora mediante l'uso delle linee di trasmissione, studiando dunque lo scattering multiplo, ma non è necessario fare ciò: questo tipo di struttura può essere pensata come un reticolo, composto da una sola cella!

Questo sicuramente funziona, ed è possibile ottenere le curva con le formule basate sulle onde di Bloch anche in questo caso, ponendo $N = 1$, però questo approccio ha anche i propri limiti: sarebbe come considerare un cambio di base, in un singolo punto; le onde di Bloch, con l'approccio da noi utilizzato, sono infatti definite su singoli punti, perdendo dunque di interesse al di fuori di essi.

3.5.4 Confronto tra reticoli di Bragg e reticoli di diffrazione

Precedentemente si è parlato di strutture di questo genere:

Questa struttura era stata introdotta nel capitolo introduttivo al resto della trattazione; questa struttura è molto complicata da studiare, dal momento che non si ha un'interfaccia piana, bensì dei rilievi (che possono sia esser realizzati mediante scavi, sia mediante la crescita di rilievi, a seconda della tecnologia che si ha a disposizione); il periodo spaziale della struttura è $d \sim \lambda$. Precedentemente, si è detto che una struttura di questo tipo, data un'onda piana incidente, ha un'onda piana riflessa (ordine 0), ma non solo: vi sono varie onde piane riflesse, a vari ordini, minori o maggiori di 0. La domanda principale che ci si potrebbe porre è: quante onde piane effettivamente sono riflesse da questa struttura? Come in pratica quasi ogni problema lineare, questa domanda ha risposte banali: 0, 1, infinite; in questo caso, le onde piane sono infinite, ma solo un numero finito è ben visibile, dal momento che il k_z associato è reale: infinite onde sono infatti evanescenti, e solo alcune presentano k_z reale e dunque effettivamente si propagano (un po' come i modi sopra taglio in una struttura guidante). Quando $d \gg \lambda$, si hanno tantissime onde piane riflesse; in alternativa, nei reticoli *subwavelength*, ossia quelli per cui $d < \lambda$, tutte le onde riflesse sono sotto taglio, tranne una.

Abbiamo ripreso l'argomento, ma cosa c'entra questo con Bragg? A prima vista, sembrerebbe nulla, ma ciò non è vero. Questo è un problema difficile da studiare, ma è possibile fare una cosa del genere:

Questo problema, in qualche modo, deve essere modellato; il punto essenziale è riconoscere la presenza di due interfacce, dividerle, e poi applicare il formalismo delle matrici scattering (qui si darà solo l'idea ma non si procederà, anche dal momento che questi calcoli non sono fattibili in forma chiusa). La prima delle due interfacce è pensabile come un reticoli di Bragg, ruotato però di 90° : questo significa che, in questo caso, la direzione di propagazione è parallela alle interfacce tra i vari strati, invece che normale. A questo punto, per questa struttura, è necessario imporre la continuità dei campi elettrici trasversali alle interfacce, con una differenza: se il problema con cui siamo abituati a che fare è 2×2 , questo è $\infty \times \infty$, essendo la struttura illimitata lungo x . Un'idea per affrontare ciò potrebbe essere quella di prolungare la periodicità anche a sinistra, nel mezzo n_0 : $n_0 = n_1$. Questo semplifica il problema.

I modi di propagazione nel mezzo omogeneo, sono onde piane, ma qui non tutte le possibili onde piane sono valide, bensì solo quelle, con un certo ϑ , con i seguenti valori discreti:

$$\xi_n = k_0 n_0 \sin \vartheta + n \frac{2\pi}{d}$$

in pratica, non tutte le onde piane “a sinistra” sono possibili, ma solamente quelli che sono legati agli ξ incidenti in questo modo. A destra, si ha a che fare con un reticolo di Bragg infinito, oppure in alternativa un array di guide d’onda planari equispaziate; in questo caso, i modi di propagazione sono caratterizzati da una certa β (costante di propagazione), e dal fatto che lo sfasamento su una cella ϕ è dato da:

$$\phi = k_0 n_0 \sin \vartheta d$$

questi sono, se si pensa bene, gli stessi simboli utilizzati per gli specchi di Bragg: le due cose, in effetti, sono quindi abbastanza vicine tra loro: questo potrebbe essere un collegamento tra le due strutture. Questo è lo sfasamento per cella; precedentemente, quello che era stato fatto era fare dei grafici in cui fissavamo ω , β , e trovavamo ϕ ; noi invece ora vogliamo trovare β funzione di ϕ e ω , dove ϕ è fissato dall’onda incidente; considero il problema come lo studio di un reticolo di Bragg, dove però fisso lo sfasamento grazie a ciò che conosco dal campo incidente, e ricavo β ; ricavato β , si impone la continuità dei campi e così via.

Commento aggiuntivo

Un commento aggiuntivo: uno specchio di Bragg può essere utilizzato come “carico” per una guida planare, in questa maniera:

Cosa si può, qualitativamente, dire su questa struttura? Data un’onda che arriva in una guida (supponiamo per esempio da destra), essa incontra discontinuità, ma discontinuità complicate da studiare: a differenza di quello che capita in una normale interfaccia piana (dove si ha una sola onda riflessa e una sola onda trasmessa), quello che si rischia di avere è una situazione in cui vi sono varie onde incidenti e scatterate da essa, ottenendo un accoppiamento tra i vari modi: dato un solo modo viaggiante, si rischia di eccitare i modi dello spettro continuo, facendo nascere una miriade continua di modi (per l’appunto dello spettro continuo) che irradiano da tutte le parti; questo significa che un osservatore, a occhio, vedrebbe della luce associata a questo rilievo, e ciò non è per niente positivo: l’energia non è confinata. Ad ogni modo, in realtà, questo è uno dei modi standard di utilizzare i reticoli di Bragg in una guida, però con attenzione: si deve fare attenzione di ridurre al minimo l’eccitazione; ad ogni modo, più o meno, questa struttura si comporta come uno dei reticoli di Bragg studiati, per quanto i calcoli esatti siano estremamente complicati da fare.

3.5.5 Analisi per piccole riflessioni

Al fine di studiare i reticoli di Bragg, ciò che si può fare è partire da un approccio completamente diverso da quello analizzato, molto più vicino a quello basato sullo studio dello scattering multiplo invece che da quello delle onde di Bloch. Il ragionamento, dunque, per quanto approssimato, sarà basato solo ed esclusivamente sulle onde di potenza, e sarà basato sull'analisi per piccole riflessioni.

Si immagini di avere a che fare con una struttura di questo tipo:
dove

$$\phi_1 = k_{z1}d_1$$

$$\phi_2 = k_{z2}d_2$$

ϕ è uno sfasamento “vero”, non “per cella”. Si introduca il coefficiente di riflessione di Fresnel tra le interfacce 1 e 2 come:

$$\Gamma_0 = \frac{Z_{\infty 2} - Z_{\infty 1}}{Z_{\infty 2} + Z_{\infty 1}}$$

mediante alcuni calcoli, basati sull'uso dell'approssimazione per piccole riflessioni, si può trovare che il coefficiente all'ingresso della struttura di N_C celle sia:

$$\Gamma(N_C) = \Gamma_0 \left(1 - e^{-j2(\phi_1 + \phi_2)}\right) \sum_{i=0}^{N_C-1} e^{-j2(\phi_1 + \phi_2)i}$$

Questa sommatoria ha un significato fisico: essa rappresenta la somma delle riflessioni. La teoria delle piccole riflessioni semplifica tutto, nel senso che trascura un contributo fondamentale: quello delle riflessioni multiple. Il termine delle riflessioni multiple, come si può intuire da un ragionamento qualitativo, è quello che dipende dal denominatore, che si trascura per l'appunto con la teoria delle piccole riflessioni: il denominatore di ciascun termine si poteva infatti studiare come la serie geometrica, i cui vari termini erano i singoli contributi delle onde riflesse, e questo per ciascuna interfaccia. Si vuole a questo punto semplificare i termini, usando un po' di algebra; si consideri z il termine della sommatoria:

$$z = e^{-j2(\phi_1 + \phi_2)}$$

dato ciò, si ha che la somma è:

$$\begin{aligned}\sum_{i=0}^{N_C-1} z^i &= z^0 + z + z^2 + z^3 + \dots + z^{N_C-1} = \\ &= \frac{z^{N_C-1}}{z-1}\end{aligned}$$

questo si può alternativamente scrivere come:

$$\begin{aligned}\frac{z^{N_C-1}}{z-1} &= \frac{z^{\frac{N_C}{2}} z^{\frac{N_C}{2}} - z^{-\frac{N_C}{2}}}{z^{\frac{1}{2}} (z^{\frac{1}{2}} - z^{-\frac{1}{2}})} = \\ &= z^{\frac{N_C-1}{2}} \frac{\sin(N_C(\phi_1 + \phi_2))}{\sin(\phi_1 + \phi_2)}\end{aligned}$$

Un caso molto interessante è quello per cui $\phi_1 = \phi_2 = \phi$; questo, per esempio, è il caso delle strutture $\lambda/4$, come visto precedentemente; si ha, in queste situazioni, che:

$$\Gamma(N_C) \sim j\Gamma_0 e^{-j(2N_C-1)\phi} \frac{\sin(2N_C\phi)}{\cos\phi}$$

Questa è una formula molto semplice e comoda da utilizzare.

Se si sovrappone questa espressione a quella esatta, ricavata mediante le onde di Bloch e quel formalismo, si vede che le due si sovrappongono in maniera quasi indistinguibile per i “lobi secondari”, mentre sono molto differenti per quanto riguarda la zona a elevata riflettività; questo accade dal momento che, per quei valori, Γ è grande, quindi la teoria sbaglia molto: non è più nel proprio range di validità.

I massimi per questa espressione sono per

$$\phi = (2n+1) \frac{\pi}{2}$$

il caso più semplice, di questi, è $n=0$, quindi $\phi = \pi/2$; ϕ , tuttavia, è sia ϕ_1 , sia ϕ_2 : questo capita quando entrambi gli spessori sono $\lambda_g/4$, proprio come atteso.

Capitolo 4

Interferometri di Fabry-Perot

4.1 Introduzione e concetti preliminari

Si è detto, precedentemente, che il formalismo delle onde di Bloch può essere utilizzato anche per strutture semplici, a poche celle; si immagini per esempio di avere a che fare con una struttura a due celle:

Si consideri, per ipotesi, $d_1 \gg d_2$: il caso di interesse, in questa situazione, è quello di riflessione totale frustrata, dunque per cui $n_2 < n_1$, e $\vartheta_1 > \vartheta_c$, in modo tale che le onde di potenza (questi discorsi sono relativi alle onde di potenza) siano evanescenti in n_2 , ma possano propagarsi in n_1 .

Al fine di analizzare queste strutture, un approccio è di nuovo quello delle formule a partire dalle quali sono stati analizzati gli specchi di Bragg, ma non è l'unico: ciò che si potrebbe fare, per quanto riguarda questa struttura, è utilizzare un approccio basato sulle cavità risonanti: questo sistema può essere modellato mediante due specchi semitrasparenti: specchi con un certo coefficiente di riflessione, piuttosto elevato, e uno di trasmissione conseguentemente piccolo; il modello in questo caso potrebbe essere **ottico**, e potrebbe essere il seguente:

Si hanno due specchi, con in mezzo un certo mezzo n , separati da distanza l . Questo oggetto è, a tutti gli effetti, un interferometro di Fabry-Perot, e, come vedremo, esso ha dei picchi di trasmissione ad alte frequenze.

Pensare che un oggetto di questo genere, con riflessione elevata e svariati strati, intuitivamente difficilmente potrà avere un coefficiente di trasmissione elevato; questo in realtà si avrà, esclusivamente, per delle frequenze particolari: per $\phi = n\pi$, infatti, si ha un certo picco di trasmissività. Questo si era visto quando si aveva uno slab, per esempio; questa, è la stessa situazione, dove però lo specchio semitrasparente ora è, invece che l'interfaccia, una cosa più complicata: un certo reticolo.

I reticoli di Bragg possono studiare anche strutture di questo tipo, e si parla di **resonant tunnelling**: si ha un tratto per cui l'onda è evanescente, e questo è legato a un fenomeno di risonanza, dal momento che la trasmissione totale avviene solamente quando d_1 è un multiplo di $\lambda_{g1}/2$.

Introdotta questo concetto con questo approccio, andiamo avanti con il problema, continuando con il modello:

Il modello che verrà utilizzato è ancora una volta basato su un circuito a parametri distribuiti: si hanno due componenti (gli specchi) caratterizzati ciascuno da una certa matrice scattering: lo specchio sinistro avrà matrice \underline{S}' , lo specchio destro \underline{S}'' . Il dielettrico centrale, n , sarà complesso, ma non nel solito senso: in questo caso, infatti, esso sarà:

$$n = \beta - j\alpha, \quad \alpha < 0$$

questo significa che il dielettrico centrale è **attivo**, ossia che può avere un guadagno: questo è, per esempio, il modello di una cavità LASER. Tutto ciò è eccitato da un'onda piana incidente da sinistra, con angolo ϑ_1 arbitrario. Il guadagno nella fattispecie è in potenza: se si considera una sola percorrenza del tratto, quindi solo l , il guadagno è di potenza; se si considera sia "andare" sia "tornare", allora è considerabile come un guadagno di tensione (essendo infatti il guadagno di tensione pari alla radice di quello di potenza).

Deve essere calcolata la matrice scattering dei due elementi in cascata; ciò si può fare subito, per quanto riguarda il S_{21} dell'interferometro totale, ricordando le formule e vedendo quindi che:

$$S_{21} = \frac{S'_{21} S''_{21} e^{-jk_z l}}{1 - S''_{11} S'_{22} e^{-j2k_z l}}$$

Questa formula è fondamentale, dunque vogliamo "sviscerarla" per bene. Prima di tutto, al fine di ricavare informazioni importanti, un'idea è quella di calcolarne il modulo quadro, trovando cioè:

$$|S_{21}|^2 = \frac{|S''_{21}|^2 |S'_{21}|^2 G_0}{|1 - G_0 S'_{22} S''_{11} e^{-j2\beta l}|}$$

dove

$$G_0 = e^{-2\alpha l}$$

e, si ricorda, questa volta $\alpha < 0$, essendo il materiale attivo. L'esponenziale quindi non ha modulo unitario, e anzi in questo caso è tale da aumentare l'ampiezza dell'onda. Considerando n_0 il mezzo di ingresso "da sinistra", si ha:

$$k_z = k_0 \sqrt{n^2 - n_0^2 \sin^2 \vartheta_i}$$

4.2 Osservazioni e calcoli sull'interferometro di Fabry-Perot

Prima di proporre alcuni calcoli e osservazioni ulteriori su questa formula, alcune ipotesi sugli specchi:

- si supponga di avere a che fare con specchi reciproci:

$$S'_{21} = S'_{12} \quad S''_{21} = S''_{12}$$

- si supponga di avere specchi senza perdite; questo implica il fatto di avere matrici \underline{S} unitarie, ma, nel dettaglio, le seguenti condizioni:

$$|S'_{11}|^2 + |S'_{21}|^2 = 1$$

$$|S''_{11}|^2 + |S''_{21}|^2 = 1$$

Si introducono inoltre i concetti di riflettività in potenza degli specchi, come:

$$R_1 = |S'_{22}|^2$$

e

$$R_2 = |S''_{11}|^2$$

Come mai proprio questi parametri? Semplice: per come è stato definito il sistema, e dal momento che la riflettività che ci interessa è quella “interna”, quella nello spazio l , sono i coefficienti di riflessione “interni” quelli importanti, quelli “visti da dentro”: quelli relativi alla porta 2 per quanto riguarda lo specchio 1, e quello relativo alla porta 1 per lo specchio 2.

Dato che valgono le ipotesi e le definizioni precedentemente scritte, è possibile scrivere che:

$$|S'_{21}|^2 = 1 - R_1$$

$$|S''_{21}|^2 = 1 - R_2$$

Si noti che questi parametri, le **riflettività**, non sono esattamente dei coefficienti di riflessione: questi sono infatti grandezze legate più che altro alle onde di potenza progressiva o regressiva, mentre le riflettività si riferiscono direttamente alla potenza; in ambito ottico, solitamente, sono più presenti le riflettività dei parametri scattering, quindi dei coefficienti riferiti alla tensione.

Riscrivendo il numeratore con queste considerazioni e sviluppando il denominatore secondo il teorema di Carnot, è possibile scrivere ciò:

$$|S_{21}|^2 = \frac{G_0(1 - R_1)(1 - R_2)}{1 + G_0^2 R_1 R_2 - 2G_0 \sqrt{R_1 R_2} \cos(\angle(S''_{11} S'_{22}) - 2\beta l)}$$

L'argomento del coseno rappresenta a tutti gli effetti un fattore di fase; esso, verrà chiamato, per comodità:

$$2\varphi \triangleq \angle(S''_{11} S'_{22}) - 2\beta l$$

2φ rappresenta, a tutti gli effetti, la fase del guadagno di anello: si tratta infatti della somma delle fasi dei due coefficienti di riflessione degli specchi, meno la fase del tragitto andando dal primo al secondo e dal secondo al primo. Proseguendo, si può vedere che vale la seguente formula, riarrangiando la presente:

$$|S_{21}|^2 = \frac{G_0(1 - R_1)(1 - R_2)}{(1 - G_0 \sqrt{R_1 R_2})^2 + 2G_0 \sqrt{R_1 R_2} (1 - \cos(2\varphi))}$$

Usando a questo punto le formule di bisezione, ricordando che:

$$1 - \cos(2\varphi) = 2 \sin^2 \left(\frac{2\varphi}{2} \right)$$

si può dire che:

$$|S_{21}|^2 = \frac{G_0(1 - R_1)(1 - R_2)}{(1 - G_0 \sqrt{R_1 R_2})^2 + 4G_0 \sqrt{R_1 R_2} \sin^2 \varphi}$$

La frequenza, in prima approssimazione, è presente esclusivamente nel termine φ , a causa della presenza di β in esso; rigorosamente parlando essa dovrebbe in realtà far variare sostanzialmente ogni termine, ma per ora si considera un caso piuttosto particolare: quello di specchi la cui matrice $\underline{\underline{S}}$ non dipende dalla frequenza; in questa situazione, sia le riflettività sia la fase dei coefficienti scattering non dipendono dalla frequenza; questo è vero in caso gli specchi per l'appunto siano indipendenti dalla frequenza ma, come si discuterà in seguito, già nel caso degli specchi di Bragg non è vero. Se d'altra parte lo specchio è solo un'interfaccia, siamo a posto.

Questo concetto è comprensibile anche solo da una prima osservazione del sistema: in generale, la dipendenza della frequenza è tanto più forte quanto più il dispositivo è grande rispetto alle λ in gioco; la variabilità della frequenza, infatti, come già detto in precedenza, è strettamente legata alla possibilità di accumulare in esso energia, cosa che a sua volta è strettamente legata alle dimensioni del sistema; la lunghezza l della linea rappresenta infatti la prima causa di variabilità, almeno date le ipotesi precedenti (specchi non dispersivi): maggiore è lo spazio, maggiore è l'energia immagazzinabile, quindi maggiore la dipendenza dalla frequenza. β invece è:

$$\beta = \operatorname{Re} \{k_z\} = \operatorname{Re} \left\{ k_0 \sqrt{n_{\text{cavity}} - n_1^2 \sin^2 \vartheta_1} \right\}$$

Al variare di β , S_{21} può presentare dei massimi; dal momento che il numeratore è costante e che il denominatore ha come parte variabile il solo seno, il massimo si avrà quando:

$$|S_{21}|_{\max}^2 = \frac{G_0(1 - R_1)(1 - R_2)}{(1 - G_0\sqrt{R_1 R_2})^2}$$

questo accade quando

$$\sin \varphi = 0$$

quindi per

$$\varphi = n\pi$$

Si noti che, tuttavia, φ rappresenta metà della fase del guadagno di anello T (essendo la fase completa 2φ); si ha dunque che la massima trasmissività si ha per:

$$\angle T = 2n\pi$$

A questo punto, $|S_{21}|_{\max}^2$ potrebbe essere maggiore, minore o uguale a 1! Si consideri una condizione particolare, per cui $G_0 = 1$: in queste condizioni, $|S_{21}|_{\max}^2 < 1$, a meno che ovviamente $R_1 = R_2$.

4.2.1 Banda dell'interferometro

Abbiamo determinato il valore massimo della trasmissività che l'interferometro può avere; questo non è tuttavia l'unico parametro interessante, dal momento che possiamo essere interessati anche a quello che capita per altri valori della frequenza.

Qual è la larghezza di banda della struttura? Beh, per definire la larghezza di banda, è necessario scegliere un parametro sul quale basarla; come fatto molto spesso in ambito elettronico, ciò che si fa è definirla a metà potenza, mediante un valore di sfasamento $\varphi_{-3\text{dB}}$. Quanto deve valere questo sfasamento? Beh, la risposta è abbastanza semplice: al denominatore si hanno due termini, uno costante e uno variabile; quando il termine variabile, dipendente dal seno, eguaglia l'altro, si ha concettualmente che il denominatore raddoppia rispetto al caso massimo; la condizione, dunque, è:

$$\sin^2(\varphi_{-3\text{dB}}) = \frac{(1 - G_0\sqrt{R_1R_2})^2}{4G_0\sqrt{R_1R_2}}$$

L'espressione può tuttavia essere modificata: $\varphi_{-3\text{dB}}$ rappresenta il valore di fase "assoluto" rispetto cui si ha un abbassamento della potenza di 3 dB; quello che si può fare, tuttavia, è scrivere questo valore di fase come il valore del massimo, più la distanza, in termini di fase, rispetto al massimo:

$$\varphi_{-3\text{dB}} = n\pi \pm \frac{\Delta\varphi_{-3\text{dB}}}{2}$$

detto $\varphi_{-3\text{dB}}$ quindi l'intero range. Il punto chiave, ricavabile a questo punto mediante le formule degli archi associati note dalla goniometria, è:

$$\sin^2 \varphi_{-3\text{dB}} = \sin^2 \left(\frac{\Delta\varphi_{-3\text{dB}}}{2} \right)$$

In generale, tuttavia, la larghezza di banda funzione della fase sarà molto piccola; non è dunque una cattiva approssimazione il fatto di confondere l'argomento del seno con il seno, ottenendo:

$$\Delta\varphi_{-3\text{dB}} \sim n\pi \pm \frac{1 - G_0\sqrt{R_1R_2}}{2\sqrt{G_0}\sqrt[4]{R_1R_2}}$$

A questo punto, facciamo un'operazione diversa: supponendo di poter variare G_0 , ossia di poter variare il "guadagno" del dielettrico attivo, dati specchi fissi, cosa capita? Beh, se $G_0 = 1$, abbiamo detto che il massimo della trasmissività è minore di 1; man mano che si cresce, il denominatore diventa sempre più piccolo, fino a quando non si assume un particolare valore:

$$G_0 = \frac{1}{\sqrt{R_1R_2}}$$

in questo caso, il denominatore va a 0, e:

$$|S_{21}|_{\text{max}}^2 \rightarrow \infty$$

Al contempo però, osservando la formula della larghezza di banda, si vede che:

$$\Delta\varphi_{-3\text{dB}} \rightarrow 0$$

Cosa significa ciò? Al fine di capirlo, si ricordi qual è la definizione di parametro scattering:

$$S_{21} \triangleq \left. \frac{b_2}{a_1} \right|_{a_2=0}$$

quindi, o a_1 è finito e b_2 è infinito, o b_2 è finito e a_1 infinitamente piccolo. Ciò che si ha, ovviamente, è la più realistica delle due, ossia la seconda situazione: se S_{21} esplose, si ha sostanzialmente un'uscita finita anche con un ingresso sostanzialmente nullo: questa è la condizione che si ha in un oscillatore. Un oscillatore, tuttavia, ha un comportamento particolare: data una certa energia, a una certa frequenza (che potrebbe anche essere la frequenza nulla, ossia la continua), si ha in uscita un segnale a un'altra frequenza; parte della potenza introdotta (non tutta, dal momento che il sistema sicuramente avrà un'efficienza meno che unitaria), verrà quindi convertita in energia in uscita sotto forma sinusoidale.

Come mai capita ciò? Come mai lo si è ottenuto facendo crescere G_0 ? La risposta è abbastanza semplice, dal momento che ancora una volta è legata al guadagno di anello: se trattiamo il massimo, al solito, abbiamo che $\angle T = 2n\pi$, quindi una delle condizioni per il criterio di Bode è rispettata; l'altra condizione, era il fatto che $|T| = 1$, e, se si ha

$$G_0 = \frac{1}{\sqrt{R_1 R_2}}$$

si ha esattamente questa condizione! Aumentare G_0 significa quindi aumentare il guadagno di anello, fino a farlo arrivare a unitario, ottenendo per il sistema la condizione di oscillazione.

Il sistema ha perdite, non nel senso ohmico, bensì nel senso "radiativo": gli specchi, avendo una riflettività non totale, irradiano parte dell'energia contenuta tra essi, dunque contenuta nella cavità; la cavità, dunque, ha delle pareti "forate" sotto il punto di vista dell'energia, ossia l'energia scappa da essa. Avere $|T|$ pari a 1 significa sostanzialmente fare in modo che l'energia che esce sia controbilanciata da quella introdotta dal materiale attivo: il guadagno ripristina l'energia, sommando in fase e con la stessa ampiezza ciò che vi era prima che l'energia sfuggisse dai "buchi".

Proseguiamo con la nostra analisi: abbiamo identificato, a $\varphi_{-3\text{dB}} = n\pi$, la posizione del primo massimo; fissato n , il massimo successivo sarà a $\varphi_{-3\text{dB}} =$

$(n + 1)\pi$. Per tradizione, l'intervallo tra due massimi è stato detto **FSR**: Free-Spectral Range. Come si può vedere facendo la sottrazione tra le $\varphi_{-3\text{dB}}$ dei due massimi, si può vedere che, in **radianti**,

$$\text{FSR} = \pi$$

Questa è una definizione alternativa a quella del Q , ossia del fattore di qualità: in strutture di questo tipo, il Q verrebbe elevatissimo, essendo la banda evidentemente molto stretta, quindi ciò che si fa è definire il parametro di **Finesse**, F , come:

$$F \triangleq \frac{\text{FSR}}{\Delta\varphi_{-3\text{dB}}}$$

ribadendo che $\text{FSR} = \pi$, ricordando la formula precedentemente scritta:

$$F = \frac{\sqrt{G_0} \sqrt[4]{R_1 R_2}}{1 - G_0 \sqrt{R_1 R_2}} \pi$$

(infatti, si ha che la formula di prima teneva conto del $\frac{1}{2}$ che veniva dai conti, ma, dal momento che ora non si ha, non lo si è messo, essendo il doppio del semitermine, per dirla alla buona).

Come si può esprimere, in frequenza, la larghezza di banda? Quella che si vuol presentare, si vuol ribadire, è un'espressione che **non tiene conto delle variazioni di fase degli specchi**, considerati ideali; questa, dunque, **non sarà sufficiente nel caso dei riflettori di Bragg**. $\varphi_{-3\text{dB}}$ è pensabile come variabile da β , il quale è

$$\beta = k_z l$$

ma quindi:

$$\Delta\varphi_{-3\text{dB}} = -\Delta\text{Re}\{k_z l\}$$

avevamo che:

$$\varphi = \frac{1}{2} \angle S''_{11} S'_{22} - \beta l$$

l'altro termine è costante, dunque ce ne possiamo fregare; si ha, quindi:

$$\Delta\varphi_{-3\text{dB}} = \Delta\beta l$$

utilizzando la formula degli incrementi finiti, è possibile legare una variazione di $\Delta\omega_{-3\text{dB}}$ a quella di β :

$$\Delta\beta = \frac{\partial\beta}{\partial\omega}\Delta\omega_{-3\text{dB}}l$$

in questo modo, abbiamo potuto esprimere $\Delta\beta$. A questo punto, si può osservare che:

$$\frac{\partial\beta}{\partial\omega} = \frac{1}{v_g}$$

quindi:

$$\Delta\beta = \frac{l}{v_g}\Delta\omega_{-3\text{dB}}$$

infine, invertendo e usando la formula di prima:

$$\Delta\omega_{-3\text{dB}} = \frac{v_g}{l}\Delta\beta = \frac{v_f}{l} \frac{1 - G_0\sqrt{R_1R_2}}{\sqrt{G_0^4 R_1R_2}}$$

Questa è l'espressione della banda, espressa in rad/s (ossia secondo un'unità di frequenza); questo è un risultato più utile, sotto il punto di vista applicativo, dato che la larghezza di banda di solito è espressa in termini della frequenza.

Analisi in transitorio del comportamento della cavità, alimentata CW

Tutto ciò che è stato fatto finora si riferisce a un'analisi del comportamento della cavità, supponendone regime sinusoidale continuo, CW (Continuous Wave): si è supposto che ci sia una certa onda incidente a una certa pulsazione ω , abbiamo calcolato per ogni frequenza $|S_{21}|^2$, abbiamo fatto il grafico in φ (o, equivalentemente, in ω), e abbiamo trovato il picco di risonanza.

In transitorio, ossia nel dominio del tempo, cosa capita? Beh, si è detto che il S_{21} esplose; il fatto che, nel dominio della frequenza, si abbia questo comportamento, è riconducibile al fatto che un sistema risonante venga alimentato esattamente alla frequenza di risonanza; in tale situazione, nel dominio del tempo l'andamento dell'uscita, dell'onda trasmessa, è qualcosa di questo tipo:

Ossia, si ha che

$$\underline{\mathcal{E}}(\underline{r}, t) \propto t \sin(\omega_0 t)$$

Come noto infatti, un sistema risonante (ideale) presenta poli sull'asse immaginario, ossia per:

$$s = \pm j\omega_0$$

se l'ingresso è un segnale monocromatico a pulsazione $\omega \neq \omega_0$, calcolando l'uscita mediante il metodo delle trasformate di Laplace come:

$$\mathcal{L}^{-1} \{H(x)X(s)\}$$

se la situazione è quella appena descritta, si hanno dei poli semplici, e la situazione tende a convergere dopo un certo tempo. Se la situazione è la stessa di prima, ma con $\omega = \omega_0$, si ha qualcosa di diverso: l'eccitazione infatti presenta gli stessi poli che presenta la funzione di trasferimento, quindi l'antitrasformata di Laplace, come si può dimostrare, porta ad avere una sinusoidale, moltiplicata per il tempo t ; questo, si noti, vale per qualsiasi sistema LTI descritto mediante il formalismo delle trasformate di Laplace, dunque mediante la funzione di trasferimento.

4.2.2 Analisi in transitorio della cavità con condizioni iniziali

Consideriamo, dopo questa osservazione, qualcosa di diverso: quello che è stato analizzato prima, volendo fare un paragone meccanico, è il problema dell'analisi in transitorio (e anche nel dominio della frequenza) di un sistema sempre alimentato, CW: un po' come avere un bambino sull'altalena, con un genitore che lo spinge alla frequenza e intensità tale da fargli fare il giro massimo: l'altalena con il proprio attrito (introdotto dai giunti arrugginiti o altro) riduce la velocità, ma i genitori la ripristinano. Il problema che si vuole invece affrontare ora è quello del pendolo carico: dato il pendolo con una certa energia potenziale all'inizio, lo si lascia in evoluzione libera, vedendo come esso si comporta.

Cosa capita, nel nostro specifico caso (della cavità)? Proviamo ad analizzarlo in questo modo:

Si immagini di aver caricato la cavità, al tempo $t = 0$, con una certa energia $W(0)$; in questo modo, è come aver messo, al suo interno, un certo numero di fotoni. L'energia dei fotoni, come noto, è $\hbar\omega_0$; si ha quindi:

$$W(0) \propto N_p(0) \frac{\hbar\omega_0}{2\pi}$$

dove ω_0 è la frequenza per cui la fase del guadagno di anello è pari a $n2\pi$, N_p è il numero dei fotoni, e quindi l'energia è il numero dei fotoni per la loro energia (\hbar è la costante di Planck non normalizzata a 2π).

A questo punto, di tutta la cavità, consideriamo solo un certo piccolo volumetto dV : dato esso, il numero di fotoni al suo interno sarà il numero di fotoni, per il volumetto, diviso il volume totale V , dal momento che si suppone uniforme la concentrazione dei fotoni nella cavità:

$$\Rightarrow N_p(0) \frac{dV}{V}$$

I fotoni, tuttavia, non stanno fermi: si suppone che essi si muovano verso destra: se essi vanno verso destra, dunque verso lo specchio 2, e vengono riflessi, a seconda della riflettività R_2 (che, si noti, è la grandezza più appropriata in questo ambito, dal momento che si sta parlando di energia, quindi di potenza, quindi non di tensione: di grandezze in altre parole quadratiche), un certo numero di essi tornerà indietro. Si tenga tuttavia conto che, essendo ciò che sta dentro lo specchio attivo, man mano che vanno verso destra, il loro numero aumenta; allo stesso tempo, dopo la riflessione (che, essendo la riflettività meno che unitaria, porterà alla riduzione del numero dei fotoni), essi torneranno indietro e il loro numero pian piano continuerà ancora ad aumentare, grazie al guadagno. Si ha quindi, in totale, che, dopo la seconda riflessione, allo specchio 1, situazione in cui un “ciclo” dei fotoni è stato ultimato:

$$N_p = N_p(0) \frac{dV}{V} G_0 R_1 R_2 G_0$$

questo è il numero totale di fotoni dopo un “giro”, ed è generalmente diverso dal numero di prima. Questo, si suppone, è minore del numero di prima. Definiamo $-dN_p$ la perdita totale di fotoni rispetto a prima, e, intuitivamente, si può calcolare che essa sia:

$$-dN_p = (1 - G_0^2 R_1 R_2) N_p$$

Questo va fatto su tutta la cavità, ma, dal momento che si dovrebbe semplicemente fare un integrale di volume, si finirebbe per trovare che tutto è costante rispetto alla variabile di integrazione, quindi si finirebbe solo per avere che

$$\int_V \frac{dV}{V} = 1$$

A questo punto, un’ultima osservazione: questa è la perdita che si ha in un ciclo; essendo l la lunghezza totale della cavità, e v_f la velocità di fase, ossia la velocità con cui supponiamo che si muovono i fotoni, si ha che il tempo è

$$\text{tempo} = \frac{2l}{v_f}$$

questo, usando la normale definizione di velocità come spazio su tempo, invertita.

Mettendo tutto assieme, se ho una perdita di fotoni in questo tempo, ho che:

$$\frac{dN_p}{dt} = \frac{1 - G_0^2 R_1 R_2}{2l} N_p v_f = \frac{N_p}{\tau_p}$$

dove τ_p è detto “tempo di vita media dei fotoni nella cavità”: si tratta della costante di tempo della cavità.

Quella che abbiamo appena ottenuto è un’equazione differenziale, di cui oltretutto conosciamo il valore iniziale: un problema di Cauchy. Questo ha inoltre una soluzione ben nota in forma analitica, ossia:

$$N_p(t) = N_p(0)e^{-\frac{t}{\tau_p}}$$

La stessa legge vale ovviamente anche per quanto concerne l’energia, proporzionale al numero dei fotoni:

$$W(t) = W(0)e^{-\frac{t}{\tau_p}}$$

L’energia, quindi, si riduce nel tempo, a causa del fatto che le pareti sono “forate”.

Osservazione finale

Sono state introdotte due caratterizzazioni della stessa cavità:

- una, in regime stazionario, CW, ottenendo una curva di risposta con una certa larghezza di banda, ottenendo un $\Delta\omega_{-3dB}$;
- una, dal punto di vista del transitorio, trovando come parametro fondamentale per la determinazione dell’andamento, τ_p .

A questo punto, un parametro è temporale, uno è nella frequenza; viene naturale chiedersi cosa capitò, quando i due vengono moltiplicati tra loro. Si prova dunque a farlo:

$$\Delta\omega_{-3dB}\tau_p = \frac{v_g}{l} \frac{1 - G_0\sqrt{R_1 R_2}}{\sqrt{G_0^4 R_1 R_2}} \frac{2l}{v_f} \frac{1}{1 - G_0^2 R_1 R_2}$$

l’ultima frazione può però essere scritta come:

$$\frac{1}{1 - G_0^2 R_1 R_2} = \frac{1}{(1 - G_0 \sqrt{R_1 R_2})(1 + G_0 \sqrt{R_1 R_2})}$$

quindi, sostituendo e semplificando, si trova:

$$\Delta\omega_{-3\text{dB}}\tau_p = \frac{2}{1 + G_0 \sqrt{R_1 R_2}} \frac{1}{\sqrt{G_0^4 R_1 R_2}} \frac{v_g}{v_f}$$

Dal momento che la risonanza di solito è molto stretta, si ha che $G_0 \sqrt{R_1 R_2} \sim 1$ (non esattamente 1 ovviamente, altrimenti si avrebbe un picco perfetto, cosa impossibile); il denominatore della prima frazione quindi verrebbe 2 circa, il secondo 1, e quindi si avrebbe:

$$\Delta\omega_{-3\text{dB}}\tau_p \sim \frac{v_g}{v_f}$$

Se $v_f = v_g$, cosa ragionevole dal momento che non si ha, per ipotesi, dispersione nel materiale

$$\Delta\omega_{-3\text{dB}}\tau_p \sim 1$$

Questo dunque pone in relazione due descrizioni in un certo senso complementari dello stesso sistema; curioso è il fatto che il prodotto sia unitario: ciò introduce una sorta di principio di indeterminazione per lo studio del sistema.

Come si può motivare tutto ciò? Beh, data una cavità con specchi molto buoni, quindi molto riflettenti τ_p è molto lungo; prima che l'energia esca, quindi, ci vuole molto tempo, dal momento che le riflettività sono molto elevate e dunque è come se i "buchi" siano piccoli. Allo stesso tempo, $\Delta\omega$ deve essere molto stretta, e questo è già visto in un altro caso: precedentemente, parlando dello *slab*, si era visto che, volendo trattare le due discontinuità come due specchi, ci sono situazioni per cui, nonostante le onde evanescenti, si ha trasmissività elevata; questa è esattamente la stessa cosa! Il massimo di S_{21} è alto in modulo, e quindi, nonostante geometricamente sembri di avere a che fare con una barriera quasi insormontabile per la luce, la trasmissività è elevata e la banda conseguentemente stretta. Il fatto che la banda sia stretta giustifica il comportamento risonante del sistema: essendo la barriera molto grande, molto difficile da superare, il "gioco delle fasi" è molto critico, nel senso che le configurazioni per cui si hanno somme in fase dei vari contributi tali da avere ciò sono molto rare; questo implica il fatto che queste particolari configurazioni si avranno per una banda molto stretta, e da qui $\Delta\omega$ sarà molto stretta.

Questo principio di indeterminazione è molto generale e vale per qualsiasi fenomeno oscillatorio: quando si ha a che fare per esempio con delle particelle

instabili, come quelle generate dagli acceleratori di particelle, per misurare un tempo di vita $\tau \sim 10^{-15}$ s, quello che si fa è fare esperimento di scattering, registrando l'ampiezza di scattering in funzione dell'energia E (che sarebbe il corrispondente della nostra ω); si hanno dei picchi al variare di E , e questa permette, mediante l'applicazione di una teoria di questo tipo, di calcolare il tempo di vita delle particelle.

Cenni alla realizzazione della cavità

In realtà, quello che si ha è un qualcosa di un poco diverso, come già detto: quello che noi facciamo infatti di solito è supporre di avere specchi piani, ma ciò non ha molto senso, dal momento che non è possibile avere degli specchi infinitamente estesi. Al fine di avere un sistema realizzabile con buona approssimazione, quello che si fa è curvare gli specchi, in modo da ottenere come specchi superfici sferiche, approssimabili al vertice mediante dei paraboloidi. In queste condizioni, si può dimostrare che i modi risonanti assumono una forma molto particolare: i modi di risonanza di una cavità composta da due specchi di questo tipo sono dei fasci gaussiani, ossia distribuzioni di campo di questo tipo:

$$\underline{E} \propto e^{-\frac{x^2}{2w^2}} H_n \left(\frac{x}{w} \right)$$

H_n tiene conto del fatto che le gaussiane possono essere di ordine superiore, ed è semplicemente il modo di indicare i polinomi di Hermite: si tratta di polinomi che moltiplicano la gaussiana.

Tenendo conto di questo fatto, il fattore di propagazione da uno specchio all'altro dovrà tenere conto dell'incurvamento, che porterà ad avere per l'appunto questo fascio gaussiano, introducendo una correzione sul fattore di propagazione.

4.3 Progetto di un interferometro di Fabry-Perot con specchi di Bragg

A questo punto ci proponiamo l'obiettivo di progettare un interferometro di Fabry-Perot i cui specchi, invece che essere gli specchi ideali di cui si parlava prima, sono reticoli di Bragg. Progettarlo significa, data f_0 a cui si vuole che l'interferometro sia trasparente (ossia la posizione del massimo di S_{21}), data la Δf_{-3dB} , dati n_1 , n_2 , ϑ_i , la polarizzazione, $R_1 = R_2$ per ipotesi (specchi identici), supponendo che le strutture siano dei quarter-wave stack e infine considerando la situazione semplificata di n_1 sia all'interno sia all'esterno

degli specchi, di dimensionare tutto il resto; essendo n_1 lo strato centrale ed essendo non attivo, ipotizzeremo $G_0 = 1$, così che tutte le formule viste in precedenza saranno valide, solo imponendo questo valore. Il punto chiave del progetto è la traduzione delle specifiche su f_0 e su Δf in valori della lunghezza l e del numero di strati dei riflettori.

Prima di tutto, f_0 dipende solamente da l ; al fine di determinare la distanza l tra i due specchi, dunque, è necessario imporre che lo sfasamento del guadagno di anello sia un multiplo di 2π . Δf è invece legato al numero di strati: la larghezza di banda infatti dipende dal $\Delta\varphi$, che a sua volta dipende dalla riflettività degli specchi: da Δf è dunque possibile risalire al numero di strati.

Dal momento che il progetto degli strati è $\lambda_g/4$, si può dire che, alla frequenza f_0 , si abbia un S_{11} così calcolabile:

$$S_{11} = \frac{\left(\frac{Z_{\infty 1}}{Z_{\infty 2}}\right)^{2N} - 1}{\left(\frac{Z_{\infty 1}}{Z_{\infty 2}}\right)^{2N} + 1}$$

Questo deriva da una semplice applicazione del modello delle linee di trasmissione: a centro banda, infatti, non è necessario ricorrere alle onde di Bloch, ma basta utilizzare questa formula, sfruttando le nozioni note dagli adattatori $\lambda/4$. Dal momento che tutto è reale, anche S_{11} è reale, dunque la fase può essere o 0 o π .

Al fine di proseguire con il progetto, cerchiamo informazioni sul RTPS, ossia sul Round Trip Phase Shift:

$$RTPS = 2\angle S_{11} - 2\beta l$$

infatti, dal momento che consideriamo gli specchi uguali, e dal momento che li consideriamo simmetrici, $S_{11} = S_{22}$, e quindi i pedici ' e '' decadono di importanza: basta sommare due volte la fase di un qualsiasi coefficiente di riflessione. β ovviamente è quello valutato in f_0 :

$$\beta_0 = \frac{2\pi f_0}{c}$$

A f_0 , si vuole che il sistema sia sul "massimo" di trasmissività; questo significa che il RTPS deve essere uguale a un multiplo di 2π . In altre parole:

$$2\angle S_{11} - 2\beta_0 l = n'2\pi$$

Dal momento che per ipotesi nella cavità vi è n_1 , questa espressione può essere riscritta in un modo diverso:

$$l = m \frac{\lambda_{g1}}{2}$$

ossia, si può dire che, essendo all'interno della cavità presente n_1 , si ha k_{z1} in essa, riferito al ϑ_1 di ingresso al sistema; a f_0 , dire che si ha la massima trasmissività, e dire che il RTPS è uguale a un multiplo di 2π , significa sostanzialmente richiedere che l sia un certo multiplo della lunghezza d'onda guidata nella struttura, nella cavità. m è un parametro che dice di “quante mezze lunghezze d'onda” si vuole fare l ; si usa di solito scegliere m bassi: $m = 1$, $m = 2$ o simili. m è sostanzialmente una sorta di “indice modale”, dal momento che permette di decidere quante oscillazioni avere; per $m = 1$, si può dire di avere il “modo fondamentale”.

Questo, per quanto riguarda f_0 . Quante celle servono invece, per tenere conto della larghezza di banda desiderata? Beh, gli specchi che consideriamo hanno un certo spessore, dunque una dimensione significativa; questo significa che, in questo caso, non si potrà più dire che le matrici scattering siano completamente indipendenti dalla frequenza.

Considerando una formula dimostrata precedentemente, si ha che:

$$\Delta\varphi_{-3dB} = \frac{1 - |S_{11}|^2}{|S_{11}|}$$

questo, dal momento che $R_1 = R_2$, e che abbiamo scritto invece di esso il parametro scattering in modulo quadro. Questa formula non va bene per il progetto, dal momento che la specifica precedentemente introdotta era in Hz, ossia era nel dominio della frequenza, mentre questo è soltanto uno sfasamento, collegato alla frequenza ma non in maniera immediatissima. Precedentemente, al fine di passare all'unità di misura della frequenza, era stato usato l'incremento finito; ora si farà la stessa cosa, tenendo tuttavia conto delle non idealità aggiuntive introdotte dagli specchi di Bragg. Usiamo dunque gli incrementi finiti:

$$\Delta\varphi_{-3dB} = \frac{d\varphi}{d\omega} 2\pi \Delta f_{-3dB}$$

φ tuttavia è semplicemente metà del RTPS; si può dunque scrivere che

$$\varphi = \angle S_{11} - \beta l$$

Si noti che ora siamo tuttavia costretti a tenere conto di S_{11} e della sua fase, dal momento che, come detto, gli specchi dipendono dalla frequenza; si ha dunque, volendo calcolare la derivata prima scritta:

$$\frac{d\varphi}{d\omega} = \frac{d}{d\omega} \angle S_{11} - \frac{d\beta}{d\omega} l = \frac{d}{d\omega} \angle S_{11} - \frac{l}{v_g}$$

A questo punto è necessario introdurre la prima approssimazione grossa: come visto precedentemente, per la teoria delle piccole riflessioni, si ha che:

$$S_{11} = j\Gamma_0 e^{-j\phi(2N-1)} \frac{\sin(2N\phi)}{\cos\phi}$$

dove ϕ è

$$\phi = k_{z1}d_1 = k_{z2}d_2$$

quindi, si può dire che:

$$\angle S_{11} = \frac{\pi}{2} - (2N-1)\phi$$

quindi

$$\frac{d}{d\omega} \angle S_{11} = -(2N-1) \frac{d\phi}{d\omega}$$

A questo punto, un po' di manipolazione:

$$\phi = k_{z1}d_1 = k_0 n_1 \cos\vartheta_1 d_1 = \frac{\omega}{c} n_1 \cos\vartheta_1 d_1$$

a questo punto, multiplico e divido per ω_0 :

$$= \frac{\omega_0}{c} \frac{\omega}{\omega_0} n_1 \cos\vartheta_1 d_1$$

si osservi tuttavia che

$$\frac{\omega_0}{c} n_1 \cos\vartheta_1 d_1 = k_{z1}|_{f=f_0} d_1$$

ma, a queste condizioni, si ha che $d_1 = \lambda_{g1}/4$, e che quindi si ha uno sfasamento complessivamente pari a $\pi/2$:

$$k_{z1}d_1 = \frac{2\pi}{\lambda_{g1}} \frac{\lambda_{g1}}{4} = \frac{\pi}{2}$$

quindi, la formula è:

$$\phi = \frac{\omega}{\omega_0} \frac{\pi}{2}$$

da qua, finalmente, è possibile dire che:

$$\begin{aligned}\frac{d}{d\omega} \angle S_{11} &= -(2N - 1) \frac{\pi}{2} \frac{1}{\omega_0} = \\ &= -\frac{2N - 1}{4f_0}\end{aligned}$$

A questo punto, è possibile sostituire ciò nell'espressione madre, e trovare la relazione tra Δf e $\Delta\varphi$:

$$\Delta\varphi_{-3\text{dB}} = -\left(\frac{2N - 1}{4f_0} - \frac{l}{v_g|_{f=f_0}}\right) 2\pi\Delta f_{-3\text{dB}}$$

Dal momento che il mezzo è non dispersivo, si ha che:

$$\frac{l}{v_g} = \frac{l}{v_f} = \frac{ln_1 \cos \vartheta_i}{c}$$

Ma, come già ricavato prima a partire dalle informazioni su f_0 , si ha che:

$$l = m \frac{\lambda_{g1}}{2}$$

quindi, sostituendo qua dentro, si può trovare che:

$$\begin{aligned}\frac{l}{v_g} &= \frac{m\pi}{\frac{\omega_0}{c} n_1 \cos \vartheta_i} = \frac{n_1 \cos \vartheta_i}{c} = \frac{m\pi}{\omega_0} = \\ &= \frac{m}{2f_0}\end{aligned}$$

Dunque, si ha che

$$\Delta\varphi_{-3\text{dB}} = -\left(\frac{2N - 1}{4} + \frac{m}{2}\right) 2\pi \frac{\Delta f_{-3\text{dB}}}{f_0}$$

Questo lega il $\Delta\varphi$ al Δf .

A questo punto, abbiamo due espressioni di $\Delta\varphi$: quella appena ricavata, e quella dipendente da $|S_{11}|$. Al fine di effettuare il progetto, da questo punto in poi, è necessario procedere iterativamente: si deve provare a sintetizzare gli specchi, a partire da diversi valori di N . Da $N = 1$, si ricava il S_{11} , quindi si trova l'espressione:

$$\overline{\Delta f_{-3\text{dB}}} = \frac{\frac{1 - |S_{11}|^2}{|S_{11}|}}{2\pi \left(\frac{2N - 1}{4} + \frac{m}{2}\right)} f_0$$

Si valuta per N questo $\overline{\Delta f_{-3\text{dB}}}$, e si effettua un confronto: se

$$\overline{\Delta f_{-3\text{dB}}} \leq \Delta f_{-3\text{dB,specifica}}$$

allora si può concludere il ciclo, dal momento che il numero di celle permette di soddisfare la specifica sulla minima larghezza di banda a - 3 dB.

4.3.1 Analisi

Una volta ultimato il progetto, seguendo i passi precedenti, è necessario effettuare l'analisi: essa si fa o utilizzando la teoria delle linee, o utilizzando la teoria degli interferometri di Fabry-Perot introdotta, basata sull'approccio con le matrici scattering. Le riflettività degli specchi possono essere ricavate mediante le nozioni precedentemente introdotte sugli specchi di Bragg, essendo per l'appunto gli specchi così realizzati.

Al fine di avere un'idea tuttavia di ciò che si deve ottenere in linea di massima, si vuole proporre un'idea qualitativa del risultato. Una curva di un interferometro di Fabry-Perot, ha un andamento di questo genere:

Questa curva è sostanzialmente basata su un'ipotesi fondamentale: il fatto che gli specchi non abbiano un comportamento particolare al variare della frequenza (dal momento che è stata trovata, a patto di avere S'_{ij} costanti rispetto alla frequenza). Nel nostro caso, tuttavia, gli specchi hanno un comportamento chiaramente dipendente dalla frequenza, dal momento che essi sono composti da reticoli di Bragg; un reticolo di Bragg ha trasmissività e riflettività di questo tipo:

La riflettività dello specchio di Bragg deriva dal fatto che si ha un fenomeno di riflessione totale frustrata, motivabile dalla presenza di onde di Bloch evanescenti nella struttura periodica.

Per avere la curva risultante, è necessario sostanzialmente giustapporre la curva del Fabry-Perot su quella degli specchi di Bragg (le curve di Bragg saranno leggermente più accentuate, rispetto a quelle del singolo, essendo gli specchi in questione due); si ha quindi ciò:

Si ha sostanzialmente un andamento caratteristico degli specchi di Bragg, con un picco di risonanza dovuto all'effetto della configurazione "Fabry-Perot": all'interno della banda attenuata dello specchio si ha un picco di trasmissione.

Come noto dalla teoria degli interferometri Fabry-Perot, dopo un certo FSR (Free-Spectral Range) si ha un altro picco di trasmissione, ma questo generalmente si trova a frequenze troppo alte per essere apprezzabile: di solito, infatti, si ha che questo secondo picco è a frequenze troppo elevate, quindi finisce per essere in un range di frequenze in cui gli specchi di Bragg

sono trasparenti, e quindi in una situazione in cui un picco non è apprezzabile, essendo la trasmissività già molto elevata: al fine di vedere un picco, è necessario trovarsi in un “panorama”.

La funzione totale è, come detto, la giustapposizione delle due: si ha quindi la giustapposizione di un andamento passa banda e di un andamento rigettabanda: questo significa che il comportamento finale è un passa banda, dove la banda è molto stretta rispetto a quella rigettata. A condizione che il FSR sia sufficientemente elevato, e quindi l (la lunghezza della cavità) ridotta, si avrà solo un picco (che peraltro è quanto richiesto dalle applicazioni tipiche di questo sistema, ossia il LASER). La larghezza della valle su cui sorgono i picchi, come si vede dalla matematica precedentemente presentata, dipende dagli specchi, e a loro volta da n_1 e n_2 : tanto più gli indici di rifrazione sono diversi, tanto più la stop-band è ampia. Esiste una formula approssimata che può stimare la stop band:

$$\frac{\Delta\omega_{\text{gap}}}{\omega_0} \sim \frac{2}{\pi} \frac{\Delta n}{n}, \quad n = \frac{n_1 + n_2}{2}, \quad \Delta n = n_1 - n_2$$

Δf potrebbe per esempio essere 1/10 della stop band (questa potrebbe essere una richiesta ragionevole).

Il comportamento degli specchi di Bragg, dunque, è buono solo nella stop band; questo significa che il materiale attivo deve comportarsi come tale esclusivamente in una sottobanda, rispetto alla stop band: se così non fosse, in uscita dagli specchi non si avrebbe una sola riga spettrale, bensì svariate componenti: tutto ciò serve per prendere, a partire da un materiale attivo, una singola frequenza, come in un oscillatore; non ha senso che il rigettabanda non riesca a rigettare tutte le frequenze, perchè in caso diverso si avrebbe a che fare con dispositivi che non si comportano correttamente, che non produrrebbero luce coerente. Quello che si vede in uscita da un LASER è lo spettro di emissione della cavità, ossia quello del materiale attivo filtrato.

Si noti che il discorso che si sta facendo riguarda solo il comportamento lungo z : trasversalmente, consideriamo le interfacce piane e onde piane, ma ciò non corrisponde alla realtà: nella realtà si possono avere gli specchi curvi, oppure un LASER in fibra ottica.

Note conclusive

Una nota conclusiva: molto spesso, per descrivere questa struttura, si sentono discorsi magari non sbagliati, ma molto *fumosi*: si suol dire che il sistema sia sostanzialmente una cavità risonante, quindi una struttura che presenta dei modi di risonanza (non perfetti, dal momento che si ha a che fare con perdite dovute all'irradiazione dell'energia verso l'esterno, quindi l'ampiezza

di oscillazione è non costante). Dato che il collegamento con l'esterno è tramite un coefficiente di trasmissione molto piccolo, si può pensare che l'onda si trovi davanti un muro impenetrabile, dunque dovrebbe essere interamente riflessa, ma grazie alla presenza di un modo di risonanza l'onda sarebbe tale da eccitare la risonanza e mediante essa arrivare dall'altra parte. Questo discorso non è per forza sbagliato, ma fare un progetto a partire da esso è sostanzialmente impossibile; utilizzare invece l'approccio basato sullo studio come se fosse (ed in effetti è) un interferometro di Fabry-Perot, basta mettere insieme l'approccio scattering con le matrici scattering degli specchi di Bragg, e tutto è finito.

Si consideri inoltre la struttura, sotto questo punto di vista:

Si immagini che $l = \lambda_{g1}/2$: si avrebbe a che fare con una struttura periodica ovunque, tranne che nello strato centrale. Si parla, di solito, di *difetto* rispetto alla struttura periodica: se infatti $l = \lambda_{g1}/4$, si ha una struttura periodica senza discontinuità, quindi sostanzialmente uno specchio; si vuol dire che questo "difetto" sia quello che permette di aggiungere un picco di risonanza (si parla di questa struttura come di un $\lambda_g/4$ shift), ma anche in questo caso la formulazione non permette di arrivare a risultati interessanti.

Capitolo 5

Guide d'onda dielettriche

5.1 Introduzione

Si vuole a questo punto introdurre qualcosa di molto diverso rispetto a prima: quanto discusso finora, infatti, era un esempio di una categoria di dispositivi noti come PBG Materials (Photonic Band-Gap); queste sono strutture periodiche, ma in **una dimensione**: tutto ciò che è stato studiato finora presenta periodicità su una sola dimensione. Negli ultimi anni, tuttavia, si è arrivati a immaginare, studiare e realizzare strutture di questo tipo:

Ciò può essere realizzato in molti modi: cilindri di dielettrico con $n_{\text{dielettrico}} > n_{\text{ambiente}}$ (per esempio immettendo questi cilindri di dielettrico in aria), o il contrario (per esempio considerando un vetro con delle aperture di forma cilindrica e riempite d'aria). In questo caso, si hanno due spazature: considerando che questo piano sia l'asse (x, y) , si avranno spazature d_x su una direzione, d_y sull'altra, e quindi si avrà a che fare con due periodicità spaziali, ossia periodicità in due direzioni.

Data un'onda elettromagnetica che si propaga con un vettore \underline{k} che sta su un piano, si avrà dunque qualcosa di diverso rispetto a prima: prima la costante di propagazione era un k_z , dove la periodicità del sistema era solamente verso z (infatti, la struttura periodica presentava la periodicità solo lungo la direzione di propagazione, la quale era per l'appunto z), e ciò portava ad avere onde che potessero viaggiare in una sola direzione, con due versi possibili, o, in altre parole, onda progressiva e onda regressiva. Ora ciò non è più vero, dal momento che \underline{k} può assumere qualsiasi valore, e quindi qualsiasi direzione nel piano, e quindi qualsiasi direzione può essere la direzione di propagazione di un'onda piana.

Ciò che capita ora è che a seconda di \underline{k} si hanno diverse curve di dispersione, ossia curve $\omega(\underline{k})$; nella fattispecie, stiamo parlando di strutture perio-

diche, dunque la \underline{k} che stiamo considerando è una \underline{k}_B , ossia una costante di propagazione di Bloch (dal momento che le onde che stiamo considerando non sono tradizionali onde piane, bensì onde di Bloch). Queste strutture, a seconda di questo \underline{k} , presentano diverse zone di gap:

Al fine di studiare le curve di dispersione in questo ambito, si deve considerare un certo contorno sul piano (x, y) , ossia una “prima cella di Brillouin”, quindi si rappresenta su questa cella un contorno indicante i vari valori assunti da \underline{k} , e infine si “distende” questo insieme di valori su un asse, ottenendo delle curve che indicano il fatto che, per certi ω , non ci sono \underline{k} reali: esattamente come le curve di dispersione relative alle curve di Bloch precedentemente introdotte.

Se dunque si manda un fascio LASER con la frequenza appartenente a questo gap, capita ciò che capitava allo specchio: si ha una forte riflessione, dal momento che all’interno della struttura periodica il \underline{k} diventa complesso, e l’onda si attenua. Ciò che si può fare è dunque realizzare specchi bidimensionali nel piano, semplicemente cambiando la spaziatura di queste inclusioni dielettriche. La stop band si può modificare, ingegnerizzare, scegliendo idoneamente d_x e d_y : è una sorta di specchio sintonizzabile, nel senso che, prima di costruirlo, se ne possono ingegnerizzare le caratteristiche.

Ci sono realizzazioni di strutture che sfruttano questi concetti, eliminando però una delle righe:

Questa struttura elimina una delle righe di barre. A cosa serve ciò? Cosa fa? Beh, questa struttura è una specie di guida d’onda: la parte di sopra si comporta come uno specchio, per onde che vi incidono dall’interno: una volta che l’onda, andando in su, trova uno specchio, viene riflessa da essa e va avanti; poi, va in giù, incontra un altro specchio, viene riflessa, e così via.

Le guide d’onda normali funzionano sul principio della riflessione totale: invece che avere un “groviera” come in questo caso, si ha una struttura continua. La cosa però ha dei grossi svantaggi: la riflessione, in una struttura di questo tipo, è totale solamente se l’angolo di incidenza sulle pareti è maggiore di un certo angolo critico, ricavabile con la solita teoria; se tuttavia si prende questa guida e la si fa curvare, l’onda, che va avanti a suon di riflessioni, potrebbe trovarsi a incidere con un angolo inferiore rispetto a quello critico, e dunque uscire dalla struttura guidante!

Se nella struttura tradizionale dunque non si possono fare curve se non molto poco accentuate, sulle strutture periodiche non si corre un rischio del genere: per come sono progettati gli specchi, essi fanno “da specchi” in ogni direzione! È dunque possibile fare percorsi anche molto curvi, senza avere perdite di irradiazione: si possono avere, in ogni caso, forti riflessioni. Ciò permette di costruire circuiti ottici con curve a raggio strettissimo, anche di

qualche μm , a patto di accettare una contropartita: ciò, da studiare, è molto complesso, e quasi sempre necessita metodi numerici.

5.1.1 Principali tipi di guide d'onda ottiche

Quello che si farà ora sarà molto simile a ciò che è stato fatto nel precedente capitolo sulle onde piane, dove però la variabile che veniva rappresentata spettralmente era x , con la sua ξ corrispondente. A partire da un approccio simile a questo, solo applicato a z e alla sua β corrispondente nel dominio spettrale, si intende studiare l'argomento "guide d'onda": come noto, l'energia elettromagnetica, se "lasciata libera", si propaga in ogni direzione; l'obiettivo di questo capitolo, come d'altra parte era, nei corsi di Campi Elettromagnetici, parlando di guide d'onda metalliche, è quindi quello di trovare strutture che possano confinare, "intubare" l'energia elettromagnetica nello spettro ottico in modo da farla propagare solo lungo una certa direzione.

In ambito ottico, la guida più conosciuta è la fibra ottica: essa è costituita da due mezzi dielettrici, per cui $n_1 > n_2$.

In questo caso, l'interfaccia, invece che essere piana, è cilindrica, tuttavia si può dimostrare che capita qualcosa di molto simile rispetto a quanto visto in ambito di onde piane, e si ha un fenomeno di riflessione totale, tale per cui si ha un campo evanescente esternamente a n_1 , e quindi il campo rimane sostanzialmente confinato solo in n_1 (detto "core") e nelle vicinanze (dal momento che nel "cladding" il campo elettromagnetico è esponenzialmente decrescente). Oltre a questi due strati, una fibra ottica poi ha un certo insieme di strati protettivi, che la rendono maneggevole, ma otticamente gli unici due strati di interesse sono il core e il cladding.

Spesso, per fare amplificatori, un altro tipo di guida ottica che si trova è il seguente:

Si considera un bulk n_2 , una regione con indice più alto, e aria fuori da tutto; il campo, anche in questo caso, è una gaussiana discendente. Anche in questo caso, la regione con $n_1 > n_2 > n_{\text{aria}}$ è quella che tiene sostanzialmente confinato il campo. La regione con n_1 può avere forma generica. L'importante, in ogni caso, è avere una regione con un $n_1 > n_2$, in modo che in qualche senso il campo rimanga confinato dentro. Questo è il modo fondamentale, dal momento che ha "solo una gobba", ma in realtà potrebbe anche avere più variazioni all'interno del cilindro centrale. Tanto più la guida è grande rispetto a λ , tanti più modi sono sopra taglio.

Le guide dielettriche sono "strane": le pareti, infatti, fanno da specchio solamente quando l'angolo di incidenza è superiore a quello critico; questo porta ad avere una differenza sostanziale rispetto alle guide d'onda metalliche: se nelle guide d'onda metalliche i possibili modi di propagazione erano

infiniti, di cui solo alcuni sopra taglio, ora (nelle guide dielettriche) non è più possibile avere a che fare con infiniti modi, dal momento che se essi sono tali da incidere sotto l'angolo critico, non si propagano, e anzi escono dalla struttura guidante: c'è sempre un numero **finito** di modi.

Altra struttura reale, spesso utilizzata, è la *struttura ridge*: essa è composta da un n_{bulk} , un n_0 , e quindi un n_{film} con un "ridge"; anche in questo caso il modo fondamentale è una sorta di gaussiana.

Il problema delle strutture presentate è che non si possono analizzare in modo semplice, analitico: di tutte le strutture realmente utilizzate, l'unica che si possa analizzare analiticamente è la fibra ottica, ma essa comunque è molto complicata da studiare.

5.2 Analisi di una struttura accademica

Dal momento che tutte le strutture proposte sono molto complicate da studiare, ciò che verrà ora fatto, al fine di introdurre i concetti, è studiare una guida d'onda che in pratica non si utilizza, ma che comunque permette di ottenere dei concetti esportabili in strutture più complicate; questa è una struttura ragionevolmente semplice da studiare.

La struttura è la seguente:

Si può vedere che si hanno sostanzialmente tre strati, come nello *slab*. La direzione di propagazione è quella lungo l'asse z , che in questo caso (come d'altra parte anche nell'ambito dello slab) è la direzione orizzontale rispetto al nostro punto di vista; x è ancora una volta l'asse verticale, y è uscente dal foglio (ma non interessante dal momento che ancora una volta si considera l'ipotesi di "problema bidimensionale"). Si hanno tre strati, ossia tre valori di indice di rifrazione: per x grandi, si ha lo strato di "cover", con un n_c corrispondente; quindi, si ha un film, con strato n_f , depositato su un substrato con indice n_s . In generale, si farà in modo che:

$$n_f > n_s > n_c$$

Si ha quindi una propagazione di onde piane, incidenti all'interno di questa struttura, che rimbalzeranno da una parete all'altra. Questo, quindi, può essere analizzato con la teoria precedentemente introdotta, supponendo una condizione di riflessione totale alle due interfacce, dove però si tenga ben presente che l'onda incide **da dentro**.

5.2.1 Analisi modale della struttura

L'obiettivo a questo punto è quello di determinare i modi della struttura. I modi, come si ricorderà, sono soluzioni delle equazioni di Maxwell in assenza di sorgenti, dunque problemi omogenei. I modi, nella fattispecie, sono particolari configurazioni di campo che hanno una dipendenza funzionale del tipo:

$$\underline{E}(x, y, z) = \underline{u}(x, y)e^{-j\beta z}$$

$$\underline{H}(x, y, z) = \underline{v}(x, y)e^{-j\beta z}$$

Ossia, si ha una certa dipendenza da x e y che dipende dalla struttura della guida, ma di essi si sa una cosa: essi si mantengono costanti sezione per sezione, a meno di un certo fattore di fase dipendente solo da z : $e^{-j\beta z}$. Di tutte le possibili topografie di campo, quindi, i modi sono quelle che hanno questo tipo di comportamento: una sorta di “traslazione rigida” del campo lungo la sezione trasversale, a meno di un fattore di fase che dipende dalla sola coordinata longitudinale, ossia dalla sola posizione lungo la direzione di propagazione dell'informazione. Ciò si può anche vedere in questo modo:

$$\underline{A}x = \lambda x$$

dove \underline{A} è l'operatore di propagazione (per esempio potrebbe essere ∇^2): applicato \underline{A} , operatore di propagazione, quindi di “traslazione” del campo, si può pensare a tutto ciò come a un problema agli autovalori, e quindi capire che λ è semplicemente l'autovalore rispetto a cui si calcola la traslazione; \underline{u} è l'autofunzione modale, ossia è quella configurazione di campo trasversale tale per cui l'operatore di traslazione, di propagazione, applicato su di esso, non lo modifichi, introducendo solo un termine di fase: quello dipendente da β . β sarà dunque l'autovalore relativo a questo problema.

Concettualmente, ciò è molto simile a quanto visto precedentemente sullo slab (la struttura era a tutti gli effetti uguale, a patto di tenere conto del fatto che essa è semplicemente “ruotata di $\pi/2$ ” sul piano xz): ai tempi si era scritto che, data un'onda incidente con un certo ϑ_i , si era legato ϑ_i a un certo ξ , e si era detto che il campo riflesso, incidente e trasmesso hanno la stessa dipendenza da $e^{-j\xi x}$, quindi ciò permetteva di avere già soddisfatte le condizioni al contorno.

Un modo (tornando alla struttura che stiamo per analizzare) è costituito da due elementi: uno è la β , la costante di propagazione, e l'altro è la relazione di impedenza, l'impedenza modale. È un po' come dire che serve autovalore

(il β) e autovettore (relazione tra campo elettrico e campo magnetico, ossia impedenza modale).

Si noti che se si conosce il campo elettrico in un punto, nello spazio libero, non è possibile calcolare il campo magnetico; questo si può invece fare, se si conosce il campo elettrico in un certo insieme di punti: vale infatti, senza voler andare troppo lontano, l'equazione di Maxwell:

$$\nabla \times \underline{E} = -j\omega\mu\underline{H}$$

conoscendo tutto, se si hanno diversi punti, è possibile calcolare le derivate e quindi trovare \underline{H} .

La relazione di impedenza precedentemente introdotta, ossia:

$$\underline{E} = Z_0 \underline{H} \times \hat{n}$$

vale solo per campi elettromagnetici di una particolarissima forma: per le onde piane! In questo modo so qual è il campo elettrico rispetto al campo magnetico. Questa cosa si ha, anche, modo per modo: per ciascun modo si definisce una particolare relazione di impedenza, e si avran diverse impedenze modali per ciascun modo.

L'obiettivo a questo punto è quello di studiare i modi di propagazione di questa guida a lastra, e per fare ciò possiamo utilizzare il formalismo delle linee di trasmissione, al fine di trovare qualcosa di semplice: z sarà come al solito la direzione di propagazione lungo la struttura guidante. In questo caso, le linee di trasmissione saranno "trasversali" (in realtà, come prima):

Si vuole chiarire un fatto: la propagazione del segnale, dell'informazione, avviene lungo z ; la variabile spettrale associata a z è la β , ossia la costante di propagazione; β , in altre parole, rappresenta il ritmo della variazione della fase, per unità di lunghezza, lungo z . Prima la costante di propagazione era caratterizzata da ξ , ora da β ; l'espressione dunque sarà:

$$k_{xf} = \sqrt{k_0^2 n_f^2 - \beta^2} \quad Z_{\text{infty}}^{\text{TE}} = \frac{\omega\mu}{k_x} \quad Z_{\text{infty}}^{\text{TM}} = \frac{k_x}{\omega\varepsilon}$$

dove

$$\beta = \frac{d\phi}{dz}$$

In sostanza, queste formule sono le stesse di prima, ottenute "rinominando gli assi".

Si vuol fare presente un fatto: le linee di trasmissione equivalenti sono non nella direzione di propagazione, ma in quella trasversale; come mai? Beh, come già si sarà capito dalla lettura dei capitoli precedenti, una linea di

trasmissione, oltre a essere il modello di un coassiale o di una bifilare, è anche un modello matematico per la soluzione di equazioni differenziali, della forma ben nota; dal momento che il nostro obiettivo è quello di trovare i modi, una tecnica per farlo è quella di utilizzare queste linee in modo che, applicando “circuitalmente” le condizioni al contorno, si possano ottenere dei risultati; questo approccio funziona ed è molto semplificato.

Al fine di proporre l’esempio di partenza, si vuole introdurre il concetto di “specchio perfetto”: nei corsi di base di Campi Elettromagnetici, l’ipotesi sotto cui si fa l’analisi delle guide d’onda è il fatto di avere pareti in PEC; in altre parole, definita una certa geometria (quella “classica” è quella con sezione trasversale rettangolare), si fa in modo da annullare il campo elettrico normale alla superficie della geometria scelta, ottenendo, di fatto, proprio la condizione PEC: Perfect Electrical Conductor. In ambito ottico, più che parlare di PEC, ha senso parlare di specchi, ossia di superfici che, data un’onda incidente, la riflettono completamente. Si vuole introdurre una coppia di definizioni più dettagliate.

- Per specchio perfetto, ossia per quello che potrebbe essere il “PEC” riportato in condizioni ottiche, si ha uno specchio che deve funzionare per ogni valore di angolo, e quindi non solo per $\vartheta_i > \vartheta_c$. Dunque, si ha:

$$\Gamma = -1 \quad \forall \vartheta_i$$

- Uno specchio perfetto è una cosa ben diversa da uno specchio ottenuto per riflessione totale: questo, infatti, ha due condizioni:

$$|\Gamma| = 1, \quad \vartheta_i > \vartheta_c$$

$$\angle \Gamma \text{ dipende da } \vartheta_i$$

Quello degli specchi perfetti è un esempio semplice da analizzare, e per questo può essere utilizzato per introdurre l’argomento.

5.2.2 Analisi di una guida a facce piane parallele

Si vuole analizzare, a questo punto, una struttura di questo tipo:

Questa è una struttura già nota da Campi Elettromagnetici, ma presenta forti analogie con quanto si vuole analizzare, dunque è un buon punto di partenza; essa verrà analizzata con il formalismo delle linee di trasmissione

caricate¹. Sostanzialmente, dunque, è come avere una linea di trasmissione chiusa su due corto-circuiti.

Al fine di modellare il sistema, dunque, possiamo scriverne le equazioni dei telegrafisti:

$$\begin{cases} -\frac{dV}{dx} = jk_x Z_\infty I \\ -\frac{dI}{dx} = jk_x Y_\infty V \end{cases}$$

eliminando la corrente e sostituendo, si trova l'equazione d'onda per la tensione:

$$\frac{d^2V}{dx^2} + k_x^2 V = 0$$

A questo punto, vogliamo trovare i modi TE di questa struttura: $E_y \neq 0$, $H_{x,z} \neq 0$. Trovare i modi, come detto, significa trovare una espressione $E_y(x, z)$ (si considera y ininfluente) con una dipendenza del tipo:

$$E_y(x, z) = V_0^+ u(x) e^{-j\beta z}$$

Ciò che si può fare per il problema è porre l'eguaglianza

$$V_0^+ u(x) = V(x)$$

nelle equazioni delle linee di trasmissione; si trova, alla fine, il seguente problema, sostituendo l'espressione di k_x :

$$\frac{d^2u}{dx^2} + (k_0^2 - \beta^2)u(x) = 0$$

(si ha solo k_0 supponendo che tra i due PEC ci sia il vuoto). Il problema non è ancora completamente definito: bisogna infatti anche introdurre le condizioni al contorno, ossia:

$$u\left(\frac{d}{2}\right) = u\left(-\frac{d}{2}\right) = 0$$

Data anche la condizione al contorno, il problema è ben posto e può essere risolto; prima di tutto, dunque, si deve trovare l'integrale generale dell'equazione differenziale, e risolverla, ma ciò è semplice, dal momento che essa è

¹Il modo più corretto sarebbe basato su una applicazione del formalismo di Marcuvitz e Schwinger, risolvendo il problema agli autovalori ($\nabla_t^2 \Phi + k_t^2 \Phi = 0$ e $\nabla_t^2 \Psi + k_t^2 \Psi = 0$), applicando le condizioni al contorno, ma il risultato è equivalente e non si vuole introdurre una matematica troppo complicata

semplicemente un'equazione delle linee di trasmissione, e quindi un'equazione d'onda, con una componente progressiva e una regressiva:

$$u(x) = u_0^+ e^{-jk_x x} + u_0^- e^{jk_x x}$$

A questo punto, su questa, impongo le condizioni al contorno:

$$0 = u_0^+ e^{-jk_x \frac{d}{2}} + u_0^- e^{jk_x \frac{d}{2}}, \quad x = \frac{d}{2}$$

$$0 = u_0^+ e^{jk_x \frac{d}{2}} + u_0^- e^{-jk_x \frac{d}{2}}, \quad x = -\frac{d}{2}$$

Questo significa avere due equazioni in due incognite, quindi si può scrivere tutto sotto forma di sistema di equazioni algebriche:

$$\begin{bmatrix} e^{-jk_x \frac{d}{2}} & e^{jk_x \frac{d}{2}} \\ e^{jk_x \frac{d}{2}} & e^{-jk_x \frac{d}{2}} \end{bmatrix} \begin{bmatrix} u_0^+ \\ u_0^- \end{bmatrix} = \underline{0}$$

Il sistema lineare è omogeneo, come del resto ci si poteva aspettare: il problema differenziale di partenza era infatti omogeneo. Al fine di avere soluzioni oltre a quella banale, è necessario chiedere che il determinante della matrice, chiamata \underline{A} , sia nullo:

$$\det \{ \underline{A} \} = e^{-jk_x d} - e^{jk_x d} = -2j \sin(k_x d) = 0$$

questa cosa si ha per

$$k_x d = m\pi, \quad m = 0, \pm 1, \pm 2, \pm 3 \dots$$

Quindi:

$$k_{x,m} = \frac{m\pi}{d}$$

questi sono i valori degli autovalori del sistema! Data $u(x)$ la autofunzione (autovettore), $k_{x,m}$ è l'autovalore.

Abbiamo trovato, per questo sistema, gli autovalori; a questo punto, è di nostro interesse trovare anche le espressioni degli autovettori: per fare ciò, è necessario risolvere il sistema lineare; si può verificare che la soluzione è:

$$u_0^- = -e^{-jk_{x,m} d} u_0^+$$

tuttavia, si può vedere che:

$$k_{x,m} d = m\pi$$

quindi

$$= -u_0^+ e^{-jm\pi}$$

al fine di effettuare l'analisi del risultato, conviene distinguere il caso di m pari a quello di m dispari.

- se m è dispari, si ha:

$$u_m(x) = 2u_0^+ \cos\left(\frac{m\pi}{d}x\right)$$

- se m è pari, si ha:

$$u_m(x) = -2ju_0^+ \sin\left(\frac{m\pi}{d}x\right)$$

si parla di u_m : si parla delle autofunzioni u associate a un particolare valore di m .

Si noti che gli interi hanno doppio segno: se invece di $m = 1$ ho $m = -1$ non c'è problema: per m dispari infatti la funzione è un coseno, e il coseno è una funzione notoriamente pari; nel caso di m pari, invece, la funzione cambia segno; questo cambia l'autovettore, ma non nella direzione! Ciò che conta di un autovettore non è tanto il segno quanto la sua direzione, quindi va comunque tutto bene.

Queste sono le autofunzioni, ma esse non sono ancora completamente specificate: ciò che si può a questo punto richiedere è il fatto che esse siano ortonormali, ossia normalizzate rispetto a una certa costante. Ortogonali in realtà lo sono già, quindi ciò che si deve richiedere è:

$$\int_{-\frac{d}{2}}^{+\frac{d}{2}} u_m(x) dx = 1$$

si ricava:

$$u_m(x) = \sqrt{\frac{2}{d}} (\sin)(\cos) \left(\frac{m\pi}{d}x\right)$$

Le distribuzioni di campo sono sinusoidali, dunque sono quelle ben note dalla teoria delle guide d'onda; m dà, sostanzialmente, solo il numero di arcate con cui si ha a che fare.

Curve di dispersione

Abbiamo trovato che:

$$k_x = \sqrt{k_0^2 - \beta^2}$$

quindi:

$$\beta = \sqrt{k_0^2 - \left(\frac{m\pi}{d}\right)^2}$$

Questo ha una conseguenza importante: data ω (e quindi k_0) come variabile, si ha ciò:

si ha che $\beta = 0$ quando il radicando è nullo; al crescere di ω , quindi, si ha un'iperbole. Essendo un'iperbole, si ha un asintoto: per $k_0 \rightarrow \infty$, infatti, l'iperbole va sostanzialmente come:

$$\beta \rightarrow \frac{\omega}{c}$$

questo è l'asintoto. Questo significa che, per $\omega \rightarrow \infty$, l'onda “se ne frega” delle pareti: essa si propaga lungo z come se fosse nel vuoto, cosa che invece non capita nel vuoto.

Per $\omega < \omega_c$, β diventa immaginario, dunque si ha una curva sostanzialmente ellittica (o circolare); essa vale:

$$\beta = -j\frac{m\pi}{d}$$

Questa, quindi, è la curva di dispersione per i modi della guida, costituita da specchi perfetti. Quella appena vista è la teoria analitica per trovare i modi della guida: di questi, quindi, si possono per esempio volere quelli per $m = 1$, e vedere che essi sono interpretabili come la somma di due onde piane, che “vanno a zig zag” nella struttura.

Osservazioni aggiuntive sull'esempio svolto

Si vuole a questo punto proporre qualche osservazione su ciò che è stato appena fatto.

Come detto, si era partiti dalla “guida a specchi perfetti”, ossia all'equivalente PEC in ambito ottico; come visto, si è messo il riferimento “in mezzo” alla struttura: questa fatto spesso è molto utile, dal momento che la struttura è visibilmente simmetrica. Il fatto di sfruttare le simmetrie di una struttura per la definizione del sistema di riferimento in cui essa viene

studiata è molto utile, in strutture complicate, dal momento che la presenza di simmetria di una struttura rispetto al sistema di riferimento in cui si calcolano le funzioni si va a rispecchiare nelle funzioni che modellano il comportamento del sistema: per esempio una sorgente potrebbe eccitare i soli modi pari di una struttura e quindi, nel caso di presenza di simmetrie di questo genere, è possibile senza alcun problema dimezzare i conti da fare.

Si vuole proporre un'analogia, a questo punto, tra il metodo dell'espansione modale e l'applicazione del principio delle immagini: la struttura che si sta studiando, infatti, potrebbe essere eccitata da una sorgente filiforme (per esempio un dipolo), posto in mezzo alla struttura, per $x = 0$ rispetto al riferimento introdotto, proprio al fine di eccitare solo alcuni modi (per esempio, in una struttura di questo genere, si ecciterebbero solo i modi pari). Un modo di studiare ciò è basato sul prendere l'espressione della funzione, proiettarlo nella base delle autofunzioni modali, quindi trovare i modi eccitati; si troverebbe, in questo caso, che solo i modi pari sono eccitati, mentre quelli dispari (quelli che per $x = 0$ sono nulli) non lo sono. Un metodo alternativo a questo è basato sull'applicazione del principio delle immagini:

In cosa consiste l'applicazione? Si deve fare qualcosa di particolare: si ha una sorgente tra due piani metallici; quello che si può fare è considerare un problema equivalente rispetto a uno dei due piani, quindi sdoppiarlo, ottenendo il secondo passo: il solito piano (nell'esempio quello in alto) che resta dov'è, idem la sorgente, quindi una seconda sorgente "girata al contrario", e un altro piano metallico, che rappresenta l'immagine del piano precedente; questo si può quindi riapplicare al piano sotto, trovando qualcosa di analogo, e così via, fino a ottenere sostanzialmente infinite sorgenti e un solo piano metallico. Il risultato, quindi, è avere infinite sorgenti equivalenti, al posto di una sola sorgente con due piani metallici.

Applicare ciò non è necessario, ma talvolta può essere utile: se si vuole calcolare il campo a una certa distanza (abbastanza lontano dalla sorgente) l'espansione modale è il metodo più consigliato, dal momento che l'espansione modale, come detto, consiste nel prendere le funzioni, espanderle in una serie, e quindi dividere in vari contributi; lontano dalla sorgente, tuttavia, dei vari contributi solo pochi si manterranno intatti, dal momento che solo pochi dei modi si propagano, mentre gli altri si attenuano, dunque le espressioni dei campi sarebbero relativamente semplici; vicino alla sorgente, invece, si avrebbe qualcosa di duale: i campi elettromagnetici vanno secondo la funzione di Hankel bidimensionale:

$$E_y \propto H_0^{(2)}(k\rho) \sim \sqrt{\frac{2}{\pi k\rho}} e^{-j(k\rho - \frac{\pi}{4})}$$

Se si è vicini alla sorgente, questo secondo metodo è più comodo dal momento che si ha a che fare con degli esponenziali (al contrario del caso in cui si è lontano, dove il decadimento è come $\frac{1}{\sqrt{\rho}}$), quindi con espressioni rapidamente convergenti a 0. Si noti che in questo metodo non ha senso parlare di modi, dal momento che in questo caso si indaga sulle espressioni del “campo totale”, non dei singoli modi, quindi quella con cui si ha a che fare è la somma degli infiniti modi, dei modi già “mescolati” tra loro.

Come già accennato in precedenza, i due metodi sono legati tra loro: la trasformazione tra espansione modale e metodo basato sulle infinite sorgenti è dato dalla formula di Poisson (la stessa che si studia in Teoria dei Segnali), e ciò permette di usare un “trucco”: quando si devono fare calcoli, conviene essere nel dominio in cui la funzione converge più rapidamente, come in questo caso: a seconda del tipo di problema, ci saranno situazioni in cui la convergenza sarà più rapida nel dominio “diretto”, altre nel dominio “trasformato”.

Note sui problemi agli autovalori

Si vuole a questo punto introdurre un’ulteriore osservazione, riguardante i problemi agli autovalori. Precedentemente, è stato proposto e risolto il seguente problema:

$$\begin{cases} \frac{d^2u}{dx^2} + (k_0^2 - \beta^2)u(x) = 0 \\ u\left(\frac{d}{2}\right) = u\left(-\frac{d}{2}\right) = 0 \end{cases}$$

questo è un problema di Dirichlet, dal momento che quando si hanno condizioni che annullano il valore del campo ai bordi, si attribuisce alle condizioni al contorno il nome “condizioni di Dirichlet”; k^2 , qui, è l’autovalore del sistema. Questo sistema è stato risolto, quindi è stato detto che

$$k = \frac{m\pi}{a}, \quad m = \dots$$

sono stati gestiti i vari casi di valori assumibili da m , e sono state fatte alcune discussioni.

Precedentemente, è stata fatta un’introduzione generale sui problemi agli autovalori finito-dimensionali, ossia su matrici; si era detto che, al fine di avere soluzioni non banali, si era imposto che

$$\det \{ \underline{\underline{A}} - \lambda \underline{\underline{I}} \} = 0$$

Come si fa a risolvere invece un problema differenziale come quello proposto? Beh, in questo caso, chiamare la funzione “eig” di MATLAB non è

certo possibile, dal momento che essa opera su matrici, e qua di matrici non ce ne sono.

Dopo aver scritto la soluzione generale dell'equazione differenziale e applicate su di essa le condizioni al contorno, era stato ricavato un sistema lineare omogeneo del tipo:

$$\begin{bmatrix} e^{-jk_x \frac{d}{2}} & e^{jk_x \frac{d}{2}} \\ e^{jk_x \frac{d}{2}} & e^{-jk_x \frac{d}{2}} \end{bmatrix} \begin{bmatrix} u_0^+ \\ u_0^- \end{bmatrix} = \underline{0}$$

Al fine di avere soluzioni non banali, quindi, si imponeva che il determinante di questo sistema fosse non nullo.

Cosa stiamo facendo ora? Beh, abbiamo un problema **ricondotto** a un problema matriciale, e ora quindi si ha a che fare con una condizione del tipo

$$\det \{ \underline{A} - \lambda \underline{I} \} = 0$$

Ma prima, il problema era diverso: si aveva un problema differenziale, con k^2 come autovalore: prima si aveva un problema **differenziale con una dipendenza non lineare dall'autovalore**: un **problema agli autovalori non lineare**.

La differenza tra questo problema e i problemi normali è il fatto che i problemi lineari si risolvono con le tecniche standard, le normali routine dei vari programmi, quelli non lineari no: di solito si ha una soluzione diversa per ciascun problema.

Un metodo di solito utilizzato è basato sullo studio (magari grafico) della curva del determinante della matrice \underline{A} funzione di λ , e si vede che si ha una funzione molto “frastagliata”, magari con picchi a pendenza elevatissima:

I picchi possono avere ripidità davvero mostruose, dunque è possibile che uno zero abbia anche valori relativamente elevati; quando si parla di calcoli analitici, uno “zero” deve essere sul serio 0, ma quando si parla di analisi numerica esso va accettato con una certa tolleranza; quando si hanno pendenze così ripide, è possibile che quello che consideriamo “zero” sia per esempio un valore pari a 150, dal momento che la precisione con cui si valuta l'asse delle ascisse si ripercuote sulla precisione dello zero.

Questo discorso vuole semplicemente proporre un fatto: anche quando i problemi agli autovalori contengono operatori differenziali, si ha in qualche modo a che fare con una matrice di cui si deve annullare il determinante; la cosa, tuttavia, è meno immediata, e non è assolutamente detto che il problema in questione sia semplice come uno dei problemi “standard”.

5.2.3 Verifica della natura modale delle espressioni ricavate

Si è detto che le espressioni $u_i(x)$ siano modi della funzione, ma come si può essere sicuri di ciò? Hanno inoltre esse proprietà interessanti?

Consideriamo nel dettaglio un modo, $u_1(x)$; dalle formule precedenti, si può scrivere che l'autofunzione modale normalizzata sia:

$$u_1(x) = \sqrt{\frac{2}{d}} \cos\left(\frac{\pi}{d}x\right)$$

Da ciò, si passi alla forma “onda progressiva più onda regressiva”: questa banalmente può essere ricavata semplicemente applicando la formula di Eulero:

$$u_1(x) = \frac{1}{2} \sqrt{\frac{2}{d}} [e^{j\frac{\pi}{d}x} + e^{-j\frac{\pi}{d}x}]$$

A questo punto ci poniamo la seguente domanda: qual è la relazione tra u e il campo elettrico nella guida? Si può dire che:

$$E_y|_{\text{modo } 1} = V_0^+ u_1(x) e^{-j\beta_1 z}$$

Un commento: V_0^+ è un coefficiente che dice “quanto il modo è eccitato”, ossia il “peso” dell'autofunzione normalizzata; $u_1(x)$ è la già nota autofunzione normalizzata; la dipendenza da z , infine, è quella propria dei modi: un esponenziale complesso, dove

$$\beta_1 = \sqrt{k_0^2 - \left(\frac{\pi}{d}\right)^2}$$

Scriviamo per esteso:

$$\begin{aligned} E_{y,1}(x) &= V_0^+ \frac{1}{2} \sqrt{\frac{2}{d}} [e^{j\frac{\pi}{d}x} + e^{-j\frac{\pi}{d}x}] e^{-j\beta_1 z} = \\ &= \frac{V_0^+}{2} \sqrt{\frac{2}{d}} [e^{-j\mathbf{k}_1^{(-)} \cdot \mathbf{r}} + e^{-j\mathbf{k}_1^{(+)} \cdot \mathbf{r}}] \end{aligned}$$

dove

$$\mathbf{k}_1^{(-)} = -\frac{\pi}{d} \hat{\mathbf{x}} + \beta_1 \hat{\mathbf{z}}$$

$$\mathbf{k}_1^{(+)} = \frac{\pi}{d} \hat{\mathbf{x}} + \beta_1 \hat{\mathbf{z}}$$

(queste espressioni sono equivalenti, come si può vedere sostituendo: si riottengono le espressioni della riga precedente).

Questa espressione sembrerebbe, a prima occhiata, riconducibile alla somma di due onde piane; per averne la certezza, tuttavia, sarebbe buona cosa verificare il modulo di \underline{k} , ossia nella fattispecie determinare che sia valida la relazione di dispersione (ricordando che si suppone di avere il vuoto tra i due specchi):

$$|\underline{k}^{(\pm)}| = k_0$$

Per capire se ciò è vero è sufficiente riprendere le definizioni:

$$\left(\frac{\pi}{d}\right)^2 + \beta_1^2 = \left(\frac{\pi}{d}\right)^2 + k_0^2 - \left(\frac{\pi}{d}\right)^2 = k_0^2$$

Bisogna quindi poi osservare che il campo elettrico sia perpendicolare a \underline{k} : \underline{k} , però, evidentemente è nel piano xz , mentre il campo elettrico giace su y , quindi ok; si dovrebbe poi fare la verifica del campo magnetico.

In conclusione, la cosa bella di tutta questa analisi è il fatto che il modo fondamentale può anche essere descritto come un'onda piana che va “verso l'alto” (componente progressiva, dal momento che si muove nel verso delle x crescenti), e una che va “verso il basso” (duale a prima). Si può immaginare che l'onda che va verso il basso sia una conseguenza di quella che va verso l'alto, per aver sbattuto contro le pareti, o viceversa si può pensare il contrario (dal momento che tutte queste sono soluzioni che si autosostengono, autoconsistente, dunque porsi questa domanda è un po' come porsi la domanda della primogenità tra uovo e gallina).

La questione interessante a questo punto è la seguente: se queste sono onde piane, allora hanno i piani a fase costante che si riflettono e propagano in un certo modo; sono tra loro paralleli questi piani? Fanno parte della stessa famiglia? La risposta deve essere “sì”, altrimenti non si avrebbe una singola onda piana. Il “metodo della congruenza di fase” permette di ricavare l'angolo ϑ con cui si propagano le due onde piane:

facendo considerazioni sulle due famiglie di onde piane e imponendo che esse appartengano alla stessa famiglia, si riesce a fare in modo da trovare l'angolo ϑ di incidenza.

Il fatto che si abbia a che fare con due onde piane dipende sostanzialmente dal fatto che il modo in questione è un TE_1 : si immagini che questo assomigli, in qualche modo, a un TE_{10} (ovviamente il secondo indice qui non ha significato, dal momento che la struttura è trasversalmente illimitata, quindi si perde il “vincolo” sul secondo pedice); nel caso si avesse a che fare con un modo TE_{11} , con tutti gli indici non nulli, le onde piane sareb-

bero 4, e si avrebbe quindi un campo che avanza come una sorta di elica, “intrecciandosi” nelle varie dimensioni.

In questa situazione, l’angolo ϑ non può essere un angolo qualunque: il fatto di avere

$$k_{x,m} = \frac{m\pi}{d}$$

introduce un “effetto di quantizzazione” (ricordando la teoria della meccanica quantistica).

Il metodo che useremo noi è differente: per capire che questa configurazione di campo è un modo, posso considerare il fatto che questo sistema si comporta come una linea di trasmissione chiusa in corto circuito; si può verificare che i valori di campo che sono stati trovati, si possono anche ottenere a partire dal principio fisico delle onde nelle linee di trasmissione: questo sistema è equivalente a una linea di trasmissione in cui si ha un’onda che parte da una delle due facce, si propaga fino all’altra, si riflette completamente, quindi torna indietro, si rifletterà completamente anche alla faccia di partenza, e così via.

Perchè questo campo elettromagnetico sia un modo, si può verificare che, dopo un giro completo:

- la fase del guadagno di anello sia multipla di 2π (altrimenti non si avrebbe una somma in fase);
- il modulo del guadagno di anello sia unitario.

La seconda condizione per noi è una novità (almeno, rispetto a quanto precedentemente analizzato nell’ambito dell’interferometro di Fabry-Perot): nel caso in cui le due condizioni siano contemporaneamente soddisfatte, il sistema è in condizione di **oscillazione**.

Se ci si pensa, un modo di propagazione è proprio un’oscillazione: i modi sono infatti soluzioni delle equazioni di Maxwell in assenza di sorgenti, ma dunque, essendo le soluzioni omogenee (senza sorgenti) delle equazioni di Maxwell delle onde, anche i modi devono essere onde, oscillazioni che si autosostengono, e ciò si può avere solamente se $|T| = 1$, $\angle T = 2n\pi$.

Fatto questo discorso introduttivo, calcoliamo il guadagno di anello per questo sistema:

$$T = e^{-jk_x d}(-1)e^{-jk_x d}(-1)$$

i (-1) sarebbero semplicemente i valori dei coefficienti di riflessione: dalla carta di Smith (o semplicemente dalla formula che permette di definire il

Γ a partire dall'impedenza normalizzata ζ) si può vedere che, con un corto circuito, il coefficiente di riflessione è proprio $\Gamma = -1$; i termini esponenziali complessi rappresentano semplicemente lo sfasamento introdotto dalla guida in un senso e nell'altro (e i percorsi sono ovviamente uguali e lunghi d); si ricorda che questa rappresentazione ha semplicemente lo scopo di determinare in maniera semplice la natura dei modi. Si può dunque banalmente scrivere che:

$$T = e^{-j2k_x d}$$

Una prima osservazione ci può dire che $|T| = 1$: questo è garantito dal fatto che k_x è un numero puramente reale, poichè se così non fosse il modulo dell'esponenziale varierebbe. Una delle due condizioni di oscillazione, quindi, è automaticamente soddisfatta.

Per quanto riguarda la seconda condizione, bisogna imporre che $\angle T = 2m\pi$; si ha dunque che:

$$\angle T = -2k_x d - \pi - \pi$$

infatti, i coefficienti di riflessione dei corto circuiti hanno modulo unitario, e fase pari a π (o $-\pi$ a seconda dei gusti). Eliminando il contributo di 2π (tanto non è influente), e togliendo il segno $-$ (tanto non conta), la condizione diventa:

$$\angle T = 2k_x d = 2m\pi$$

da qui, quindi, si ottiene che:

$$k_{x,m} = m \frac{\pi}{d}$$

Fissato k_x , anche β è fissato. β , come si discuterà anche in seguito, è piuttosto importante, dal momento che è la componente del vettore d'onda lungo z , ossia parallelamente all'interfaccia (esattamente come era ξ quando si studiava la struttura *slab*); si ha:

$$\beta_m = \sqrt{k_0^2 - \left(\frac{m\pi}{d}\right)^2}, \quad m = \pm 1, \pm 2 \dots$$

Fissata la frequenza, esiste solo un numero finito di modi con β_m reale, ossia di "modi sopra taglio": in altre parole, di β_m per cui il radicando è positivo. Esiste, invece, un'infinità di modi sotto taglio, per cui quindi β_m è immaginario. Si può ovviamente definire anche un'impedenza modale: per il modo TE, si ha:

$$Z_{\infty}^{\text{TE}} = \frac{\omega\mu}{k_{z,m}}$$

quindi, se $k_{z,m}$ è immaginario, il modo sotto taglio (come visto precedentemente) non porta energia, essendo evanescente.

Che differenza si ha tra la guida metallica e la struttura *slab*? Sostanzialmente, si ha che ora la dipendenza trasversale del campo è proporzionale a $u(x)$, mentre con lo slab era $e^{-j\xi x}$; la cosa bella, tuttavia, è il fatto che la dipendenza longitudinale è la stessa.

Prima, si aveva:

$$k_z = \sqrt{k_0^2 n_i^2 - \xi^2}$$

Come mai si è appena trovata l'analogia tra ξ e k_z ? La risposta è la seguente: l'obiettivo è sostanzialmente determinare in che modo il campo si propaghi lungo z (che, per l'appunto, è la "direzione longitudinale", o "direzione di propagazione"); al fine di determinare informazioni lungo la direzione longitudinale, tuttavia, è necessario prima conoscere il comportamento del campo nella sezione trasversale a quella di propagazione (come d'altra parte si fa anche nelle guide d'onda: a partire da informazioni sulla sezione trasversale, sui campi nella sezione trasversale, si possono ricavare i parametri della linea modale): si possono quindi ricavare $k_{z,i}$ per i vari mezzi, e $Z_{\infty,i}$, a partire dalle informazioni trasversali. Nel caso dello slab, le informazioni sui campi trasversali erano nascoste in ξ , mentre ora lo sono in $k_{x,m}$, mediante $u(x)$; come prima, dunque, a partire da $k_{x,m}$, è possibile determinare le costanti di propagazione longitudinali (che in questo caso sono semplicemente le β_m), e quindi utilizzare le linee di trasmissione "longitudinali" per determinare il campo che si propaga nella direzione longitudinale. Si hanno dunque **due** linee di trasmissione

- linee di trasmissione "verticali", relative ai campi trasversali: esse servono sostanzialmente per determinare le informazioni dei modi di propagazione;
- linee di trasmissione "orizzontali": una volta trovate le informazioni dei campi trasversali, che permettono per l'appunto di determinare i parametri delle linee che stiamo discutendo in questo punto, è possibile mediante queste determinare il valore del modo (e dunque del campo) per vari valori di z .

Una volta trovati i modi, cioè i valori di $k_{x,m}$, le relative costanti di propagazione e impedenze modali, si devono utilizzare in una linea di trasmissione nella direzione z .

Le linee di trasmissione qui vengono usate semplicemente al fine di risolvere problemi matematici mediante un approccio circuitale: una sorta di promemoria al fine di poter rievocare con semplicità un certo insieme di trasformazioni matematiche. Si può pensare, in realtà, che la teoria delle linee di trasmissione sia un “vestito circuitale” che si costruisce sul metodo matematico di separazione delle variabili: un modo di trattare equazioni differenziali mediante la soluzione di circuiti; questo approccio può essere utilizzato non solo in ambito di sistemi cartesiani, ma anche usando linee di trasmissione polari, o sferiche: le superfici che definiscono una struttura possono essere superfici $r = \text{costante}$, $\vartheta = \text{costante}$, e così via; in ambito di coordinate cartesiane, l’unica possibilità è quella di avere dei parallelepipedi, ma in ambito di coordinate sferiche o cilindriche è possibile definire tubi, cilindri, coni.

5.3 Studio della struttura *asymmetric slab*

Ciò che è stato finora fatto è sostanzialmente lo studio matematico di una struttura molto idealizzata: una guida a “specchi perfetti”. Cerchiamo, a questo punto, di importare tutte le nozioni apprese su un caso più realistico: una vera guida dielettrica (che in realtà è già stata introdotta ed è, comunque, un’idealizzazione delle strutture effettivamente utilizzate).

Questo tipo di struttura di solito ha $n_c = 1$ (ossia, il “cover” è l’aria), n_s è invece un dielettrico con valore intermedio tra l’aria (il cover) e il film, che ha indice di rifrazione n_f .

Come è possibile, in una struttura del genere, avere modi guidati? L’idea è piuttosto semplice: data l’ipotesi appena introdotta, è necessario sfruttare il principio della riflessione totale per intrappolare nel film l’energia; quello che si farà ora, quindi, è usare ancora una volta una linea di trasmissione “trasversale”, come strumento matematico, come “vestito circuitale”, per la determinazione dei modi (ossia di quella che prima era la $k_{x,m}$, a partire dalla quale poi si può studiare la propagazione nella direzione longitudinale del campo). L’ipotesi fondamentale, al fine di avere energia intrappolata all’interno del substrato, è il fatto che le onde in substrato e cover dovranno essere evanescenti.

Analizzeremo solo il caso dei modi TE (ossia per cui l’unica componente di campo elettrico non nulla è E_y); in questo caso, la struttura non è simmetrica, dunque il sistema di riferimento può avere origine posizionata per esempio sull’interfaccia tra cover e film: ora non vi sono più particolari simmetrie da introdurre.

Usando il formalismo modale i cui parametri verranno ricavati mediante la linea di trasmissione “verticale”, “trasversale”, si può immediatamente dire

che:

$$E_y(x, z) = V_0^+ u(x) e^{-j\beta x}$$

Si noti che il dominio in cui questo sistema è definito è $x \in (-\infty, +\infty)$: n_s infatti esiste da $-\infty$, e n_c fino a $+\infty$, quindi questa struttura è una “guida d’onda aperta”; questo sembrerebbe problematico, dal momento che la struttura in questo modo non presenterebbe un dominio chiuso e quindi sembrerebbe infinita, ma non è così: nel cover e nel substrato, infatti, il campo tende asintoticamente a 0 molto rapidamente, quindi come sia effettivamente fatto il substrato non ci interessa: il campo sarà presto nullo, e tutto sarà, almeno in linea di principio, indipendente dal substrato in lontananza dalle giunzioni.

Si tenga ben presente un altro fatto: supponendo che il modo abbia una forma del tipo

questo, quando si propaga lungo z , va immaginato come un blocco che si muove “alla stessa velocità”; quando si parla di “velocità” in questo ambito si parla di “sfasamento”, nel senso che per ogni valore di z ciascun valore del campo per ciascuna x viene sfasato dello stesso angolo. Questo si ha per la natura intrinseca del modo: il modo infatti è una configurazione di campo che, a meno di un certo termine di fase, si preserva identica in ogni sezione trasversale alla direzione di propagazione (z), quindi la β dovrà essere **unica**: non è che si abbia una β per il substrato, una per il film, una per il cover, altrimenti si avrebbe che nelle varie sezioni z il modo subirebbe sfasamenti diversi, e ciò è contro la definizione stessa di modo.

Nel caso della guida realizzata mediante specchi perfetti, sugli specchi il campo era immediatamente nullo, mentre qua non è così: qua si ha un valore non nullo che esponenzialmente decresce fino a zero, essendo i modi evanescenti al di fuori del film. Il fatto che β sia non nullo implica una cosa che a prima vista potrebbe essere sbagliata ma in realtà non lo è assolutamente: l’energia, al di fuori del “tubo”, del film, non è zero, e anzi si propaga nello spazio; in realtà, a causa della presenza di questa “linguetta” (data dalla coda dell’esponenziale), si ha dell’energia che si può propagare, ma solo nella direzione parallela all’interfaccia: nella direzione normale all’interfaccia (dunque lungo \underline{x}) non si ha propagazione di energia attiva; l’energia comunque non è “confinata” tra le interfacce, dal momento che c’è questa sorta di “appendice” che si propaga sempre nella direzione z (e non nella direzione “verticale”).

Il fatto che β sia costante è sostanzialmente analogo al fatto che, nel caso delle onde piane incidenti sullo *slab*, ξ era costante: ciò, sia in questo caso sia in quello appena nominato, dipende dal fatto che, per ipotesi, le

superfici di discontinuità tra i mezzi sono lisce, introducendo un'invarianza per traslazione lungo la direzione x prima, z ora. Facciamo più nel dettaglio un'analisi delle differenze delle due strutture:

- ξ nel caso dello *slab* rappresentava **sia** l'etichetta modale, ossia ciò che rappresenta il modo, la configurazione di campo trasversale, **sia** l'elemento "invariante per traslazione", essendo, in questa configurazione, le interfacce direzionate lungo x (cosa che ha permesso per l'appunto di considerare la rappresentazione spettrale mediante ξ , che in questo caso è sempre uguale, ed è la componente lungo x del vettore d'onda);
- nel caso della struttura guidante (*asymmetric slab*), la direzione per cui si ha invarianza per traslazione in questo caso è z , e la variabile spettrale (in altre parole, la componente lungo la direzione z del vettore d'onda, componente che in questa struttura si mantiene costante in ogni punto) è β ; ciò che rappresenta "l'etichetta modale", a partire dalla quale si ricava la costante di propagazione lungo la linea (che coincide con la componente "invariante") è $k_{x,m}$.

Procediamo a questo punto con i calcoli: al fine di calcolare le $k_{x,i}$ nei vari casi (cover, film, substrato), si usano sostanzialmente le stesse formule di prima:

$$k_{x,s} = \sqrt{k_0^2 n_s^2 - \beta^2} \quad k_{x,f} = \sqrt{k_0^2 n_f^2 - \beta^2} \quad k_{x,c} = \sqrt{k_0^2 n_c^2 - \beta^2}$$

A questo punto, al fine di effettuare alcuni calcoli, si evidenzia k_0 :

$$k_{x,s} = k_0^2 \sqrt{n_s^2 - n_{\text{eff}}^2} \quad k_{x,f} = k_0^2 \sqrt{n_f^2 - n_{\text{eff}}^2} \quad k_{x,c} = k_0^2 \sqrt{n_c^2 - n_{\text{eff}}^2}$$

dove

$$n_{\text{eff}} \triangleq \frac{\beta}{k_0}$$

questo n_{eff} è una sorta di indice di rifrazione equivalente (esso è consistente con il termine a cui si sottrae); questo è esattamente quello che si aveva, in ambito di microstrisce, quando si parlava di $\sqrt{\varepsilon_{\text{eff}}}$.

Detto tutto ciò, si è detto che si vuole fare in modo che, fuori dal film, i modi siano evanescenti; questo si traduce, in altre parole, con il fatto che $k_{x,s}$ e $k_{x,c}$ devono essere immaginari, mentre $k_{x,f}$ deve essere reale; questo, quindi, si andrà a riflettere sul valore dei radicandi. Si consideri n_{eff} variabile, e i vari $n_{c,f,s}$ fissati dal problema (i substrati sono fissati a priori); si può dire

che l'unico range di valori per cui la condizione precedentemente espressa (costante di propagazione lungo x intesa come "indice modale" reale per il film, immaginaria per gli altri) sia valida è il seguente:

ossia

$$n_s < n_{\text{eff}} < n_f$$

Spesso sul disegno si usa introdurre delle barre verticali", per indicare l'insieme dei valori accettabili per gli n_i ; questo ha un'analogia interessante con la meccanica quantistica, applicato al profilo dell'indice di rifrazione: si indica un certo valore di n_i , e si vede, ruotando di 90° il disegno, un disegno sostanzialmente identico a quello delle buche di potenziale, dove la riga (ora orizzontale) indica, in ambito quantistico, un livello di energia; in ambito quantistico, l'energia di uno stato legato deve essere al di sotto dell'energia potenziale, delle "pianure" alte, ma devono essere al di sopra del minimo. Nel caso le righe finissero "a sinistra" del valore di n_s , nel substrato l'onda non sarebbe più evanescente, quindi non avrebbe senso l'ipotesi iniziale: non si rispetterebbero più le condizioni sui $k_{x,i}$.

Abbiamo a questo punto definito l'intervallo dei valori che possono essere assunti da n_{eff} , a condizione di avere i fenomeni di riflessione totale; introduciamo, fatta questa ipotesi, il concetto di guadagno di anello su questa guida, calcolandolo:

$$T = e^{-jk_{x,f}d} \Gamma_A e^{-jk_{x,f}d} \Gamma_B$$

dove $\Gamma_{A,B}$ sono i coefficienti di riflessione "visti da dentro": l'onda infatti gira sempre "all'interno della cavità", per questo motivo si ha ciò. Il fatto che i moduli dei coefficienti di riflessione siano unitari è una conseguenza automatica della presenza della riflessione totale, quindi del fatto che

$$n_f > n_{\text{eff}} > n_c$$

del fatto che le onde dunque nel cover e nel substrato sono evanescenti.

Se infatti $n_{\text{eff}} > n_f$, si avrebbe $k_{x,f}$ immaginario, quindi gli esponenziali avrebbero tutti modulo diverso da 1. Se invece $n_{\text{eff}} < n_s$, si ha che Γ_B ha modulo unitario, ma Γ_A ha modulo meno che unitario: in questo modo, discutendo il guadagno di anello, abbiamo verificato la validità del già introdotto range di variabilità di n_{eff} .

Questa, era una condizione; per avere oscillazioni, tuttavia, non è sufficiente ciò che è stato detto: $\angle T$ deve infatti essere un multiplo di 2π ; quindi:

$$\angle T = 2m\pi = -2k_{x,f}d + \angle \Gamma_A + \angle \Gamma_B$$

Al fine di procedere analiticamente, è necessario trovare le espressioni delle fasi dei due coefficienti di riflessione; queste in realtà sono semplici da ricavare, dal momento che sono già note dalla teoria dei coefficienti di Fresnel, precedentemente introdotta:

$$\angle\Gamma_B = 2\arctan\left(\frac{\sqrt{\beta^2 - k_0^2 n_c^2}}{\sqrt{k_0^2 n_f^2 - \beta^2}}\right)$$

(dove, si noti, è stato scambiato l'ordine al numeratore, dal momento che si ha l'imposizione dell'onda immaginaria sul cover, quindi si vuole ottenere qualcosa nella forma $-j|k_{x,c}|$). Ricordando la definizione di indice di rifrazione efficace, si ottiene:

$$\angle\Gamma_B = 2\arctan\left(\frac{\sqrt{n_{\text{eff}}^2 - n_c^2}}{\sqrt{n_f^2 - n_{\text{eff}}^2}}\right)$$

Per $\angle\Gamma_A$ si ha poi qualcosa di molto simile:

$$\angle\Gamma_A = 2\arctan\left(\frac{\sqrt{n_{\text{eff}}^2 - n_s^2}}{\sqrt{n_f^2 - n_{\text{eff}}^2}}\right)$$

A questo punto, se introduco le espressioni delle fasi in questo modo, ricordando che

$$k_0 = \frac{2\pi}{\lambda_0}$$

ottengo la seguente equazione:

$$\frac{2\pi d}{\lambda_0} \sqrt{n_f^2 - n_{\text{eff}}^2} = \arctan\left(\frac{\sqrt{n_{\text{eff}}^2 - n_c^2}}{\sqrt{n_f^2 - n_{\text{eff}}^2}}\right) + \arctan\left(\frac{\sqrt{n_{\text{eff}}^2 - n_s^2}}{\sqrt{n_f^2 - n_{\text{eff}}^2}}\right) + \nu\pi$$

dove ν è una versione “modificata” di m , tenendo conto che si è diviso per 2, e di eventuali termini additivi o sottrattivi. Si ha:

$$\nu = 0, 1, 2, \dots$$

Questa equazione rappresenta la relazione di dispersione per la struttura! In essa, il modo fondamentale della struttura, si ha per $\nu = 0$.

Precedentemente, nel caso della guida a specchi perfetti, era stata scritta un'equazione del tipo:

$$k_x d = m\pi$$

anche questa, in realtà, può essere scritta come un'equazione di dispersione, sviluppando k_x :

$$\sqrt{k_0^2 - \beta^2 d} = m\pi \implies dk_0 \sqrt{1 - n_{\text{eff}}^2} = m\pi$$

quindi

$$\frac{2\pi d}{\lambda_0} \sqrt{1 - n_{\text{eff}}^2} = m\pi, \quad m = 1, 2, 3\dots$$

questa è la relazione di dispersione per specchi perfetti mentre quella, più complicata, scritta sopra, quella di uno *slab* asimmetrico.

Si ha a che fare con un'equazione in due variabili: λ_0 e n_{eff} ; le due possibilità, quindi, sono o risolverla in n_{eff} , o in λ_0 ; avrebbe ovviamente più senso ottenere $n_{\text{eff}}(\lambda_0)$, dal momento che, quando si lavora per esempio con dei LASER, quello che si fa di solito è fissare la frequenza di lavoro, quindi determinare di conseguenza l'indice di rifrazione efficace. Nel caso degli specchi perfetti, risolvere entrambi i casi è piuttosto facile, mentre, nel caso "vero", avere n_{eff} è impossibile, se non in casi semplificati, senza ricorrere a metodi di risoluzione numerici.

Proponiamo, a questo punto, alcune considerazioni sui vari membri dell'equazione: definiamo

$$B \triangleq \arctan \left(\frac{\sqrt{n_{\text{eff}}^2 - n_{\text{c}}^2}}{\sqrt{n_{\text{f}}^2 - n_{\text{eff}}^2}} \right)$$

$$A \triangleq \arctan \left(\frac{\sqrt{n_{\text{eff}}^2 - n_{\text{s}}^2}}{\sqrt{n_{\text{f}}^2 - n_{\text{eff}}^2}} \right)$$

A questo punto, alcune casistiche:

- se n_{eff} tende a n_{s} , il numeratore di A tende a 0, quindi $A \rightarrow 0$; B invece tende a un certo valore, non particolarmente importante, ma quantificabile mediante la sostituzione $n_{\text{eff}} \rightarrow n_{\text{s}}$:

$$B \rightarrow \sqrt{\frac{n_{\text{s}}^2 - n_{\text{c}}^2}{n_{\text{f}}^2 - n_{\text{s}}^2}}$$

nel caso di guida simmetrica, poi, questo termine tende a 0;

- nel caso in cui $n_{\text{eff}} \rightarrow n_{\text{f}}$, si ha il **denominatore** che si annulla, quindi la radice quadrata tende comunque a $+\infty$, e quindi:

$$A \rightarrow \frac{\pi}{2}, \quad B \rightarrow \frac{\pi}{2}$$

Volendo dunque risolvere il “metodo semplice”, ossia volendo ottenere, dato n_{eff} come variabile indipendente, $k_0 d|_{\nu=0}$, si ha qualcosa di questo tipo:

questo, supposto di considerare il modo fondamentale: se $n_{\text{eff}} \rightarrow n_f$, si ha che il secondo membro tende a π ; d'altra parte, se $n_{\text{eff}} \rightarrow n_s$, a primo membro si ha ancora la radice, a meno che la struttura non sia simmetrica (in tal caso, il punto va a 0). Volendo disegnare la stessa figura, ribaltando gli assi (cosa che si può fare banalmente con MATLAB), si vedrà questo:

Questa operazione esteticamente è assolutamente perfetta, ma concettualmente è piuttosto scadente: di fatto le operazioni fatte son sempre le stesse, solo ribaltando gli assi. Questo discorso inoltre completa il caso $\nu = 0$, ma non gli altri: se $\nu \neq 0$, si ha che il secondo membro, rispetto a quello usato, è incrementato di $\pi, 2\pi, 3\pi$, e così via; la conclusione, dunque, è il fatto che l'insieme delle curve di dispersione ha sostanzialmente questo tipo di andamento:

Su quest'ultima figura, è possibile fare delle considerazioni: fissato n_{eff} , quali sono le frequenze a cui i vari modi di propagazione hanno questo indice di rifrazione, e quindi quella velocità di fase, coefficiente di propagazione? I modi sono infiniti: le curve derivano da ν che va da 0 a ∞ , quindi a frequenze via via più grandi ci sarà sempre un punto che incrocerà il valore di n_{eff} ; questo, sostanzialmente, è ciò che si ottiene dal “metodo semplice” prima fatto; la domanda interessante è: dato un LASER che lavora a una certa λ_0 , quanti sono i modi attivi, alla frequenza a essa corrispondente? Rispondere a questa domanda sostanzialmente significa, data l'ultima figura, tracciare una retta verticale, e vedere quante intersezioni abbia, ossia quante curve essa intersechi; volendo avere una guida monomodale, la frequenza deve essere tale da avere meno che la frequenza per cui scatta il secondo modo.

Un grafico che spesso si ritrova in letteratura² è il seguente:

di solito i grafici vengono presentati in forma normalizzata, dunque si hanno dei parametri di normalizzazione; in questo, si hanno V, a, b :

$$V = k_0 d \sqrt{n_f^2 - n_s^2}$$

non si considera il cover, dal momento che non è fondamentale: quello che conta è infatti il substrato, nel senso che quando un'onda smette di essere evanescente, di essere confinata, il problema risiede non nel cover, ma nel substrato (dal momento che $n_s > n_c$, per ipotesi). Una grandezza normalizzata per n_{eff} invece è:

²Kogelnik H. - pag. 66 - Fig. 2.8

$$b = \frac{n_{\text{eff}}^2 - n_s^2}{n_f^2 - n_s^2}$$

al minimo, $b = 0$; al massimo, $b = 1$.

Vengono poi rappresentate non curve singole, ma gruppi di curve: per $\nu = 0$, per esempio, si può vedere che si ha a che fare con tre curve. Il parametro a , detto “parametro di asimmetria”, è quello che permette di differenziare queste curve; per la polarizzazione TE,

$$a^{\text{TE}} = \frac{n_s^2 - n_c^2}{n_f^2 - n_s^2}$$

questo parametro, come il nome suggerisce, quantifica l’asimmetria della struttura: se la struttura è simmetrica, $a = 0$. Nel caso $a = 0$, la curva passa per l’origine: questo è l’unico caso per cui a frequenze pressochè nulle si ha comunque un modo di propagazione; in altre parole, per frequenze asintoticamente tendenti a 0, si ha un modo per $\nu = 0$; b sarebbe nullo, quindi si avrebbe $n_{\text{eff}} = n_s$. Questo significa che, almeno in linea di principio, una guida dielettrica per l’ottica potrebbe anche trasmettere onde non nel campo dell’ottico, per esempio a 1 MHz (questa cosa in realtà va presa con le pinze dal momento che non è proprio così). Nel caso delle onde TM, si ha qualcosa di un po’ diverso:

$$a^{\text{TM}} = a^{\text{TE}} \left(\frac{n_f}{n_c} \right)^4$$

il grafico dunque è universale: moltiplicando per questo termine correttivo, è possibile utilizzare questo grafico per entrambe le polarizzazioni.

Si può dimostrare che queste curve partano a tangente orizzontale; il comportamento è, tuttavia, veramente poco evidente, quindi non ci si aspetti di osservarlo bene.

Note finali

Quali sono, di queste, le curve valide per la modellazione della guida a specchi perfetti paralleli? Beh, si ha:

$$\beta = \sqrt{k_0^2 - \left(\frac{m\pi}{d} \right)^2} = k_0 \sqrt{1 - \left(\frac{m\pi}{k_0 d} \right)^2}$$

da qui:

$$n_{\text{eff}} = \frac{\beta}{k_0} = \sqrt{1 - \left(\frac{m\pi}{k_0 d}\right)^2}$$

Volendo farne un grafico, si avrebbe ciò:
per $f \rightarrow \infty$, sapendo quanto vale V :

$$V = k_0 d$$

(essendo la guida a specchi perfetti paralleli vuota), se $k_0 d \rightarrow \infty$, ho

$$n_{\text{eff}} = \sqrt{1} = 1$$

quindi, il valore limite di n_{eff} è 1, per tutti i modi. A Campi Elettromagnetici spesso si presentano dei grafici del tipo:

$$\frac{k_z}{k_0} = \frac{Z^{\text{TM}}}{Z_0} = \frac{v_g}{c}$$

queste quantità sono semplicemente gli indici di rifrazione efficaci, ossia le curve appena proposte! Questo significa che, per tutte le guide d'onda, si hanno andamenti di questo tipo: l'andamento generale è sempre questo.

Note per la determinazione mediante MATLAB delle funzioni modali

Come detto precedentemente, al fine di determinare le funzioni modali si utilizza il modello delle “linee di trasmissione trasversali”, ossia quello per cui si hanno le linee di trasmissione dei vari strati, e mediante esse si determinano i vari parametri della linea lungo z , quella con la quale si può studiare la propagazione dello stato elettrico. La cosa interessante, a questo punto, è il fatto che è possibile studiare le onde progressive e le onde regressive nella “linea trasversale”, dal momento che esse hanno un significato ben preciso: quello delle “funzioni modali”; in altre parole, il diagramma d'onda stazionaria all'interno della struttura è coincidente con le funzioni modali che si stanno descrivendo; volendo dunque disegnare gli andamenti dei campi elettrici o magnetici trasversali, relativi a un particolare modo, è sufficiente disegnare il diagramma d'onda stazionaria della linea trasversale.

Tutto ciò si può fare, supponendo di conoscere i vari valori di $k_{x,f,s,c}$, fissato n_{eff} e ν .

Da cosa si capisce quale sia il modo? Tutte le formule di teoria delle linee che si usano, sono sempre le stesse: l'unica cosa che cambia è ν ; a seconda del valore di frequenza, si avranno uno o più modi attivi: uno, due, tre o n modi,

dal momento che “mettere una frequenza più grande” significa semplicemente che, quando la frequenza è più grande, k_x è più grande, dunque a parità di movimento fisico sulla linea ho un maggiore sfasamento, un maggiore numero di giri sulla carta di Smith!

Un ultimo dettaglio: la funzione $V(x)$ del diagramma d’onda stazionario (che si può ricavare mediante la teoria delle linee di trasmissione), è reale o complessa? Beh, ci sono esponenziali complessi, dunque la prima risposta che si potrebbe dare di sicuro è “complessa”. La questione tuttavia non è per niente triviale: se si studia sotto un punto di vista matematico questo problema, essendo il sistema senza perdite, si ha a che fare con un operatore differenziale autoaggiunto, ma quindi un teorema afferma che gli autovettori, di conseguenza, devono essere reali. Gli autovettori, semplicemente, sono le funzioni modali, dunque esse devono essere reali. Il teorema non è sbagliato: se si studia $V(x)$, essa è una funzione fatta di numeri, con una parte reale e una parte immaginaria; se si studia tuttavia la fase, si vede che essa è costante a tratti: si ha sostanzialmente che:

$$V(x) = \alpha u(x)$$

dove $u(x)$ è una funzione reale che può essere positiva o negativa, quindi che procura un salto pari a π nella fase, e α è un generico numero complesso. La fase del numero complesso è fissa, mentre quella di $u(x)$ varia “a scatti”. Il fatto che la fase sia costante ma non nulla, deriva dal fatto che di solito i grafici si fanno normalizzati rispetto a un qualche parametro, per esempio V_{A+}^+ , che potrebbe essere posto uguale a 1, per esempio. Ciò che si può fare, al fine di rendere reale la funzione, è prendere la funzione, valutarla in un certo punto, e dividere ogni punto della tensione per questo valore: in questo modo, si compensa in ogni punto il fattore di fase, dal momento che esso è costante.

5.3.1 Note aggiuntive

A questo punto, al fine di comprendere meglio ciò di cui si sta parlando, verranno introdotte alcune note aggiuntive sulle guide d’onda dielettriche, rispetto a quanto già fatto.

Avendo a disposizione un software in grado di determinare i modi che si possono propagare in una struttura guidante dielettrica per una λ_0 (dunque per una frequenza) fissata, al variare dello spessore d del film e del profilo degli n_i , si potrebbe notare che diminuendo d man mano diminuisce il numero di modi guidati: ciò si potrebbe vedere con il disegno presentante l’analogia con la meccanica quantistica, identificando ciascun modo con un livello ener-

getico consentito nella “buca di potenziale”: questo è esattamente quello che capita qua, parlando di modi di propagazione invece che di livelli di energia. Volendo avere un’analogia con le strutture guidanti metalliche (per quanto, come vedremo tra breve, essa sarebbe inappropriata in senso stretto), si può considerare questa situazione:

Questa è quella situazione che, in ambito di guide d’onda metalliche, viene detta “frequenza pari alla frequenza di taglio del primo modo superiore”: questo significa che il modo fondamentale è ben attivo, mentre il secondo non lo è ancora, per quanto vi sia; come si dirà meglio tra breve e dimostrerà in seguito, si vedrà che parlare di “taglio” è assolutamente inappropriato, nell’ambito di guide d’onda dielettriche.

Considerando un singolo modo di propagazione (il fondamentale), esso avrà più o meno la forma di una gaussiana, ma asimmetrica: nel cover infatti il campo elettromagnetico tende a zero tanto più rapidamente rispetto a quanto avviene nel substrato tanto maggiore è l’asimmetria tra cover e substrato (si ricordi che si suppone sempre che $n_c < n_s$). Si può parlare, come si fa spesso per esempio parlando di effetto pelle, di una “costante di spazio”, che permette di avere un’idea di quale sia la rapidità con cui il campo si annulla al crescere di x ; si noti tuttavia che questo tipo di fenomenologia non è assolutamente legata all’effetto pelle, in quanto l’effetto pelle è originato dalla presenza di una dissipazione, dovuta a una conducibilità parassita presente in un dielettrico; qui si ha a che fare con attenuazioni del campo, originate tuttavia non da dissipazioni.

Si può osservare, utilizzando questo software, ancora una cosa: avendo a che fare con una guida simmetrica, dunque con $n_c = n_s$, anche nel caso di d molto piccolo (teoricamente nullo), si avrebbe comunque a che fare con un modo che si propaga; questo significa, in altre parole, che è possibile avere a che fare con un modo a frequenza molto bassa, teoricamente TEM, avendo una guida anche grossa (ma simmetrica). La domanda a questo punto è: è possibile trasmettere su di una guida d’onda dielettrica, dunque nata per gestire frequenze ottiche, un segnale a 100 MHz?

La risposta è la seguente:

Sì, sicuramente sì, con però una grossa controindicazione: quando si parla di guide d’onda, si parla di strutture che mantengono il campo elettromagnetico confinato al loro interno, avendo un’onda evanescente al di fuori di esse; in questo caso l’onda evanescente in effetti c’è, ma è estremamente alta anche al di fuori della guida: l’onda dunque trasporta potenza parallelamente all’interfaccia, ma anche fuori dalla guida, a una distanza non trascurabile; questo significa, in altre parole, che si ha molta più potenza fuori dalla guida che non dentro di essa.

Il rapporto tra la potenza all'interno del film e quella al di fuori viene calcolata mediante un coefficiente C_c , detto “coefficiente di confinamento”, così quantificabile:

$$C_c = \frac{\operatorname{Re} \left\{ \int_0^d E_y H_x^* dx \right\}}{\operatorname{Re} \left\{ \int_{-\infty}^{+\infty} E_y H_x^* dx \right\}}$$

Il modo formalmente è guidato, ma praticamente è inutilizzabile, dal momento che si ha a che fare con molta energia transitante al di fuori della struttura dielettrica; una guida d'onda (metallica o meno) deve essere pensata come a una specie di tubo dell'acqua: l'acqua deve stare dentro al tubo, non fuori da esso!

5.4 Numero di modi presenti in una guida dielettrica

Date le strutture (metalliche e dielettriche) finora studiate, vorremmo a questo punto spiegare meglio un concetto introdotto ma non motivato: il fatto che in una guida dielettrica si ha a che fare con un numero finito di modi (guidati) attivi.

In una guida metallica, ciò che si ha è, a seconda della frequenza alla quale si eccita la struttura, un certo numero di modi sopra taglio, e un certo numero di modi sotto taglio; alla frequenza critica f_c , infatti, si ha $k_z = 0$; al di sotto di questa, k_z è puramente immaginario, mentre al di sopra di questa k_z è puramente reale. In questo caso, quando il modo ha $n_{\text{eff}} = n_s$, il β , che nell'ambito della propagazione ha lo stesso ruolo di k_z , non è nullo, bensì è:

$$\beta = k_0 n_s$$

Se il valore della frequenza scende al di sotto del valore della “frequenza critica”, non è che il modo esista e abbia una costante di propagazione immaginaria, ma **non esiste proprio**: non esiste il concetto di modo sotto taglio, nel senso che i modi scompaiono, non esistono proprio, e non sono semplicemente attenuati.

Si noti che la spaziatura in termini di indice di rifrazione non è costante: tanto più la guida è grande, tanto più numerosi sono i modi guidati, e la spaziatura non è uniforme: in n_{eff} , quanto più grande è l'indice efficace, tanto minore è la distanza tra gli n_{eff} di due modi. Se si aumenta la profondità, il salto tra esterno e interno della guida, il numero dei modi aumenta; idem, se λ si riducesse (cosa che coincide con il far crescere d). Il parametro V

precedentemente definito è un buon “parametro globale”, dal momento che si può vedere che questi tre parametri sono tutti contenuti in esso, e quindi si vede che tutti contribuiscono a far crescere V .

Si immagini a questo punto di avere a che fare con una lastra di cristallo, per esempio di un tavolino, spesso un centimetro, per luce visibile: un centimetro contro $1 \mu\text{m}$ dà luogo a 10000 lunghezze d’onda, quindi, supponendo (ci sono formule più precise) che:

$$N_{\text{modi}} \sim \frac{d}{\lambda}$$

si ha a che fare con 10000 modi circa, ossia 10000 possibili valori di n_{eff} , spaziati tra 1 e 1,5 (cover e film, dove il film è la silice, la lastra di vetro); in una situazione del genere, l’intervallo tra due n_{eff} è quasi nullo, dunque in una struttura del genere si potrebbe quasi dire che “qualsiasi numero” n_{eff} rappresenti un modo. In termini fisici, quello che sostanzialmente capita in una situazione del genere è la sparizione del concetto di quantizzazione (come l’elettrone nella scatola: una particella classica in una buca di potenziale può avere qualsiasi valore energetico, non solo valori discreti; solo quando l’energia è molto piccola rispetto all’altezza delle pareti della buca di potenziale, si ha a che fare con gli effetti di quantizzazione). Solitamente, le strutture guidanti si usano in monomodalità: la condizione normale è quella per cui il campo decade esponenzialmente, al di fuori della guida, in maniera rapida.

Una volta introdotto il problema, introduciamo un discorso più formale. Questo può essere sostanzialmente basato su un metodo più tradizionale, ossia il metodo che è stato usato per ricavare i modi sulla guida a specchi perfetti paralleli: in quell’ambito infatti sono state usate le “linee di trasmissione verticali”, che sostanzialmente sostituivano il formalismo di Marcuvitz e Schwinger, introducendo un’equazione d’onda risolubile circuitalmente; nell’ambito della guida dielettrica invece era stato usato un approccio più vicino all’automatica, definendo un modo come una configurazione di campo per cui $T = 1$.

Consideriamo a questo punto la seguente linea di trasmissione:

Nel caso della guida a specchi perfetti paralleli, il problema era stato ricondotto (essendo coincidente) a quello di una guida a piani metallici paralleli, dunque a una “linea verticale” caricata su dei corto circuiti; il problema dunque, in termini di equazioni differenziali, era il seguente:

$$\begin{cases} \frac{d^2V}{dx^2} + k_x^2 V = 0 \\ V(0) = V(d) = 0 \end{cases}$$

in questo caso, risolvendo, si era trovato:

$$k_x = \frac{m\pi}{d}$$

Qui, invece, si ha sempre a che fare con delle equazioni delle linee di trasmissione (differenziando i k_x , dal momento che se ne hanno tre diversi, a seconda del mezzo in cui ci si trova, differenziando dunque cover, film e substrato), quindi con equazioni del tipo:

$$\frac{d^2V}{dx^2} + k_{x,i}^2 V = 0$$

ma con condizioni al contorno diverse. Quali sono, ora, queste condizioni al contorno? Beh, si ha a che fare con due giunzioni, ma esse devono essere solo delle discontinuità di impedenza: V e I (tensione e corrente totali) devono essere funzioni continue; si deve inoltre avere, al fine di avere un sistema fisicamente sensato, che le tensioni e le correnti siano limitate, ossia che non divergano per $x \rightarrow \pm\infty$; il sistema da risolvere quindi è:

$$\begin{cases} \frac{d^2V}{dx^2} + k_{x,i}^2 V = 0 \\ V, I \text{ continue in } x = \{0, d\} \\ V, I \text{ limitate per } x \rightarrow \pm\infty \end{cases}$$

Questi due sistemi di equazioni sembrano molto simili tra loro, ma in realtà uno ha un numero infinito di soluzioni, l'altro un numero finito di soluzioni; come mai? Per rispondere a questa domanda, è necessario procedere con i calcoli.

Prima di tutto, come fatto precedentemente, è necessario scrivere l'espressione dell'integrale generale, quindi applicare su di essa le condizioni al contorno. Si avrà:

- per $x < 0$,

$$V(x) = V_{A^-}^+ e^{-jk_{x,c}x} + V_{A^-}^- e^{+jk_{x,c}x}$$

$$I(x) = \frac{V_{A^-}^+}{Z_{\infty c}} e^{-jk_{x,c}x} - \frac{V_{A^-}^-}{Z_{\infty c}} e^{+jk_{x,c}x}$$

- per $0 < x < d$:

$$V(x) = V_{A^+}^+ e^{-jk_{x,f}x} + V_{A^+}^- e^{+jk_{x,f}x}$$

$$I(x) = \frac{V_{A^+}^+}{Z_{\infty f}} e^{-jk_{x,f}x} - \frac{V_{A^+}^-}{Z_{\infty f}} e^{+jk_{x,f}x}$$

- per $x > d$:

$$V(x) = V_{B^+}^+ e^{-jk_{x,s}x} + V_{B^+}^- e^{+jk_{x,s}x}$$

$$I(x) = \frac{V_{B^+}^+}{Z_{\infty s}} e^{-jk_{x,s}x} - \frac{V_{B^+}^-}{Z_{\infty s}} e^{+jk_{x,s}x}$$

Questo è, per i vari range di x , l'integrale generale dell'espressione; le costanti arbitrarie qui sono i valori iniziali di tensioni progressive e regressive, dunque si ha a che fare con sei "costanti arbitrarie" (si parla di "arbitrarie", rispetto alla soluzione generale: una volta imposte le condizioni al contorno, esse non sono assolutamente più arbitrarie!).

A questo punto, vogliamo risolvere il sistema, in due particolari situazioni: quella dei "modi guidati", e quella dei modi "non guidati".

5.4.1 Modi guidati

Quando si parla di modi "guidati", si considera, per ipotesi, di avere $k_{x,s}$, $k_{x,c}$ immaginari, $k_{x,f}$ reale: in questo modo la potenza è confinata nel film, mentre al di fuori di esso le onde sono evanescenti. In queste situazioni, quindi:

- $k_{x,s} = -j|k_{x,s}|$;
- $k_{x,c} = -j|k_{x,c}|$.

Sostituendo queste espressioni in quelle dell'integrale generale, si trova che $V_{A^-}^+ = V_{B^+}^- = 0$: gli esponenziali moltiplicati per queste due costanti, infatti, sostituendo le espressioni, esploderebbero, rispettivamente al decrescere e al crescere di x . Questa osservazione, basata sull'imposizione della limitatezza dei campi nello spazio, ci ha permesso di determinare due delle sei costanti arbitrarie: ora, di sei costanti, ne abbiamo solo più quattro da determinare. La condizione al contorno che rimane, a questo punto, è quella della continuità:

$$\begin{cases} V(A^-) = V(A^+) \\ I(A^-) = I(A^+) \\ V(B^-) = V(B^+) \\ I(B^-) = I(B^+) \end{cases}$$

Nel caso della guida a specchi piani paralleli, si aveva a che fare con solo due equazioni in due incognite; ora, invece, si hanno quattro equazioni in quattro incognite, ottenendo un sistema che può essere scritto mediante

una matrice 4×4 ; la matrice $\underline{\underline{A}}$, che non verrà esplicitata, conterrà degli esponenziali.

Al fine di avere soluzione non banale, è necessario imporre $\det \{\underline{\underline{A}}\} = 0$. Senza voler ricavare la soluzione con questo procedimento, in questo caso si ha che $\underline{\underline{A}}$ è funzione sia della frequenza, sia di n_{eff} ; imporre il determinante nullo significa soddisfare una certa $f(\omega, n_{\text{eff}})$, ossia una certa equazione di dispersione: quella equazione che è stata ricavata per vie traverse, giocando sul guadagno di anello. Questa equazione, come noto dall'analisi precedentemente fatta su di essa, ha un numero finito di soluzioni: i modi guidati.

Tutta la questione del numero finito di modi sta nella parola “guidati”: se i modi sono infatti guidati, si ha a che fare con costanti di propagazione immaginarie, cosa che porta ad ottenere l'equazione di dispersione precedentemente scritta, la quale ha solo soluzioni in numero finito.

5.4.2 Modi non guidati

È necessario che le costanti di propagazione siano immaginarie? In realtà no: è necessario solo se si vuole avere a che fare con soluzioni guidate, ossia con soluzioni, configurazioni di campo, la cui energia sia confinata nel film. Se avessimo n_{eff} minore di $n_{s,c}$, avremo a che fare con $k_{x,i}$ reali.

Data questa ipotesi, perdiamo la necessità della condizione al contorno di limitatezza: a queste condizioni, infatti, il campo è sempre limitato, dal momento che tutti gli esponenziali sono complessi, quindi si ha a che fare con funzioni dal modulo sempre unitario. Ciò porta a non avere più solo quattro costanti da determinare, bensì sei, per quanto le equazioni derivanti dall'imposizione della continuità continuino a rimanere quattro. In altre parole, in queste situazioni, il sistema è **sottodeterminato**: sempre omogeneo (dal momento che senza sorgenti), ma sottodeterminato; questo significa che, al fine di avere soluzioni non banali, non è necessario imporre un ipotetico determinante (la matrice è rettangolare, dunque non avrebbe senso farlo) a zero.

Come si può procedere, in questo caso? Quello che si può fare è prendere due delle costanti arbitrarie, portarle “a secondo membro”, quindi ottenendo un vettore con le due costanti arbitrarie, trattato come se fosse un termine noto; in questo caso, quindi, fissando del tutto arbitrariamente due costanti, le altre quattro sono la soluzione; il sistema in questo modo non è più omogeneo, ma quindi la necessità di imporre il determinante (in questo caso esistente) a zero non è presente; questo significa che, dal momento che non si definisce nessuna funzione dei parametri ω, n_{eff} , non si ha nessuna relazione di dispersione!

Quanto vale, in questa situazione, n_{eff} ? Qualsiasi valore reale (e anche immaginario)! Infatti:

$$\beta = k_0 n_{\text{eff}}$$

quindi, n_{eff} varia con continuità. Questi modi appartengono allo spettro continuo dell'operatore: ora lo spettro non è solo discreto, ma può variare con la potenza del continuo, avendo dunque infiniti valori di n_{eff} , quindi infiniti modi possibili.

Questo permette di chiudere il discorso precedente: una guida dielettrica ha un numero finito di **modi guidati**, mentre ha infiniti modi non guidati, che esistono sempre, per ogni valore di frequenza. Qual è la forma di questi modi? Semplice: si immagini di considerare gli studi fatti sugli *slab* (in ambito di studio dell'ottica dei mezzi stratificati): quando si visualizza il diagramma di onda stazionaria, supponendo che la linea di trasmissione in questione abbia il significato di “linea di trasmissione trasversale”, ossia quello di “linea per la determinazione delle configurazioni di campo dei modi”, quello che si vede è proprio un modo dello spettro continuo: campi non decadenti in nessuna delle zone!

Questa configurazione di campo rispetta tutte le proprietà che un modo deve avere: è limitata, “periodica” in x , costante in z a meno del fattore di fase $e^{-j\beta z}$.

Quella visualizzabile con l'esempio dello *slab* in realtà mostra alcuni dei modi non guidati, ma non tutti: nel dettaglio, i modi considerati sono quelli per cui ξ è maggiore di zero; queste sono solo parte delle soluzioni: volendo usare lo slab per capire il concetto, le soluzioni che servono sono quelle con onde incidenti “da sinistra”, “da destra”, e con valori di ξ (ossia di ϑ_i) positivi e negativi: quelle presentate nell'esempio qualitativo, dunque sono sostanzialmente “un quarto” di tutte le soluzioni.

5.4.3 Note conclusive sulle guide d'onda dielettriche

Si vogliono a questo punto proporre alcune note conclusive sugli argomenti presentati.

Guida a specchi perfetti paralleli

Per questa struttura è presente una formula molto semplice, in grado di fornire il numero dei modi guidati presenti (sopra taglio) a una certa frequenza, o meglio per una certa λ ; si ha, infatti:

$$M = \frac{2d}{\lambda}$$

Un discorso importante da fare è quello della **velocità di gruppo**: v_g . Questa è importantissima, dal momento che quello che si introduce nella guida non è un segnale analogico monocromatico, bensì un segnale digitale; esso può essere pensato come una certa portante, con una sagoma gaussiana.

Un impulso nel modo fondamentale viaggia con la velocità di gruppo: la velocità di fase è infatti la velocità della portante, ma quello che contiene l'informazione non è la portante, che viene scartata dal ricevitore, bensì l'involuppo! Esso si propaga secondo la velocità di gruppo, ma dunque è necessario che essa sia unica", al fine che l'informazione non venga modificata; per questo motivo, è necessario avere a che fare con guide monomodali: se si ha a che fare con guide multimodali infatti si ha una diversa velocità di gruppo per ciascuno dei modi, solo che i modi non sono disaccoppiabili da un ricevitore: essi sono configurazioni di campo che vengono "mischiate", senza la possibilità di poterle disaccoppiare (a causa dell'accoppiamento modale, per cui parte della potenza di un modo viene trasferita sugli altri), quindi sarebbe buona cosa avere solo il modo fondamentale, nella propria struttura. Finché la rete in fibra ottica deve essere usata per cablare reti piccole, come l'interno di un edificio, la multimodalità è accettabile, dal momento che la differenza di v_g non si fa sentire più di tanto, nel senso che non riesce a disallineare più di tanto l'impulso, mantenendo dunque relativamente bassa l'interferenza intersimbolica; questa non è accettabile quando le tratte sono tuttavia lunghe.

Nell'ambito delle microonde solitamente le guide vengono utilizzate quasi sempre in monomodalità; il motivo per cui la cosa è più critica in ambito ottico è il fatto che le λ in ambito ottico sono molto ridotte, rispetto a quelle in ambito microonde: fare una guida monomodale in ottica significa avere una guida con dimensioni tecnologicamente ottime (bassissime incertezze) nell'ordine dei μm .

Il numero dei modi, va valutato a partire dall'apertura numerica, parametro sul quale si dedicherà una sottosezione.

Apertura numerica (NA)

L'apertura numerica è un parametro che deriva dall'applicazione del modello dell'ottica geometrica. Si consideri una struttura di questo tipo:

A sinistra, la guida viene troncata, e a sinistra si ha il vuoto (indice n_0); supposta simmetrica (per semplicità) la guida, si manda dentro essa un'onda piana, con un certo angolo ϑ_0 (rispetto alla normale); l'onda trasmessa avrà

un vettore “più inclinato”, dal momento che $n_f > n_0$, quindi si ha riflessione verso l’alto, quindi riflessione verso il basso da qui per riflessione totale, e così si ha un modo guidato all’interno della struttura.

Si supponga a questo punto che il raggio interno arrivi all’angolo limite: ϑ_0 sarà dunque d’ora in avanti l’angolo limite di incidenza, ossia il massimo angolo per cui il raggio rimane confinato nella struttura, senza sfuggirne. Dato φ l’angolo di incidenza sull’interfaccia tra core e cladding, si ha che:

$$n_f \sin \vartheta_t = n_f \sin(90 - \varphi) = n_f \cos \varphi = n_f \sqrt{1 - \sin^2 \varphi}$$

Se φ è φ_c , ossia l’angolo critico rispetto alla normale alle pareti della guida, si ha che:

$$\varphi_c = \arcsin \left(\frac{n_c}{n_f} \right) \implies \sin \varphi = \frac{n_c}{n_f}$$

quindi:

$$n_f \sqrt{1 - \sin^2 \varphi} = n_f \sqrt{1 - \left(\frac{n_c}{n_f} \right)^2} = \sqrt{n_f^2 - n_c^2} \triangleq \text{NA}$$

Questa è l’apertura numerica della guida dielettrica; ϑ_0 è detto “acceptance angle”, ossia è il massimo angolo per cui i raggi rimangono intrappolati nella struttura, senza poter “scappare fuori”.

Questo è un discorso di ottica geometrica: si parla di raggi, come si fa usualmente in ottica geometrica, siamo nelle condizioni di applicare questo modello? Non è detto: l’ottica geometrica è una buona approssimazione solo quando la dimensione dell’oggetto in cui si vuole applicare il modello è grande rispetto alla lunghezza d’onda λ ; se d’altra parte la guida fosse monomodale, $d \sim \lambda$, dunque l’ottica geometrica non si potrebbe usare; la contraddizione è evidente anche dal momento che sembrerebbe di poter fissare ϑ_0 arbitrariamente, a patto di stare nel core, ma noi abbiamo a disposizione solo valori discreti, ben definiti di n_{eff} , e ciascuno di questi valori (si può agevolmente dimostrare) è associato a un ben definito valore di ϑ_0 : possiamo dunque usare qualsiasi valore di angolo? Sembra di avere a che fare con una contraddizione, a meno che la guida non sia molto multimodale: se d è molto grande, i modi sono tantissimi, quindi sono spazati l’uno dall’altro molto poco, e il discorso è più credibile (è possibile scegliere qualsiasi angolo e bene o male vi sarà un n_{eff} associato, per il discorso prima fatto).

L’apertura numerica è un parametro che permette di capire quanti modi sono presenti all’interno della guida, e non solo: la radice quadrata interviene nel V precedentemente definito; normalmente, NA si impiega perchè si dice

che, se NA è grande, è “facile accoppiare il campo dentro”: essendo NA grande si hanno tanti modi, quindi non è necessario scegliere ϑ_0 con precisione: l’angolo di accettazione può essere molto grande.

Si immagini di avere una struttura in vetro e aria: $n_c = 1$, $n_f = 1,5$; viene fuori, in termini di apertura numerica:

$$\text{NA} = \sqrt{2,25 - 1} = \sqrt{1,25} \sim 1,1$$

ma questo è $n_0 \sin \vartheta_0$, dunque $\vartheta_0 \sim 90^\circ$: qualsiasi sia l’angolo del raggio, esso viene intrappolato al suo interno.

Le guide dielettriche non funzionano in realtà con salti così grandi: in realtà $n_c \sim n_f$: si può avere per esempio $n_f = 1,5$, $n_c = 1,495$; questo significa che l’angolo critico, essendo il rapporto così, sarà di 88° : i raggi viaggiano praticamente orizzontali nella fibra!

Esiste un metodo, basato sul tener conto del basso “gradino” di indice di rifrazione, che permette di studiare in maniera semplificata le fibre ottiche: approssimazione **weakly-guided**, ossia “debolmente guidata”: ciò porta a semplificazioni nello studio delle comunicazioni.

Una figura che spesso si trova è la seguente:

Questa è una rappresentazione alternativa delle curve di dispersione: essa presenta $\beta(\omega)$; il risultato è dunque il fascio di curve, in cui si ha a che fare con le curve da noi presentate, “rigirate” e “infilate nella struttura ad angolo”.

Velocità di gruppo e Goos-Haenchen effect

Anche nel caso delle guide dielettriche è possibile trovare un’espressione della velocità di gruppo v_g ; questa però ha una certa espressione, piuttosto complicata, giustificabile mediante lo studio di un particolare effetto che si può avere: l’effetto Goos-Haenchen. Si consideri la seguente rappresentazione:

Data una guida dielettrica, in realtà, la riflessione non avviene esattamente nella modalità precedentemente mostrata, ma in una modalità un poco diversa: si ha uno “spostamento”, un certo d tra il punto di incidenza e il punto di riflessione.

Questo non è esatto, o quantomeno non è quello che dovrebbe capitare con una vera onda piana: dire che si ha questo d è infatti sciocco, dal momento che, quando si ha a che fare con questi “raggi”, non è che se ne abbia solo uno per volta: se ne hanno infiniti e su una dimensione estesa; il campo non è dato da un raggio, da uno spaghetti, da una bacchetta: il campo non esiste solo in punti singoli, ma esiste in molti punti! Dire quale dei raggi si sposti, è sostanzialmente insensato, impossibile.

L’idea più formale è: il coefficiente di riflessione all’interfaccia dielettrica ha una certa fase; dire che esso introduce una certa fase coincide sostanzial-

mente con il dire che $\angle\Gamma_s$ introduce uno “spostamento”, come se la riflessione avvenisse a un “punto più alto”:

In realtà dire ciò è insensato dal momento che questo “spostamento” non è univoco: non si hanno formule che, al variare dell’angolo o del Γ o di altro, lo spostamento cambia, dunque non esiste una formula in grado di introdurre una qualche correzione alla guida.

Un modo più serio di osservare questo fenomeno è basato sul mandare all’interfaccia non un raggio, ma un fascio gaussiano, prodotto per esempio con un LASER: questo è un esperimento che permette di vedere che il fascio riflesso “parte con una traslazione”; in questo caso, ciò che avviene è legato al fatto che un fascio è fatto da tante onde piane, quindi quello che si ha qui è una sorta di “ritardo di gruppo”, “traslazione di gruppo”.

Eccitazione di guide d’onda dielettriche

Si è fatto cenno all’eccitazione di una guida d’onda dielettrica, supponendo che in qualche modo “da sinistra” provenga un’onda piana, che rimanga in essa confinata. Esistono altri modi, per effettuare questa eccitazione?

Una domanda preliminare: data una situazione di questo genere, si ha un modo guidato?

La risposta è no: affinché un modo sia **guidato**, si deve avere un campo evanescente sia nel cover, sia nel substrato; in questo caso la situazione sarebbe corretta nel substrato, ma non nel cover; si può pensare a questa situazione come assolutamente analoga a quella dello *slab*, in cui si ha una certa eccitazione “da sinistra” (in questo caso “dall’alto”), quindi un certo insieme di ondulazioni nel film, infine una coda evanescente nel cover.

Si tenga presente che:

$$\beta = k_0 n_c \sin \vartheta_i$$

in totale analogia con lo slab: di fatto, β è esattamente identico a ξ , dunque anche la sua relazione con ϑ_i lo sarà; si ha che:

$$n_{\text{eff}} = \frac{\beta}{k_0} = n_c \sin \vartheta_i$$

questo è un numero sicuramente inferiore o uguale a n_c , ma quindi che non rispetta, per ipotesi, la condizione $n_s < n_{\text{eff}} < n_f$: ovviamente, in questa situazione, il modo non potrà essere guidato.

Altra osservazione potrebbe essere quella per cui il modo è un modo dello spettro continuo: essendo un modo dello spettro continuo, non può anche essere un modo discreto.

Ora è stato spiegato come **non** si eccita una guida d'onda, però possiamo intuire, da qua, come si fa a eccitarla! Si è detto che il problema è sostanzialmente il fatto che in n_c si avrebbe un campo non evanescente; questo suggerisce che la via per introdurre l'alimentazione è quella di mandare, in qualche maniera, un'onda evanescente nel cover, che ecciti un modo guidato nella struttura.

Prima di tutto, si fissi una certa f_0 ; a f_0 è associato k_0 , come ben noto:

$$k_0 = \frac{2\pi f_0}{c}$$

a questo punto, sappiamo che relazione si ha tra β e ϑ_i . Quello che si può fare, al fine di introdurre onda evanescente, è il seguente stratagemma:

Si ha un prisma, sul quale si incide una certa onda piana, con incidenza solitamente normale; il fatto dell'incidenza normale non è necessario, ma, dal momento che solitamente sul prisma si introducono degli strati antiriflesso, in modo da ridurre le perdite di disadattamento e aumentando dunque l'efficienza del sistema, essi sono solitamente migliori o comunque più facili da progettare per incidenza normale. L'onda entra, e quindi prosegue, fino a incidere con la "base" del triangolo, del prisma, con un certo angolo ϑ_i ; questo è un angolo molto ben definito, dal momento che, fissata f_0 e utilizzando la relazione di dispersione (la quale, si ricorda, ha soluzioni, dunque n_{eff} , ben definite, discrete), è possibile ricavare n_{eff} e quindi il β che si ha quando si vuole eccitare un particolare modo. Invertendo quindi:

$$\beta_{\text{modo guidato}} = k_0 n_{\text{prisma}} \sin \vartheta_i$$

è possibile trovare l'angolo e, con un po' di goniometria, trovare tutti gli angoli del prisma e definirne la geometria. Supponendo che il prisma abbia quindi un n simile a quello del film, si ha una situazione di questo tipo:

un fenomeno di FTR: riflessione totale frustrata. Cosa capita, in pratica, nella struttura guidante? Qualcosa di questo genere:

L'onda in questo modo entra, va in basso (si noti che questo è il contributo dei soli k_x , quindi è "verticale" perchè non stiamo ancora tenendo conto di β , che si occupa della "propagazione lungo z "), incontra la parete, e, essendo evanescente nel substrato, si ha riflessione totale (non frustrata) e torna indietro. Quando torna indietro, tuttavia, si ha qualcosa di "spiacevole": una volta iniettata l'onda nella struttura guidante, si vorrebbe che questa vi rimanesse confinata, ma in realtà ciò non accade: si ha infatti una situazione analoga a quella di partenza, con la "freccia girata al contrario": come l'onda tendeva a entrare nella guida, ora tende a entrare nel prisma, "sgocciolando" fuori dalla guida (almeno in parte). Questo è il motivo per

cui il prisma deve essere corto: fino a quando c'è il prisma, si rientra in una situazione in cui si ha riflessione totale frustrata, e quindi “sgocciolamenti” di energia dalla struttura guidante.

La situazione totale è la seguente:

Il fatto di avere questi “sgocciolamenti” riduce l'efficienza del sistema: parte dell'energia viene “riflessa” dal sistema risultante, dal momento che viene “trasmessa al di fuori della guida”. Dal momento che l'energia non esce più, il modo è “stabile”: si propaga secondo la propria legge, tipica dei modi guidati (dipendente da $e^{-j\beta z}$). Questo tipo di accoppiamento è basato sul “accoppiatore a prisma”, per ovvie motivazioni; il prisma è solamente un trucco per costruire un'onda piana evanescente tra prisma e guida; in altre parole, questo meccanismo converte un'onda piana omogenea in un'onda piana evanescente.

***m*-line spectroscopy**

Il meccanismo basato su questo prisma è stato utilizzato per alimentare la struttura, per introdurre una sorgente, ma in realtà può anche essere utilizzato per l'obiettivo opposto: per “accoppiare in fuori” della potenza:

In questo modo, si ha che il modo che arriva viene inviato “in fuori”, e, grazie alla riflessione totale frustrata, parte dell'energia del modo esce sotto forma di onda piana; l'angolo con cui esce dipende ovviamente dal β del modo che è stato “accoppiato in fuori”. La tecnica appena descritta è detta *m*-line spectroscopy, e, al fine di comprenderne il nome, si consideri il seguente esempio:

Si immagini di avere a che fare con tre modi, con costanti di propagazione $\beta_1, \beta_2, \beta_3$, tutti e tre eccitati e che si propagano, ciascuno con un certo peso, nella struttura guidante; se si introduce il “prisma di uscita”, il campo sulla faccia inferiore, che “esce” per riflessione totale frustrata, si può pensare come la sovrapposizione di tre onde piane:

Ciascun modo è infatti associato a un β_i (che sono le costanti di propagazione), il quale è associato a un ϑ di incidenza sulle pareti; è dunque ragionevole pensare che sia possibile questo tipo di scomposizione; considerando ciascun modo uno per volta, a seconda di β varia l'angolo, dunque la “pendenza” dei raggi. Al di fuori del prisma, si avrà una situazione di questo tipo:

Di questi tre modi, quello fondamentale è quello con n_{eff} più grande: più modi ci sono, più quelli “superiori” hanno n_{eff} elevato; questa cosa è ovvia, vedendo le curve di dispersione: se si hanno più modi, quello che aveva la “frequenza critica” più bassa è quello che è “più attivo”, e che quindi ha

un n_{eff} sicuramente più elevato; in altre parole, è quello per cui la curva di dispersione è “più alta”.

Dal momento che si ha il legame tra β e ϑ , e che

$$\beta = k_0 n_{\text{eff}}$$

allora è ovvio che il modo ad angolo più basso è quello con n_{eff} più elevato. In altre parole, il modo “più verticale” è il modo di indice massimo che riesce ancora a rimanere confinato nell’interfaccia: se infatti un modo è “molto ripido” allora esso “esce”! Il modo di indice più alto è quello massimamente verticale.

Ora che abbiamo queste tre onde uscenti dal prisma, è possibile per esempio proiettarle su uno schermo; il risultato, quindi, è il seguente:

Si hanno delle “macchie di luce”, “allungate e allargate” in due dimensioni. Il disegno andrebbe interpretato “ruotato”, come se il foglio di carta fosse “con il piano del foglio uscente dal foglio”.

I fenomeni sono due: un “allungamento” e un “allargamento”.

- Come mai si ha l’allungamento? Prima di tutto, si tenga ben presente che questo “allungamento” è sull’asse y ; questa informazione è interessante, dal momento che la struttura che stiamo considerando è una guida **planare**, ma in pratica una guida che occupi tutto y non esiste: dovrebbe infatti essere di dimensioni infinite. Aldilà di questo aspetto, spesso trascurabile, c’è da tenere conto che, quando si eccita, quello che si utilizza è un campo in cui $\frac{\partial}{\partial y}$ non è nulla: il LASER, che si usa per eccitare la struttura, ha infatti un fascio tondo, che in qualche modo eccita i campi; lateralmente tuttavia la guida non fa nulla, quindi la caratteristica del modo dipende solamente dal fascio buttato dentro, non dalla guida. Questo non è un modo: i modi di propagazione che sono stati ricavati non sono in realtà tutti i modi, dal momento che sono stati ricavati, nella nostra analisi, solo quelli che **non dipendono da y** ; esistono modi che dipendono per esempio linearmente da y , o altri, e noi non ne abbiamo tenuto conto: abbiamo solo considerato i modi che vanno “dritti”, non quelli che “deviano”, andando un po’ “ad angolo” in tutte le direzioni possibili. Il fascio gaussiano, rispetto a y , è come se fosse fatto di tante onde piane, ognuna delle quali dà luogo a un modo di propagazione “a zig zag”, introducendo un’infinità di onde piane, nella direzione y , che descrivono l’allargamento del fascio su y . Da sopra, si vede ciò:

la macchia si “allarga”, dando luogo all’allungamento sullo schermo: questo è dovuto al fascio gaussiano che si allarga.

- Per quanto riguarda la “dimensione stretta”, l’allargamento, esso deriva dal fatto che il prisma ha un’ampiezza finita, e non enorme come potrebbe essere quella della guida planare, rispetto alle larghezze. Questa questione è sostanzialmente riconducibile, nel dominio del tempo, a questa situazione:

quando si ha una sinusoidale che parte da $t = 0$ e finisce per un certo tempo, lo spettro è una sinc, traslata alla frequenza della sinusoidale. La larghezza del picco è sostanzialmente collegato al numero di cicli contenuti nell’intervallo in cui la sinusoidale è non nulla; se vi fossero milioni di cicli, la sinc sarebbe sostanzialmente una δ , ma se vi fosse un numero basso di cicli, ciò non sarebbe vero. Qualcosa di analogo nel dominio dello spazio è quello che accade con le antenne ad apertura: esse producono non un’onda piana, ma un’onda sferica. Un’antenna a parabola è ciò che produce il campo che più assomiglia a un’onda piana, ma solo sulla superficie della antenna: si può pensare di avere un’onda piana, per una “porta spaziale” delle dimensioni dell’apertura. Questa “finestratura” è quella che porta ad avere un lobo di una certa dimensione, nel diagramma di irradiazione, e la stessa cosa avviene nel caso del prisma: il fatto di avere un prisma finito, “finestra” l’onda piana, ottenendo questa sorta di allargamento che in un certo senso ricorda in effetti quello di un diagramma di irradiazione.

Questi ellissi, per quanto allungati e allargati, vengono chiamati “lines”: da qua parte del nome. m è dipendente dal modo: noi abbiamo usato ν , ma in realtà nei primi testi si utilizzava m , e da qua il nome. Per quanto riguarda il nome “spectroscopy”, esso deriva dal fatto che, in un certo senso, si fa un’analisi di tipo spettroscopico: usualmente si parla di spettroscopia quando si fa l’analisi nel dominio della frequenza; in questo caso, invece di avere un’analisi nel dominio della frequenza, si ha una sorta di “dominio dei modi”, e la spettroscopia è spaziale: una sorta di spettroscopia spaziale.

In effetti un prisma usualmente fa spettroscopia, ma non in questo senso: di solito, data luce bianca, la si scompone in diverse luci, luci dai diversi colori, quindi con diverse λ_i , quindi posizionando i ricevitori in modo da prendere una o un’altra componente, si hanno informazioni sulle varie componenti. Questa idea può essere usata in molti modi, per esempio per misurare gli indici di rifrazione della stessa struttura (ad esempio n_f): introducendo vari modi nella struttura, si misurano i vari β , e quindi si può risalire a n_f mediante il metodo dei minimi quadrati!

Sono state analizzate sostanzialmente due tecniche per eccitare modi in una guida; solitamente i prismi non si usano, dal momento che non si usano nemmeno strutture di questo tipo in pratica; quello che si fa di solito è

comunque utilizzare altre strutture per alimentare, come nel *BUTT* coupling: “accoppiamento da davanti”; per eccitare un modo si potrebbe per esempio usare una fibra ottica, tenendo comunque conto del fatto che anche questo tipo di eccitazione ha perdite: il modo della fibra è diverso dal modo della struttura, quindi la differenza tra i due darà riflessioni.

5.4.4 Strutture planari

Una barra di dielettrico rettangolare non è semplice da studiare, analiticamente: una guida metallica sicuramente è semplice, dal momento che il fatto di avere a che fare con strutture metalliche permette di utilizzare il principio della separazione delle variabili; una struttura di questo tipo, invece, non è per niente semplice da studiare:

Una cosa del genere non si può studiare analiticamente: a seconda dei “tagli” che si fanno si hanno diverse strutture; gli angoli, poi, sono il problema grosso; se nella guida rettangolare è come avere il prodotto cartesiano di due strutture, e ora non è più così: a seconda dei tagli, si può o meno incontrare il core (in alcuni tagli si ha solo il cladding, in altre anche il core).

5.4.5 Metodo dell’indice di rifrazione efficace

A questo punto si vuole presentare un metodo (approssimato) per l’analisi di strutture guidanti tridimensionali, ossia nelle quali non si ha solo confinamento in una delle due coordinate trasversali, bensì in entrambe. Queste strutture da studiare solitamente, come già anticipato in precedenza, sono molto difficili (e, in molti casi, analiticamente impossibili). Quello che si vuole presentare ora è un metodo approssimato ma molto interessante poichè di applicazione relativamente semplice.

Si immagini di avere a che fare con una guida **ridge** dielettrica, ossia con una struttura di questo tipo:

In questa struttura applicare il metodo di separazione delle variabili, al fine di ottenere un risultato esatto, è impossibile; quello che noi faremo, tuttavia, è applicarlo comunque, non curandoci dell’approssimazione introdotta (almeno per ora). Quella disegnata è la sezione trasversale della guida: i modi sono calcolati rispetto a questa, e poi l’invarianza per traslazione si ha lungo la direzione z , “entrante” nel foglio.

Volendo presentare questo metodo, i passi da seguire (nel caso della struttura ridge) sono ora presentati.

1. Si incomincia effettuando l’analisi lungo x (asse verticale, in questo caso): questo, dal momento che è evidente la presenza di tre mezzi,

dunque di una “guida”. La struttura è divisibile in tre parti, essendo tre i mezzi: tutte e tre sono guide planari, ma quelle a sinistra e a destra sono larghe f , quella centrale è larga h , dove $h > f$.

2. Dall’analisi delle tre strutture su x si ricavano gli n_{eff} per le tre parti: tre indici di rifrazione efficace (che in realtà sono 2, essendo due strutture identiche tra loro): N_f , relativo agli strati di spessore f , e N_h , relativo agli strati di spessore h . Per fare ciò, prima di tutto si deve risolvere l’equazione di dispersione, per esempio utilizzando le frequenze normalizzate:

$$V_f = k_0 f \sqrt{n_f^2 - n_s^2}$$

$$V_h = k_0 h \sqrt{n_f^2 - n_s^2}$$

(in alternativa usando l’equazione di dispersione). Da qui, si possono ricavare i valori di b_f e b_h (che sarebbero gli n_{eff} normalizzati), usando una delle varie curve con $a^{\text{TE,TM}}$ o direttamente risolvendo le equazioni di dispersione.

3. Una volta ottenuto tutto ciò, si invertano le seguenti equazioni:

$$b_f = \frac{N_f^2 - n_s^2}{n_f^2 - n_s^2}$$

$$b_h = \frac{N_h^2 - n_s^2}{n_f^2 - n_s^2}$$

dove N_f e N_h sono degli indici di rifrazione **efficaci**.

4. Una volta trovati gli indici di rifrazione efficaci, si può trattare la struttura lungo y come un’altra guida dielettrica, dove però gli indici di rifrazione non sono “fisici”, “propri della materia”, bensì sono proprio quelli efficaci; lo strato centrale è più largo, avrà un $N_h > N_f$, e quindi sarà il “film” della guida. Questa situazione è equivalentemente analoga all’analisi di una lastra di vetro, dove però gli indici di rifrazione non sono fisici, bensì efficaci; dal momento che, tuttavia, anche quelli che stiamo chiamando “fisici” rappresentano una media del comportamento della materia rispetto all’eccitazione del campo, si può dire che non ci stiamo sbagliando molto considerando un’approssimazione di questo genere.

5. Applicando lo stesso procedimento agli indici efficaci, risolvendo ancora una volta l'equazione di dispersione (o usando le curve di dispersione), si ricavano i vari valori, e si ottiene alla fine di tutto:

$$b_a = \frac{N_a^2 - N_f^2}{N_h^2 - N_f^2}$$

infatti, in questo caso, si ricorda che N_h rappresenta il “film” della struttura guidante equivalente, N_f il “substrato”.

Per concludere su questo metodo, esso è approssimato, dunque non ha senso studiare questo N_a considerando troppe cifre decimali; esso, ad ogni modo, è sicuramente valido per avere un'idea di massima su quale sia l'andamento di β (costante di propagazione) per la struttura. Dove sta l'approssimazione del metodo? Per quale motivo esso produce risultati sostanzialmente validi? La risposta è: questo metodo applica il metodo della separazione delle variabili, in un ambito dove le funzioni non sono a variabili separabili; il risultato per questo motivo, di sicuro, non potrà essere esatto! Se questo risultato sarebbe infatti esatto nel caso di una guida rettangolare metallica, situazione in cui si può pensare alla struttura come data dal prodotto cartesiano di due domini, qua non è assolutamente possibile fare considerazioni di questo genere. Il motivo per cui tuttavia il risultato è accettabile, è che l'errore è piccolo dal momento che l'errore sta in questa regione:

In questa zona, il campo è evanescente, ma non solo per una causa, bensì per due: in questa zona il campo è evanescente sia perchè si esce dal dente, sia dalla guida laterale: è come avere un campo “evanescente al quadrato”: come avere il campo dato dal prodotto di due campi evanescenti, dunque molto piccolo; il risultato è il fatto che quindi, essendo trascurato questo campo, il risultato finale è accettabile.

5.5 Teoria dei modi accoppiati

L'argomento conclusivo per la trattazione è quello della teoria dei modi accoppiati.

Si parla di modi di strutture guidanti, dunque di modi guidati, facendo in principio soprattutto riferimento alle nozioni dei corsi di Campi Elettromagnetici. In tal senso, si può pensare che parlare di “modi che si accoppiano” è un'affermazione autocontraddittoria: non solo i modi sono funzioni ortogonali tra loro, ma sono autofunzioni rispetto all'operatore che descrive la propagazione nella struttura guidante. Si propone a questo punto rapidamente il

procedimento che si segue per lo studio della propagazione guidata, al fine di comprendere meglio questa cosa.

5.5.1 Cenni al formalismo di Marcuvitz e Schwinger

Lo studio della propagazione guidata parte sostanzialmente dalla soluzione di un sistema di equazioni, note come “equazioni di Marcuvitz e Schwinger”:

$$\begin{cases} -\frac{\partial \underline{E}_t}{\partial z} = j\omega\mu \left[\frac{\nabla_t \nabla_t}{k^2} + \underline{\underline{I}} \right] \cdot (\underline{H}_t \times \hat{z}) \\ -\frac{\partial \underline{H}_t}{\partial z} = j\omega\varepsilon \left[\frac{\nabla_t \nabla_t}{k^2} + \underline{\underline{I}} \right] \cdot (\hat{z} \times \underline{E}_t) \end{cases}$$

Questo sistema è strutturato nella seguente maniera: si ha da una parte la derivata rispetto a z di un campo trasversale, dall'altra parte le derivate rispetto a x e y . L'idea a questo punto è quello di considerare l'espansione modale dei campi, ossia di esprimere i campi in questa forma:

$$\begin{cases} \underline{E}_t = \sum_n V_n(z) \underline{e}_n(\underline{\varrho}) \\ \underline{H}_t = \sum_n I_n(z) \underline{h}_n(\underline{\varrho}) \end{cases}$$

Una nota sul significato di tutto ciò: queste funzioni \underline{e} e \underline{h} non sono una successione ortogonale a caso, ma hanno una particolare proprietà: esse sono una successione ortonormale di autofunzioni dell'operatore $\nabla_t \nabla_t$; questo significa, in altre parole, che non solo sono funzioni ortonormali, nel senso che il prodotto scalare è nullo a meno che non si faccia il prodotto di due funzioni uguali, ma anche che la soluzione di questo sistema è, sostituendo e proiettando³ i campi su ciascuna delle autofunzioni:

$$\begin{cases} -\frac{dV_n}{dz} = jk_{z,n} Z_{\infty n} I_n \quad \forall n \\ -\frac{dI_n}{dz} = jk_{z,n} Y_{\infty n} V_n \quad \forall n \end{cases}$$

Cosa significa tutto ciò? Ciascuna di queste equazioni vale $\forall n$, ma si noti che, quando si effettua la proiezione:

$$\langle \underline{E}_t | \underline{e}_n \rangle = \int_{\Sigma} \underline{E}_t \underline{e}_n^* d\underline{\varrho}$$

quello che si ottiene è che, essendo \underline{E}_t dato dalla somma di autofunzioni ortogonali, ciascuno dei sistemi di equazioni (delle linee) appena presentati dipende solo ed esclusivamente dal valore di n considerato, non da $n+1$ o $n-1$ o altri: ciascun sistema fa storia a sè. In altre parole, volendo considerare un sistema in forma matriciale (sistema che sarebbe $\infty \times \infty$), in questo caso

³ossia applicando il procedimento di Gram-Schmidt, data l'ipotesi di ortogonalità

esso sarebbe diagonale, dal momento che è stata sfruttata l'ipotesi di avere una base di autofunzioni e non solo una base: il fatto che sono autofunzioni implica che ciascuna proiezione dà luogo a un singolo elemento del sistema, dipendente da un solo valore di n : un sistema diagonale. In altre parole, quando proietto il \underline{E}_t su \underline{e}_1 , il risultato dipenderà solo da $n = 1$; in altre parole ancora, in questa situazione, si vede chiaramente che i modi sono uno disaccoppiato dagli altri: dipendenza di ciascuna linea modale da solo un valore di n ; ciascun n fa vita a sè. Questo non dovrebbe stupire: avendo usato i campi come somma delle funzioni modali, il risultato è proprio la diagonalizzazione.

Caso non standard: guida non standard, con funzioni modali sconosciute

Si immagini a questo punto, per esempio, di avere a che fare con una situazione di questo genere:

Cosa abbiamo, in questa situazione? In un qualche senso, si ha a che fare con una guida che “assomiglia” a una standard: a una guida d'onda rettangolare. Questa non è tuttavia esattamente una guida rettangolare, a causa per esempio del cilindretto di dielettrico in essa inserito. In questo caso, quello che si potrebbe fare, dal momento che le funzioni modali non sono note, è utilizzare, come **funzioni di espansione** per i campi, le stesse funzioni \underline{e}_n e \underline{h}_n di prima. Si noti, tuttavia, che in questo caso queste sono sicuramente una successione ortogonale (come lo erano prima, lo sono tutt'ora, dal momento che l'ortogonalità è una proprietà delle funzioni), però non più di autofunzioni: in questo caso, questa successione non è più una base di autofunzioni per il problema, di conseguenza il sistema risultante non sarà più diagonale. In altre parole, quello che si avrà sono linee in cui la proiezione non dipende solo più da un singolo valore di n , bensì anche da altri valori.

Fisicamente, cosa significa tutto ciò? Si immagini di espandere il campo, proiettandolo nella base di modi che si vuole considerare; quando si studia il contributo di un modo nel sistema, quello che si avrà in esso non sarà più un semplice modo singolo, bensì diversi modi⁴: dal momento che la proiezione non dà più luogo a funzioni dipendenti da un singolo n , quello che si ha a questo punto è un sistema per cui, data l'eccitazione di un modo, dopo se ne avranno svariati che si propagano. Questo significa “accoppiamento modale”: utilizzando come funzioni di espansione per una struttura le funzioni modali di una struttura simile (non è detto che si debbano usare funzioni relative

⁴il primo modo del sistema in studio, per esempio, sarà costituito da una combinazione lineare delle varie funzioni di espansione della struttura ad esso simile, non da una sola di esse, dal momento che esse non sono più autofunzioni

a strutture “simili”, ma c’è da usare “furbizia” nella scelta delle funzioni di espansione, quindi tendenzialmente sarà così), introducendo un modo, per il fatto che non si ha più a che fare con autofunzioni, il sistema sarà non più diagonale, ma sparso: un modo trasferisce, in altre parole, la propria potenza ad altri modi.

Questo accade dal momento che quello che si fa, in realtà, è ciò:

$$\langle \underline{e}_m, \left[\frac{\nabla_t \nabla_t}{k^2} + \underline{\mathcal{I}} \right] \underline{e}_n \rangle$$

questo dettaglio prima non era stato scritto dal momento che, nel caso precedente, essendo \underline{e}_n un’autofunzione, il risultato dall’applicazione dell’operatore⁵ era un vettore (funzione) parallelo a \underline{e}_n : autofunzione! Ora, invece, si deve tenere conto della rotazione prodotta dall’applicazione dell’operatore al vettore.

Si considera quindi un matricione \underline{A} (che sarebbe $\infty \times \infty$), lo si diagonalizza, e quindi si trova \underline{M} : \underline{M} è la matrice degli autovettori relativi a \underline{A} , e quindi si ha:

$$\begin{bmatrix} M_{11} \\ M_{21} \\ M_{31} \\ \vdots \end{bmatrix}$$

In questo modo, si ottiene:

$$\underline{E}_t = \sum_n M_{n,1} \underline{e}_n$$

ossia, si prendono i numeri della prima colonna (ricordando che le colonne han significato di autovettori), li moltiplico per la funzione di espansione, e si ottiene un campo funzione di (x, y) . \underline{E}_t è il primo modo della struttura “stravagante”, descritto nella base delle \underline{e}_n , ossia quelle che erano le autofunzioni relative alla struttura precedente. Il problema è numerico, dal momento che la ricerca degli autovalori e degli autovettori va, generalmente, fatta numericamente (anche perchè si ha a che fare con matrici enormi). Usare i modi della “guida standard” per descrivere quelli della guida “non standard” significa usare il **metodo dei modi accoppiati**. I modi della

⁵si noti che l’operatore in realtà tiene conto della geometria: le autofunzioni sono relative all’operatore, con una geometria ben definita; la geometria, nel dettaglio, si definisce al momento della definizione del prodotto scalare: è contenuta nella Σ su cui si effettua l’integrazione

struttura “non standard” son dati dall’accoppiamento di diversi modi della struttura “standard”

Esempio matematico per comprendere il problema

Si supponga di dover studiare il seguente problema differenziale:

$$\frac{d^2 f}{dx^2} (k^2 + x^2) f = \lambda f$$

questo problema, come si può vedere, assomiglia molto a quest’altro problema:

$$\frac{d^2 f}{dx^2} + k^2 f = \lambda f$$

Si supponga di voler trovare gli autovalori di questo problema “perturbato”, nel dominio $x \in [0, a]$.

Come si fa? Beh, se non ci fosse la x^2 , il problema sarebbe facile; basterebbe utilizzare il set di autofunzioni relative al problema “semplice”, definito come:

$$f = \sum_n c_n \sin\left(\frac{n\pi}{a}x\right)$$

Queste, per il nostro problema “strano”, non sono più un set di autofunzioni; quello che si può fare, tuttavia, è sostituire le espressioni (derivando e così via), ottenendo il seguente sistema:

$$\sum \left[-\left(\frac{m\pi}{a}\right)^2 c_n \sin\left(\frac{m\pi}{a}x\right) \right] + \sum_n (k^2 + x^2) c_m \sin\left(\frac{m\pi}{a}x\right) = \lambda f$$

A questo punto, posso portare λf a sinistra, quindi proiettare ambo i membri, $\forall m$, per $\sin\left(\frac{n\pi}{a}x\right)$: ciò prima aveva senso, dal momento che i seni, per $m \neq n$, sono una base ortonormale; quello che si ottiene è ciò:

$$-\left(\frac{n\pi}{a}\right)^2 c_n + (k^2 - \lambda^2) c_n + \sum_m \int \sin\left(\frac{m\pi}{a}x\right) x^2 \sin\left(\frac{n\pi}{a}x\right) dx$$

Infatti, nella prima parte dell’espressione tutti i termini si annullano, per $m \neq n$, dal momento che l’integrale (integrale di proiezione) veniva sempre nullo; nell’ultimo caso, tuttavia, la presenza del x^2 rovina la cosa, dal momento che in questo caso l’integrale non è più nullo, quindi si ha a

che fare con altri termini; volendo esprimere questo sistema mediante una matrice, nel caso “standard” essa verrà diagonale, ma nel caso analizzato no, a causa di questo x^2 che “perturba il problema”, rendendolo difficile da risolvere rispetto a quello di partenza.

In altre parole, quando si parla di accoppiamento modale, si parla dell'accoppiamento che si ha tra le funzioni di espansione scelte! Che non saranno di sicuro autofunzioni rispetto al problema “non standard”, ma magari potrebbero esserlo per un problema ad esso simile.

5.5.2 Studio dell'accoppiamento modale su guide ottiche

Si consideri a questo punto un problema di reale interesse: l'accoppiamento modale tra guide ottiche. Si considerino dunque due guide slab a distanza non infinita: quali sono i modi di **questa** struttura?

Questa è, a tutti gli effetti, una singola guida d'onda, dal momento che la struttura, per quanto particolare, è tale da confinare il campo entro una regione spaziale finita; d'altra parte, però, questa ricorda anche molto la forma di due guide d'onda, messe una di fianco all'altro; da ingegneri, l'idea potrebbe essere quella di recuperare gli “attrezzi” già noti, al fine di costruire una teoria in modo relativamente semplice.

Si può ancora dire che le funzioni modali di questa struttura sono quelle della singola guida d'onda? Assolutamente no: è proprio quello che è stato spiegato nella sottosezione precedente! Infatti, queste funzioni possono essere sicuramente usate come funzioni di espansione, ma queste non saranno autofunzioni” per questa struttura, dunque si avrà “accoppiamento modale”.

Un'idea, che verrà confermata a posteriori, è quella di studiare il sistema come un sistema 2×2 : questo fatto è per ora non motivato, dal momento che la struttura supporta un certo insieme di modi (contando quelli dello spettro continuo, infiniti), però di essi se ne considera solo uno in una struttura, un altro nell'altra. Il sistema, descritto nel maggior dettaglio, è il seguente:

Dato campo elettromagnetico \underline{E} , esso è descrivibile da:

$$\underline{E} = a_1(z)u_1(y)$$

u_1 è la funzione modale relativa alla **singola** guida d'onda, nella struttura 1; a_1 è la funzione dell'ampiezza, al variare di z , del campo nella guida. Si può dimostrare che le equazioni differenziali che, se risolte, soddisfano l'andamento di queste ampiezze, sono:

$$\frac{da_1}{dz} = -j\mathcal{C}_{21}e^{j(\Delta\beta)z}a_2(z)$$

$$\frac{da_2}{dz} = -j\mathcal{C}_{12}e^{-j(\Delta\beta)z}a_1(z)$$

Si dice che $\mathcal{C}_{12} = \mathcal{C}_{21} = \mathcal{C}$ è il “coefficiente di accoppiamento”; esso è definito come:

$$\mathcal{C}_{21} = \frac{1}{2} (n_2^2 - n_1^2) \frac{k_0}{\beta_1} \int_a^{a+d} u_1(y)u_2(y)dy$$

dove u_1 e u_2 sono le due funzioni diseguate: quelle che erano, per le singole strutture considerate infinitamente distanti tra loro, le autofunzioni modali, ora usate da funzioni di espansione. Si noti, da questa definizione, che tanto più le guide sono lontane, tanto più \mathcal{C} è piccolo: si ha infatti a che fare con il prodotto u_1u_2 , ma se u_1 (nel caso di \mathcal{C}_{21}) è sempre ampio uguale, valutato tra d e $a + d$, u_2 in quell'intervallo è solo una coda evanescente; tanto più lontane sono le strutture guidanti, tanto più attenuato sarà il campo, nell'onda evanescente, quindi piccolo, e il valore del prodotto conseguentemente piccolo.

Si è inoltre parlato di un $\Delta\beta$: esso è definito come:

$$\Delta\beta = \beta_1 - \beta_2$$

Questa è la differenza delle costanti di propagazione delle due funzioni di espansione, considerando le guide isolate: una relativa a u_1 , l'altra a u_2 .

Questo è un sistema di equazioni differenziali, ma non a coefficienti costanti; si può tuttavia dimostrare che, con un opportuno cambio di variabili, il sistema diventi un sistema di ODE, ottenendo quindi una matrice di numeri invece che di variabili. Facendo i vari conti, si può dimostrare che le potenze nelle due guide sono:

$$P_1(z) = \frac{1}{2} |a_1|^2 = P_1(0) \left[\cos^2(\gamma z) + \left(\frac{\Delta\beta}{2\gamma} \right)^2 \sin^2(\gamma z) \right]$$

$$P_2(z) = \frac{1}{2} |a_2|^2 = P_1(0) \frac{|\mathcal{C}_{12}|^2}{\gamma^2} \sin^2(\gamma z)$$

dove

$$\gamma = \sqrt{\left(\frac{\Delta\beta}{2} \right)^2 + \mathcal{C}^2}$$

Cosa vuol dire tutto ciò? Questa è la soluzione del problema, supponendo che la potenza sia “infilata solo nella guida 1”: questo significa che all'inizio,

per $z = 0$, la guida 1 è carica, la guida 2 è scarica, e poi al variare di z si lascia cambiare la situazione; quello che capita, quindi, è che u_1 trasporta potenza alla guida 2.

Si noti che, in generale, si ha che

$$P_1(z) + P_2(z) = P_1(0)$$

questo è buono: significa che, non essendoci perdite, la potenza si conserva; questo è evidente, con $\Delta\beta = 0$ ($\beta_1 = \beta_2$): in questo caso, infatti, le due guide (considerate isolatamente) sono identiche, u_1 e u_2 hanno la stessa costante di propagazione, e quindi si ha la condizione di sincronismo, o *phase matching*:

$$\gamma = \mathcal{C}$$

e quindi:

$$P_1(z) = P_1(0) \cos^2(\mathcal{C}z)$$

$$P_2(z) = P_1(0) \sin^2(\mathcal{C}z)$$

Facendo un grafico di ciò, si otterrebbe qualcosa di questo genere:

Si può definire una lunghezza L_0 , detta “lunghezza di accoppiamento”: essa vale

$$L_0 = \frac{\pi}{2\mathcal{C}}$$

Supponendo che la prima guida sia completamente carica di potenza, dopo L_0 la potenza lascerebbe completamente la guida 1 per andare nella guida 2; dopo un'altra L_0 tutto tornerebbe come prima, e così via.

Cosa capita invece, se $\Delta\beta \neq 0$? Bisognerebbe disegnare la funzione, ma quello che si vedrebbe, in sostanza, è qualcosa del genere:

il trasferimento di potenza in effetti si avrebbe, ma non completo, dal momento che il fenomeno è del tutto analogo, ma meno efficace; in questo caso, inoltre, la posizione su z del minimo varia:

$$L = \frac{\pi}{\gamma}$$

Facendo crescere ulteriormente $\Delta\beta$, quello che si avrebbe è ciò:

Con dei $\Delta\beta$ significativi, l'oscillazione, dunque il trasferimento di potenza da una struttura all'altra, è molto piccola; questo, cambiando solo il dissincronismo, e non la distanza tra le guide.

Al fine di avere accoppiamento in quantità importanti, quindi, vi sono due condizioni necessarie:

- un buon \mathcal{C} , dunque una distanza ridotta tra le due guide;
- il fatto che i due modi delle guide disaccoppiate siano il più possibile sincroni; questo si rispecchia nel richiedere che le due guide siano il più simili possibile.

Questa ultima affermazione motiva il fatto che la presenza di due soli modi accoppiati sia sufficiente per studiare questo problema con un'approssimazione comunque ragionevole: i modi delle due guide non possono sicuramente essere tutti sincroni tra loro, e anzi è già difficile che presentino buoni livelli di sincronismo anche solo due dei modi; i modi fuori sincronismo, come si può vedere, presentano solamente oscillazioni molto piccole di trasferimento di potenza, quindi sono trascurabili rispetto a quelli sincroni. Trascurando i modi dissincroni, si trascurano solo piccole oscillazioni sulle oscillazioni:

Applicazioni dell'accoppiamento modale: interferometri di Mach-Zehnder

Questo sistema di accoppiare guide è una cosa buona o cattiva? Beh, sostanzialmente, dipende: se si vogliono avere due guide con due connessioni diverse, per esempio due guide trasportanti telefonate, questo è crosstalk, quindi è una cosa cattiva; tuttavia, questo tipo di fenomenologia può essere utilizzato per realizzare dispositivi ottici, come per esempio accoppiatori direzionali:

Se si prendono delle guide, e si “avvicinano” in modo tale da avere una lunghezza inferiore di L_0 , si può scegliere, a seconda della distanza per cui le guide sono tra loro vicine, quanta potenza accoppiare; questo per esempio può essere utilizzato per realizzare degli accoppiatori a 3 dB, o con altri rapporti di partizione, a scelta (progettando L di conseguenza).

Quella che si vuole a questo punto proporre è una particolare applicazione di queste strutture: gli interferometri di Mach-Zehnder. Si consideri ciò:

Da un punto di vista di principio, gli specchi (1) e (3) servono solamente per “guidare” i raggi: lo specchio (2), semitrasparente, nel quale entra il segnale, il raggio di partenza, divide (per esempio a 3 dB) in due la potenza, che viene guidata da (1) e (3) a (4), dove i segnali si sommano e ricombinano. Se le distanze percorse sono tutte uguali, allora i segnali si sommano in fase; se invece c'è una differenza pari a $\lambda/2$, si ha uno sfasamento pari a π , e quindi i segnali si sottraggono.

Questo dispositivo è molto interessante dal momento che permette di convertire una modulazione di fase in una modulazione di ampiezza: otticamente, mediante materiali **elettroottici**, è possibile variare l'indice di rifrazione mediante un campo elettrico statico; girando la manopola di un potenziometro, è possibile aggiustare l'indice di rifrazione. Avendo poi la possibilità di cambiare la tensione dinamicamente, è possibile cambiare dinamicamente anche l'indice di rifrazione, e facendo così il k , che vale

$$k = k_0 n$$

cambiando n nel tempo, cambia anche k , e questa è quindi una modulazione di fase; applicando il segnale così modulato in fase in un Mach-Zehnder, la modulazione di fase diventa una modulazione di ampiezza, e viene così rilevata, mediante la conversione in una modulazione di ampiezza.

Il motivo per cui si parla di ciò, è il fatto che è possibile realizzare tutto ciò in ottica integrata:

Questo è, a tutti gli effetti, un accoppiatore di Mach-Zehnder: in questo caso gli specchi perfetti non servono più, dal momento che essi servivano solamente per fare da "guidaggio"; gli specchi "semitrasparenti" sono invece realizzati semplicemente con le zone delle guide vicine tra loro. Le due lunghezze delle guide sono diverse e, a seconda delle differenze di lunghezze, i segnali si sommano in fase, in controfase, o come voglio; utilizzando poi materiali elettroottici, posso cambiare lo sfasamento a parità di lunghezza.

Potrei fare cose ancora più complicate: se sulla guida superiore vi sono segnali non monocromatici, con due diverse λ_i , è possibile progettare il sistema per fare in modo che da una parte esca λ_1 , dall'altro λ_2 , sfruttando l'accoppiamento modale.

Presenza di dissincronismo

Si è parlato dell'importanza del sincronismo, ma di fatto non si ha la garanzia di avere sincronismo: volendo avere a tutti i costi avere trasferimento di energia da una struttura all'altra, si devono **portare i modi al sincronismo**, mediante il seguente trucco:

Date due guide, una con u_1 da costante di propagazione β_1 , l'altra u_2 con β_2 , l'idea è mettere, nella regione intermedia, un reticolo di Bragg: cambiare, periodicamente, l'indice di rifrazione; dato d il periodo del reticolo, è necessario che:

$$\beta_1 = \beta_2 + \frac{2\pi}{d}$$

supposto che $\beta_1 > \beta_2$: la differenza tra i due β fornisce la lunghezza del periodo del reticolo. In questa maniera, si sono portati i modi al sincronismo.

Questa cosa richiama un concetto di meccanica quantistica: quando si ha un atomo che ha un elettrone esterno, in un livello di energia E_1 , che può stare anche in un E_2 , si ha che:

$$E_1 - E_2 = h\nu$$

ossia, si ha emissione/assorbimento di un fotone con frequenza proporzionale alla differenza delle energie. Ciò è molto simile a quanto scritto prima:

$$\beta_1 - \beta_2 = \frac{2\pi}{d}$$

frequenza spaziale e frequenza temporale, ma le due cose coincidono, concettualmente: questo accade perchè l'analisi semplice dell'emissione/assorbimento di un atomo si formula con un'equazione i cui coefficienti sono simili a quelle di questa situazione: due guide con in mezzo il reticolo.

La cosa si può portare all'estremo: supponendo di avere a che fare con una sola guida, monomodale (dove in realtà per monomodale si intende il fatto che si ha sia un modo progressivo, sia un modo regressivo), si può fare ciò:

Come noto, modo progressivo e modo regressivo sono due onde, con costanti β_1 e $-\beta_1$: le onde sono ruotate di π , quindi

$$\Delta\beta = 2\beta_1$$

I due modi sono estremamente dissincroni; posso, tuttavia, utilizzare un reticolo, al fine di “ri-sincronizzare” il campo; a seconda di dove lo metto, esso interagisce con una diversa porzione di campo: potrei sia metterlo dentro la guida (e farlo interagire con i campi confinati), sia al di fuori (farlo interagire con i modi evanescenti, ma quindi sarebbe meno efficace). A parte ciò, quello che si può fare, è imporre che:

$$\Delta\beta = 2\beta_1 = \frac{2\pi}{d}$$

da qua

$$d = \frac{2\pi}{2\beta_1} = \frac{\pi}{\beta_1} = \frac{\pi}{\frac{2\pi}{\lambda_g}} = \frac{\lambda_g}{2}$$

Questo significa che il reticolo che porta il modo progressivo a essere sincrono con il modo regressivo, deve avere periodo pari a $\lambda_g/2$: il periodo che

dà luogo ad accoppiamento sincrono è questo. Se ci si pensa, gli specchi di Bragg avevano lo stesso periodo: due volte $\lambda_g/4$, che, sommati, portavano ad avere lo stesso sfasamento! Un'altra maniera, diversa, per capire il funzionamento degli specchi di Bragg, è proprio questo: quando un'onda si riflette, quello che capita è che si ha un accoppiamento tra l'onda progressiva, e un'onda regressiva nella stessa struttura: lo specchio di fatto introduce un trasferimento di potenza tra il modo già presente nella guida, e un modo che sarà ruotato di π . Lo specchio di Bragg, è un oggetto che produce una forte riflessione, dunque un buon trasferimento di potenza da onda incidente a onda riflessa, e ciò può essere descritto mediante l'accoppiamento modale: i due modi, in questo senso, possono per questo da esso anche essere portati al sincronismo. In generale quindi, se

$$\beta_1 - \beta_2 = n \frac{2\pi}{d}$$

con un n generico (prima n è stato fissato a 1), si ha la condizione soddisfatta.