

Dispense relative all'insegnamento
“Fibre ottiche e fotonica” (cod. 84472)

Corso di Laurea Triennale in
Ingegneria Elettronica e Tecnologie dell'Informazione
(classe L-8)

A cura di Mirco Raffetto

Ultimo aggiornamento: 4 marzo 2020
Aggiornamento precedente: 12 luglio 2019

Capitolo 1

Introduzione, motivazioni e cenni storici

Il titolo dell'insegnamento fa riferimento alle fibre ottiche e alla fotonica.

Cosa siano le fibre ottiche dovrebbe essere ben noto a tutti. Ad esempio, in uno dei nostri testi di riferimento [1] (pp. 15-16) si afferma: “An optical fiber is a light conduit made of two concentric glass (or plastic) cylinders”. Quasi tutti gli aspetti relativi alla propagazione di campi elettromagnetici nelle fibre ottiche o, più in generale, nelle guide dielettriche, possono essere studiati, con ottima approssimazione, usando la teoria classica dei campi elettromagnetici [1] (p. 445).

Più in generale, quando un campo elettromagnetico interagisce con la materia, in particolare modo quando la sua frequenza è elevata, come nei casi di interesse, non è più possibile trascurare gli effetti della quantizzazione. Alcuni fenomeni di interesse potranno essere compresi continuando a utilizzare la descrizione classica del campo elettromagnetico e introducendo la quantizzazione a livello atomico [2] (p. 3). Nei casi più complessi sarà necessario considerare che anche il campo elettromagnetico è quantizzato [1] (p. 448), [2] (p. 3), [3] (p. 4).

In questo insegnamento i fenomeni quantistici più complessi verranno descritti solo in modo qualitativo.

I quanti del campo elettromagnetico sono noti con il termine di fotoni [2] (p. 5) e la fotonica è la scienza o la tecnologia che si occupa di generare, controllare e rilevare fotoni [1] (p. XI).

Si potrebbero richiamare moltissime applicazioni, dai sistemi ottici di telecomunicazione agli apparati utilizzati in ambito medicale, e citare molti contributi per spiegare l'importanza delle fibre ottiche e della fotonica.

Per raggiungere l'obiettivo, tuttavia, mi paiono sufficienti un paio di citazioni:

- la tecnologia “lightwave”, insieme alla microelettronica, è stata considerata un fattore determinante per l'avvento della “information age” [4] (p. 1),
- la fotonica nel senso più ampio del termine ha un notevolissimo impatto sull'economia mondiale; si stima che il mercato globale della fotonica abbia un valore di 300 miliardi di euro; le previsioni per il 2020 parlano di un valore del mercato superiore a 600 miliardi di euro, con una crescita nell'industria fotonica più che doppia rispetto a quella del PIL (prodotto interno lordo) mondiale tra il 2005 e il 2011 [5] (link “Why Light Matters”, sezione “Economic Impact”).

Da un punto di vista storico, gli eventi più significativi, per gli argomenti di interesse in questo insegnamento, sono [4] (sezione 1.1), [1] (pp. 1085-1087):

- 1865 nel testo “A Dynamical Theory of the Electromagnetic Field”, pubblicato da James Clerk Maxwell, appaiono per la prima volta le equazioni di Maxwell al completo ed in forma differenziale (la notazione moderna più comune fu sviluppata successivamente da Oliver Heaviside)

- 1887 verifica sperimentale delle equazioni di Maxwell e dell'esistenza delle onde elettromagnetiche
- 1900 Max Planck introduce l'ipotesi quantistica per i valori dell'energia delle onde elettromagnetiche
- 1905 Einstein pubblica il suo lavoro sull'effetto fotoelettrico (fotoni)
- 1922 effetto Compton (o Compton scattering: la diffusione di onde elettromagnetiche di lunghezza d'onda maggiore rispetto a quella che incide su elettroni)
- seconda metà degli anni 20: formalizzazione della meccanica quantistica
- 1927 Dirac pubblica il lavoro sulla quantizzazione del campo elettromagnetico
- 1930?-1960 invenzione del LASER
- 1966 proposto utilizzo fibre ottiche
- 1970 ottenute fibre ottiche con attenuazione inferiore a 20 dB/Km
- 1970 invenzione laser a semiconduttore
- anni '80 del secolo scorso, studio e introduzione degli amplificatori ottici.

Capitolo 2

Propagazione nelle guide dielettriche

Nella seconda parte di questo insegnamento considereremo diversi componenti: dai laser ai led, per la generazione di onde elettromagnetiche, dai modulatori elettroottici, ai multiplexer e ai demultiplexer, per la trasmissione delle onde elettromagnetiche, per giungere agli amplificatori ottici e ai fotodiodi, per la propagazione e la ricezione dei segnali. Per comprendere il funzionamento di questi componenti sarà necessario introdurre qualche considerazione quantistica [1] (p. 445), [2] (pp. 3-6), [3] (pp. 3-4).

La prima parte dell'insegnamento sarà, invece, dedicata alla propagazione in guida dielettrica o fibra ottica. Questa parte potrà essere basata sulla teoria classica del campo elettromagnetico.

Le nozioni fondamentali di interesse per questo studio possono essere espresse in modo chiaro se si considerano modelli di strutture guidanti rettilinee, che consentono il confinamento trasversale del campo elettromagnetico, e in cui non sono presenti sorgenti del campo o variazioni geometriche o dei materiali coinvolti. Le nozioni potranno quindi essere dedotte facendo uso di sistemi di riferimento cilindrici, assumendo come coordinata assiale la coordinata z .

Assumeremo anche che i mezzi coinvolti siano tutti lineari, in modo da poter lavorare in regime sinusoidale permanente, utilizzando la trasformata di Fourier. Si useranno i cosiddetti fasori, scalari e vettoriali, assumendo un fattore comune a tutti i vettori di campo pari a $e^{j\omega t}$. I materiali lineari considerati inizialmente sono completamente caratterizzati da $\mu = \mu_0\mu_r$ (permeabilità magnetica (assoluta, del vuoto, relativa)) e $\varepsilon = \varepsilon_0\varepsilon_r$ (permettività elettrica (assoluta, del vuoto, relativa)). Spesso, nel contesto di interesse, si fa riferimento al cosiddetto indice di rifrazione $n = \sqrt{\mu_r\varepsilon_r}$.

Nella letteratura tecnica di riferimento per questi argomenti viene spesso utilizzato il termine “modo” (fibre multimodali o monomodali, dispersione modale, ecc.). Di cosa si tratta? I modi sono campi elettromagnetici monocromatici che soddisfano le equazioni di Maxwell (omogenee, per assenza di sorgenti), le relazioni costitutive e le condizioni al contorno della struttura guidante a cui fanno riferimento e che presentano una dipendenza dalla coordinata assiale di tutte le componenti di campo del tipo $e^{-\gamma z}$ (per l'onda progressiva; con segno opposto dell'esponente per l'onda regressiva). $\gamma \in \mathbb{C}$ è la cosiddetta costante di propagazione. La sua parte reale, α , è detta costante di attenuazione mentre il coefficiente dell'immaginaria, β , è detto costante di fase. Che caratteristiche hanno? Per rispondere a questa domanda abbiamo bisogno di calcolare i campi elettromagnetici che costituiscono i modi stessi. Per far ciò è necessario partire, come sempre quando si ha a che fare con l'elettromagnetismo classico, dalle equazioni di Maxwell e dalle relazioni costitutive. Una risposta alla domanda precedente richiederà in ogni caso un lavoro non semplicissimo.

Osservazione 1 *Nel seguito considereremo sempre onde caratterizzate dalla presenza di uno dei fattori di propagazione, $e^{\gamma z}$ o $e^{-\gamma z}$. Tenendo conto di ciò, non è restrittivo evitare di considerare un'ulteriore ambiguità di segno su β ed è possibile assumere $\beta \geq 0$ in ogni caso, identificando con*

ciò, se β è strettamente positivo, l'onda progressiva (regressiva) con quella che presenta il fattore di propagazione $e^{-\gamma z}$ ($e^{\gamma z}$).

Si verifica immediatamente, inoltre, che, in presenza di mezzi passivi, se $\beta > 0$ ($\beta < 0$) si deve avere $\alpha \geq 0$ ($\alpha \leq 0$), indipendentemente dal segno dell'esponente del fattore di propagazione $e^{\pm\gamma z}$.

Infine, per le stesse ragioni di prima, se $\beta = 0$, la presenza dei due fattori di propagazione ci consente di evitare di considerare $\alpha < 0$.

Prima di addentrarci nei calcoli che consentiranno di rispondere alla domanda di interesse effettuiamo alcune semplici valutazioni che seguono direttamente dalla definizione di modo. Per un modo, infatti, le equazioni di Maxwell di rotore integrate con le relazioni costitutive diventano, per la sola onda progressiva [6] (p. 12):

$$\frac{\partial H_z}{\partial y} + \gamma H_y = j\omega\epsilon E_x \quad (2.1)$$

$$-\frac{\partial H_z}{\partial x} - \gamma H_x = j\omega\epsilon E_y \quad (2.2)$$

$$\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} = j\omega\epsilon E_z \quad (2.3)$$

$$\frac{\partial E_z}{\partial y} + \gamma E_y = -j\omega\mu H_x \quad (2.4)$$

$$\frac{\partial E_z}{\partial x} + \gamma E_x = j\omega\mu H_y \quad (2.5)$$

$$\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = -j\omega\mu H_z. \quad (2.6)$$

Per l'onda regressiva il segno davanti a γ si inverte.

Osservazione 2 Il fattore comune $e^{-\gamma z}$ (per l'onda progressiva; $e^{\gamma z}$ per quella regressiva) presente in tutte le componenti cartesiane dei campi elettrico e magnetico è diverso da zero per ogni valore di $\gamma \in \mathbb{C}$ e $z \in \mathbb{R}$. Quindi può essere semplificato. Con questa semplificazione tutte le quantità coinvolte nelle equazioni (2.1)-(2.6) dipendono al più da x e y .

Assumendo $\omega \neq 0$, nelle regioni dove $n \neq 0$ si deduce che le componenti assiali del campo elettrico e magnetico (del modo) consentono di esprimere le componenti cartesiane trasversali degli stessi campi, purchè si abbia $\omega^2\mu_0\epsilon_0 n^2 + \gamma^2 \neq 0$ [6] (pp. 12-13):

$$E_x = -\frac{\gamma}{\omega^2\mu\epsilon + \gamma^2} \frac{\partial E_z}{\partial x} - \frac{j\omega\mu}{\omega^2\mu\epsilon + \gamma^2} \frac{\partial H_z}{\partial y} \quad (2.7)$$

$$E_y = -\frac{\gamma}{\omega^2\mu\epsilon + \gamma^2} \frac{\partial E_z}{\partial y} + \frac{j\omega\mu}{\omega^2\mu\epsilon + \gamma^2} \frac{\partial H_z}{\partial x} \quad (2.8)$$

$$H_x = -\frac{\gamma}{\omega^2\mu\epsilon + \gamma^2} \frac{\partial H_z}{\partial x} + \frac{j\omega\epsilon}{\omega^2\mu\epsilon + \gamma^2} \frac{\partial E_z}{\partial y} \quad (2.9)$$

$$H_y = -\frac{\gamma}{\omega^2\mu\epsilon + \gamma^2} \frac{\partial H_z}{\partial y} - \frac{j\omega\epsilon}{\omega^2\mu\epsilon + \gamma^2} \frac{\partial E_z}{\partial x}. \quad (2.10)$$

Anche in queste equazioni bisogna cambiare il segno di γ se si considera il fattore di propagazione $e^{\gamma z}$ a posto di $e^{-\gamma z}$.

Esercizio 1 Dalle equazioni (2.2) e (2.4) dedurre le equazioni (2.9) e (2.8). Analogamente dalle equazioni (2.1) e (2.5) dedurre le equazioni (2.10) e (2.7).

Soluzione. Sostituendo il primo membro dell'equazione (2.2) nell'equazione (2.4) moltiplicata per $j\omega\epsilon$ deduciamo

$$j\omega\epsilon \frac{\partial E_z}{\partial y} - \gamma \frac{\partial H_z}{\partial x} = (\omega^2\mu\epsilon + \gamma^2)H_x, \quad (2.11)$$

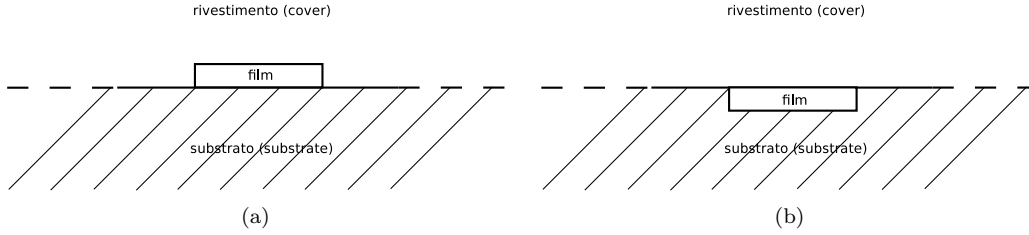


Figura 2.1: Alcuni semplici esempi di strutture guidanti dielettriche: le cosiddette “raised strip” (a) e “embedded strip” (b).

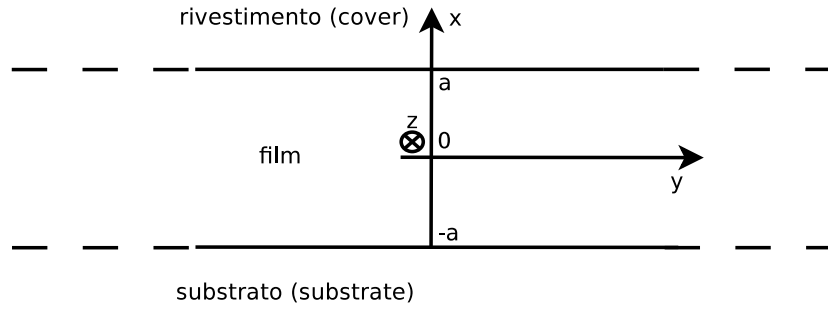


Figura 2.2: Un modello semplificato di guida dielettrica: la cosiddetta “lastra dielettrica”.

da cui segue immediatamente l'equazione (2.9) ovunque si abbia $\omega^2 \mu_0 \varepsilon_0 n^2 + \gamma^2 \neq 0$.

Sostituendo il primo membro dell'equazione (2.4) nell'equazione (2.2) moltiplicata per $-j\omega\mu$ deduciamo

$$j\omega\mu \frac{\partial H_z}{\partial x} - \gamma \frac{\partial E_z}{\partial y} = (\omega^2 \mu \varepsilon + \gamma^2) E_y, \quad (2.12)$$

da cui segue immediatamente l'equazione (2.8) ovunque si abbia $\omega^2 \mu_0 \varepsilon_0 n^2 + \gamma^2 \neq 0$.

Per le altre due equazioni si procede in modo duale.

Ciò giustifica la suddivisione dei modi in classi disgiunte ed esaustive basata sulle sole componenti assiali E_z e H_z : modi TEM ($E_z = 0$ e $H_z = 0$), modi TE ($E_z = 0$ e $H_z \neq 0$), modi TM ($H_z = 0$ e $E_z \neq 0$) e modi ibridi ($E_z \neq 0$ e $H_z \neq 0$).

Le stesse equazioni giustificano anche il fatto che, per trovare i modi, spesso ci si limita a valutare le sole componenti assiali (ciò riduce la complessità del problema in modo molto significativo: non si hanno più tre incognite (ad esempio le componenti cartesiane del campo elettrico), definite su domini tridimensionali, ma due incognite (E_z e H_z) definite sulla sezione della struttura guidante di interesse). Queste considerazioni non dovrebbero essere valide per i modi TEM.

A questo punto potremmo proseguire con i dettagli dei calcoli e si potrebbero valutare i modi delle strutture guidanti di maggior interesse nelle applicazioni: le fibre ottiche, le guide dielettriche o le fibre a cristallo fotonico. Alcuni esempi di guide dielettriche di interesse nelle applicazioni sono riportati nella figura 2.1.

Tuttavia, per capire le più importanti caratteristiche dei modi non è necessario studiare le guide dielettriche più complicate. Strutture guidanti più semplici consentono di ridurre la complessità matematica del problema, semplificando da un lato l'esposizione e mantenendo, dall'altro, una notevole generalità dei risultati ottenuti [6] (p. 305).

Per queste ragioni, consideriamo solo il seguente esempio di struttura guidante dielettrica: la cosiddetta lastra dielettrica. La sezione trasversale della lastra dielettrica e le convenzioni che utilizzeremo in questa parte del corso sono riportate in Figura 2.2

2.1 Modi di una lastra dielettrica

Nel caso della lastra dielettrica avremo in ogni caso a che fare con regioni contenenti dielettrici omogenei.

Come indicato in Figura 2.2, disponiamo il sistema cartesiano di riferimento in modo che la lastra dielettrica risulti geometricamente e dielettricamente indipendente dalla coordinata y .

Nelle situazioni di maggior interesse pratico le strutture guidanti sono in grado di “guidare” il campo nella direzione z . La lastra dielettrica non è in grado di fare la stessa cosa, visto che presenta una simmetria di rotazione attorno all’asse x . Per questa ragione se ne studia la sua capacità guidante nel solo piano (x, z) (scegliendo, implicitamente, nel piano (y, z) come direzione privilegiata l’asse z) e, così facendo, la trattazione che segue si semplifica molto. L’analisi dei modi della lastra dielettrica, quindi, si limiterà allo studio dei modi indipendenti dalla coordinata y .

Per semplificare il più possibile l’analisi senza ridurre in modo significativo la portata dei risultati, considereremo la lastra dielettrica simmetrica, assumendo cioè che il substrato e il rivestimento siano fatti con lo stesso materiale isotropo caratterizzato da ε_1 e μ_1 . Il “film”, detto anche “core”, è caratterizzato da ε_2 e μ_2 .

Con lo stesso spirito, ipotizzeremo che i mezzi siano privi di perdite e, quindi, $\varepsilon_1, \mu_1, \varepsilon_2, \mu_2 \in \mathbb{R}$.

La lastra dielettrica, nelle configurazioni più simili a quelle che vengono considerate nelle strutture guidanti trasversalmente limitate, presenta $\mu_2\varepsilon_2 > \mu_1\varepsilon_1 > 0$, $\mu_1 > 0$ e $\mu_2 > 0$. Nel seguito considereremo solo questi casi. Si potrebbe generalizzare l’analisi anche ai casi di strutture non guidanti, in cui $\mu_1\varepsilon_1 \geq \mu_2\varepsilon_2$, o di strutture composte da metamateriali ideali, in cui $\frac{\mu_2}{\mu_1} < 0$. Si evita di fare ciò per mantenere il più possibile semplice l’analisi, considerando al tempo stesso i casi più rilevanti da un punto di vista applicativo.

Modi TEM

Iniziamo con l’affermare che i modi TEM di interesse (cioè non identicamente nulli) non esistono se $\omega \neq 0$, $\mu_1\varepsilon_1 \neq \mu_2\varepsilon_2$ ed entrambe queste quantità sono non nulle. Queste condizioni sono sempre verificate in tutte le situazioni di interesse pratico.

Partendo dall’equazione (2.3) ((2.6)), sostituendo $E_z = 0 = H_z$ e considerando tutti le componenti indipendenti da y , si deduce che in ogni regione di omogeneità (con l’analisi classica le equazioni di Maxwell di rotore non possono essere ritenute valide sulle interfacce dielettriche) la componente H_y (E_y) è costante a tratti. Per la continuità delle componenti tangenziali ciò implica $H_y = C_1$ ($E_y = C_2$).

Se moltiplico l’equazione (2.5) ((2.2)) per $j\omega\varepsilon$ ($j\omega\mu$), tenendo conto che $E_z = 0$ ($H_z = 0$), ottengo $\gamma j\omega\varepsilon E_x = -\omega^2\mu\varepsilon H_y$ ($-\gamma j\omega\mu H_x = -\omega^2\mu\varepsilon E_y$). Sostituendo $j\omega\varepsilon E_x$ ($-j\omega\mu H_x$), ricavato dalla equazione (2.1) ((2.4)) tenendo conto che $H_z = 0$ ($E_z = 0$), si ricava

$$(\omega^2\mu\varepsilon + \gamma^2)H_y = 0 \quad ((\omega^2\mu\varepsilon + \gamma^2)E_y = 0). \quad (2.13)$$

I risultati ottenuti quindi implicano

$$(\omega^2\mu\varepsilon + \gamma^2)C_1 = 0 \quad ((\omega^2\mu\varepsilon + \gamma^2)C_2 = 0) \quad (2.14)$$

in ogni regione di omogeneità. A questo punto è evidente che, in presenza di una discontinuità tale per cui $\omega^2\mu_1\varepsilon_1 \neq \omega^2\mu_2\varepsilon_2$, l’equazione precedente implica $C_1 = 0$ ($C_2 = 0$) e ciò a sua volta implica $H_y = 0$ ($E_y = 0$). Se $\varepsilon \neq 0$ ($\mu \neq 0$), dall’equazione (2.1) ((2.4)) si deduce infine $E_x = 0$ ($H_x = 0$). Tenendo conto che a priori $E_z = 0 = H_z$ si conclude.

Altri modi

In ogni regione di omogeneità, sostituendo i secondi membri delle equazioni (2.9) e (2.10), sotto le ipotesi che ne garantiscono la validità, nella relazione (2.3), si deduce

$$\frac{\partial^2 E_z}{\partial x^2} + \frac{\partial^2 E_z}{\partial y^2} + (\omega^2\mu\varepsilon + \gamma^2)E_z = 0. \quad (2.15)$$

Analogamente, sostituendo le (2.7) e (2.8) nella (2.6) si ricava

$$\frac{\partial^2 H_z}{\partial x^2} + \frac{\partial^2 H_z}{\partial y^2} + (\omega^2 \mu \varepsilon + \gamma^2) H_z = 0. \quad (2.16)$$

Definiamo $k_{ci}^2 = \omega^2 \mu_i \varepsilon_i + \gamma^2$, con pedice $i = 1$ nel rivestimento o nel substrato e $i = 2$ nel film. Le ipotesi considerate sui mezzi consentono immediatamente di dedurre $k_{c1} \neq k_{c2}$ se $\omega \neq 0$.

Con le ipotesi e le notazioni introdotte, le equazioni che determinano il comportamento delle componenti assiali di un qualunque modo TE, TM o ibrido in ogni regione di omogeneità diventano:

$$\frac{d^2 E_z}{dx^2} + k_{c1}^2 E_z = 0, \quad x < -a, \quad x > a, \quad (2.17)$$

$$\frac{d^2 E_z}{dx^2} + k_{c2}^2 E_z = 0, \quad -a < x < a, \quad (2.18)$$

$$\frac{d^2 H_z}{dx^2} + k_{c1}^2 H_z = 0, \quad x < -a, \quad x > a, \quad (2.19)$$

$$\frac{d^2 H_z}{dx^2} + k_{c2}^2 H_z = 0, \quad -a < x < a. \quad (2.20)$$

Per determinare i modi di interesse, a queste equazioni vanno aggiunte le condizioni di continuità delle componenti tangenziali alle interfacce dielettriche e le condizioni al contorno. Tutti questi vincoli devono essere espressi in termini di E_z e H_z per poter ottenere una formulazione coerente (equazioni e vincoli espressi in termini di E_z e H_z). Le componenti E_z e H_z sono tangenziali alle interfacce. Quindi otteniamo i seguenti vincoli:

$$E_z|_{x=-a^-} = E_z|_{x=-a^+} \quad (2.21)$$

$$E_z|_{x=a^-} = E_z|_{x=a^+}. \quad (2.22)$$

$$H_z|_{x=-a^-} = H_z|_{x=-a^+} \quad (2.23)$$

$$H_z|_{x=a^-} = H_z|_{x=a^+}. \quad (2.24)$$

Quando il procedimento proposto ha senso, necessariamente avremo $k_{c1} \neq 0$ e $k_{c2} \neq 0$. Ciò consente di utilizzare le espressioni (2.7) - (2.10). In particolare, si deduce:

$$H_y = -\frac{j\omega\varepsilon_i}{k_{ci}^2} \frac{dE_z}{dx}, \quad (2.25)$$

$$E_y = \frac{j\omega\mu_i}{k_{ci}^2} \frac{dH_z}{dx}, \quad (2.26)$$

e, in presenza del fattore di propagazione $e^{-\gamma z}$,

$$E_x = -\frac{\gamma}{k_{ci}^2} \frac{dE_z}{dx}, \quad (2.27)$$

$$H_x = -\frac{\gamma}{k_{ci}^2} \frac{dH_z}{dx}. \quad (2.28)$$

In presenza del fattore di propagazione $e^{\gamma z}$ nelle ultime due equazioni si deve cambiare segno.

Quindi le condizioni da imporre alle interfacce sono

$$\frac{\mu_1}{k_{c1}^2} \frac{dH_z}{dx} \Big|_{x=-a^-} = \frac{\mu_2}{k_{c2}^2} \frac{dH_z}{dx} \Big|_{x=-a^+}, \quad (2.29)$$

$$\frac{\mu_2}{k_{c2}^2} \frac{dH_z}{dx} \Big|_{x=a^-} = \frac{\mu_1}{k_{c1}^2} \frac{dH_z}{dx} \Big|_{x=a^+}, \quad (2.30)$$

$$\frac{\varepsilon_1}{k_{c1}^2} \frac{dE_z}{dx} \Big|_{x=-a^-} = \frac{\varepsilon_2}{k_{c2}^2} \frac{dE_z}{dx} \Big|_{x=-a^+}, \quad (2.31)$$

$$\frac{\varepsilon_2}{k_{c2}^2} \frac{dE_z}{dx} \Big|_{x=a^-} = \frac{\varepsilon_1}{k_{c1}^2} \frac{dE_z}{dx} \Big|_{x=a^+}. \quad (2.32)$$

Si noti che i quattro precedenti vincoli garantiscono, grazie alla validità delle equazioni (2.27) e (2.28) la continuità delle componenti normali dei campi spostamento elettrico e induzione magnetica.

Per completare la formulazione rimangono da considerare le condizioni al contorno (all'infinito). Può aver senso considerare

$$\lim_{|x| \rightarrow +\infty} |E_z| \rightarrow 0, \quad (2.33)$$

$$\lim_{|x| \rightarrow +\infty} |H_z| \rightarrow 0 \quad (2.34)$$

oppure semplicemente

$$|E_z| < C, \quad C \in \mathbb{R}, \quad C > 0, \quad (2.35)$$

$$|H_z| < C, \quad C \in \mathbb{R}, \quad C > 0. \quad (2.36)$$

Non si hanno ragioni in questo modello per dire che solo uno dei due tipi di condizioni al contorno è ragionevole. Visto che le condizioni al contorno potrebbero avere un impatto significativo sulle proprietà dei modi di interesse, si ritiene utile analizzare il problema considerando entrambi i tipi di condizioni al contorno. Verificheremo che i due tipi di condizioni al contorno hanno in effetti una importanza fondamentale nel determinare proprietà significative dei modi.

A questo punto se $k_{c1} \neq 0$ e $k_{c2} \neq 0$ il problema al contorno che consente di determinare E_z e H_z nella lastra dielettrica simmetrica è completamente definito dalle equazioni (2.17), (2.18), (2.19), (2.20) e dai vincoli (2.21), (2.22), (2.23), (2.24), (2.29), (2.30), (2.31), (2.32), (2.33), (2.34), (2.35), (2.36).

Modi TE con condizione al contorno (2.34)

Tutte le equazioni, le condizioni di interfaccia e all'infinito che coinvolgono il solo E_z sono automaticamente soddisfatte. Rimangono da verificare o soddisfare le equazioni e i vincoli che coinvolgono il solo H_z . Per trovare H_z , nota la formulazione con equazioni, condizioni di interfaccia e condizioni al contorno, si procede con un procedimento standard per la ricerca dei modi. Prima si trovano gli integrali generali delle equazioni differenziali omogenee e poi si impongono i vincoli. Le espressioni degli integrali generali possono utilizzare le funzioni seno e coseno o gli esponenziali. Per ragioni che diventeranno subito evidenti esprimiamo l'integrale generale dell'equazione relativa al film con seni e coseni e quello dell'equazione relativa ai dielettrici esterni con esponenziali. In particolare,

$$H_z(x) = C_{12} \cos(k_{c2}x) + C_{22} \sin(k_{c2}x), \quad |x| < a \quad (2.37)$$

$$H_z(x) = C_{11} e^{jk_{c1}x} + C_{21} e^{-jk_{c1}x}, \quad x > a \quad (2.38)$$

$$H_z(x) = C_{13} e^{jk_{c1}x} + C_{23} e^{-jk_{c1}x}, \quad x < -a. \quad (2.39)$$

L'espressione complessiva di H_z presenta sette parametri complessi liberi (per ogni valore di frequenza fissata, oltre alle costanti C_{ij} , $i = 1, 2$, $j = 1, 2, 3$, c'è solo γ , che determina i valori k_{ci}^2 , $i = 1, 2$) e risulta essere vincolata da due condizioni al contorno (una per $x \rightarrow +\infty$ e una per $x \rightarrow -\infty$) e quattro di interfaccia. Come sempre rimarrà una delle costanti C_{ij} libera e si determineranno tutte le altre (si ricordi che un modo determina automaticamente uno spazio vettoriale di soluzioni; quindi è escluso che la costante che rimane libera possa essere γ).

In generale, per $k_{ci} \in \mathbb{C}$, si ha $jk_{ci} = -Im(k_{ci}) + jRe(k_{ci})$. Nel rivestimento ($x > a$) e nel substrato ($x < -a$) non si ottiene il decadimento richiesto se $k_{c1} \in \mathbb{R}$ (in questo caso si verifica immediatamente che sarebbe necessario avere $C_{11} = C_{21} = 0$ e $C_{13} = C_{23} = 0$ e ciò implica la nullità di H_z sia nel rivestimento che nel substrato. Le condizioni alle interfacce impongono la nullità di H_z e di $\frac{dH_z}{dx}$ sia per $x = a^-$ che per $x = -a^+$ e ciò implica $H_z = 0$ identicamente nel

film; quindi si avrebbe $H_z = 0$ ovunque e il modo non sarebbe un TE ma un modo TEM, che si studia con altre tecniche (si veda la corrispondente sottosezione)). Allora $Im(k_{c1}) \neq 0$. Per il resto, al momento non abbiamo alcun controllo su k_{c1} . Se $Im(k_{c1}) > 0$ (< 0), si ha il decadimento richiesto se e solo se $C_{21} = 0$ ($C_{11} = 0$). In entrambi i casi,

$$H_z(x) = C_1 e^{\pm j k_{c1} x}, \quad x > a \quad (2.40)$$

dove il segno $+$ si ottiene se $Im(k_{c1}) > 0$ mentre il segno $-$ si deduce quando $Im(k_{c1}) < 0$. Analogamente, per $x < -a$, si deduce $Im(k_{c1}) \neq 0$ e

$$H_z(x) = C_3 e^{\mp j k_{c1} x}, \quad x < -a \quad (2.41)$$

(con $-$ se $Im(k_{c1}) > 0$, $+$ se $Im(k_{c1}) < 0$).

Le due condizioni al contorno hanno consentito di ridurre da sette a cinque il numero di parametri complessi liberi di H_z .

Osservazione 3 Le deduzioni precedenti su k_{c1} ($Im(k_{c1}) \neq 0$) implicano banalmente $k_{c1} \neq 0$. E' possibile dedurre che anche k_{c2} è diverso da zero se $\omega \neq 0$. Procedendo per assurdo assumiamo $\omega \neq 0$ e $k_{c2} = 0$. Otteniamo $0 = k_{c2}^2 = \omega^2 \mu_2 \varepsilon_2 + \gamma^2$ da cui segue $\gamma^2 = -\omega^2 \mu_2 \varepsilon_2 \neq 0$. Inoltre, sotto l'ipotesi assurda considerata, k_{c1} risulta essere puramente immaginario poichè $k_{c1}^2 = \omega^2 \mu_1 \varepsilon_1 + \gamma^2 = \omega^2 \mu_1 \varepsilon_1 - \omega^2 \mu_2 \varepsilon_2 < 0$, visto che $\mu_2 \varepsilon_2 > \mu_1 \varepsilon_1$.

Dalle (2.11) nel film si ottiene $\frac{\partial H_z}{\partial x} = -\frac{k_{c2}^2}{\gamma} H_x$, visto che $\gamma \neq 0$. L'ipotesi che abbiamo assunto ($k_{c2} = 0$) e vogliamo mostrare essere assurda consente quindi di concludere che $H_z(x)$ è costante nel film.

Tenendo conto di ciò, della continuità di H_z alle interfacce dielettriche e delle espressioni (2.40) e (2.41) si deduce $H_z(x) = C_1 e^{\pm j k_{c1} a}$ nel film e $C_3 = C_1$.

Per la (2.6) si ottiene $\frac{\partial E_y}{\partial x} = -j\omega\mu H_z$ per tutti i campi indipendenti da y . Nel film, quindi, si deduce $\frac{\partial E_y}{\partial x} = -j\omega\mu_2 C_1 e^{\pm j k_{c1} a}$ e, quindi, $E_y = -j\omega\mu_2 C_1 e^{\pm j k_{c1} a} x + C_4$.

La componente assiale H_z determina E_y nel rivestimento e nel substrato, dato che $k_{c1} \neq 0$. Si ha $E_y = \frac{j\omega\mu_1}{k_{c1}^2} \frac{dH_z}{dx}$. Dalla (2.40) si deduce, per il rivestimento, $E_y = \frac{j\omega\mu_1}{k_{c1}^2} C_1 (\pm j k_{c1}) e^{\pm j k_{c1} x}$. Analogamente, dalla (2.41), tenendo conto del fatto che $C_3 = C_1$, per il substrato si ottiene: $E_y = \frac{j\omega\mu_1}{k_{c1}^2} C_1 (\mp j k_{c1}) e^{\mp j k_{c1} x}$.

Imponendo la continuità alle interfacce dielettriche per la componente E_y si ottiene

$$-j\omega\mu_2 C_1 e^{\pm j k_{c1} a} + C_4 = \frac{j\omega\mu_1}{k_{c1}^2} C_1 (\pm j k_{c1}) e^{\pm j k_{c1} a} \quad (2.42)$$

su $x = a$ e

$$j\omega\mu_2 C_1 e^{\pm j k_{c1} a} + C_4 = \frac{j\omega\mu_1}{k_{c1}^2} C_1 (\mp j k_{c1}) e^{\pm j k_{c1} a} \quad (2.43)$$

su $x = -a$.

Sommando termine a termine le due equazioni si deduce immediatamente che $C_4 = 0$. Le due equazioni precedenti con $C_4 = 0$ diventano identiche a meno del segno e possiamo considerare solo la (2.43) (con $C_4 = 0$).

Il coefficiente C_1 è necessariamente diverso da zero. In caso contrario, infatti, si avrebbe $H_z = 0$ ovunque e il modo non sarebbe un TE ma un TEM. Abbiamo già visto, in particolare, che nelle ipotesi considerate ($\mu_2 \varepsilon_2 > \mu_1 \varepsilon_1$) tale modo è necessariamente identicamente nullo e non è di alcun interesse.

Allora, semplificando tutti i termini non nulli, l'equazione (2.43) con $C_4 = 0$ diventa

$$\mu_2 a = \mp j \frac{\mu_1}{k_{c1}}. \quad (2.44)$$

Ma è noto che $k_{c1} = j Im(k_{c1})$. Quindi

$$\frac{\mu_2}{\mu_1} a = \mp \frac{1}{Im(k_{c1})}. \quad (2.45)$$

Osservando che il primo membro è sempre positivo mentre il secondo è sempre negativo si ottiene il risultato voluto: l'equazione (2.43) può essere soddisfatta se e solo se $C_1 = 0$ e tale modo identicamente nullo non è di alcun interesse.

Osservazione 4 A questo punto abbiamo provato che per i modi TE guidati non identicamente nulli si ha $k_{c1} \neq 0$ e $k_{c2} \neq 0$ e possiamo essere certi che il procedimento seguito, che si basa su H_z , è corretto.

Rimangono da forzare le condizioni alle interfacce dielettriche. Dai vincoli (2.23) e (2.24) si deduce

$$C_3 e^{\pm j k_{c1} a} = C_{12} \cos(k_{c2} a) - C_{22} \sin(k_{c2} a) \quad (2.46)$$

e

$$C_1 e^{\pm j k_{c1} a} = C_{12} \cos(k_{c2} a) + C_{22} \sin(k_{c2} a). \quad (2.47)$$

Queste equazioni rendono possibile la determinazione delle costanti C_1 e C_3 , ad esempio:

$$C_1 = e^{\mp j k_{c1} a} (C_{12} \cos(k_{c2} a) + C_{22} \sin(k_{c2} a)), \quad (2.48)$$

$$C_3 = e^{\mp j k_{c1} a} (C_{12} \cos(k_{c2} a) - C_{22} \sin(k_{c2} a)) \quad (2.49)$$

che, in ultima analisi, determinano H_z nel rivestimento e nel substrato a meno del parametro tuttora libero k_{c1} . Anche le loro derivate ne risultano determinate. In particolare, si ha:

$$\frac{dH_z}{dx} = C_1 (\pm j k_{c1}) e^{\pm j k_{c1} x} = e^{\mp j k_{c1} a} (C_{12} \cos(k_{c2} a) + C_{22} \sin(k_{c2} a)) (\pm j k_{c1}) e^{\pm j k_{c1} x}, \quad x > a \quad (2.50)$$

(il segno $+$ si ottiene se $\text{Im}(k_{c1}) > 0$ mentre il segno $-$ si deduce quando $\text{Im}(k_{c1}) < 0$) e

$$\frac{dH_z}{dx} = C_3 (\mp j k_{c1}) e^{\mp j k_{c1} x} = e^{\mp j k_{c1} a} (C_{12} \cos(k_{c2} a) - C_{22} \sin(k_{c2} a)) (\mp j k_{c1}) e^{\mp j k_{c1} x}, \quad x < -a \quad (2.51)$$

(con $-$ se $\text{Im}(k_{c1}) > 0$, $+$ se $\text{Im}(k_{c1}) < 0$).

Per $x = a^+$ la (2.50) diventa

$$\frac{dH_z}{dx} \Big|_{x=a^+} = (C_{12} \cos(k_{c2} a) + C_{22} \sin(k_{c2} a)) (\pm j k_{c1}), \quad (2.52)$$

con il segno $+$ se $\text{Im}(k_{c1}) > 0$ e il segno $-$ se $\text{Im}(k_{c1}) < 0$. Analogamente, per $x = -a^-$ la (2.51) diventa

$$\frac{dH_z}{dx} \Big|_{x=-a^-} = (C_{12} \cos(k_{c2} a) - C_{22} \sin(k_{c2} a)) (\mp j k_{c1}), \quad (2.53)$$

con il segno $-$ se $\text{Im}(k_{c1}) > 0$ e il segno $+$ se $\text{Im}(k_{c1}) < 0$.

Con queste deduzioni le condizioni di interfaccia (2.29) e (2.30) diventano, rispettivamente

$$(\mp j \frac{\mu_1}{k_{c1}}) (C_{12} \cos(k_{c2} a) - C_{22} \sin(k_{c2} a)) = \frac{\mu_2}{k_{c2}} (C_{12} \sin(k_{c2} a) + C_{22} \cos(k_{c2} a)) \quad (2.54)$$

$$(\pm j \frac{\mu_1}{k_{c1}}) (C_{12} \cos(k_{c2} a) + C_{22} \sin(k_{c2} a)) = \frac{\mu_2}{k_{c2}} (-C_{12} \sin(k_{c2} a) + C_{22} \cos(k_{c2} a)), \quad (2.55)$$

dove il segno più in alto (basso) nei termini con ambiguità di segno in entrambe le equazioni si applica quando $\text{Im}(k_{c1}) > 0$ ($\text{Im}(k_{c1}) < 0$). Si tenga presente che è noto che $\text{Im}(k_{c1}) \neq 0$. Riordinando i termini abbiamo dedotto

$$\begin{cases} (\mp j \frac{\mu_1}{k_{c1}} \cos(k_{c2} a) - \frac{\mu_2}{k_{c2}} \sin(k_{c2} a)) C_{12} - (\mp j \frac{\mu_1}{k_{c1}} \sin(k_{c2} a) + \frac{\mu_2}{k_{c2}} \cos(k_{c2} a)) C_{22} = 0 \\ (\pm j \frac{\mu_1}{k_{c1}} \cos(k_{c2} a) + \frac{\mu_2}{k_{c2}} \sin(k_{c2} a)) C_{12} + (\pm j \frac{\mu_1}{k_{c1}} \sin(k_{c2} a) - \frac{\mu_2}{k_{c2}} \cos(k_{c2} a)) C_{22} = 0. \end{cases} \quad (2.56)$$

Siamo interessati ad ottenere una soluzione non nulla al sistema omogeneo (2.56) nelle due indeterminate C_{12} e C_{22} (la soluzione $C_{12} = C_{22} = 0$ implicherebbe $H_z = 0$ identicamente nel mezzo

2 e, usando le equazioni (2.48), (2.49), (2.40) e (2.41), nei mezzi 1 e 3, ciò comporterebbe $H_z = 0$ ovunque e il modo non sarebbe un TE ma un modo TEM; è noto che tali modi si studiano con altre tecniche (si veda la corrispondente sottosezione)). Dobbiamo quindi forzare il determinante della matrice ad essere nullo. Ne segue un vincolo che coinvolge i soli k_{ci} , $i = 1, 2$, a cui si aggiungono le relazioni che legano i k_{ci} tra loro.

Per semplificare la notazione definiamo $A_{11} = \mp \frac{j\mu_1}{k_{c1}} \cos(k_{c2}a) - \frac{\mu_2}{k_{c2}} \sin(k_{c2}a)$ e $A_{12} = \pm \frac{j\mu_1}{k_{c1}} \sin(k_{c2}a) - \frac{\mu_2}{k_{c2}} \cos(k_{c2}a)$ e osserviamo che il sistema precedente può essere riscritto come segue

$$\begin{cases} A_{11}C_{12} + A_{12}C_{22} = 0 \\ -A_{11}C_{12} + A_{12}C_{22} = 0. \end{cases} \quad (2.57)$$

Ripristinando la notazione estesa, la condizione che impone il determinante a zero si riduce quindi a

$$\left(\mp \frac{j\mu_1}{k_{c1}} \cos(k_{c2}a) - \frac{\mu_2}{k_{c2}} \sin(k_{c2}a) \right) \cdot \left(\pm \frac{j\mu_1}{k_{c1}} \sin(k_{c2}a) - \frac{\mu_2}{k_{c2}} \cos(k_{c2}a) \right) = 0. \quad (2.58)$$

Si può subito osservare che non è possibile ottenere soluzioni per $k_{c2} \in \mathbb{C}$ che comportano $\cos(k_{c2}a) = 0$ o $\sin(k_{c2}a) = 0$. Infatti le funzioni seno e coseno presentano degli zeri solo sull'asse reale del piano complesso e gli zeri di una funzione non coincidono mai con gli zeri dell'altra e viceversa. Allora, in presenza di $\sin(k_{c2}a) = 0$ e $\cos(k_{c2}a) \neq 0$ o di $\cos(k_{c2}a) = 0$ e $\sin(k_{c2}a) \neq 0$ il determinante si annulla solo se $\mu_1\mu_2 = 0$ e ciò non è consentito dalle ipotesi.

Con i risultati ottenuti è possibile moltiplicare la parentesi a sinistra (destra) dell'equazione (2.58) per $\frac{k_{c1}k_{c2}}{\mu_1 \sin(k_{c2}a)}$ ($\frac{k_{c1}k_{c2}}{\mu_1 \cos(k_{c2}a)}$) senza alterare le soluzioni dell'equazione da risolvere (k_{c1} , k_{c2} , μ_1 , $\sin(k_{c2}a)$ e $\cos(k_{c2}a)$ sono non nulli). Così facendo ottengo

$$\left(\mp jk_{c2} \frac{1}{\tan(k_{c2}a)} - \frac{\mu_2 k_{c1}}{\mu_1} \right) \cdot \left(\pm jk_{c2} \tan(k_{c2}a) - \frac{\mu_2 k_{c1}}{\mu_1} \right) = 0. \quad (2.59)$$

Non è possibile trovare k_{c1} e k_{c2} che rendono entrambe le parentesi del primo membro dell'equazione (2.59) nulle. Infatti, se ciò accadesse si avrebbe

$$\pm jk_{c2} \frac{1}{\tan(k_{c2}a)} = -\frac{\mu_2 k_{c1}}{\mu_1} = \mp jk_{c2} \tan(k_{c2}a) \quad (2.60)$$

che implicherebbe $\cos^2(k_{c2}a) + \sin^2(k_{c2}a) = 0$ e ciò è impossibile $\forall k_{c2} \in \mathbb{C}$, come si vede utilizzando per le funzioni seno e coseno le corrispondenti espressioni basate sugli esponenziali (usando $\cos(z) = \frac{1}{2}(e^{jz} + e^{-jz})$ e $\sin(z) = \frac{1}{2j}(e^{jz} - e^{-jz})$ si deduce $\cos^2(z) = \frac{1}{4}(e^{2jz} + e^{-2jz} + 2)$ e $\sin^2(z) = -\frac{1}{4}(e^{2jz} + e^{-2jz} - 2)$ e quindi $\cos^2(z) + \sin^2(z) = 1$, $\forall z \in \mathbb{C}$).

Allora, fissata ω bisogna cercare γ che risolve l'una o l'altra delle equazioni di dispersione seguenti:

$$\begin{cases} \pm k_{c2} \frac{1}{\tan(k_{c2}a)} = j \frac{\mu_2 k_{c1}}{\mu_1} \\ k_{c1}^2 = \omega^2(\mu_1 \varepsilon_1 - \mu_2 \varepsilon_2) + k_{c2}^2, \end{cases} \quad (2.61)$$

$$\begin{cases} \mp k_{c2} \tan(k_{c2}a) = j \frac{\mu_2 k_{c1}}{\mu_1} \\ k_{c1}^2 = \omega^2(\mu_1 \varepsilon_1 - \mu_2 \varepsilon_2) + k_{c2}^2, \end{cases} \quad (2.62)$$

spesso denominate relazioni o equazioni caratteristiche o di dispersione per i modi TE guidati e indipendenti da y dello slab dielettrico simmetrico. Si ricorda che l'ambiguità di segno ancora presente nelle equazioni caratteristiche deve essere gestita utilizzando il segno più in alto (basso) quando $Im(k_{c1}) > 0$ ($Im(k_{c1}) < 0$).

Dalle relazioni (2.61)₂ e (2.62)₂ in entrambi i casi è possibile ricavare k_{c1} per dedurre una unica equazione nella sola indeterminata k_{c2} . È evidente che così facendo si lascia indeterminato il segno di k_{c1} e, quindi, il segno della sua parte immaginaria. Entrambi i segni di k_{c1} sono legittimi ma si ricordi che, nelle equazioni (2.61)₁ e (2.62)₁, è necessario, per ogni scelta di tale segno, utilizzare l'alternativa \pm richiesta dal procedimento.

Osservando che $jk_{c1} = -Im(k_{c1}) + jRe(k_{c1}) = \sqrt{(jk_{c1})^2} = \sqrt{-k_{c1}^2} = \sqrt{\omega^2(\mu_2 \varepsilon_2 - \mu_1 \varepsilon_1) - k_{c2}^2}$, è possibile dedurre che se si prende la radice con parte reale positiva, come si fa di solito, si sceglie

implicitamente $Im(k_{c1})$ negativo. Per questa ragione, con la convenzione indicata sull'operazione di estrazione della radice, dobbiamo scegliere il segno in basso nelle due equazioni di dispersione. Così facendo si ottiene:

$$k_{c2}a \frac{1}{\tan(k_{c2}a)} = -\frac{\mu_2}{\mu_1} \sqrt{\omega^2 a^2 (\mu_2 \varepsilon_2 - \mu_1 \varepsilon_1) - k_{c2}^2 a^2}, \quad (2.63)$$

$$k_{c2}a \tan(k_{c2}a) = \frac{\mu_2}{\mu_1} \sqrt{\omega^2 a^2 (\mu_2 \varepsilon_2 - \mu_1 \varepsilon_1) - k_{c2}^2 a^2}. \quad (2.64)$$

Osservazione 5 *Da questo momento in avanti non abbiamo più una ambiguità di segno ma dobbiamo ricordare la convenzione sull'operazione radice!*

Si tratta di equazioni nella indeterminata complessa $k_{c2}a$, una volta stabilito il valore di ω (e la struttura guidante di interesse). Purtroppo non è possibile trovare una forma esplicita di γ in funzione di ω . Tale funzione può essere trovata solo con l'ausilio di un calcolatore.

Nonostante le difficoltà indicate, alcune proprietà del numero d'onda trasversale k_{c2} possono essere dedotte immediatamente. Si consideri, ad esempio, la possibilità di ottenere soluzioni per $k_{c2}^2 \in (-\infty, 0)$. In questo caso si avrebbe k_{c2} immaginario puro. Quindi $k_{c2} = jIm(k_{c2})$ con $Im(k_{c2}) \neq 0$. In queste condizioni

$$\tan(k_{c2}a) = \frac{\sin(k_{c2}a)}{\cos(k_{c2}a)} = \frac{e^{jk_{c2}a} - e^{-jk_{c2}a}}{2j} \frac{2}{e^{jk_{c2}a} + e^{-jk_{c2}a}} = -j \frac{e^{-Im(k_{c2})a} - e^{Im(k_{c2})a}}{e^{-Im(k_{c2})a} + e^{Im(k_{c2})a}}. \quad (2.65)$$

Ne segue che

$$k_{c2}a \tan(k_{c2}a) = jIm(k_{c2})a \left(-j \frac{e^{-Im(k_{c2})a} - e^{Im(k_{c2})a}}{e^{-Im(k_{c2})a} + e^{Im(k_{c2})a}} \right) = Im(k_{c2})a \frac{e^{-Im(k_{c2})a} - e^{Im(k_{c2})a}}{e^{-Im(k_{c2})a} + e^{Im(k_{c2})a}} \quad (2.66)$$

e

$$\frac{k_{c2}a}{\tan(k_{c2}a)} = jIm(k_{c2})a \frac{1}{\left(-j \frac{e^{-Im(k_{c2})a} - e^{Im(k_{c2})a}}{e^{-Im(k_{c2})a} + e^{Im(k_{c2})a}} \right)} = -Im(k_{c2})a \frac{e^{-Im(k_{c2})a} + e^{Im(k_{c2})a}}{e^{-Im(k_{c2})a} - e^{Im(k_{c2})a}}. \quad (2.67)$$

Per ogni valore reale di $Im(k_{c2})a$ la funzione riportata nella equazione (2.66) è sempre negativa mentre quella riportata nell'equazione (2.67) è sempre positiva. Nelle equazioni di dispersione la funzione riportata nella equazione (2.66) va eguagliata con la funzione $\frac{\mu_2}{\mu_1} \sqrt{\omega^2 a^2 (\mu_2 \varepsilon_2 - \mu_1 \varepsilon_1) - k_{c2}^2 a^2} = \frac{\mu_2}{\mu_1} \sqrt{\omega^2 a^2 (\mu_2 \varepsilon_2 - \mu_1 \varepsilon_1) + (Im(k_{c2}))^2 a^2}$. Per le ipotesi riportate all'inizio ($\frac{\mu_2}{\mu_1} > 0$ e $\mu_2 \varepsilon_2 - \mu_1 \varepsilon_1 > 0$) il secondo membro dell'equazione di dispersione (2.64) è strettamente positivo. Quindi l'equazione di dispersione (2.64) non ammette soluzioni per $k_{c2}^2 \in (-\infty, 0)$ nei casi indicati.

Per le stesse ragioni, tenendo conto dei cambi di segno di entrambi i membri, anche l'equazione di dispersione (2.63) non ammette soluzioni per $k_{c2}^2 \in (-\infty, 0)$ negli stessi casi.

Un altro set di valori di k_{c2}^2 che può essere gestito con semplicità è dato dall'insieme $Im(k_{c2}^2) \neq 0$. Visto che $\omega^2 \mu_2 \varepsilon_2 \in \mathbb{R}$ si deduce immediatamente che $Im(k_{c2}^2) \neq 0$ implica $Im(\gamma^2) \neq 0$. Ma $Im(\gamma^2) \neq 0$ implica α e β entrambi non nulli (se uno dei due fosse nullo avremmo $\gamma^2 \in \mathbb{R}$). Dalla Osservazione 1, deduciamo che $\beta \geq 0$ e $\alpha \geq 0$. Mettendo insieme le ultime considerazioni deduciamo, in particolare, $\beta > 0$ e $\alpha > 0$.

Quindi, in particolare

$$Im(k_{c2}^2) \neq 0 \implies \alpha > 0. \quad (2.68)$$

Allora il modo, propagandosi in una delle due direzioni lungo l'asse z , presenta necessariamente un decadimento esponenziale. Ma i modi guidati considerati decadono esponenzialmente lungo l'asse x e sono indipendenti dalla coordinata y . Queste proprietà sono incompatibili con il principio di conservazione dell'energia, per uno slab dielettrico senza perdite. Faremo una verifica esplicita di ciò con l'esercizio 2.

A questo punto abbiamo verificato che non è possibile considerare $k_{c2}^2 \in (-\infty, 0)$ e $Im(k_{c2}^2) \neq 0$ (in aggiunta a $k_{c2}^2 = 0$). Ne segue che $k_{c2}^2 \in (0, \infty)$, da cui segue $k_{c2} \in \mathbb{R}$, $k_{c2} \neq 0$, $\gamma^2 \in \mathbb{R}$ e

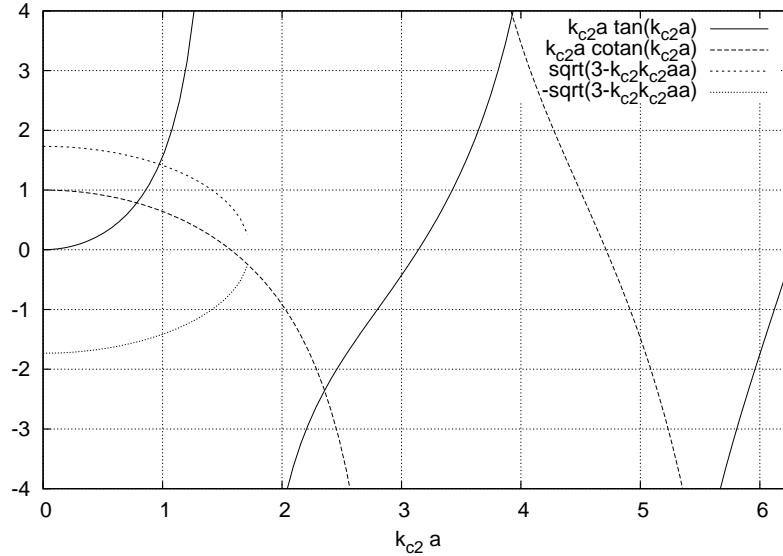


Figura 2.3: Andamenti tipici dei termini coinvolti nelle equazioni di dispersione per i modi TE guidati indipendenti da y pari e dispari di uno slab dielettrico in presenza di dielettrici privi di perdite, al variare di $k_{c2}a \in \mathbb{R}$. I grafici sono ottenuti per $\mu_2 = \mu_1$ e ω tale che $\omega^2 a^2 (\mu_2 \varepsilon_2 - \mu_1 \varepsilon_1) = 3$.

$k_{c1}^2 \in \mathbb{R}$. Quest'ultima deduzione implica $Im(k_{c1}) = 0$ se $k_{c1}^2 \in [0, \infty)$ ed è noto che tali valori non consentono di soddisfare la condizione al contorno (2.34). Inoltre $\gamma^2 \in \mathbb{R}$ può verificarsi se $\gamma^2 \geq 0$ (ovvero $\gamma = \alpha$ e $\beta = 0$), e si tornerebbe ad avere k_{c1}^2 reale e positivo, o se $\gamma^2 < 0$. Solo in quest'ultimo caso, che comporta $\gamma = j\beta$ e $\alpha = 0$, è possibile soddisfare il vincolo relativo a $Im(k_{c1})$ non nullo. In particolare, è necessario avere $\omega^2 \mu_1 \varepsilon_1 < \beta^2 < \omega^2 \mu_2 \varepsilon_2$ (il $<$ a sinistra (destra) segue dalle proprietà di k_{c1} (k_{c2}); si noti che la disuguaglianza precedente può essere verificata solo se una delle ipotesi riportate all'inizio ($\mu_1 \varepsilon_1 < \mu_2 \varepsilon_2$) è verificata).

Per valori reali non nulli di k_{c2} le equazioni di dispersione possono essere risolte graficamente. In figura 2.3 vengono riportati gli andamenti tipici dei termini coinvolti nelle due equazioni di dispersione, al variare di $k_{c2}a \in \mathbb{R}$. Tenendo presente che $k_{c2}a \tan(k_{c2}a)$, $k_{c2}a \frac{1}{\tan(k_{c2}a)}$ e $\sqrt{\omega^2 a^2 (\mu_2 \varepsilon_2 - \mu_1 \varepsilon_1) - k_{c2}^2 a^2}$ sono funzioni pari di $k_{c2}a$ possiamo limitare l'analisi ai soli valori positivi di $k_{c2}a$. A questo proposito, si tenga presente che, nota una soluzione con $k_{c2} > 0$ il modo corrispondente alla soluzione ottenuta per $-k_{c2}$ non fornisce un modo indipendente da quello che si deduce con la soluzione di partenza (k_{c2}^2 è ovviamente lo stesso e, quindi, γ^2 è lo stesso e possiamo ottenere lo stesso γ ; vedremo che anche le espressioni delle componenti dei campi sono inalterate o cambiano segno, rimanendo chiaramente all'interno dello stesso spazio vettoriale).

Ciascuna delle intersezioni dei grafici indicati fornisce un valore di $k_{c2}a$ e, quindi, di k_{c2} che soddisfa la corrispondente equazione di dispersione. Da ciò si deducono i valori di γ e k_{c1} che determinano il corrispondente modo TE guidato, indipendente da y , dello slab dielettrico simmetrico e senza perdite. Si tenga presente che la radice quadrata presente nel membro di destra delle equazioni di dispersione (2.63) e (2.64) fornisce valori reali solo per $|k_{c2}a| < \omega a \sqrt{\mu_2 \varepsilon_2 - \mu_1 \varepsilon_1} = \frac{\omega a}{c_0} \sqrt{n_2^2 - n_1^2}$ ($c_0 = \frac{1}{\sqrt{\mu_0 \varepsilon_0}}$ è la velocità della luce nel vuoto). L'analisi grafica può quindi essere limitata al range di $k_{c2}a$ che va da 0 a $\frac{\omega a}{c_0} \sqrt{n_2^2 - n_1^2}$. E' evidente che, in ogni caso, per ogni valore di ω fissato, è possibile ottenere solo un numero finito di modi guidati.

A questo punto abbiamo capito che, fissato ω , siamo in grado di trovare γ che soddisfa una e una sola delle equazioni di dispersione. Se l'equazione (2.61) è soddisfatta (e non è soddisfatta l'equazione (2.62)) ci siamo posti nella condizione di avere $A_{11} = 0$ e $A_{12} \neq 0$ nel sistema (2.57). Quindi la coppia di coefficienti (C_{12}, C_{22}) che individuano, insieme a γ , completamente il modo

è data da ($C_{12}, C_{22} = 0$), con C_{12} libero di variare (è noto che i modi determinano uno spazio vettoriale di campi).

Viceversa, se, fissato ω , siamo in grado di trovare γ che soddisfa l'equazione (2.62) (e non soddisfa l'equazione (2.61)) allora ci siamo posti nella condizione di avere $A_{12} = 0$ e $A_{11} \neq 0$ nel sistema (2.57) e quindi la coppia di coefficienti (C_{12}, C_{22}) che individuano, insieme a γ , completamente il modo è data da ($C_{12} = 0, C_{22}$), con C_{22} libero di variare.

Osservazione 6 *Le deduzioni precedenti su A_{11} , A_{12} , C_{12} , C_{22} sono riconducibili alle considerazioni riportate tra l'equazione (2.57) e l'equazione (2.61) e non sono in alcun modo influenzate dalle restrizioni dedotte su k_{c2}^2 a valle della Osservazione 5.*

Nel primo dei due casi considerati si ottiene, utilizzando le relazioni (2.37), (2.48), (2.40), (2.49), (2.41), (2.50), (2.26), (2.51), (2.28),

$$H_z(x) = C_{12} \cos(k_{c2}x), \quad |x| < a \quad (2.69)$$

$$H_z(x) = e^{jk_{c1}a} C_{12} \cos(k_{c2}a) e^{-jk_{c1}x}, \quad x > a \quad (2.70)$$

$$H_z(x) = e^{jk_{c1}a} C_{12} \cos(k_{c2}a) e^{jk_{c1}x}, \quad x < -a \quad (2.71)$$

$$H_x(x) = \frac{\gamma}{k_{c2}} C_{12} \sin(k_{c2}x), \quad |x| < a \quad (2.72)$$

$$H_x(x) = \frac{j\gamma}{k_{c1}} e^{jk_{c1}a} C_{12} \cos(k_{c2}a) e^{-jk_{c1}x}, \quad x > a \quad (2.73)$$

$$H_x(x) = -\frac{j\gamma}{k_{c1}} e^{jk_{c1}a} C_{12} \cos(k_{c2}a) e^{jk_{c1}x}, \quad x < -a \quad (2.74)$$

$$E_y(x) = -\frac{j\omega\mu_2}{k_{c2}} C_{12} \sin(k_{c2}x), \quad |x| < a \quad (2.75)$$

$$E_y(x) = \frac{\omega\mu_1}{k_{c1}} e^{jk_{c1}a} C_{12} \cos(k_{c2}a) e^{-jk_{c1}x}, \quad x > a \quad (2.76)$$

$$E_y(x) = -\frac{\omega\mu_1}{k_{c1}} e^{jk_{c1}a} C_{12} \cos(k_{c2}a) e^{jk_{c1}x}, \quad x < -a. \quad (2.77)$$

Le espressioni precedenti sono state dedotte tenendo anche conto del fatto che, con le convenzioni utilizzate, dobbiamo scegliere il segno in basso nelle espressioni indicate. Inoltre si è fatto riferimento alle onde che presentano il fattore di propagazione $e^{-\gamma z}$. Si ricorda che se si intende fare riferimento alle onde che presentano il fattore di propagazione $e^{\gamma z}$, è necessario cambiare il segno di γ e, quindi, di H_x .

Nel secondo dei due casi considerati con le stesse relazioni e convenzioni si ottiene

$$H_z(x) = C_{22} \sin(k_{c2}x), \quad |x| < a \quad (2.78)$$

$$H_z(x) = e^{jk_{c1}a} C_{22} \sin(k_{c2}a) e^{-jk_{c1}x}, \quad x > a \quad (2.79)$$

$$H_z(x) = -e^{jk_{c1}a} C_{22} \sin(k_{c2}a) e^{jk_{c1}x}, \quad x < -a \quad (2.80)$$

$$H_x(x) = -\frac{\gamma}{k_{c2}} C_{22} \cos(k_{c2}x), \quad |x| < a \quad (2.81)$$

$$H_x(x) = \frac{j\gamma}{k_{c1}} e^{jk_{c1}a} C_{22} \sin(k_{c2}a) e^{-jk_{c1}x}, \quad x > a \quad (2.82)$$

$$H_x(x) = \frac{j\gamma}{k_{c1}} e^{jk_{c1}a} C_{22} \sin(k_{c2}a) e^{jk_{c1}x}, \quad x < -a \quad (2.83)$$

$$E_y(x) = \frac{j\omega\mu_2}{k_{c2}} C_{22} \cos(k_{c2}x), \quad |x| < a \quad (2.84)$$

$$E_y(x) = \frac{\omega\mu_1}{k_{c1}} e^{jk_{c1}a} C_{22} \sin(k_{c2}a) e^{-jk_{c1}x}, \quad x > a \quad (2.85)$$

$$E_y(x) = \frac{\omega\mu_1}{k_{c1}} e^{jk_{c1}a} C_{22} \sin(k_{c2}a) e^{jk_{c1}x}, \quad x < -a. \quad (2.86)$$

Si tenga presente che tutti i modi TE indipendenti dalla coordinata y presentano $E_x = 0 = H_y$.

Anche se la terminologia non è univoca [7] (p. 264), i modi della prima (seconda) classe vengono spesso denominati modi dispari (pari), dando più importanza alla parità delle componenti trasversali (la parità di quelle assiali è opposta).

Osservazione 7 Tenendo conto dell'Osservazione 6 e del fatto che le espressioni delle componenti di campo sono state dedotte utilizzando le relazioni (2.37), (2.48), (2.40), (2.49), (2.41), (2.50), (2.26), (2.51), (2.28), che risultano essere corrette indipendentemente dal valore assunto da k_{c2} , si deduce che le espressioni (2.69)- (2.77) e (2.78)- (2.86) sono valide per le componenti di campo dei modi TE qui considerati anche se k_{c2} fosse tale da determinare $k_{c2}^2 \in (-\infty, 0)$ o $\text{Im}(k_{c2}^2) \neq 0$.

Esercizio 2 Considerando le espressioni (2.81) - (2.86), calcolare la potenza attiva trasmessa da un modo TE guidato, pari e progressivo attraverso la striscia $(x, y, z = \bar{z})$, $x \in (-\infty, \infty)$, $y \in (0, 1)$.

Soluzione. La potenza attiva trasmessa da un modo TE pari o dispari attraverso la striscia indicata è data da (si tenga presente che le componenti trasversali sono entrambe pari o entrambe dispari e il loro prodotto è sempre pari)

$$\begin{aligned} P(\bar{z}) &= \frac{1}{2} \int_0^1 dy \int_{-\infty}^{\infty} \text{Re}(\mathbf{E}(x) e^{-\gamma \bar{z}} \times \mathbf{H}^*(x) (e^{-\gamma \bar{z}})^* \cdot \hat{\mathbf{z}}) dx = \\ &= \frac{1}{2} [(\hat{\mathbf{y}} \times \hat{\mathbf{x}}) \cdot \hat{\mathbf{z}}] e^{-2\alpha \bar{z}} \int_{-\infty}^{\infty} \text{Re}(E_y(x) H_x^*(x)) dx = -\frac{1}{2} e^{-2\alpha \bar{z}} \int_{-\infty}^{\infty} \text{Re}(E_y(x) H_x^*(x)) dx = \\ &= -e^{-2\alpha \bar{z}} \int_0^{\infty} \text{Re}(E_y(x) H_x^*(x)) dx = \\ &= -e^{-2\alpha \bar{z}} \int_0^a \text{Re}(E_y(x) H_x^*(x)) dx - e^{-2\alpha \bar{z}} \int_a^{\infty} \text{Re}(E_y(x) H_x^*(x)) dx. \end{aligned} \quad (2.87)$$

Per quel che riguarda il primo dei due integrali presenti nell'ultimo membro, considerando, come indicato nel testo dell'esercizio, le espressioni delle componenti per i modi pari,

$$\begin{aligned} \int_0^a \text{Re}(E_y(x) H_x^*(x)) dx &= \\ &= \int_0^a \text{Re}\left(\frac{j\omega\mu_2}{k_{c2}} C_{22} \cos(k_{c2}x) (-1) \frac{\gamma^*}{k_{c2}^*} C_{22}^* (\cos(k_{c2}x))^*\right) dx \\ &= \int_0^a \text{Re}\left(\frac{-j\omega\mu_2\gamma^* |C_{22}|^2}{|k_{c2}|^2} |\cos(k_{c2}x)|^2\right) dx. \end{aligned} \quad (2.88)$$

Considerando che $\mu_2 \in \mathbb{R}$ e che $\text{Re}(j\gamma^*) = \beta$, si deduce

$$\int_0^a \text{Re}(E_y(x) H_x^*(x)) dx = -\frac{\omega\mu_2\beta |C_{22}|^2}{|k_{c2}|^2} \int_0^a |\cos(k_{c2}x)|^2 dx. \quad (2.89)$$

Il secondo integrale di interesse è pari a

$$\begin{aligned} \int_a^{\infty} \text{Re}(E_y(x) H_x^*(x)) dx &= \\ &= \int_a^{\infty} \text{Re}\left(\frac{\omega\mu_1}{k_{c1}} e^{jk_{c1}a} C_{22} \sin(k_{c2}a) e^{-jk_{c1}x} \frac{j^* \gamma^*}{k_{c1}^*} (e^{jk_{c1}a})^* C_{22}^* (\sin(k_{c2}a))^* (e^{-jk_{c1}x})^*\right) dx. \end{aligned} \quad (2.90)$$

Considerando che $\mu_1 \in \mathbb{R}$, $e^{jk_{c1}a} (e^{jk_{c1}a})^* = e^{-2\text{Im}(k_{c1})a}$, $e^{-jk_{c1}x} (e^{-jk_{c1}x})^* = e^{2\text{Im}(k_{c1})x}$ e che $\text{Re}(j^* \gamma^*) = -\beta$, si deduce

$$\begin{aligned} \int_a^{\infty} \text{Re}(E_y(x) H_x^*(x)) dx &= \\ &= -\frac{\omega\mu_1\beta |C_{22}|^2 |\sin(k_{c2}a)|^2 e^{-2\text{Im}(k_{c1})a}}{|k_{c1}|^2} \int_a^{\infty} e^{2\text{Im}(k_{c1})x} dx. \end{aligned} \quad (2.91)$$

Infine, dato che secondo le nostre convenzioni $\text{Im}(k_{c1}) < 0$, $\int_a^\infty e^{2\text{Im}(k_{c1})x} dx = -\frac{1}{2\text{Im}(k_{c1})} e^{2\text{Im}(k_{c1})a}$ si ottiene

$$\int_a^\infty \text{Re}(E_y(x)H_x^*(x)) dx = \frac{\omega\mu_1\beta|C_{22}|^2|\sin(k_{c2}a)|^2}{2\text{Im}(k_{c1})|k_{c1}|^2}. \quad (2.92)$$

Entrambi gli integrali calcolati sono negativi. Ne segue che

$$P(\bar{z}) = \omega\beta|C_{22}|^2 e^{-2\alpha\bar{z}} \left(\frac{\mu_2}{|k_{c2}|^2} \int_0^a |\cos(k_{c2}x)|^2 dx - \frac{\mu_1|\sin(k_{c2}a)|^2}{2\text{Im}(k_{c1})|k_{c1}|^2} \right). \quad (2.93)$$

A questo punto, ricordando che $\omega > 0$ per ipotesi, $\beta > 0$ (verificato in precedenza sia per $k_{c2}^2 > 0$ che per $\text{Im}(k_{c2}^2) \neq 0$; inoltre se $k_{c2}^2 < 0$ si ottiene $\gamma^2 < -\omega^2\mu_2\varepsilon_2 < 0$ e, quindi, nuovamente $\beta \neq 0$, da cui, per l'osservazione 1, $\beta > 0$), $|C_{22}|^2 > 0$ (altrimenti avremmo a che fare con un modo identicamente nullo che non si studia), $\sin(k_{c2}a) \neq 0$, $\text{Im}(k_{c1}) < 0$ (e quindi $|k_{c1}| \neq 0$), ottengo un valore di $P(\bar{z})$ strettamente positivo se $\mu_i > 0$, $i = 1, 2$. Evitiamo, per mantenere semplice l'analisi, considerando al tempo stesso il caso più interessante per le applicazioni, di indagare quello che succede se l'ultima ipotesi non è verificata.

Osservazione 8 Potremmo dedurre un risultato analogo per i modi TE guidati, dispari, progressivi.

Osservazione 9 Si possono dedurre risultati analoghi per i modi TE guidati, pari o dispari, regressivi pur di valutare il flusso del vettore di Poynting in direzione $-\hat{z}$ (il segno di γ fa cambiare il segno dei due integrali ma ciò è compensato dal cambio di segno della normale). Ciò è coerente con l'interpretazione fisica della trasmissione di potenza attiva nella direzione $-\hat{z}$.

Osservazione 10 L'esercizio precedente mette in evidenza che la potenza attiva trasmessa da un modo è determinata dalle sole componenti trasversali del campo del modo stesso.

Osservazione 11 Come indicato in precedenza, l'esercizio 2 mette in evidenza che $P(z)$ decade esponenzialmente a mano a mano che il modo si propaga (in una delle due direzioni, a seconda che si consideri l'onda progressiva o quella regressiva) se $\alpha > 0$. Una tale riduzione di potenza attiva non è compatibile con il principio di conservazione dell'energia, valido anche per fenomeni elettromagnetici, visto che la potenza attiva che entra dalla striscia posta da uno dei due lati (quello di sinistra, su $z = z_1$, se l'onda è progressiva; quello di destra, su $z = z_2$, se l'onda è regressiva) non esce dalla striscia posta dall'altro lato, non è stata dissipata nella struttura guidante (mezzi senza perdite), non è irradiata in direzione x (si può trovare un \bar{x} sufficientemente grande che determina un flusso del vettore di Poynting attraverso la striscia $x = \bar{x}$, di larghezza pari alla distanza $|z_2 - z_1|$ in direzione z e di un metro lungo l'asse y , più piccolo di qualunque valore positivo) e non può uscire da ogni superficie a y costante, visto che il versore normale a tali superfici è proporzionale a \hat{y} e i modi TE considerati presentano un campo elettrico sempre polarizzato lungo y .

Allora per tali modi $\alpha > 0$ è ipotesi assurda. Quindi, per l'implicazione (2.68) $\text{Im}(k_{c2}^2) \neq 0$ è ipotesi assurda.

A questo punto l'analisi di interesse può considerarsi completa da un punto di vista analitico. Da un punto di vista pratico, tuttavia, può essere opportuno mettere in evidenza i seguenti aspetti:

- dati $\mathbb{R} \ni \omega, a, \mu_i, \varepsilon_i, i = 1, 2$ si trovi $m \in \mathbb{N}, m \geq 0$:

$$(2m-1)\frac{\pi}{2} \leq \omega a \sqrt{\mu_2\varepsilon_2 - \mu_1\varepsilon_1} < (2m+1)\frac{\pi}{2} \quad \left(m\pi \leq \omega a \sqrt{\mu_2\varepsilon_2 - \mu_1\varepsilon_1} < (m+1)\pi \right).$$

Le soluzioni distinte dell'equazione di dispersione (2.63) ((2.64)) sono in numero pari a $m(m+1)$, $m \in \mathbb{N}, m \geq 0$. Ciascuna di tali $m(m+1)$ soluzioni corrisponde a un modo TE guidato, dispari (pari), indipendente da y , dello slab dielettrico simmetrico e senza perdite; il corrispondente modo viene indicato con il simbolo $\text{TE}_n^{(o)}$ ($\text{TE}_n^{(e)}$) (l'apice “o” (“e”) sta per “odd” (“even”)), $n \in \mathbb{N}, m \geq n \geq 1$ ($m \geq n \geq 0$).

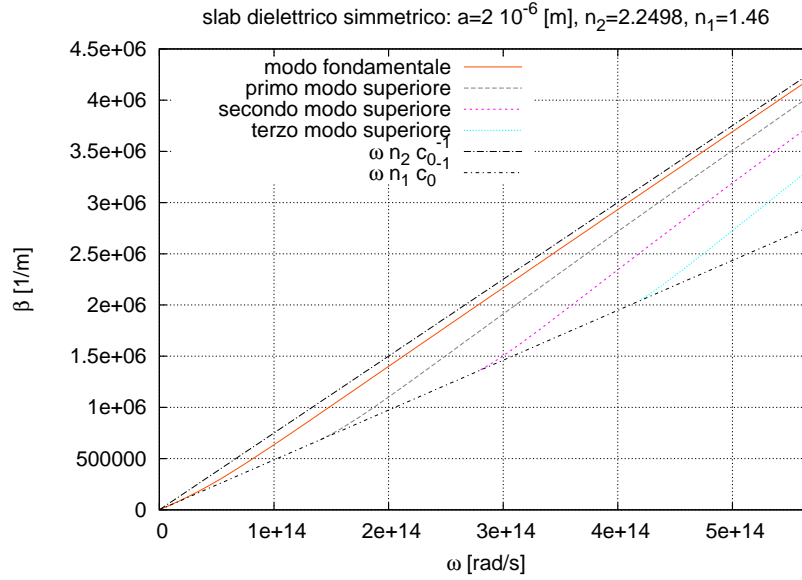


Figura 2.4: Andamento di $\beta(\omega)$ per i primi modi TE guidati indipendenti da y di uno slab simmetrico caratterizzato da $a = 2$ micron, $n_1 = 1.46$, $n_2 = 2.2498$, $\mu_2 = \mu_1$.

- per ogni modo $\text{TE}_n^{(o)}$ ($\text{TE}_n^{(e)}$) esiste $\omega_n \in \mathbb{R}$,

$$\omega_n = \frac{(2n-1)\pi}{2a\sqrt{n_2^2 - n_1^2}}c_0, \quad n \in \mathbb{N}, n \geq 1 \quad (\omega_n = \frac{n\pi}{a\sqrt{n_2^2 - n_1^2}}c_0, \quad n \in \mathbb{N}, n \geq 0) \quad (2.94)$$

detta pulsazione di taglio (del modo considerato). Il corrispondente modo esiste ed è ben definito $\forall \omega > \omega_n$. Non esiste $\forall \omega \leq \omega_n$ (a partire dalla frequenza di taglio e al di sotto di questa i modi guidati perdono la loro individualità (svaniscono); in particolare, per $\omega = \omega_n$ si ottiene $k_{c1}^2 = 0$ e il modo corrispondente non può più essere considerato un modo guidato (si veda l'Osservazione 4)). Si noti che il modo pari $\text{TE}_0^{(e)}$ esiste $\forall \omega > 0$ ($\omega_0 = 0$) ed è detto modo fondamentale o dominante. Gli altri modi sono detti modi superiori.

Ad esempio, considerando $a = 2$ micron, $n_1 = 1.46$, $n_2 = 2.2498$, $\mu_2 = \mu_1$, si ottiene $\omega_1 = 137.555 \cdot 10^{12}$ [rad/s] ($f_1 = 21.8926$ THz) per il modo guidato $\text{TE}_1^{(o)}$. La frequenza di taglio raddoppia per il $\text{TE}_1^{(e)}$, triplica per il $\text{TE}_2^{(o)}$, quadruplica per il $\text{TE}_2^{(e)}$ e così via. Gli stessi valori si ottengono se si cambia struttura guidante mantenendo costante $a\sqrt{n_2^2 - n_1^2}$. Ad esempio, gli stessi valori di frequenze di taglio di ottengono per $a = 4$ micron, $n_1 = 1.46$, $n_2 = 1.692365$, $\mu_2 = \mu_1$ o per $a = 20$ micron, $n_1 = 1.46$, $n_2 = 1.47$, $\mu_2 = \mu_1$.

- con un calcolatore è possibile graficare l'andamento di $\beta(\omega)$ per ogni modo $\text{TE}_n^{(o)}$ ($\text{TE}_n^{(e)}$) $n \in \mathbb{N}$, $n \geq 1$ ($n \geq 0$).

Ad esempio, considerando $a = 2$ micron, $n_1 = 1.46$, $n_2 = 2.2498$, $\mu_2 = \mu_1$, si ottiene il diagramma di dispersione rappresentato in figura 2.4 (assumendo i due mezzi non dispersivi nel tempo). Le figure 2.5 e 2.6 riportano i corrispondenti andamenti grafici quando le dimensioni geometriche e l'indice di rifrazione del film sono modificati in modo da non alterare il prodotto $a\sqrt{n_2^2 - n_1^2}$ rispetto al caso considerato in figura 2.4.

Si tenga presente che il cosiddetto “fractional refractive-index change”, definito [1] (p. 327) come $\Delta = \frac{n_2^2 - n_1^2}{2n_2^2} = \frac{(n_2 - n_1)(n_2 + n_1)}{2n_2^2} \simeq \frac{(n_2 - n_1)2n_2}{2n_2^2} \simeq \frac{n_2 - n_1}{n_2}$ risulta essere molto piccolo, nell'intervallo $[0.001, 0.02]$, per i tipici materiali utilizzati per realizzare fibre ottiche (fused silica glass (SiO_2), con indici di rifrazione $\simeq 1.45$). Da questo punto di vista il diagramma riportato in figura 2.6 è quello più vicino a un caso realistico.

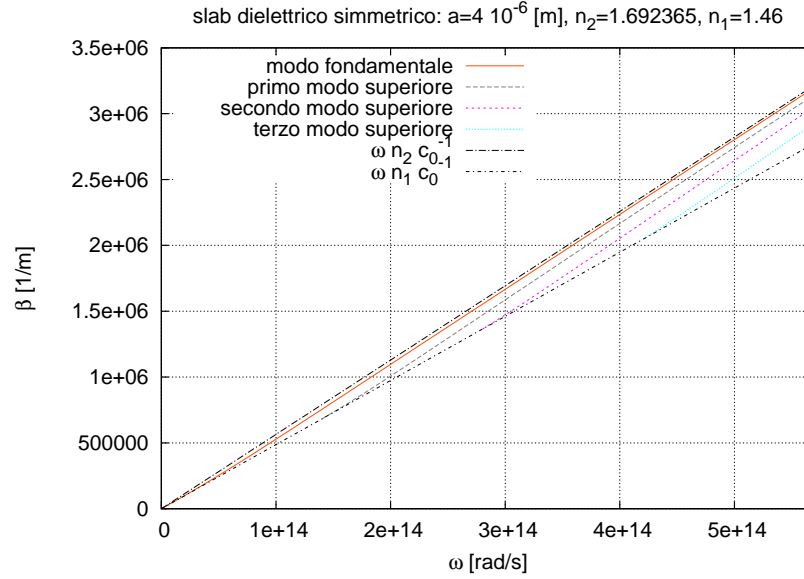


Figura 2.5: Andamento di $\beta(\omega)$ per i primi modi TE guidati indipendenti da y di uno slab simmetrico caratterizzato da $a = 4$ micron, $n_1 = 1.46$, $n_2 = 1.692365$, $\mu_2 = \mu_1$.

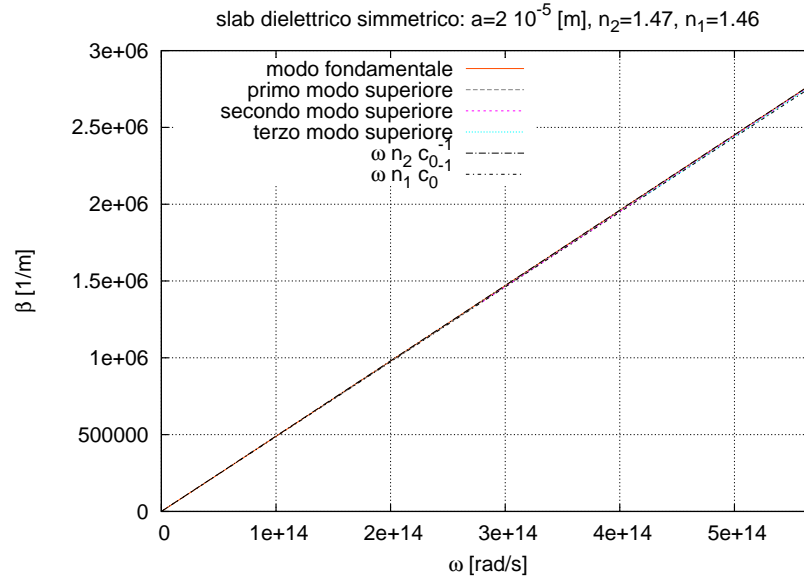


Figura 2.6: Andamento di $\beta(\omega)$ per i primi modi TE guidati indipendenti da y di uno slab simmetrico caratterizzato da $a = 20$ micron, $n_1 = 1.46$, $n_2 = 1.47$, $\mu_2 = \mu_1$.

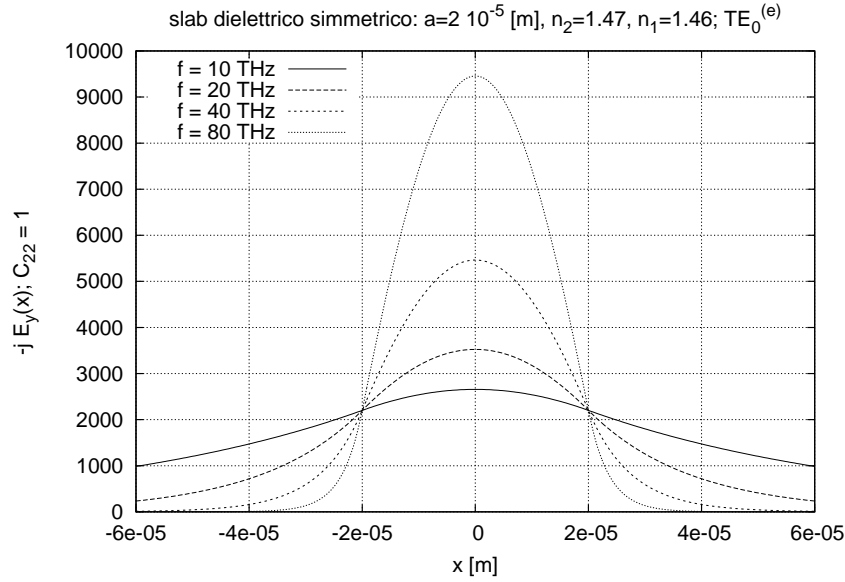


Figura 2.7: Andamento di $-jE_y(x)$ per il modo fondamentale $TE_0^{(e)}$ di uno slab simmetrico caratterizzato da $a = 20$ micron, $n_1 = 1.46$, $n_2 = 1.47$, $\mu_2 = \mu_1$. Il campo è stato valutato ponendo $C_{22} = 1$.

- fissato $n \in \mathbb{N}$, $n > 0$ nel caso di modo dispari ($n \geq 0$ nel caso di modo pari), si ha $k_{c2}a \in ((2n-1)\frac{\pi}{2}, n\pi)$ ($k_{c2}a \in (n\pi, n\pi + \frac{\pi}{2})$). Si noti che $k_{c2}x \in (-k_{c2}a, k_{c2}a)$. Inoltre, per $\omega \rightarrow \omega_n^+$ si ha $k_{c2}a \rightarrow (2n-1)\frac{\pi}{2} = \omega_n a \sqrt{\mu_2 \varepsilon_2 - \mu_1 \varepsilon_1}$ ($k_{c2}a \rightarrow n\pi = \omega_n a \sqrt{\mu_2 \varepsilon_2 - \mu_1 \varepsilon_1}$) e $Im(k_{c1}) \rightarrow 0$ (visto che $k_{c1}^2 = \omega^2(\mu_1 \varepsilon_1 - \mu_2 \varepsilon_2) + k_{c2}^2$). Infine, da $\beta^2 = \frac{\omega^2}{c_0^2} n_2^2 - k_{c2}^2$ si deduce, in entrambi i casi, $\lim_{\omega \rightarrow \omega_n^+} \beta(\omega) = \frac{\omega}{c_0} n_1$. Ciò implica che $\beta(\omega)$ è vicino alla retta $\frac{\omega}{c_0} n_1$ per $\omega \rightarrow \omega_n^+$. Per $\omega \rightarrow +\infty$ si ha $k_{c2}a \rightarrow n\pi$ ($k_{c2}a \rightarrow n\pi + \frac{\pi}{2}$), $\beta(\omega) \simeq \frac{\omega}{c_0} n_2$ e $|Im(k_{c1})| \rightarrow \infty$. Con queste considerazioni è possibile comprendere alcuni dettagli dei grafici riportati in Figura 2.4.

Per ottenere gli andamenti grafici delle componenti del modo TE guidato $TE_0^{(e)}$ ($TE_1^{(o)}$), riportati nelle figure 2.7 e 2.8 (2.9 e 2.10), qualche ulteriore considerazione quantitativa è necessaria. Ad esempio, se consideriamo una lastra avente $a = 20$ micron, $n_1 = 1.46$, $n_2 = 1.47$, $\mu_2 = \mu_1 = \mu_0$, per $f = 20$ THz ($f = 25$ THz) si ottiene $\omega \simeq 125.664 \cdot 10^{12}$ ($\omega \simeq 157.080 \cdot 10^{12}$) a cui corrisponde $k_{c2} \simeq 0.448112 \cdot 10^5$ ($k_{c2} \simeq 0.880653 \cdot 10^5$), $k_{c1} \simeq -j0.560334 \cdot 10^5$ ($k_{c1} \simeq -j0.169832 \cdot 10^5$), $\frac{\omega \mu_2}{k_{c2}} \simeq 3523.99$ ($\frac{\omega \mu_2}{k_{c2}} \simeq 2241.43$), $\beta = 0.614529 \cdot 10^6$ ($\beta = 0.765174 \cdot 10^6$) e $\frac{\gamma}{k_{c2}} = j \frac{\beta}{k_{c2}} = j13.7137$ ($\frac{\gamma}{k_{c2}} = j \frac{\beta}{k_{c2}} = j8.68871$).

Si tenga inoltre presente che $\frac{dE_y}{dx}$ è continua se $\mu_2 = \mu_1$ e presenta una discontinuità in caso contrario (come si deduce immediatamente dall'equazione (2.6) tenendo presente la continuità di H_z e l'indipendenza dalla coordinata y di tutte le componenti di campo dei modi qui considerati).

Le stesse figure mostrano gli andamenti anche a diversi valori di frequenza. Il lettore potrà svolgere autonomamente i calcoli che consentono di dedurre i corrispondenti andamenti. Gli andamenti delle componenti H_x dei modi alle diverse frequenze non sono riportati perchè, a meno di fattori costanti, coincidono con gli andamenti di E_y . Si verifichi, in particolare, che nel caso considerato, in cui $\mu_2 = \mu_1$, se si moltiplicano i coefficienti iniziali delle espressioni di E_y , $\frac{j\omega\mu_2}{k_{c2}}$ e $\frac{\omega\mu_1}{k_{c1}}$, per $\frac{-\beta}{\omega\mu_2}$ si ottengono, rispettivamente, $-\frac{\gamma}{k_{c2}}$ e $\frac{j\gamma}{k_{c1}}$, i coefficienti iniziali delle espressioni di H_x . Nel caso più generale in cui $\mu_2 \neq \mu_1$ la componente H_x presenterebbe un salto alle interfacce dielettriche che la componente E_y non può avere. Per completezza, la figura 2.11 riporta l'andamento di E_y del modo $TE_1^{(e)}$ a diverse frequenze. Si noti che, quando

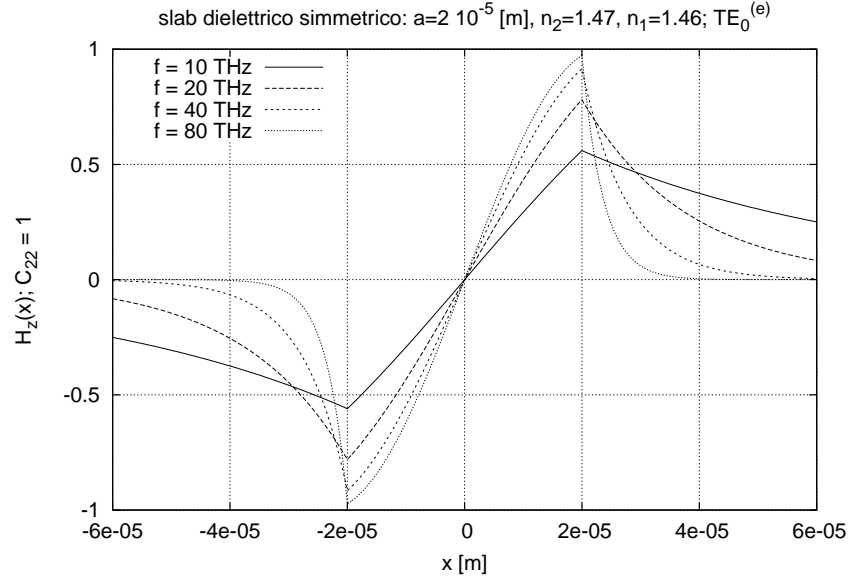


Figura 2.8: Andamento di $H_z(x)$ per il modo fondamentale $TE_0^{(e)}$ di uno slab simmetrico caratterizzato da $a = 20$ micron, $n_1 = 1.46$, $n_2 = 1.47$, $\mu_2 = \mu_1$. Il campo è stato valutato ponendo $C_{22} = 1$.

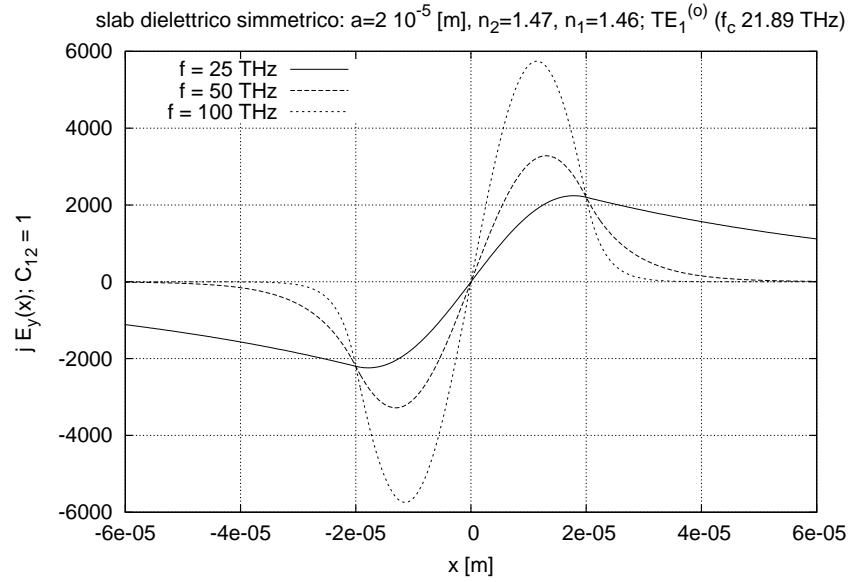


Figura 2.9: Andamento di $jE_y(x)$ per il primo modo superiore $TE_1^{(o)}$ di uno slab simmetrico caratterizzato da $a = 20$ micron, $n_1 = 1.46$, $n_2 = 1.47$, $\mu_2 = \mu_1$. Il campo è stato valutato ponendo $C_{12} = 1$.

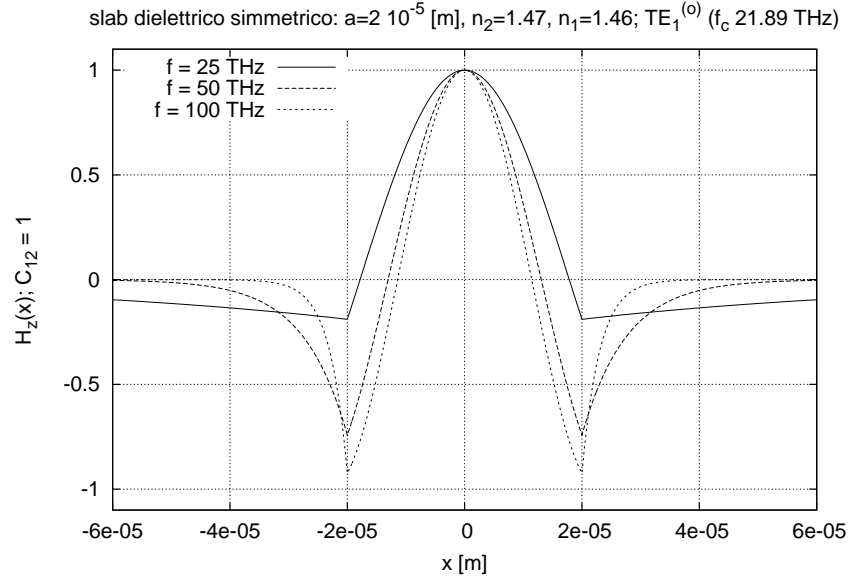


Figura 2.10: Andamento di $H_z(x)$ per il primo modo superiore $TE_1^{(o)}$ di uno slab simmetrico caratterizzato da $a = 20$ micron, $n_1 = 1.46$, $n_2 = 1.47$, $\mu_2 = \mu_1$. Il campo è stato valutato ponendo $C_{12} = 1$.

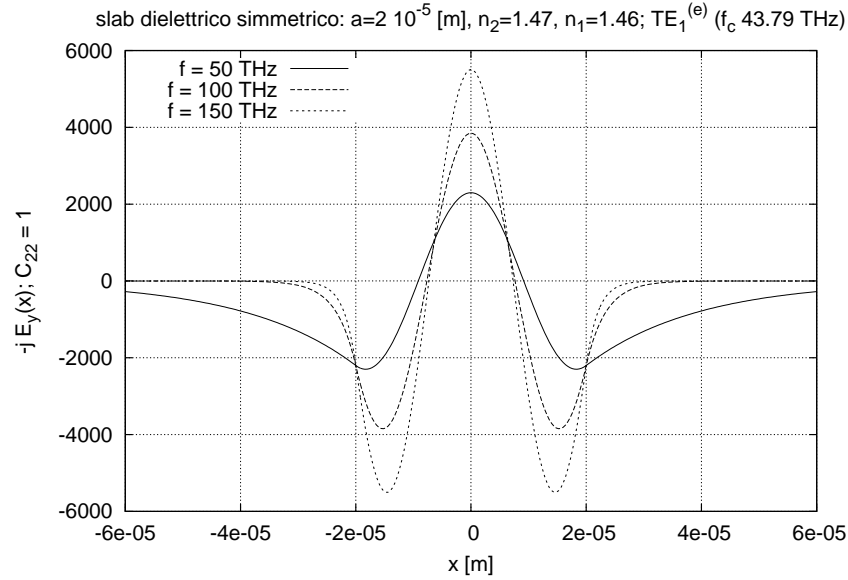


Figura 2.11: Andamento di $-jE_y(x)$ per il secondo modo superiore $TE_1^{(e)}$ di uno slab simmetrico caratterizzato da $a = 20$ micron, $n_1 = 1.46$, $n_2 = 1.47$, $\mu_2 = \mu_1$. Il campo è stato valutato ponendo $C_{22} = 1$.

esistono, i modi $\text{TE}_n^{(e)}$ o $\text{TE}_n^{(o)}$ presentano delle oscillazioni nel film aventi un frequenza spaziale che cresce con n , a causa della crescita di k_{c2} con n .

- dagli andamenti grafici che possono essere dedotti segue che i campi dei modi guidati variano la loro forma trasversale; se la frequenza cresce migliora il confinamento dei campi, ottenendo un confinamento ideale al limite per $\omega \rightarrow +\infty$; il confinamento si riduce e il campo rimane esponenzialmente confinato in un intorno del film per ogni valore finito di frequenza superiore alla frequenza di taglio; il confinamento tende ad annullarsi al limite per ω che tende (da sopra) alla frequenza di taglio del modo considerato.

I risultati per mezzi con poche perdite rimangono ingegneristicamente corretti. In particolare, ogni modo guidato presenterà necessariamente una attenuazione non nulla.

Esercizio 3 Per una lastra dielettrica debolmente guidante, si valutino le ampiezze relative delle componenti trasversali dei modi TE pari confrontandole con quella della componente assiale del campo magnetico.

Soluzione. Per il decadimento esponenziale di tutte le componenti di campo nel rivestimento e nel substrato, le ampiezze di interesse raggiungono il loro valore massimo nel film (si vedano le espressioni (2.78), (2.79), (2.80), (2.81), (2.82), (2.83), (2.84), (2.85) e (2.86)). Inoltre le componenti trasversali assumono l'ampiezza massima proprio per $x = 0$. In particolare, per $x = 0$ si ha $|E_y| = \frac{\omega\mu_2}{k_{c2}}|C_{22}|$ e $|H_x| = \frac{\beta}{k_{c2}}|C_{22}|$. La componente assiale H_z , infine, risulta avere in ogni caso un modulo strettamente inferiore a $|C_{22}|$.

Tenendo presente che $\beta > \frac{\omega n_1}{c_0}$ e che $k_{c2}a < \frac{\omega a}{c_0} \sqrt{n_2^2 - n_1^2}$ si conclude che

$$\frac{|H_x|}{|H_z|} > \frac{n_1}{\sqrt{n_2^2 - n_1^2}}. \quad (2.95)$$

Ad esempio per $n_1 = 1.46$ e $n_2 = 1.47$ si ottiene che il membro di destra della precedente disuguaglianza è pari a circa 8.529.

Analogamente, per la componente trasversale del campo elettrico si ottiene

$$\frac{|E_y|}{|H_z|} > \frac{\mu_2 c_0}{\sqrt{n_2^2 - n_1^2}} = \sqrt{\frac{\mu_0}{\epsilon_0}} \frac{\mu_{r2}}{\sqrt{n_2^2 - n_1^2}}. \quad (2.96)$$

Anche questo secondo rapporto, a meno della usuale presenza dell'impedenza intrinseca del vuoto $\sqrt{\frac{\mu_0}{\epsilon_0}}$ nei legami tra le ampiezze delle componenti del campo elettrico e del campo magnetico, è significativamente superiore a 1 (con i valori indicati, se, come accade di solito, $\mu_{r2} = 1$, si ottiene $\frac{1}{\sqrt{n_2^2 - n_1^2}} \simeq 5.842$).

Per queste ragioni, tenendo presente che lo stesso risultato si ottiene per i modi TE dispari, nelle lastre dielettriche debolmente guidanti si usa spesso dire che le componenti trasversali sono di ampiezza molto maggiore di quella assiale.

Alcune possibili interpretazioni in termini di ottica geometrica

Nel film la componente E_y di un modo TE pari è data dall'espressione (2.84). La si può esprimere in termini di esponenziali complessi. Tenendo conto che, per quanto visto, $\gamma = j\beta$ è possibile esprimere la componente E_y come funzione di x e z . Si ottiene:

$$E_y(x, z) = \frac{j\omega\mu_2}{2k_{c2}} C_{22} (e^{jk_{c2}x} + e^{-jk_{c2}x}) e^{-j\beta z}. \quad (2.97)$$

È evidente che nel film il campo elettromagnetico si può esprimere come la sovrapposizione di due onde piane. Il vettore d'onda è $k_{c2}\hat{\mathbf{x}} + \beta\hat{\mathbf{z}}$ per un'onda e $-k_{c2}\hat{\mathbf{x}} + \beta\hat{\mathbf{z}}$ per l'altra. L'angolo δ ($-\delta$) formato dal primo (secondo) di questi vettori con la normale alle interfacce, misurato, a partire dalla parte della normale interna al film, in senso orario, è quindi tale che $\tan(\delta) = \frac{\beta}{k_{c2}}$. La figura 2.12 riporta le informazioni salienti per questo tipo di analisi.

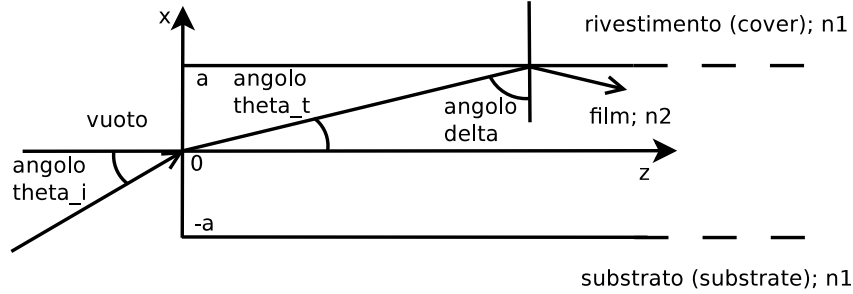


Figura 2.12: Geometria dello slab dielettrico simmetrico e angoli di interesse.

Al “cut-off”, per il modo $TE_m^{(e)}$ è noto che $k_{c2}|_{\omega=\omega_m^+} = \omega_m \sqrt{\mu_2 \varepsilon_2 - \mu_1 \varepsilon_1}$, $\beta|_{\omega=\omega_m^+} = \omega_m \sqrt{\mu_1 \varepsilon_1}$. Complessivamente, al taglio si deduce che $\tan(\delta)|_{\omega=\omega_m^+} = \frac{\beta|_{\omega=\omega_m^+}}{k_{c2}|_{\omega=\omega_m^+}} = \frac{\sqrt{\mu_1 \varepsilon_1}}{\sqrt{\mu_2 \varepsilon_2 - \mu_1 \varepsilon_1}} = \frac{n_1}{\sqrt{n_2^2 - n_1^2}}$. Ricordando il fenomeno della riflessione totale ($\sin(\delta) = \frac{n_1}{n_2}$ da cui $\tan(\delta) = \frac{\sin(\delta)}{\cos(\delta)} = \frac{\sin(\delta)}{\sqrt{1 - \sin^2(\delta)}} = \frac{\frac{n_1}{n_2}}{\sqrt{1 - \frac{n_1^2}{n_2^2}}} = \frac{n_1}{\sqrt{n_2^2 - n_1^2}}$), si deduce che alla frequenza di taglio le onde piane che si propagano nel film incidono sull'interfaccia con un angolo pari all'angolo limite per la riflessione totale (si vedano le pp. 20, 310, 311, di [6]).

Per valori di frequenza superiori a quella di cut-off, in presenza di mezzi non dispersivi nel tempo, dagli andamenti grafici riportati nelle figure 2.4, 2.5, 2.6, si deduce che β cresce più che linearmente con ω mentre k_{c2} satura a un valore finito (si veda, ad esempio, la figura 2.3). Ciò implica che il rapporto $\frac{\beta}{k_{c2}}$ cresce con ω e non è limitato e, quindi, δ cresce con ω e tende a $\pi/2$ per $\omega \rightarrow +\infty$.

In conclusione, i modi guidati dello slab simmetrico possono essere considerati come delle onde piane monocromatiche che si propagano in direzioni oblique rispetto all'asse z e che subiscono continue riflessioni totali alle due interfacce dielettriche. Nelle condizioni indicate (di riflessione totale) non è una sorpresa notare che il campo nel rivestimento e nel substrato è un campo evanescente [6] (p. 311).

Per eccitare un modo guidato di uno slab simmetrico con della luce esterna è conveniente limitare l'angolo di incidenza θ_i , che la luce forma con l'asse z (si veda nuovamente la figura 2.12), a un valore tale da avere $\sin(\theta_i) \leq \sqrt{n_2^2 - n_1^2}$. Considerando la legge di Snell, infatti, si ha $\frac{\sin(\theta_i)}{\sin(\theta_t)} = \frac{n_2}{n_{\text{vuoto}}} = n_2$. Inoltre, dalla figura 2.12, si deduce $\sin(\theta_t) = \cos(\delta)$. Infine, in condizioni di riflessione totale alle interfacce film-rivestimento e film-substrato, si ha $\sin(\delta) \geq \frac{n_1}{n_2}$ e quindi $\cos(\delta) = \sqrt{1 - \sin^2(\delta)} \leq \sqrt{1 - \frac{n_1^2}{n_2^2}} = \frac{\sqrt{n_2^2 - n_1^2}}{n_2}$. Dalle deduzioni precedenti segue che $\sin(\theta_i) = n_2 \sin(\theta_t) = n_2 \cos(\delta) \leq n_2 \frac{\sqrt{n_2^2 - n_1^2}}{n_2}$, che coincide con il risultato fornito in precedenza. Al parametro $\sqrt{n_2^2 - n_1^2}$, che caratterizza l'angolo massimo di incidenza della luce esterna che consente di eccitare un modo guidato, viene dato il nome di apertura numerica.

Questa spiegazione funziona anche per altre strutture guidanti, come la fibra ottica, dove, a causa della complicata geometria, il fenomeno della riflessione totale non è così immediato da dedurre [6] (p. 311).

Funziona anche per il modo fondamentale dello slab, che non presenta taglio, se si considera, nelle espressioni precedenti $\omega_0 = 0$. Ovviamente, per tale modo non è possibile violare la condizione di riflessione totale.

In ogni caso si tenga presente che l'ottica geometrica non consente di ottenere tutti i risultati deducibili in termini di campi elettromagnetici [6] (p. 311), [1] (pp. 2-3) (ad esempio, perchè per alcuni modi sia possibile violare la condizione di riflessione totale, presentando il fenomeno del

taglio, e per altri no).

Modi TE con condizione al contorno (2.36)

In una delle sottosezioni precedenti sono stati considerati i modi che si deducono per $Im(k_{c1}) \neq 0$ e che presentano un decadimento trasversale esponenziale per $|x| \rightarrow \infty$. In particolare, se $Im(k_{c1}) \neq 0$ la condizione (2.36) esclude la possibilità di ottenere un esponenziale crescente e lascia i soli termini che coinvolgono esponenziali decrescenti. Quindi si ritorna all'analisi precedente.

Dei modi completamente indipendenti sono invece possibili se $Im(k_{c1}) = 0$, ovvero se $k_{c1} \in \mathbb{R}$. Dato che non esistono altri casi da considerare ci concentriamo su questa condizione. Si noti che $k_{c1} \in \mathbb{R}$ implica $k_{c1}^2 \in \mathbb{R}$, $k_{c1}^2 \geq 0$.

Il caso $k_{c1}^2 = 0$ va gestito a parte. $k_{c1}^2 = 0$ implica $\gamma \neq 0$ se $\omega \neq 0$ e $\mu_1 \varepsilon_1 \neq 0$. Quindi, direttamente dall'equazione (2.11) per campi indipendenti dalla coordinata y , tenendo conto che $\gamma \neq 0$, si deduce $\frac{\partial H_z}{\partial x} = 0$ nel rivestimento e nel substrato e, di conseguenza, $H_z = C_1$ nel rivestimento e $H_z = C_3$ nel substrato. Sempre per campi indipendenti dalla coordinata y , per l'equazione (2.6), nel rivestimento deduciamo $E_y = -j\omega\mu_1 C_1 x + C_4$; nel substrato si ottiene $E_y = -j\omega\mu_1 C_3 x + C_5$. Se $C_1 \neq 0$ o $C_3 \neq 0$ tali campi divergono nel rivestimento o nel substrato e tale condizione viene considerata inammissibile (si tenga presente che, avendo $k_{c1} = 0$, dobbiamo accettare che le condizioni al contorno non possano essere completamente espresse in termini di componenti assiali; la limitatezza qui assunta ha la stessa dignità di quelle considerate nelle condizioni al contorno (2.35) o (2.36)). Ne segue che i modi TE con componenti limitate e indipendenti dalla coordinata y che soddisfano la condizione $k_{c1}^2 = 0$ devono necessariamente avere $C_1 = 0 = C_3$ e ciò implica $H_z = 0$ nel rivestimento e nel substrato (oltre che $E_y = C_4$ nel rivestimento e $E_y = C_5$ nel substrato e quindi, per la (2.5) e ricordando che $k_{c1}^2 = 0$ e quindi $\gamma^2 = -\omega^2 \mu_1 \varepsilon_1$, $H_x = -\frac{n_1}{c_0 \mu_1} C_4$ nel rivestimento e $H_x = -\frac{n_1}{c_0 \mu_1} C_5$ nel substrato). Nel film, la componente H_z assume la solita espressione $H_z(x) = C_{12} \cos(k_{c2}x) + C_{22} \sin(k_{c2}x)$, $|x| < a$ (si veda l'equazione (2.37)), visto che $k_{c2} \neq 0$ ($k_{c2}^2 = \omega^2(\mu_2 \varepsilon_2 - \mu_1 \varepsilon_1) + k_{c1}^2 = \omega^2(\mu_2 \varepsilon_2 - \mu_1 \varepsilon_1) > 0$ nelle ipotesi considerate). Dalla continuità di H_z seguono $C_{12} \cos(k_{c2}a) + C_{22} \sin(k_{c2}a) = 0$ e $C_{12} \cos(k_{c2}a) - C_{22} \sin(k_{c2}a) = 0$. Si tratta di un sistema algebrico lineare omogeneo nelle indeterminate C_{12} e C_{22} . La soluzione nulla $C_{12} = 0 = C_{22}$ non è ammissibile poichè implicherebbe H_z identicamente nullo e i modi considerati sarebbero dei TEM (che si studiano con altre tecniche; si veda la corrispondente sottosezione). Una soluzione non nulla del sistema algebrico lineare omogeneo si ottiene se e solo se $\cos(k_{c2}a) \sin(k_{c2}a) = 0$. Avendo verificato che $\cos^2(z) + \sin^2(z) = 1$, $\forall z \in \mathbb{C}$, non è possibile avere $\cos(k_{c2}a) = 0$ e $\sin(k_{c2}a) = 0$. Se $\cos(k_{c2}a) = 0$ ($\sin(k_{c2}a) = 0$) si deduce $\sin(k_{c2}a) \neq 0$ ($\cos(k_{c2}a) \neq 0$) e, quindi, $C_{22} = 0$ ($C_{12} = 0$). Da ciò segue $H_z = C_{12} \cos(k_{c2}x)$ ($H_z = C_{22} \sin(k_{c2}x)$) nel film e $k_{c2}a = \pi/2 + n\pi$ ($k_{c2}a = n\pi$), $n \in \mathbb{N}$ ($n \in \mathbb{N}$, $n \neq 0$, visto che con $n = 0$ si avrebbe $k_{c2} = 0$ e si otterrebbe nuovamente H_z identicamente nullo). Ma si è visto che $k_{c2}^2 = \omega^2(\mu_2 \varepsilon_2 - \mu_1 \varepsilon_1)$ e quindi $\pi/2 + n\pi = \omega a \sqrt{\mu_2 \varepsilon_2 - \mu_1 \varepsilon_1}$, ($n\pi = \omega a \sqrt{\mu_2 \varepsilon_2 - \mu_1 \varepsilon_1}$). Di conseguenza i modi TE con componenti limitate e indipendenti dalla coordinata y che soddisfano la condizione $k_{c1}^2 = 0$ possono esistere solo per valori di frequenza angolare particolari, pari a $\omega_n = \frac{(\pi/2 + n\pi)c_0}{a\sqrt{n^2 - n_1^2}}$, $n \in \mathbb{N}$ ($\omega_n = \frac{n\pi c_0}{a\sqrt{n^2 - n_1^2}}$, $n \in \mathbb{N}$, $n \neq 0$). Tali valori di frequenza angolare coincidono con le pulsazioni di taglio dei modi TE guidati, riportate nelle equazioni (2.94), se si esclude la pulsazione di taglio nulla del modo fondamentale. Per tali valori di frequenza si deve verificare che un modo TE possa esistere. Per farlo, non dobbiamo fare altro che determinare coerentemente E_y nel film e verificare che complessivamente E_y risulti essere continuo alle interfacce. Ma nel film, avendo $k_{c2} \neq 0$, la componente H_z determina univocamente le altre componenti trasversali. Ne segue, per l'equazione (2.8), $E_y = \frac{j\omega\mu_2}{k_{c2}^2} \frac{dH_z}{dx} = -\frac{j\omega\mu_2}{k_{c2}} C_{12} \sin(k_{c2}x)$ ($E_y = \frac{j\omega\mu_2}{k_{c2}} C_{22} \cos(k_{c2}x)$). La continuità di E_y alle interfacce dielettriche può essere sempre verificata grazie all'arbitrarietà delle costanti C_4 e C_5 (H_x risulta automaticamente individuata ovunque ma su essa non abbiamo requisiti di continuità; inoltre per i modi di interesse ovviamente E_z si pone a zero e le componenti E_x e H_y possono essere poste a zero; il modo, quindi, è completamente definito alle pulsazioni angolari indicate).

Osservazione 12 *I modi TE con componenti limitate e indipendenti dalla coordinata y che soddisfano la condizione $k_{c1}^2 = 0$ possono esistere solo per valori di frequenza angolare pari alla pulsazione di taglio dei modi TE guidati. Tali modi sono il limite in atto dei modi guidati. Non possono essere inclusi nella sezione dedicata ai modi TE guidati poichè non presentano un decadimento a mano a mano che ci si allontana dal film. Nel rivestimento e nel substrato tali modi sono delle vere e proprie onde TEM, che si propagano senza attenuazione (nello slab senza perdite) in direzione assiale e sono trasversalmente uniformi (limitatamente al rivestimento e al substrato). Per queste caratteristiche non sono identificabili né come modi irradianti né come modi evanescenti (si vedano le definizioni che seguono per queste nuove famiglie di modi).*

Resta da valutare cosa succede quando $k_{c1}^2 > 0$. Sapendo che $k_{c1}^2 = \omega^2 \mu_1 \varepsilon_1 + \gamma^2$ si ottiene $\gamma^2 \in \mathbb{R}$. Ciò implica, da un lato, $k_{c2}^2 = \omega^2 \mu_2 \varepsilon_2 + \gamma^2 \in \mathbb{R}$ e, dall'altro, $\gamma = \alpha$ ($\beta = 0$) o $\gamma = j\beta$ ($\alpha = 0$). Nel secondo dei due casi in “or” si ottiene $k_{c1}^2 = \omega^2 \mu_1 \varepsilon_1 - \beta^2 > 0$ se e solo se $\beta < \omega \sqrt{\mu_1 \varepsilon_1}$. Con questo vincolo si deduce $k_{c2}^2 > 0$ se vale la solita ipotesi ($\mu_2 \varepsilon_2 - \mu_1 \varepsilon_1 > 0$). Nel primo caso, invece, $k_{c1}^2 = \omega^2 \mu_1 \varepsilon_1 + \alpha^2 \geq \omega^2 \mu_1 \varepsilon_1$ e $k_{c2}^2 = \omega^2 \mu_2 \varepsilon_2 + \alpha^2 \geq \omega^2 \mu_2 \varepsilon_2$.

I modi oggetto di indagine in questo paragrafo caratterizzati da $\gamma = j\beta$ e $\alpha = 0$ vengono detti modi irradianti (TE radiation modes, in inglese). Per l'osservazione 1 possiamo evitare di considerare $\beta < 0$. Si chiamano modi evanescenti quelli caratterizzati da $\gamma = \alpha$ e $\beta = 0$ (TE evanescent modes, in inglese). Anche in questo caso, per l'osservazione 1 possiamo evitare di considerare $\alpha < 0$. Il modo con $\gamma = 0$ sembra appartenere a entrambe le classi di modi. Per fare ordine e mantenere le classi disgiunte tale modo viene rimosso dalla classe dei modi evanescenti. Lo considereremo un modo irradiente.

Come si trovano questi modi (irradianti o evanescenti) e che caratteristiche hanno? Dato che $k_{c1}^2 > 0$ implica $k_{ci} \neq 0$, $i = 1, 2$, l'algoritmo basato sulle componenti assiali può essere utilizzato. In particolare, non cambiano le espressioni degli integrali generali (2.37), (2.38), (2.39). Quindi, anche in questo caso l'espressione complessiva di H_z presenta sette parametri liberi. Tuttavia, ora le condizioni al contorno non impongono alcuna restrizione al campo considerato nel rivestimento e nel substrato (nella sottosezione precedente, al contrario, le condizioni al contorno consentivano di ridurre da sette a cinque il numero di parametri liberi).

Per ogni ω di interesse si consideri un qualunque $0 \leq \beta < \omega \sqrt{\mu_1 \varepsilon_1}$ e $\alpha = 0$ per i modi irradianti e un qualunque $\alpha > 0$ e $\beta = 0$ per i modi evanescenti. Con le scelte indicate per uno dei modi irradianti e uno dei modi evanescenti si determinano i valori di k_{c1}^2 e k_{c2}^2 . Per la presenza di esponenziali con entrambi i segni dell'esponente non è restrittivo considerare $k_{c1} > 0$. Nel film, per k_{c2} reale, l'espressione $C_{12} \cos(k_{c2}x) + C_{22} \sin(k_{c2}x)$ può essere sostituita da $\frac{C_{12} + C_{22}}{2} e^{jk_{c2}x} + \frac{C_{12} - C_{22}}{2} e^{-jk_{c2}x}$ e, ripetendo il ragionamento precedente, si può assumere $k_{c2} > 0$. Ciò è in linea con il comportamento usualmente considerato per la funzione radice quadrata.

Con questi parametri fissati i vincoli di continuità sulle interfacce dielettriche si traducono in quattro vincoli lineari omogenei nelle sei indeterminate C_{ij} , $i = 1, 2$, $j = 1, 2, 3$. Tale sistema ammette sempre soluzioni non banali. Quindi esiste una infinità continua di modi TE irradianti e una infinità continua di modi TE evanescenti.

Esercizio 4 *Considerando le espressioni (2.37), (2.38), (2.39) per il campo H_z nelle tre regioni, assumendo i valori reali e positivi di k_{c1} e k_{c2} noti e non nulli, e forzando le quattro condizioni alle interfacce dielettriche (2.23), (2.24), (2.29) e (2.30), si determinino le espressioni di $H_z(x)$, $E_y(x)$ e $H_x(x)$.*

Soluzione. *La condizione (2.23) implica*

$$C_{13}e^{-jk_{c1}a} + C_{23}e^{jk_{c1}a} = C_{12} \cos(k_{c2}a) - C_{22} \sin(k_{c2}a). \quad (2.98)$$

In modo duale la condizione (2.24) comporta

$$C_{11}e^{jk_{c1}a} + C_{21}e^{-jk_{c1}a} = C_{12} \cos(k_{c2}a) + C_{22} \sin(k_{c2}a). \quad (2.99)$$

$e^{-jk_{c1}a} \neq 0$ in ogni caso e dalle equazioni precedenti posso ricavare

$$C_{13} = -C_{23}e^{2jk_{c1}a} + (C_{12} \cos(k_{c2}a) - C_{22} \sin(k_{c2}a))e^{jk_{c1}a} \quad (2.100)$$

e

$$C_{21} = -C_{11}e^{2jk_{c1}a} + (C_{12}\cos(k_{c2}a) + C_{22}\sin(k_{c2}a))e^{jk_{c1}a}. \quad (2.101)$$

La condizione (2.29) che impone la continuità di E_y per $x = -a$ implica

$$\begin{aligned} \frac{\mu_1}{k_{c1}^2} \frac{dH_z}{dx} \Big|_{x=-a^-} &= \frac{\mu_1}{k_{c1}^2} (C_{13}(jk_{c1})e^{-jk_{c1}a} + C_{23}(-jk_{c1})e^{jk_{c1}a}) = \\ &= \frac{\mu_2}{k_{c2}^2} \frac{dH_z}{dx} \Big|_{x=-a^+} = \frac{\mu_2}{k_{c2}^2} (C_{12}k_{c2}\sin(k_{c2}a) + C_{22}k_{c2}\cos(k_{c2}a)) \end{aligned} \quad (2.102)$$

da cui si deduce

$$\frac{\mu_1}{jk_{c1}} (C_{23}e^{jk_{c1}a} - C_{13}e^{-jk_{c1}a}) = \frac{\mu_2}{k_{c2}} (C_{12}\sin(k_{c2}a) + C_{22}\cos(k_{c2}a)). \quad (2.103)$$

Sostituendo l'espressione di C_{13} , data dalla relazione (2.100), nell'equazione precedente si ottiene:

$$\frac{\mu_1}{jk_{c1}} (C_{23}e^{jk_{c1}a} + C_{23}e^{jk_{c1}a} - (C_{12}\cos(k_{c2}a) - C_{22}\sin(k_{c2}a))) = \frac{\mu_2}{k_{c2}} (C_{12}\sin(k_{c2}a) + C_{22}\cos(k_{c2}a)) \quad (2.104)$$

da cui segue

$$2C_{23}e^{jk_{c1}a} = (C_{12}\cos(k_{c2}a) - C_{22}\sin(k_{c2}a)) + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} (C_{12}\sin(k_{c2}a) + C_{22}\cos(k_{c2}a)) \quad (2.105)$$

e quindi

$$C_{23} = \frac{1}{2}e^{-jk_{c1}a} \left((C_{12} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{22}) \cos(k_{c2}a) + (-C_{22} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{12}) \sin(k_{c2}a) \right) \quad (2.106)$$

che esprime C_{23} in funzione di C_{12} e C_{22} . Sostituendo questa espressione nella relazione (2.100) si ricava:

$$\begin{aligned} C_{13} &= (C_{12}\cos(k_{c2}a) - C_{22}\sin(k_{c2}a))e^{jk_{c1}a} + \\ &- \frac{1}{2}e^{jk_{c1}a} \left((C_{12} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{22}) \cos(k_{c2}a) + (-C_{22} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{12}) \sin(k_{c2}a) \right) \end{aligned} \quad (2.107)$$

e da questa segue la relazione che esprime C_{13} in funzione di C_{12} e C_{22} :

$$C_{13} = -\frac{1}{2}e^{jk_{c1}a} \left((-C_{12} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{22}) \cos(k_{c2}a) + (C_{22} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{12}) \sin(k_{c2}a) \right). \quad (2.108)$$

In modo analogo, la condizione (2.30) che impone la continuità di E_y per $x = a$ implica

$$\begin{aligned} \frac{\mu_1}{k_{c1}^2} \frac{dH_z}{dx} \Big|_{x=a^+} &= \frac{\mu_1}{k_{c1}^2} (C_{11}(jk_{c1})e^{jk_{c1}a} + C_{21}(-jk_{c1})e^{-jk_{c1}a}) = \\ &= \frac{\mu_2}{k_{c2}^2} \frac{dH_z}{dx} \Big|_{x=a^-} = \frac{\mu_2}{k_{c2}^2} (-C_{12}k_{c2}\sin(k_{c2}a) + C_{22}k_{c2}\cos(k_{c2}a)) \end{aligned} \quad (2.109)$$

da cui si deduce

$$\frac{\mu_1}{jk_{c1}} (C_{21}e^{-jk_{c1}a} - C_{11}e^{jk_{c1}a}) = \frac{\mu_2}{k_{c2}} (-C_{12}\sin(k_{c2}a) + C_{22}\cos(k_{c2}a)). \quad (2.110)$$

Sostituendo l'espressione di C_{21} , data dalla relazione (2.101), nell'equazione precedente si ottiene:

$$\frac{\mu_1}{jk_{c1}} (-C_{11}e^{jk_{c1}a} - C_{11}e^{jk_{c1}a} + (C_{12}\cos(k_{c2}a) + C_{22}\sin(k_{c2}a))) = \frac{\mu_2}{k_{c2}} (-C_{12}\sin(k_{c2}a) + C_{22}\cos(k_{c2}a)). \quad (2.111)$$

da cui segue

$$2C_{11}e^{jk_{c1}a} = (C_{12} \cos(k_{c2}a) + C_{22} \sin(k_{c2}a)) + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} (C_{12} \sin(k_{c2}a) - C_{22} \cos(k_{c2}a)) \quad (2.112)$$

e quindi

$$C_{11} = \frac{1}{2}e^{-jk_{c1}a} \left((C_{12} - \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{22}) \cos(k_{c2}a) + (C_{22} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{12}) \sin(k_{c2}a) \right) \quad (2.113)$$

che esprime C_{11} in funzione di C_{12} e C_{22} . Sostituendo questa espressione nella relazione (2.101) si ricava:

$$C_{21} = (C_{12} \cos(k_{c2}a) + C_{22} \sin(k_{c2}a))e^{jk_{c1}a} + \frac{1}{2}e^{jk_{c1}a} \left((C_{12} - \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{22}) \cos(k_{c2}a) + (C_{22} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{12}) \sin(k_{c2}a) \right) \quad (2.114)$$

e da questa segue la relazione che esprime C_{21} in funzione di C_{12} e C_{22} :

$$C_{21} = \frac{1}{2}e^{jk_{c1}a} \left((C_{12} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{22}) \cos(k_{c2}a) + (C_{22} - \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{12}) \sin(k_{c2}a) \right). \quad (2.115)$$

A questo punto si possono dedurre le espressioni finali di $H_z(x)$, $E_y(x)$ e $H_x(x)$ in funzione dei soli parametri C_{12} e C_{22}

$$H_z(x) = \begin{cases} \frac{1}{2}e^{jk_{c1}(x-a)} \left((C_{12} - \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{22}) \cos(k_{c2}a) + (C_{22} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{12}) \sin(k_{c2}a) \right) + \\ \quad + \frac{1}{2}e^{-jk_{c1}(x-a)} \left((C_{12} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{22}) \cos(k_{c2}a) + (C_{22} - \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{12}) \sin(k_{c2}a) \right) & x > a \\ C_{12} \cos(k_{c2}x) + C_{22} \sin(k_{c2}x) & |x| < a \\ -\frac{1}{2}e^{jk_{c1}(x+a)} \left((-C_{12} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{22}) \cos(k_{c2}a) + (C_{22} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{12}) \sin(k_{c2}a) \right) + \\ \quad + \frac{1}{2}e^{-jk_{c1}(x+a)} \left((C_{12} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{22}) \cos(k_{c2}a) + (-C_{22} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{12}) \sin(k_{c2}a) \right) & x < -a \end{cases} \quad (2.116)$$

$$E_y(x) = \begin{cases} \frac{\omega\mu_1}{2k_{c1}} e^{jk_{c1}(x-a)} \left(- (C_{12} - \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{22}) \cos(k_{c2}a) - (C_{22} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{12}) \sin(k_{c2}a) \right) + \\ \quad + \frac{\omega\mu_1}{2k_{c1}} e^{-jk_{c1}(x-a)} \left((C_{12} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{22}) \cos(k_{c2}a) + (C_{22} - \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{12}) \sin(k_{c2}a) \right) & x > a \\ \frac{j\omega\mu_2}{k_{c2}} (-C_{12} \sin(k_{c2}x) + C_{22} \cos(k_{c2}x)) & |x| < a \\ \frac{\omega\mu_1}{2k_{c1}} e^{jk_{c1}(x+a)} \left((C_{22} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{12}) \sin(k_{c2}a) - (C_{12} - \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{22}) \cos(k_{c2}a) \right) + \\ \quad + \frac{\omega\mu_1}{2k_{c1}} e^{-jk_{c1}(x+a)} \left((C_{12} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{22}) \cos(k_{c2}a) + (-C_{22} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{12}) \sin(k_{c2}a) \right) & x < -a \end{cases} \quad (2.117)$$

$$H_x(x) = \begin{cases} -\frac{j\gamma}{2k_{c1}} e^{jk_{c1}(x-a)} \left((C_{12} - \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{22}) \cos(k_{c2}a) + (C_{22} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{12}) \sin(k_{c2}a) \right) + \\ \quad + \frac{j\gamma}{2k_{c1}} e^{-jk_{c1}(x-a)} \left((C_{12} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{22}) \cos(k_{c2}a) + (C_{22} - \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{12}) \sin(k_{c2}a) \right) & x > a \\ \frac{\gamma}{k_{c2}} (C_{12} \sin(k_{c2}x) - C_{22} \cos(k_{c2}x)) & |x| < a \\ \frac{j\gamma}{2k_{c1}} e^{jk_{c1}(x+a)} \left((C_{22} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{12}) \sin(k_{c2}a) - (C_{12} - \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{22}) \cos(k_{c2}a) \right) + \\ \quad + \frac{j\gamma}{2k_{c1}} e^{-jk_{c1}(x+a)} \left((C_{12} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{22}) \cos(k_{c2}a) + (-C_{22} + \frac{jk_{c1}\mu_2}{\mu_1 k_{c2}} C_{12}) \sin(k_{c2}a) \right) & x < -a \end{cases} \quad (2.118)$$

Osservazione 13 La soluzione dell'esercizio 4 consente di suddividere i modi TE irradianti e evanescenti in modi pari e dispari.

Si vede facilmente che, per $|x| > a$, i modi irradianti corrispondono a onde che si propagano in direzioni oblique rispetto a quella dell'asse z , mentre i modi evanescenti corrispondono a onde il cui vettore di attenuazione è diretto lungo l'asse z mentre il vettore di fase è diretto lungo l'asse x [7] (p. 262).

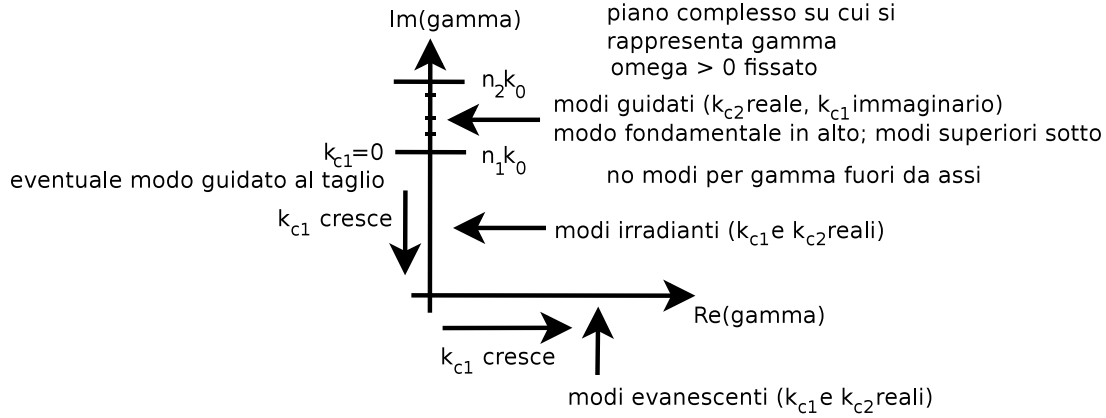


Figura 2.13: Valori ammissibili di γ per i modi TE indipendenti da y di una lastra dielettrica in presenza di mezzi privi di perdite, per un valore di ω fissato. Vengono fornite informazioni anche sui valori di k_{c1} (i valori di k_{c2} sono facilmente deducibili). k_0 indica $\omega\sqrt{\mu_0\epsilon_0}$. In corrispondenza di $\beta = n_1k_0$ si ha ovviamente $k_{c1} = 0$ e $k_{c2} = k_0\sqrt{n_2^2 - n_1^2}$. Nell'origine $\gamma = 0$, $k_{c1} = n_1k_0$ e $k_{c2} = n_2k_0$.

Può essere utile, ad esempio da un punto di vista mnemonico, rappresentare sul piano complesso il luogo dei valori di γ ammissibili per i modi TE indipendenti da y della lastra dielettrica senza perdite. La figura 2.13 indica graficamente tali valori ammissibili (per una rappresentazione analoga si veda [6] (p. 321)).

Modi TM

In molti testi, compresi [6] e [7], la lastra dielettrica viene studiata considerando $\mu_2 = \mu_1$. In queste condizioni la determinazione dei modi TM richiede uno sviluppo indipendente, poichè non è direttamente deducibile dai risultati ottenuti per i modi TE.

Al contrario, avendo considerato parametri costitutivi generici, i modi TM possono dedursi dai modi TE, sostanzialmente per dualità (bisogna scambiare \mathbf{E} , \mathbf{H} , ϵ e μ con \mathbf{H} , $-\mathbf{E}$, μ e ϵ , rispettivamente; con questo approccio, ovviamente, i modi TM di una lastra dielettrica si deducono dai TE della lastra dielettrica duale, ottenuta da quella di interesse scambiando, come indicato in precedenza, ϵ con μ e μ con ϵ).

Modi ibridi

Supponiamo di conoscere, a una certa frequenza, un modo ibrido indipendente dalla coordinata y la cui componente assiale H_z soddisfa le condizioni (2.34) o (2.36). Supponiamo che E_z soddisfi una delle condizioni corrispondenti. Non ipotizziamo che la condizione all'infinito sia necessariamente dello stesso tipo per le due componenti assiali. Dalla conoscenza del modo ibrido possiamo determinare certamente i valori di ω , γ e i campi $E_z(x)$ e $H_z(x)$. Nella solita ipotesi considerata all'inizio ($\mu_2\epsilon_2 > \mu_1\epsilon_1$) solo uno dei due k_{ci} può essere nullo. Per semplicità, in questa sezione evitiamo di considerare i casi in cui uno dei due k_{ci} si annulla. Questi casi potrebbero essere gestiti come indicato in precedenza. Se assumiamo che entrambi i k_{ci} siano non nulli, le equazioni per E_z e H_z nelle tre regioni, le due possibili condizioni all'infinito (modi guidati o irradianti-evanescenti) per E_z e H_z e le condizioni di interfaccia, sempre per E_z e H_z , coinvolgono sempre una sola componente assiale; inoltre si tratta in ogni caso di vincoli lineari e omogenei.

Allora, dalla conoscenza del modo ibrido è possibile dedurre due modi, un TE e un TM caratterizzati, rispettivamente, da ω , γ , H_z , $E_z = 0$ o ω , γ , $H_z = 0$, E_z . Si tratta di modi degeneri (a ω fissata hanno lo stesso γ) la cui sovrapposizione determina il modo ibrido di partenza.

Viceversa, noti due modi indipendenti da y , un TE e un TM, che alla pulsazione angolare di interesse sono degeneri, si possono verificare contemporaneamente le equazioni, le due possibili condizioni all'infinito (modi guidati o irradianti-evanescenti) e le condizioni di interfaccia, visto che questi vincoli su E_z e H_z coinvolgono sempre una sola delle componenti assiali. Allora, nuovamente, la sovrapposizione dei modi indicati determina un modo ibrido.

2.2 Ortogonalità e completezza dei modi: eccitazione dei campi e propagazione in termini di modi

Considerando un ordinamento unico per i modi guidati pari e dispari ($n = 0$ per il modo fondamentale $TE_0^{(e)}$, $n = 1$ per il primo modo superiore $TE_1^{(o)}$, $n = 2$ per il secondo modo superiore $TE_2^{(e)}$, e così via), si può verificare la validità della seguente relazione [6] (pp. 324-325):

$$\frac{\gamma_m}{2\omega\mu_0} \int_{-\infty}^{\infty} \frac{1}{\mu_r} E_{y,m}(x) E_{y,n}^*(x) dx = P\delta_{mn}, \quad (2.119)$$

dove δ_{mn} è la delta di Kronecker e P è la potenza del modo TE quando $m = n$. $E_{y,m}(x)$ rappresenta la componente y del campo elettrico per il modo TE guidato individuato dal pedice m , alla pulsazione ω , e γ_m rappresenta, coerentemente, la costante di propagazione del modo alla stessa pulsazione. Il parametro ω è implicitamente fissato e uguale per tutte le quantità coinvolte.

La stessa relazione vale se si considera un modo guidato e un modo irradiante o evanescente. In particolare, indicando con $E_{y,\tilde{k}_{c1}}(x)$ la componente y del campo elettrico per il modo TE irradiante o evanescente individuato da $k_{c1} = \tilde{k}_{c1}$, si ottiene

$$\int_{-\infty}^{\infty} \frac{1}{\mu_r} E_{y,m}(x) E_{y,\tilde{k}_{c1}}^*(x) dx = 0. \quad (2.120)$$

Per due modi TE irradianti o evanescenti caratterizzati da $k_{c1} = \bar{k}_{c1}$ e $k_{c1} = \tilde{k}_{c1}$ vale una relazione di ortogonalità analoga:

$$\frac{\gamma(\bar{k}_{c1})}{2\omega\mu_0} \int_{-\infty}^{\infty} \frac{1}{\mu_r} E_{y,\bar{k}_{c1}}(x) E_{y,\tilde{k}_{c1}}^*(x) dx = P\delta(\bar{k}_{c1} - \tilde{k}_{c1}). \quad (2.121)$$

Osservazione 14 Negli spazi di funzioni di una sola variabile, definite in \mathbb{R} e che risultano essere a quadrato integrabile, si definisce un prodotto scalare tra le funzioni f e g dello spazio come segue

$$(f, g) = \int_{-\infty}^{\infty} p(x) f(x) g^*(x) dx, \quad (2.122)$$

se p è una funzione a valori reali e positivi. Ciò fa capire perchè le proprietà integrali precedenti dei modi sono indicate con il termine “ortogonalità dei modi”: i prodotti scalari tra modi diversi sono sempre nulli e quindi tali modi sono ortogonali secondo il prodotto scalare considerato. Nelle nostre espressioni la funzione di pesatura p è pari a $\frac{1}{\mu_r}$ e risulta essere a valori reali e positivi.

Osservazione 15 La relazione di ortogonalità (2.121) si può comprendere tenendo presente l'espressione (2.117) di $E_y(x)$ per i modi irradianti ed evanescenti, in cui compaiono esponenziali di tipo $e^{\pm jk_{c1}x}$ nel rivestimento e nel substrato e $e^{\pm jk_{c2}x}$ nel film, con k_{c1} e k_{c2} reali e positivi, e l'ortogonalità delle funzioni esponenziali con esponente puramente immaginario nota dagli studi sulla trasformata di Fourier (si veda, ad esempio, l'equazione (23) di p. 61 di [8]).

Ogni andamento della componente $E_y(x, z = z_0)$ di un campo elettromagnetico monocromatico nello slab dielettrico simmetrico senza perdite considerato può essere espanso in termini di modi TE come segue

$$E_y(x, z = z_0) = \left(\sum_{n=0}^{N-1} C_n E_{y,n}(x) \right) + \int_0^{\infty} C_{k_{c1}} E_{y,k_{c1}}(x) dk_{c1}, \quad (2.123)$$

dove N rappresenta il numero di modi TE guidati. I coefficienti dell’espansione possono essere determinati usando le relazioni di ortogonalità viste in precedenza e sono dati da

$$C_n = \frac{\gamma_n}{2\omega\mu_0 P} \int_{-\infty}^{\infty} \frac{1}{\mu_r} E_y(x, z = z_0) E_{y,n}^*(x) dx \quad (2.124)$$

e

$$C_{k_{c1}} = \frac{\gamma(k_{c1})}{2\omega\mu_0 P} \int_{-\infty}^{\infty} \frac{1}{\mu_r} E_y(x, z = z_0) E_{y,k_{c1}}^*(x) dx. \quad (2.125)$$

Osservazione 16 *La completezza del set di esponenziali $\{e^{j\omega t}\}_{\omega \in \mathbb{R}}$, che sta alla base della teoria della trasformata di Fourier, e la presenza, nelle espressioni di E_y , degli esponenziali $e^{\pm jk_{c1}x}$ nel rivestimento e nel substrato e $e^{\pm jk_{c2}x}$ nel film, con k_{c1} e k_{c2} reali, rende meno ostica l’espressione (2.123).*

Nota l’espansione della componente $E_y(x, z = z_0)$, la componente stessa rimane determinata per ogni $z > z_0$. Infatti, da tale sezione ogni modo si sviluppa in z secondo quanto prescritto dalla sua costante di propagazione γ . L’espansione per $z = z_1 > z_0$ sarà

$$E_y(x, z = z_1) = \left(\sum_{n=0}^{N-1} C_n E_{y,n}(x) e^{-\gamma_n(z_1-z_0)} \right) + \int_0^{\infty} C_{k_{c1}} E_{y,k_{c1}}(x) e^{-\gamma(k_{c1})(z_1-z_0)} dk_{c1}. \quad (2.126)$$

Osservazione 17 *Per semplicità si è evitato di considerare cosa succede nel caso più generale in cui sono presenti onde progressive e regressive.*

Considerazioni analoghe si applicano per la componente H_x . Le componenti H_y ed E_x possono invece essere espresse in termini di modi TM, con un procedimento analogo.

Le considerazioni precedenti sono fondamentali per determinare l’eccitazione di un campo nella struttura guidante di interesse da parte di una sorgente (ad esempio, un light-emitting diode o un laser diode). Se la sorgente determina una distribuzione di campo che è identica a quella di un modo, solo quel modo verrà eccitato [1] (p. 313). Più in generale, una sorgente ecciterà diversi modi con diverse ampiezze e fasi [1] (p. 313).

Osservazione 18 *Nella precedente analisi non sono stati considerati sviluppi “modali” delle componenti assiali dei campi elettromagnetici su una determinata sezione della struttura guidante perchè la prescrizione di una tale condizione al contorno non consente di individuare il campo elettromagnetico in alcuna regione. Viceversa, la prescrizione della parte trasversale del campo elettrico o di quello magnetico su una sezione (se sono coinvolte solo onde progressive; su due sezioni, nel caso generale) individua univocamente il campo elettromagnetico a valle della sezione stessa (o tra le due sezioni considerate, nel caso generale).*

2.3 Onde (o modi) leaky

Abbiamo visto che un numero finito di modi guidati presenta il fenomeno del taglio (ideale, nella lastra dielettrica senza perdite considerata). Al di sotto della frequenza di taglio i modi guidati perdono la loro individualità. Se si cerca di “seguire” un modo guidato al di sotto della frequenza di taglio cercando le soluzioni dell’equazione di dispersione nel piano complesso (per k_{c2}) e rilassando i vincoli sul decadimento esponenziale del campo guidato nel rivestimento e substrato si ottiene una infinità numerabile di cosiddetti modi “leaky”. A questo proposito si tenga presente che i modi guidati sono in numero finito ma i rami delle funzioni periodiche presenti nelle equazioni di dispersione sono un’infinità numerabile e da ciascuno di essi si deduce, con il procedimento descritto, un modo leaky. Queste onde presentano una attenuazione lungo l’asse z [7] (p. 270) (perchè k_{c2} presenta, almeno vicino al taglio, sia parte reale, simile a quella che aveva appena sopra al taglio, che parte immaginaria non nulla, anche se piccola; ciò implica che k_{c2} abbia una fase piccola ma non nulla e ciò a sua volta implica che γ^2 non può essere reale negativo) e irradiano

potenza verso l'infinito in direzione trasversale (da ciò deriva il nome “onde leaky”, visto che, in termini di propagazione in direzione z , si ottiene una vera e propria perdita di potenza). Nelle vicinanze delle frequenze di taglio questi modi consentono di ottenere ottime approssimazioni del campo di radiazione in termini di uno o di poche onde leaky. Le eventuali perdite rendono graduale il fenomeno del taglio e rendono più credibili i modi leaky [7] (pp. 269-270). Si tenga presente che, a discapito del nome utilizzato, i modi leaky non sono modi [7] (pp. 269-270) e spesso sono indicati con il termine più appropriato di “onde leaky”. Molte altre considerazioni sulle onde leaky possono essere trovate in [9] (pp. 31-43).

2.4 Esercitazione in laboratorio con COMSOL Multiphysics relativa alla eccitazione dei modi di uno slab dielettrico e alla propagazione in termini di modi

L'esercitazione inizia con una breve descrizione del simulatore numerico commerciale COMSOL Multiphysics. Quindi si descrive il modo con cui il simulatore gestisce la simulazione di un problema elettromagnetico: geometria, dielettrici, condizioni al contorno, sorgenti, equazioni, soluzione, presentazione dei risultati.

Gli slab dielettrici sono strutture guidanti trasversalmente illimitate. Il simulatore, invece, richiede che la formulazione del problema sia posta su domini limitati (in questo caso bidimensionali). Per questa ragione e per semplicità considereremo sempre una regione di indagine numerica rettangolare. Si approssimano i problemi di interesse ponendo, sulle pareti laterali (parallele alle interfacce dello slab) introdotte artificialmente, delle condizioni al contorno dette di assorbimento (che assorbono i campi elettromagnetici incidenti, idealmente senza rifletterli, evitando di far percepire la presenza di queste parti fittizie del contorno). Tali condizioni al contorno si comportano come dei buoni assorbitori quando non sono poste in prossimità delle zone in cui i campi elettromagnetici sono molto intensi. Per questa ragione le pareti laterali vanno sempre poste a una certa distanza dal film. Tale distanza è determinata anche dalle caratteristiche della struttura guidante che si intende simulare. Per evitare di dover gestire domini di indagine numerica trasversalmente molto ampi, che comporterebbero un significativo aumento dei tempi di calcolo, in tutte le simulazioni eviteremo di considerare slab dielettrici debolmente guidanti. Considereremo, in particolare, $n_1 = 1.46$, $n_2 = 2.25$, $a = 2 \cdot 10^{-6}$, $\mu_2 = \mu_1 = \mu_0$ e assumeremo che i dielettrici siano temporalmente non dispersivi. I valori considerati non limiteranno in alcun modo l'importanza dell'esercitazione.

Se si vuole simulare una struttura guidante di estensione infinita nella direzione di propagazione, sulla parete ortogonale alla direzione di propagazione z , posta su $z = z_{\text{massimo}}$, verrà forzata anche in questo caso una condizione al contorno di assorbimento.

Nel primo set di simulazioni si eccitano singoli modi guidati o alcuni modi guidati, forzando, sulla parete posta su $z = z_{\text{minimo}}$, il campo di interesse, ad esempio in termini della componente H_x (tangenziale alla parte di frontiera del dominio qui considerata). Ciò è possibile conoscendo analiticamente gli andamenti delle diverse componenti dei modi. Così facendo si riescono, ad esempio, a mettere in evidenza il diverso confinamento dei campi a diversa frequenza o le diverse velocità di propagazione dei modi. In particolare,

- nel primo gruppo di simulazioni di questo set si eccita il solo modo $\text{TE}_0^{(e)}$, per $f = 10, 11, \dots, 65$ THz e si verifica il diverso confinamento del campo al variare della frequenza (modello COMSOL: slab_dielettrico_mod0.TE0e_1.primaparte.mph)
- nel secondo gruppo di simulazioni di questo set si eccitano i modi $\text{TE}_0^{(e)}$ e $\text{TE}_1^{(o)}$, per $f = 25, 26, \dots, 65$ THz; per tali valori di frequenza entrambi i modi guidati considerati hanno senso e si verifica il diverso pattern di interferenza tra i modi al variare della frequenza (modello COMSOL: slab_dielettrico_mod0.TE0e_TE1o_2.primaparte.mph)

- nel terzo gruppo di simulazioni di questo set si eccitano i modi $TE_0^{(e)}$, $TE_1^{(o)}$ e $TE_1^{(e)}$ per $f = 45, 46, \dots, 65$ THz; per tali valori di frequenza i modi guidati considerati hanno senso e si verifica il complesso pattern di interferenza tra i modi che varia al variare della frequenza (modello COMSOL: slab_dielettrico_modi_TE0e-TE1o-TE1e_3.primaparte.mph)

Nella seconda parte dell'esercitazione si forza la componente E_y con l'andamento caratteristico della componente E_y del modo $TE_0^{(e)}$. In particolare, si considera l'andamento che si ottiene a un particolare valore di frequenza che determina: $E_y = \cos(k_{c2}x)$ per $|x| < a$, $E_y = \frac{\sqrt{2}}{2} e^{Im(k_{c1})(x-a)}$ per $x > a$ e $E_y = \frac{\sqrt{2}}{2} e^{-Im(k_{c1})(x+a)}$ per $x < -a$, con k_{c2} tale che $k_{c2}a = \frac{\pi}{4}$ e $Im(k_{c1}) = -k_{c2}$. Conoscendo il fenomeno del confinamento possiamo dedurre che la componente forzata può eccitare il modo fondamentale a una e una sola frequenza.

Si valutano gli effetti di tale eccitazione a diverse frequenze, a partire da $f = 2$ THz fino a $f = 30$ THz con passo di 2 THz, discutendo l'eccitazione dei diversi modi e la composizione del campo risultante in termini di sovrapposizione di modi guidati ed evanescenti o irradianti. Gli studenti dovrebbero utilizzare i risultati ottenuti con il simulatore per valutare qualitativamente a che frequenza la condizione al contorno forzata eccita prevalentemente il modo fondamentale. Dovrebbero quindi utilizzare il programma che risolve l'equazione di dispersione per verificare la bontà della deduzione qualitativa precedente.

Come ulteriore analisi gli studenti dovrebbero valutare in maggior dettaglio quello che è lecito attendersi a bassa frequenza, ad esempio per $f = 2$ THz. In questo caso, sapendo che per i modi guidati $k_{c2}a \in (0, \frac{\omega a}{c_0} \sqrt{n_2^2 - n_1^2})$ (come visto nella parte dedicata alla risoluzione grafica delle equazioni di dispersione), si deduce $0 < k_{c2}a < \frac{2\pi 2 \cdot 10^{12} \cdot 2 \cdot 10^{-6}}{c_0} \sqrt{2.25^2 - 1.46^2} \simeq 0.14352$. Inoltre, per i modi irradianti $k_{c2}a \in (\frac{\omega a}{c_0} \sqrt{n_2^2 - n_1^2}, \frac{\omega a}{c_0} n_2)$; per il caso considerato, in particolare, $k_{c2}a < \frac{2\pi 2 \cdot 10^{12} \cdot 2 \cdot 10^{-6}}{c_0} \cdot 2.25 \simeq 0.18863$. A questo punto è evidente che l'unico modo guidato presente alla frequenza considerata e l'insieme continuo dei modi irradianti, anch'esso considerato alla frequenza di $f = 2$ THz, non sono sufficienti a generare il campo che viene forzato su una parte del contorno. Da ciò segue immediatamente, tenendo conto della completezza del set modale, che l'eccitazione forzata a $f = 2$ THz comporta la nascita di modi evanescenti.

Quindi si potrebbe cercare di valutare l'attenuazione dei modi evanescenti certamente eccitati e la distanza (normalizzata con a) che comporta un decadimento superiore al 99 per cento. Si tenga presente, in ogni caso, che è estremamente difficile, se non con una analisi di dettaglio, capire quali modi evanescenti siano realmente eccitati.

Il modello COMSOL da considerare è slab_dielettrico_campo_elettrico_forzato_1.secondaparte.mph.

Quindi si passa al caso di eccitazione di un modo guidato superiore, il $TE_1^{(o)}$. Si forza sul contorno la distribuzione di E_y valutata per $\omega = 139 \cdot 10^{12}$ (per questo modo, nello slab considerato, si ha $\omega_c = 137.534 \cdot 10^{12}$). Si valuti quello che succede per $\omega = 139 \cdot 10^{12}$, $\omega = 137.5 \cdot 10^{12}$ e ω ancora inferiore e si consideri la semplificazione che si deduce ragionando in termini di onda leaky. Il modello COMSOL da considerare è slab_dielettrico_campo_elettrico_onda_leaky_3.secondaparte.mph.

Si ripete l'analisi per un campo al contorno “sbagliato” di tipo $E_y = \sin(\frac{\pi}{2a}x)$ nel film, $e^{\frac{x-a}{4a}}$ per $x > a$ e $-e^{\frac{x+a}{4a}}$ per $x < -a$, valutando le conseguenze determinate da tale eccitazione. L'eccitazione non corrisponde all'andamento di E_y di alcun modo, visto che nel film si è forzato l'andamento “al taglio” mentre nel rivestimento e nel substrato si considera un decadimento esponenziale. Il modello COMSOL da considerare in questo caso è slab_dielettrico_campo_elettrico_sbagliato_2.secondaparte.mph.

Un altro caso interessante, che gli studenti potrebbero gestire autonomamente, considera una condizione al contorno che presenta nel film una frequenza spaziale molto più alta di quella consentita dal modo fondamentale.

Quindi si passa a eccitazioni di tipo gaussiano, con gaussiana centrata in prossimità dello slab (slab_dielettrico_campo_elettrico_gaussiano_vicino_centro_1.terzaparte.mph.) o lontana da questo (slab_dielettrico_campo_elettrico_gaussiano_lontano_centro_2.terzaparte.mph.), per capire cosa può succedere nel caso la sorgente non sia allineata con la struttura guidante.

Per capire l'importanza della teoria modale sui circuiti in generale si considerano gli effetti delle discontinuità in un quarto set di simulazioni (slab_dielettrico_con_discontinuita_centrata_modulo_TE0e_quartaparte.mph e slab_dielettrico_con_discontinuita_non_centrata_modulo_TE0e_quartaparte.mph).

2.5 Caratteristiche delle guide dielettriche più importanti

Le guide dielettriche più importanti sono le fibre ottiche, soprattutto per le trasmissioni a grande distanza, che storicamente possono essere classificate in step-index e graded-index, le “holey fibers” e le fibre a cristallo fotonico (photonic crystal fibers), e, infine, le guide dielettriche per l'ottica integrata (come la embedded strip o la raised strip; si veda la figura 2.1).

L'importanza dei modi, l'ortogonalità e la completezza del set modale rimane la stessa, anche se le loro caratteristiche e le ultime proprietà indicate verranno definite con espressioni più complicate. Con ciò si capisce che il materiale presentato per la lastra dielettrica simmetrica costituisce in realtà il fondamento di una teoria generale della propagazione guidata.

2.5.1 Fibre ottiche

La geometria a cui si fa riferimento per queste strutture guidanti è cilindrica a base circolare, costituita da almeno tre cerchi concentrici. Il cilindro centrale viene chiamato “core” (o nucleo) ed in tutte le applicazioni è il mezzo con l'indice di rifrazione più alto. All'esterno del nucleo abbiamo il cosiddetto “cladding” (o rivestimento o mantello). Il mantello è poi ricoperto con il cosiddetto “jacket” o “coating” (o guaina). Se la fibra fosse utilizzata in modo isolato si potrebbe considerare un modello con, all'esterno del jacket, uno strato infinito di aria o vuoto. Molte fibre, in realtà sono utilizzate all'interno di cavi, in cui possono essere presenti più fibre e altre strutture cilindriche aventi lo scopo di irrobustire meccanicamente il cavo stesso.

In pratica, per quel che riguarda i modi guidati, per frequenze non troppo prossime alle frequenze di taglio, il cladding è così esteso che il modello che viene considerato negli sviluppi analitici è composto dal core e da un cladding di estensione trasversale infinita.

Se le due regioni sono omogenee si parla di fibre ottiche step-index. In realtà, spesso con tale termine si fa riferimento alla sola omogeneità del core, dando per scontata l'omogeneità del cladding. Viceversa, una “graded-index optical fiber” (fibra ottica a gradiente di indice [7] (p. 287)) presenta un indice di rifrazione nel core che varia con continuità, tipicamente riducendosi a mano a mano che ci si allontana dal centro.

Il modello più semplice di una fibra ottica step-index è completamente definito dai valori dell'indice di rifrazione del mantello e del nucleo e dal diametro di quest'ultimo. Di solito l'indice di rifrazione del nucleo è contenuto nel range $[1.44, 1.46]$ [1] (p. 327) e il valore del fractional-refractive index change è contenuto nel range $[0.001, 0.02]$ [1] (p. 327) (il range di n del cladding viene individuato di conseguenza). I diametri del core di alcune delle fibre ottiche step-index standardizzate sono (in μm): 8, 50, 62.5, 85, 100 [1] (p. 327). I corrispondenti diametri del cladding sono (sempre in μm): 125, 125, 125, 125, 140 [1] (p. 327).

Osservazione 19 Negli ultimi due o tre dei cinque casi indicati l'assunzione di avere un cladding di dimensione trasversale infinita non è verificata, se non per frequenze che garantiscono un confinamento significativo dei modi guidati di interesse.

Cenni sui modi delle fibre ottiche step-index

I modi TEM non esistono poichè l'indice di rifrazione del core è diverso da quello del cladding. Rimangono i modi TE, TM e ibridi. Anche in questo caso i modi TM si possono dedurre dai modi TE per dualità, come abbiamo visto per la lastra dielettrica. I modi ibridi, invece, non possono essere considerati ridondanti, come accadeva nello slab, considerandoli una sovrapposizione di modi TE e TM degeneri, perchè la continuità di E_φ e H_φ comporta vincoli espressi in termini di E_z e H_z che non risultano soddisfatti se si annulla una delle due componenti, se non nel caso in cui E_z e H_z risultino essere indipendenti dalla coordinata angolare. Ad esempio, il modo guidato

fondamentale, noto come modo ibrido HE_{11} , non presenta una frequenza di taglio (come il modo fondamentale della lastra dielettrica) e sotto la frequenza di taglio dei primi modi guidati superiori, il TE_{01} e il TM_{01} , che risultano essere non degeneri per tutte le frequenze superiori alla frequenza di taglio [7] (p. 275 e p. 272), non esiste alcun modo guidato TE o TM. Per completezza, si tenga presente che la frequenza di taglio dei modi TE_{01} e TM_{01} è pari a $\frac{\omega}{c_0} a \sqrt{n_2^2 - n_1^2} = 2.405$, dove a indica il raggio del core, n_2 il suo indice di rifrazione, n_1 quello del cladding [7] (p. 275 e p. 272). Può valer la pena ricordare che per la lastra dielettrica simmetrica senza perdite il primo modo superiore, $TE_1^{(o)}$, presenta il taglio per $\frac{\omega}{c_0} a \sqrt{n_2^2 - n_1^2} = \frac{\pi}{2}$, dove a , in questo caso, è pari a metà dello spessore della lastra.

Anche per le fibre ottiche step-index tutte le componenti di campo dei modi sono esprimibili a variabili separate.

Per le tre classi di modi (TE, TM e ibridi) esistono modi guidati, irradianti e evanescenti. Nel seguito forniamo qualche ulteriore considerazione sui soli modi guidati.

I modi guidati nella fibra ottica step-index sono sempre in numero finito, per ogni frequenza fissata. Tali modi, come visto in precedenza, sono individuati da due pedici interi, m e n . Il primo dei due, m , definisce la dipendenza di tutte le componenti del modo dalla coordinata angolare φ : essa sarà data da $\cos(m\varphi)$, $\sin(m\varphi)$ o $e^{\pm jm\varphi}$. I modi TE e TM esistono solo per $m = 0$. I modi ibridi solo per $m \neq 0$. Il pedice $n \geq 1$ descrive la variazione delle componenti in direzione radiale. In particolare, $|e_z|$ e $|h_z|$ presentano n massimi ciascuno in direzione radiale, inclusi gli estremi, se esistono, nell'origine.

I modi ibridi EH sono così chiamati perchè la loro pulsazione di taglio viene definita da una formula $J_m(\frac{\omega}{c_0} a \sqrt{n_2^2 - n_1^2}) = 0$ che coincide con quella che definisce la pulsazione di taglio dei modi TE di una guida d'onda circolare (data dalla condizione $J_m(\frac{\omega}{c_0} na) = 0$). Una sigla alternativa per i modi TM è E (facendo riferimento alla componente assiale non nulla anzichè a quella nulla come nella denominazione usuale). Per questo motivo i modi ibridi indicati sono individuati dalla sigla EH. Esistono due classi distinte di modi ibridi guidati. Alla classe dei modi ibridi non individuata con la sigla EH viene assegnata la sigla HE.

Si noti, infine, che tutti i modi ibridi definiscono intrinsecamente uno spazio vettoriale di modi di dimensione due. Per la simmetria azimutale della struttura, ogni modo che presenta una dipendenza angolare di tipo $\cos(m\varphi)$ con $m \neq 0$, può presentare anche la dipendenza di tipo $\sin(m\varphi)$. Questa è una proprietà vera in generale, in tutte le strutture guidanti che presentano una simmetria di rotazione. E' evidente che il discorso non si applica ai modi TE_{0n} e TM_{0n} , che sono indipendenti da φ (e una rotazione non determina un modo indipendente).

Il modo dominante HE_{11} presenta nel nucleo parti trasversali dei campi elettrico e magnetico che sono mutuamente ortogonali e, con ottima approssimazione, polarizzati linearmente (convenzionalmente in direzione x e y).

Quando il fractional refractive-index change è piccolo, come accade in tutte le applicazioni pratiche, le parti indicate sono approssimativamente polarizzate linearmente anche nel mantello. In caso contrario, quest'ultima proprietà non è verificata.

Sempre nel caso di fibre debolmente guidanti i primi modi superiori, TE_{01} , TM_{01} e HE_{21} , presentano delle curve di dispersione praticamente coincidenti. Ciascuno di questi modi non è linearmente polarizzato ma, nelle condizioni viste, si possono definire delle combinazioni lineari di modi che formano campi (che risultano essere approssimativamente) polarizzati linearmente nel nucleo. Ciò si verifica anche per i successivi gruppi di modi superiori. In conclusione, i modi delle fibre ottiche step-index debolmente guidanti sono approssimativamente polarizzati linearmente o è possibile combinare due modi quasi degeneri in modo che formino dei campi linearmente polarizzati, almeno nel nucleo. Questi campi (che non sono modi, formalmente), vengono chiamati modi LP (acronimo di Linearly Polarized). Esistono i modi LP_{mn} , $m, n \in \mathbb{N}$, $n > 0$. Il modo fondamentale è LP_{01} (due modi degeneri HE_{11}). I primi modi superiori sono LP_{11} (quattro modi degeneri); seguono gli LP_{21} (quattro modi degeneri) e gli LP_{02} (due modi degeneri) [10] (pp. 231-243).

Osservazione 20 Il significato dei pedici m e n per i modi LP è lo stesso di prima ma bisogna tener presente che per i modi LP si fa riferimento alla componente dominante del campo trasversale

(ad esempio E_x per il modo LP_{01} polarizzato lungo x). Nel caso dei modi TE , TM o ibridi, invece, i pedici descrivono il comportamento di tutte le componenti cilindriche dei campi. Per queste ragioni, non è assurdo poter ottenere un modo LP_{01} (con $m = 0$) da due modi HE_{11} (con $m = 1$).

Esercizio 5 Trovare la banda monomodale per una fibra ottica “step-index” caratterizzata da $n_2 = 1.45$, $\Delta = 0.001$ o $\Delta = 0.02$, $a = 4 \mu\text{m}$ o $a = 42.5 \mu\text{m}$. Valutare ω_c , f_c e $\lambda_{0,c}$ (pulsazione, frequenza e lunghezza d’onda nel vuoto al taglio) del primo modo superiore e gli effetti del parametro a e dell’apertura numerica.

Cenni sulle fibre ottiche “graded-index”

In passato, quando si producevano fibre ottiche di scarsa qualità e con dimensioni del nucleo significative (multimodali) aveva un senso ottimizzare il comportamento dell’indice di rifrazione del nucleo per migliorare le prestazioni, ad esempio in termini di dispersione (si veda la sezione 2.7). Al giorno d’oggi, con il significativo miglioramento dei processi produttivi, si producono fibre ottiche con un diametro del core sufficiente piccolo da essere utilizzate nelle bande di funzionamento monomodale (si usa dire, per brevità, fibra ottica monomodale o unimodale (in inglese, “single-mode fiber”)); ad esempio la fibra ottica step-index standardizzata con diametro del core da $8 \mu\text{m}$. Anche al giorno d’oggi, vengono studiate fibre ottiche “graded-index” che consentono di [1] (p. 356):

- azzerare la dispersione cromatica per due valori di lunghezza d’onda riducendo la dispersione per lunghezze d’onda intermedie, come nel caso delle cosiddette “dispersion-flattened fibers”
- indurre una dispersione opposta a quella dovuta alla usuali fibre ottiche step-index, definendo delle fibre ottiche a compensazione della dispersione (“dispersion compensating fibers”); queste fibre, collegate in serie a un tratto di fibra usuale, consente di azzerare la dispersione del tratto composto dai due tipi di fibre collegate in cascata.

Si tenga presente che queste fibre ottiche non vengono utilizzate su larga scala perchè la loro realizzazione (per modificare opportunamente l’andamento dell’indice di rifrazione) comporta l’uso di dopanti (ad esempio il titanio, il germanio o il boro) che tipicamente fanno salire i valori della costante di attenuazione dei modi guidanti di interesse.

Cenni su “holey fibers” e “photonic crystal fibers”

Principio di funzionamento delle fibre bucherellate: indice di rifrazione medio del mantello; “pattern” dei bucherelli può essere in linea di principio ottimizzato per ottenere delle prestazioni particolari (banda monomodale molto ampia (“endless single-mode guidance” [1] (p. 361)) o fibre con la capacità di gestire potenze del segnale significative (“large mode-area” [1] (p. 361))).

Principio di funzionamento delle fibre a cristallo fotonico: propagazione in strutture periodiche (modi di Bloch-Floquet; bande in cui non è possibile la propagazione (“photonic bandgaps” o “forbidden bands”); proprietà per fotoni analoga a quella degli elettroni in un reticolo cristallino (banda energetica tra la banda di valenza e quella di conduzione, ad esempio), tenendo presente, come vedremo, che una banda di frequenze per i fotoni corrisponde a una banda di energie; da ciò segue il nome “photonic crystal fibers”). Potenzialità significative grazie alla possibilità teorica di ottimizzare la periodicità dei fori nel mantello. [1] (p. 244).

Cenni sulle guide dielettriche per l’ottica integrata

Le più importanti guide dielettriche per l’ottica integrata sono la “embedded strip” e la “raised strip” (si veda la figura 2.1). Altre varianti sono le cosiddette “rib” o “ridge” e la “strip loaded” [1] (p. 310). Al giorno d’oggi venfono fabbricate diffondendo del titanio in un substrato di niobato di litio (LiNbO_3), al fine di aumentare il valore dell’indice di rifrazione nella zona della “strip”. Le guide dielettriche di questo tipo vengono anche realizzate con materiali semiconduttori, utilizzando una ben consolidata tecnologia [1] (p. 310).

2.6 Attenuazione nelle guide dielettriche

Da un punto di vista tecnico, per semplici campi monocromatici come le onde piane uniformi o i modi di una struttura guidante, la gestione dell’attenuazione è molto semplice. Infatti i campi indicati presentano un fattore di attenuazione pari a $e^{-\alpha z}$ (consideriamo gli effetti sulle sole onde progressive, caratterizzate dal fattore di propagazione $e^{-\gamma z}$; le considerazioni che seguono si replicano senza alcuna difficoltà per le onde regressive). In potenza il corrispondente fattore è pari a $e^{-2\alpha z}$. Ciò significa che se indico con $P(z)$ la densità di potenza attiva per unità di area del campo considerato si avrà

$$P(z) = P(0)e^{-2\alpha z}. \quad (2.127)$$

E’ noto che l’unità di misura del coefficiente di attenuazione α è m^{-1} .

Il fattore di attenuazione può ridurre in modo molto significativo l’ampiezza del campo elettromagnetico. In questi casi spesso si usano unità logaritmiche per valutarne l’effetto. In particolare, si definisce il cosiddetto coefficiente di attenuazione α_l come segue:

$$\alpha_l = 10 \log_{10} \left(\frac{P(z=0)}{P(z=l)} \right) = 10 \log_{10} (e^{2\alpha l}) = \alpha l 20 \log_{10}(e) \simeq \alpha l 8.685889638 \quad [\text{dB}]. \quad (2.128)$$

Sono comunemente usati i coefficienti di attenuazione per l pari a 1 cm ($\alpha_{1\text{ cm}}$ [dB]; ad esempio nell’ottica integrata), a 1 m o a 1 km ($\alpha_{1\text{ km}}$ [dB]; ad esempio per le fibre ottiche; valori tipici per le moderne fibre monomodali $\alpha_{1\text{ km}} \leq 0.2$ [dB]).

Esercizio 6 *Valutare il coefficiente di attenuazione di un tratto di 100 km di fibra ottica in “fused silica glass” che per $\lambda_0 = 1.55$ [μm] presenta $\alpha_{1\text{ km}} \leq 0.15$ [dB]. Valutare la costante di attenuazione.*

Esercizio 7 *Tenendo presente i valori su scala lineare corrispondenti ai valori notevoli di 0, 3 e 10 [dB], determinare i valori su scala lineare corrispondenti a ogni possibile valore intero di [dB].*

Da un punto di vista fisico, l’attenuazione viene determinata da fenomeni distinti: l’assorbimento e lo “scattering”. L’assorbimento determina la “sparizione” di fotoni, che cedono la propria energia al materiale, dove viene convertita in energia termica (vibrazionale di atomi e molecole), meccanica (ad esempio, di rotazione o deformazione di molecole o cristalli), di altra natura (ad esempio, acustica o chimica) o in energia assorbita da atomi e molecole che subiscono una transizione verso stati energetici più elevati (che può innescare successivamente tutti le conversioni indicate o fenomeni di luminescenza). Nella visione campistica classica si ottiene un passaggio di energia dai campi al materiale e di solito non ci si preoccupa di come tale energia si manifesti nel materiale stesso. Lo scattering, nelle sue versioni più semplici, non determina uno scambio di energia duraturo tra fotoni (campo elettromagnetico) e materia ma alcuni fotoni (il campo elettromagnetico) vengono assorbiti e reirradiati in direzioni diverse da quella di interesse (il campo elettromagnetico viene diffuso in diverse direzioni) e, ai fini della propagazione in una direzione specifica, ciò determina una riduzione dell’intensità del campo elettromagnetico assimilabile ad una attenuazione.

Da un punto di vista quantitativo, questi fenomeni hanno effetti che dipendono dal materiale. Ad esempio per il SiO_2 con cui si realizzano comunemente le fibre ottiche monomodali, l’assorbimento con conversione termica dell’energia avviene fino a frequenze nella banda dell’infrarosso mentre l’assorbimento con transizione verso stati energetici più elevati si manifesta, in modo sempre più significativo, a partire dalle frequenze nella banda della radiazione ultravioletta. Lo scattering di Rayleigh è lo scattering elastico (senza assorbimento di energia) di un campo elettromagnetico provocato da particelle piccole rispetto alla lunghezza d’onda del campo stesso. In termini di intensità (densità di potenza), lo scattering di Rayleigh è proporzionale a $\frac{1}{\lambda_0^4}$.

Esercizio 8 *Gli isolanti sono caratterizzati da un’ampiezza della banda proibita superiore a 3 eV [1] (p. 493). Si valuti il corrispondente valore di frequenza minima di un fotone che consente a un elettrone della banda di valenza di passare nella banda di conduzione.*

Esercizio 9 Per il SiO_2 determinare l'andamento grafico del coefficiente di attenuazione (in dB/km) in funzione di λ_0 , considerando gli assorbimenti in banda UV e IR e lo scattering di Rayleigh.

Esercizio 10 Utilizzando il fenomeno dello scattering di Rayleigh e le sue proprietà, trovare una giustificazione per il colore azzurro del cielo nelle giornate di sole.

Esercizio 11 Utilizzando il fenomeno dello scattering di Rayleigh e le sue proprietà, trovare una giustificazione per il colore rosso del cielo in direzione del sole al tramonto, sempre nelle giornate di sole.

In aggiunta a queste cause di attenuazione, possono essere presenti ioni OH o droganti che aumentano l'assorbimento o lo scattering. Ad esempio, l'assorbimento degli ioni OH ha giocato un ruolo significativo in passato e tuttora le lunghezze d'onda di utilizzo delle fibre ottiche vengono determinate per evitare l'assorbimento da ioni OH (fibre operanti in seconda finestra, intorno ai 1310 nm, o in terza finestra, intorno ai 1550 nm). I dopanti, invece, in generale aumentano le perdite per scattering.

Altre cause di attenuazione sono le randomiche variazioni della sezione delle fibre ottiche determinate dai processi di produzione e le necessarie curvature introdotte nella posa dei cavi di fibre (eccitazione di modi irradianti o evanescenti che comportano una riduzione della potenza dei modi guidati).

Il valore di attenuazione, inoltre, dipende tipicamente dal modo guidato considerato, visto che le diverse distribuzioni del campo determinano diversi effetti in generale.

Per ridurre l'attenuazione, si stanno studiando materiali con un assorbimento di tipo termico che inizia a più bassa frequenza (ad esempio a 2600 nm anziché a 1600 nm) [1] (pp. 350-351).

2.7 Dispersione nelle guide dielettriche

Con il termine “dispersione” si intende far riferimento al fatto che la velocità di propagazione delle onde elettromagnetiche in generale dipende dalla loro frequenza. Il termine è chiaramente correlato alle cosiddette equazioni di dispersione, ai diagrammi di dispersione, ai mezzi dispersivi (temporalmente).

Per capire il fenomeno della dispersione consideriamo, come al solito, un caso semplice: la propagazione di un'onda piana uniforme in un mezzo omogeneo infinito privo di perdite (che sono state considerate nel paragrafo 2.6) e con parametri costitutivi ε_r e μ_r che risultano essere funzioni della frequenza (o se si vuole con indice di rifrazione $n = n(\omega)$). L'onda piana considerata è polarizzata linearmente e possiamo ragionare su quello che succede, propagandosi, all'unica componente del campo elettrico o di quello magnetico. Possiamo quindi considerare una quantità scalare che si propaga con $\beta(\omega) = \frac{\omega}{c_0} n(\omega)$.

Le valutazioni effettuate nel caso di onda piana uniforme potranno essere estese alla propagazione di ogni singolo modo di una struttura guidante, pur di tener conto del corretto andamento di $\beta(\omega)$ e del fatto che l'andamento dei campi di ogni singolo modo dipende da ω (cosa che non succede per le onde piane uniformi). Queste difficoltà verranno affrontate dopo aver capito quali sono gli effetti della dispersione sulla propagazione delle onde piane uniformi.

Consideriamo quindi una quantità scalare $u(z, t)$ e supponiamo che in $z = 0$ tale segnale sia generato da un modulatore di ampiezza. Tale segnale è certamente di interesse nelle comunicazioni ottiche e, più in generale, nelle comunicazioni digitali. Inoltre presenta uno spettro che da un lato non è banale al punto da non consentire di mettere in evidenza il fenomeno di interesse e dall'altro non complica troppo l'analisi del fenomeno stesso.

Avremo

$$u(z = 0, t) = a(t)e^{j\omega_0 t}, \quad (2.129)$$

dove $a(t)$ è il segnale modulante e ω_0 è la pulsazione del segnale modulato.

Osservazione 21 Considereremo genericamente segnali complessi per evitare di appesantire troppo la notazione. Una modulazione più realistica, utilizza, ad esempio, il segnale modulato $\cos(\omega_0 t) = \frac{e^{j\omega_0 t} + e^{-j\omega_0 t}}{2}$, e, come è evidente, saremmo costretti ad usare una notazione più pesante.

Vogliamo capire cosa diventa il segnale $u(z = l, t)$ per $l > 0$. Mettiamo subito in evidenza che non è banale trovare una risposta. Se si cercasse di comprendere il fenomeno interamente nel dominio del tempo sarebbe necessario trovare la risposta all'impulso $h(z = l, t)$ quando $u(z = 0, t) = \delta(t)$ e poi fare la solita convoluzione. Purtroppo, non è affatto semplice trovare $h(z = l, t)$. Anzi, uno dei modi di farlo è proprio quello di ricorrere all'analisi nel dominio della frequenza. Per questa ragione l'approccio più conveniente al problema è quello che effettua l'analisi nel dominio della frequenza. In effetti sappiamo che ogni tono sinusoidale di pulsazione $\bar{\omega}$ del segnale $u(z = 0, t)$ si propaga con il fattore di propagazione $e^{-j\beta(\bar{\omega})z}$. Utilizzando la trasformata di Fourier, decomponiamo $u(z = 0, t)$ in toni sinusoidali, determiniamo, per ciascuno di questi, l'effetto della propagazione sulla tratta $(0, l)$, e, con la trasformata inversa di Fourier, ricostruiamo il segnale $u(z, t)$.

Sia $F(\omega)$ la trasformata di Fourier del generico segnale $f(t)$:

$$F(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt, \quad (2.130)$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) e^{j\omega t} d\omega, \quad (2.131)$$

Avremo:

$$U(z = 0, \omega) = \int_{-\infty}^{+\infty} u(z = 0, t) e^{-j\omega t} dt = \int_{-\infty}^{+\infty} a(t) e^{-j(\omega - \omega_0)t} dt = A(\omega - \omega_0), \quad (2.132)$$

$$U(z = l, \omega) = U(z = 0, \omega) e^{-j\beta(\omega)l} = A(\omega - \omega_0) e^{-j\beta(\omega)l}, \quad (2.133)$$

$$u(z = l, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} U(z = l, \omega) e^{j\omega t} d\omega = \frac{1}{2\pi} \int_{-\infty}^{+\infty} A(\omega - \omega_0) e^{-j\beta(\omega)l} e^{j\omega t} d\omega. \quad (2.134)$$

In particolare, ponendo $w = \omega - \omega_0$ nell'ultima equazione, si ottiene:

$$u(z = l, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} A(w) e^{-j\beta(w + \omega_0)l} e^{j(w + \omega_0)t} dw. \quad (2.135)$$

Consideriamo il seguente sviluppo di Taylor della funzione $\beta(\omega)$

$$\beta(\omega) = \beta(\omega_0) + \beta'(\omega_0)(\omega - \omega_0) + \frac{1}{2}\beta''(\omega_0)(\omega - \omega_0)^2, \quad (2.136)$$

che fornisce risultati soddisfacenti quando i termini successivi a quelli considerati non danno contributi significativi per il segnale modulante $a(t)$ di interesse (di banda limitata).

Esercizio 12 Dedurre le unità di misura di $\beta'(\omega_0)$ e $\beta''(\omega_0)$.

Sostituendo tale sviluppo nell'equazione (2.135) si ottiene

$$\begin{aligned} u(z = l, t) &= \\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} A(w) e^{-j\beta(\omega_0)l} e^{-j\beta'(\omega_0)wl} e^{-j\beta''(\omega_0)w^2 \frac{l}{2}} e^{j\omega t} e^{j\omega_0 t} dw = \\ &= e^{j\omega_0(t - \frac{\beta(\omega_0)l}{\omega_0})} \frac{1}{2\pi} \int_{-\infty}^{+\infty} (A(w) e^{-j\beta''(\omega_0)w^2 \frac{l}{2}}) e^{jw(t - \beta'(\omega_0)l)} dw = \\ &= e^{j\omega_0(t - \frac{\beta(\omega_0)l}{\omega_0})} a_d(t - \beta'(\omega_0)l), \end{aligned} \quad (2.137)$$

dove a_d è l'antitrasformata di Fourier di $A_d(\omega) = A(\omega) e^{-j\beta''(\omega_0)\omega^2 \frac{l}{2}}$.

Si noti che, in generale, la modulante subisce un ritardo pari a $\frac{\beta(\omega_0)}{\omega_0}l$, detto ritardo di fase. Ciò significa che si propaga con velocità $v_f(\omega_0) = \frac{\omega_0}{\beta(\omega_0)}$, detta velocità di fase. La modulante viene distorta e subisce un ritardo pari a $\beta'(\omega_0)l$, detto ritardo di gruppo. Ciò significa che si propaga con velocità $v_g(\omega_0) = \frac{1}{\beta'(\omega_0)}$, detta velocità di gruppo.

In generale le velocità di fase e di gruppo hanno valori diversi e dipendono dal valore considerato di ω_0 .

Se, in particolare, la derivata seconda valutata in ω_0 è trascurabile, si ottiene $A_d(w) = A(w)$ e, di conseguenza

$$u(z = l, t) = e^{j\omega_0\left(t - \frac{l}{v_f(\omega_0)}\right)} a\left(t - \frac{l}{v_g(\omega_0)}\right). \quad (2.138)$$

In questo caso, quindi, la modulante non viene distorta. Nel caso ancora più particolare in cui si abbia $\beta''(\omega_0) = 0$ e $v_f(\omega_0) = v_g(\omega_0)$ si deduce

$$u(z = l, t) = e^{j\omega_0\left(t - \frac{l}{v_f(\omega_0)}\right)} a\left(t - \frac{l}{v_f(\omega_0)}\right) = u\left(z = 0, t - \frac{l}{v_f(\omega_0)}\right) \quad (2.139)$$

e quindi, in questo caso, il segnale trasmesso, propagandosi, rimane identico a se stesso e subisce un semplice ritardo. In altre parole, in questo caso particolare, non c'è nulla di nuovo rispetto a quello che succede per i segnali puramente sinusoidali.

Esercizio 13 Verificare che il segnale, propagandosi, rimane identico a se stesso quando in un intorno di ω_0 contenente il supporto di $A(\omega - \omega_0)$ si ha $\beta(\omega) = \frac{\beta(\omega_0)}{\omega_0}\omega$.

Osservazione 22 Per un'onda elettromagnetica piana in un mezzo omogeneo senza perdite caratterizzato da $\varepsilon(\omega)$, $\mu(\omega)$, $\sigma(\omega) = 0$, con ε e μ a valori reali, si ha $\beta(\omega) = \omega\sqrt{\mu(\omega)\varepsilon(\omega)} = \frac{\omega}{c_0}n(\omega)$. Per ottenere un'onda "indeformabile" è quindi necessario che nell'intorno di ω_0 contenente la banda del segnale modulato AM il mezzo non sia dispersivo nel tempo.

Nel caso generale, la distorsione subita dalla modulante durante la propagazione è equivalente a quella data da un filtro lineare avente funzione di trasferimento:

$$H(\omega) = e^{-j\beta''(\omega_0)\omega^2 \frac{l}{2}} \quad (2.140)$$

con modulo unitario e fase quadratica in ω .

Osservazione 23 Per i casi in cui il problema di specifico interesse suggerisce l'utilizzo di più termini dello sviluppo di Taylor (2.136), la funzione di trasferimento del filtro lineare da considerare diventa

$$H(\omega) = \prod_{i=2}^N e^{-j\beta^{(i)}(\omega_0)\omega^i \frac{l}{i!}}, \quad N \geq 2, \quad (2.141)$$

con ovvio significato dei simboli. Il modulo della funzione di trasferimento rimane unitario ma il comportamento della fase diventa molto più complesso.

Per capire il fenomeno della dispersione dobbiamo valutare gli effetti distorsivi della propagazione sul segnale modulante. Tale valutazione è complessa, in generale. Tuttavia, in alcuni casi significativi, è possibile portare a termine l'analisi. Le informazioni dedotte da questa analisi potranno essere estrapolate fino ad ottenere informazioni preziose in generale [11] (p. 329).

L'andamento gaussiano della modulante è quello che consente di portare a termine il procedimento senza troppe difficoltà analitiche. Si consideri, in particolare una modulante gaussiana:

$$a(t) = Ce^{-\pi b^2 t^2}, \quad C, b \in \mathbb{R}, b \neq 0. \quad (2.142)$$

Il coefficiente b si misura in $[s^{-1}]$.

La corrispondente trasformata di Fourier [8] (p. 656), anche in presenza di $b \in \mathbb{C}$, $Re(b^2) > 0$, è pari a:

$$A(\omega) = \frac{C}{b} e^{-\frac{\omega^2}{4\pi b^2}}. \quad (2.143)$$

L'effetto della propagazione determina

$$A_d(\omega) = A(\omega) e^{-j\beta''(\omega_0)\omega^2 \frac{l}{2}} = \frac{C}{b} e^{-\frac{\omega^2}{4\pi b^2}} e^{-j\beta''(\omega_0)\omega^2 \frac{l}{2}} = \frac{C_1}{b_1} e^{-\frac{\omega^2}{4\pi b_1^2}} \quad (2.144)$$

avendo definito

$$b_1^2 = \frac{b^2}{1 + j\beta''(\omega_0)2\pi b^2 l} \quad (2.145)$$

e

$$C_1 = C \frac{b_1}{b}. \quad (2.146)$$

Da ciò si deduce

$$a_d(t) = C_1 e^{-\pi b_1^2 t^2}, \quad C_1, b_1 \in \mathbb{C}. \quad (2.147)$$

Quindi, l'effetto della distorsione dovuta alla propagazione trasforma la modulante gaussiana a parametri reali in una modulante gaussiana a parametri complessi.

Esercizio 14 Assumendo $b \in \mathbb{R}$, si verifichi la validità delle seguenti relazioni:

$$Re(b_1^2) = \frac{b^2}{1 + (\beta''(\omega_0)2\pi b^2 l)^2}, \quad (2.148)$$

$$Im(b_1^2) = -\frac{\beta''(\omega_0)2\pi b^4 l}{1 + (\beta''(\omega_0)2\pi b^2 l)^2}, \quad (2.149)$$

$$|b_1^2| = \frac{b^2}{\sqrt{1 + (\beta''(\omega_0)2\pi b^2 l)^2}}, \quad (2.150)$$

$$\varphi_{b_1^2} = -\arctan(\beta''(\omega_0)2\pi b^2 l), \quad (2.151)$$

$$|C_1| = \frac{|C|}{(1 + (\beta''(\omega_0)2\pi b^2 l)^2)^{\frac{1}{4}}}, \quad (2.152)$$

$$\varphi_{C_1} = \varphi_C - \frac{1}{2} \arctan(\beta''(\omega_0)2\pi b^2 l) \quad \text{se } b > 0. \quad (2.153)$$

A questo punto l'effetto della propagazione su un segnale modulato AM con modulante gaussiana è completamente determinato. E' sufficiente sostituire l'espressione di $a_d(t)$ ottenuta dall'equazione (2.147) nell'equazione (2.137) per ottenere $u(z = l, t)$. In particolare,

$$\begin{aligned} u(z = l, t) &= \\ &= |C_1| e^{j\varphi_{C_1}} e^{j\omega_0 \left(t - \frac{l}{v_f(\omega_0)}\right)} e^{-\pi b_1^2 \left(t - \frac{l}{v_g(\omega_0)}\right)^2} = \\ &= |C_1| e^{-\pi Re(b_1^2) \left(t - \frac{l}{v_g(\omega_0)}\right)^2} e^{j\varphi_{C_1}} e^{j\omega_0 \left(t - \frac{l}{v_f(\omega_0)}\right)} e^{-j\pi Im(b_1^2) \left(t - \frac{l}{v_g(\omega_0)}\right)^2}. \end{aligned} \quad (2.154)$$

2.7.1 Effetto della propagazione sulla ampiezza e durata del segnale modulante

Uno dei modi usuali di definire la durata di un segnale è quello che valuta l'intervallo di tempo tra gli istanti a metà potenza del valore massimo, dove per potenza istantanea di un segnale $x(t)$ si intende $p_x(t) = |x(t)|^2$. In alternativa al fattore 1/2 si usa spesso anche il fattore 1/e. Si parla rispettivamente di FWHM_x (full-width at half maximum) o FW1/eM_x (full-width at 1/e maximum).

Esercizio 15 Si consideri $a(t) = Ce^{-\pi b^2 t^2}$, $C, b \in \mathbb{R}$, e si verifichi la validità della relazione

$$FW1/eM_a = \frac{\sqrt{2}}{\sqrt{\pi}|b|}. \quad (2.155)$$

Osservazione 24 Il termine “durata” fa chiaramente riferimento a un segnale nel dominio del tempo. Quando lo stesso concetto si riferisce a uno spettro il termine più utilizzato è “larghezza di banda”.

Si noti che, visto che lo spettro di un segnale gaussiano è a sua volta una gaussiana, l'esercizio 15 e l'equazione (2.143) consentono di dedurre che la larghezza di banda del segnale $a(t)$ è pari a

$$FW1/eM_A = 2\sqrt{2\pi}|b|. \quad (2.156)$$

Le espressioni (2.155) e (2.156) mettono in evidenza il fatto, noto in generale, che la larghezza di banda di un segnale è inversamente proporzionale alla sua durata. In particolare, nel caso considerato si ha $FW1/eM_a \cdot FW1/eM_A = 4$.

Esercizio 16 Si consideri $a_d(t) = C_1 e^{-\pi b_1^2 t^2}$, $C_1, b_1 \in \mathbb{C}$, e si verifichi la validità della relazione

$$\begin{aligned} FW1/eM_{a_d} &= \frac{\sqrt{2}}{\sqrt{\pi}\sqrt{Re(b_1^2)}} = \frac{\sqrt{2}}{\sqrt{\pi}|b|} \sqrt{1 + (\beta''(\omega_0)2\pi b^2 l)^2} = \\ &= FW1/eM_a \sqrt{1 + (\beta''(\omega_0)2\pi b^2 l)^2} = FW1/eM_a \sqrt{1 + \frac{16(\beta''(\omega_0))^2 l^2}{FW1/eM_a^4}}. \end{aligned} \quad (2.157)$$

Quando $\beta''(\omega_0) \neq 0$ si definisce $l_0 > 0$:

$$l_0^2 = \frac{1}{(\beta''(\omega_0)2\pi b^2)^2} = \frac{FW1/eM_a^4}{16(\beta''(\omega_0))^2} = \frac{16}{(\beta''(\omega_0))^2 FW1/eM_A^4}. \quad (2.158)$$

Dal risultato dell'esercizio 16 si deduce quindi

$$FW1/eM_{a_d} = FW1/eM_a \sqrt{1 + \frac{l^2}{l_0^2}}. \quad (2.159)$$

E' evidente che in presenza di dispersione si ha sempre $FW1/eM_{a_d} > FW1/eM_a$ per ogni $l > 0$. Quindi, l'impulso gaussiano, propagandosi, si allarga (e ovviamente accumula ritardo).

Per $l = l_0$ l'allargamento è pari a un fattore $\sqrt{2}$ mentre per $l \gg l_0$ si ottiene $FW1/eM_{a_d} = FW1/eM_a \frac{l}{l_0} = \frac{4l\beta''(\omega_0)}{FW1/eM_a} = l\beta''(\omega_0) \cdot FW1/eM_A$.

Osservazione 25 Tenendo presente che $\beta''(\omega_0)$ interviene nello sviluppo di Taylor di $\beta(\omega)$ attraverso il fattore $(\omega - \omega_0)^2$, è ragionevole che $\beta''(\omega_0)$ distorca il segnale in modo proporzionale $FW1/eM_A$. Questa considerazione è indipendente dal fatto che il segnale modulante sia gaussiano. Si noti che, nel caso particolare di segnale modulato AM con modulante gaussiana, l_0 risulta essere inversamente proporzionale a $|\beta''(\omega_0)| FW1/eM_A^2$.

Dall'equazione (2.152) e dalla definizione di l_0 si deduce inoltre

$$|C_1| = \frac{|C|}{\left(1 + \frac{l^2}{l_0^2}\right)^{\frac{1}{4}}}. \quad (2.160)$$

Si deduce che il fattore che determina l'allungamento della durata dell'impulso gaussiano determina contemporaneamente una riduzione della sua ampiezza (in assenza di attenuazione si conserva la potenza complessiva del segnale).

2.7.2 Effetto della propagazione sulla frequenza istantanea del segnale

La cosiddetta pulsazione istantanea del segnale $u(z = l, t)$ (si veda l'equazione (2.154)) è data [8] (p. 231) dalla derivata nel tempo della fase. Si deduce

$$\omega(t) = \omega_0 - 2\pi \text{Im}(b_1^2) \left(t - \frac{l}{v_g(\omega_0)} \right). \quad (2.161)$$

Ciò significa che il segnale $u(z = l, t)$ è un segnale modulato AM non classico, con una modulante gaussiana e con una portante che risulta avere una pulsazione dipendente dal tempo.

In particolare, se $\text{Im}(b_1^2) < 0$ si ottiene un valore di pulsazione istantanea $\omega(t)$ maggiore (minore) di ω_0 per $(t - \frac{l}{v_g(\omega_0)}) > 0$ ($(t - \frac{l}{v_g(\omega_0)}) < 0$). In questo caso si usa dire che il segnale è “up-chirped” [1] (p. 942). Quando $\text{Im}(b_1^2) > 0$ si ottengono variazioni opposte del valore della pulsazione istantanea e il segnale è detto “down-chirped” [1] (p. 942).

Tenendo conto delle equazioni (2.149) e (2.158) si deduce, se $\beta''(\omega_0) > 0$

$$\text{Im}(b_1^2) = -\frac{b^2 \frac{l}{l_0}}{1 + \frac{l^2}{l_0^2}}. \quad (2.162)$$

Osservazione 26 *Ad esempio, per $l = l_0$ si ottiene $\text{Im}(b_1^2) = -\frac{b^2}{2}$ e per $l = 10l_0$ o $l = \frac{l_0}{10}$ si deduce $\text{Im}(b_1^2) \simeq -\frac{b^2}{10}$.*

Osservazione 27 *Per $t = \frac{l}{v_g(\omega_0)}$ si ottiene sempre la massima ampiezza della modulante gaussiana. Quando ciò si verifica la pulsazione istantanea della portante è sempre pari alla pulsazione della portante in $z = 0$; al contrario, la fase in tali condizioni è pari a φ_{c1} .*

2.7.3 Simulazioni e risultati

Nelle prime simulazioni considereremo $u(z = 0, t) = a(t) \cos(\omega_0 t)$, dove $\omega_0 = 2\pi f_0$ e $f_0 = 100$ [THz], corrispondente a un periodo T_0 pari a 0.01 ps. La modulante gaussiana $a(t)$ è data da $e^{-\frac{(t-100T_0)^2}{50T_0^2}}$, presenta un massimo di ampiezza unitaria ($C = 1$) per $t = 100T_0$ ed è definita in modo da avere una durata FW1/eM_a pari a $10T_0$ ($b^2 = \frac{1}{50\pi T_0^2}$).

Per ricostruire fedelmente i segnali si considera un tempo di campionamento pari a $\frac{T_0}{16}$ (in alternativa $\frac{T_0}{32}$). Per poter gestire tutte le simulazioni di interesse faremo un'analisi considerando 131072 campioni (il corrispondente intervallo temporale analizzato nelle simulazioni è pari a $8192T_0$ (in alternativa $4096T_0$)).

Le simulazioni vengono fatte definendo il segnale a tempo discreto in ingresso. Quindi viene valutata la DFT del segnale e, per ogni valore di frequenza, viene applicato il fattore di propagazione alla DFT del segnale di ingresso, ottenendo la DFT del segnale di uscita (facendo attenzione in modo da ottenere una DFT hermitiana simmetrica). Quindi si applica l'algoritmo di inversione della DFT ottenendo il segnale tempo discreto in uscita (a valle del tratto di propagazione).

Considereremo tre andamenti di $\beta(\omega)$, fissando in ogni caso $\alpha = 0$, come detto in precedenza.

Nel primo caso consideriamo $\beta(\omega) = \frac{1}{c_0}\omega$, ovvero l'andamento di $\beta(\omega)$ tipico delle onde piane uniformi nel vuoto.

I risultati ottenuti in questo caso sono riportati nelle figure 2.14, 2.15 e 2.16.

In particolare, in figura 2.14 viene riportato l'andamento del segnale in ingresso; in figura 2.15 la corrispondente potenza istantanea consente di valutare graficamente la durata del segnale di ingresso; infine, in figura 2.16 viene riportato il segnale di uscita che, dopo aver coperto una distanza pari a $l = 2 \cdot 0.299792458$ [mm] presenta un ritardo, rispetto al segnale di ingresso, di $200T_0$. Si nota che il segnale in uscita è una semplice copia ritardata del segnale in ingresso.

Esercizio 17 *Utilizzando la figura 2.15 si valuti graficamente FW1/eM_a .*

Nel secondo caso viene considerato $\beta(\omega) = \frac{39}{38} \frac{\omega_0}{c_0} + \frac{20}{19} \frac{1}{c_0} (\omega - \omega_0)$.

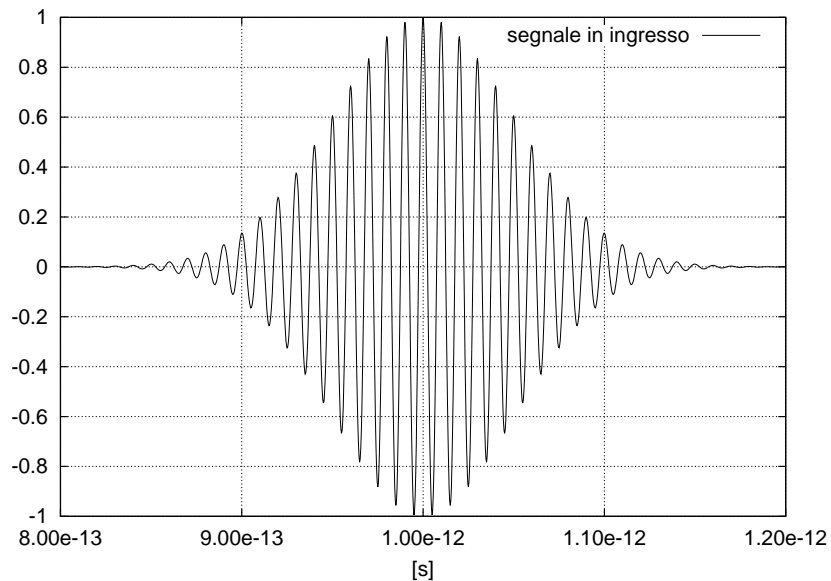


Figura 2.14: Andamento nel tempo del segnale di ingresso. Una portante di frequenza $f_0 = 100$ [THz] viene modulata da un segnale gaussiano.

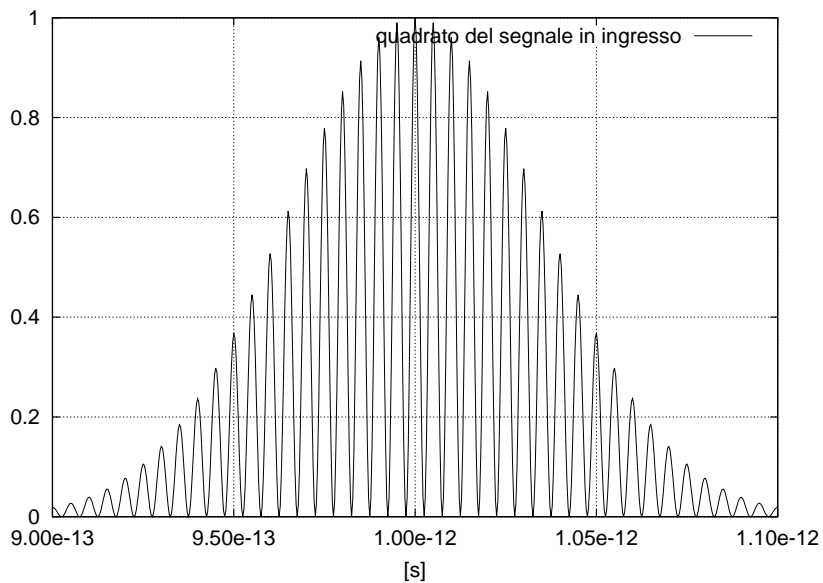


Figura 2.15: Andamento nel tempo della potenza istantanea del segnale di ingresso.

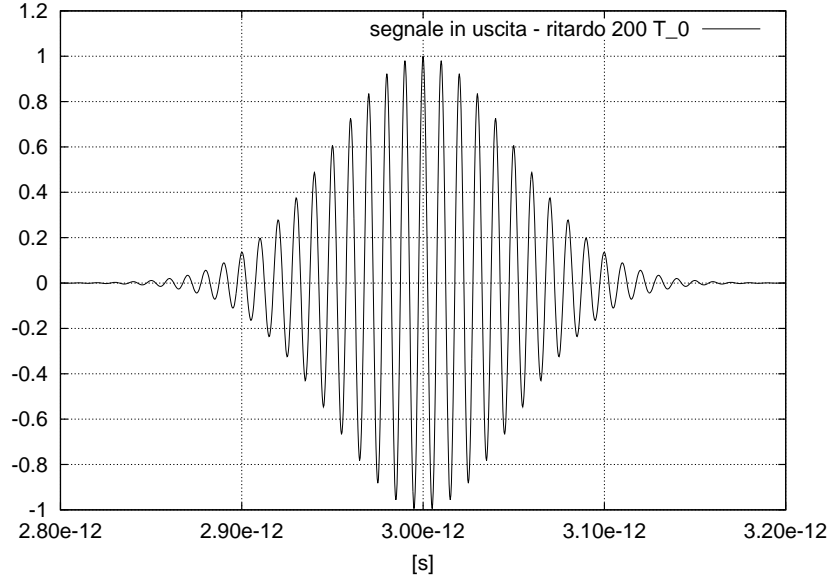


Figura 2.16: Andamento nel tempo del segnale in uscita, dopo aver percorso una tratta che comporta un ritardo di $200 T_0$. Il segnale è stato ottenuto assumendo che la propagazione si verifichi in un mezzo che determina $\beta(\omega) = \frac{1}{c_0}\omega$ e $\alpha = 0$.

Esercizio 18 Verificare che, per l'andamento di $\beta(\omega)$ indicato, si ottiene $v_g = \frac{19}{20} c_0$ e $v_f = \frac{38}{39} c_0$ per $\omega = \omega_0$.

Esercizio 19 Verificare che la distanza minima che comporta una differenza di tempi di propagazione pari a $\frac{T_0}{2}$ per segnali che si propagano con velocità pari rispettivamente a v_1 e v_2 , $v_1 > v_2$, è uguale a $l_m = \frac{T_0}{2} \frac{v_1 v_2}{v_1 - v_2}$. Quindi verificare che $l_m = 19 \lambda_0$ quando $v_1 = \frac{38}{39} c_0$ e $v_2 = \frac{19}{20} c_0$.

I risultati ottenuti in questo caso sono riportati nelle figure 2.17 e 2.18. In particolare, in figura 2.17 (2.18) viene riportato l'andamento del segnale dopo che questo ha percorso una tratta pari a $3 l_m$ ($6 l_m$), dove $l_m = 19 \lambda_0 = 19 \frac{c_0}{f_0} = 19 \cdot 2.99792458 \cdot 10^{-6} = 56.96056702 \text{ } [\mu\text{m}]$. Le figure mettono in evidenza la diversa velocità di propagazione della portante e del segnale modulante. Lo stesso effetto può essere facilmente rilevato considerando una tratta di lunghezza pari a $l = \frac{l_m}{2}$.

Esercizio 20 Che effetto si ottiene considerando $l = \frac{l_m}{2}$?

Nel terzo e ultimo caso consideriamo $\beta(\omega)$ in modo da ottenere $v_f(\omega_0) = v_g(\omega_0) = c_0$ e l_0 pari alla distanza percorsa dalla modulante in $100 T_0$.

Esercizio 21 Verificare che il polinomio di Taylor di secondo grado (2.136) è quasi completamente definito grazie alla conoscenza di ω_0 , v_f , v_g , l_0 e $FW1/eM_a$. L'unico parametro indeterminato è il segno di $\beta''(\omega_0)$.

Esercizio 22 Con i dati a disposizione ricavare $\beta(\omega_0)$, $\beta'(\omega_0)$ e $\beta''(\omega_0)$ ($\simeq 2.09584502 \cdot 10^6$, $\simeq 0.333564095 \cdot 10^{-8}$, $\simeq 8.3391024 \cdot 10^{-24}$).

In figura 2.19 viene riportato l'andamento nel tempo del segnale di uscita, per $l = \frac{l_0}{10}$. L'effetto distorsivo di $\beta''(\omega_0)$ (scelto di segno negativo) su una tale tratta è praticamente trascurabile, come si deduce dal fatto che $\sqrt{1 + \frac{l^2}{l_0^2}} \simeq 1.005$.

Quando $l = l_0$ lo stesso fattore diventa pari a $\sqrt{2} \simeq 1.414$. Tale fattore determina un incremento della durata del segnale modulante in uscita e la riduzione della sua potenza di picco. La figura

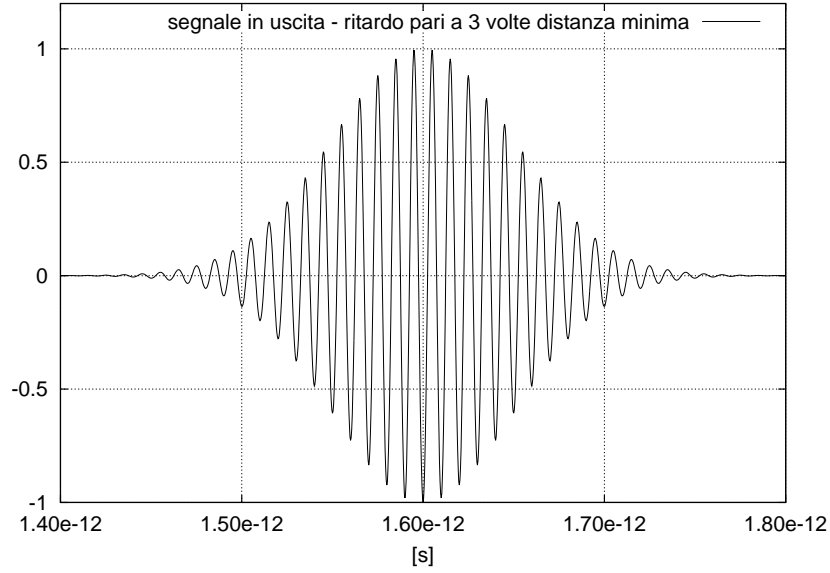


Figura 2.17: Andamento nel tempo del segnale in uscita, dopo aver percorso una tratta pari a $3l_m$. Il segnale è stato ottenuto assumendo che la propagazione si verifichi in un mezzo che determina $\beta(\omega) = \frac{39\omega_0}{38c_0} + \frac{20}{19c_0}(\omega - \omega_0)$ e $\alpha = 0$.

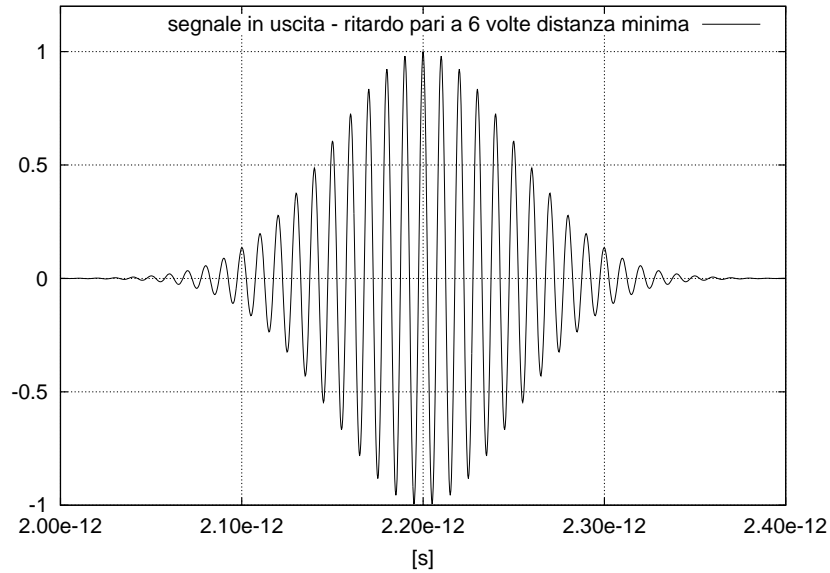


Figura 2.18: Andamento nel tempo del segnale in uscita, dopo aver percorso una tratta pari a $6l_m$. Il segnale è stato ottenuto assumendo che la propagazione si verifichi in un mezzo che determina $\beta(\omega) = \frac{39\omega_0}{38c_0} + \frac{20}{19c_0}(\omega - \omega_0)$ e $\alpha = 0$.

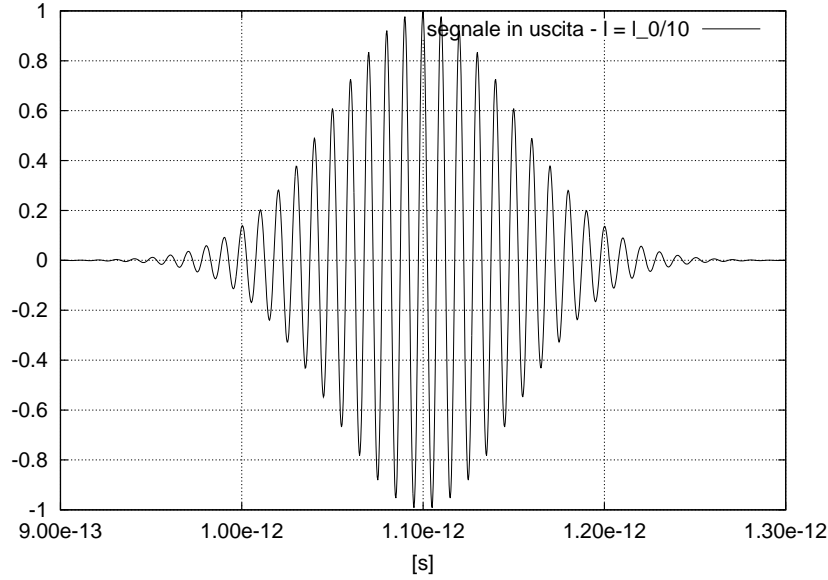


Figura 2.19: Andamento nel tempo del segnale di uscita, per $l = \frac{l_0}{10}$. Il segnale è stato ottenuto assumendo che la propagazione si verifichi in un mezzo che determina $v_f(\omega_0) = v_g(\omega_0) = c_0$, $l_0 = 100 T_0 c_0$, segno di $\beta''(\omega_0)$ negativo e $\alpha = 0$.

2.20, che riporta l'andamento nel tempo del segnale di uscita e della sua potenza, per $l = l_0$, consente di valutare graficamente sia la durata che l'ampiezza in termini di potenza di picco. La stessa figura mette inoltre in evidenza che il segnale in uscita è lievemente “down-chirped” (in linea con l'ipotesi $\beta''(\omega_0) < 0$ che determina $\text{Im}(b_1^2) > 0$). In particolare, dall'equazione (2.162), per $l = l_0$ si deduce $\text{Im}(b_1^2) = \frac{b^2}{2} = \frac{1}{\pi \text{FWI}/eMa^2}$ e, per l'equazione (2.161), considerando $\text{FWI}/eMa = 10 T_0$, la pulsazione istantanea risulta essere pari a $\omega(t) = \omega_0 - \frac{1}{50 T_0^2} \left(t - \frac{l}{v_g(\omega_0)} \right)$.

Esercizio 23 Si valuti la pulsazione istantanea che si ottiene nel caso considerato per diversi valori di $\left(t - \frac{l}{v_g(\omega_0)} \right)$.

Esercizio 24 Si determini il parametro che consente di rendere più evidente la variazione di pulsazione istantanea.

Esercizio 25 In presenza di un segnale modulante gaussiano, si valuti in che modo è necessario modificare il valore del parametro b se si vogliono limitare gli effetti della dispersione su una tratta di lunghezza l sempre maggiore. Si valutino le ricadute in termini di bit al secondo che possono essere trasmessi su tratte sempre più lunghe per limitare gli effetti della dispersione.

Infine, la figura 2.21 riporta l'andamento nel tempo del segnale di uscita e della sua potenza per $l = 10 l_0$. E' possibile, anche in questo caso, valutare graficamente sia la durata che l'ampiezza in termini di potenza di picco. La variazione della pulsazione istantanea risulta un po' meno evidente. Per concludere, le figure 2.22 e 2.23 riportano gli andamenti nel tempo del segnale di ingresso e di uscita che si ottengono quando la modulante è un segnale a trapezio isoscele che mantiene il valore massimo per $10 T_0$ e il cui supporto è pari a $15 T_0$, nel caso in cui il segnale si propaghi su una tratta di lunghezza $2.99792458 \cdot 10^{-4}$. Le caratteristiche di $\beta(\omega)$ sono le stesse del caso precedente e il valore del ritardo di gruppo è coerente con il valore impostato della velocità di gruppo.

Con modulante trapezoidale non si riescono a dedurre i dettagli degli evidenti effetti della dispersione.

Infine, si tenga presente che gli effetti della propagazione sui segnali sono in generale di più difficile interpretazione a causa dei fenomeni dissipativi che determinano $\alpha = \alpha(\omega)$.

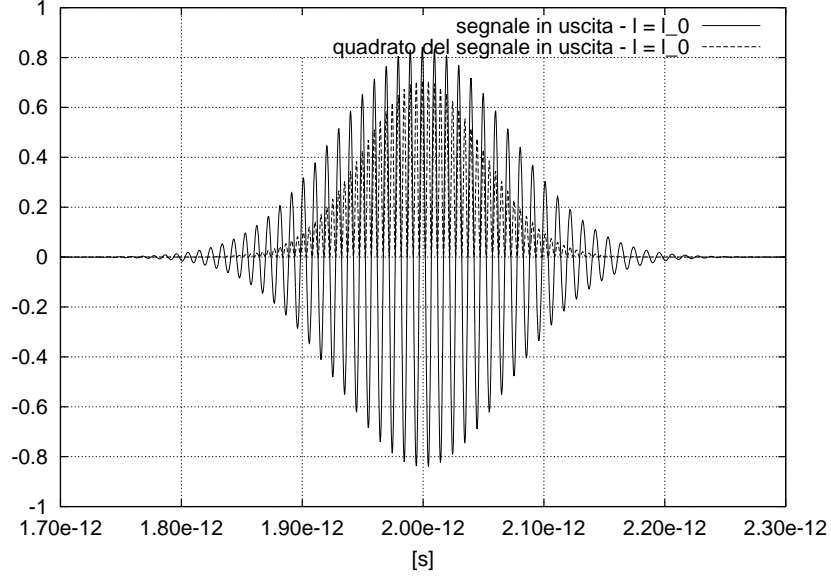


Figura 2.20: Andamento nel tempo del segnale di uscita e della sua potenza, per $l = l_0$. Il segnale è stato ottenuto assumendo che la propagazione si verifichi in un mezzo che determina $v_f(\omega_0) = v_g(\omega_0) = c_0$, $l_0 = 100 T_0 c_0$, segno di $\beta''(\omega_0)$ negativo e $\alpha = 0$.

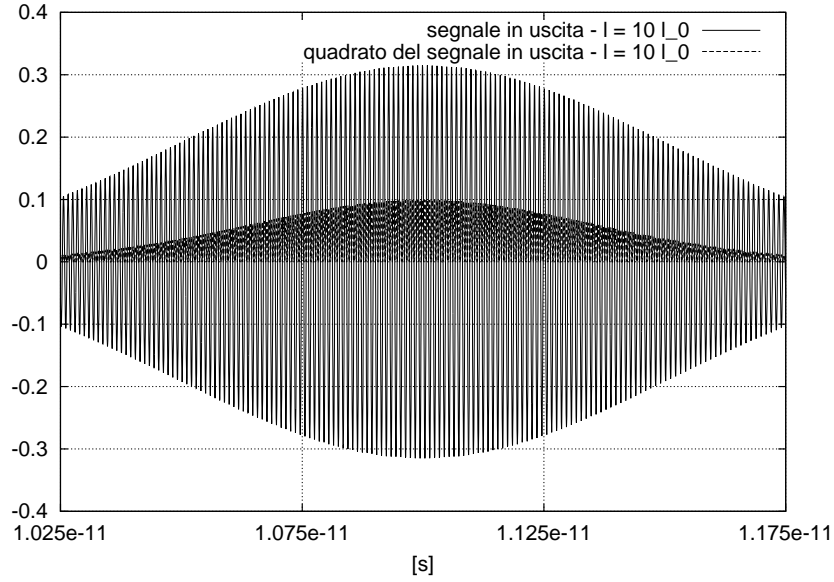


Figura 2.21: Andamento nel tempo del segnale di uscita e della sua potenza, per $l = 10l_0$. Il segnale è stato ottenuto assumendo che la propagazione si verifichi in un mezzo che determina $v_f(\omega_0) = v_g(\omega_0) = c_0$, $l_0 = 100 T_0 c_0$, segno di $\beta''(\omega_0)$ negativo e $\alpha = 0$.

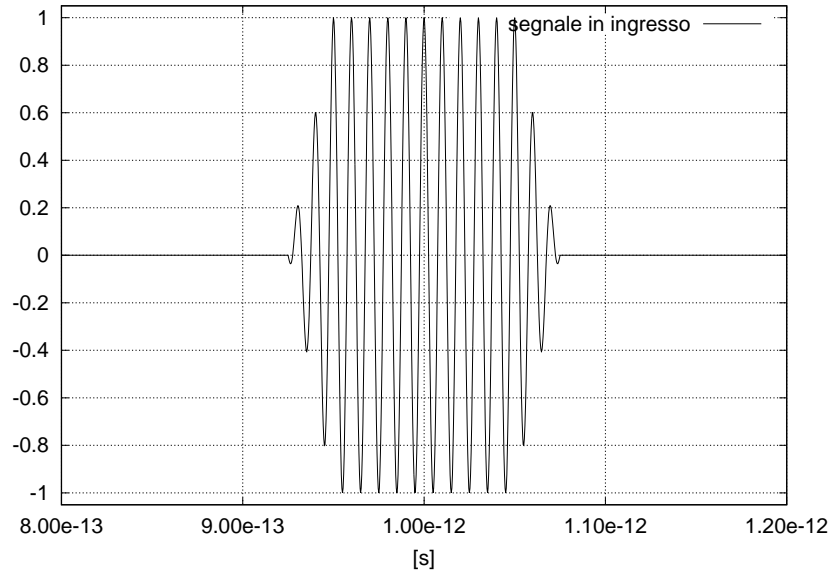


Figura 2.22: Andamento nel tempo del segnale di ingresso quando la modulante è un segnale a trapezio isoscele che mantiene il valore massimo per $10 T_0$ e il cui supporto è pari a $15 T_0$.

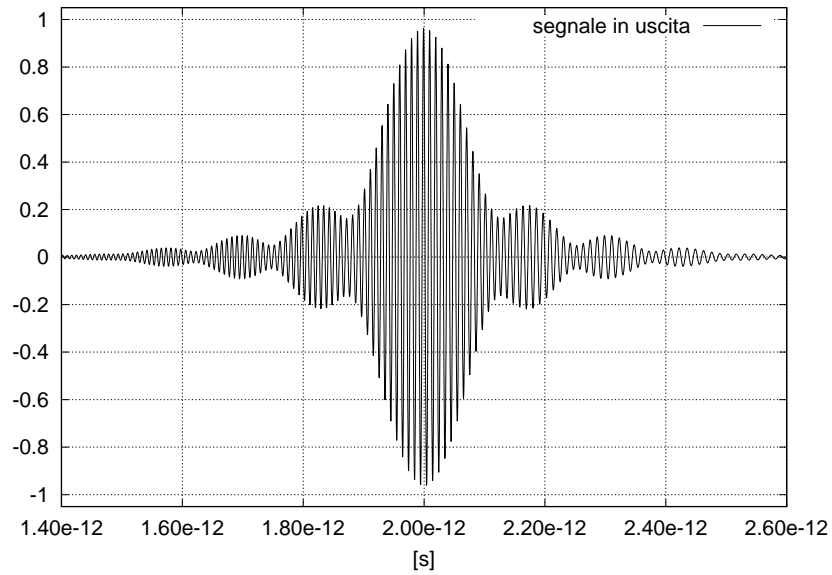


Figura 2.23: Andamento nel tempo del segnale di uscita quando la modulante è un segnale trapezoidale che mantiene il valore massimo per $10 T_0$ e il cui supporto è pari a $15 T_0$, nel caso in cui il segnale si propaghi su una tratta di lunghezza $2.99792458 \cdot 10^{-4}$.

2.7.4 Parametri utilizzati per valutare la dispersione

Il parametro fondamentale che determina la distorsione della modulante è $\beta''(\omega_0)$ e per questo motivo prende il nome di coefficiente di dispersione [$\text{s}^2 \text{ m}^{-1} \text{ rad}^{-2}$]. Tale coefficiente valuta come varia con ω , per $\omega = \omega_0$, il reciproco della velocità di gruppo (visto che $\frac{1}{v_g(\omega_0)} = \beta'(\omega_0)$), quantità che coincide con il ritardo di gruppo τ_1 [s] per percorrere una distanza di 1 m. Visto che in ottica si fa spesso uso di λ_0 (più che della frequenza o della frequenza angolare), il parametro che determina la distorsione della modulante può essere dato in termini di $\frac{d\tau_1}{d\lambda_0}|_{\lambda_0=\bar{\lambda}_0}$ [s m^{-1}] (o equivalenti; ad esempio [ps nm^{-1}]), dove $\bar{\lambda}_0$ è il valore di λ_0 che si ottiene per $\omega = \omega_0$.

2.7.5 Effetto della dispersione sulla propagazione modale

Dispersione nelle strutture guidanti:

- modale, per eccitazione multimodale (di più modi guidati) (potenza ottica distribuita su più modi che viaggiano in generale con velocità di gruppo differenti)
- dei materiali; i materiali sono in generale temporalmente dispersivi e i loro parametri costitutivi variano con la lunghezza d'onda, determinando una variazione delle curve di dispersione che si otterrebbero in presenza di mezzi non dispersivi
- della struttura guidante; i modi guidati si propagano in generale con curve di dispersione non lineari; questo effetto diventa quasi trascurabile per strutture debolmente guidanti
- la dispersione dei materiali e quella della struttura guidante di solito sono trattate insieme nella cosiddetta dispersione cromatica; per strutture debolmente guidanti la dispersione cromatica è dominata da quella dei materiali
- della polarizzazione ad esempio per i modi LP_{01} polarizzati lungo x o lungo y a causa delle imperfezioni nella realizzazione della sezione circolare della fibra o della sua alterazione nella fase di posatura (ad esempio, quando si posano creando delle curve); su una distanza fissa si introduce il cosiddetto “differential group delay” [ps]; al variare della distanza, per effetti statistici la distanza stessa interviene sotto radice quadrata e si introduce un “polarization-mode dispersion coefficient” che si misura in [$\text{ps km}^{-\frac{1}{2}}$]
- dispersione non lineare; in generale molto complessa da gestire; in alcune circostanze la dispersione non lineare compensa esattamente la dispersione lineare consentendo una propagazione senza distorsione degli impulsi ottici, che in questo caso prendono il nome di “solitoni”.

Esercizio 26 *E' possibile progettare un ricevitore che annulli gli effetti negativi della dispersione lineare e della attenuazione di un tratto di guida dielettrica?*

Capitolo 3

Componenti ottici

In questa seconda parte dell'insegnamento studieremo diversi componenti: gli accoppiatori direzionali, i divisori di potenza, i multiplexer e i demultiplexer per Wavelength Division Multiplexing (WDM), i modulatori di fase elettroottici, i modulatori di ampiezza o "switch" elettroottici, i led, i laser, gli amplificatori ottici, i fotorivelatori.

I primi di questi possono essere studiati in modo classico. I led, i laser, gli amplificatori ottici e i fotorivelatori richiedono lo studio dell'interazione dei fotoni con la materia.

3.1 Lastre dielettriche accoppiate

Importanza argomento: per realizzare componenti come power splitters, mux-demux per WDM, switch.

Figure: switch, power splitter e mux-demux per WDM.

Dalle figure si evince l'importanza di capire cosa succede quando abbiamo guide dielettriche vicine.

Per capire ciò non è necessario considerare configurazioni realistiche che rendono l'analisi complicata. Come sempre consideriamo il modello di lastre accoppiate più semplice che, da un lato, consente di sviluppare l'argomento senza difficoltà analitiche eccessive, e, dall'altro, non semplifica troppo il problema consentendo di capire l'argomento di interesse.

Per quanto indicato, considereremo lastre dielettriche accoppiate.

Disegno geometria e mezzi: lastra 1 tra b e $b+2a$, mezzo film caratterizzato da n_2 ; lastra 2 tra $-b-2a$ e $-b$, mezzo film caratterizzato da n_3 ; sopra la lastra 1, sotto la lastra 2 e tra le lastre è presente un mezzo caratterizzato da n_1 . Tutti i mezzi sono assunti essere privi di perdite (e quindi $n_i \in \mathbb{R}$, $i = 1, 2, 3$). Per definizione $\Omega_1 = (-\infty, -b-2a) \cup (-b, b) \cup (b+2a, \infty)$, $\Omega_2 = (b, b+2a)$, $\Omega_3 = (-b-2a, -b)$. Inoltre, per avere delle lastre dielettriche guidanti, $n_2 > n_1$ e $n_3 > n_1$.

Possibile approccio: si tratta di una struttura guidanti invariante in y ; possiamo trovare i modi indipendenti da y , come abbiamo fatto per la lastra dielettrica. Questo è l'approccio rigoroso al problema ma comporta uno sforzo significativo, visto che la struttura guidante è più complessa dell'unica struttura guidante che abbiamo studiato in tutti i dettagli: la lastra dielettrica.

Per questo motivo non useremo l'approccio rigoroso e cercheremo di definire una tecnica approssimata che sfrutta il più possibile le nostre conoscenze sui modi guidati della singola lastra dielettrica. In quest'ottica, al fine di semplificare il più possibile il problema mantenendo inalterata l'utilità dei risultati, consideriamo quello che succede quando nelle lastre dielettriche sono presenti modi guidati riconducibili al modo fondamentale delle singole lastre. Assumeremo, inoltre, la sola presenza delle onde progressive.

Osserviamo subito che, per $b \rightarrow \infty$, l'accoppiamento tra le lastre si annulla e la presenza delle sole onde progressive dei modi fondamentali consente di scrivere il campo elettrico nella struttura complessiva come $\mathbf{E}(x, z) = E_y(x, z)\hat{\mathbf{y}}$, dove:

$$E_y(x, z) = c_1 e_{1y}(x) e^{-j\beta_1 z} + c_2 e_{2y}(x) e^{-j\beta_2 z}. \quad (3.1)$$

Nell'equazione i coefficienti c_m , $m = 1, 2$, sono complessi, β_m , $m = 1, 2$, sono le costanti di fase dei modi fondamentali e $e_{my}(x)$, $m = 1, 2$, è l'andamento a "campana" della componente y del campo elettrico del $TE_0^{(e)}$ della lastra m .

Se b non è troppo piccolo e la frequenza non è troppo bassa si può ritenere che l'accoppiamento tra i modi fondamentali delle due lastre non sia troppo intenso e che i due modi interagiscano debolmente (ipotesi detta di debole accoppiamento). In questa ipotesi è ragionevole confermare la "struttura" del campo, espresso come nell'equazione (3.1) in termini di una combinazione lineare di modi fondamentali delle due lastre dielettriche. Si ammette però che il debole accoppiamento determini una variazione in z dei coefficienti della combinazione lineare. In questa ipotesi si ottiene

$$E_y(x, z) = c_1(z)e_{1y}(x)e^{-j\beta_1 z} + c_2(z)e_{2y}(x)e^{-j\beta_2 z}. \quad (3.2)$$

Per risolvere il problema di interesse dobbiamo quindi trovare le funzioni a valori complessi $c_m(z)$, $m = 1, 2$.

3.1.1 Deduzione del problema differenziale che individua le funzioni incognite $c_m(z)$, $m = 1, 2$

L'ipotesi formulata riguarda la forma di E_y . Per questa ragione ricordiamo che il campo elettrico in qualunque regione omogenea soddisfa l'equazione di Helmholtz

$$\nabla^2 \mathbf{E}(x, z) + \frac{\omega^2}{c_0^2} n_i^2 \mathbf{E}(x, z) = 0, \quad (x, z) \in \Omega_i \times \mathbb{R}, i = 1, 2, 3. \quad (3.3)$$

Se per ipotesi il campo elettrico presenta la sola componente y si deduce

$$\frac{\partial^2}{\partial x^2} E_y(x, z) + \frac{\partial^2}{\partial z^2} E_y(x, z) + \frac{\omega^2}{c_0^2} n_i^2 E_y(x, z) = 0, \quad (3.4)$$

per $(x, z) \in \Omega_i \times \mathbb{R}$, $i = 1, 2, 3$.

Coerentemente con l'ipotesi relativa all'espressione di E_y , calcoliamo le derivate parziali e imponiamo alla componente E_y considerata di soddisfare l'equazione (3.4).

Per quel che riguarda la derivata parziale rispetto a x la valutazione è semplice essendo presente in ogni addendo una sola funzione della variabile x . Si ottiene

$$\frac{\partial^2}{\partial x^2} E_y(x, z) = \sum_{m=1}^2 c_m(z) e^{-j\beta_m z} \frac{d^2}{dx^2} e_{my}(x). \quad (3.5)$$

La derivata parziale rispetto a z prevede di derivare in ogni addendo il prodotto di due funzioni. Si ha $(a(z)b(z))' = a'(z)b(z) + a(z)b'(z)$ e quindi $(a(z)b(z))'' = (a'(z)b(z) + a(z)b'(z))' = a''(z)b(z) + 2a'(z)b'(z) + a(z)b''(z)$. Si deduce allora

$$\frac{\partial^2}{\partial z^2} E_y(x, z) = \sum_{m=1}^2 e_{my}(x) \left[e^{-j\beta_m z} \left(\frac{d^2}{dz^2} c_m(z) \right) - 2j\beta_m e^{-j\beta_m z} \left(\frac{d}{dz} c_m(z) \right) - \beta_m^2 e^{-j\beta_m z} c_m(z) \right]. \quad (3.6)$$

Come visto in precedenza, in caso di accoppiamento nullo, anche le derivate prime e seconde in z dei coefficienti della combinazione lineare si annullano. In presenza di debole accoppiamento si può invece assumere che i coefficienti della combinazione lineare varino ma lo facciano molto più lentamente in z dei singoli campi modali, che dipendono da z con il solito fattore di propagazione. Per questa ragione, nella parentesi quadra presente a secondo membro dell'equazione (3.6) il termine con β_m^2 è relativamente grande, quello con β_m e con la derivata prima rispetto a z di $c_m(z)$ ha una ampiezza intermedia e quello con la derivata seconda dei coefficienti $c_m(z)$ è relativamente piccolo.

Ne segue che l'espressione (3.6) può essere approssimata con la seguente:

$$\frac{\partial^2}{\partial z^2} E_y(x, z) = - \sum_{m=1}^2 e_{my}(x) \left[2j\beta_m e^{-j\beta_m z} \left(\frac{d}{dz} c_m(z) \right) + \beta_m^2 e^{-j\beta_m z} c_m(z) \right]. \quad (3.7)$$

Sostituendo le espressioni (3.5) e (3.7), per le derivate seconde, e (3.2), per il campo, nell'equazione (3.4) si deduce

$$\begin{aligned} \sum_{m=1}^2 \left\{ c_m(z) e^{-j\beta_m z} \frac{d^2}{dx^2} e_{my}(x) + \right. \\ \left. - e_{my}(x) \left[2j\beta_m e^{-j\beta_m z} \left(\frac{d}{dz} c_m(z) \right) + \beta_m^2 e^{-j\beta_m z} c_m(z) \right] + \right. \\ \left. + \frac{\omega^2}{c_0^2} n_i^2 c_m(z) e_{my}(x) e^{-j\beta_m z} \right\} = 0, \end{aligned} \quad (3.8)$$

per $(x, z) \in \Omega_i \times \mathbb{R}$, $i = 1, 2, 3$.

Le singole regioni $\Omega_i \times \mathbb{R}$, $i = 1, 2, 3$, sono omogenee. In particolare, nella regione $\Omega_1 \times \mathbb{R}$ è presente lo stesso mezzo che si avrebbe in presenza di una sola struttura guidante. Quindi, entrambi i campi $e_{my}(x) e^{-j\beta_m z}$ soddisfano l'equazione (3.4) da cui si deduce, semplificando il fattore di propagazione, comune a tutti i termini e non nullo $\forall \beta_m \in \mathbb{R}$ e $\forall z \in \mathbb{R}$:

$$\frac{d^2}{dx^2} e_{my}(x) + \left(\frac{\omega^2}{c_0^2} n_1^2 - \beta_m^2 \right) e_{my}(x) = 0, \quad (3.9)$$

$m = 1, 2$. Lo stesso ragionamento si applica per e_{1y} in $\Omega_2 \times \mathbb{R}$ e per e_{2y} in $\Omega_3 \times \mathbb{R}$; ne segue che

$$\frac{d^2}{dx^2} e_{1y}(x) + \left(\frac{\omega^2}{c_0^2} n_2^2 - \beta_1^2 \right) e_{1y}(x) = 0 \quad (3.10)$$

e

$$\frac{d^2}{dx^2} e_{2y}(x) + \left(\frac{\omega^2}{c_0^2} n_3^2 - \beta_2^2 \right) e_{2y}(x) = 0. \quad (3.11)$$

Viceversa, il modo e_{1y} in $\Omega_3 \times \mathbb{R}$, pur soddisfacendo la stessa equazione, lo fa in presenza di un mezzo diverso da quello che è presente nella configurazione di interesse. In particolare si ha

$$\frac{d^2}{dx^2} e_{1y}(x) + \left(\frac{\omega^2}{c_0^2} n_1^2 - \beta_1^2 \right) e_{1y}(x) = 0 \quad (3.12)$$

e quindi

$$\frac{d^2}{dx^2} e_{1y}(x) + \left(\frac{\omega^2}{c_0^2} n_3^2 - \beta_1^2 \right) e_{1y}(x) = \frac{\omega^2}{c_0^2} (n_3^2 - n_1^2) e_{1y}(x). \quad (3.13)$$

Analogamente, il modo e_{2y} in $\Omega_2 \times \mathbb{R}$ soddisfa

$$\frac{d^2}{dx^2} e_{2y}(x) + \left(\frac{\omega^2}{c_0^2} n_2^2 - \beta_2^2 \right) e_{2y}(x) = \frac{\omega^2}{c_0^2} (n_2^2 - n_1^2) e_{2y}(x). \quad (3.14)$$

Sostituendo quanto dedotto dall'equazione (3.9) nell'equazione (3.8) per $(x, z) \in \Omega_1 \times \mathbb{R}$ si ottiene:

$$- \sum_{m=1}^2 e_{my}(x) 2j\beta_m e^{-j\beta_m z} \left(\frac{d}{dz} c_m(z) \right) = 0. \quad (3.15)$$

Analogamente, sostituendo quanto dedotto dalle equazioni (3.10) e (3.14) in $\Omega_2 \times \mathbb{R}$, rispettivamente per e_{1y} e per e_{2y} , nell'equazione (3.8) si ottiene:

$$\frac{\omega^2}{c_0^2} (n_2^2 - n_1^2) e_{2y}(x) e^{-j\beta_2 z} c_2(z) - \sum_{m=1}^2 e_{my}(x) 2j\beta_m e^{-j\beta_m z} \left(\frac{d}{dz} c_m(z) \right) = 0, \quad (3.16)$$

per $(x, z) \in \Omega_2 \times \mathbb{R}$.

In modo duale si ottiene:

$$\frac{\omega^2}{c_0^2} (n_3^2 - n_1^2) e_{1y}(x) e^{-j\beta_1 z} c_1(z) - \sum_{m=1}^2 e_{my}(x) 2j\beta_m e^{-j\beta_m z} \left(\frac{d}{dz} c_m(z) \right) = 0, \quad (3.17)$$

per $(x, z) \in \Omega_3 \times \mathbb{R}$.

Le equazioni (3.15), (3.16), (3.17) possono essere riscritte in modo sintetico come segue:

$$\begin{aligned} & \frac{\omega^2}{c_0^2} (n_2^2 - n_1^2) e_{2y}(x) e^{-j\beta_2 z} c_2(z) W_2(x) + \\ & \frac{\omega^2}{c_0^2} (n_3^2 - n_1^2) e_{1y}(x) e^{-j\beta_1 z} c_1(z) W_3(x) + \\ & - \sum_{m=1}^2 e_{my}(x) 2j\beta_m e^{-j\beta_m z} \left(\frac{d}{dz} c_m(z) \right) = 0, \quad \forall (x, z) \in \mathbb{R}^2, \end{aligned} \quad (3.18)$$

dove le $W_i(x)$ sono le cosiddette funzioni caratteristiche degli insiemi Ω_i e sono pari a 1 in Ω_i e valgono zero altrove, $i = 2, 3$.

Le incognite sono due e sono funzioni della sola variabile z . Per questo motivo proiettiamo la precedente equazione, valida in \mathbb{R}^2 , rispettivamente su e_{1y} e e_{2y} (l'operazione di proiezione va intesa nel senso di $L^2(\mathbb{R})$). Così facendo si ottiene:

$$\begin{aligned} & \frac{\omega^2}{c_0^2} (n_2^2 - n_1^2) e^{-j\beta_2 z} c_2(z) \int_b^{b+2a} e_{2y}(x) e_{1y}^*(x) dx + \\ & \frac{\omega^2}{c_0^2} (n_3^2 - n_1^2) e^{-j\beta_1 z} c_1(z) \int_{-b-2a}^{-b} e_{1y}(x) e_{1y}^*(x) dx + \\ & - \sum_{m=1}^2 2j\beta_m e^{-j\beta_m z} \left(\frac{d}{dz} c_m(z) \right) \int_{-\infty}^{\infty} e_{my}(x) e_{1y}^*(x) dx = 0, \quad \forall z \in \mathbb{R}, \end{aligned} \quad (3.19)$$

$$\begin{aligned} & \frac{\omega^2}{c_0^2} (n_2^2 - n_1^2) e^{-j\beta_2 z} c_2(z) \int_b^{b+2a} e_{2y}(x) e_{2y}^*(x) dx + \\ & \frac{\omega^2}{c_0^2} (n_3^2 - n_1^2) e^{-j\beta_1 z} c_1(z) \int_{-b-2a}^{-b} e_{1y}(x) e_{2y}^*(x) dx + \\ & - \sum_{m=1}^2 2j\beta_m e^{-j\beta_m z} \left(\frac{d}{dz} c_m(z) \right) \int_{-\infty}^{\infty} e_{my}(x) e_{2y}^*(x) dx = 0, \quad \forall z \in \mathbb{R}. \end{aligned} \quad (3.20)$$

Gli integrali sono dei numeri (gli integrali impropri sono tutti convergenti grazie al decadimento esponenziale delle funzioni presenti nelle integrande) e le due equazioni ottenute costituiscono un sistema di equazioni differenziali ordinarie del primo ordine a coefficienti non costanti nelle incognite $c_i(z)$, $i = 1, 2$.

Si considerino i seguenti simboli per gli integrali di interesse: $I_{2,21}$, $I_{3,11}$, $I_{\mathbb{R},11}$, $I_{\mathbb{R},21}$, $I_{2,22}$, $I_{3,12}$, $I_{\mathbb{R},12}$, $I_{\mathbb{R},22}$; il primo pedice fa riferimento in ogni caso al dominio di integrazione (quando è pari a i significa che il dominio di integrazione è Ω_i e quando è pari a \mathbb{R} significa che è tutto \mathbb{R}) mentre gli altri pedici individuano il campo che compare rispettivamente in prima o seconda posizione nell'integranda.

Senza perdita di generalità, i campi $e_{my}(x)$, $m = 1, 2$, possono essere considerati reali e positivi (ad esempio, usando le espressioni (2.84), (2.85) e (2.86) con C_{22} puramente immaginario si ottiene e_{1y} a valori reali e usando $C_{22} = -j|C_{22}|$ si ottiene a valori positivi, ipotizzando di prendere, come fatto a suo tempo, k_{c2} reale e positivo e k_{c1} con $Im(k_{c1})$ negativo). Ciò consente di concludere che i valori degli integrali $I_{2,21}$, $I_{3,11}$, $I_{\mathbb{R},11}$, $I_{\mathbb{R},21}$, $I_{2,22}$, $I_{3,12}$, $I_{\mathbb{R},12}$ e $I_{\mathbb{R},22}$ sono tutti reali e positivi.

Inoltre è possibile scegliere la loro ampiezza in modo che le funzioni $e_{my}(x)$, $m = 1, 2$, siano normalizzate, determinando $I_{\mathbb{R},11} = I_{\mathbb{R},22} = 1$. Infine, nell'ipotesi di debole accoppiamento, $I_{3,11}$ e $I_{2,22}$ sono numeri reali positivi estremamente piccoli e trascurabili rispetto agli altri integrali e inoltre $I_{\mathbb{R},21}$ e $I_{\mathbb{R},12}$ risultano essere molto più piccoli di $I_{\mathbb{R},11}$ e $I_{\mathbb{R},22}$ e quindi dell'unità.

Si deduce quindi che il sistema di equazioni differenziali ordinarie del primo ordine a coefficienti non costanti nelle incognite $c_i(z)$, $i = 1, 2$, può essere riscritto, in modo approssimato ma con ottima approssimazione, come segue, per tutti i $z \in \mathbb{R}$:

$$\begin{cases} \frac{\omega^2}{c_0^2} (n_2^2 - n_1^2) e^{-j\beta_2 z} c_2(z) I_{2,21} - 2j\beta_1 e^{-j\beta_1 z} \left(\frac{d}{dz} c_1(z) \right) - 2j\beta_2 e^{-j\beta_2 z} \left(\frac{d}{dz} c_2(z) \right) I_{\mathbb{R},21} = 0 \\ \frac{\omega^2}{c_0^2} (n_3^2 - n_1^2) e^{-j\beta_1 z} c_1(z) I_{3,12} - 2j\beta_1 e^{-j\beta_1 z} \left(\frac{d}{dz} c_1(z) \right) I_{\mathbb{R},12} - 2j\beta_2 e^{-j\beta_2 z} \left(\frac{d}{dz} c_2(z) \right) = 0 \end{cases} \quad (3.21)$$

A questo punto è evidente che se alla prima equazione sottraggo la seconda equazione del sistema moltiplicata per $I_{\mathbb{R},21}$ e se alla seconda equazione sottraggo la prima moltiplicata per $I_{\mathbb{R},12}$ ottengo un sistema di equazioni differenziali del primo ordine in cui compare una sola derivata prima. In particolare, si ottiene:

$$\begin{cases} \frac{\omega^2}{c_0^2} (n_2^2 - n_1^2) e^{-j\beta_2 z} c_2(z) I_{2,21} - \frac{\omega^2}{c_0^2} (n_3^2 - n_1^2) e^{-j\beta_1 z} c_1(z) I_{3,12} I_{\mathbb{R},21} - 2j\beta_1 e^{-j\beta_1 z} \left(\frac{d}{dz} c_1(z) \right) (1 - I_{\mathbb{R},12} I_{\mathbb{R},21}) = 0 \\ \frac{\omega^2}{c_0^2} (n_3^2 - n_1^2) e^{-j\beta_1 z} c_1(z) I_{3,12} - \frac{\omega^2}{c_0^2} (n_2^2 - n_1^2) e^{-j\beta_2 z} c_2(z) I_{2,21} I_{\mathbb{R},12} - 2j\beta_2 e^{-j\beta_2 z} \left(\frac{d}{dz} c_2(z) \right) (1 - I_{\mathbb{R},21} I_{\mathbb{R},12}) = 0 \end{cases} \quad (3.22)$$

Le funzioni o i numeri che moltiplicano le derivate prime nelle due equazioni del sistema sono certamente non nulle (per $\omega > 0$ si ha $\beta_i > 0$, $i = 1, 2$; inoltre $I_{\mathbb{R},21} I_{\mathbb{R},12} < 1$; infine $e^{-j\beta_i z} \neq 0$ per ogni valore di β_i e z). Allora possiamo riscrivere il sistema in forma normale come segue:

$$\begin{cases} \frac{d}{dz} c_1(z) = \frac{\omega^2 (n_2^2 - n_1^2) I_{2,21}}{c_0^2 2j\beta_1 (1 - I_{\mathbb{R},12} I_{\mathbb{R},21})} e^{-j(\beta_2 - \beta_1)z} c_2(z) - \frac{\omega^2 (n_3^2 - n_1^2) I_{3,12} I_{\mathbb{R},21}}{c_0^2 2j\beta_1 (1 - I_{\mathbb{R},12} I_{\mathbb{R},21})} c_1(z) \\ \frac{d}{dz} c_2(z) = \frac{\omega^2 (n_3^2 - n_1^2) I_{3,12}}{c_0^2 2j\beta_2 (1 - I_{\mathbb{R},21} I_{\mathbb{R},12})} e^{-j(\beta_1 - \beta_2)z} c_1(z) - \frac{\omega^2 (n_2^2 - n_1^2) I_{2,21} I_{\mathbb{R},12}}{c_0^2 2j\beta_2 (1 - I_{\mathbb{R},21} I_{\mathbb{R},12})} c_2(z) \end{cases} \quad (3.23)$$

In situazioni in cui $n_2 \simeq n_3$, sotto l'ipotesi di debole accoppiamento, si deduce $I_{2,21} \simeq I_{3,12}$; si tratta di quantità piccole. Inoltre, per quanto indicato in precedenza, $I_{\mathbb{R},21}$ e $I_{\mathbb{R},12}$ sono piccoli rispetto all'unità. Allora, nell'ipotesi di debole accoppiamento tra lastre dielettriche non troppo diverse, possiamo semplificare ancora il sistema di equazioni differenziali in forma normale, visto che è possibile trascurare i termini al cui numeratore compare il prodotto di integrali. Così facendo si ottiene, isolando un paio di coefficienti reali e strettamente positivi (per ipotesi $n_2 > n_1$ e $n_3 > n_1$) a secondo membro:

$$\begin{cases} \frac{d}{dz} c_1(z) = -j \frac{\omega^2 (n_2^2 - n_1^2) I_{2,21}}{c_0^2 2\beta_1 (1 - I_{\mathbb{R},12} I_{\mathbb{R},21})} e^{-j(\beta_2 - \beta_1)z} c_2(z) = -j C_{21} e^{-j(\beta_2 - \beta_1)z} c_2(z) \\ \frac{d}{dz} c_2(z) = -j \frac{\omega^2 (n_3^2 - n_1^2) I_{3,12}}{c_0^2 2\beta_2 (1 - I_{\mathbb{R},21} I_{\mathbb{R},12})} e^{-j(\beta_1 - \beta_2)z} c_1(z) = -j C_{12} e^{-j(\beta_1 - \beta_2)z} c_1(z) \end{cases} \quad (3.24)$$

Per risolvere il sistema di equazioni differenziali ordinarie del primo ordine così dedotto servono condizioni iniziali o al contorno. Se limitiamo la regione di accoppiamento delle lastre dielettriche il sistema di equazioni dedotto può essere considerato su un intervallo dell'asse z , (z_1, z_2) , $z_2 > z_1$, e possiamo considerare condizioni iniziali o al contorno per il sistema stesso. Avendo semplificato l'analisi considerando solo onde progressive, per coerenza dobbiamo accettare che le condizioni siano solo iniziali (nel senso che dovranno essere forzati i valori delle soluzioni per $z = z_1$ e si dovranno studiare gli effetti dell'accoppiamento per $z \in (z_1, z_2)$). Nel seguito porremo sempre, senza alcuna perdita di generalità, $z_1 = 0$.

3.1.2 Soluzione del problema differenziale che individua le funzioni incognite $c_m(z)$, $m = 1, 2$

Il sistema di equazioni differenziali presenta tre parametri caratteristici: $\beta_1 - \beta_2, C_{21}, C_{12} \in \mathbb{R}$, $C_{21} > 0, C_{12} > 0$. Definiamo $d = \frac{\beta_1 - \beta_2}{2} \in \mathbb{R}$ e $r = \sqrt{d^2 + C_{21} C_{12}} \in \mathbb{R}$, $r > 0$.

Risolvendo il problema differenziale con condizioni iniziali $c_1(0)$ e $c_2(0)$ si ottiene:

$$c_1(z) = e^{jdz} \left(\cos(rz) - j \frac{d}{r} \sin(rz) \right) c_1(0) - j \frac{C_{21}}{r} \sin(rz) e^{jdz} c_2(0) \quad (3.25)$$

$$c_2(z) = -j \frac{C_{12}}{r} \sin(rz) e^{-jdz} c_1(0) + e^{-jdz} \left(\cos(rz) + j \frac{d}{r} \sin(rz) \right) c_2(0). \quad (3.26)$$

Esercizio 27 Si verifichi che la coppia di funzioni data dalle equazioni (3.25) e (3.26) costituisce l'unica soluzione del sistema (3.24).

3.1.3 Considerazioni applicative

Gli aspetti più interessanti possono essere messi in evidenza considerando una condizione iniziale particolarmente semplice; quella in cui si pone $c_2(0) = 0$ e $c_1(0) \neq 0$.

In queste ipotesi la soluzione si semplifica e si ottiene:

$$c_1(z) = e^{jdz} \left(\cos(rz) - j \frac{d}{r} \sin(rz) \right) c_1(0) \quad (3.27)$$

$$c_2(z) = -j \frac{C_{12}}{r} \sin(rz) e^{-jdz} c_1(0). \quad (3.28)$$

Inoltre le valutazioni più significative riguardano la potenza attiva dei segnali modali nei due slab e per questo si valutano le quantità $|c_i(z)|^2$, $i = 1, 2$:

$$|c_1(z)|^2 = \left(\cos^2(rz) + \frac{d^2}{r^2} \sin^2(rz) \right) |c_1(0)|^2 \quad (3.29)$$

$$|c_2(z)|^2 = \frac{C_{12}^2}{r^2} \sin^2(rz) |c_1(0)|^2. \quad (3.30)$$

Si consideri, in particolare, il caso “phase matched”. Per definizione si ha $n_2 = n_3$ e quindi $\beta_1 = \beta_2$. Da ciò segue $d = 0$, $C_{12} = C_{21}$ e $r^2 = C_{12}^2$. Si deducono quindi le seguenti espressioni semplificate di $|c_i(z)|^2$, $i = 1, 2$:

$$|c_1(z)|^2 = \cos^2(rz) |c_1(0)|^2 \quad (3.31)$$

$$|c_2(z)|^2 = \sin^2(rz) |c_1(0)|^2. \quad (3.32)$$

Andamento grafico di $|c_1(z)|^2$ e $|c_2(z)|^2$ per $|c_1(0)| = 1$.

Fenomeno fisico: scambio di potenza attiva tra i due slab debolmente accoppiati.

Possibile funzionamento come power splitter o switch (parametri costruttivi nel caso phase matched: n_2 , n_1 , a , b e parametri dipendenti dall'eccitazione esterna λ_0 e quindi ω).

Effetto della frequenza (focusing dei modi guidati) e funzionamento come parte di mux e demux per WDM.

Caso “phase mismatched” in termini di $|c_2(z)|^2$ con oscillazioni di ampiezza inferiore (senza verifica) a $|c_1(0)|^2$.

Anche nel caso generale, se non si ha la possibilità di calcolare i coefficienti d , C_{12} e C_{21} , si possono ottenere delle stime di tali quantità se è possibile osservare l'andamento di $|c_2(z)|^2$ si tratta di lastre dielettriche debolmente accoppiate sufficientemente lunga. Infatti, in tal caso, si può dedurre il valore di r e di C_{12} . Una volta stimato il valore di r , dall'osservazione di $|c_1(z)|^2$ si deduce il valore di d e dai valori ottenuti si trova anche C_{21} .

Applicazioni interessanti; lo sarebbero decisamente di più se avessimo la possibilità di realizzare componenti il cui funzionamento viene controllato elettronicamente. Il prossimo paragrafo si occuperà proprio di questo.

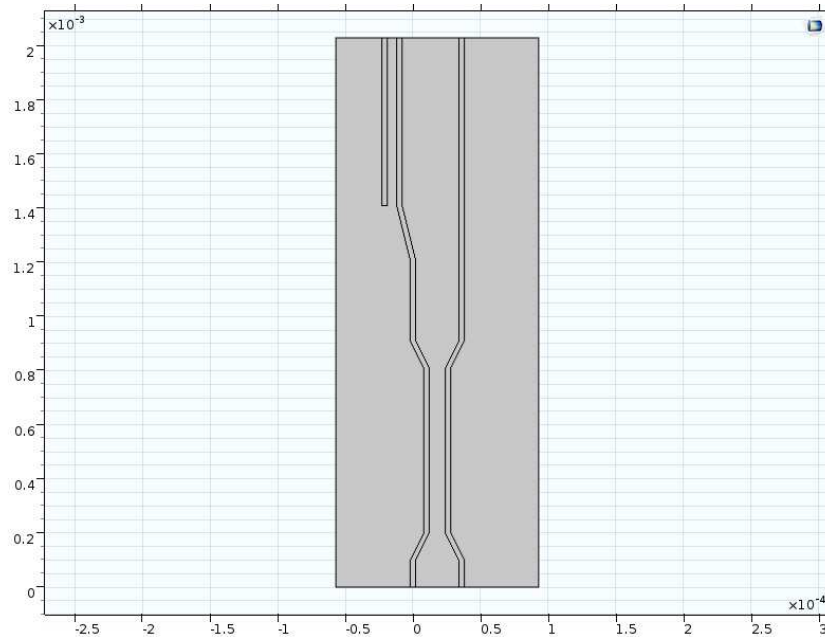


Figura 3.1: Geometria di un multiplexer/demultiplexer realizzato con lastre dielettriche accoppiate.

3.1.4 Simulazioni e risultati

Le figure 3.1 riporta la geometria bidimensionale di un demultiplexer realizzato con lastre dielettriche accoppiate. In particolare, la larghezza di tutti i film mostrati in figura è pari a $4 \mu\text{m}$ ($a = 2 \mu\text{m}$), e il loro indice di rifrazione è in ogni caso pari a 2.25 ($n_2 = n_3 = 2.25$, caso “phase matched”). La materia restante è caratterizzata da $n_1 = 1.46$.

Non vengono considerate lastre dielettriche debolmente guidanti per evitare di considerare domini di indagine numerica particolarmente grandi, che renderebbero la simulazione del dispositivo numericamente onerosa.

Il dispositivo viene alimentato dalla porta di ingresso situata in basso a sinistra (si veda la figura 3.1) con il modo fondamentale a 12 e a 18 THz. Vengono fatte due simulazioni indipendenti ma per linearità il risultato sarebbe lo stesso se si considerasse presente il modo fondamentale sia a 12 THz sia a 18 THz.

I risultati ottenuti con il simulatore COMSOL Multiphysics sono riportati nelle figure 3.2, quando $f = 12 \text{ THz}$, e 3.3, quando $f = 18 \text{ THz}$.

Le figure mettono in evidenza che il segnale in ingresso viene inviato a porte diverse a seconda della λ_0 (frequenza) del segnale stesso.

Nel tratto più in basso, dove le lastre presentano una separazione minima di $12 \mu\text{m}$ (che determina $b = 6 \mu\text{m}$), le lastre si sviluppano in parallelo su una lunghezza di $590 \mu\text{m}$. Ciò è sufficiente a determinare il completo switch del segnale ottico a 12 THz.

Lo stesso tratto non determina un rilevabile travaso di potenza ottica a 18 THz (si veda la figura 3.3). Lo switch a 18 THz avviene nel tratto accoppiato più in alto dove la separazione tra le lastre è pari a $7 \mu\text{m}$ (in questo caso $b = 3.5 \mu\text{m}$). Il risultato si ottiene su una lunghezza di $600 \mu\text{m}$.

Ovviamente lo stesso dispositivo, utilizzato in direzione opposta (dall’alto verso il basso), può funzionare da multiplexer (se si eccita il modo fondamentale dalla porta in alto a sinistra a 18 THz e quella in alto a destra a 12 THz).

Esercizio 28 Cosa succederebbe se la porta d’ingresso in basso a sinistra venisse alimentata dal modo fondamentale a 30 THz? (Suggerimento: l’ulteriore confinamento del modo fundamenta-

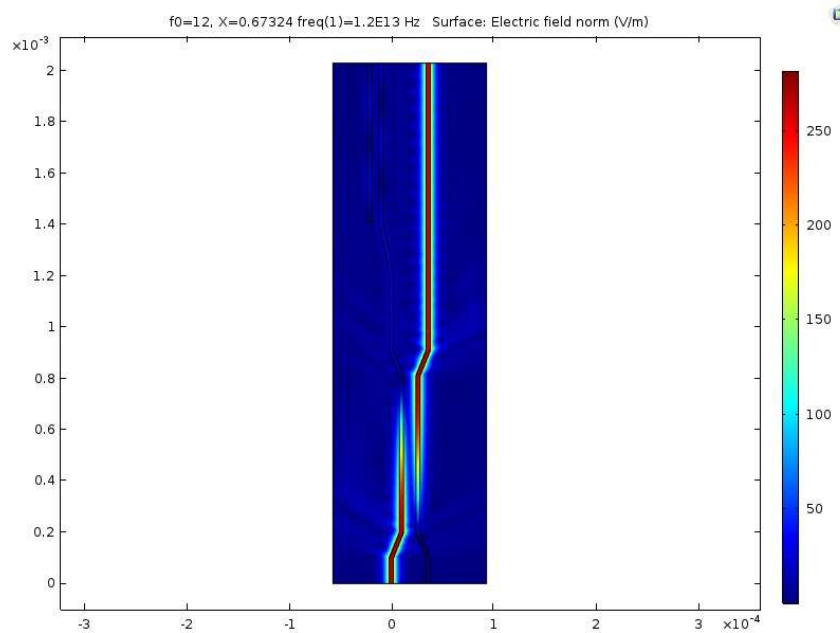


Figura 3.2: A 12 THz le lastre dielettriche che si trovano più in basso consentono di ottenere il totale trasferimento della potenza da uno slab all'altro e la potenza ottica esce dalla porta più a destra tra quelle di uscita.

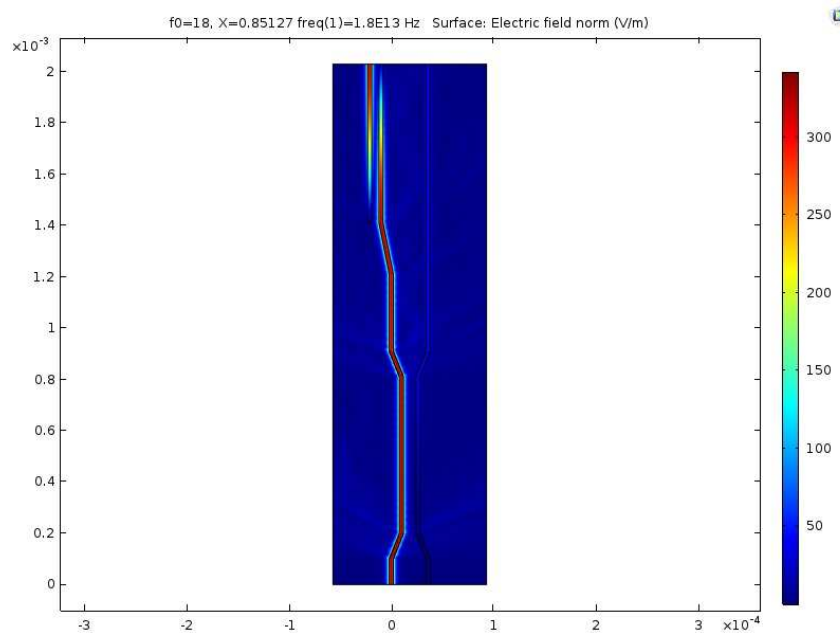


Figura 3.3: A 18 THz il trasferimento di potenza tra le lastre dielettriche che si trovano in basso risulta invece essere trascurabile. A questa frequenza il modo fondamentale procede nello slab di sinistra ed entra nel secondo dispositivo a lastre dielettriche accoppiate. Quest'ultimo è dimensionato in modo da determinare lo switch della potenza ottica, che esce dalla porta più a sinistra tra quelle di uscita.

le ridurrebbe l'accoppiamento a un livello tale da rendere entrambi i tratti di lastre dielettriche accoppiate non sufficienti a realizzare un significativo travaso di potenza).

3.2 Effetto elettroottico

Di cosa si tratta: campo elettrico statico altera il comportamento della materia. Ciò determina, in particolare, una variazione del valore dell'indice di rifrazione della materia stessa e tale variazione, come sappiamo, può avere effetti davvero significativi sulla propagazione dei modi.

Motivo fisico: polarizzazione elettrica determina ε_r e quindi n . Polarizzazione elettronica, atomica-molecolare e per orientamento delle molecole polari. Effetto elettroottico presente in tutti i mezzi (tranne che nel (modello del) vuoto) ma spesso trascurabile. In alcuni casi non lo è. Esempio notevole: niobato di litio. Tali materiali sono detti elettroottici.

Analisi quantitativa: per ragioni fisiche l'indice di rifrazione è funzione regolare di campo elettrico statico che sollecita il materiale. Allora è possibile espansione di McLaurin dell'indice di rifrazione:

$$n(x, y, z, E_x, E_y, E_z) = n(x, y, z, 0, 0, 0) + \sum_{\alpha=x,y,z} \frac{\partial n}{\partial E_\alpha} \Big|_{(x,y,z,0,0,0)} E_\alpha + \frac{1}{2} \sum_{\alpha,\beta=x,y,z} \frac{\partial^2 n}{\partial E_\alpha \partial E_\beta} \Big|_{(x,y,z,0,0,0)} E_\alpha E_\beta + \dots \quad (3.33)$$

L'espansione può essere troncata alle derivate seconde perchè non sono noti mezzi in cui siano di qualche rilevanza le derivate terze.

Per capire il fenomeno è sufficiente considerare quello che succede quando il campo elettrico statico ha una sola componente cartesiana non nulla, ad esempio E_x . In questo caso si ha:

$$n(x, y, z, E_x) = n(x, y, z, 0) + \frac{\partial n}{\partial E_x} \Big|_{(x,y,z,0)} E_x + \frac{1}{2} \frac{\partial^2 n}{\partial E_x^2} \Big|_{(x,y,z,0)} E_x^2. \quad (3.34)$$

Materiali Pockels ed effetto elettroottico Pockels: mezzi in cui i termini con le derivate prime sono non nulli e quelli con le derivate seconde sono completamente trascurabili. Esempio: niobato di litio. Prime valutazioni quantitative: per $|E_x|$ dell'ordine di 10^6 [V/m] si ottengono variazioni di n nel range $[10^{-6}, 10^{-4}]$.

Materiali Kerr ed effetto elettroottico Kerr. Alcuni materiali con forti simmetrie (gas, liquidi e alcuni cristalli) hanno n che è una funzione pari di E_x (nella espansione di McLaurin semplificata) e, di conseguenza, il termine con la derivata prima si annulla e il termine dominante che determina l'effetto elettroottico è quello con la derivata seconda. Esempi: alcuni liquidi polari, tipo il nitrobenzene ($C_6H_5NO_2$). Richiedono tensioni tipicamente superiori a quelle utilizzate nei mezzi Pockels per ottenere effetti apprezzabili. L'effetto Kerr di solito è presente anche nei mezzi Pockels ma, in questi casi, di solito è trascurabile. Prime valutazioni quantitative: per $|E_x|$ dell'ordine di 10^6 [V/m] si ottengono variazioni di n nel range $[10^{-6}, 10^{-2}]$ nei mezzi Kerr cristallini e nel range $[10^{-10}, 10^{-7}]$ nei liquidi.

Esercizio 29 Si consideri una lastra dielettrica debolmente guidante realizzata su substrato di $LiNbO_3$ diffondendo titanio. Si assuma che localmente, dove è presente il grosso dell'energia del modo fondamentale per $\lambda_0 = 1.55 \mu m$, gli indici di rifrazione siano $n_1 \simeq 2.22$ e $n_2 \simeq 2.25$ e possano essere alterati di $\Delta n_i = 2.5 \cdot 10^{-5} n_i$, $i = 1, 2$, in presenza del campo elettrostatico. Valutare la variazione di fase che subisce il modo fondamentale propagandosi su una tratta di lastra lunga 10000λ , in assenza di campo elettrostatico e in sua presenza.

Esercizio 30 Valutare la dimensione di un dispositivo di lunghezza pari a 10000λ , per $\lambda_0 = 1.55$ micron e $n_1 \simeq 2.22$, $n_2 \simeq 2.25$.

Applicazioni: modulatore di fase binario: realizzazione pratica in ottica integrata (fibra, embedded strip in niobato di litio e titanio diffuso, fibra, elettrodi molto lunghi e molto ravvicinati,

posti sull'interfaccia niobato di litio-aria, valutazione del campo elettrico statico); estensione del funzionamento al caso di campo quasi-statico e valutazione della durata di un periodo di un segnale a 100 GHz rispetto a segnali ottici a 200 THz; realizzati modulatori di fase binari a 100 Gb/s; modulatore di ampiezza realizzato con modulatore di fase e interferometro Mach-Zehnder.

Effetto acustooptico simile ad effetto elettroottico, con la differenza che in questo caso la sollecitazione è di tipo elastico (onda acustica). Non svilupperemo l'argomento.

3.3 Proprietà essenziali dei fotoni

Alcuni risultati sperimentali (radiazione da corpo nero [12] (pp. 580-586); effetto fotoelettrico; [12] (pp. 586-590); effetto Compton [12] (pp. 590-594)) hanno portato alla conclusione che la teoria di Maxwell è sbagliata [13] (p. 184, nota a piè pagina). In particolare, gli esperimenti hanno messo in evidenza che tale teoria non è in grado di descrivere in modo accurato l'interazione del campo elettromagnetico con la materia. Così nacque la fisica quantistica. Inizialmente si è sviluppata la meccanica quantistica, cioè la teoria quantistica degli atomi. Al giorno d'oggi la teoria classica dei campi elettromagnetici viene considerata come un limite della cosiddetta elettrodinamica quantistica (o, in inglese, quantum electrodynamics (QED), che include, come dice il termine, la teoria quantistica del campo elettromagnetico) [11] (p. 1), che a sua volta è stata integrata con la teoria delle interazioni deboli e oggi è inclusa nella cosiddetta teoria elettrodebole [1] (p. 445) (o, in inglese, electroweak theory). E' curioso osservare che alcuni dei fenomeni che hanno portato alla nascita delle teorie quantistiche e alla quantizzazione stessa del campo elettromagnetico (ad esempio l'effetto fotoelettrico) possano in realtà essere compresi trattando i soli atomi come oggetti quantizzati (facendo ricorso al modello semi classico, in cui il campo elettromagnetico viene considerato una classica onda) [2] (p. 3).

Nell'interazione con la materia il campo elettromagnetico deve essere descritto in termini di quanti del campo elettromagnetico, noti come fotoni (termine introdotto da Lewis nel 1926 [3] (p. 1)).

Un fotone va inteso ed è sperimentalmente rilevabile come un quanto indivisibile del campo elettromagnetico. La sua energia è pari ad hf (formula suggerita da Einstein nell'interpretazione dell'effetto fotoelettrico del 1905) e con quantità di moto di modulo $\frac{h}{\lambda_0}$, come segue da teoria classica dei campi elettromagnetici. Presenta massa a riposo nulla e si muove sempre a velocità c_0 . Polarizzazione-spin. Fotoni caratterizzati dalla stessa energia, quantità di moto e spin sono indistinguibili. Trattandosi di oggetti quantistici tutti i fenomeni ad essi connessi sono intrinsecamente aleatori, e le descrizioni degli eventi saranno sempre di tipo probabilistico. Fenomeni come interferenza e diffrazione restano validi e sono più convenientemente interpretati in termini ondulatori.

Esercizio 31 *Determinare la densità media dei fotoni, n_m , e il cosiddetto “mean photon flux density”, Φ , che corrisponde al valor medio della densità del flusso dei fotoni e si misura in $[m^{-2} s^{-1}]$, a una distanza r di 100 [Km] da un'antenna isotropa che irradia nel vuoto $P_{ir} = 100$ [W] (potenza attiva) alla frequenza di 100 [MHz]. (suggerimento: per ogni metro quadrato della superficie sferica di raggio r centrata sull'antenna e per ogni secondo viene trasmessa una energia di $\frac{P_{ir}}{4\pi r^2}$ $[J m^{-2} s^{-1}]$ e quindi transitano $\Phi = \frac{P_{ir}}{4\pi r^2 hf}$ fotoni. La densità media dei fotoni è quindi $n_m = \frac{\Phi}{c_0}$ perchè in un secondo attraversano ogni metro quadrato di superficie tutti i fotoni a distanza inferiore o uguale a c_0 . Può essere più intuitivo ragionare sull'energia trasmessa attraverso un metro quadrato di superficie in $dt = \frac{1}{c_0}$ [s] che, divisa per l'energia del singolo fotone, determina direttamente n_m).*

Esercizio 32 *Una lampadina irradia nel vuoto $P_{ir} = 100$ [W] (potenza attiva) di luce verde ($\lambda_0 = 500$ [nm]), distribuendola uniformemente nello spazio (radiazione isotropa). Quanti fotoni giungono a una distanza r di 2 [m] dalla lampadina in un secondo per ogni centimetro quadrato? E sulla sezione trasversale di un atomo, considerando un diametro tipico pari a 100 [pm]? Quanto vale la densità media dei fotoni a due metri di distanza in $[cm^{-3}]$? (suggerimento: come prima, per ogni centimetro quadrato della superficie sferica di raggio r centrata sulla lampadina e per ogni*

secondo viene trasmessa un'energia di $\frac{P_{tr}}{4\pi r^2} \simeq 2 [J m^{-2} s^{-1}]$ e quindi giungono $\Phi = \frac{P_{tr}}{4\pi r^2 h f} \simeq 5 \cdot 10^{18}$ fotoni per metro quadrato e per secondo. La sezione trasversale dell'atomo considerato è pari a $\frac{\pi}{4} 10^{-20} \simeq 0.785 \cdot 10^{-20} [m^2]$ e quindi il numero di fotoni che arrivano sull'atomo considerato in un secondo è pari a $\simeq 5 \cdot 10^{18} \cdot 0.785 \cdot 10^{-20} \simeq 3.93 \cdot 10^{-2}$. Quindi sul singolo atomo considerato arriva in media un fotone ogni circa 25 [s]. $n_m = \frac{\Phi}{c_0} \simeq 1.67 \cdot 10^{10} [m^{-3}]$.

Osservazione 28 L'esercizio precedente mette in evidenza che ogni atomo riceve, nelle condizioni di illuminazione indicate, un fotone ogni 25 secondi. In queste condizioni, quindi, risulta poco probabile l'interazione di uno stesso atomo con più fotoni in un breve intervallo di tempo. Per rendere probabile il fenomeno su scale temporali dello stesso ordine di grandezza della dinamica atomica ($10^{-17}, 10^{-13}$) [s] servirebbe una potenza enorme, che darebbe ben altri effetti macroscopici.

3.4 Fotoni e materia

Come è stato indicato all'inizio del capitolo, la comprensione del funzionamento di alcuni componenti ottici richiede lo studio dell'interazione dei fotoni con la materia.

I fotoni interagiscono naturalmente con cariche elettriche, essendo per definizione i quanti del campo elettromagnetico.

L'interazione di fotoni e materia viene studiata in Quantum ElectroDynamics (QED) [1] (p. 445), [14].

Noi procederemo in modo semplice, senza sviluppare i dettagli della QED.

Cariche elettriche possono essere non vincolate o vincolate.

Le prime sono di maggior interesse in astrofisica e determinano una interazione fotone-carica elettrica libera. In particolare, i fotoni incidenti a basse energie (quando il campo elettromagnetico dell'onda incidente è a frequenza non troppo elevata) accelerano la carica elettrica libera inducendo l'emissione di radiazione della stessa frequenza dell'onda incidente. In questo modo l'onda incidente viene diffusa e si parla di scattering Thomson. Quando i fotoni incidenti hanno energia maggiore lo scattering è detto Compton ed è un fenomeno non classico che si gestisce imponendo la conservazione dell'energia e della quantità di moto nell'urto tra un fotone e una carica elettrica (tipicamente un elettrone). L'interazione determina una riduzione della frequenza dei fotoni scatterati rispetto a quella dei fotoni incidenti.

Più interessante per noi è il caso di interazione tra fotoni e cariche elettriche vincolate ad atomi, molecole (anche di gas o liquidi) o cristalli (solidi e solidi amorfi, che non presentano una disposizione ordinata delle posizioni degli atomi o delle molecole che lo costituiscono su distanze significative ma che, ai fini delle interazioni e delle frequenze di interesse, sono ad essi assimilabili).

Questa interazione è possibile ma è soggetta a vincoli (conservazione dell'energia o dei momenti lineare ed angolare, ad esempio, dove ciascuna di queste quantità delle particelle subatomiche è quantizzata). Gli aspetti più importanti dell'interazione dei fotoni con la materia (ma non tutti! ad esempio per i LED con direct bandgap semiconductors vs indirect bandgap semiconductors intervengono considerazioni sulla quantità di moto) sono riconducibili a valutazioni di tipo energetico (cioè alla legge di conservazione dell'energia e alla quantizzazione dell'energia delle particelle subatomiche). Per questa ragione, inizialmente cercheremo di capire i livelli quantizzati ammessi in atomi isolati, molecole o solidi.

3.4.1 Livelli energetici associati a differenti tipi di materia

Energia elettrica di configurazione di cariche elettriche a riposo nel vuoto, U_{ec} .

Caso di cariche elementari a riposo di segno opposto come configurazione di interesse per lo studio dell'atomo di idrogeno: $U_{ec} = -\frac{1}{4\pi\epsilon_0} \frac{e^2}{r}$, dove e indica la carica elementare ($1.6021766208(98) \cdot 10^{-19}$ [C]) e r indica la distanza tra le cariche elementari. Si noti che U_{ec} è negativa e cresce a mano a mano che si allontanano le cariche.

Se si considera un sistema di riferimento solidale ad una delle due cariche, è possibile che l'altra carica sia in movimento. Questa carica in moto possiede una energia cinetica, U_c .

Valutazione classica dell'energia cinetica di una carica elementare negativa che ruota intorno a una carica elementare positiva con orbita circolare di raggio r [12] (p. 597) ($x(t) = r \cos(\omega t)$, $y(t) = r \sin(\omega t)$, $x'(t) = -r\omega \sin(\omega t)$, $y'(t) = r\omega \cos(\omega t)$, $x''(t) = -r\omega^2 \cos(\omega t)$, $y''(t) = -r\omega^2 \sin(\omega t)$, $|\mathbf{f}| = \frac{e^2}{4\pi\epsilon_0 r^2} = m|\mathbf{a}| = mr\omega^2$ da cui $mr^2\omega^2 = \frac{e^2}{4\pi\epsilon_0 r}$; $|\mathbf{v}| = r\omega$ e quindi $U_c = \frac{1}{2}m|\mathbf{v}|^2 = \frac{1}{2}mr^2\omega^2 = \frac{1}{8\pi\epsilon_0} \frac{e^2}{r}$; m è la massa a riposo dell'elettrone, pari a $\simeq 9.109 \cdot 10^{-31}$ [Kg]).

$U_t = U_{ec} + U_c$ fornisce l'energia complessiva del sistema di due cariche nel vuoto, di cui una in movimento rispetto a quella considerata a riposo nel sistema di riferimento scelto.

Visto che l'energia cinetica è associata alla carica in moto, spesso si associa l'intera energia U_t a tale carica e non al sistema di due cariche considerato.

Nel modello atomico di Bohr si ha quindi $U_t = U_{ec} + U_c = -\frac{1}{8\pi\epsilon_0} \frac{e^2}{r}$.

Dato il raggio r sono determinate U_t , U_{ec} , U_c , la velocità, la quantità di moto e il momento angolare.

Ipotesi di Bohr per la quantizzazione del raggio dell'atomo di idrogeno ($L = n \frac{h}{2\pi}$) e De Broglie ($p = \frac{h}{\lambda}$ dalla formula di p per i fotoni trovata con l'elettromagnetismo classico (!) e $2\pi r = n\lambda$ (se c'è onda di uno stato stazionario dell'elettrone, qualunque sia il senso attribuito all'onda (solo dopo Pauli (!)), deve essere periodica), $n \in \mathbb{N}$, $n > 0$, da cui $p = \frac{hn}{2\pi r}$; ma $L = pr$ e quindi ipotesi di Bohr immediata; inoltre $2U_c m = m^2 v^2$ da cui $p = mv = \sqrt{2U_c m} = \frac{e}{2} \sqrt{\frac{m}{\pi\epsilon_0 r}}$ e quindi $\frac{hn}{2\pi r} = \frac{e}{2} \sqrt{\frac{m}{\pi\epsilon_0 r}}$ da cui $\frac{hn}{\sqrt{\pi\epsilon_0 r}} = e \sqrt{\frac{m}{\epsilon_0}}$ e quindi...). Formula raggio dell'atomo di idrogeno $r = \frac{\epsilon_0 h^2 n^2}{e^2 \pi m}$. Il valore del raggio dell'orbita dell'elettrone nel modello di Bohr dell'atomo di idrogeno è quindi quantizzato!

Conseguenza: è quantizzata anche l'energia complessiva dell'atomo, che assume i seguenti valori $U_{t,n} = -\frac{me^4}{8\epsilon_0^2 h^2 n^2}$ [15] (p. 431), [12] (pp. 595-617). Anche U_{ec} , U_c , la velocità di rotazione dell'elettrone, p , L e la frequenza di rotazione dell'elettrone intorno al nucleo assumono valori quantizzati.

La figura 3.4 mostra l'andamento grafico di U_t dell'atomo di idrogeno secondo Bohr in funzione di n . Se n potesse variare con continuità si otterrebbe la curva continua riportata in figura 3.4. Ciò ovviamente non accade e i valori permessi di U_t per l'atomo di idrogeno secondo Bohr sono messi in evidenza dalle rette orizzontali. n è noto come numero quantico principale.

All'elettrone dell'atomo di idrogeno sono consentiti i soli salti che gli consentono di passare da un livello energetico permesso a un altro livello energetico permesso. In questi salti, se l'energia dello stato stazionario finale U_{t,n_2} è inferiore a quella dello stato di partenza U_{t,n_1} ($n_1 > n_2$), l'atomo di idrogeno può emettere un fotone di frequenza: $hf = U_{t,n_1} - U_{t,n_2}$. Ne segue che le frequenze dei fotoni emessi dall'atomo di idrogeno sono $\frac{me^4}{8\epsilon_0^2 h^3} \left(\frac{1}{n_2^2} - \frac{1}{n_1^2} \right)$, dove $n_i \in \mathbb{N}$, $n_i > 0$, $i = 1, 2$, e $n_1 > n_2$.

Questi valori sono in accordo con i risultati sperimentali relativi allo spettro (frequenza dei fotoni) dell'atomo di idrogeno [15] (p. 432), [12] (p. 600).

E' possibile anche il fenomeno opposto: illuminando gli atomi di idrogeno è possibile che dei fotoni di frequenza opportuna vengano assorbiti dagli atomi stessi. In questo fenomeno di assorbimento un elettrone, assorbendo integralmente un fotone indivisibile, passa da uno stato stazionario a energia inferiore ad uno a energia superiore.

Entrambi questi fenomeni di emissione o assorbimento sono rappresentati graficamente in figura 3.5.

Esercizio 33 Calcolare l'energia massima necessaria ad allontanare definitivamente l'elettrone dal protone dell'atomo di idrogeno secondo Bohr (suggerimento: usare la formula precedente per $n_2 = 1$ e $n_1 = \infty$; risultato: $\simeq 2.17 \cdot 10^{-18}$ [J] corrispondente a un fotone di $\lambda_0 = 91.7$ [nm] (nella banda ultravioletta)).

Esercizio 34 Calcolare la quantità di moto di un elettrone che si trova nello stato stazionario determinato da $n = 1$ (suggerimento: $U_c \simeq 2.17 \cdot 10^{-18}$, $p = \sqrt{2U_c m} \simeq 2 \cdot 10^{-24}$).

Esercizio 35 Calcolare la quantità di moto di un fotone di lunghezza d'onda nel vuoto λ_0 pari a 100 [nm] (non molto diversa da quella di un fotone in grado di ionizzare l'atomo di idrogeno; suggerimento: $p = \frac{h}{\lambda_0} \simeq 6.626 \cdot 10^{-27}$).

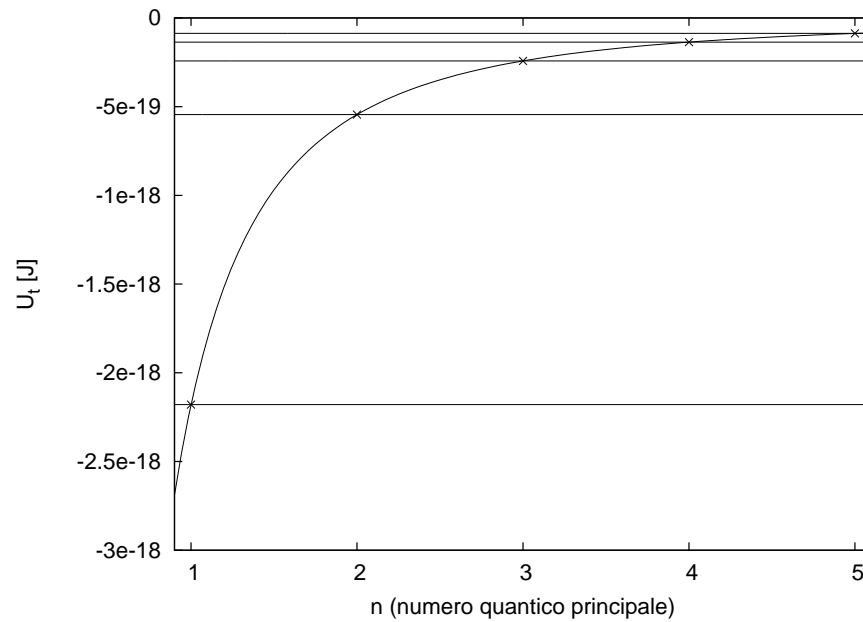


Figura 3.4: Andamento dell'energia totale dell'atomo di idrogeno secondo Bohr in funzione del numero quantico principale. La curva continua fornisce l'andamento non quantizzato che si otterrebbe se il numero quantico principale potesse variare con continuità. I punti sulla curva e le corrispondenti linee orizzontali forniscono i valori quantizzati della stessa quantità. Come unità di misura per l'energia si fa spesso riferimento all'elettronvolt, pari a $\simeq 1.602 \cdot 10^{-19}$ [J].

Le orbite sono tracciate senza considerare i loro possibili raggi

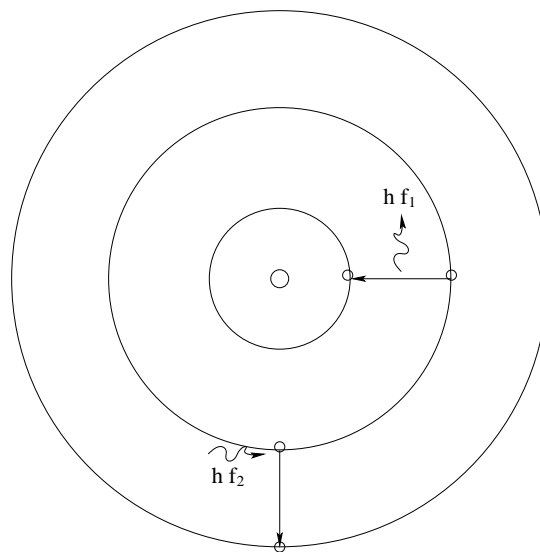


Figura 3.5: Rappresentazione grafica dell'atomo di idrogeno secondo Bohr in diversi possibili stati stazionari e possibili transizioni tra stati stazionari. I valori dei raggi non sono in scala.

Osservazione 29 *Gli esercizi 34 e 35 chiariscono perchè nei processi di emissione o assorbimento di fotoni da parte di cariche vincolate ad atomi, molecole o cristalli è spesso possibile trascurare la variazione della quantità di moto delle cariche stesse [1] (p. 664).*

Osservazione 30 *La banda visibile comprende la radiazione di λ_0 tra i 400 e i 700 [nm] ((630,700) rosso, (590,630) arancione, (570,590) giallo, (500,570) verde, (450,500) blu, (400,450) viola); la banda ultravioletto si estende dai 10 ai 400 [nm]; i raggi X da 1 [pm] a 10 [nm], i raggi γ hanno una $\lambda_0 < 1$ [pm]. Dal lato opposto della banda visibile abbiamo la banda dell'infrarosso, che si estende da 700 [nm] a diverse centinaia di μm].*

Esercizio 36 *Calcolare il raggio dell'orbita dell'elettrone dell'atomo di Bohr quando si trova nello stato stazionario determinato da $n = 1$ ($r_1 \simeq 52.93$ [pm]).*

Osservazione 31 *E' noto che il diametro degli atomi neutri delle prime sei righe della tabella degli elementi è contenuto nel range (60,600) [pm] [16] (pp. 815-822).*

Con questa descrizione quasi classica non si possono spiegare tutti i dettagli dello spettro degli atomi di idrogeno, anche perchè sono stati considerati solo i fenomeni dominanti, trascurando tutte le altre proprietà delle particelle coinvolte (spin, momenti di dipolo magnetico di tipo intrinseco, ecc.; inoltre, una transizione da uno stato stazionario corrispondente a $n = 2$ a uno stato stazionario corrispondente a $n = 1$, comporta la riduzione dell'energia $U_{t,n}$, coerentemente col fatto che può essere emesso un fotone, e corrispondentemente un aumento della quantità di moto p e, da questo punto di vista, la coerenza viene meno). Ciò che nel modello viene percepito come un livello energetico può essere in realtà costituito da livelli energetici consentiti molto ravvicinati (struttura fine e iperfine dello spettro). Per questi dettagli è necessario far ricorso alla teoria quantistica.

Atomi con più elettroni: situazione simile alla precedente, nei casi più generali considerati (livelli energetici raggruppati attorno a livelli energetici principali, determinati dal numero quantico principale).

Molecole: i livelli energetici di una molecola vengono determinati da tre tipi di contributo che determinano salti energetici che ricadono in diverse regioni dello spettro. Energia dovuta alle rotazioni delle molecole (transizioni tra livelli energetici diversi di energia "rotazionale" determinano emissioni a microonde o nel lontano infrarosso), energia dovuta alle vibrazioni delle molecole (transizioni tra livelli energetici diversi di questo tipo determinano emissioni nell'infrarosso) e energia dovuta ai livelli energetici degli elettroni (le cui transizioni determinano emissione nel visibile e nell'ultravioletto).

Liquidi e gas come moltitudine di singole molecole.

Solidi: sono possibili oscillazioni di singoli atomi attorno alle loro posizioni di equilibrio, deformazioni elastiche dei reticoli, a cui corrispondono livelli energetici quantizzati corrispondenti a energia di tipo meccanico, acusto-elastica se le deformazioni si propagano, ecc. Esistono anche i livelli energetici permessi agli elettroni. Questi sono praticamente coincidenti con quelli del singolo atomo costituente il reticolo, quando si considerano gli elettroni più interni. Invece gli elettroni più esterni possono assumere livelli di energia così ravvicinati da creare delle vere e proprie bande [17] (pp. 27-34). La cosiddetta banda di conduzione è la banda a più basso livello energetico che risulta essere non occupata o parzialmente occupata da elettroni. La cosiddetta banda di valenza è la banda a più alto livello energetico totalmente occupata. La banda energetica compresa tra la banda di valenza e quella di conduzione è la cosiddetta banda vietata.

Gli elettroni che si trovano ai livelli energetici più elevati sono meno legati ai nuclei atomici e possono comportarsi come portatori mobili di carica. In particolare, lo sono tutti gli elettroni della banda di conduzione. Al contrario, gli elettroni della banda di valenza non sono liberi di muoversi e non si comportano come portatori mobili di carica elettrica.

I metalli presentano una banda, di conduzione, parzialmente occupata a tutte le temperature.

I semiconduttori intrinseci a temperatura assoluta nulla hanno una banda di valenza totalmente occupata e una banda di conduzione completamente vuota. A temperatura superiore l'energia termica è sufficiente a spostare alcuni elettroni dalla banda di valenza a quella di conduzione.

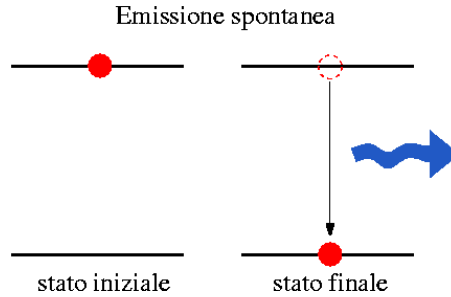


Figura 3.6: Il fenomeno dell'emissione spontanea.

Le bande energetiche degli elettroni degli isolanti hanno le stesse caratteristiche di quelle dei semiconduttori intrinseci. L'unica importante differenza riguarda l'ampiezza della banda vietata, superiore a 3 [eV] per gli isolanti e ≤ 3 [eV] per i semiconduttori intrinseci [1] (pp. 492-493). E' proprio la differente ampiezza della banda vietata a rendere meno probabile, a ogni temperatura diversa dallo zero assoluto, il passaggio di un elettrone dalla banda di valenza a quella di conduzione negli isolanti rispetto a quello che succede nei semiconduttori (ed è per questo che a temperature ordinarie la densità dei portatori mobili nei semiconduttori è circa un miliardo di volte superiore a quella dei buoni isolanti).

Ad esempio, per il silicio si ha una banda vietata di $\simeq 1.1$ [eV]; per l'arseniuro di gallio (GaAs) tale banda ha un'ampiezza di $\simeq 1.42$ [eV]; per il diamante si passa a $\simeq 5.5$ [eV].

Esercizio 37 Determinare la lunghezza d'onda del fotone potenzialmente emesso nel passaggio di un elettrone dalla banda di valenza alla banda di conduzione, nel caso del silicio, dell'arseniuro di gallio, del diamante (Si: $f = \frac{1.1 \cdot 1.6 \cdot 10^{-19}}{6.626 \cdot 10^{-34}} \simeq 0.2656 \cdot 10^{15}$, corrispondente a $\lambda_0 \simeq \frac{3 \cdot 10^8}{0.2656 \cdot 10^{15}} \simeq 11.29 \cdot 10^{-7}$ [m]. Per GaAs $\lambda_0 \simeq 11.29 \cdot 10^{-7} \frac{1.1}{1.42} \simeq 875$ [nm]. Per diamante $\lambda_0 \simeq 11.29 \cdot 10^{-7} \frac{1.1}{5.5} \simeq 226$ [nm].).

3.4.2 Interazione di fotoni e materia: fenomeni fisici elementari descrivibili in termini energetici

Come abbiamo visto la materia a cui siamo maggiormente interessati può assumere solo livelli energetici quantizzati. Una riduzione del livello energetico di atomi, molecole o di singoli portatori di carica in un reticolo cristallino può determinare l'emissione di un fotone. Se la riduzione di energia avviene tra il livello energetico E_1 e il livello energetico E_2 , con $E_1 > E_2$, allora, se viene emesso un fotone, si avrà $hf = E_1 - E_2$. Quando la riduzione del livello energetico avviene spontaneamente con corrispondente emissione di un fotone si parla di emissione spontanea. Il fenomeno è graficamente descritto nella figura 3.6. Non sono presenti novità rispetto a quanto anticipato considerando l'atomo di idrogeno.

Quando è possibile, una riduzione analoga del livello energetico può venir stimolata dalla presenza di un fotone. In questo caso è possibile che si crei una clone del fotone presente (stessa frequenza, quantità di moto, polarizzazione-spin, ecc.), in un processo noto come emissione stimolata (si veda la figura 3.7).

Il fotone che può determinare l'emissione stimolata può anche essere assorbito, in un processo opposto noto col termine assorbimento (del tutto analogo a quello presentato considerando l'atomo di idrogeno) (si veda la figura 3.7).

I fenomeni elementari descritti sono fondamentali per la realizzazione di sorgenti ottiche, amplificatori ottici, fotorivelatori. Per i LED sono particolarmente importanti l'emissione spontanea e l'assorbimento. Per i LASER e gli amplificatori ottici l'emissione stimolata e l'assorbimento. Per i fotorivelatori l'assorbimento.

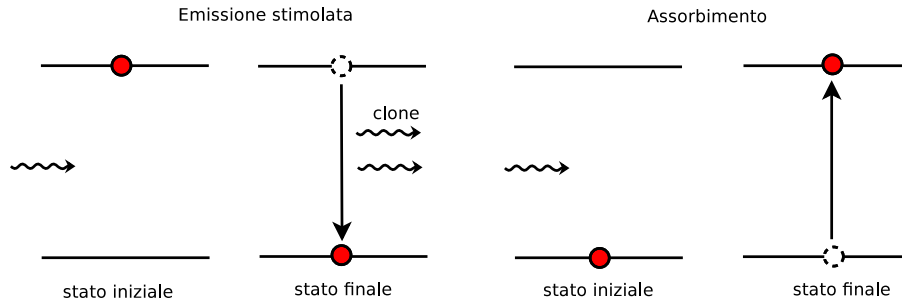


Figura 3.7: I fenomeni dell'emissione stimolata e dell'assorbimento.

3.4.3 Cicli di fenomeni fisici elementari descrivibili in termini energetici

I fenomeni elementari possono susseguirsi senza troppi vincoli, rispettando le leggi di conservazione, in particolare quella dell'energia. In figura 3.8 vengono riportati alcuni dei cicli possibili.

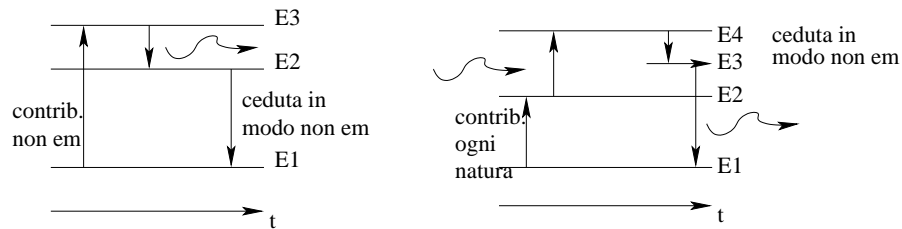


Figura 3.8: E' possibile osservare sequenze molto varie di fenomeni elementari. In figura sono riportati solo un paio di esempi.

Luminescenza: fenomeni che terminano con emissione di fotoni (da energia termica e termografia, cinetica-catodica e vecchie TV, chimica e lightstick (rottura chiusura ermetica), bioluminescenza e lucciole, sonoluminescenza da assorbimento di onde acustiche, elettroluminescenza (LED), fotoluminescenza che inizia con assorbimento di fotoni e termina con emissione di fotoni (fluorescenza è ciclo veloce; fosforescenza è ciclo lento (ordine secondi)).

Fenomeni che terminano con assorbimento di fotoni (conversione di energia dei fotoni (e non solo) in energia vibrazionale, rotatoria, degli elettroni, acustica, luminosa).

Alcuni cicli fanno riferimento a fenomeni di scattering. Per quel che riguarda lo scattering di Rayleigh, abbiamo già incontrato il fenomeno parlando di attenuazione nelle guide dielettriche ottiche. In termini di fotoni è come se per ogni fotone assorbito venisse rimesso in direzione casuale un fotone avente la stessa frequenza di quello assorbito. Ciò avviene su bande estremamente ampie ed è come se venisse coinvolta una transizione tra un livello energetico reale e uno virtuale (in pratica il fenomeno può avvenire anche in presenza di materiali che non ammetterebbero l'interazione considerata se venissero considerati i soli livelli energetici permessi).

Lo scattering di Raman è un processo in cui un fotone di energia $h f_i$ viene assorbito da un mezzo non lineare e viene subito emesso a energia $h f_e$, con $f_i \neq f_e$. Se $f_e < f_i$ si parla di Stokes scattering. Anti-Stokes scattering in caso contrario. La differenza tra l'energia del fotone assorbito e quella del fotone emesso viene fornita o ceduta in termini di energia meccanica (rotazionale o oscillatoria).

In generale, un'onda elettromagnetica viene scatterata in termini di Rayleigh e Raman scattering (Stokes e anti-Stokes) e l'onda scatterata presenta le frequenze dell'onda incidente insieme a frequenze inferiori e superiori. Lo spostamento delle frequenze dovuto al Raman scattering dipende dal tipo di materiale. Può essere molto limitato, in presenza di materiali con uno spettro vibrazionale stretto, o ampio, in presenza di materiali con uno spettro vibrazionale ampio.

Lo scattering di Brillouin è simile al Raman scattering. L'unica differenza è dovuta al fatto che lo scambio di energia avviene per mezzo di fononi (onde acustiche) invece che in termini di vibrazioni.

Abbiamo già incontrato lo scattering Thomson e Compton. Non sono particolarmente rilevanti per questo insegnamento.

I fenomeni elementari e le sequenze di fenomeni viste possono essere utilizzati per realizzare importanti componenti. Questi vengono descritti nei paragrafi seguenti.

3.5 Light Emitting Diodes e Semiconductor Optical Amplifiers

Per ottenere dei dispositivi utilizzabili come sorgenti elettromagnetiche può essere utile favorire l'emissione spontanea di fotoni da parte della materia. Un qualunque tipo di luminescenza può essere utilizzato. Un tipo di luminescenza che abbiamo imparato a controllare bene è quello indicato con il termine elettroluminescenza.

I conduttori hanno la banda di conduzione parzialmente occupata e la banda inferiore totalmente occupata; per questo sono possibili transizioni elettroniche solo tra diversi livelli della banda di conduzione, che determinano un salto energetico corrispondente a una frequenza molto bassa rispetto a quelle di interesse in questo insegnamento. Questo tipo di emissione può essere di interesse per problemi di radiazione di onde radio. In effetti molte antenne trasmettenti sono costruite utilizzando dei conduttori.

Se limitiamo il nostro interesse alle sorgenti ottiche, allora sembra ragionevole escludere dai materiali di nostro interesse i conduttori. I cristalli di semiconduttore, che a temperatura ambiente presentano densità di portatori mobili di carica molto superiori a quelle degli isolanti, come indicato in precedenza, sono di particolare interesse sia perchè rendono più probabile il fenomeno stesso rispetto agli isolanti, sia perchè il salto energetico corrisponde all'emissione di fotoni nelle bande di maggior interesse applicativo (vicino IR, visibile, UV vicino).

Senza interventi esterni, tuttavia, i cicli senz'altro presenti di transizioni dovuta a fenomeni non elettromagnetici, emissione spontanea, emissione stimolata e assorbimento non danno risultati apprezzabili.

Per giungere a livelli di emissione significativi consideriamo un diodo (giunzione p-n) [17] (pp. 35-111).

Consideriamo una giunzione p-n senza sollecitazioni esterne (a circuito aperto). Fenomeno della diffusione dei portatori. Creazione della regione di svuotamento (depletion layer) dei portatori mobili e presenza di ioni donatori e accettori dai due lati della giunzione. La regione ha uno spessore di $\simeq 0.5 \text{ } [\mu\text{m}]$ [17] (pp. 70-71). Creazione di un campo elettrostatico. Si giunge a situazione di equilibrio tra diffusione e drift. Andamento grafico delle densità delle lacune e degli elettroni lungo un asse che percorre da un capo all'altro il dispositivo.

In modo analogo, se consideriamo una sollecitazione esterna che determina una polarizzazione inversa della giunzione si determina un ampliamento della regione di svuotamento e un aumento dell'intensità del campo elettrostatico che contrasta il fenomeno di diffusione dei portatori. Andamento grafico delle densità delle lacune e degli elettroni lungo un asse che percorre da un capo all'altro il dispositivo.

Nei casi indicati non si riesce ad ottenere una zona in cui si abbiano contemporaneamente una elevata densità di elettroni e di lacune.

Diverso è il caso in cui la giunzione p-n viene polarizzata direttamente. Al crescere della tensione continua con cui il diodo viene polarizzato direttamente si verifica un progressivo restringimento della regione di svuotamento e una progressiva riduzione dell'intensità del campo elettrostatico nella zona di svuotamento. Andamento grafico delle densità delle lacune e degli elettroni lungo un asse che percorre da un capo all'altro il dispositivo. Si viene a determinare una regione, in un intorno della giunzione, dove sono presenti densità molto elevate di elettroni e lacune.

Se p e n sono le densità di lacune ed elettroni ($[m^{-3}]$) (le prime sono cariche equivalenti positive libere di muoversi nella banda di valenza e le seconde sono portatori mobili di carica elettrica negativa presenti nella banda di conduzione; non ripeteremo queste considerazioni nel seguito) il cosiddetto tasso di ricombinazione $t_r [m^{-3} s^{-1}]$ è pari a

$$t_r = r n p, \quad (3.35)$$

dove r è il cosiddetto coefficiente di ricombinazione ($[m^3 s^{-1}]$). Le ricombinazioni elettroni-lacune non sono tutte radiative (con emissione di fotoni). In particolare, $r = r_r + r_{nr}$, dove r_r (r_{nr}) è il cosiddetto coefficiente di ricombinazione radiativa (non radiativa). Per ovvie ragioni, quando si ragiona in termini di sorgenti ottiche si definisce la cosiddetta internal quantum efficiency, cioè il rendimento delle ricombinazioni in termini di emissione di fotoni, come segue:

$$\eta_i = \frac{r_r}{r}. \quad (3.36)$$

La quantità d di fotoni emessa per metro cubo e per secondo $[m^{-3} s^{-1}]$ è data da

$$d = t_r \eta_i = r n p \eta_i = r_r n p. \quad (3.37)$$

A temperatura ambiente, per il silicio si ha $r_r = 10^{-15} [cm^3 s^{-1}]$ e $\eta_i = 10^{-5}$. Corrispondentemente, per l'arseniuro di gallio $r_r = 10^{-10} [cm^3 s^{-1}]$, $\eta_i = 0.5$. Si noti che il coefficiente di ricombinazione per i due materiali indicati è simile ($r = 10^{-10}$ per il silicio e $r = 2 \cdot 10^{-10}$ per l'arseniuro di gallio) ma in termini di emissione di fotoni il loro comportamento è molto diverso (a parità di densità p e n si ottiene una differenza tra i valori di d data da un fattore 10^{-5}).

Ciò è dovuto al fatto che il silicio è un indirect bandgap semiconductor mentre il GaAs è un direct bandgap semiconductor. Questo è l'unico aspetto della nostra analisi in cui interviene la legge di conservazione della quantità di moto. I valori di energia e quantità di moto di elettroni e lacune sono certamente legati tra loro (abbiamo già incontrato la relazione $p = \sqrt{2U_c m}$ e, nel caso dell'atomo di idrogeno, si può facilmente dedurre il legame $p = p(U_t)$). Il legame per elettroni e lacune in un cristallo di semiconduttore è ovviamente influenzato dal cristallo considerato. Per i semiconduttori a bandgap diretto, il minimo di energia della banda di conduzione e il massimo di energia della banda di valenza si ottengono per lo stesso valore di quantità di moto. Al contrario, per i semiconduttori a bandgap indiretto i massimi e i minimi indicati si ottengono per valori di quantità di moto diversi. I livelli energetici più densamente popolati dagli elettroni (lacune) nella banda di conduzione (valenza) sono quelli più bassi (alti). Ne segue che le ricombinazioni elettrone-lacuna nel caso di semiconduttori a bandgap diretto può avvenire con variazioni trascurabili della quantità di moto mentre ciò è impossibile per i semiconduttori a bandgap inverso. Ma è noto che i fotoni nelle bande di maggior interesse hanno una quantità di moto molto piccola rispetto a quella di elettroni e lacune (estrapolando quello che abbiamo visto per l'atomo di idrogeno). Quindi una ricombinazione elettrone-lacuna radiativa può avvenire senza problemi nei semiconduttori a bandgap diretto mentre deve avvenire con l'intervento di un'altra particella (che fornisca il salto necessario di quantità di moto) in un indirect bandgap semiconductors. Questo ulteriore vincolo rende molto meno probabile l'evento. Con ciò si giustificano i valori molto diversi di η_i del silicio e dell'arseniuro di gallio.

Si tenga presente che le stesse difficoltà non si verificano quando il fenomeno di interesse è l'assorbimento di fotoni perchè un elettrone passa dalla banda di valenza a quella di conduzione andando in un livello energetico più elevato rispetto al minimo della banda di conduzione e poi eventualmente decade per altra via variando nel tempo la sua quantità di moto nel processo di riduzione della sua energia. Per questo potremo realizzare fotorivelatori usando semiconduttori sia a bandgap inverso sia a bandgap diretto.

Esercizio 38 A temperatura ambiente la concentrazione di elettroni e lacune per l'arseniuro di gallio intrinseco è pari a $\simeq 1.8 \cdot 10^{12} [m^{-3}]$. Si determini il numero di fotoni generato per ogni centimetro cubo e per ogni secondo (suggerimento: si ha $p = n = 1.8 \cdot 10^6 [cm^{-3}]$ e $r_r = 10^{-10} [cm^3 s^{-1}]$. Si ottengono quindi $r_r n p = 324$ fotoni per centimetro cubo per secondo.).

Esercizio 39 Si valuti la potenza ottica emessa per ogni centimetro cubo di arseniuro di gallio (suggerimento: assumendo che tutti i fotoni vengano emessi in modo da avere energia pari a 1.42 [eV] si deduce: $324\,1.42\,1.6\,10^{-19} \simeq 7.4\,10^{-17} [W\,cm^{-3}]$).

Osservazione 32 I diodi LED non devono essere molto spessi in direzione ortogonale alla giunzione dato che la parte otticamente attiva è quella intorno alla giunzione stessa. L'eventuale materia aggiuntiva in direzione ortogonale alla giunzione determinerà un incremento dell'assorbimento di fotoni, contribuendo a deteriorarne il rendimento.

Esercizio 40 In presenza di polarizzazione diretta di un diodo realizzato con arseniuro di gallio si può giungere, in un intorno della giunzione, a una concentrazione di elettroni e lacune pari a $\simeq 10^{17} [cm^{-3}]$. Si determini anche in questo caso il numero di fotoni generato per ogni centimetro cubo e per ogni secondo. Si valuti anche la potenza ottica emessa per ogni centimetro cubo di arseniuro di gallio. (suggerimento: si considera un incremento della concentrazione di elettroni e lacune di $\simeq 5\,10^{10}$ e ciò si traduce in un incremento del numero di fotoni emesso determinato da un fattore $\simeq 25\,10^{20}$).

Complessivamente, per un diodo si ottiene la seguente ragionevole legge di proporzionalità, per i fotoni emessi ogni secondo: $\Phi \propto \eta_i i$. I fotoni emessi hanno frequenza abbastanza concentrata su livelli corrispondenti a valori leggermente superiori a quelli determinati dall'ampiezza della banda vietata, ma non si ha alcun controllo in generale nè sulla direzione di propagazione dei fotoni nè sulla loro polarizzazione.

E' chiaro che se la zona attiva di un LED viene illuminata si possono determinare fenomeni di emissione stimolata e il LED si può comportare come un amplificatore ottico a semiconduttore (SOA, Semiconductor Optical Amplifier).

3.6 Principio di funzionamento dei LASER diodes

Se a un LED si aggiungono due specchi parzialmente riflettenti, che si realizzano tagliando il cristallo lungo i suoi piani cristallini e sfruttando il salto di indice di rifrazione che si ottiene tra il cristallo e l'aria circostante, si ottiene un oscillatore ottico, detto LASER (LASER diode o diode LASER). LASER è l'acronimo di Light Amplification by Stimulated Emission of Radiation.

Gli specchi parzialmente riflettenti realizzano una cavità risonante Fabry-Perot, che consente la presenza di onde di particolari frequenze e con particolari direzioni di propagazione (quindi seleziona l'energia e la quantità di moto in senso vettoriale). Il diodo genera campi elettromagnetici aventi un set ampio di frequenze e direzioni di propagazione. Tali campi vengono quasi completamente dispersi ai fini del funzionamento del laser diode. Per campi di frequenza opportuna e con quantità di moto opportuna (diretta da uno specchio all'altro) la presenza della cavità non ne contrasta l'esistenza. Consente loro di sopravvivere e di amplificarsi grazie al fenomeno dell'emissione stimolata.

3.7 Principio di funzionamento degli Erbium Doped Fiber Amplifiers

Usano fotoluminescenza in fibra drogata all'erbio (Er). Il principio di funzionamento con tre livelli di “pumping” è il seguente: grazie a una cosiddetta “pompa ottica”, ad esempio a 980 nm, si porta il livello energetico più elevato (dei tre di interesse) a popolarsi; si verifica quindi una transizione veloce non radiativa dal terzo livello energetico al secondo, con conseguente incremento della popolazione al secondo livello energetico; la transizione diretta dal terzo al primo livello energetico è poco probabile perchè è molto veloce la transizione dal terzo al secondo livello; inoltre la transizione spontanea tra il secondo e il primo livello energetico non avviene molto rapidamente; infine questa transizione corrisponde all'emissione di fotoni a 1550 [nm]; quindi, facendo propagare nel tratto di fibra drogata anche il segnale a 1550 [nm], nella condizione indicata di grande popolazione di

portatori che occupa il secondo livello energetico, posso ottenere un grosso effetto di emissione stimolata con conseguente amplificazione ottica del segnale.

La luce a 980 [nm] può essere iniettata dalla stessa parte del segnale (forward direction), dalla parte opposta (backward direction) o da entrambe le parti (bidirectional). L'ultima soluzione, anche se più complicata da realizzare, è la migliore da un punto di vista ottico in quanto consente di ottenere maggior uniformità del campo a 980 [nm] lungo il tratto di fibra drogata all'erbio e ciò si traduce in una maggior efficienza e una minor figura di rumore dell'amplificatore.

Esempio: un tratto di 50 [m] di fibra ottica in "silica glass" viene drogata con 300 ppm di ossido di erbio (Er_2O_3) e consente 30 dB di amplificazione nelle bande 1530-1565 [nm] o 1565-1625 [nm] (importantissimi per sistemi DWDM).

3.8 Principio di funzionamento dei Raman Fiber Amplifiers

I Raman Fiber Amplifiers utilizzano lo scattering di Raman stimolato e il fenomeno è simile a quello che avviene negli Erbium Doped Fiber Amplifiers: si utilizza una pompa ottica con fotoni a energia $h f_p$ e un segnale in fibra con fotoni di energia $h f_s$. Entrambi vengono fatti entrare in un mezzo ottico non lineare, con cui viene realizzato un tratto di fibra ottica. I fotoni di energia $h f_p$ fanno andare a un livello energetico virtuale elevato gli atomi del mezzo non lineare e il segnale stimola l'emissione di fotoni copia. La differenza tra l'energia dei fotoni assorbiti e quella dei fotoni emessi viene ceduta alla fibra in termini di energia vibrazionale.

La banda dei Raman Fiber Amplifiers è molto ampia poichè non è legata a livelli energetici reali (ad esempio quelli del mezzo drogante negli Erbium Doped Fiber Amplifiers) ma fa coinvolgere a livelli energetici virtuali e fenomeni vibrazionali che possono avvenire su bande enormi.

La pompa ottica può essere iniettata in forward direction, backward direction o bidirectional, come negli Erbium Doped Fiber Amplifiers.

I Raman Fiber Amplifiers sono migliori degli Erbium Doped Fiber Amplifiers in termini di banda ma sono peggiori in termini di guadagno ed efficienza.

3.9 Fotorivelatori (Photodetectors)

Sono dispositivi che servono a misurare o a far misurare il flusso di fotoni o la potenza ottica.

Esistono due grandi classi di fotorivelatori: termici o basati sull'effetto fotoelettrico.

I fotorivelatori termici convertono l'energia dei fotoni in energia termica e consentono una valutazione indiretta della potenza ottica (viene misurata la temperatura della parte che assorbe i fotoni). Si tratta di fotorivelatori generalmente inefficienti e relativamente lenti. Non ne approfondiremo l'analisi.

I fotorivelatori basati sull'effetto fotoelettrico sfruttano la possibile presenza di un numero maggiore di portatori mobili di cariche elettriche che si verifica quando dei fotoni vengono assorbiti, se i fotoni stessi sono sufficientemente energetici.

L'effetto fotoelettrico può essere interno o esterno. Quando viene utilizzato l'effetto fotoelettrico esterno si ottengono degli elettroni all'esterno del materiale che possono muoversi se il materiale si trova in un ampolla vuota in presenza di opportune differenze di potenziale che ne possono guidare il movimento. Quando si sfrutta l'effetto fotoelettrico interno si utilizza il maggior numero di portatori mobili di carica elettrica che fanno crescere la conducibilità del materiale (fotoconducibilità o "photoconductivity").

3.9.1 Fotorivelatori ad effetto fotoelettrico esterno

Indichiamo con S_e la quantità minima di energia necessaria per svincolare gli elettroni del materiale al più alto livello energetico fortemente popolato del materiale stesso. Se un fotone di energia $h f_i$ viene assorbito da tali elettroni, otteniamo, se $h f_i > S_e$, un elettrone libero esterno con energia cinetica $C_e = h f_i - S_e$. Gli elettroni che si trovano a livello energetico inferiore, se riescono a liberarsi ed uscire dal materiale, hanno un'energia cinetica inferiore.

Il valore di S_e più basso per i metalli si ha nel cesio (Cs), pari a circa 2 [eV]. Ciò determina la comparsa dell'effetto fotoelettrico esterno per fotoni di $\lambda_0 \leq 620$ [nm].

I più semplici fotorivelatori basati sull'effetto fotoelettrico esterno sono realizzati come tubi a vuoto e sono detti fotodiodi a vuoto o phototubes. Il catodo può essere opaco e lavorare in modo riflettente o essere semitrasparente e lavorare in modo trasmissivo. Grazie alla differenza di potenziale tra catodo e anodo si ottiene una corrente, proporzionale alla potenza ottica incidente. Se la tensione è sufficientemente elevata gli elettroni emessi acquistano una grande energia cinetica e possono, con un fenomeno di emissione secondaria, creare un flusso di numerosi elettroni. Così si realizzano i fotomoltiplicatori che possono rilevare e contare i singoli fotoni.

3.9.2 Fotorivelatori ad effetto fotoelettrico interno

L'effetto fotoelettrico interno viene sfruttato nei cosiddetti “photoconductive detectors”, nei fotodiodi e negli “avalanche photodiodes”.

Se ci limitiamo a considerare semiconduttori, intrinseci o meno, per ogni fotone assorbito si può creare una coppia elettrone-lacuna. La conducibilità elettrica del materiale cresce quindi in proporzione al flusso di fotoni (è noto che $\sigma = \mu q n$ per singola famiglia di portatori mobili di carica elettrica). La creazione di coppie elettroni-lacune determina un aumento del fenomeno della ricombinazione di lacune ed elettroni liberi, fenomeno la cui intensità cresce con le concentrazioni stesse di lacune ed elettroni liberi. All'equilibrio un flusso di fotoni crea concentrazioni di lacune ed elettroni liberi proporzionali al flusso stesso.

Una tensione esterna può generare una corrente elettrica misurabile in un circuito esterno, che varia in modo proporzionale al flusso di fotoni. Alternativamente si misura una caduta di potenziale ai capi di una resistenza collegata in serie. Il rendimento η con cui un fotone crea una coppia elettrone libero-lacuna è detto quantum efficiency. η dipende fortemente da λ_0 . Non è difficile in questo caso trovare semiconduttori che per salti da banda di valenza a banda di conduzione possano operare nella banda dell'infrarosso. Purtroppo per $\lambda_0 \geq 2$ [μm] l'energia dei fotoni è talmente piccola che per materiali con un band gap tanto piccolo la sola eccitazione termica è sufficiente a creare coppie elettroni liberi-lacune. In questi casi si realizzano photoconductors solo se vengono opportunamente raffreddati. Ad esempio il germanio drogato con gallio consente la creazione di elettroni liberi-lacune anche per $\lambda_0 \geq 100$ [μm] ma in questo caso si usa una refrigerazione ad elio liquido a quattro gradi Kelvin.

I fotodiodi sono diodi la cui corrente inversa cresce quando viene illuminato da fotoni. Se si considera ad esempio una giunzione p-n con bias inverso e tale giunzione viene investita da fotoni le coppie di portatori generate si muovono in direzione non randomica solo se create nello strato di svuotamento e in un suo piccolo intorno e la corrente creata per effetto fotoelettrico è sempre nella direzione inversa, dalla regione n a quella p. I fotodiodi sono generalmente più veloci dei photoconductive detectors perchè il campo elettrico nel depletion layer è molto intenso e determina una velocità superiore per i portatori fotogenerati. Si migliorano le prestazioni considerando diodi p-i-n che allargano il depletion layer aumentando la regione (il suo volume, in particolare) utile alla conversione fotoni-portatori. Un tipico indium gallium arsenide p-i-n photodetectors che opera a 1550 [nm] ha una quantum efficiency pari a circa 0.75.

Si usano anche Schottky-barrier photodiodes, cioè metal-conductor photodiodes. Il metallo è usato al posto della parte p o n. Si ottengono fotodiodi con tempi di risposta dell'ordine dei [ps] e banda dell'ordine dei 100 [GHz].

Gli avalanche photodiodes sono reverse-biased photodiodes con tensione di bias alta in cui si crea un campo elettrico molto intenso nella giunzione. I portatori di carica fotogenerati possono acquisire, grazie al capo elettrico, una energia sufficiente a generare nuovi portatori in un processo denominato ionizzazione da impatto.

Indice

1	Introduzione, motivazioni e cenni storici	2
2	Propagazione nelle guide dielettriche	4
2.1	Modi di una lastra dielettrica	7
2.2	Ortogonalità e completezza dei modi: eccitazione dei campi e propagazione in termini di modi	30
2.3	Onde (o modi) leaky	31
2.4	Esercitazione in laboratorio con COMSOL Multiphysics relativa alla eccitazione dei modi di uno slab dielettrico e alla propagazione in termini di modi	32
2.5	Caratteristiche delle guide dielettriche più importanti	34
2.5.1	Fibre ottiche	34
2.6	Attenuazione nelle guide dielettriche	37
2.7	Dispersione nelle guide dielettriche	38
2.7.1	Effetto della propagazione sulla ampiezza e durata del segnale modulante	41
2.7.2	Effetto della propagazione sulla frequenza istantanea del segnale	43
2.7.3	Simulazioni e risultati	43
2.7.4	Parametri utilizzati per valutare la dispersione	50
2.7.5	Effetto della dispersione sulla propagazione modale	50
3	Componenti ottici	51
3.1	Lastre dielettriche accoppiate	51
3.1.1	Deduzione del problema differenziale che individua le funzioni incognite $c_m(z)$, $m = 1, 2$	52
3.1.2	Soluzione del problema differenziale che individua le funzioni incognite $c_m(z)$, $m = 1, 2$	55
3.1.3	Considerazioni applicative	56
3.1.4	Simulazioni e risultati	57
3.2	Effetto elettroottico	59
3.3	Proprietà essenziali dei fotoni	60
3.4	Fotoni e materia	61
3.4.1	Livelli energetici associati a differenti tipi di materia	61
3.4.2	Interazione di fotoni e materia: fenomeni fisici elementari descrivibili in termini energetici	65
3.4.3	Cicli di fenomeni fisici elementari descrivibili in termini energetici	66
3.5	Light Emitting Diodes e Semiconductor Optical Amplifiers	67
3.6	Principio di funzionamento dei LASER diodes	69
3.7	Principio di funzionamento degli Erbium Doped Fiber Amplifiers	69
3.8	Principio di funzionamento dei Raman Fiber Amplifiers	70
3.9	Fotorivelatori (Photodetectors)	70
3.9.1	Fotorivelatori ad effetto fotoelettrico esterno	70
3.9.2	Fotorivelatori ad effetto fotoelettrico interno	71

Bibliografia

- [1] B. E. A. Saleh and M. C. Teich. *Fundamentals of Photonics*, volume 1. Wiley, New York, USA, 2nd edition, 2007.
- [2] M. Fox. *Quantum optics: an introduction*. Oxford University Press, Oxford, UK, 2006.
- [3] R. Loudon. *The quantum theory of light*. Oxford University Press, Oxford, UK, 3rd edition, 2000.
- [4] G. P. Agrawal. *Fiber-optic communication systems*, volume 1. Wiley interscience, New York, USA, 3rd edition, 2002.
- [5] 2015 INTERNATIONAL YEAR OF LIGHT and Light-Based Technologies. <http://www.light2015.org/Home.html>.
- [6] D. Marcuse. *Light transmission optics*, volume 1. Van Nostrand Reinhold Company, New York, USA, 1st edition, 1972.
- [7] C. G. Someda. *Onde elettromagnetiche*. UTET, Turin, Italy, 1986. In Italian.
- [8] A. B. Carlson. *Communication Systems*. McGraw-Hill, New York, 3rd edition, 1986.
- [9] D. Marcuse. *Theory of dielectric optical waveguides*, volume 1. Academic Press Company, New York, USA, 1st edition, 1974.
- [10] C.-L. Chen. *Foundations for guided-wave optics*. Wiley, New Jersey, USA, 1st edition, 2007.
- [11] J. D. Jackson. *Classical electrodynamics*. Wiley, New York, 3rd edition, 1999.
- [12] D. Halliday and R. Resnick. *Fisica 2*. Casa Editrice Ambrosiana, Milan, Italy, 1982.
- [13] R. P. Feynman. *QED. La strana teoria della luce e della materia*. Gli Adelphi, 2003.
- [14] C. Cohen-Tannoudji, J. Dupont-Roc, and G. Grynberg. *Photons and Atoms - Introduction to Quantum Electrodynamics*. Wiley, New York, 1997.
- [15] D. Pescetti. *Elettromagnetismo*. Piccin, Padova, Italy, 1985.
- [16] B. A. Averill and P. Eldredge. *Principles of General Chemistry (v. 1.0)*. Creative Commons licensed, 2012.
- [17] J. Millman and C. C. Halkias. *Microelettronica*. Boringhieri, Torino, 1987.