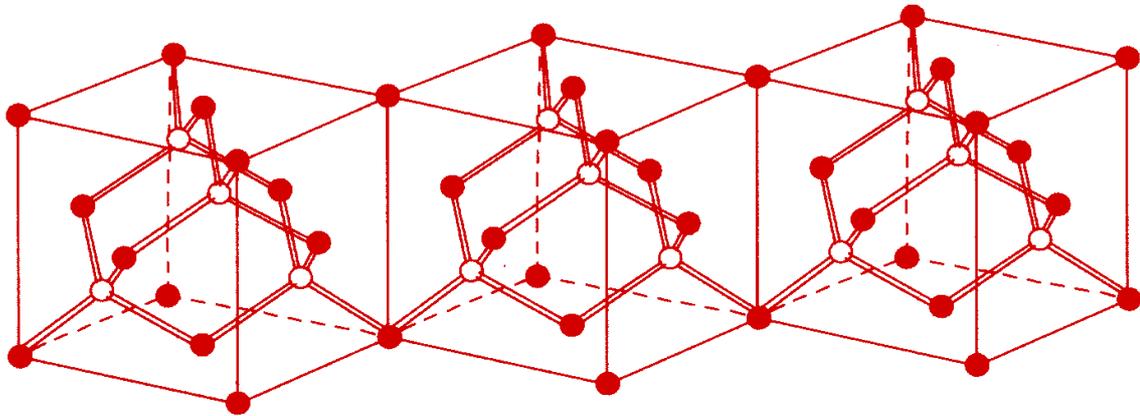


PHYSIQUE DES SEMI-CONDUCTEURS II

Applications Spéciales



(Phys 233)

Marc Haelterman
Service d'Optique et d'Acoustique
Faculté des Sciences Appliquées
Université Libre de Bruxelles

Avant Propos

Ce cours a pour but d'enseigner les éléments de base de la physique des interactions lumière-matière dans les semi-conducteurs et de présenter leurs applications les plus courantes dans le domaine de la génération et de la détection de lumière. Puisque, d'une part, le cours s'adresse à des étudiants de formations variées et que, d'autre part, la physique des semi-conducteurs fait appel à des notions de mécanique quantique élaborées, il m'a fallu débiter le cours avec un rappel des notions élémentaires de la physique des semi-conducteurs et du rayonnement électromagnétique. L'exposé de ces notions a été volontairement simplifié (parfois au dépend de la rigueur scientifique) de façon à établir le meilleur compromis entre le temps qui y est consacré et une bonne compréhension du cours. De même, le cours dans sa globalité est conçu de manière à développer une bonne compréhension des mécanismes physiques qui sont à la base des applications optoélectroniques des semi-conducteurs.

Une importance particulière est donc donnée au contenu physique des développements théoriques plutôt qu'à leur aspect mathématique. Pour cette raison, les développements théoriques sont, parfois au détriment de leur rigueur, fortement simplifiés par rapport à ceux rencontrés dans la plupart des livres traitant de l'optique des semi-conducteurs.

Le cours est structuré selon un fil conducteur dont la logique permet de mettre en évidence l'énorme progrès réalisé par l'homme dans le domaine de la maîtrise de la lumière. Le cours passe, en effet, progressivement de la description des sources de rayonnement thermique aux sources laser à semi-conducteur. Ces dernières constituent les sources de lumière les plus performantes aussi bien du point de vue de leur rendement que pour la qualité de leur rayonnement. C'est, en particulier, grâce aux lasers à semi-conducteur que les télécommunications optiques (dont on connaît l'importance aujourd'hui) ont pu être développées. L'atout majeur des télécommunications optiques est qu'elles sont basées sur la transmission de données par modulation d'une onde porteuse dont la fréquence est extrêmement élevée ($\sim 10^{15}$ Hz) et permet donc d'atteindre des taux de transmission inaccessibles aux techniques classiques (ondes hertziennes ou microondes). Ainsi, en ce qui concerne les applications, la motivation principale de ce cours est de montrer comment la connaissance et la maîtrise des cristaux semi-conducteurs ont permis la réalisation de générateurs et détecteurs de telles ondes porteuses pour les télécommunications.

février 1998

Marc Haelterman

1. INTRODUCTION

Ce chapitre est divisé en deux parties. La première donne une introduction aux *télécommunications optiques*. Son but est de montrer que le *laser* à semi-conducteur constitue le dispositif de base des télécommunications optiques. Cette première partie permet ainsi de motiver le cours en montrant que la maîtrise des matériaux semi-conducteurs est à l'origine des télécommunications modernes. La deuxième partie est consacrée à un rappel des principes de base de la physique de semi-conducteurs. Nous passons brièvement en revue les modèles de *l'électron libre* et de *l'électron quasi-libre* d'un cristal. Ceci nous permet d'introduire les concepts de base de la physique des semi-conducteurs tels que la *structure de bande* des diagrammes d'énergie. Des notions, utiles pour la compréhension des interactions lumière-matière, telles que la *densité d'états électroniques* par unité de volume et d'énergie sont également revues.

1.1 Introduction aux télécommunications optiques

Les télécommunications à fibre optique ont été développées à un rythme croissant depuis maintenant plus de 20 ans. La maturité de cette nouvelle technologie est bien illustrée par la vitesse à laquelle les liaisons optiques sont actuellement déployées au travers des continents et des océans. Les progrès réalisés ces dernières décades peuvent être mesurés en remarquant que les premières liaisons optiques trans-océaniques pouvaient transmettre jusqu'à 80.000 conversations téléphoniques simultanées alors que les premiers câbles coaxiaux trans-océaniques voyaient leur capacité de transmission saturer à 36 voix simultanées. Cette progression est le résultat de nombreux efforts de développement de la technologie des fibres optiques et des dispositifs électro-optiques semi-conducteurs : les diodes électroluminescentes, les photodiodes et surtout les lasers. Dans ce qui suit nous allons voir pourquoi le laser à semi-conducteur joue un rôle si important en télécommunications optiques.

1.1.2 La fibre optique

Les télécommunications optiques peuvent être vues comme étant une extension des télécommunications à ondes hertziennes au domaine des fréquences ultra-hautes de l'optique. Le simple principe de modulation d'ondes porteuses pour la transmission d'information étant toujours appliqué, l'utilisation de porteuses de fréquence optique (10^{15} Hz) permet des taux de transmission beaucoup plus élevés qu'avec les ondes hertziennes dont les fréquences ne dépassent pas l'échelle des Gigahertz (microondes). Les deux obstacles majeurs qu'il a fallu surmonter pour accéder aux télécommunications optiques sont d'une part la réalisation d'une source d'onde lumineuse monochromatique cohérente dont le rôle est de remplacer l'antenne émettrice des télécommunications hertziennes et d'autre part la réalisation de guides d'onde optiques jouant le rôle du câble coaxial pour les liaisons point à point. L'antenne émettrice pour les télécommunications optiques est le laser à semi-conducteur tandis que le câble véhiculant les ondes est la fibre optique.

Le principe du guidage d'une onde lumineuse par une fibre optique est très simple. Il est illustré à la figure 1 ci-dessous. La fibre est un cylindre étiré de verre pur (silice, SiO_2) dopé en son centre avec un ou plusieurs autres composés tels que GeO_2 ou P_2O_5 de manière à ce que l'indice de réfraction y soit légèrement supérieur (voir Fig.1a). La zone centrale d'indice plus élevé résultante s'appelle le coeur de la fibre (core). La zone périphérique entourant le coeur s'appelle la gaine de la fibre (cladding). Pour assurer une bonne tenue mécanique de la fibre, une gaine extérieure en matériau polymère est placée à sa surface (jacket). Le guidage de la lumière s'effectue dans le coeur par réflexion interne totale sur l'interface coeur-gaine (voir Fig.1b). La réflexion totale a lieu pour autant que l'angle d'incidence sur l'interface soit suffisamment rasant (plus précisément, supérieur à l'angle critique de réflexion totale). Puisque la différence

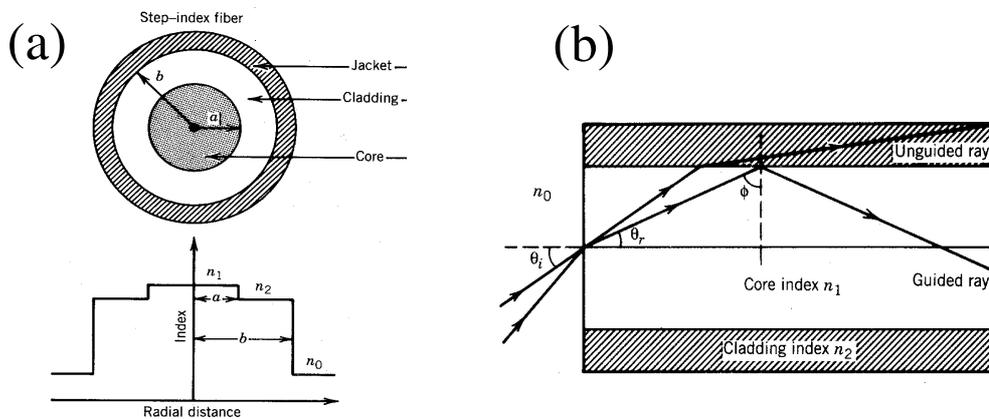


Fig.1: Illustration schématique de (a) la section transverse d'une fibre optique et de son profil d'indice de réfraction; (b) la section longitudinale d'une fibre optique, d'un rayon lumineux guidé et d'un rayon lumineux non guidé (voir texte).

d'indice entre le coeur et la gaine est très faible (de l'ordre de 10^{-2} - 10^{-3}), l'angle critique de réflexion totale est proche de 90° . Dès lors, seules les ondes lumineuses ayant une incidence quasiment normale à la surface d'entrée de la fibre seront guidées (θ_i petit, voir Fig.1b). Ceci indique que l'onde lumineuse transportant l'information devra être bien collimatée (unidirectionnelle) pour pouvoir être injectée et guidée dans la fibre sans trop de pertes. Comme nous le verrons dans la suite du cours, ce point a une grande importance pratique pour le choix des sources lumineuses utilisées en télécommunications optiques. Nous verrons notamment, que ce simple critère géométrique de qualité d'une source ne peut être raisonnablement satisfait que par les lasers. Ayant une distribution angulaire de rayonnement relativement large, les autres sources lumineuses, telles que les diodes électroluminescentes, conduisent à des pertes importantes d'énergie lumineuse lors de l'injection dans les fibres optiques.

1.1.3 La dispersion chromatique des fibres optiques

Le principe de la transmission d'information par fibre optique est illustré schématiquement à la figure 2. Idéalement, l'onde porteuse optique est générée par un laser à semi-conducteur. L'information y est inscrite à l'aide d'un modulateur électro-

optique (dans le cas de l'illustration : par modulation d'amplitude) selon un codage binaire. L'onde modulée est alors injectée dans une fibre et se propage jusqu'en bout de liaison où elle est détectée par une photodiode (jouant le rôle d'antenne réceptrice). La quantité d'informations transmises dépend du taux de modulation applicable aux ondes optiques. Puisque la fréquence de la porteuse optique est de l'ordre de 10^{15} Hz, on pourrait imaginer moduler jusqu'au Terahertz (10^{12} Hz) sans rencontrer de problème (~1000 oscillations du champ par impulsion). Pour se donner une idée du potentiel de l'optique il faut réaliser qu'une ligne de transmission à fibre optique basée sur un taux de modulation d'un Terahertz serait équivalente à plus de 150 millions de lignes téléphoniques classiques (64 KHz).

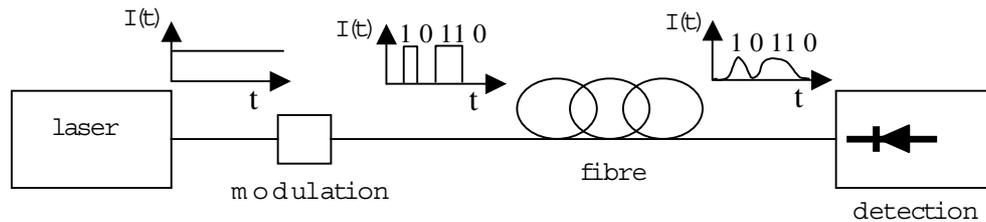


Fig.2: Illustration schématique d'une ligne de transmission de données à fibre optique.

Il existe malheureusement une limitation d'origine physique importante qui ne permet pas d'atteindre de tels taux de transmission avec les techniques actuelles. Il s'agit de la dispersion chromatique des fibres optiques. La dispersion chromatique est due au fait que l'indice de réfraction n de la silice constituant la fibre optique est une fonction fortement dépendante de la fréquence de l'onde porteuse, soit, $n = n(\omega)$ où ω est la fréquence de la lumière (dans le reste du cours nous notons ω la fréquence angulaire de la lumière $\omega = 2\pi f$). De par la définition de l'indice de réfraction, ceci signifie que si la source lumineuse utilisée n'est pas parfaitement monochromatique, chaque composante de fréquence différente se propage à une vitesse différente. Dès lors les impulsions lumineuses constituant l'information à transmettre se déforment en cours de propagation. Ce phénomène est illustré à la figure 3 où l'on voit les oscillations du champ électrique de la lumière (et leur enveloppe) au sein d'une impulsion avant et après propagation dans une fibre optique. Sur de grandes distances

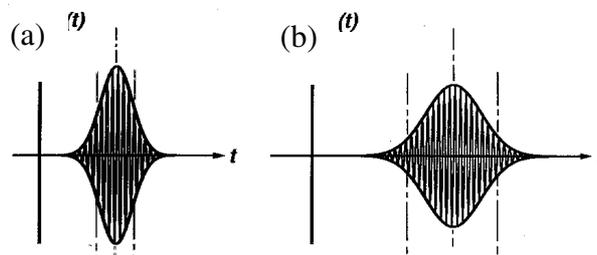


Fig.3: Représentation schématique de l'effet de la dispersion chromatique sur une impulsion lumineuse, (a) impulsion initiale; (b) impulsion après propagation dans une fibre optique.

de propagation, la déformation peut devenir si grande que les impulsions ne seraient plus distinguables entre elles et l'information serait perdue. Il est donc important

d'utiliser une source qui soit la plus monochromatique possible de façon à limiter les effets nuisibles de la dispersion chromatique sur les transmissions à grandes distances. C'est la raison pour laquelle les lasers sont utilisés en pratique. Certains lasers performants sont capables de générer des ondes quasiment monochromatiques et leur utilisation permet de minimiser les effets de la dispersion.

Il faut néanmoins garder à l'esprit que, même avec une source monochromatique idéale, la dispersion opère toujours puisque le spectre d'un signal modulé (c'est à dire, contenant de l'information) a toujours une largeur non nulle. La limite théorique de la largeur du spectre d'un signal est donnée par la relation $\Delta\nu\Delta t = 2\pi$, où Δt est la largeur des impulsions constituant le signal.

Cette brève introduction aux télécommunications optiques nous permet de comprendre l'importance du rôle du laser dans les télécommunications optiques. Par rapport aux critères de directivité, de monochromaticité et de rendement énergétique, le laser à semi-conducteur est sans conteste le dispositif le plus performant que l'homme ait pu réaliser dans sa quête vers la maîtrise de la lumière. Tout au long de ce cours, nous passerons en revue les différentes étapes qui ont été nécessaires à la réalisation du laser. Nous verrons les connaissances scientifiques et techniques qu'il a fallu rassembler dans le domaine de la physique de l'état solide et de l'ingénierie des matériaux pour arriver à cette source parfaite de lumière qui est déjà en train de révolutionner, à l'échelle planétaire, les communications entre les hommes.

1.2 Rappel de Physique des Semi-conducteurs

1.2.1 Introduction : Formalisme

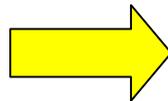
Le formalisme utilisé pour la description des propriétés des matériaux semi-conducteurs est celui de la mécanique quantique. La description détaillée et rigoureuse des S-C exige des développements théoriques relativement élaborés faisant appel, notamment, à des concepts abstraits (p.ex. Groupes de symétrie de la théorie des groupes). Néanmoins, dans le cadre du cours nous n'aborderons que les propriétés optiques élémentaires des S-C. Celles-ci peuvent être décrites et expliquées à partir de raisonnements simples faisant appel aux notions élémentaires de la mécanique quantique.

Pour comprendre les propriétés optiques des S-C abordées dans le cours il nous suffira de nous rappeler qu'en mécanique quantique l'électron est décrit comme étant une onde. Dans le cas particulier de l'électron libre il s'agit d'une onde plane dont le vecteur d'onde est proportionnel à la quantité de mouvement de l'électron. La constante de proportionnalité étant la constante de Planck divisée par 2π . La fréquence de l'onde est donnée par l'énergie de l'électron; celui-ci étant libre, il s'agit de son énergie cinétique proportionnelle au carré de la quantité de mouvement. La relation entre la fréquence (énergie) et le vecteur d'onde (quantité de mouvement) s'appelle la relation de dispersion. Dans le cas de l'électron libre, nous avons donc une relation de dispersion parabolique.

Électron libre (**classique**)

particule

$$\left\{ \begin{array}{l} p = m_e v \\ E = \frac{1}{2} m_e v^2 = \frac{1}{2m_e} p^2 \end{array} \right.$$



Postulats de la M.Q.

onde

$$\left\{ \begin{array}{l} p = \hbar k \\ E = \hbar \omega = \frac{\hbar^2 k^2}{2m_e} \end{array} \right.$$

Electron libre = onde plane : $\psi = e^{ikx} e^{-i\omega t}$

Les postulats de la mécanique quantique établissent le principe d'équivalence. Ce principe conduit à l'équation de Schrödinger lorsque l'on traduit la relation de dispersion parabolique

Principe d'équivalence :

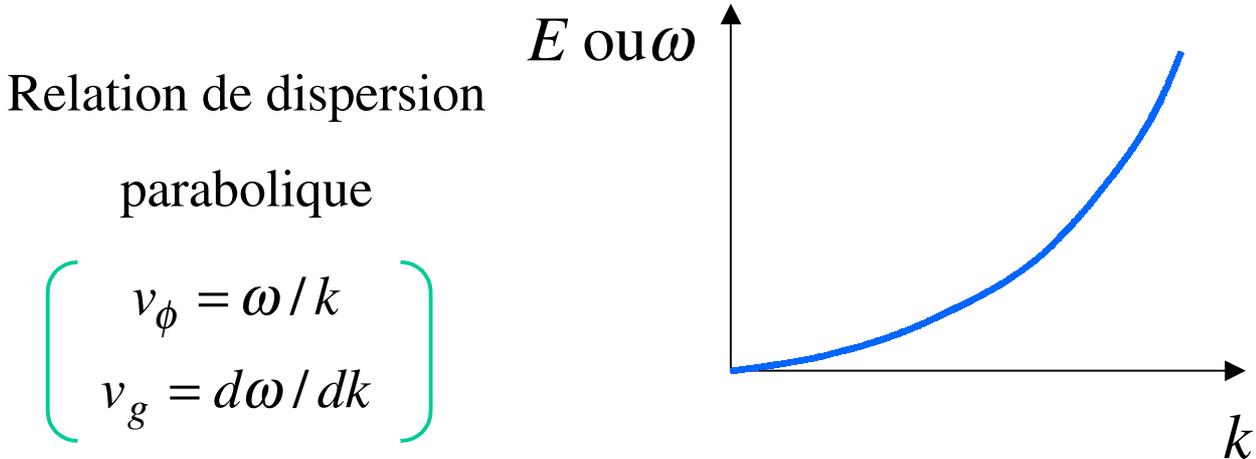
$$\left\{ \begin{array}{l} p \rightarrow -i\hbar \partial_x \\ E \rightarrow i\hbar \partial_t \end{array} \right.$$

Equation de Schrödinger :

$$i\hbar \partial_t \psi = \frac{-\hbar^2}{2m_e} \partial_{xx}^2 \psi$$

(Pour la simplicité de l'exposé nous écrivons l'équation de Schrödinger avec une seule dimension spatiale)

Nous représentons ici la relation de dispersion parabolique car on verra qu'elle joue un rôle central dans la description des S-C. La courbe de dispersion permet d'obtenir la vitesse de phase de l'onde associée à l'électron ainsi que sa vitesse de groupe (à partir de la pente de la courbe). C'est la vitesse de groupe qui représente la vitesse physique réelle de l'électron (on trouve en effet la relation classique $v_g = p/m_e$)



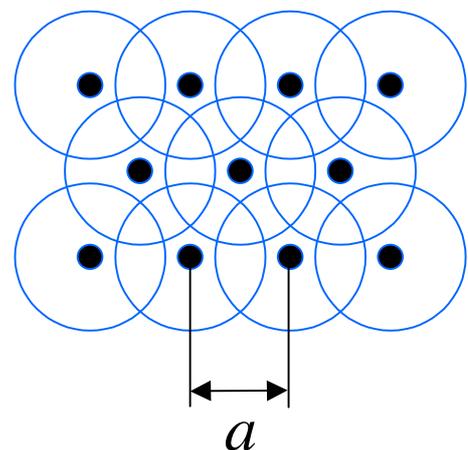
1.2.2 Modèle de l'électron libre des S-C

Les semi-conducteurs sont des matériaux cristallins. Les électrons des couches périphériques des atomes du cristal sont libérés et peuvent se déplacer, ce qui conduit à la conductivité électrique, propriété physique essentielle des semi-conducteurs. La description la plus simple des propriétés électroniques d'un cristal est obtenue à l'aide du modèle de l'électron libre. Ce modèle consiste à négliger l'effet du potentiel électrostatique des atomes (ionisés) sur les électrons libres.

Pour décrire cette situation, il suffit donc de considérer l'équation de Schrödinger de l'électron en l'absence de potentiel (on retrouve alors le cas de l'électron libre vu ci-dessus). Considérons un électron d'énergie E donnée. La fonction d'onde de cet électron sera une onde plane de fréquence ω déterminée par l'énergie E . La fréquence étant donnée, l'équation de Schrödinger devient l'équation de Schrödinger stationnaire dont la solution, en l'absence de potentiel, est $\phi = e^{ikx}$ (où k est lié à l'énergie E par la relation de dispersion parabolique)

Eq. de Schrödinger stationnaire : $\psi = \phi e^{-i\omega t}$, $E = \hbar\omega$

$$E\phi = \frac{-\hbar^2}{2m_e} \partial_{xx}^2 \phi + \cancel{V(x)}\phi$$
$$\left\{ \begin{array}{l} \phi = e^{ikx} \\ E = \frac{\hbar^2 k^2}{2m_e} \end{array} \right.$$

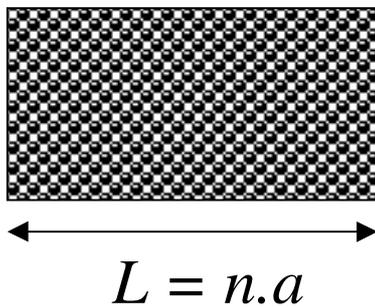


Représentation schématique d'un cristal (a : constante de réseau)

Dans le modèle de l'électron libre nous ne tenons pas compte du potentiel des atomes du réseau, nous prenons seulement en compte la nature périodique du milieu dans lequel les électrons libres se déplacent. Pour ce faire nous imposons à la fonction d'onde de l'électron d'avoir une périodicité dont la période est un multiple de la constante de réseau. A cette fin, on définit une cellule élémentaire de longueur $L = na$ (ou n est un entier). Le choix de L (ou n) est arbitraire, tous les résultats qui suivent étant indépendants de ce choix. La seule chose importante est que n soit grand.

Imposer la périodicité revient à dire que le vecteur d'onde k de toute fonction d'onde électronique doit répondre à la condition $k_m = 2m\pi/L$. Ceci signifie simplement que les électrons ne peuvent pas avoir n'importe quelle vitesse (ou énergie cinétique) dans un milieu cristallin : la relation de dispersion devient discrète.

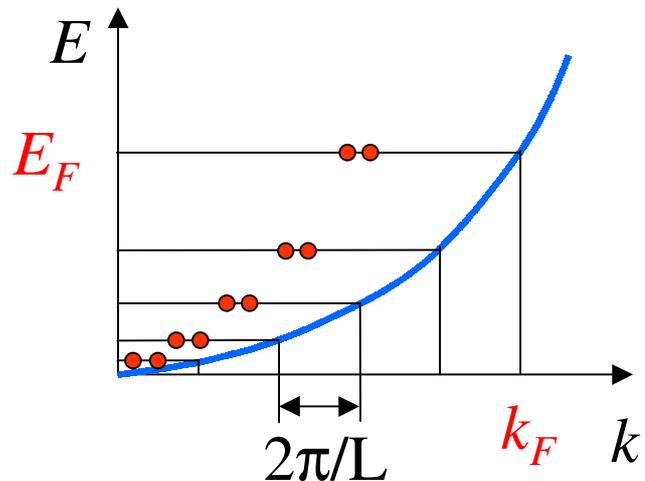
Conditions aux limites périodiques



$$\begin{cases} \phi_m = \exp(ik_m x) \\ k_m = \frac{2m\pi}{L} \end{cases}$$

Relation de dispersion
discrète

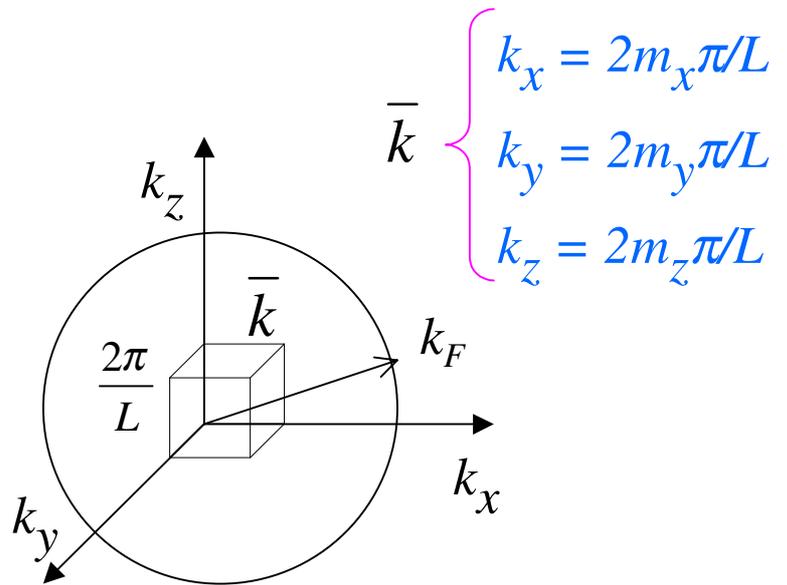
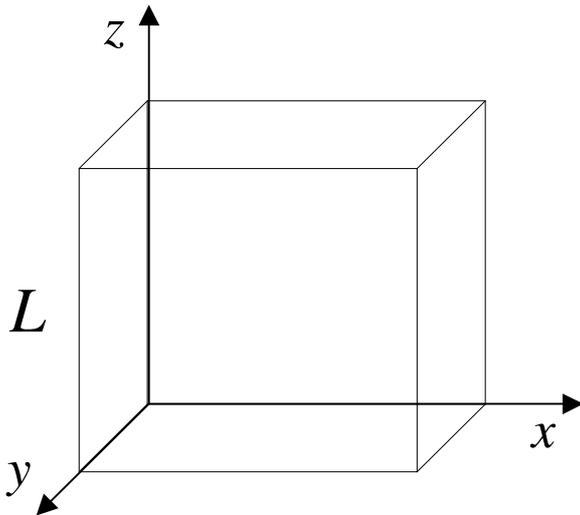
⊕ Principe de Pauli :
spin: $2 e^-$ par état



Chaque valeur de k_m (ou E_m) est associée à un état électronique possible pour un électron dans le cristal. D'après le principe de Pauli, au maximum deux électrons de spin opposé peuvent occuper un tel état. Les électrons libres dans le cristal s'organiseront donc en formant un remplissage de la courbe de dispersion discrète de façon à remplir tous les états d'énergie inférieure par paire d'électrons. Cette organisation est analogue à celle qui a lieu autour d'un noyau pour former un atome. Le remplissage des états se termine au niveau de Fermi E_F (ou k_F). Par définition, le niveau de Fermi est le niveau d'énergie atteint par le dernier électron disponible lors du remplissage des états en l'absence d'agitation thermique.

On peut facilement calculer le niveau de Fermi pour un cristal donné à partir de la relation de dispersion. En effet, le nombre d'électrons libérés dans le cristal est égal à la valence de l'atome multipliée par le nombre d'atomes; connaissant la constante de réseau on peut calculer la densité d'atome et donc également la densité électronique n_e . Dans une cellule élémentaire cubique de côté L le nombre d'électrons est donc $N_e = n_e L^3$. On peut dès lors calculer le vecteur d'onde de Fermi en sachant que : (1)- les électrons sont disposés régulièrement dans l'espace réciproque (espace des vecteurs d'onde) et y forment un empilement cubique dont la maille élémentaire a un côté de $2\pi/L$, (2)- l'empilement se fait préférentiellement sur les états d'énergie inférieure (donc sur une sphère dont le centre est l'origine) jusqu'au vecteur d'onde de Fermi, par définition.

3 dimensions :



$$\bar{k} \begin{cases} k_x = 2m_x\pi/L \\ k_y = 2m_y\pi/L \\ k_z = 2m_z\pi/L \end{cases}$$

Niveau de Fermi :

$$N_e = n_e L^3$$

$$N_e = 2 \cdot \frac{\frac{4}{3}\pi k_F^3}{\left(\frac{2\pi}{L}\right)^3}$$

$N_e = \text{nbr d'e}^-$, $n_e = \text{densité d'e}^-$

(2 → spin/Pauli)

$$k_F = (3\pi^2 n_e)^{1/3}$$

(k_F indépendant de L)

Notez que k_F est indépendant du choix de la longueur L de la cellule élémentaire et ne dépend que de la densité d'électrons (ceci est en accord avec l'idée du choix arbitraire de L mentionné plus haut).

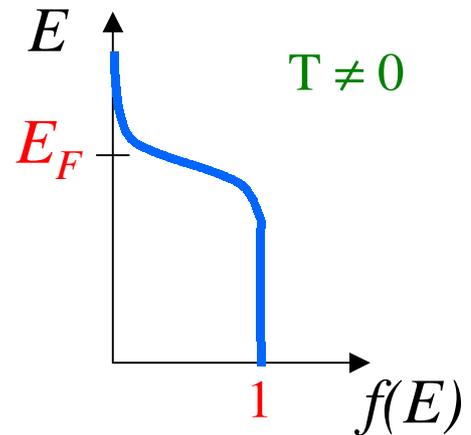
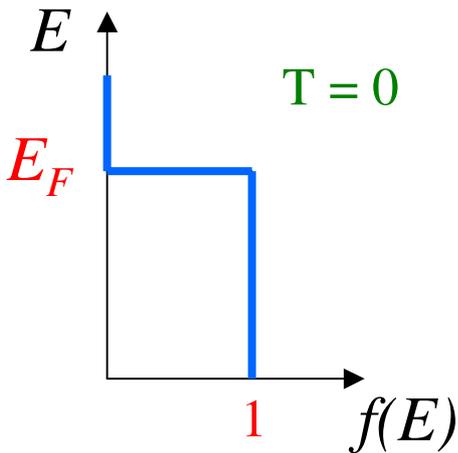
Connaissant k_F on peut facilement connaître le niveau de Fermi E_F à partir de la relation de dispersion parabolique:

$$E_F = \frac{\hbar^2 k_F^2}{2m_e} = \frac{\hbar^2}{2m_e} (3\pi^2 n_e)^{2/3}$$

Ce type de raisonnement, basé sur le fait que l'empilement des états dans l'espace réciproque est simple (cubique, dans l'exemple simplifié ci-dessus) est souvent utile pour la caractérisation des propriétés des S-C. Nous verrons que cela nous sera utile, entre autre, pour l'étude du spectre d'absorption optique des S-C.

En raison de l'agitation thermique, les dernières couches avant le niveau de Fermi ne se remplissent pas complètement car les électrons qui les occuperaient sont libres de se déplacer sur les états non-occupés supérieurs; ce qu'ils font suite à des échanges d'énergie avec les phonons du réseau cristallin et les autres électrons (agitation thermique). Les lois statistiques de la mécanique quantique nous apprennent que la répartition énergétique des électrons due à l'agitation thermique est donnée par la loi de distribution de Fermi-Dirac (appelée ici la distribution de Fermi pour la simplicité) :

Distribution de Fermi: $f(E) = 1 / \{ \exp[(E - E_F) / k_B T] + 1 \}$

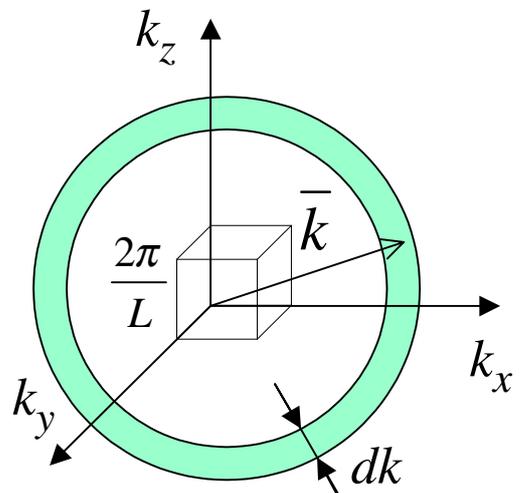


Densité d'états

Comme nous le verrons plus loin, la densité d'états par unité de volume et d'énergie est un autre concept important pour la caractérisation des propriétés optiques des S-C. Nous la calculons ici pour le cas simple d'un cristal de maille cubique. Comme pour le calcul du niveau de Fermi, le calcul de la densité d'états est basé sur le compte aisé du nombre d'états dans l'espace réciproque. Comme on veut connaître le nombre d'électrons par unité d'énergie on calcule d'abord le nombre d'électrons dans l'espace réciproque k et l'on applique ensuite la relation de dispersion parabolique. Ci-dessous, dN_e est le nombre d'électrons contenus dans la couche sphérique de rayon k et d'épaisseur dk correspondant à une énergie donnée E et un intervalle d'énergie dE . En appliquant le calcul au cas d'un volume unitaire ($L=1$) dN_e devient l'accroissement de densité électronique par unité de volume dn_e . Le calcul de la densité d'états $\rho = dn_e/dE$ est donc très simple.

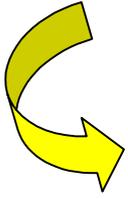
$$\rho = \frac{dn_e}{dE}$$

$$dN_e = 2 \cdot \frac{4\pi k^2 dk}{\left(\frac{2\pi}{L}\right)^3} = \frac{k^2}{\pi^2} dk$$



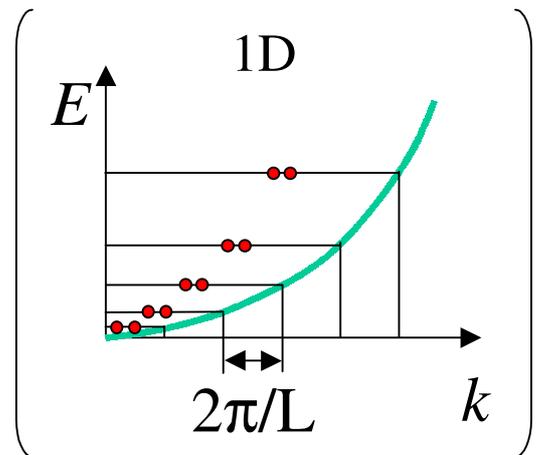
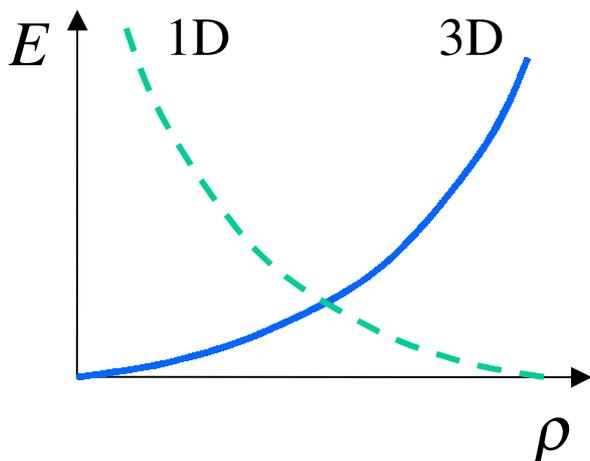
$L = 1 \rightarrow dN_e = dn_e$

$$E = \frac{\hbar^2 k^2}{2m_e} \rightarrow dE = \frac{\hbar^2}{2m_e} 2k dk \rightarrow dk = \frac{1}{2} \sqrt{\frac{2m_e}{\hbar^2 E}} dE$$



$$\frac{dn_e}{dE} = \rho = \frac{1}{2\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \sqrt{E}$$

La densité d'états est donc proportionnelle à la racine carrée de l'énergie : plus on monte en énergie plus le nombre d'états par unité d'énergie est grand. Ce qui n'est pas nécessairement intuitif puisque dans le modèle uni-dimensionnel simplifié, la densité apparente d'états diminue avec l'énergie (voir schéma entre parenthèse et la ligne en tirets sur la figure ci-dessous). La différence entre 1D et 3D vient simplement du fait que le nombre d'états dans l'intervalle dk est constant en 1D et augmente avec le carré de k en 3D.



La description des propriétés des cristaux fournie par le modèle de l'électron libre est simpliste et très limitée dans son application. S'il permet d'introduire, comme vu ci-dessus, les concepts élémentaires liés à l'agencement des états électroniques (organisation des électrons) dans les cristaux, le modèle de l'électron libre ne permet pas d'expliquer les propriétés importantes des cristaux telles que leurs propriétés de conduction électrique.

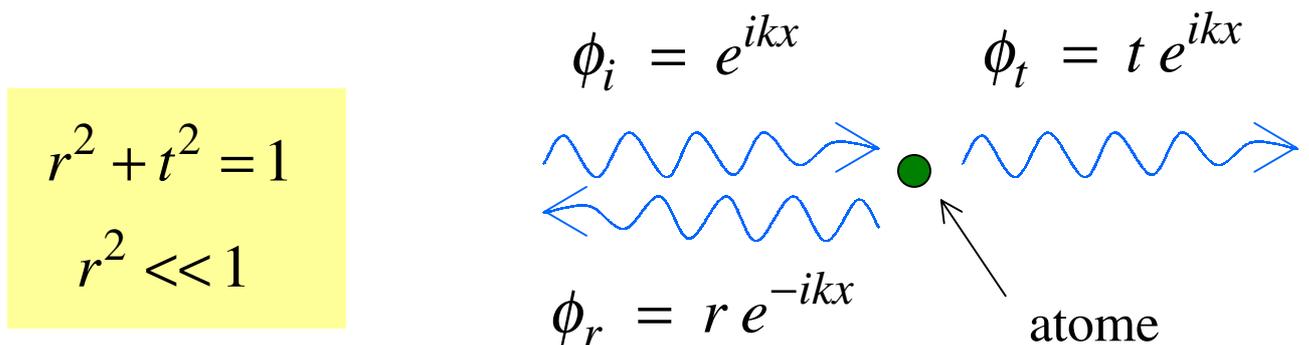
Pour expliquer les propriétés de conduction des cristaux il faut prendre en compte l'interaction des électrons libres du cristal avec le réseau d'atomes (ions) le constituant. Le modèle le plus simple qui permet cela est le modèle de l'électron quasi-libre.

2. Modèle de l'électron quasi-libre

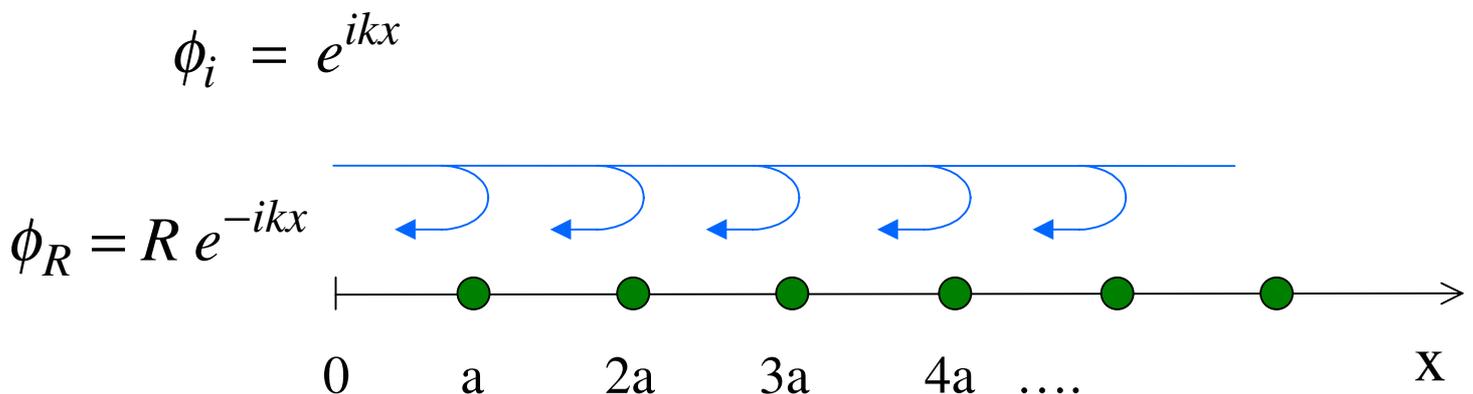
Le modèle de l'électron quasi-libre est obtenu en considérant que l'interaction des électrons avec le réseau cristallin est très faible. Dans le cadre du présent rappel de physique des S-C, nous ne développerons pas ce modèle rigoureusement; nous nous contenterons d'une description simple et intuitive des mécanismes responsables des propriétés de conduction des S-C.

Approche intuitive :

Puisque les électrons sont représentés par des ondes en mécanique quantique, l'interaction d'un électron avec un atome du réseau cristallin peut être modélisée comme une simple réflexion partielle de l'onde sur l'atome. Le modèle de l'électron quasi-libre est basé sur l'hypothèse d'une faible interaction ce qui revient à dire que le coefficient de réflexion correspondant r est petit, $r \ll 1$. Par conservation du nombre d'électrons le coefficient de transmission associé t doit répondre à la relation $r^2 + t^2 = 1$. La fonction d'onde représentant l'électron en interaction avec l'atome sera donc composée de trois ondes: une onde incidente d'amplitude unitaire et de vecteur d'onde $+k$ (par convention), une onde réfléchie d'amplitude r et de vecteur d'onde $-k$ et d'une onde transmise d'amplitude t et de vecteur d'onde $+k$.



Pour un électron se propageant dans un cristal, la réflexion a lieu de manière périodique avec une période déterminée par la constante de réseau. La situation est illustrée de manière schématique ci-dessous pour un modèle de cristal à une dimension de constante de réseau a . Même si elles sont très légères comme supposé ici, les réflexions multiples ont un impact extrêmement important sur le comportement des électrons. Pour le voir, calculons l'amplitude R de l'onde réfléchie en $x = 0$ résultante de toutes les réflexions multiples ayant lieu en $x > 0$. Ce calcul doit bien entendu prendre en compte la phase de chacune des composantes réfléchies par les atomes successifs. Pour simplifier le calcul, nous négligeons les réflexions secondaires (une fois qu'une onde est réfléchie par un atome elle n'est plus réfléchie par les atomes qu'elle rencontrent en se propageant vers $x = 0$).



Composante réfléchie en $x = a$ Composante réfléchie en $x = 2a$ (transmise 2 fois) Composante réfléchie en $x = 3a$ (transmise 4 fois) t^2 est inférieur mais proche de l'unité puisque l'interaction avec l'atome est faible

$$R = re^{2ika} + t^2 re^{4ika} + t^4 re^{6ika} + \dots$$

$(t^2 \leq 1)$

$$R = re^{2ika} \left(1 + t^2 e^{2ika} + t^4 e^{4ika} + \dots \right) = \frac{re^{2ika}}{1 - t^2 e^{2ika}}$$

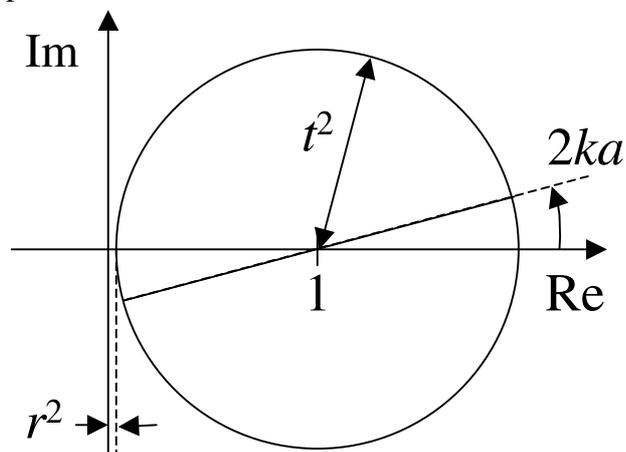
La série ainsi obtenue converge car $t^2 < 1$. La fonction résultante est périodique en k , sa période étant de $\Delta k = \pi/a$. Puisque t^2 est proche de l'unité, le dénominateur de la fonction prend une valeur proche de zéro pour $k.a = m\pi$ (m étant un entier). Ceci indique un phénomène de résonance : la réflexion de l'électron sur les sites atomiques présente des pics pour des valeurs bien précises de sa vitesse (ou énergie cinétique, ou quantité de mouvement). Entre les pics le dénominateur prend des valeurs de l'ordre de l'unité et la réflexion R est donc beaucoup plus faible. La valeur exacte de la réflexion R n'est pas fournie par la fonction ci-dessus car nous avons négligé les réflexions secondaires. Le calcul exact conduit à un facteur r^2 au numérateur plutôt que r , tel qu'indiqué ci-dessous. La réflexion R vaut donc l'unité aux résonances $k = m\pi/a$ et est proche de zéro entre celles-ci. Ce phénomène de résonance électronique n'est rien d'autre que le phénomène de réflexion de Bragg.

Avec réflexions secondaires :

$$R = \frac{r^2 e^{2ika}}{1 - t^2 e^{2ika}}$$

$R = 1$ si $k.a = m\pi$

Représentation schématique du dénominateur de la fonction de réflexion R dans le plan complexe permettant de visualiser la notion de résonance.



Ce résultat est remarquable puisqu'il indique que malgré l'hypothèse d'interaction faible avec les atomes, les électrons peuvent être fortement influencés par le réseau cristallin lorsque leur quantité de mouvement est proche de la condition de résonance. Comme nous le verrons ci-dessus c'est ce phénomène de résonance (ou de réflexion de Bragg) qui est à la base des propriétés de conduction des cristaux.

Une description détaillée des propriétés électroniques des S-C exige la résolution de l'équation de Schrödinger avec le potentiel électrostatique périodique correspondant aux atomes ionisés du cristal. Cette résolution n'est pas aisée et fait appel à des notions élaborées (notamment de la théorie des groupes) dépassant de loin le cadre du présent rappel. Néanmoins afin de bien comprendre l'origine physique des propriétés électroniques des S-C, nous donnons ci-dessous une méthode de résolution approchée extrêmement simplifiée (et non rigoureuse) permettant de rendre compte de ce qui se produit soit exactement à la résonance soit complètement en dehors de celle-ci mais pas dans une situation intermédiaire.

Résolution de l'équation de Schrödinger :

Considérons pour simplifier l'équation de Schrödinger stationnaire à une dimension spatiale. La résolution de cette équation fournit les fonctions d'onde représentant les états stationnaires des électrons dans le S-C. Elle permet donc en principe de connaître les propriétés électroniques des S-C.

L'équation de Schrödinger stationnaire est une équation aux valeurs propres aux dérivées partielles dont l'énergie E constitue la valeur propre.

$$\frac{-\hbar^2}{2m_e} \partial_{xx}^2 \phi + V(x)\phi = E\phi$$

Dans le cas du réseau cristallin uni-dimensionnel le potentiel a la forme d'une simple fonction périodique décomposable en série de Fourier (si l'origine $x=0$ est placée par convention sur un des atomes du cristal, la symétrie conduit à une série de Fourier « cosinus »).

$$V(x) = -V_1 \cos\left(\frac{2\pi}{a}x\right) + V_2 \cos\left(\frac{4\pi}{a}x\right) + \dots \quad \text{où } V_1 > 0$$

où V_n sont les coefficients du développement de Fourier ou « harmoniques » du potentiel électrostatique du cristal.

L'idée de la méthode de résolution simplifiée est d'imposer une forme à la fonction d'onde stationnaire solution de l'équation de Schrödinger. En s'inspirant du phénomène de réflexion de Bragg vu ci-dessus, on peut supposer que la fonction d'onde est la somme d'une onde progressive (de vecteur d'onde $+k$) et d'une onde réfléchie (de vecteur d'onde $-k$) de coefficient R . Le problème revient alors à calculer la valeur de R associée à une énergie E en substituant cette fonction dans l'équation.

Avec solution de la forme : $\phi = e^{ikx} + R e^{-ikx}$


$$\frac{\hbar^2 k^2}{2m_e} (e^{ikx} + R e^{-ikx}) - \frac{V_1}{2} \left(e^{i\frac{2\pi}{a}x} + e^{-i\frac{2\pi}{a}x} \right) (e^{ikx} + R e^{-ikx}) = E (e^{ikx} + R e^{-ikx})$$

Pour éliminer la variable spatiale x , on intègre cette relation sur x après avoir multiplié par $\exp(-ikx)$. Les intégrales des termes rapidement oscillants tels que $\exp(2ikx)$ sont bien entendu nulles. On voit alors que seul le terme oscillant en $2\pi/a-2k$ est susceptible de contribuer au terme de potentiel de l'équation:

Intégration :

$$\frac{1}{L} \int_{-L/2}^{L/2} \dots \times e^{-ikx} dx \quad \text{et} \quad \frac{1}{L} \int_{-L/2}^{L/2} e^{\pm 2ikx} dx = 0, \quad \frac{1}{L} \int_{-L/2}^{L/2} e^{\pm 2i\left(\frac{\pi}{a}+k\right)x} dx = 0$$

Les intégrales se font sur la longueur L tendant vers l'infini (procédure de moyennage).

L'équation de Schrödinger devient donc :



$$\frac{\hbar^2 k^2}{2m_e} - \frac{V_1}{2} R \frac{1}{L} \int_{-L/2}^{L/2} e^{2i\left(\frac{\pi}{a}-k\right)x} dx = E$$

Pour les valeurs de k éloignées de π/a , le dernier terme de potentiel restant est aussi nul, dans ce cas on voit que l'énergie est donnée pas la relation de dispersion classique rencontrée plus haut pour l'électron libre. Ce n'est que lorsque $k = \pi/a$ que le terme de potentiel a une contribution non nulle.

$$\text{Soit } k \neq \frac{\pi}{a} \quad \rightarrow \quad \frac{\hbar^2 k^2}{2m_e} = E \quad \text{(électron libre)}$$
$$\text{Soit } k = \frac{\pi}{a} \quad \rightarrow \quad \frac{\hbar^2 k^2}{2m_e} - \frac{V_1}{2} R = E$$

Cette dernière relation ne fournit encore ni la fonction d'onde (soit la valeur de R) ni la valeur propre associée E . Celles-ci peuvent être obtenue en effectuant la même procédure d'intégration appliquée cette fois après multiplication de l'équation de Schrödinger par $\exp(+ikx)$. On voit dans ce cas que seul le terme oscillant en $-2\pi/a+2k$ est susceptible de contribuer au terme de potentiel de l'équation:

Intégration :

$$\frac{1}{L} \int_{-L/2}^{L/2} \dots \times e^{ikx} dx \quad \text{et} \quad \frac{1}{L} \int_{-L/2}^{L/2} e^{\pm 2ikx} dx = 0, \quad \frac{1}{L} \int_{-L/2}^{L/2} e^{\pm 2i\left(\frac{\pi}{a}+k\right)x} dx = 0$$



L'équation de Schrödinger devient donc cette fois:

$$\frac{\hbar^2 k^2}{2m_e} R - \frac{V_1}{2} \frac{1}{L} \int_{-L/2}^{L/2} e^{-2i\left(\frac{\pi}{a}-k\right)x} dx = R E$$

La conclusion est dès lors la même : pour les valeurs de k éloignées de π/a , le dernier terme de potentiel restant est nul, dans ce cas on voit que l'énergie est donnée pas la relation de dispersion classique rencontrée plus haut pour l'électron libre. Ce n'est que lorsque $k = \pi/a$ que le terme de potentiel a une contribution non nulle.

$$\text{Soit } k \neq \frac{\pi}{a} \quad \rightarrow \quad \frac{\hbar^2 k^2}{2m_e} = E \quad \text{(électron libre)}$$

$$\text{Soit } k = \frac{\pi}{a} \quad \rightarrow \quad \frac{\hbar^2 k^2}{2m_e} R - \frac{V_1}{2} = E R$$

Notons que la condition $k = \pi/a$ constitue un cas particulier de la condition de réflexion de Bragg vue ci-dessus (il s'agit de la réflexion de Bragg d'ordre le plus bas). On retrouve donc bien le résultat de l'approche intuitive qui dit que, d'une part, à la résonance les électrons subissent une forte interaction avec le réseau cristallin et d'autre part, loin de la résonance les électrons ne voient pas le réseau (ne sont pas réfléchis) et se comportent comme des électrons libres.

Voyons maintenant ce que donne plus précisément la résolution de l'équation de Schrödinger à la résonance. Les deux procédures d'intégration donnent deux équations pour les deux inconnues du problème:



À la résonance : 2 équations à 2 inconnues

$$\left\{ \begin{array}{l} \frac{\hbar^2 k^2}{2m_e} - \frac{V_1}{2} R = E \\ \frac{\hbar^2 k^2}{2m_e} R - \frac{V_1}{2} = E R \end{array} \right. \quad \rightarrow \quad R^2 = 1$$

$$R = 1 \quad \rightarrow \quad E = \frac{\hbar^2 k^2}{2m_e} - \frac{V_1}{2}$$

$$R = -1 \quad \rightarrow \quad E = \frac{\hbar^2 k^2}{2m_e} + \frac{V_1}{2}$$

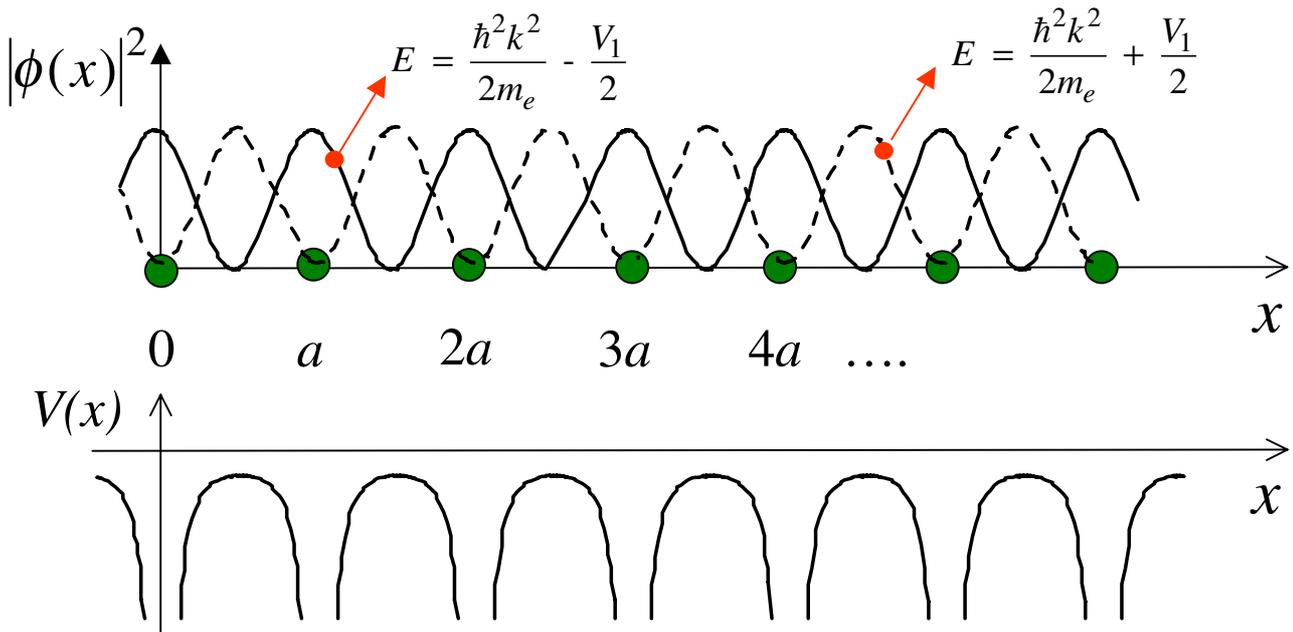
En accord avec l'approche intuitive, on voit que la réflexion est unitaire à la résonance. La résolution de l'équation de Schrödinger nous apporte néanmoins plus de précision: nous voyons que la résonance est caractérisée par deux fonctions d'onde d'énergies différentes correspondant à une réflexion sans déphasage ($R=1$) et une réflexion avec déphasage de π ($R=-1$). Leur interprétation physique est très simple: Lorsque $R=1$, les ondes incidente et réfléchie forment une onde stationnaire avec les maxima localisés sur les atomes. Cette situation correspond donc à une probabilité de présence maximale là où il est naturel de la trouver puisque les électrons sont attirés par les atomes ionisés. Il est donc logique que cette situation corresponde à la valeur la plus faible de l'énergie : $E = \hbar^2 k^2 / 2m_e - V_1 / 2$.

Dans l'autre cas ($R=-1$) l'onde stationnaire a ses maxima entre les atomes ionisés. Situation qui maximise l'énergie potentielle des électrons; il est donc logique de trouver l'énergie la plus élevée pour cette situation : $E = \hbar^2 k^2 / 2m_e + V_1 / 2$. Cette discussion est résumée et illustrée ci-dessous.

Interprétation :

si $k = \pi/a$

$$\phi = e^{ikx} + R e^{-ikx} \quad \left\{ \begin{array}{l} R = 1 \rightarrow |\phi|^2 \div \cos^2(kx) \text{ ———} \\ R = -1 \rightarrow |\phi|^2 \div \sin^2(kx) \text{ - - -} \end{array} \right.$$



En dehors de la résonance, on a vu dans l'approche intuitive que la réflexion est négligeable et que les électrons se comportent comme des électrons libres. C'est à dire que leur fonction d'onde associée est une onde progressive seule (sans réflexion). Une telle fonction d'onde représente une probabilité de présence uniforme. Le potentiel électrostatique moyen vu par les électrons est donc nul et il est logique de retrouver l'énergie de l'électron libre (notons que, par souci de simplicité, on n'a pas pris en compte la composante de fréquence spatiale nulle dans le développement de Fourier du potentiel, l'ajout d'un potentiel constant étant sans effets physiques. C'est pourquoi l'on retrouve, hors résonance, l'énergie de l'électron libre sans la présence d'une constante arbitraire inutile).

si $k \neq \pi/a$

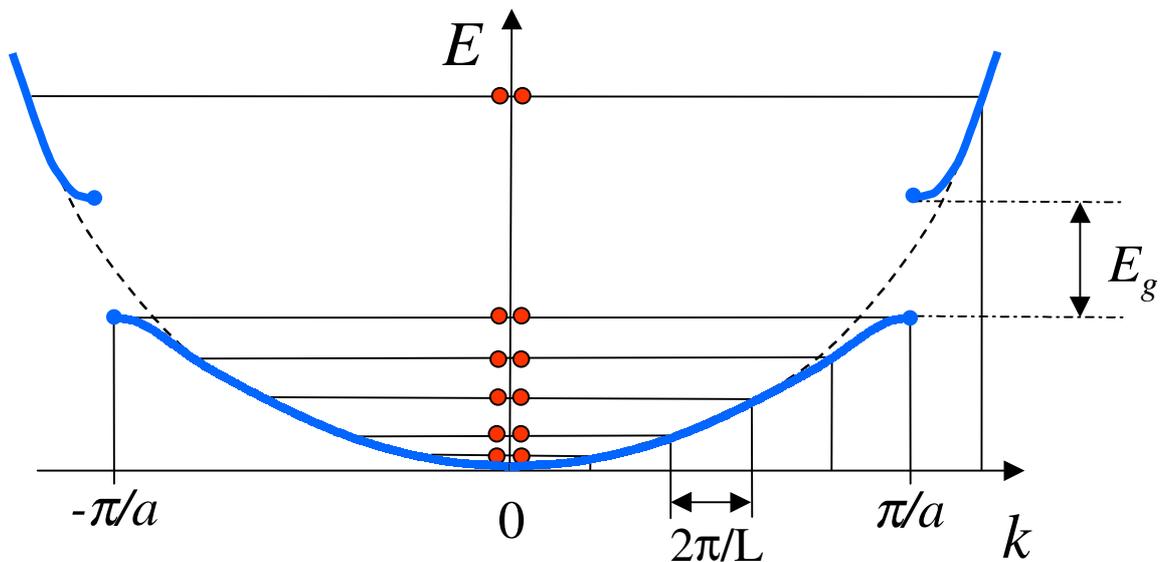
$$R \approx 0 \rightarrow \phi \approx e^{ikx}$$

$$|\phi|^2 = 1 \quad \Rightarrow \quad \text{potentiel moyen nul} \quad \Rightarrow \quad E = \frac{\hbar^2 k^2}{2m_e}$$

De cette discussion on peut conclure que pour des petites valeurs de k (soit $k < \pi/a$) les électrons se comportent comme des électrons libres et la courbe de dispersion est donc parabolique. Quand on se rapproche de la résonance on peut s'attendre à ce que d'abord la situation d'énergie la plus faible soit rencontrée, c'est à dire que la réflexion des électrons sur le réseau apparaît sans déphasage et la courbe de dispersion est modifiée de manière à trouver l'énergie $E = \hbar^2 k^2 / 2m_e - V_1 / 2$ à la résonance. En forçant encore k à augmenter on passe abruptement à la situation d'énergie élevée $E = \hbar^2 k^2 / 2m_e + V_1 / 2$ et au fur et à mesure que l'on s'éloigne de la résonance la courbe de dispersion se rapproche de celle de l'électron libre. La courbe de dispersion résultante est représentée schématiquement à la figure ci-dessous pour un cristal à 1D.

Conductivité des cristaux :

1 dimension



Cette courbe de dispersion fait apparaître la notion de bandes d'énergie permises et interdites qui est à l'origine des propriétés de conduction des S-C. Pour bien comprendre pourquoi l'existence de bandes interdites conditionne les propriétés de conduction, il suffit de considérer le nombre d'électrons que peut contenir une bande permise. A cette fin nous utilisons une fois de plus le comptage des états dans l'espace réciproque k . Le nombre d'états possibles occupés par les électrons est donné par la largeur de la bande permise $2\pi/a$ divisée la largeur de la cellule élémentaire $2\pi/L$, le tout multiplié par 2 pour le nombre de spins.

Capacité d'une bande

$$\begin{aligned}
 & \text{spin} \quad \nearrow \quad \nearrow \quad \text{largeur de la bande} \\
 - \text{ nombre d'états} &= 2 \frac{\frac{2\pi}{a}}{\frac{2\pi}{L}} = 2 \frac{L}{a} = 2 \cdot \text{nbre d'atomes} \\
 & \quad \quad \quad \searrow \quad \text{largeur d'une cellule}
 \end{aligned}$$

Or, le nombre d'électrons disponibles dans le cristal est donné par la valence (le nombre d'électrons de la dernière couche atomique) multipliée par le nombre d'atomes du cristal. Pour le cristal 1D ce nombre est L/a . On voit donc qu'un cristal d'atomes monovalents tel que le sodium aura sa bande permise à moitié vide, tandis que celle-ci sera remplie pour un cristal d'atomes de valence 2 tel que le magnésium.

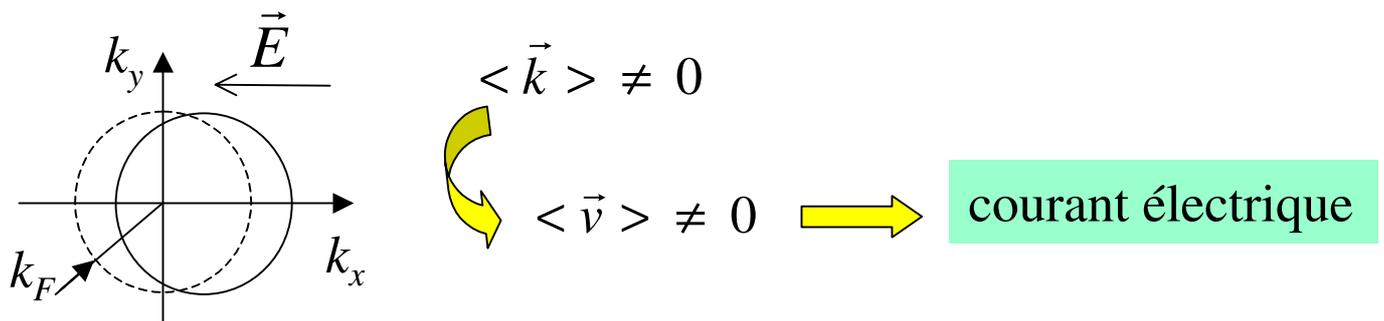
- nombre d'électrons dans le cristal : **valence × nbre d'atomes**

Na : 1 électron $s \implies$ nbre d' $e^- = L/a \implies$ bande à moitié vide

Mg : 2 électron $s \implies$ nbre d' $e^- = 2L/a \implies$ bande remplie

En l'absence de champ électrique extérieur, les électrons empilés dans la bande permise du cristal ont des vitesses qui peuvent être différentes de zéro (car k est proportionnel à la quantité de mouvement $m_e v$ et les électrons ont des valeurs de k qui s'étalent sur la zone de Brillouin) mais la moyenne de la vitesse sur tous les électrons est nulle car la bande permise est symétrique par rapport à l'origine $k=0$.

Si un champ électrique appliqué de l'extérieur (par une différence de potentiel aux extrémités d'un morceau de cristal), les électrons subiront une force qui les fera bouger dans le sens opposé au champ. Ce mouvement, qui engendrera un courant électrique, se fait à une certaine vitesse moyenne $\langle v \rangle$ (constante, en raison des collisions avec le réseau cristallin qui donne lieu à une friction effective) correspondant à une valeur moyenne de k soit $\langle k \rangle$ non nulle. Si la bande permise n'est pas remplie rien n'empêche d'avoir une valeur moyenne de k non nulle car rien n'empêche les électrons de s'organiser de manière à créer une distribution asymétrique dans l'espace réciproque. Par contre si la bande permise est remplie, toutes les valeurs de k sont occupées et il est impossible d'asymétriser la distribution dans l'espace réciproque. On voit donc ainsi qu'un courant électrique ne peut pas s'établir dans un cristal à atome de valence 2. Ce raisonnement peut bien entendu être généralisé à des valences supérieures à 2. Il y aura alors remplissage complet ou partiel des bandes permises supérieures et les conclusions seront les mêmes.



Ceci n'est applicable que si l'on ne tient pas compte de l'agitation thermique. En effet, s'il y a une certaine agitation thermique, comme c'est toujours le cas en pratique, les électrons pourront recevoir une énergie suffisante que pour pouvoir passer dans la bande permise supérieure (il faut pour cela que l'énergie thermique ne soit pas négligeable devant l'énergie de la bande interdite appelée GAP). Dans ce cas la distribution des états électroniques peut devenir asymétrique et un courant électrique peut s'établir.

Avant de considérer ce point de vue, il est intéressant de calculer le vecteur d'onde de Fermi pour un cristal de valence et de constante de réseau données. Celui-ci peut être aisément calculé en remarquant que, par définition du niveau de Fermi, les états électroniques occupent la bande permise de $-k_F$ à $+k_F$, comme exprimé dans la formule ci-dessous :

Agitation thermique et distribution de Fermi-Dirac

Vecteur d'onde de Fermi

1 dimension

$$\text{nbre d'états occupés} = \text{valence} \frac{L}{a} \equiv 2 \frac{2k_F}{\frac{2\pi}{L}}$$

$$k_F = \frac{\text{valence}}{2} \times \frac{\pi}{a} \quad \longrightarrow \quad \text{valence paire : } k_F \text{ est au gap}$$

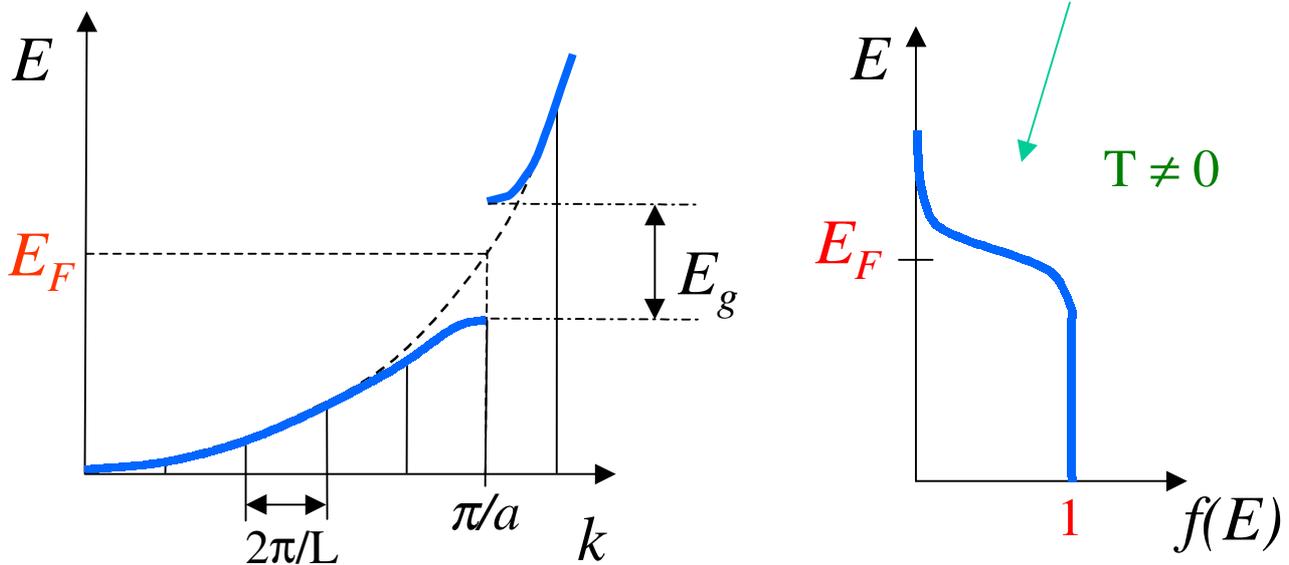
On constate donc que le vecteur d'onde de Fermi se trouve à la limite de la zone de Brillouin pour une valence paire. Dans ce cas, l'énergie de Fermi se trouve donc dans le gap.

En effet, par définition E_F est donnée par :

$$E_F = \frac{\hbar^2 k_F^2}{2m_e}$$

$$f(E) = \frac{1}{\exp[(E - E_F)/k_B T] + 1}$$

La distribution de Fermi $f(E)$ indique que les électrons des couches supérieures peuvent effectivement avoir une énergie suffisante que pour franchir le gap et atteindre la bande permise supérieure. Si l'énergie thermique n'est pas négligeable devant le gap on pourra avoir conduction. Or, l'énergie thermique à température ambiante (300 K) vaut $2.6 \cdot 10^{-2} \text{ eV}$.



si $E_g \approx k_B T \longrightarrow$ semi-conducteur

$$k_B = 1.38 \cdot 10^{-23} \text{ J/K}, T = 300 \text{ K} \rightarrow k_B T = 4.14 \cdot 10^{-21} \text{ J} = 2.6 \cdot 10^{-2} \text{ eV}$$

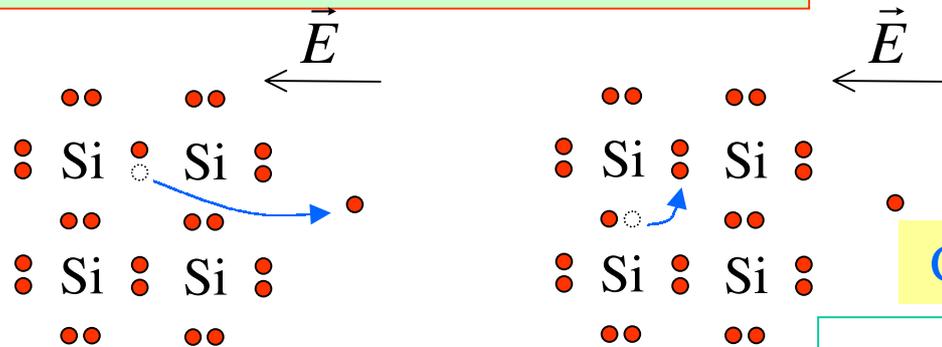
Si le gap est de l'ordre de 1 eV , la conduction sera faible mais non négligeable. C'est, par définition, le cas des cristaux semi-conducteurs. Ci-dessous nous considérons, après les éléments des deux premières colonnes du tableau de Mendeleev Na et Mg, des éléments des colonnes suivantes. L'Aluminium a une valence 3; il y a donc une bande et demie remplie et il y a donc conduction (possibilité d'avoir une distribution asymétrique des k). Le Silicium a une valence 4 correspondant à 2 bandes remplies. Le gap séparant la bande occupée supérieure et la bande permise suivante est de 1.13 eV . Il s'agit donc d'un semi-conducteur. Le Carbone a aussi une valence 4 mais son gap est de 5.5 eV et est considéré comme isolant (le diamant).

{	Al $\rightarrow 3e^-$	conducteur
	Si $\rightarrow 4e^- : E_g = 1.13 \text{ eV}$	semi-conducteur
	C $\rightarrow 4e^- : E_g = 5.5 \text{ eV}$	isolant

Nous terminerons ce rappel en passant brièvement en revue différents concepts de physique des S-C utiles à l'étude de leurs propriétés optiques.

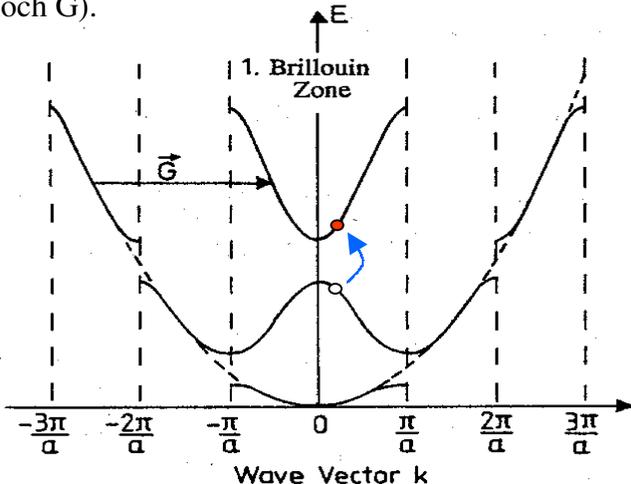
Voyons tout d'abord le concept de bandes de conduction et de valence. Nous l'illustrons ci-dessous avec l'exemple du Silicium. Les quatre électrons de valence de tous les atomes remplissent deux bandes permises et occupent donc des états donnés qui ne peuvent être quittés que grâce à l'énergie thermique. Les électrons qui reçoivent l'énergie suffisante pour surmonter le gap passe dans la bande permise supérieure appelée bande de conduction car ce sont les électrons libres de cette bande qui contribuent à la conduction. Par opposition, la bande permise des états liés est appelée bande de valence. Cette dénomination est trompeuse car, en réalité, l'asymétrisation de la distribution des électrons dans l'espace réciproque (responsable de la conduction) concerne également les électrons de la bande de valence. En effet, si d'un côté, les électrons sont montés sur la bande supérieure, de l'autre côté les électrons ont libéré des états électroniques qui se trouvent donc vacants. Ceci peut être visualisé dans l'espace réel grâce au schéma ci-dessous où l'on voit la place vacante laissée par l'électron libéré. Cette place vacante pourra être occupée par un autre électron de valence qui lui-même créera une place vacante à un autre endroit. Les électrons de valence peuvent donc également se déplacer grâce à ces places vacantes et ainsi participer à la conduction. Le mouvement des électrons de valence crée un mouvement apparent des places vacantes. Pour des raisons de simplicité de modélisation des propriétés des S-C ils s'avère plus simple de décrire le mouvement des place vacantes que l'on considère comme des particules à part entière que l'on appelle les « trous ». La justification de cette manière de procéder est illustrée par l'exemple de la mesure de la quantité de mouvement moyenne des électrons de la bande de valence dans laquelle un trou est présent. La quantité de mouvement moyenne est donnée par la quantité de mouvement de la bande pleine (quantité nulle puisque la distribution des e^- est alors symétrique) moins la quantité de mouvement de l'électron absent qui a formé le trou, soit un total de $-k_e$. Cette valeur est la quantité de mouvement du trou (plus simple à exprimée que la somme des k sur les $n-1$ états occupés). Ce raisonnement peut être facilement appliqué à l'énergie dont le signe sera également inversé pour les trous. Un trou aura toujours tendance à monter dans la bande de conduction pour minimiser l'énergie du système. Il en résulte que la distribution de Fermi pour les trous est inversée par rapport à celle des électrons.

Bandes de conduction et de valence



Conduction des trous

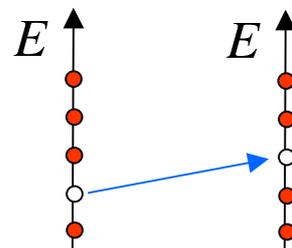
La figure ci-dessous indique simplement que d'après la théorie rigoureuse de la physique des cristaux, l'on peut décrire les bandes en les ramenant toutes dans la première zone de Brillouin (cf. Invariance des fonctions d'onde électroniques par rapport à une translation d'un vecteur de Bloch G).



$$\langle \vec{k} \rangle = \sum_{n-1 \text{ états}} \vec{k}_n$$

$$\langle \vec{k} \rangle = \sum_{n \text{ états}} \vec{k}_n - \vec{k}_e = -\vec{k}_e$$

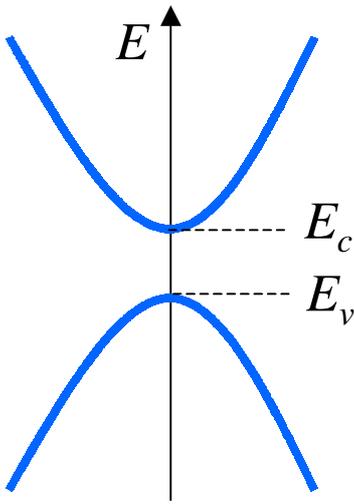
0



$$\Delta E_t = -\Delta E_e$$

Au voisinage du centre de la zone de Brillouin les courbes de dispersion pour les électrons de conduction apparaissent donc paraboliques comme pour l'électron libre vu plus haut. La seule différence avec l'électron libre est que la courbure de la courbe de dispersion n'est pas nécessairement donnée par la masse de l'électron et la constante de Planck. La courbure peut en fait prendre n'importe quelle valeur dépendant de la manière dont se forme le gap. Comme les propriétés (quantiques) des électrons sont données à partir de l'application du principe de transposition (postulat de la mécanique quantique) à la relation de dispersion (voir plus haut pour l'électron libre), on peut conclure qu'un semi-conducteur se comporte comme un nuage d'électrons libres dont la masse est fixée par la courbure de la courbe de dispersion. On introduit ainsi la masse effective des électrons comme indiqué ci-dessous. Le même raisonnement est appliqué pour les trous responsables de la conduction de la bande de valence.

Masse effective



$$E = E_c + \frac{1}{2} k^2 \left. \frac{\partial^2 E}{\partial k^2} \right|_{k=0}$$

$$\frac{1}{2} k^2 \left. \frac{\partial^2 E}{\partial k^2} \right|_{k=0} = \frac{\hbar^2}{m_c}$$

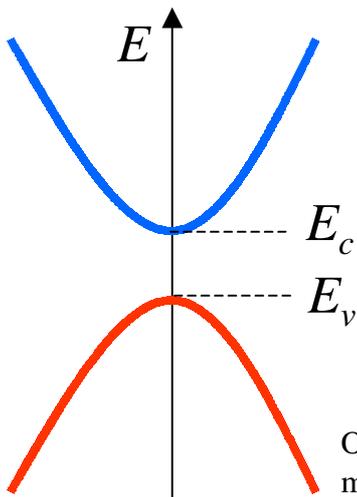
$$E = \frac{\hbar^2 k^2}{2m_c} \quad (\text{avec } \hbar k = p \text{ on a } E = \frac{p^2}{2m_c})$$

$$(\text{idem pour trous : } -\frac{1}{2} k^2 \left. \frac{\partial^2 E}{\partial k^2} \right|_{k=0} = \frac{\hbar^2}{m_v})$$

S-C = gaz d'e⁻ et trous libres de masses m_c et m_v

A partir de ce principe, les quantités caractérisant les électrons (et les trous) dans les S-C peuvent être calculées aisément en appliquant simplement les calculs faits pour les électrons libres. Ainsi, la densité d'électrons ou de trous (ou plus simplement la densité de porteurs) se calcule à partir du compte du nombre d'états électroniques dans l'espace réciproque (compte du nombre d'états dans une couche sphérique d'épaisseur dk , etc...). On arrive ainsi à la forme parabolique de la densité d'état par unité de volume et d'énergie. L'unique différence avec l'électron libre est que l'origine des énergie est maintenant le bas de la bande de conduction, soit E_c .

Densité d'états



$$E = E_c + \frac{\hbar^2 k^2}{2m_c}$$

$$dk = \frac{1}{2} \sqrt{\frac{2m_c}{\hbar^2 (E - E_c)}} dE, \quad dN_e = \frac{4\pi k^2 dk}{8\pi^3 / L^3}$$

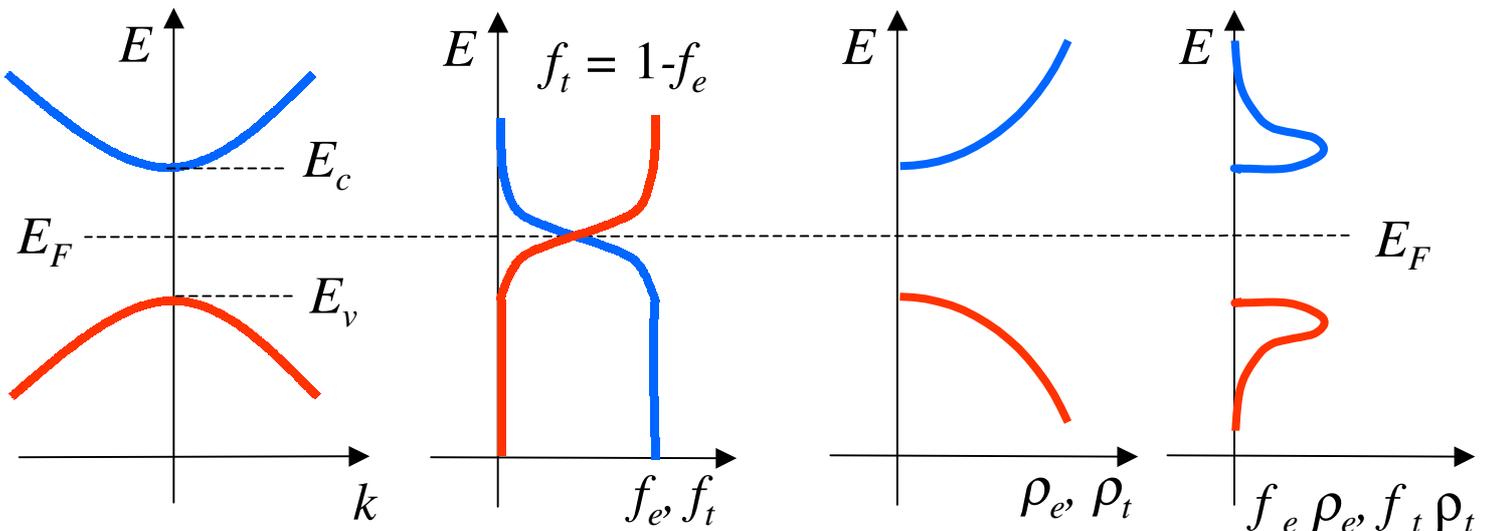
$$\frac{dn_e}{dE} = \rho_e = \frac{1}{2\pi^2} \left(\frac{2m_c}{\hbar^2} \right)^{3/2} \sqrt{E - E_c}$$

On applique le même calcul aux trous :

$$\left. \text{Trous : } \rho_t = \frac{1}{2\pi^2} \left(\frac{2m_v}{\hbar^2} \right)^{3/2} \sqrt{E_v - E} \right\}$$

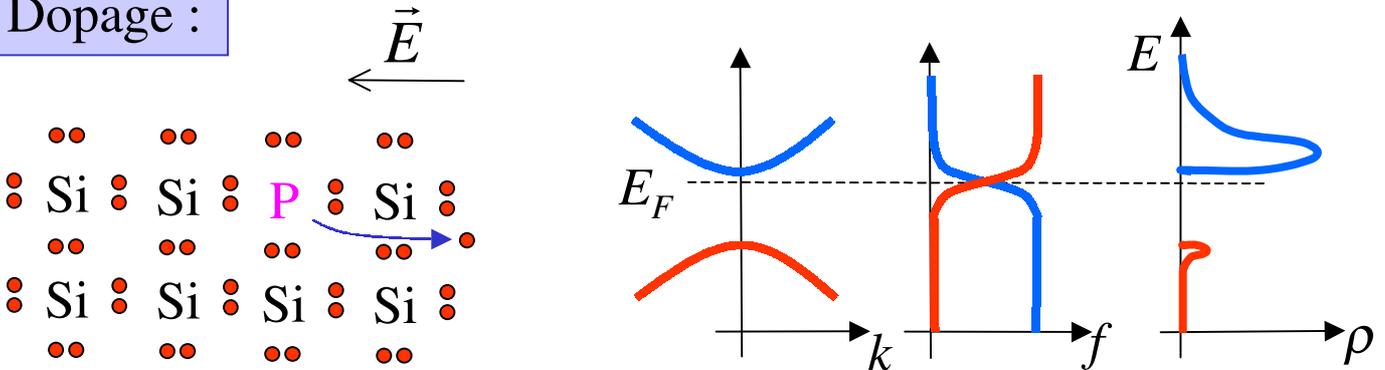
En guise de résumé, nous pouvons donc dire qu'un semi-conducteur se comporte comme un nuage de porteurs de charges libres (négative et positives) de masses fixées par la courbure des courbes de dispersion des bandes de conduction et de valence. Le nombre d'électrons et de trous dans ces bandes est fixé par la distribution de Fermi (inversée pour les trous); le niveau de Fermi se situant au milieu du gap. Au zéro absolu, il n'y a donc pas de porteurs libres dans un S-C. La densité d'états par unité de volume et d'énergie est parabolique (comme pour l'électron libre). Les produits $\rho_e f_e$ et $\rho_t f_t$ sont les densités d'électrons et trous par unité de volume et d'énergie. La densité de porteurs libres par unité de volume est dès lors donnée par l'intégrale sur E de la densité d'états ρ et de la distribution de Fermi f .

Résumé :



Si l'on dope le S-C avec des atomes de valence supérieure (ou inférieure) le nombre d'électrons (ou de trous) dans la bande de conduction (ou de valence) peut augmenter considérablement et peut dès lors fortement augmenter la conduction. Ci-dessous est illustré le cas du dopage de type n par insertion d'atome de Phosphore (valence 5) dans une matrice cristalline de Silicium. Il en résulte un plus grand nombre d'électrons dans la bande de conduction, c'est à dire, un niveau de Fermi plus élevé et donc une plus grande densité d'électrons libres par unité de volume.

Dopage :



Ceci termine le rappel des notions de base de physique des Semi-conducteurs nécessaires à la bonne compréhension de leurs propriétés optiques. Dans le chapitre suivant nous passons en revue brièvement les notions de physique qui sont à la base des interactions lumière-matière.

2. INTERACTIONS LUMIERE-MATIERE

Avant de considérer les interactions lumière-matière proprement dites, nous faisons ci-dessous un bref rappel des formalismes utilisés pour décrire la lumière. D'un point de vue classique (c-à-d, non quantique) la lumière est constituée d'ondes électromagnétiques et est donc décrite par les équations de Maxwell combinées aux relations constitutives caractérisant le milieu dans lequel elle se propage (elle sont données ici pour le vide).

A. Lumière = onde électromagnétique \longrightarrow Equations de Maxwell

$$\begin{aligned}\text{rot } E &= -\mu_o \frac{\partial H}{\partial t} \\ \text{rot } H &= \epsilon_o \frac{\partial E}{\partial t}\end{aligned}$$

Dans le vide :

$$\begin{aligned}\text{div } E &= 0 \\ \text{div } H &= 0\end{aligned}$$

La combinaison des deux équations de Maxwell fournit l'équation d'onde dont la solution la plus simple est une onde périodique progressive (onde de propagation) caractérisée par un vecteur d'onde lié à la fréquence par la vitesse de la lumière c . Comme l'équation d'onde est une équation linéaire, toutes ses solutions peuvent être décomposées en somme de telles ondes périodiques (ou modes électromagnétiques). Les ondes sont décrites ici en notation complexe symbolique (il faut ajouter le complexe conjugué pour décrire un signal physique réel).



$$\text{rot rot } E = -\mu_o \epsilon_o \frac{\partial^2 E}{\partial t^2} \longrightarrow \text{grad div } E - \Delta E = -\frac{1}{c^2} \frac{\partial^2 E}{\partial t^2}$$

1D

$$\frac{\partial^2 E}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 E}{\partial t^2} = 0$$

$$\text{solution : } E = A e^{i(kx - \omega t)} + c.c.$$

$$(\text{où } k = \omega / c)$$

L'énergie transportée par l'onde E est donnée par le vecteur de Poynting $S = E \times H$ et est donc simplement proportionnelle au carré de l'amplitude du champ. Comme l'amplitude A peut avoir n'importe quelle valeur, l'énergie de l'onde électromagnétique (E-M) de fréquence donnée ω , est une quantité qui peut varier de manière continue.

$$\text{Energie transportée : } \mathbf{E} \div |A|^2 \quad (A : \text{variable continue})$$

Comme nous le verrons dans la suite, une variation continue de l'énergie d'une onde E-M est incompatible avec la réalité expérimentale. L'expérience a en effet montré que la lumière doit être décrite en terme de quanta d'énergie appelés photons. Ce sont les expériences de Planck sur le rayonnement des corps noirs (ou rayonnement thermique) ainsi que la description par Einstein de l'effet photo-électrique qui ont mené à cette conclusion. Ces deux types d'expériences ont mené à la conclusion que l'énergie de la lumière à une fréquence ω donnée ne peut varier que par quanta $h\nu$ où h est la constante de Planck.

B. Lumière = photons

$$\text{Planck : } \Delta E = h\nu = \hbar\omega$$

Cette observation a mené à la notion de seconde quantification qui désigne la quantification du champ électromagnétique. Cette notion dépasse de loin le cadre de ce cours, mais nous pouvons en faire une description simplifiée qui en donnera une vue intuitive. L'idée est de résoudre l'équation d'onde par la méthode de séparation des variables, soit $E(x,t) = f(x) q(t)$. Puisque l'équation d'onde est linéaire, on peut choisir pour $f(x)$ un mode périodique unique d'amplitude unitaire soit $f(x) = \exp(ikx)$. L'amplitude du champ E-M sera donnée par la fonction $q(t)$. Avec cette forme de $f(x)$ l'équation en q prend la forme de l'équation de l'oscillateur harmonique avec $\omega^2 = k^2 c^2$ comme constante de rappel. L'idée est alors d'établir pour cette équation une analogie formelle avec la mécanique et en particulier avec la mécanique quantique. On associe $q(t)$ à la coordonnée d'un point massique (de masse unitaire) soumis à une force proportionnelle à la coordonnée elle-même. En d'autres termes, on a affaire à un point massique plongé dans un potentiel harmonique (parabolique) $V = \omega^2 q^2/2$. La trajectoire classique du point dans ce potentiel est une simple sinusoïde dont l'amplitude peut prendre toute les valeurs (variable continue). Cette sinusoïde représente les variations périodiques temporelles de l'onde E-M telle qu'on la visualise en électromagnétisme classique.

Quantification du champ E-M

$$E(x,t) = e^{ikx} \cdot q(t)$$



$$\frac{d^2 q}{dt^2} = -\omega^2 q$$

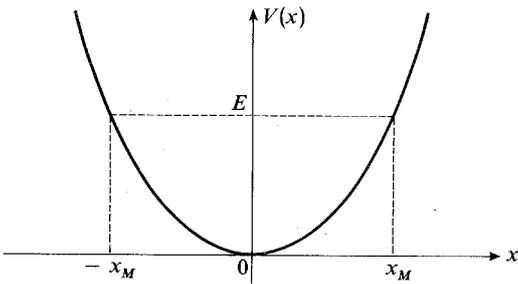
$$(\omega = kc)$$

Analogie à la mécanique :
l'oscillateur harmonique

$$F = m.a = -\frac{\partial V}{\partial q}, \quad V = \frac{1}{2} \omega^2 q^2$$

L'énergie totale de l'oscillateur est donnée par l'Hamiltonien du système donné ci-dessous. Tout comme en mécanique quantique, on applique à l'oscillateur le principe de transposition en associant à la quantité de mouvement $p = dq(t)/dt$ l'opérateur de dérivée première par rapport à la position q . Tout comme en mécanique l'application du principe de transposition à l'Hamiltonien conduit à l'équation de Schrödinger stationnaire. Comme les fonctions d'ondes et les niveaux d'énergie de l'oscillateur harmonique sont bien connus on peut directement en tirer des conclusions : l'énergie de l'oscillateur est quantifiée et est donnée par $E_n = \hbar\omega(n+1/2)$. Chacun de ces niveaux d'énergie correspond à une fonction d'onde $\phi(q)$ dont le module carré donne la probabilité de présence du point massique en q . Comme q représente ici l'amplitude du champ E-M, on constate que dans le cadre de la quantification du champ l'amplitude d'une onde E-M n'est plus déterminée de manière univoque (la sinusoïde évoquée ci-dessus) et que l'on ne peut qu'en définir la valeur de manière probabiliste. La quantification du champ consiste à adopter ce point de vue comme une réalité physique. On associe l'énergie $E_n = \hbar\omega(n+1/2)$ à l'énergie réelle du champ E-M et l'on associe à chaque quanta $\hbar\omega$ un photon. Le nombre quantique n donne donc le nombre de photon du champ E-M. On voit que, en l'absence de photon soit $n = 0$, l'énergie du champ E-M est non-nulle (en mécanique quantique l'oscillateur n'est jamais au repos). Cette énergie résiduelle représente ce que l'on appelle les fluctuations du vide. Comme on le verra par la suite, c'est une quantité qui pose problème dans certaines situations et il faut parfois la supprimer de manière artificielle (ceci est lié à ce que l'on appelle les problèmes de renormalisation en électrodynamique quantique).

Quantification :



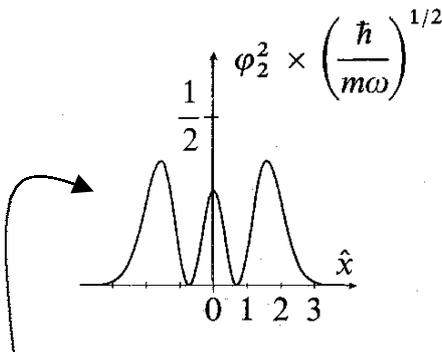
$$H = \frac{1}{2} \dot{q}^2 + \frac{1}{2} \omega^2 q^2 \quad \dot{q} = p \equiv -i\hbar \partial_q$$

$$E\phi = -\frac{\hbar^2}{2} \partial_{qq}^2 \phi + \frac{1}{2} \omega^2 q^2 \phi$$

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right)$$

n = nombre de photons à la fréquence ω

$\left(\frac{1}{2} \hbar\omega : \text{fluctuations du vide} \right)$



Fonction d'onde de l'oscillateur donnant la probabilité de trouver l'amplitude à une valeur donnée q (notée ici x)

Bien qu'il y ait maintenant une incertitude sur la valeur de l'amplitude de champ E-M à un instant donné, on peut définir l'amplitude du champ E-M comme on le fait d'habitude en mécanique quantique, c'est à dire, en prenant la valeur moyenne de la position pour un état donné (soit, un nombre n de photons donnés). La moyenne est donnée par l'intégrale sur q de la probabilité de présence multipliée par q . Le calcul appliqué avec les fonctions d'onde de l'oscillateur harmonique montre que l'amplitude moyenne A_n est proportionnelle à la racine carrée de l'énergie E_n . On retrouve ainsi la relation classique entre l'amplitude du champ E-M et son énergie (c-à-d, énergie proportionnelle au carré de l'amplitude, cf. vecteur de Poynting)

Incertainitude sur

$$\text{l'amplitude du champ : } \langle \phi_n | q | \phi_n^* \rangle = A_n \quad \text{et} \quad E_n \div A_n^2$$

Ces notions élémentaires d'électrodynamique quantique nous permettent d'étudier les mécanismes de rayonnement lumineux (c-à-d, d'émission de lumière par la matière). Le rayonnement thermique est la forme de rayonnement E-M qui fait intervenir le mécanisme d'interaction lumière-matière le plus simple que l'on puisse trouver. Il s'agit simplement du rayonnement émis par un corps matériel lorsque celui-ci est chauffé (c'est à dire, à une température supérieure au zéro absolu). Nous commencerons donc par l'étude de ce rayonnement en guise d'introduction à l'étude de l'émission de lumière par les S-C.

Le rayonnement thermique

L'étude du rayonnement thermique est basée sur une loi universelle de la thermodynamique des systèmes à grand nombre de particules. Il s'agit de la loi de Boltzmann. Cette loi donne la probabilité de trouver une particule (molécule d'un gaz, par exemple) à une énergie donnée. Il s'agit ici de l'appliquer aux photons. Pour cela on considère (avec un grand effort d'abstraction) que l'espace vide est rempli d'oscillateurs harmoniques représentant les oscillations du champ E-M et que ces oscillateurs répondent à la loi de Boltzmann. Considérons tout d'abord le champ E-M lumineux à une fréquence donnée ω . Nous généraliserons par la suite les résultats à un continuum de fréquence. Nous appliquons la loi de Boltzmann aux oscillateurs quantifiés dont l'énergie est donnée par E_n comme décrit ci-dessus. La loi de Boltzmann dit que si les oscillateurs E-M sont en équilibre thermique la probabilité de les trouver à l'énergie E_n est donnée par l'expression suivante.

Loi de Boltzmann : probabilité $p(E)$ d'avoir un oscillateur à l'énergie E

$$p_n = \frac{e^{-E_n/k_B T}}{\sum_{n=0}^{\infty} e^{-E_n/k_B T}}$$

où $E_n = \hbar\omega(n + \frac{1}{2})$

T est la température du système à l'équilibre thermique et k_B est la constante de Boltzmann (notons que l'équilibre thermique du champ E-M dans le vide ne peut être considéré rigoureusement que dans un enceinte matérielle vide dont les parois sont chauffées à la température T).

Où la somme au dénominateur est prise sur tous les états d'énergies possibles (infinité dénombrable).

L'énergie moyenne du champ E-M à la fréquence ω sera donc donnée par l'expression suivante:

Energie moyenne contenue dans le mode ω : $\langle E \rangle = \sum_n E_n p_n$

$$\langle E \rangle = \frac{\sum_n \hbar\omega(n + \frac{1}{2}) e^{-\hbar\omega(n + \frac{1}{2})/k_B T}}{\sum_n e^{-\hbar\omega(n + \frac{1}{2})/k_B T}} = \hbar\omega \frac{\sum_n n e^{-\hbar\omega n/k_B T}}{\sum_n e^{-\hbar\omega n/k_B T}} + \frac{1}{2} \hbar\omega$$

Pour simplifier les notations nous introduisons ci-dessous la quantité x (qui n'est rien d'autre que le facteur de Boltzmann).

$$x \equiv e^{-\hbar\omega/k_B T} \Rightarrow \langle E \rangle = \hbar\omega \frac{\sum_n n x^n}{\sum_n x^n} + \frac{1}{2} \hbar\omega = \hbar\omega x \frac{\sum_n n x^{n-1}}{\sum_n x^n} + \frac{1}{2} \hbar\omega$$

Le dénominateur est une somme infinie qui converge si x est plus petit que 1 ce qui est bien entendu toujours le cas. Remarquons également que la somme au numérateur est la dérivée de la somme au dénominateur, soit $1/(1-x)^2$.

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}, \quad \text{si } |x| < 1$$

Nous obtenons donc l'expression suivante pour l'énergie moyenne du champ E-M à la température T et à la fréquence ω .

$$\langle E \rangle = \hbar\omega x \frac{1}{(1-x)} + \frac{1}{2}\hbar\omega = \hbar\omega \frac{1}{1/x-1} + \frac{1}{2}\hbar\omega = \frac{\hbar\omega}{e^{\hbar\omega/k_B T} - 1} + \frac{1}{2}\hbar\omega$$

On trouve donc pour le nombre de photons moyens :

Nombre de photons moyen à la fréquence ω :

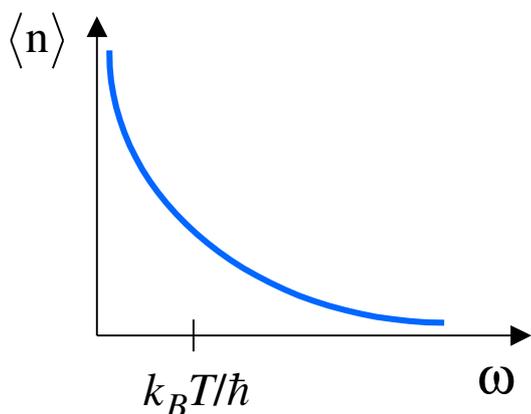
$$\langle n \rangle = \frac{1}{e^{\hbar\omega/k_B T} - 1}$$

On peut alors définir le nombre moyen de photons correspondant à cette énergie en y substituant la forme de l'énergie de l'oscillateur harmonique :

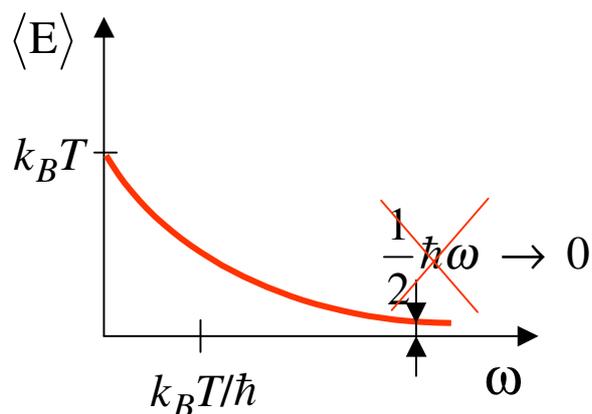


$$\langle E \rangle \equiv \hbar\omega \left(\langle n \rangle + \frac{1}{2} \right)$$

Etudions maintenant le nombre moyen de photons en fonction de leur fréquence. Nous voyons sur le graphe ci-dessous que n diverge en $\omega = 0$ et suit une diminution exponentielle déterminée par la température (dans la limite abstraite du zéro absolu, seul les photons de fréquence nulle existent). C'est la distribution de Bose-Einstein décrivant la statistique des photons (et autres bosons). Contrairement aux électrons (fermions) les photons peuvent s'accumuler indéfiniment à basse énergie. L'énergie du champ E-M à l'équilibre thermique, quant à elle, suit une exponentielle négative à partir de la valeur $k_B T$ jusqu'à la valeur asymptotique de $\hbar\omega/2$. Cette asymptote horizontale pose problème car elle rend infinie l'intégrale de l'énergie par rapport à la fréquence, ce qui indiquerait que le champ E-M à l'équilibre thermique contient une énergie infinie. Pour éviter cette absurdité physique, il faut supprimer artificiellement le terme $\hbar\omega/2$ dans l'expression de l'énergie moyenne du champ E-M (nous reverrons ce problème plus loin).



distribution de Bose-Einstein

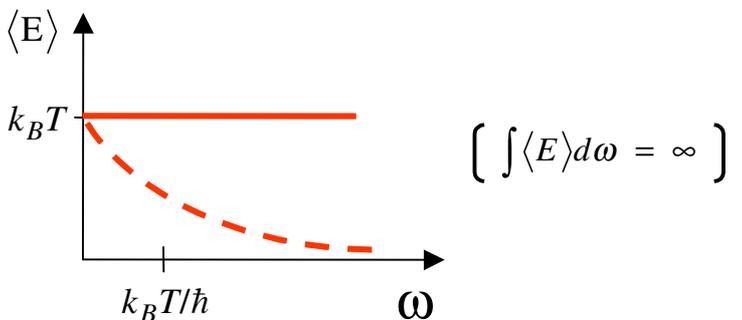


$$\left[\langle E \rangle = \hbar\omega \langle n \rangle \rightarrow \int \langle E \rangle d\omega < \infty \right]$$

Il est intéressant de comparer ce résultat à celui fourni par l'approche classique suivie par Rayleigh et Jeans à la fin du 19^{ème} siècle pour décrire la thermodynamique du champ E-M. Sur base des équations de Maxwell décrivant les ondes E-M, Rayleigh et Jeans considéraient que l'espace était rempli d'oscillateurs classiques (oscillations périodiques du champ E-M solutions des équations de Maxwell). A ces oscillateurs classiques ils appliquaient les lois de la thermodynamique et en particulier la loi de Boltzmann afin de trouver la manière dont les photons se distribuent en énergie (fréquence) à l'équilibre thermique et de comparer le résultat théorique aux mesures spectrales du champ E-M émis par un corps noir (enceinte matérielle vide dont les parois sont chauffées à la température T). Comme les oscillateurs sont classiques, leur énergie n'est pas quantifiée et le calcul de la valeur moyenne de l'énergie du champ E-M s'effectue par une intégrale de l'énergie avec le facteur de Boltzmann comme pondération (probabilité de trouver l'oscillateur à la fréquence E). L'intégrale est facile à calculer et donne pour résultat que l'énergie moyenne du champ E-M à une fréquence donnée (énergie donnée) est une constante indépendante de cette fréquence. On est donc confronté dans ce cas à une énergie intégrée infinie.

Calcul classique (Rayleigh-Jeans)

Ensemble d'oscillateurs d'énergie E à l'équilibre thermique

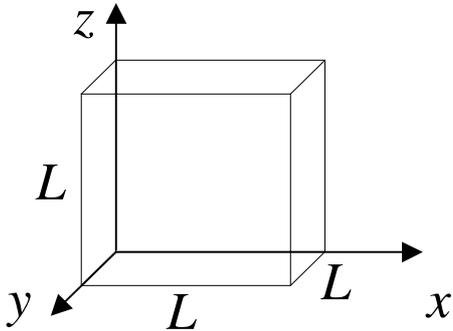
$$\langle E \rangle = \frac{\int E e^{-E/k_B T} dE}{\int e^{-E/k_B T} dE} = k_B T$$


Voyons maintenant ce problème plus en détail en considérant le concept de mesure spectrale (ou plus simplement le concept de spectre E-M). Lorsque l'on fait la mesure d'un spectre lumineux, on fait passer la lumière dans un élément dispersif qui sépare spatialement les différentes composantes fréquentielles du champ E-M qui la constitue. Plus le nombre de photons sera grand à une fréquence donnée, plus l'intensité mesurée sera importante à cette fréquence. Mais ce n'est pas seulement le nombre de photons qui détermine l'intensité mesurée à une fréquence donnée. En effet, la mesure ne peut pas se faire à une fréquence unique mais dans une fenêtre spectrale de dimension finie. Ceci signifie que la mesure s'effectue d'office sur plusieurs modes du champ E-M (mode = oscillation de fréquence ω donnée). Or, les modes du champ E-M sont organisés d'une manière analogue aux états électronique dans un cristal ce qui signifie que leur densité ne sera pas nécessairement constante mais dépendra de la fréquence même des modes (cf. densité d'énergie $\rho_e(E)$ pour les états électroniques). Ceci a pour effet que, selon la fréquence, un plus ou moins grand nombre de modes sera pris dans la fenêtre spectrale de l'appareil de mesure (spectromètre). Si ce nombre de modes est grand (grande densité), il y aura une plus grande intensité mesurée à nombres de photons égaux dans chacun des modes ω . Pour décrire le spectre E-M (ou toute mesure spectrale) il faut donc considérer le produit du nombre de photons dans chaque mode ω par la densité spectrale de modes cette fréquence. Ce produit nous fournira la *distribution spectrale* du rayonnement thermique. Cette notion de distribution spectrale sera rencontrée plus loin pour la description des propriétés optiques des semi-conducteurs.

Voyons donc tout d'abord la manière dont les modes du champ E-M s'organisent dans l'espace vide. A cette fin considérons, comme nous l'avons fait pour l'étude des états électroniques, que les photons sont à l'équilibre thermique dans une enceinte cubique de côté L à la température T et dont les parois sont absorbantes (corps noir). Les parois étant absorbantes, le champ E-M y est nul. La résolution des équations de Maxwell nous indique que le champ E-M peut être décomposé en ondes planes de fréquence donnée ω (soit, en les modes du champ E-M). Dans l'enceinte les ondes planes font des aller-retour et forment des ondes stationnaires ($\sin(kx)$ à une dimension) dont l'amplitude doit être nulle aux parois en raison de l'hypothèse d'absorption. Les vecteurs d'onde k possibles ne peuvent donc que prendre les valeurs $k_n = n \pi/L$ et puisque les fréquences sont liées aux vecteurs d'onde par la vitesse de la lumière nous aurons un ensemble discret de modes du champ E-M : $\omega_n = k_n c$.

Distribution spectrale du rayonnement thermique

Densité spectrale du champ E-M

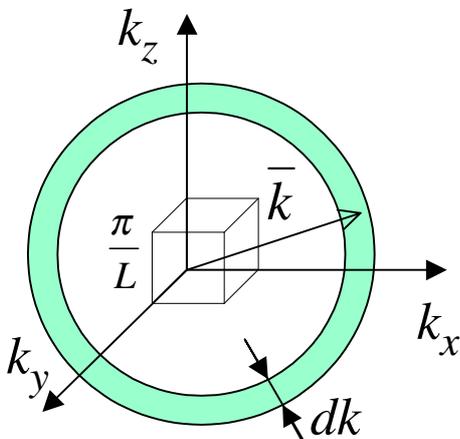


Parois absorbantes (équilibre thermique)

$$E = 0 \text{ en } x, y, z = 0, L$$

$$E = \sin k_n x, \quad k_n = n \pi / L \quad \longrightarrow \quad \omega_n = k_n c$$

De manière analogue aux états électroniques dans les cristaux, les photons s'organisent donc dans l'espace des k en formant un empilement cubique de maille π/L . Il est donc aisé d'en faire le compte afin de trouver la densité spectrale de modes du champ E-M, c'est à dire la densité de modes par unité de fréquence ω (ou d'énergie $h\omega$) des photons. A cette fin nous comptons le nombre de modes compris dans une pellicule sphérique de rayon k et d'épaisseur dk . A partir de la relation linéaire entre k et ω on peut alors facilement calculer la densité spectrale de modes. En faisant ce calcul pour un volume unitaire nous obtenons la densité spectrale par unité de volume et de fréquence $\rho(\omega) = dn_\omega/d\omega$.



$$dN_\omega = \frac{4\pi k^2 dk}{\left(\frac{\pi}{L}\right)^3} = \frac{4k^2}{\pi^2} dk = \frac{4\omega^2}{\pi^2 c^3} d\omega$$

$L = 1 \rightarrow dN_\omega = dn_\omega$

Densité spectrale de modes

$$\rho(\omega) = \frac{dn_\omega}{d\omega} = \frac{4}{\pi^2 c^3} \omega^2$$

La densité spectrale d'énergie émise dans l'enceinte est donc alors donnée par le produit de l'énergie moyenne contenue dans chaque mode (proportionnelle au nombre de photons du mode) multipliée par la densité spectrale de mode.

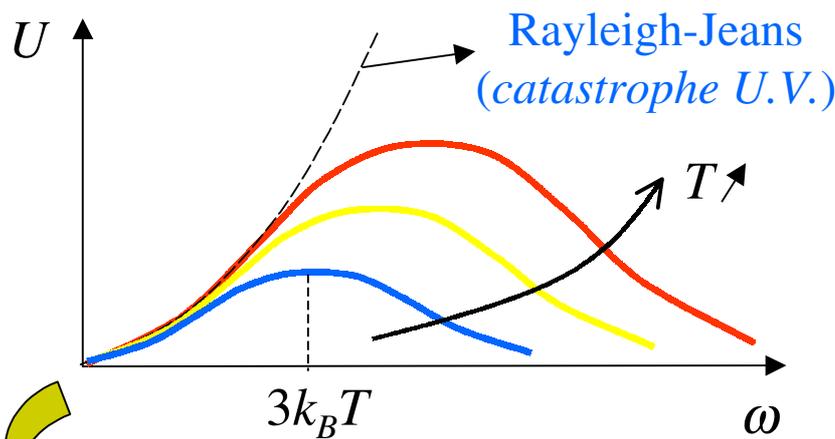
$$\text{Densité spectrale d'énergie émise : } U(\omega) = \langle E \rangle_\omega \rho(\omega)$$

C'est la grandeur mesurable (à l'aide d'un spectromètre) permettant de caractériser une source d'émission lumineuse quant à sa distribution de fréquence (ou d'énergie)

En remplaçant les valeurs de l'énergie moyenne et de la densité de modes calculées plus haut nous obtenons la loi de la radiation de Planck qui caractérise le contenu spectral des source de rayonnement thermique (rigoureusement, ce calcul n'est applicable qu'aux corps noirs, mais on peut en généraliser les résultats à toute forme de source thermique de lumière, c'est à dire, des sources formées par des matériaux soumis à un chauffage et non nécessairement à l'équilibre thermique, cf. lampes à filament). On constate que le spectre des sources thermiques a la forme d'une courbe en cloche dont la largeur vaut à peu près l'énergie thermique $k_B T$ et dont le maximum se trouve à peu près à $3 k_B T$. Comme l'indique le calcul ci-dessous, pour atteindre des longueurs d'onde visibles (inférieure à 1 micron), il faut chauffer l'enceinte jusqu'à des températures de l'ordre de plusieurs milliers de degrés Kelvin. Ceci implique donc naturellement que toute source thermique émettant dans le visible aura une très grande largeur spectrale et ne sera dès lors pas applicable à la transmission des données dans les fibres optiques (pour limiter la dispersion chromatique il faut un spectre le plus étroit possible).

Loi de la radiation de Planck

$$U(\omega) = \frac{4}{\pi^2 c^3 \hbar^2} \frac{\hbar^3 \omega^3}{e^{\hbar\omega/k_B T} - 1}$$



$$\left\{ \begin{array}{l} \text{Soit } \lambda = 1\mu\text{m} \\ \omega = 2\pi c/\lambda = 6\pi \cdot 10^{14} \text{ s}^{-1} \\ \hbar\omega = 6\pi \cdot 10^{-20} = 3k_B T \end{array} \right.$$

$T \approx 4000 \text{ K}$

Grande largeur spectrale ⊕ Isotropie ($\approx 10^{-8}$ photons/mode)

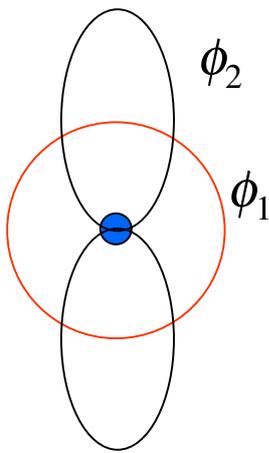
De plus, une source thermique est par essence même isotrope, ce qui signifie que son injection dans un guide d'onde (fibre optique) n'est pas efficace. A ce propos il est intéressant de rappeler que pour utiliser une fibre optique de manière optimale, il faut idéalement que la lumière soit émise non seulement selon une seule fréquence (mode temporel) mais également selon un seul mode spatial (car la fibre optique utilisée en pratique est mono-mode) c'est à dire un seul vecteur d'onde (une seule direction de propagation à mettre dans l'axe de la fibre). Avec une source thermique, en raison de la grande largeur spectrale et l'isotropie, on peut estimer que le nombre de photons moyen par mode spatio-temporel (une fréquence et une direction) est inférieur à l'unité et vaut 10^{-8} (ce qui ne convient donc pas à la transmission des données par fibre optique).

Transitions électroniques radiatives

La première chose qui vient à l'esprit pour palier aux désavantages des sources de lumière thermiques est d'utiliser les transitions électroniques des atomes constituant la matière. Comme de telles transitions correspondent à une valeur précise de l'énergie (ou fréquence) des photons émis, on peut facilement imaginer que leur utilisation permet de réduire la largeur des spectres lumineux par rapport aux spectres thermiques. De plus, un grand nombre d'atomes dans la nature présentent des transitions dont les énergies correspondent à des longueurs d'onde visible ce qui permet d'envisager leur utilisation pour les transmissions optiques.

Introduction : transitions atomiques

Il est intéressant de développer une vue intuitive du phénomène d'émission lumineuse par transition atomique. Considérons les deux premiers niveaux électroniques d'un atome. A l'équilibre, l'atome se trouve dans son état fondamental (état d'énergie le plus bas que peut avoir l'électron) dont la fonction d'onde a une fréquence ω_1 fixée par l'énergie $E_1 = \hbar\omega_1$. En vertu des lois de la mécanique quantique, cet état représente un nuage électronique diffus stationnaire donné par le module carré de la fonction d'onde complexe. Cet état quantique correspond donc à une absence de mouvement des électrons (s'il y avait mouvement de charge, il y aurait rayonnement synchrotron et l'atome s'effondrerait). Il en est de même pour l'état excité d'énergie E_2 . Par contre, l'état excité est instable car son énergie n'est pas minimale. Si l'atome est dans cet état il aura donc tendance à redescendre dans l'état fondamental.



vue intuitive du problème

$$\psi_1 = \phi_1 e^{-i\omega_1 t} \quad , \quad \psi_2 = \phi_2 e^{-i\omega_2 t}$$

$$\left[\omega_1 = \frac{E_1}{\hbar} \right] \quad \left[\omega_2 = \frac{E_2}{\hbar} \right]$$

- état fondamental : $|\psi_1|^2 = |\phi_1|^2$ (stable)

↪ pas de variations

Lors de la transition le système passe par une superposition des deux fonctions d'onde. Nous supposons ci-dessous le début de la transition pour lequel l'amplitude de l'état fondamental est petite (ε) et celle de l'état excité est encore proche de l'unité. Le module carré de la fonction d'onde totale présente un battement à la fréquence $\omega_1 - \omega_2$. Ce battement correspond à un mouvement du nuage électronique, ou plus précisément à un dipôle oscillant et est donc accompagné d'émission d'une onde E-M de fréquence $\omega_1 - \omega_2$.

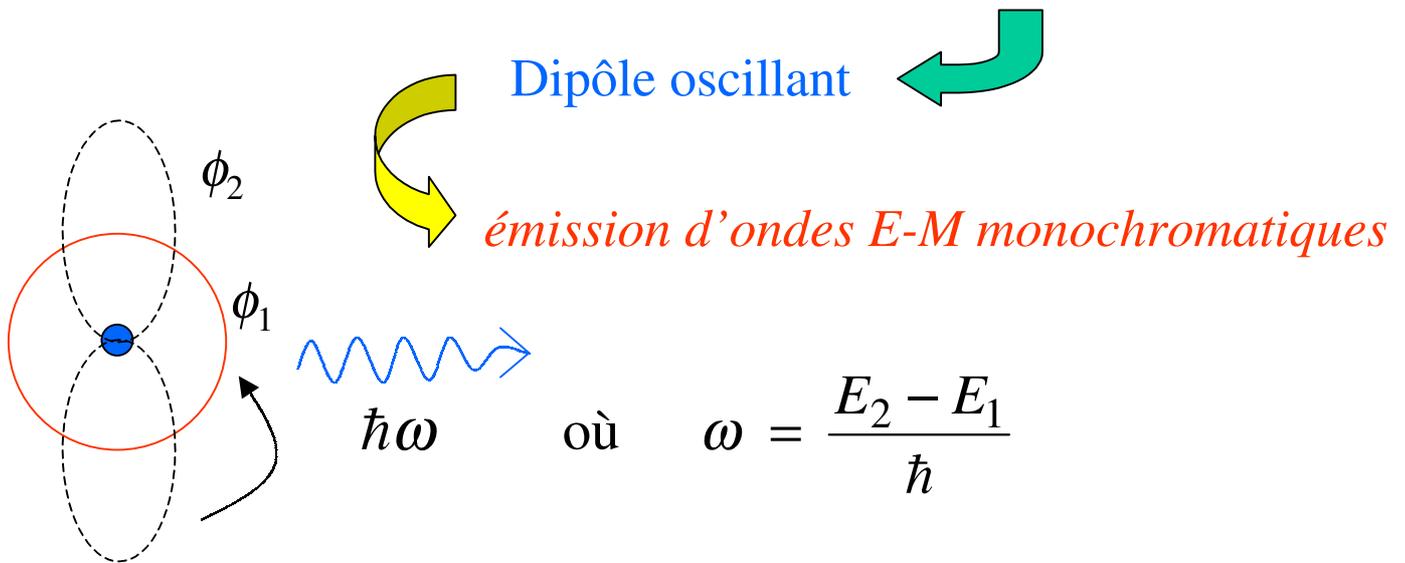
- état excité instable :

$$\psi = f_1 \phi_1 e^{-i\omega_1 t} + f_2 \phi_2 e^{-i\omega_2 t} + \dots$$

$$f_2 = 1 \quad \text{et} \quad f_1 = \varepsilon$$

Le rayonnement émis prend son énergie au système électronique qui se rapproche donc de plus en plus de l'état fondamental. Une fois à l'état fondamental le dipôle n'oscille plus et l'énergie transmise au champ E-M correspond exactement à la différence d'énergie des deux niveaux électroniques. Le paquet d'onde ainsi transmis par l'atome peut être vu comme un photon. Dans cette représentation du problème de transition électronique radiative, l'atome apparaît comme une antenne émettrice de photons au même titre qu'une antenne d'onde Hertzienne. Cette vue semi-classique non-rigoureuse du processus de transition radiative montre pourquoi l'énergie du photon est directement liée à sa fréquence par la constante de Planck.

$$|\psi|^2 = \varepsilon^2 |\phi_1|^2 + |\phi_2|^2 + 2 \operatorname{Re}(\varepsilon \phi_1 \phi_2^*) e^{-i(\omega_1 - \omega_2)t} + \dots$$



Exemples : lampes Na, Hg, Ne, ... ($\approx 10^{-2}$ photons/mode)

De telles transitions atomiques dans le visible sont exploitées pour les sources lumineuses telles que les lampes au Sodium qui ont remplacé les lampes à filament pour les applications à l'éclairage à grande échelle (cf. éclairage routier). Ces sources ont le désavantage d'être isotropes comme les sources thermiques mais leur largeur spectrale est tellement plus petite que l'on gagne à peu près six ordres de grandeur pour le nombre de photons moyen par mode spatio-temporel qui est de l'ordre de 10^{-2} . Comme nous le verrons, ce sont les transitions électroniques qui sont exploitées dans les semi-conducteurs pour l'émission et la détection de lumière. L'avantage des semi-conducteurs est qu'ils présentent une grande concentration d'électrons (et trous) libres qui à l'état excité forment autant de dipôles émetteurs de lumière. Pour bien se rendre compte de cet avantage il faut comparer les densités de porteurs dans les S-C avec les densités d'atomes (isolés) des gaz utilisés en pratique pour générer la lumière par transition atomique. De plus l'ingénierie des matériaux semi-conducteurs permet la réalisation de dispositifs bien adaptés à chaque type d'application (des sources de rayonnement anisotrope peuvent, par exemple, être facilement réalisées avec des dispositifs S-C de géométrie particulière afin d'obtenir un grand nombre de photons par mode spatio-temporel).

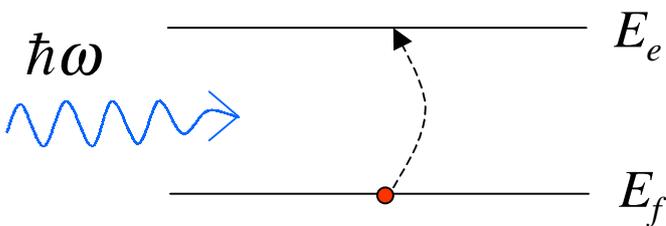
Absorption dans les semi-conducteurs

Nous abordons maintenant les propriétés optiques de S-C. La première question que l'on peut se poser est de savoir ce qui se produit lorsque l'on éclaire un matériau S-C. Comme nous le montrent beaucoup d'exemples dans la nature, les cristaux non conducteurs (non métalliques) sont souvent transparents ou colorés. La coloration indique qu'une partie du spectre visible y est absorbée en raison de mécanismes d'interaction lumière-matière spécifiques. L'étude des spectres d'absorption donne des informations sur ces mécanismes. Nous allons ici étudier les mécanismes de base de l'absorption dans les S-C. Pour cela nous commençons par un bref rappel de la notion d'absorption de lumière par les atomes.

Introduction : absorption d'un photon par un atome

L'absorption d'un photon par un atome est le processus inverse de la transition radiative. Soit un groupe de photons de fréquence ω correspondant à la différence d'énergie de deux niveaux électroniques d'un système d'atomes avec lesquels ils interagissent. La théorie de la mécanique quantique caractérise l'absorption en terme de probabilité d'absorption P_a d'un photon par un des atomes par unité de temps. Celle-ci est proportionnelle au coefficient d'Einstein B représentant en quelque sorte la force de l'interaction entre le photon et l'atome (pour les initiés: B est l'élément de matrice du dipôle électrique formé par les états concernés par la transition). La probabilité d'absorption est bien entendu également proportionnelle au nombre de photons à la fréquence ω ainsi qu'à la probabilité d'occupation de l'état fondamental de l'atome. Cette probabilité est fournie par la distribution de Fermi à l'équilibre thermique. De même la probabilité d'absorption sera proportionnelle à la probabilité que l'état d'énergie supérieure soit libre (soit $1 - f_e$ où f_e est la valeur de la distribution de Fermi à l'énergie de l'état supérieur ; cf. l'état d'énergie supérieure peut être excité avec une certaine probabilité en raison de l'agitation thermique (ceci sera d'autant plus marqué que l'énergie thermique $k_B T$ se rapproche de l'énergie de la transition)).

Système à deux niveaux



- B : coefficient d'Einstein
- $n(\omega)$: nbre de photons à ω
- f : probabilité d'occupation

Probabilité d'absorption P_a

$$P_a = B \cdot n(\omega) \cdot f_f \cdot (1 - f_e)$$

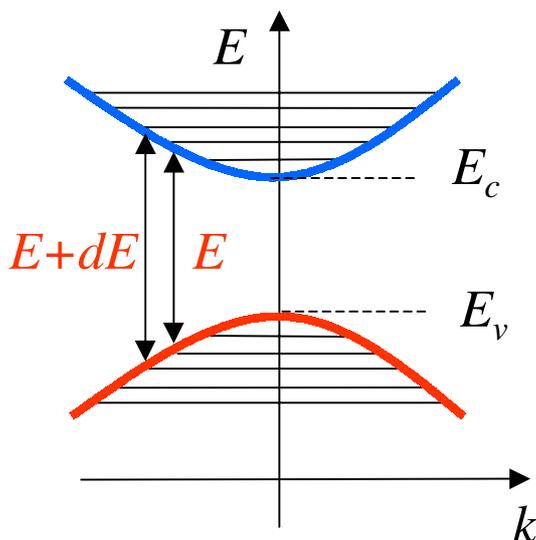
$$B \div |\langle \phi_f | W_{DE} | \phi_e \rangle|^2$$

$$n(\omega) \div |E|^2, [n(\omega) = \langle n \rangle_\omega]$$

$$\left(\begin{array}{c} \vec{E} \\ \vec{r} \\ e^- \\ + \end{array} \right) \quad W_{DE} = e \vec{r} \cdot \vec{E}$$

Absorption dans les S-C : *quasi-continuum d'états*

Voyons maintenant ce que devient l'absorption dans les S-C. Contrairement aux atomes isolés, les états électroniques dans les S-C sont en grand nombre et se regroupent en bandes où ils forment un quasi-continuum d'états. La bande de valence étant presque pleine et la bande de conduction étant presque vide les transitions électroniques entre états d'une même bande seront négligeables. Ce sont surtout les transitions entre bande de valence (pleine) et bande de conduction (vide) qui auront une importance pour le phénomène d'absorption. Comme dans ce cas les transitions ont lieu entre deux continuum d'états il faut caractériser l'absorption à partir de la notion de taux d'absorption par unité d'énergie (car on ne peut plus définir des niveaux énergétiques isolés comme dans le cas de l'atome). Nous considérons donc le nombre de transitions pouvant se produire avec une énergie comprise entre E et $E+dE$ et nous définissons le taux d'absorption par unité de temps et d'énergie r_a . Par unité de temps, le nombre de transitions (ou d'absorptions de photons) ayant une énergie comprise entre E et $E+dE$ vaut par définition $r_a dE$. Celui-ci est donné par la somme des probabilités P_a de transitions ayant lieu entre tous les niveaux de la bande de valence E_f et les niveaux et de la bande de conduction E_e séparés d'une énergie comprise entre E et $E+dE$. Ici E_f et E_e jouent le rôle des états fondamental et excité dans l'exemple de l'atome isolé considéré ci-dessus. L'expression de P_a est donc identique à celle de l'atome si ce n'est que maintenant on ne considère plus un flux de photons à une fréquence unique, mais un ensemble de photons caractérisé par une densité spectrale $\rho(\omega)$ (en raison du continuum d'états, considérer une fréquence unique n'a plus de sens).



Taux d'absorption par unité d'énergie r_a

$$r_a dE = \sum_{E < (E_e - E_f) < E + dE} P_a$$

$$r_a dE = \sum_{E < (E_e - E_f) < E + dE} B \rho(\omega) f_f (1 - f_e)$$

Le nombre de transitions d'énergie comprise entre E et $E+dE$ est très grand vu que ces transitions peuvent partir de n'importe quel niveau dans la bande de valence pour autant que l'énergie E soit suffisante que pour atteindre la bande de conduction. Pour simplifier le calcul nous devons donc recourir à des simplifications. Nous supposons que le coefficient d'Einstein B est le même pour toutes les transitions de bande à bande. De même nous considérerons que les distributions de Fermi dans les deux bandes sont uniformes (ce qui peut-être justifié dans un S-C intrinsèque pour lequel le niveau de Fermi est au milieu du gap). La somme sur toutes les transitions se réduit donc au nombre de transitions lui-même. Soit dN_t le nombre de transitions d'énergie comprises entre E et $E+dE$.

Soit : $B = \text{cste}$ et $f_f = f_v, f_e = f_c$



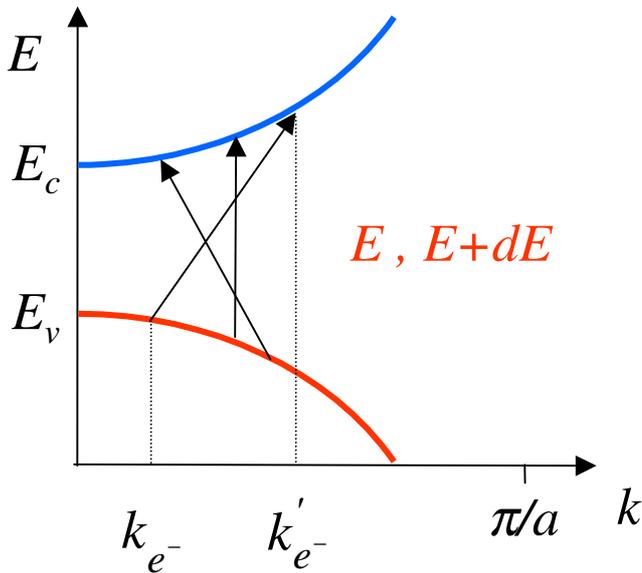
$$r_a dE = B \rho(\omega) f_v (1 - f_c)$$

↗ Nbre de transitions

$$\sum_{E < (E_e - E_f) < E + dE} 1 \quad dN_t$$

Calcul du nombre de transitions dN_t

Un raisonnement simple nous permet de simplifier fortement le calcul du nombre de transition dN_t . Comme une transition est une interaction entre deux particules, un électron et un photon, nous savons qu'elle doit répondre à la loi de conservation de la quantité de mouvement (la loi de conservation de l'énergie étant d'office satisfaite en considérant implicitement que l'énergie des photons est égale à l'énergie de la transition $h\omega = E$). La quantité de mouvement des photons et des électrons est donnée par hk , c'est à dire, proportionnelle au vecteur d'onde de leur fonction d'onde associée. Pour être significatif, le changement de quantité de mouvement d'un électron doit être non-négligeable devant la demi largeur de la zone de Brillouin soit $\pi/a \approx 10^{10} \text{ m}^{-1}$. Or, la quantité de mouvement d'un photon dans le visible est donné par $2\pi/\lambda \approx 10^6 \text{ m}^{-1}$ ce qui est de 4 ordres de grandeur inférieur à la largeur de la zone de Brillouin. On peut donc considérer dans une première approximation que les transitions électronique induites par les photons ont lieu sans changement de la quantité de mouvement. On dit que les transitions sont dans ce cas « verticales ». Dans le langage de la mécanique quantique on appelle cette restriction une règle de sélection (sélection parmi toutes les transitions possibles). Cette règle de sélection limite bien entendu considérablement le nombre de transitions à prendre en compte, comme le suggère le diagramme ci-dessous où seule une des trois transitions d'énergie E représentées par des flèches sera à prendre en compte.



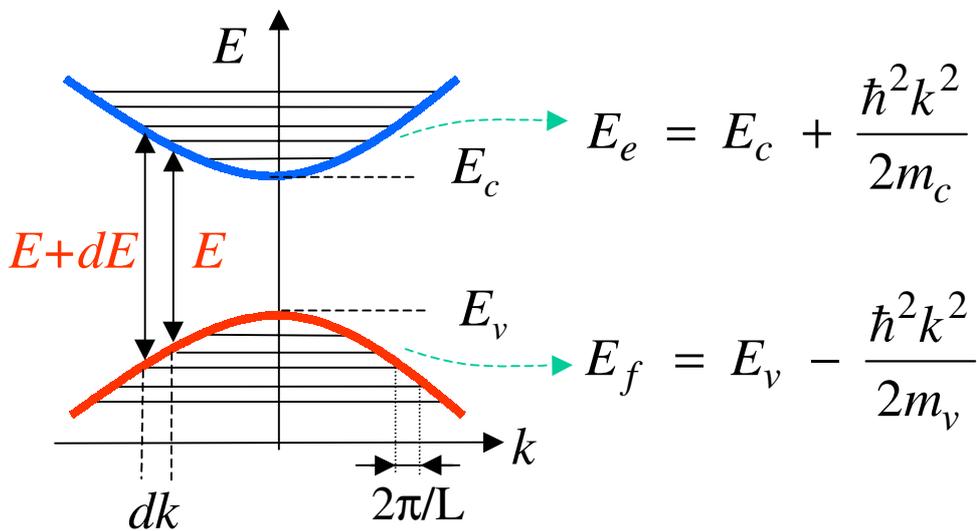
Quantité de mouvement

$$\left\{ \begin{array}{l} k_{e^-} \approx \frac{\pi}{a} \approx 10^{10} \text{ m}^{-1} \\ k_p \approx \frac{2\pi}{\lambda} \approx 10^6 \text{ m}^{-1} \end{array} \right.$$

Règle de sélection

$$k_{e'^-} = k_{e^-} + k_p \approx k_{e^-}$$

Le calcul du nombre de transitions peut alors se faire (comme on l'a fait plus haut pour d'autres grandeurs) en se basant sur le fait que l'organisation des électrons dans l'espace réciproque est un simple empilement cubique. Puisque les transitions sont verticales, à l'intervalle d'énergie de transition $[E, E+dE]$ correspond un intervalle de vecteur d'onde dk déterminé par la courbure des bandes de valence et de conduction (voir schéma ci-dessus où l'on voit que une énergie de transition E correspond à une valeur unique de k). Calculons donc tout d'abord le nombre de transitions qui peuvent avoir lieu dans l'intervalle $[k, k+dk]$. Celui-ci est simplement donné par le rapport du volume de la pellicule sphérique d'épaisseur dk et le volume de la cellule élémentaire, soit $8\pi^3$ pour un volume unitaire de S-C.



Nombre de transitions : $dN_t = 2 \cdot \frac{4\pi k^2 dk}{8\pi^3} = \frac{k^2}{\pi^2} dk$

Comme dit plus haut, la valeur de k donnée correspond une énergie de transition E unique fixée par la courbure des bandes. On a vu dans le rappel de physique des S-C que les bandes pouvaient être décrite par des paraboles en introduisant la notion de masse effective comme indiqué par les expressions de E_e et E_f ci-dessus. Le lien entre l'énergie de la transition E et la valeur de k est donc immédiat. Il est donné dans le calcul ci-dessous où l'on a introduit une nouvelle masse effective m^* .

Energie de la transition à valeur de k donnée :

$$E = E_e - E_f = E_c - E_v + \frac{\hbar^2 k^2}{2} \left(\frac{1}{m_c} + \frac{1}{m_v} \right)$$

$$E = E_g + \frac{\hbar^2 k^2}{2m^*} \quad \text{où} \quad \frac{1}{m^*} = \left(\frac{1}{m_c} + \frac{1}{m_v} \right)$$

La différentiation de cette expression fourni le lien entre dE et dk , ce qui nous permet de calculer aisément le nombre de transitions possibles dans dE et donc également le taux d'absorption par unité d'énergie et de temps dont l'expression est rappelée ci-dessous.

$$\Rightarrow dk = \frac{1}{2} \sqrt{\frac{2m^*}{\hbar^2 (E - E_g)}} dE \quad \left(\text{et } dN_t = \frac{k^2}{\pi^2} dk \right)$$

Taux d'absorption

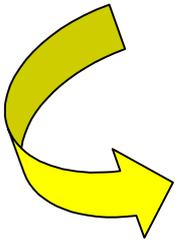
$$r_a dE = B \rho(\omega) f_v (1 - f_c) \overbrace{\sum_{E < (E_e - E_g) < E + dE} 1}^{dN_t}$$

En introduisant la notion de densité de transition par unité d'énergie comme indiqué ci-dessous, nous pouvons écrire le taux d'absorption par unité d'énergie et de temps. Sa forme montre le rôle de la force de transition (coefficient d'Einstein représentant l'efficacité de l'interaction photon-électron), la densité de photons, les probabilité de présence des électrons dans la bande de valence (états bas de la transition) la probabilité d'avoir des états libres dans la bande de conduction (états excités de la transition) et finalement le rôle de la forme des bandes au travers de la densité de transition. Notons pour ce dernier point que la densité de transition est nulle à des énergies inférieures à celle du gap et qu'elle a une dépendance parabolique en l'énergie au delà du gap en raison de la forme parabolique des bandes dans l'espace réciproque.

Densité de transitions

$$\rho_t(E) \equiv \frac{dN_t}{dE}$$

$$\rho_t(E) = \frac{1}{2\pi^2} \frac{(2m^*)^{3/2}}{\hbar^3} \sqrt{E - E_g}$$



Taux d'absorption

$$r_a = B \rho(\omega) f_v (1 - f_c) \rho_t(\omega)$$

Par définition, l'intégrale sur la fréquence (où l'énergie) du taux d'absorption est le taux de décroissance du nombre de photons se propageant dans le S-C. Cette interprétation physique sera mieux comprise si l'on considère un faisceau monochromatique de n_ω photons. Pour un tel faisceau, la densité de photons par unité d'énergie (ou densité spectrale de photons) devient une distribution de Dirac. On peut alors définir le taux d'absorption par unité de temps R_a comme indiqué ci-dessous. Le graphique de R_a montre que les photons ne sont pas absorbés si leur énergie est inférieure au gap et que leur absorption augmente de manière parabolique au delà du gap. Cette forme du profil spectrale de l'absorption est caractéristique des semi-conducteurs.

Interprétation physique

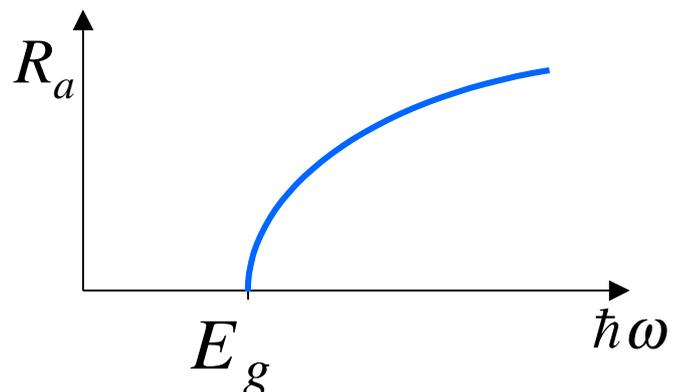
faisceau monochromatique

$$\frac{dn_{ph}}{dt} = - \int r_a(\hbar\omega') d\hbar\omega'$$

$$\rho(\omega') = n_\omega \delta(\hbar\omega' - \hbar\omega)$$

$$\frac{dn_{ph}}{dt} = - R_a(\omega) \quad [s^{-1}]$$

où $R_a = B n_\omega f_v (1 - f_c) \rho_t(\omega)$

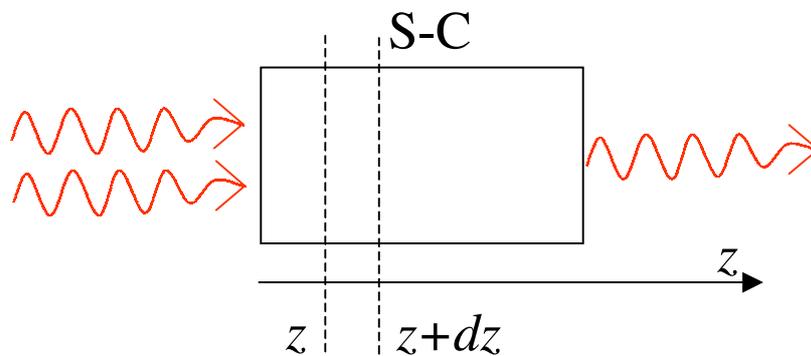


Coefficient d'absorption

Voyons comment on procède en pratique pour mesurer le spectre d'absorption d'un S-C. La mesure peut se faire à l'aide d'un faisceau monochromatique unidirectionnel dont on fait varier la fréquence. Considérons un paquet de photons à un endroit donné dans le faisceau. Ce paquet se propage à la vitesse de la lumière dans le S-C qui vaut $v = c/n$ où n est l'indice de réfraction du S-C. De ce fait le taux de disparition temporel des photons se traduit par une décroissance du nombre de photons le long de l'axe de propagation, soit l'axe z . Le taux de décroissance spatial sera lié au taux de décroissance temporel par la vitesse de la lumière $v = c/n$ comme indiqué ci-dessous. On appelle la quantité $R_d/(vn_\omega)$ le coefficient d'absorption α du S-C.

$$n_{ph} = n_\omega$$

Rayonnement monochromatique unidirectionnel



$$\frac{dn_\omega}{dt} = -R_a(\omega) \quad \longrightarrow \quad \frac{dn_\omega}{dz} = ? \quad , \quad dz = v dt \quad \text{où} \quad v = \frac{c}{n}$$

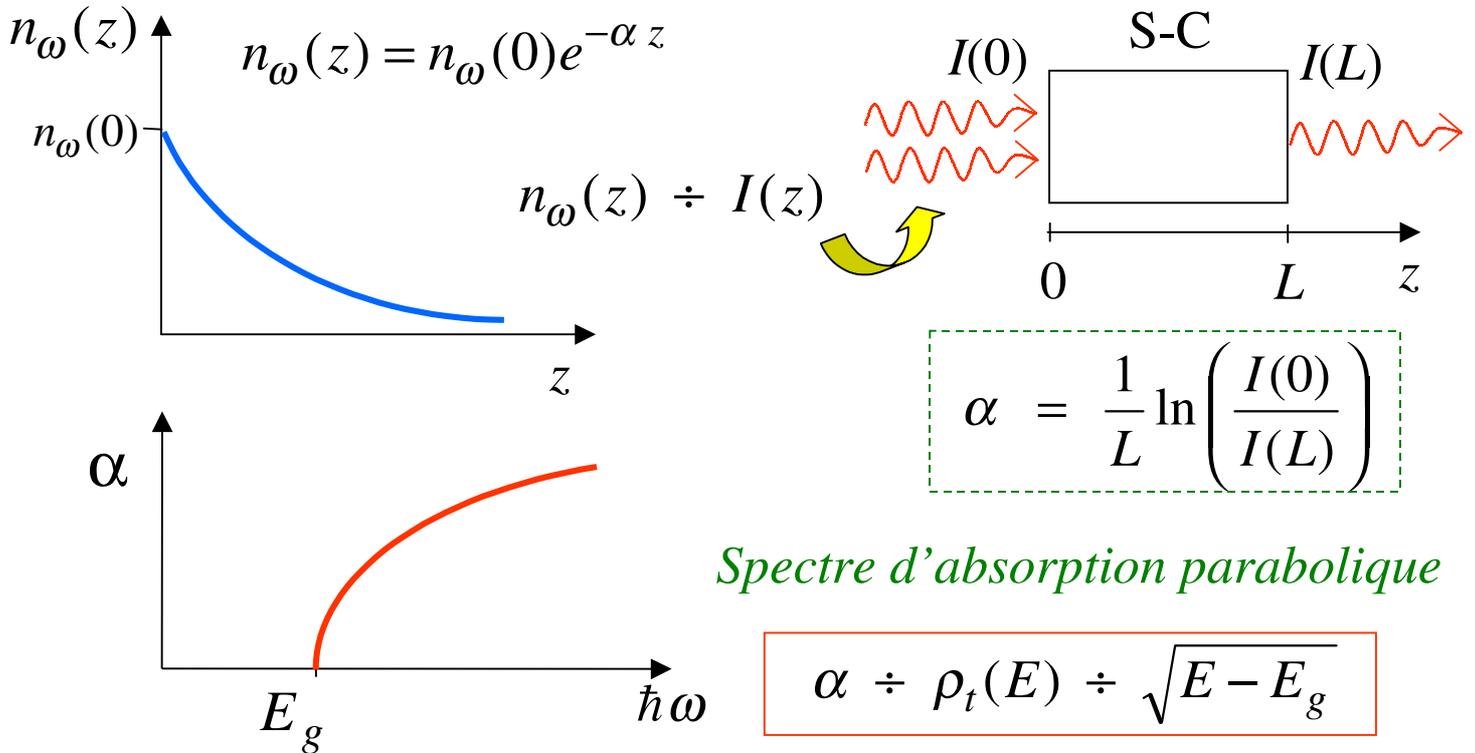
$$\frac{dn_\omega}{dz} = \frac{1}{v} \frac{dn_\omega}{dt} = -\frac{R_a}{v} = -\frac{1}{v} B n_\omega f_\nu (1 - f_c) \rho_t(\omega)$$

Coefficient d'absorption

$$\frac{dn_\omega}{dz} \equiv -\alpha n_\omega$$

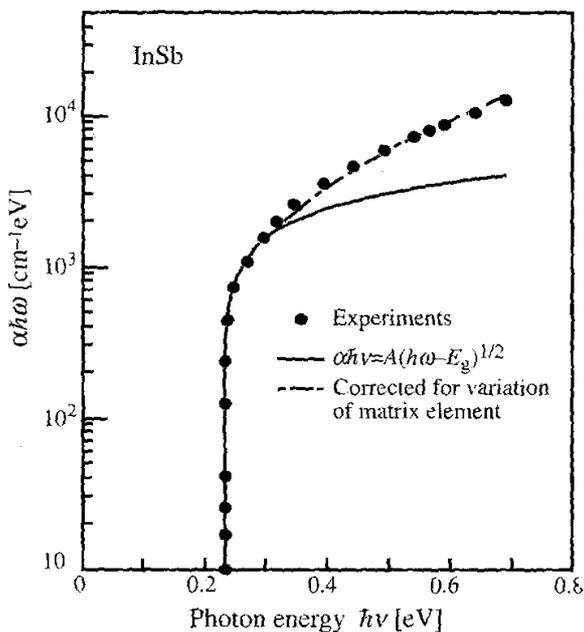
$$\alpha = \frac{1}{v} B f_\nu (1 - f_c) \rho_t(\omega)$$

En pratique, la mesure du coefficient d'absorption se fait en mesurant le rapport des intensités incidente et transmise d'un faisceau laser monochromatique traversant un échantillon de S-C de longueur L (l'intensité étant proportionnelle au flux de photons). En faisant varier la longueur d'onde de la lumière, on peut obtenir le spectre d'absorption (α en fonction de la fréquence ω). D'après ce que nous avons vu ci-dessus, pour un S-C idéal intrinsèque ce spectre aura un seuil à l'énergie du gap et sera de forme parabolique.

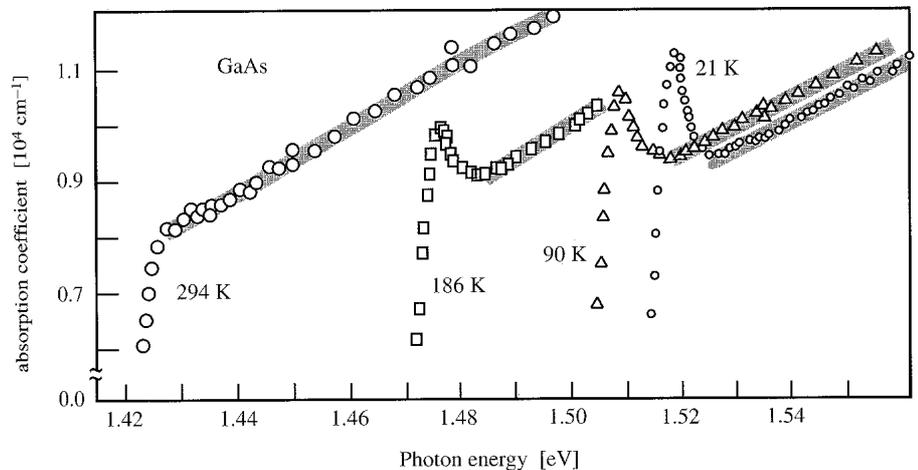


Comparaison théorie-expérience :

Les figures ci-dessous représentent les spectres d'absorption théoriques et expérimentaux. Sur la figure de gauche on voit la comparaison théorie-expérience pour l'InSb (Antimoniure d'indium). La courbe continue correspond à la théorie exposée ci-dessus, tandis que la courbe pointillée est basé sur un calcul prenant en compte la dépendance fréquentielle du coefficient d'Einstein B (ou élément de matrice de la transition).



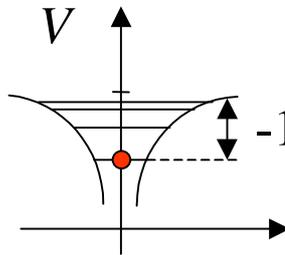
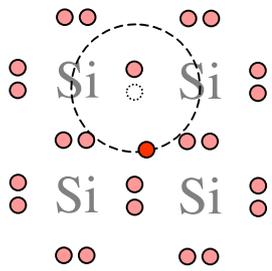
La figure de droite montre l'apparition à basse température de pics d'absorption au niveau de l'énergie du gap. Ces pics sont dus à l'existence d'excitons dans les S-C.



Pics excitoniques

Exciton: *État lié hydrogénoïde*

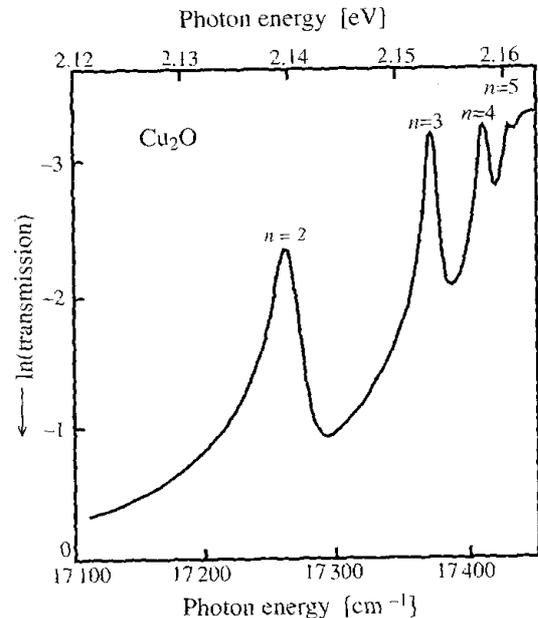
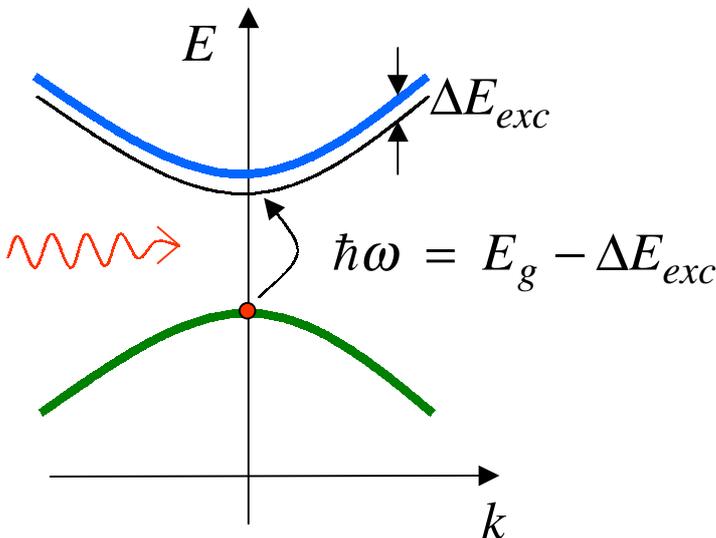
L'exciton est un état lié entre un électrons libre de la bande de conduction et un trou (de charge positive). Cet état lié est identique à un atome d'hydrogène car il est créé par le même potentiel électrostatique. On pourrait donc s'attendre à ce que, comme pour l'hydrogène, l'énergie de liaison soit de l'ordre de 13.6 eV (pour l'état fondamental). En raison de l'effet d'écran (la présence d'un grand nombre de porteurs de charges à proximité de la paire électron-trou fait que la permittivité du vide doit être remplacée par un permittivité beaucoup plus élevée) et de la masse réduite des porteurs dans les S-C, les énergie de liaisons sont beaucoup plus faibles et sont de l'ordre de l'énergie thermique (à température ambiante 300K) soit plus ou moins 10^{-2} eV. Les électrons de la bande de conduction ont une énergie minimum E_c (bas de la bande de conduction) mais lorsqu'ils forment des excitons leur énergie sera diminuée de l'énergie de liaison. Ceci signifie que l'on trouvera dans le gap des niveaux d'énergie légèrement inférieurs à E_c . A la bande de conduction de densité d'état parabolique il faut donc ajouter les niveaux d'énergie discrets des excitons (chaque niveau correspondant à un état quantique de l'exciton : l'état fondamental, et les états excités d'énergie de liaison plus faible (comme pour l'atome d'hydrogène, il y en a théoriquement une série infinie)). La théorie présentée ci-dessus devrait être modifiée pour prendre en compte les niveaux excitoniques dans la densité de transition $\rho_i(\hbar\omega)$ qui ne serait alors plus parabolique mais présenterait des pic aux basses énergies. Le spectre d'absorption ne sera donc plus simplement parabolique mais présentera des pics correspondant aux excitons. D'un point de vue physique il est facile de comprendre l'existence de ces pics dans le spectre d'absorption en se disant qu'un photon peut créer une paire électron-trou qui soit directement liée sous forme d'un exciton (la théorie ci-dessus ne prévoit que la possibilité de créer des paires de porteurs libres). Les photons responsables de ce mécanisme auront obligatoirement une énergie légèrement inférieur à l'énergie du gap. Ce qui explique la présence de pics dans les spectres d'absorption à basse température. A température ambiante ce mécanisme n'est en effet pas possible puisque les excitons sont automatiquement détruits par l'agitation thermique.



Effet écran et masse réduite

$$\Delta E_{exc} \sim 10^{-2} \text{ eV}$$

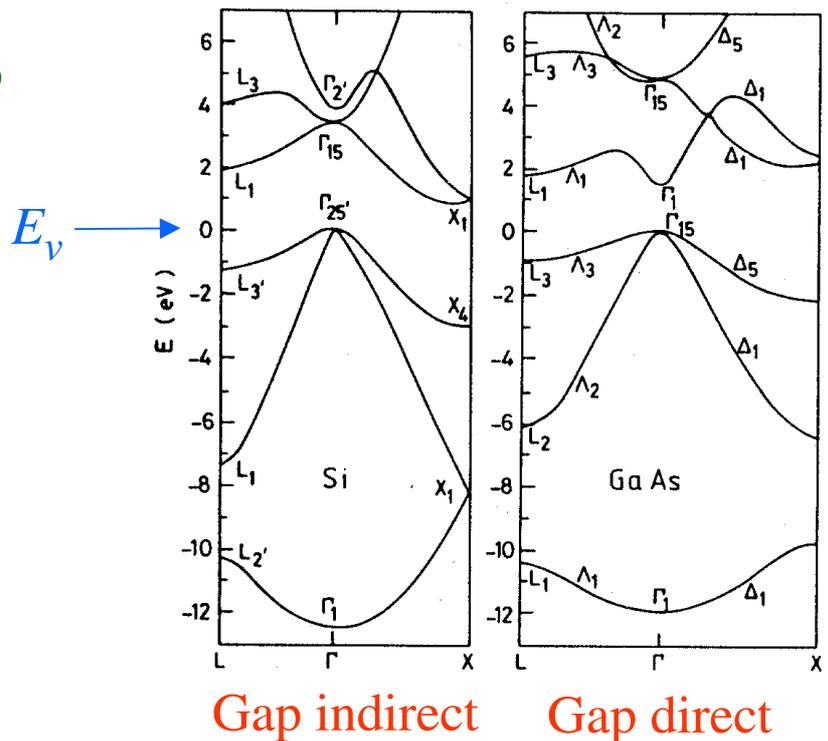
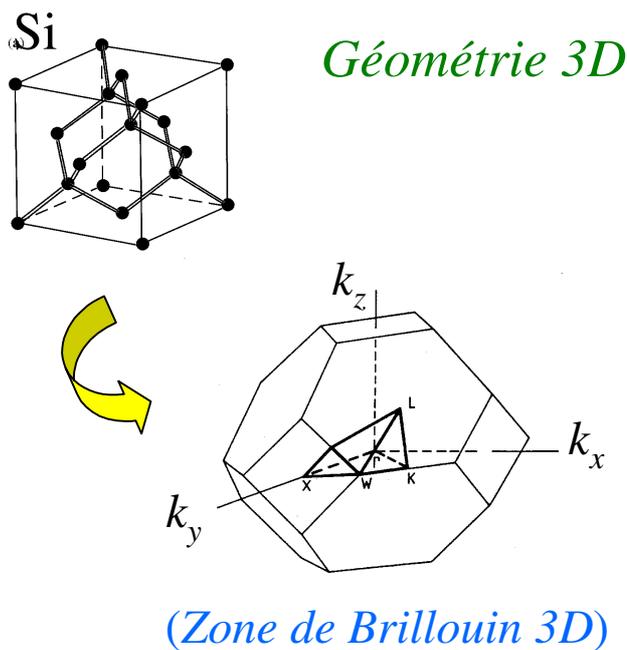
(cf. *Énergie thermique*)



La figure de droite ci dessus, montre l'existence des états excités (série de Rydberg) des excitons à partir du spectre d'absorption de l'oxyde de cuivre à très basse température (hydrogène liquide).

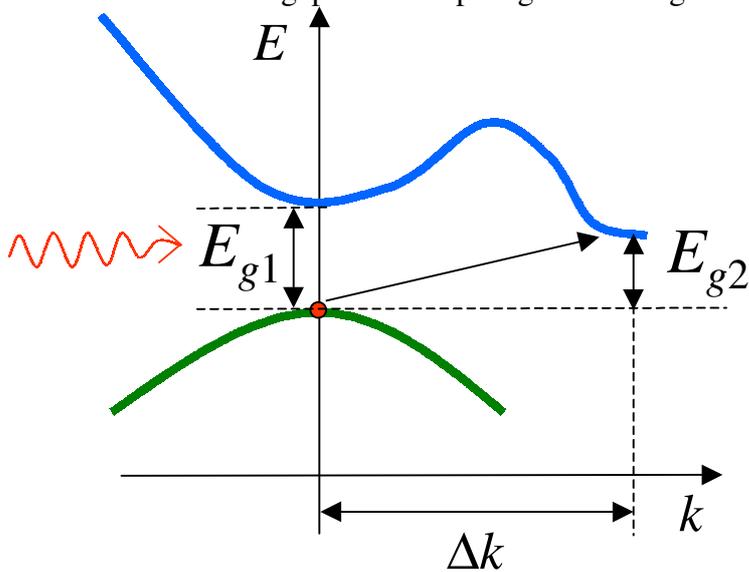
Absorption dans les S-C à gap indirect

La maille élémentaire du cristal de Silicium représentée ci-dessous illustre la complexité de la structure des S-C par rapport au modèle à une dimension (ou son extension à 3 dimensions conduisant à un empilement cubique dans l'espace réciproque) considéré dans le rappel du début du cours. Comme on le voit ci-dessous, la maille élémentaire du Silicium est composée de structures tétraédrales. La structure la plus élémentaire répétée de manière périodique est constituée de deux atomes (elle est représentée par les bâtonnets joignant les sommets au centre du tétraèdre). Il y a donc deux atomes par cellule élémentaire et donc 8 électrons de valence. Ces électrons de valence remplissent donc les 4 premières bandes permises (et non 2 comme le suggérait l'approche intuitive exposée au début du cours qui est basée sur la présence d'un atome unique par cellule élémentaire). Les cellules élémentaires (les bâtonnets) sont répétés selon une structure cubique face centrée possédant des périodicités différentes selon les différentes directions principales du cristal. Il en résulte une structure complexe de l'empilement dans l'espace réciproque dont la première zone de Brillouin est représentée ci-dessous également. De cette complexité de structure découle une complexité de la structure de bande d'énergie des S-C. Comme illustré ci-dessous, les bandes peuvent avoir plusieurs extrema et peuvent être dégénérées à certains endroits de la zone de Brillouin. L'explication de ces structures de bande sort du cadre de ce cours. La seule chose qui nous intéresse ici est la possibilité d'avoir un minimum absolu de la bande de conduction au bord de la zone de Brillouin et non pas en son centre comme nous le suggérait l'approche intuitive du début du cours. C'est le cas pour le Silicium comme illustré dans le diagramme de bande de gauche ci-dessous (l'énergie zéro indique le sommet de la 4^{ème} bande de valence). On dit dans ce cas que le S-C présente un gap indirect. Le diagramme de droite montre le cas de l'Arseniure de Gallium qui lui présente un gap direct tel que prévu par la théorie élémentaire.



L'appellation gap « indirect » est justifiée par le raisonnement suivant : les électrons libres de la bande de conduction ont toujours tendance à se placer au minimum de la bande de conduction pour minimiser leur énergie. En l'absence d'excitation extérieure (pas de photons incidents) ces électrons descendront dans la bande de valence là où ils trouveront de la place c'est à dire en son sommet qui lui est toujours au centre de la zone de Brillouin. Pour descendre dans la bande de valence les électrons d'un S-C à gap indirect doivent donc changer fortement leur quantité de mouvement. Comme nous l'avons vu plus haut, ceci ne peut pas être fait à partir de l'échange de quantité de mouvement avec les photons (les photons ont une quantité de mouvement de 4 ordres de grandeur inférieur à la largeur de la zone de Brillouin). La seule manière pour les électrons de modifier leur quantité de mouvement est de l'échanger avec les phonons. Ils descendront donc par l'intermédiaire d'un mécanisme additionnel. En ce sens la transition n'est pas directe ce qui justifie la terminologie. Pour les S-C à gap direct, les électrons peuvent redescendre directement dans la bande de valence en émettant un photon (pour la conservation de l'énergie).

Un S-C à gap indirect a bien entendu aussi un gap direct d'énergie supérieure (comme indiqué ci-dessous). Mais, puisque ce sont les transitions obliques (c-à-d, avec modification de quantité de mouvement) qui sont favorisées énergétiquement (les électrons se placent au minimum absolu de la bande de conduction) et qui ont donc lieu en pratique, on appelle le gap indirect le gap « fondamental » pour le différencier du gap direct de plus grande énergie.

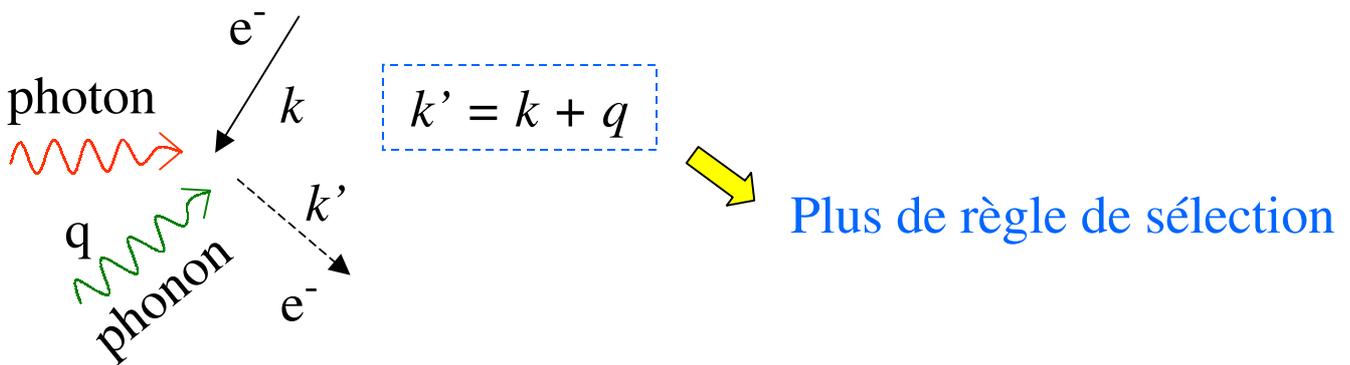


Transitions obliques favorisées énergétiquement

$$E_{g2} = \text{gap fondamental} = E_{gf}$$

La figure ci-dessus illustre l'absorption dans un S-C à gap indirect. Lorsque les photons ont une énergie inférieure au gap direct, ils pourront être absorbés par transition électronique vers le minimum absolu de la bande de conduction pour autant que leur énergie soit supérieure au gap fondamental. Pour que de telles transitions puissent avoir lieu les électrons doivent, non seulement échanger leur énergie avec les photons, mais doivent aussi échanger leur quantité de mouvement avec un phonon. Comme les phonons sont toujours présents en plus ou moins grand nombre (dépendant de l'agitation thermique du réseau cristallin), de telles transitions peuvent toujours avoir lieu. Il n'y a donc plus de règle de sélection pour les transitions électroniques dans un S-C à gap indirect (contrairement au S-C à gap direct où l'on a vu que la règle de sélection est la conservation de la quantité de mouvement).

Conservation de la quantité de mouvement : rôle des phonons

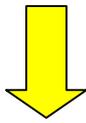


Le calcul du taux d'absorption de photons dans les S-C à gap indirect est donc à priori plus complexe que dans le cas des S-C à gap direct où le calcul se ramenait au compte du nombre de transitions dans l'espace réciproque. On peut néanmoins facilement donner une idée de l'allure du spectre d'absorption d'un S-C à gap indirect en se disant que le nombre de transitions possibles est simplement proportionnel au produit du nombre d'états (occupés) de la bande de valence par le nombre d'états (libres) de la bande de conduction.

Calcul du nombre de transitions dans un S-C à gap indirect

Pour une énergie de photon donnée $h\omega$ le nombre de transitions (ou la probabilité de transition) entre deux niveaux d'énergie donnés sera proportionnel au produit des densités d'états en ces deux niveaux (puisque'il n'y a pas de règle de sélection seul ce produit peut intervenir dans le taux de transition). Pour l'énergie du photon $h\omega$ donnée il y a un ensemble continu de transitions possibles faisant intervenir des états libres d'énergie plus ou moins élevée dans la bande de conduction. Pour désigner le niveau d'arrivée dans la bande de conduction nous introduisons sa hauteur dans la bande soit ε . Le taux total de transitions sera alors proportionnel à l'intégrale du produit des densités d'états en ε pour la bande conduction et en $h\omega - E_{gf} - \varepsilon$ pour la bande de valence. L'intégrale sur ε étant prise de 0 à $h\omega - E_{gf}$ (puisque'au delà de cette valeur ε serait trop grand car il correspondrait à un état de départ hors de la bande de valence).

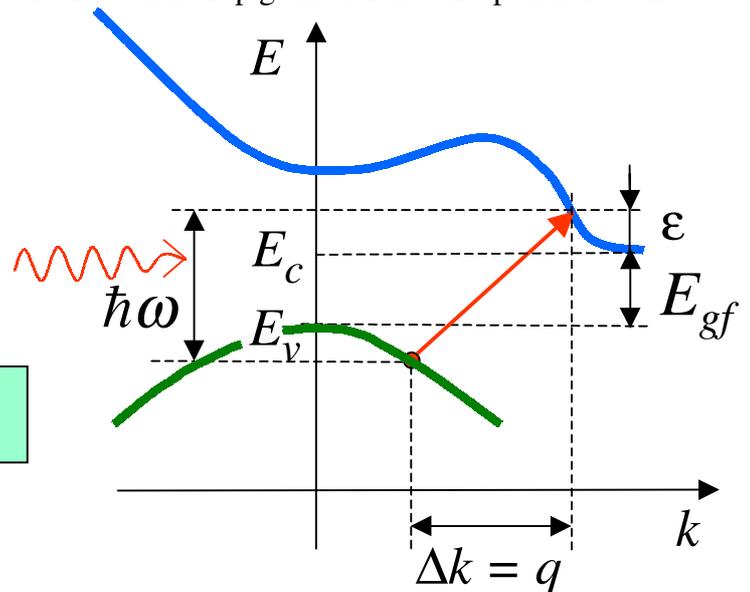
Toutes transitions permises



probabilité de transition :

$$p_t \div \rho_c \times \rho_v$$

Densité d'états



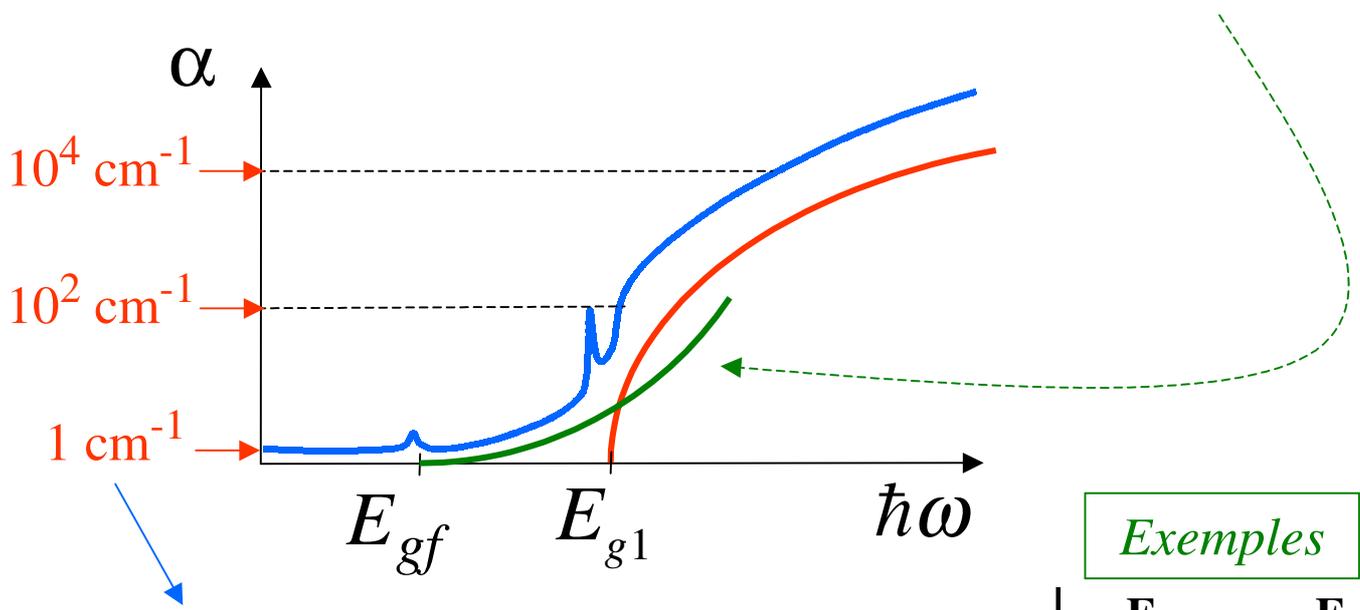
$$\left\{ \begin{array}{l} \rho_c = \frac{1}{2\pi^2} \left(\frac{2m_c}{\hbar^2} \right)^{3/2} \sqrt{E - E_c} \rightarrow \rho_c = \frac{1}{2\pi^2} \left(\frac{2m_c}{\hbar^2} \right)^{3/2} \sqrt{\varepsilon} \\ \rho_v = \frac{1}{2\pi^2} \left(\frac{2m_v}{\hbar^2} \right)^{3/2} \sqrt{\hbar\omega - E_{gf} - \varepsilon} \end{array} \right.$$



$$\rho_t(\hbar\omega) \div \int_0^{\hbar\omega - E_{gf}} \sqrt{\varepsilon(\hbar\omega - E_{gf} - \varepsilon)} d\varepsilon = \frac{\pi}{8} (\hbar\omega - E_{gf})^2$$

On arrive donc à la conclusion que la densité de transition (nombre de transitions dans un intervalle d'énergie dE divisé par dE , voir plus haut) est proportionnelle au carré de l'écart entre l'énergie du photon et le gap fondamental. Le spectre à donc une allure parabolique inversée par rapport au cas des S-C à gap direct qui eux avaient un spectre proportionnel à la racine carrée de $h\omega - E_g$.

Bien entendu, lorsque l'énergie du photon atteint l'énergie du gap direct du S-C, l'absorption sera dominée par les transitions directes qui sont beaucoup plus probables que les transition indirectes (puisque ces dernières font intervenir un processus à trois particules : l'électron, le photon et le phonon pour la conservation de la quantité de mouvement). L'allure générale des spectres d'absorption dans les S-C est donc celle représentée ci-dessous où l'on voit en E_{gf} un premier pic excitonique (si la température est basse) suivit d'une croissance quadratique jusqu'à l'approche du gap direct où l'on voit un deuxième pic (ou une série de pics) excitonique (si la température est basse) suivit de la croissance parabolique inversée caractéristique de la transition entre bandes avec conservation de la quantité de mouvement.



Absorption porteurs libres et phonons

En deçà du gap fondamental on a bien entendu une absorption résiduelle due aux transitions internes à la bande de valence, la diffusion sur les porteurs libres (diffusion Rayleigh), diffusion avec les phonons etc...

Exemples

	E_{gf}	E_{g1}
Si	1.13 eV	3.4 eV
Ge	0.76 eV	0.88 eV

Notez que le Silicium et le Germanium sont tout deux des S-C à gap indirect

Notez également les ordres de grandeur de l'absorption dans un S-C. Au delà du gap direct elle est typiquement de l'ordre de 10^4 cm^{-1} . Ce qui signifie que l'intensité de la lumière s'annule virtuellement sur quelques microns d'épaisseur de S-C. Une valeur typique pour un pic excitonique est 10^2 cm^{-1} . Tandis que l'absorption en deçà du gap fondamental est typiquement inférieure à 1 cm^{-1} .

Notons pour terminer ce chapitre sur l'absorption dans les S-C que la notion d'absorption vue ici se limite aux S-C à l'équilibre, c'est à dire, les S-C qui ne sont pas soumis à une excitation extérieur. Or, pour faire la mesure du coefficient d'absorption nous avons vu qu'il faut faire passer un faisceau laser au travers d'un échantillon de S-C. Ce faisceau laser provoque des transitions vers la bande de conduction et le S-C n'est donc plus rigoureusement à l'équilibre. Les mesures d'absorption doivent donc se faire à très faible intensité. Si ce n'était pas le cas et que l'intensité lumineuse était forte, un grand nombre d'électrons pourraient passer dans la bande de valence à un point tel que la bande de valence se dépeuplerait et plus aucune transition ne pourrait avoir lieu faisant croire alors que le S-C n'est plus absorbant. Nous verrons ce problème en détail plus loin avec la notion d'absorption saturable dans les S-C.

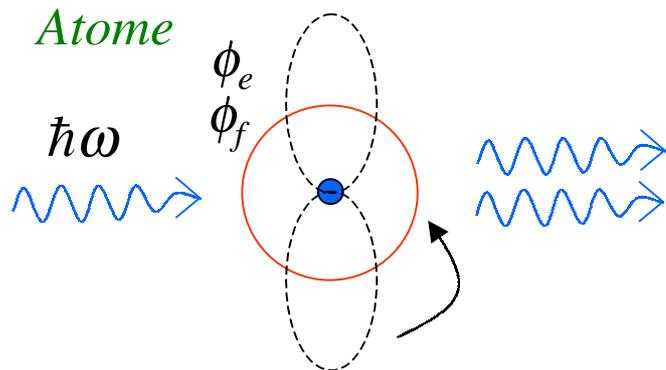
Nous verrons également plus loin comment l'absorption de photon dans les S-C peut être mise en application pour la détection de lumière.

Emission stimulée dans les S-C

Lorsqu'un S-C est excité d'une quelconque manière (par passage d'un courant dans une jonction $p-n$ ou par absorption de lumière comme nous l'avons vu ci-dessus), c'est à dire, que des électrons se trouvent dans la bande de conduction, ces derniers redescendent dans la bande de valence de manière spontanée car le système aura toujours tendance à retrouver l'équilibre (état d'énergie minimale). En redescendant les électrons libèrent une énergie (égale à l'énergie du gap) qui peut être émise sous forme d'un photon. C'est le processus inverse de l'absorption vue ci-dessus, il s'appelle processus d'émission spontanée.

L'émission spontanée peut prendre un certain temps pour s'effectuer. La désexcitation de électrons se fait avec un temps caractéristique de chaque S-C appelé « temps de vie radiative » (cf. temps de vie de l'état excité en mécanique quantique). Néanmoins, un électron peut être désexcité de manière active par passage d'un photon d'énergie égale à l'énergie de la transition. Dans ce cas l'électron se désexcite en émettant un photon identique au photon incident. Il s'agit du processus d'émission stimulée. Vu la similitude qui existe entre les traitements mathématiques de l'émission stimulée et de l'absorption nous commençons ici par décrire l'émission stimulée dans les S-C. Nous verrons ensuite l'émission spontanée.

Tout comme nous l'avons fait pour l'absorption nous commençons avec un bref rappel de la notion d'émission stimulée dans les système atomique. L'atome dans son état excité est instable et à une durée de vie limitée qui peut cependant être non négligeable. Si un photon d'énergie égale à celle de la transition rencontre l'atome dans son état excité celui-ci se désexcitera en émettant un photon jumeau. En mécanique quantique, ce processus est caractérisé par une probabilité par unité de temps P_{st} . Cette probabilité est proportionnelle au coefficient d'Einstein (élément de matrice de la transition), à la densité de photons $n(\omega)$ et à la probabilité d'avoir un atome dans l'état excité f_e , et que son état fondamental soit libre $(1 - f_f)$.

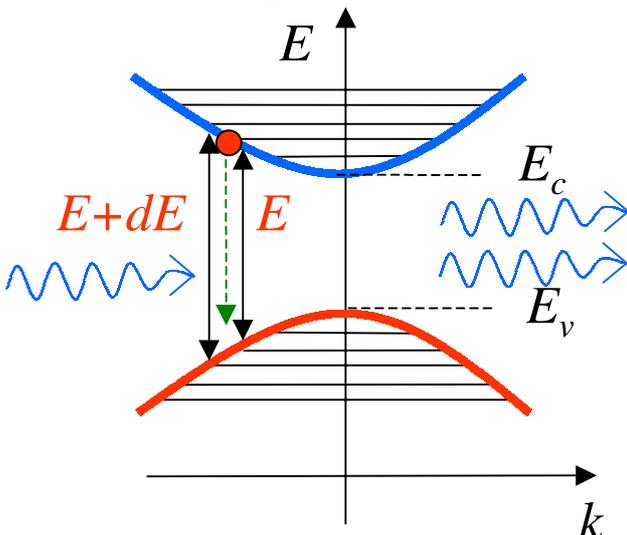


Probabilité de transition P_{st}

$$P_{st} = B \cdot n(\omega) \cdot f_e \cdot (1 - f_f) \quad [\text{s}^{-1}]$$

Pour étendre cette théorie aux S-C il suffit de refaire le raisonnement adopté pour le calcul du taux d'absorption d'un S-C à gap direct. On introduit donc le taux d'émission stimulée par unité d'énergie r_{st} en disant que $r_{st}dE$ est le nombre de transitions par unité de temps ayant lieu dans l'intervalle d'énergie dE . On fait alors l'hypothèse que B et les distributions de Fermi sont constants dans les bandes, le calcul se limite alors au compte du nombre de transitions dans dE c'est à dire le calcul de la densité de transition tout comme on l'a fait dans le calcul du taux d'absorption pour un S-C à gap direct.

S-C à Gap direct, $\Delta k = 0$



Taux d'émission stimulée par unité d'énergie r_{st}

$$r_{st} dE = \sum_{E < (E_e - E_f) < E + dE} P_{st}$$

$$r_{st} dE = \sum_{E < (E_e - E_f) < E + dE} B \rho(\omega) f_e (1 - f_f)$$

$$r_{st} = B \rho(\omega) f_c (1 - f_v) \rho_t(\omega)$$

Comme on l'a fait pour l'absorption, nous pouvons définir le taux d'émission stimulée pour un S-C éclairé par un faisceau monochromatique (c-à-d, $\rho(\omega) = n_\omega \delta(\omega - \omega')$), soit R_{st} .

Faisceau monochromatique : $R_{st} = B n_\omega f_c (1 - f_v) \rho_t(\omega)$

Une différence essentielle entre l'absorption et l'émission stimulée est que cette dernière n'est non négligeable que si la population d'électrons dans la bande de conduction f_c est significative et la population dans la bande de valence f_v est faible ce qui n'est pas une situation rencontrée naturellement, c'est à dire, lorsque le système est proche de l'équilibre thermique. Lorsque le système est proche de l'équilibre thermique, seule l'absorption semble être à considérer. Mais l'équilibre thermique est une limite abstraite pour un S-C soumis à un éclairage puisque chaque photon absorbé crée une paire électron-trou (et donc augmente f_c et diminue f_v) qui ne demande qu'à se recombiner. Le processus d'émission stimulée est donc indissociable de l'absorption.

Le taux de variation du nombre total de photons dans un matériau S-C illuminé est donc donné par $R_{st} - R_a$.



Variation du nombre de photons : $\frac{dn_{ph}}{dt} = R_{st}(\omega) - R_a(\omega)$

Comme mentionné ci dessus, près de l'équilibre thermique seule l'absorption joue un rôle.

- A l'équilibre thermique $f_c \approx 0$ et $f_v \approx 1$ \longrightarrow $R_{st} \approx 0$

Par contre dans une situation plus générale R_{st} doit être pris en compte et le taux de variation de photons devient :

- Hors équilibre : $\frac{dn_{ph}}{dt} = -B n_\omega (f_v - f_c) \rho_t(\omega)$

On peut dès lors redéfinir le coefficient d'absorption α en le généralisant aux systèmes hors de l'équilibre

Coefficient d'absorption

$$\alpha = \frac{1}{v} B (f_v - f_c) \rho_t(\omega)$$

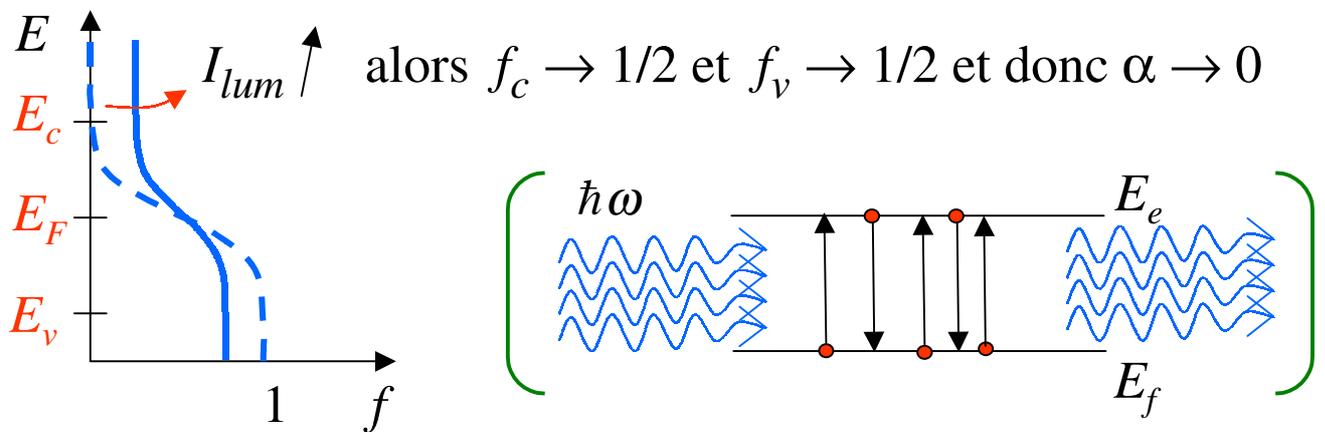
Nous voyons que lorsque les populations des bandes de conduction et de valence s'équilibrent, $f_c = f_v = 1/2$, l'absorption devient nulle. Notons que les propriétés spectrales de l'absorption sont indépendante du degré d'excitation et restent donc celles étudiées au chapitre précédent même dans les systèmes hors de l'équilibre (en particulier : spectre parabolique à seuil).

Absorption saturable :

Comme mentionné plus haut, un S-C est mis hors d'équilibre s'il est éclairé par un faisceau lumineux dont la fréquence est supérieure ou égale à la fréquence du seuil d'absorption (photons d'énergie supérieure ou égale au gap du S-C). En effet, si la lumière est absorbée, f_c augmente et f_v diminue il s'en suit donc une diminution de l'absorption elle-même. Si un grand nombre de photons sont incidents sur un S-C, beaucoup de transitions de la bande de valence vers la bande de conduction auront lieu. Il y aura donc beaucoup d'électrons dans la bande de conduction et de trous libérés dans la bande de valence c'est à dire que le taux d'émission stimulée sera très élevé au point de pouvoir compenser le taux d'absorption. Il y a compensation exacte lorsque l'intensité lumineuse est suffisamment grande que pour avoir égalité de populations dans les deux bandes.

Ceci est illustré de manière schématique ci dessous :

Principe: si la lumière est absorbée $f_c \uparrow$ et $f_v \downarrow$ et donc $\alpha \downarrow$

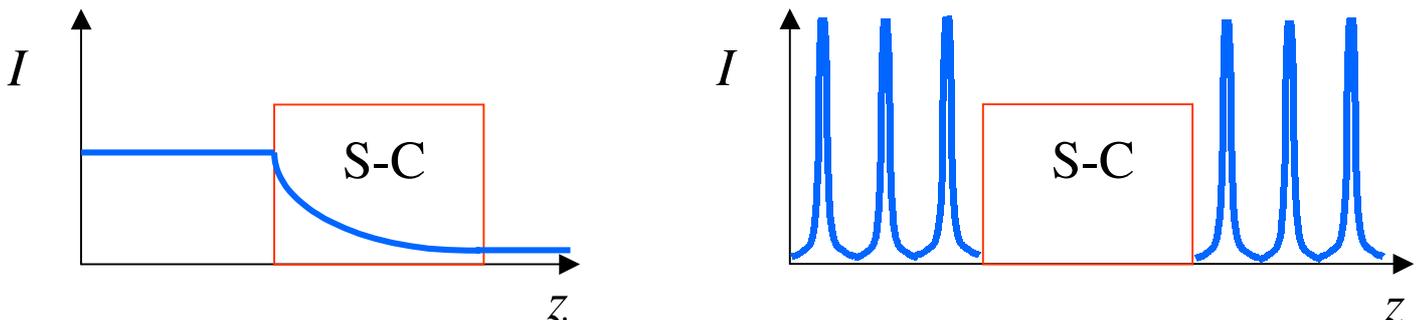


Ce phénomène de diminution de l'absorption par équilibrage entre l'absorption et l'émission stimulée s'appelle l'*absorption saturable*. Ce phénomène est général, il est caractéristique de l'absorption par transitions électroniques dans les systèmes quantiques. Il est souvent décrit mathématiquement à l'aide de la formule phénoménologique donnée ci-dessous où I_s est l'intensité de saturation (intensité pour laquelle le coefficient d'absorption est réduit de moitié). Cette formule est en fait exacte (fournie par la mécanique quantique) pour les systèmes atomiques lorsque l'on considère seulement la présence des deux états électroniques impliqués dans la transition.

Formulation phénoménologique :

$$\alpha(I) = \frac{\alpha_0}{1 + I/I_s}$$

Illustration :

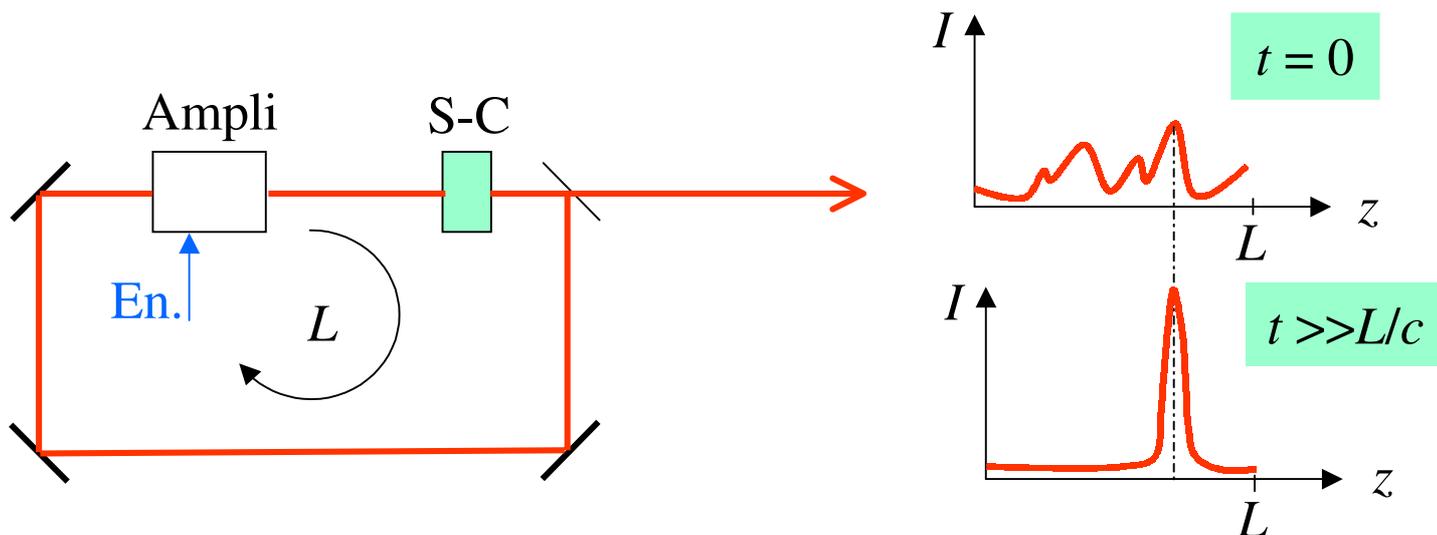


Nous montrons ci-dessus l'effet que peut avoir l'absorption saturable sur la transmission d'un faisceau laser dans un S-C : considérons deux faisceaux laser de même intensité moyenne, mais l'un est continu et l'autre pulsé. Comme le faisceau continu a une intensité moyenne plus basse (bien inférieure à I_s), il peut être absorbé alors que le faisceau pulsé est transmis puisque l'intensité des impulsions peut être beaucoup plus élevée que I_s .

Application : *génération d'impulsions courtes dans les lasers*

Nous illustrons ici brièvement l'intérêt que peut avoir le phénomène d'absorption saturable en pratique. Considérons un laser, c'est à dire, comme on le verra plus loin, une cavité optique contenant un amplificateur optique. L'amplificateur optique est un système quantique dans lequel on a réalisé une inversion de population des états électroniques (par exemple un S-C dans lequel $f_c > f_v$). Le système émet donc des photons spontanés. Ces photons spontanés sont alors amplifiés par émission stimulée et un faisceau lumineux intense est généré dans la cavité. Si rien n'est fait pour contrôler les variations temporelles du faisceau, celui-ci apparaîtra fortement bruité puisqu'il est né des photons spontanés qui par définition sont émis de façon aléatoire. Pour donner une structure temporelle cohérente au faisceau laser on place parfois un absorbant saturable dans la cavité du laser. Celui-ci possède une transmission d'autant plus élevée que l'intensité du faisceau est élevée. C'est donc à l'endroit où l'on trouve l'intensité la plus élevée dans le faisceau que l'on aura le moins de perte de photons. Autrement dit c'est au point où l'intensité est la plus élevée que le gain global du système (amplificateur + pertes de la cavité) est le plus grand. En raison de cette rétroaction positive, le pic initial d'intensité (voir le schéma ci-dessous pour $t=0$) se voit augmenter de manière importante au détriment de toutes les autres parties du faisceau. Il en résulte, en régime (c-à d, en $t \gg L/c$ où c est la vitesse de la lumière et L est la longueur de la cavité), qu'une seule impulsion lumineuse de forte intensité se forme dans la cavité du laser. Cette impulsion tourne dans la cavité et crée donc au miroir de sortie un signal pulsé périodique (et donc régulier et non plus aléatoire comme en l'absence d'absorbant saturable). Ce type de laser est appelé *laser à blocage de modes passif*. C'est une configuration très utilisée en pratique. Notez qu'en pratique pour obtenir des impulsions à l'échelle de la picoseconde, on n'utilise pas un semi-conducteur comme absorbant saturable (temps de réponse trop lent de l'ordre de la nanoseconde). On utilise plutôt des matériaux organiques de type *colorant*.

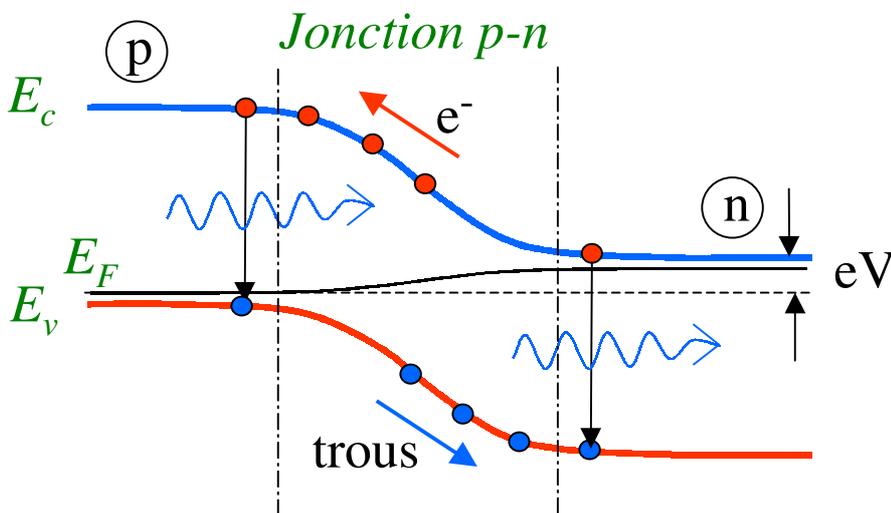
Laser à absorbant saturable



Gain Optique dans les Semi-conducteurs

Dans l'illustration qui précède nous avons introduit la notion d'amplificateur optique qui est à la base du fonctionnement du laser (à cette occasion je rappelle que le mot laser est l'acronyme de : light amplification by stimulated emission of radiation). L'amplification optique n'est rien d'autre qu'une absorption négative, c'est à dire un gain de photons plutôt qu'une perte. Pour réaliser l'amplification optique il faut donc un système pour lequel le coefficient d'absorption α soit négatif. Un simple coup d'oeil sur la formule donnant α dans les S-C permet de dire que l'amplification optique est possible dans les S-C. En effet, il suffit pour cela que l'on inverse l'équilibre naturel des électrons dans les bandes de conduction et de valence et que $f_c > f_v$ de sorte que $\alpha < 0$. Une telle inversion (appelée inversion de population) peut être facilement réalisée dans les S-C grâce à l'injection de porteurs minoritaires ayant lieu dans une jonction $p-n$ polarisée en sens passant. En effet si la polarisation de la jonction est élevée, le courant de porteurs minoritaires sera élevé et on peut espérer que du côté p suffisamment d'électrons soient injectés et suffisamment de trous retirés que pour réaliser la condition $f_c > f_v$ (même raisonnement pour le côté n).

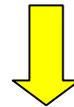
Absorption négative



Près de la région de charge d'espace :

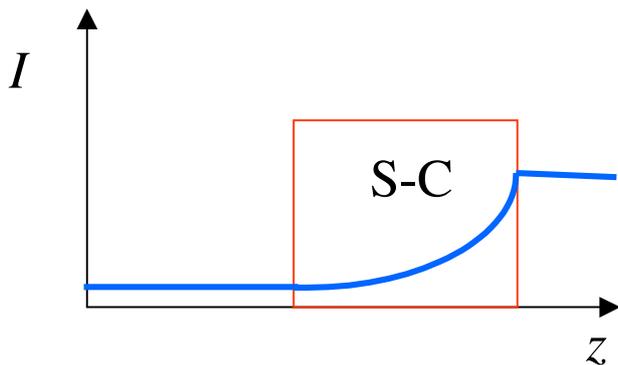
$$f_c > f_v$$

et donc $\alpha < 0$



Amplification

$$\frac{dn_\omega}{dz} \equiv -\alpha n_\omega$$



$$\frac{dI}{dz} = |\alpha| I$$



Nous donnons ci-dessus une représentation schématique de l'amplification optique qui correspond à une croissance exponentielle de l'intensité lumineuse dans le S-C avec inversion de population (notez que l'intensité lumineuse I est proportionnelle au nombre de photons n_ω).

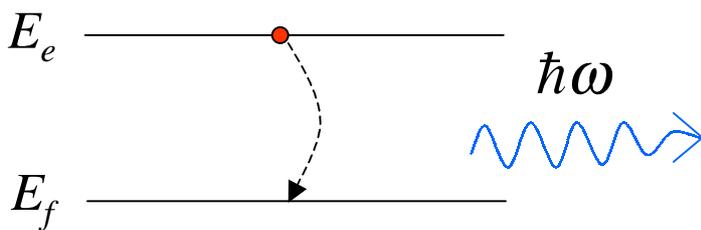
Dans le chapitre consacré aux lasers S-C nous verrons en détail comment réaliser une inversion de population dans une jonction $p-n$ afin d'obtenir un gain optique dans les S-C.

Émission spontanée dans les S-C

L'émission spontanée de photons dans les S-C joue un rôle important dans les dispositifs de génération de lumière (diodes électro-luminescentes et lasers). C'est la première cause d'émission de lumière dans les S-C. En effet, pour générer de la lumière par émission stimulée il faut que celle-ci soit supérieure à l'absorption. Il faut donc réaliser une inversion de population qui, comme nous le verrons, n'est pas facile à obtenir en pratique car elle correspond à une situation très éloignée de l'équilibre. Seule l'émission spontanée peut conduire à l'émission de lumière dans les systèmes peu excités tels que les diodes électro-luminescentes.

Après ce que nous avons déjà vu sur l'absorption et l'émission stimulée, la description de l'émission spontanée est très aisée. En effet, dans son interprétation quantique l'émission spontanée ressemble beaucoup à l'émission stimulée : la désexcitation d'un atome par émission spontanée d'un photon est vue comme une émission stimulée par les fluctuations quantiques du vide. En d'autres termes, on considère que dans l'espace vide il y a toujours un certain niveau de vibration du champ électromagnétique et que ces vibrations sont la cause de déstabilisation de l'état excité de l'atome. Dans le formalisme de la mécanique quantique les choses s'expriment comme si un photon virtuel stimulait la transition vers le bas. On retrouve ce photon virtuel dans l'expression de la probabilité de transition par émission spontanée de l'atome, P_{sp} où le nombre de photons n_ω a été remplacé par l'unité (la signification de B, f_c et f_v sont les mêmes que pour l'émission stimulée).

- *Systeme atomique à deux niveaux*



Probabilité d'émission P_{sp}

$$P_{sp} = B \cdot f_e \cdot (1 - f_v) \cdot [1]$$

Photon virtuel
(fluctuations du vide)

Le passage à la description de l'émission spontanée pour les S-C à gap direct est alors le même que pour l'absorption et l'émission stimulée. Pour le taux d'apparition de photons spontanés à une fréquence ω donnée l'on retrouve la même expression que pour R_{st} à la seule différence que maintenant la densité de photons n_ω n'apparaît pas (ce qui était attendu puisqu'il s'agit d'un processus spontané indépendant du nombre de photons environnants).

- *Semi-conducteurs*

Taux d'émission spontanée

(idem R_a et R_{st})

$$R_{sp} = B f_c (1 - f_v) \rho_t(\omega)$$

Taux de radiation total dans les S-C

Maintenant que nous avons passé en revue tous les mécanismes d'émission de photons dans les S-C, nous pouvons exprimer le taux de radiation total de photons à la fréquence ω , $R_{tot}(\omega)$ comme étant la somme de toutes les contributions: l'émission stimulée, l'absorption et l'émission spontanée. Nous appelons $R_{st} - R_a$ la contribution cohérente car elle représente des mécanismes qui ne modifient pas la cohérence d'un faisceau lumineux : les photons sont soit perdus (absorption) soit multipliés en conservant leur phase (les photons stimulés sont les photons jumeaux des photons incidents, y compris en ce qui concerne leur phase). Par opposition R_{sp} est la contribution incohérente puisque les photons générés spontanément n'ont aucun lien avec les photons environnants. En particulier, ils peuvent être émis avec n'importe quelle phase et peuvent donc créer des interférences constructives ou destructives avec les photons environnants. Les photons spontanés constituent donc une source de bruit.

$$R_{tot}(\omega) \equiv \frac{dn_{ph}}{dt} = \underbrace{R_{st}(\omega) - R_a(\omega)}_{\text{Contribution cohérente}} + \underbrace{R_{sp}(\omega)}_{\text{Contribution incohérente}}$$

$$\left\{ \begin{array}{l} R_a = Bn_{\omega}f_v(1-f_c)\rho_t(\omega) \\ R_{st} = Bn_{\omega}f_c(1-f_v)\rho_t(\omega) \\ R_{sp} = Bf_c(1-f_v)\rho_t(\omega) \end{array} \right.$$

Taux de radiation cohérente : $R_{coh} \equiv R_{st} - R_a$

Calcul du gain d'amplification

Le rôle néfaste de l'émission spontanée peut être illustré par le calcul du gain du nombre de photons dans un système amplificateur. La variation temporelle du nombre de photons est donnée par $R_{tot}(\omega)$. Comme indiqué ci-dessous, nous exprimons cette quantité en mettant la contribution cohérente en évidence et en introduisant le coefficient d'absorption α et le facteur $m = f_c(1-f_v)/(f_c - f_v)$.

$$R_{tot} = R_{coh} \left(1 + \frac{R_{sp}}{R_{coh}} \right) \quad \text{où} \quad R_{coh} = Bn_{\omega}(f_c - f_v)\rho_t(\omega) = -\alpha n_{\omega}v$$

$$\text{et} \quad \frac{R_{sp}}{R_{coh}} = \frac{f_c(1-f_v)}{n_{\omega}(f_c - f_v)} = \frac{m}{n_{\omega}}$$

$R_{tot} = -\alpha v (n_{\omega} + m)$

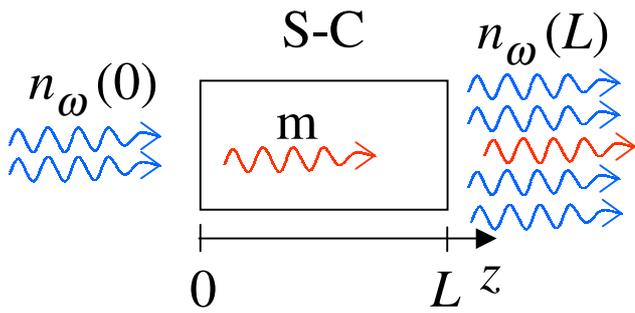
Le taux de croissance spatial du nombre de photons à la fréquence ω est donc donné par :

Amplification :

$$\frac{dn_{\omega}}{dz} = \frac{1}{v} \frac{dn_{\omega}}{dt} = \frac{R_{tot}}{v} = -\alpha(n_{\omega} + m)$$

Cette expression comporte deux contributions : la contribution cohérente $-\alpha n_{\omega}$ et la contribution incohérente $-\alpha m$.

L'intégration de l'équation en n_ω permet de calculer facilement le gain du nombre photons si l'on considère un nombre de photons $n_\omega(0)$ à l'entrée de l'amplificateur comme condition initiale. La contribution cohérente donne le terme $n_\omega(0)\exp(-\alpha L)$ ce qui permet de définir le gain du système $G = \exp(-\alpha L)$. La contribution incohérente à l'accroissement du nombre de photons conduit au terme $m[\exp(-\alpha L) - 1]$. Cette quantité représente le nombre de photons incohérents (de phase aléatoire, source de bruit) à la sortie du système. Si le gain est grand devant l'unité, $G \gg 1$ (ce qui est souvent le cas en pratique puisque lorsque l'on amplifie on désire toujours amplifier le plus possible), ce nombre peut être approché par mG . Cette expression particulièrement simple permet de dire que le niveau de bruit d'un amplificateur est équivalent à ce que donnerait l'amplification de m photons incohérents par un amplificateur idéal sans bruit de gain G . Nous verrons plus loin quelle est la valeur de m rencontrée en pratique.



$$\frac{dn_\omega}{n_\omega + m} = -\alpha dz$$

$$\ln\left(\frac{n_\omega(L) + m}{n_\omega(0) + m}\right) = -\alpha L$$

$$n_\omega(L) = \underbrace{n_\omega(0) e^{-\alpha L}}_{\text{Contribution cohérente}} + \underbrace{m (e^{-\alpha L} - 1)}_{\text{Contribution incohérente} \equiv \text{bruit} \approx mG \text{ (si } G \gg 1)}$$

Contribution cohérente

Contribution incohérente \equiv bruit $\approx mG$ (si $G \gg 1$)

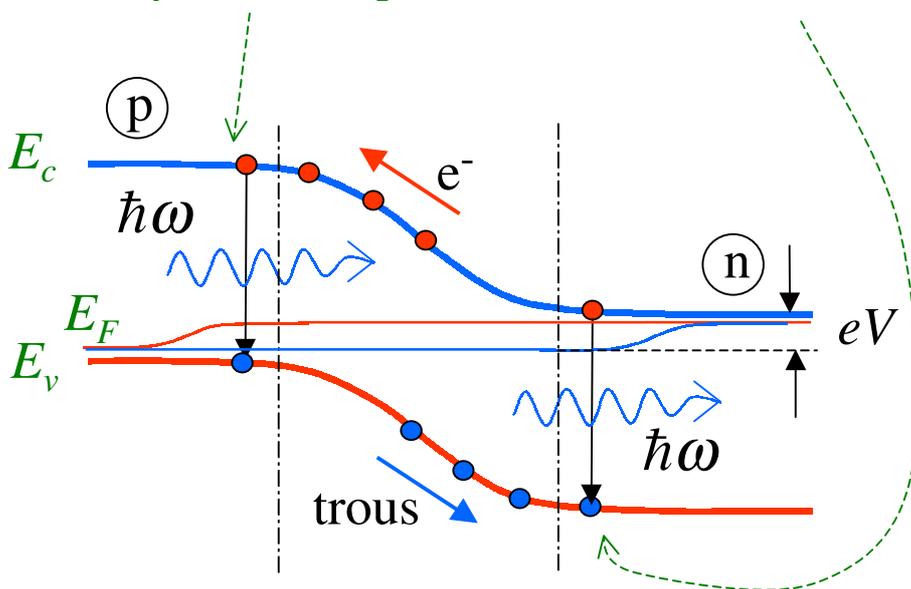


« m photons incohérents à l'entrée »

Inversion de population dans une jonction p-n

Voyons maintenant en détail comment réaliser l'inversion de population dans une jonction $p-n$. Lorsque la jonction est polarisée en sens passant, les porteurs minoritaires sont injectés avec une certaine énergie cinétique. Cette énergie est dissipée par collisions électrons-électrons ou électrons-phonons (interaction avec le réseau cristallin) jusqu'à thermalisation (c-à-d, la vitesse des électrons atteint celle qu'elle serait à équilibre thermique : distribution des vitesses de moyenne nulle et d'écart-type fixé par la température). Cette thermalisation s'effectue en des temps extrêmement courts de l'ordre du dixième de picoseconde. Ces temps sont beaucoup plus courts que les temps de vie radiative habituellement rencontrés dans les S-C (temps de vie radiative : temps moyen mis par les paires électron-trou pour se recombiner par émission spontanée de photon). Ceci signifie que les électrons du côté p et les trous du côté n sont accumulés et l'on peut donc imaginer réaliser de cette manière l'inversion de population.

Injection de porteurs minoritaires et thermalisation



Collisions :

$$\left\{ \begin{array}{l} e^- - e^- \\ \text{phonons} - e^- \end{array} \right.$$

Temps caractéristique :

$$\tau_{th} \approx 10^{-13} \text{ sec}$$

Durée de vie radiative

$$\tau_r \approx 10^{-9} \text{ sec}$$

Quasi-niveaux de Fermi

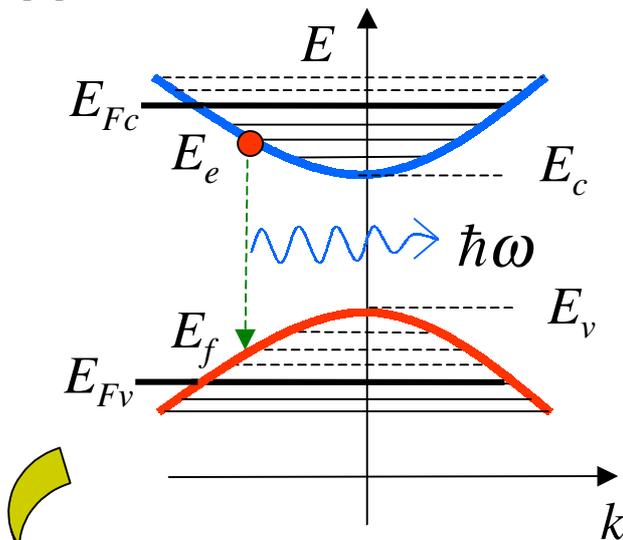
Comme les porteurs minoritaires ont le temps de se thermaliser on peut les considérer comme étant à l'équilibre thermique. Ce qui est une vue de l'esprit puisque à l'équilibre thermique les électrons du côté p sont bien entendu pour la plupart dans la bande de valence (idem pour les trous du côté n). Néanmoins à l'échelle de temps des temps de vie radiative on peut parler d'un quasi-équilibre thermique. Ceci nous permet d'introduire la notion de quasi-niveau de Fermi. Le quasi-niveau de Fermi des électrons du côté p est le niveau d'énergie correspondant aux dernières couches (états électroniques) occupées par les électrons injectés. Les électrons injectés sont en effet empilés selon le principe de Pauli sur les états quantiques du cristal de la même manière que dans un S-C à l'équilibre thermique et ils remplissent donc la bande de conduction jusqu'à un certain niveau, c'est ce niveau que l'on appelle quasi-niveau de Fermi. Le même raisonnement peut bien entendu être fait pour les trous du côté n .

Dans une jonction polarisée l'énergie du quasi-niveau de Fermi peut être évaluée en se rappelant que l'énergie de Fermi représente le potentiel électrochimique des électrons (potentiel tenant compte de l'affinité chimique pour les dopants et du potentiel électrostatique). Donc si l'on néglige la perte d'énergie cinétique des électrons injectés on peut conclure que le quasi-niveau de Fermi est donné par la différence de potentiel appliquée aux bornes de la jonction, comme indiqué sur le schéma ci-dessus : le niveau de Fermi se divise en deux niveaux distincts pour les électrons et les trous injectés au niveau de la jonction, ce sont les quasi-niveaux de Fermi des électrons et des trous minoritaires en quasi-équilibre.

Le schéma de bande ci-dessous, représente les électrons thermalisés du côté p . Comme il s'agit d'un S-C dopé de type p , le niveau de Fermi E_F est normalement proche du sommet de la bande de valence. Mais les électrons libres injectés s'empilent sur les états de la bande de conduction jusqu'au quasi-niveau de Fermi E_{Fc} (c pour « bande de conduction ») déterminé par la différence de potentiel appliquée. La bande inférieure représente la même situation pour les trous du côté n : les trous thermalisés sont accumulés au sommet de la bande de valence. La région de la jonction $p-n$ où cette situation est rencontrée est en réalité très limitée. En effet, si l'on se place du côté p loin de la zone de charge d'espace, on aura les électrons thermalisés de la bande de conduction qui seront face à une bande de valence pleine d'électrons c'est à dire sans trous pour la recombinaison radiative des paires électron-trou nécessaire à l'émission stimulée. Il en est de même pour les trous du côté n : ceux-ci ne peuvent se recombiner facilement puisqu'ils ont face à eux une bande de conduction contenant peu d'électrons. Dans ces conditions on ne peut pas avoir beaucoup d'émission stimulée, c'est à dire qu'on ne peut pas avoir de gain positif.

En fait, pour avoir une inversion de population il faut réaliser la situation représentée par le schéma de bande ci-dessous. Une telle situation peut se rencontrer dans la zone de charge d'espace, là où les flux d'électrons et de trous se croisent. Dans ce cas les électrons et trou thermalisés se retrouvent et peuvent se recombiner en masse pour donner lieu à une émission stimulée importante et donc un gain positif.

Nous avons vu plus haut que pour avoir un gain positif il faut satisfaire la condition $f_c > f_v$. En écrivant la distribution de Fermi avec les quasi-niveaux de Fermi E_{Fc} , E_{Fv} cette condition se traduit par l'inégalité $E_{Fc} - E_{Fv} > \hbar\omega$ ($\hbar\omega$ étant l'énergie de la transition soit $E_e - E_f$). En se rappelant que la différence des niveaux de Fermi $E_{Fc} - E_{Fv}$ est à peu près donnée par le potentiel électrique appliqué soit eV , la condition pour avoir inversion de population est $eV > \hbar\omega$.

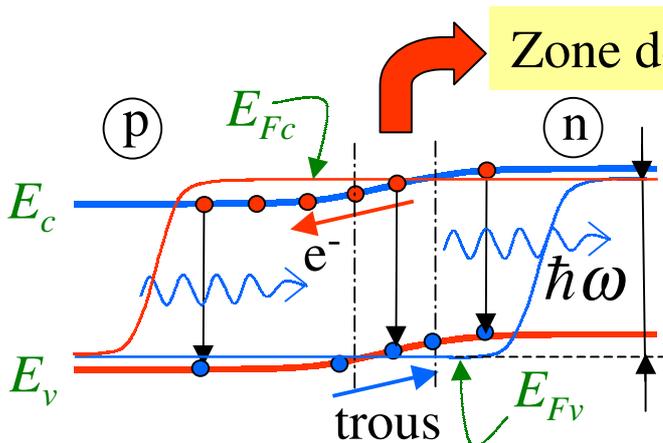


Gain ($\alpha < 0$) si $f_c > f_v$

$$\left\{ \begin{aligned} f_c &= \frac{1}{e^{(E_e - E_{Fc})/k_B T} + 1} \\ f_v &= \frac{1}{e^{(E_f - E_{Fv})/k_B T} + 1} \end{aligned} \right.$$

$$E_e - E_{Fc} < E_f - E_{Fv} \implies E_e - E_f = \hbar\omega < E_{Fc} - E_{Fv}$$

Sachant que les électrons thermalisés se trouvent au creux de la bande de conduction et que les trous thermalisés sont au sommet de la bande de valence, on peut dire que $\hbar\omega$ est à peu près égal à l'énergie du gap E_g , la condition d'inversion de population devient donc $eV > E_g$.



$$eV > \hbar\omega = E_g$$

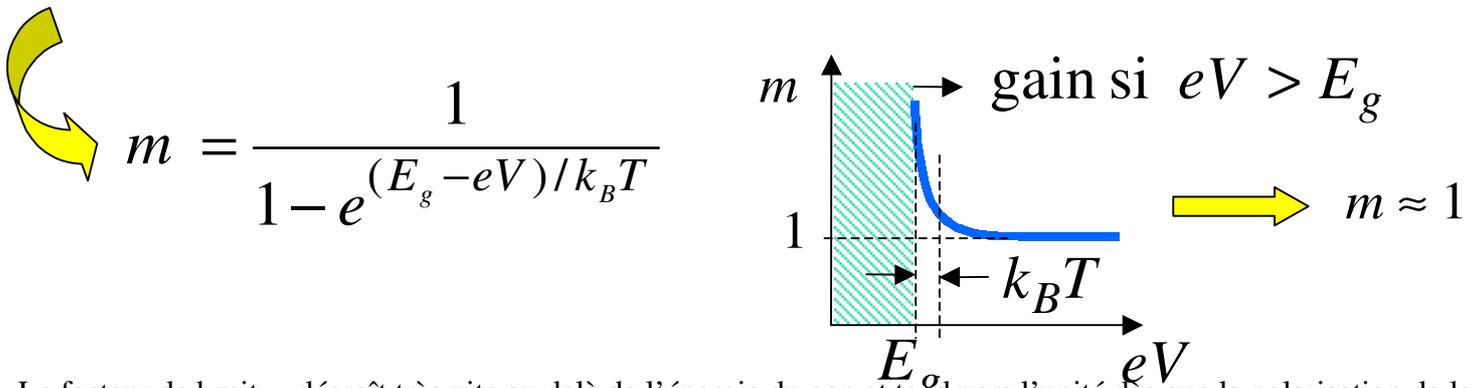
eV La figure ci-contre indique la zone où la condition d'inversion de population est effectivement rencontrée. On constate bien que pour avoir une telle situation la ddp appliquée doit être supérieure au gap (d'où l'inversion des courbures des bande de la jonction)

Calcul du bruit d'amplification

Les notions d'inversion de population et de quasi-niveaux de Fermi nous permettent de caractériser plus en détail le gain optique dans les S-C. Considérons tout d'abord le bruit d'amplification des photons spontanés. Nous avons vu que ce bruit d'un amplificateur réel de gain G pouvait s'interpréter comme étant le bruit causé par m photons à l'entrée d'un amplificateur idéal de même gain. Il est donc important de connaître la valeur de m pour un amplificateur réel.

Si l'on exprime la valeur de m en fonction des distributions de Fermi dans lesquelles on remplace l'énergie de la transition par $E_e - E_f$ par E_g et la différence des quasi-niveaux de Fermi $E_{Fc} - E_{Fv}$ par eV , nous obtenons l'expression donnée ci-dessous :

$$m \equiv \frac{f_c(1-f_v)}{(f_c-f_v)} = \frac{1/f_v - 1}{1/f_v - 1/f_c} = \frac{e^{(E_f - E_{Fv})/k_B T}}{e^{(E_f - E_{Fv})/k_B T} - e^{(E_e - E_{Fc})/k_B T}}$$



Le facteur de bruit m décroît très vite au delà de l'énergie du gap et tend vers l'unité dès que la polarisation de la jonction est plus grande que le gap d'une quantité supérieure à l'énergie thermique (typiquement un centième de l'énergie du gap). On peut donc dire en pratique que m vaut l'unité et l'expression du nombre de photons à la sortie d'un amplificateur de gain G est donnée par la formule ci-dessous :

$$n_\omega(L) = \underbrace{n_\omega(0) G}_{\text{Contribution cohérente}} + \underbrace{G - 1}_{\text{Contribution incohérente} \equiv \text{bruit} \approx G \text{ (si } G \gg 1)}$$

« 1 photon incohérent à l'entrée »

Le bruit d'un amplificateur peut donc être caractérisé en disant qu'il est équivalent à ce que l'on obtiendrait si l'on plaçait **un** photon incohérent (par rapport au signal à amplifier) à l'entrée de l'amplificateur idéal sans bruit. Cette caractérisation du bruit d'un amplificateur est très générale et ne s'applique pas seulement aux amplificateurs à S-C.

Distribution spectrale du gain

Les notions de quasi-niveaux de Fermi nous permet également d'étudier facilement le spectre de gain d'un amplificateur S-C. Nous avons vu que le gain est l'inverse du coefficient d'absorption soit $-\alpha$. Or nous avons vu plus haut que le spectre d'absorption est donné par la densité de transition $\rho_t(\omega)$. Ce résultat n'est bien entendu pas applicable au gain puisque le gain s'obtient en dehors de l'équilibre, c'est à dire que les distributions de Fermi ne peuvent plus être considérées comme constantes. Celles-ci varient significativement au sein d'une bande puisque les quasi-niveaux de Fermi se trouvent au beau milieu de ces bandes. Le calcul du spectre de gain est donc, a priori, beaucoup plus compliqué que celui de l'absorption.

Pour le simplifier nous allons considérer une jonction $p-n$ à semi-conducteur fortement dopés. Lorsque le dopage est fort (10^{18} cm^{-3} ou plus) le S-C ne se comporte plus du tout comme un semi-conducteur intrinsèque en raison d'une perte significative de la périodicité du potentiel vu par les électrons. Il en résulte que le raisonnement fait sur la conservation de la quantité de mouvement pour arriver à la règle de sélection dans les S-C à gap direct ne s'applique plus. En pratique il n'y a pas de règle de sélection et les transitions peuvent avoir lieu entre tous les états des bandes de conduction et de valence. Nous allons donc pouvoir adopter le raisonnement fait pour l'absorption dans les S-C à gap indirect et dire simplement que le taux de transition est proportionnel au produit des densités d'états des deux bandes. La différence avec le calcul de l'absorption des S-C à gap indirect est que l'on doit maintenant tenir compte de la variation des distribution de Fermi au sein des bandes. Si l'on considère la transition entre deux niveaux discrets donnés (cf. transition atomique), le coefficient d'absorption sera proportionnel au coefficient d'Einstein B et la différence des valeurs de la distribution de Fermi au deux niveaux de la transition soit $f_v - f_c$ dans les notations adoptées pour les S-C. Pour deux états au sein du continuum d'états de densités ρ_c et ρ_v , on aura donc un coefficient d'absorption proportionnel à $\rho_c \rho_v (f_v - f_c)$. Il s'agit donc simplement d'intégrer cette quantité sur tous les états des bandes. Pour cela nous introduisons la profondeur ε de l'état excité dans la bande de conduction comme on l'a fait pour le calcul de α pour les S-C à gap indirect.

S-C idéal

$$\alpha = \frac{1}{v} B (f_v - f_c) \rho_t(\omega)$$

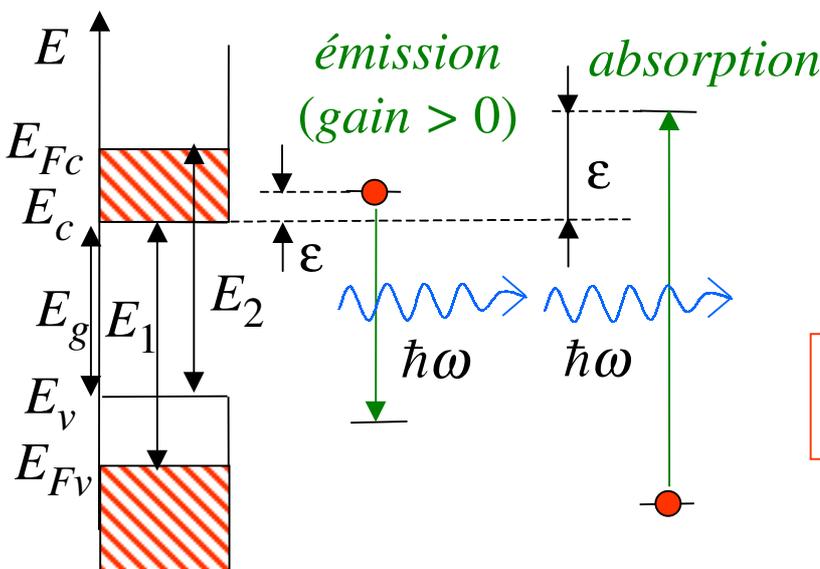
Hypothèse : S-C fortement dopé

pas de règle de transition

$$\alpha \div \int \rho_c(\varepsilon) \rho_v(\varepsilon) (f_v - f_c) d\varepsilon$$

Nous simplifions alors encore le calcul en approchant les densités par des distributions carrées comme indiqué sur le schéma ci-dessous où nous définissons également les quantités E_1 et E_2 utiles pour la simplicité des notations. Les densités carrées nous permettent de les remplacer par l'introduction des bornes d'intégration.

Simplification : densités $\rho_c, \rho_v =$ distributions carrées



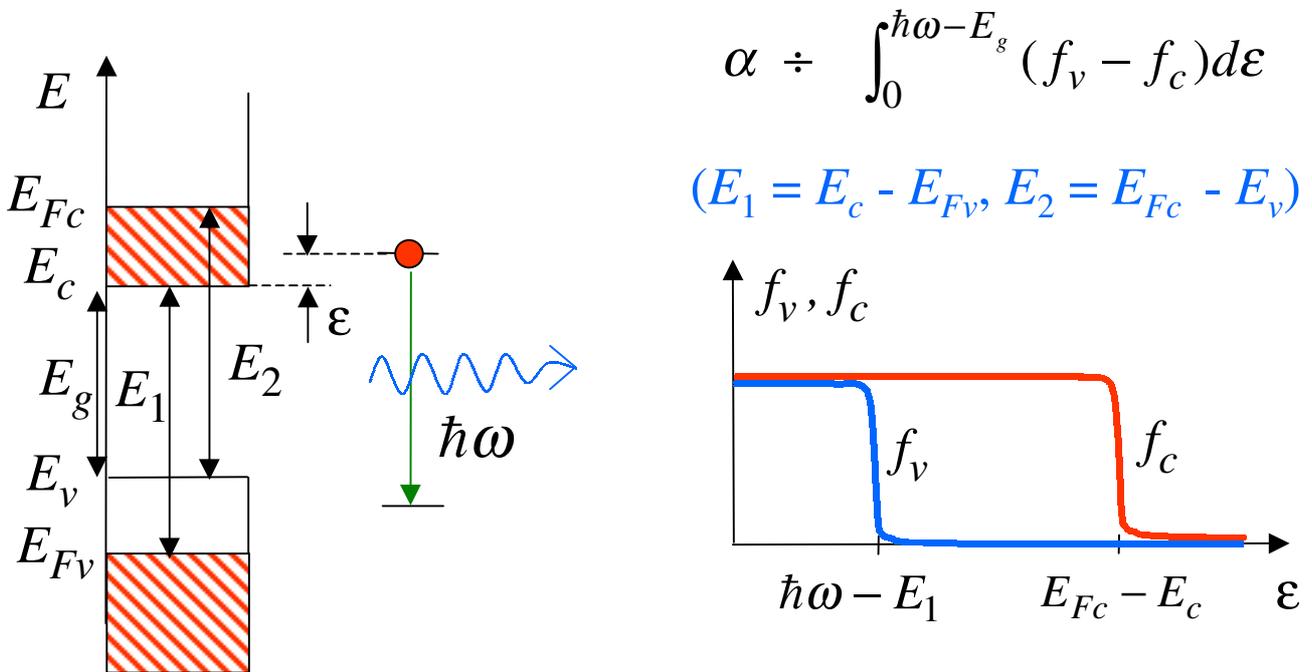
$$\left\{ \begin{array}{l} E_1 = E_c - E_{Fv} \\ E_2 = E_{Fc} - E_v \end{array} \right.$$

$$\alpha \div \int_0^{\hbar\omega - E_g} (f_v - f_c) d\varepsilon$$

Avant de faire le calcul, précisons tout d'abord la signification physique du spectre de gain. Le spectre de gain donne le gain d'un S-C à population inversée en fonction de la fréquence de la lumière (énergie des photons) à amplifier. Si la fréquence est légèrement supérieure au gap on peut s'attendre à ce que les photons puissent induire des transitions vers le bas et le système présentera donc de l'émission stimulée et un gain positif (c'est la situation schématisée à la figure ci-dessus sous le nom « *émission (gain > 0)* »). Par contre si la fréquence est beaucoup plus grande que le gap elle ne pourrait qu'induire des transitions vers le haut et le système apparaît alors absorbant (c'est la situation schématisée à la figure ci-dessus sous le nom « *absorption* »). On voit donc que le gain peut passer d'une valeur positive à une valeur négative en augmentant la fréquence de la lumière. Un tel comportement est propre aux S-C et montre l'importance de la caractérisation spectrale du gain dans les S-C.

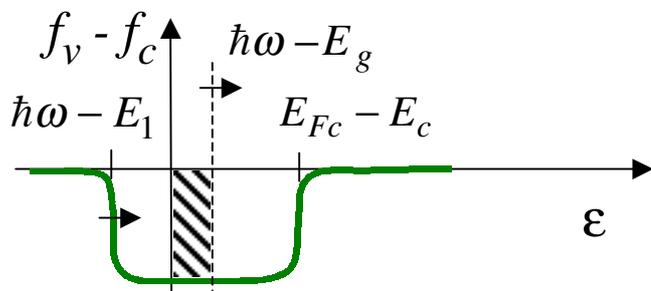
Calcul du spectre de gain

Sur base des simplifications introduites plus haut, le calcul du spectre de gain est particulièrement simple. Il suffit de traduire les distributions de Fermi en terme de ε : les quasi-niveaux de Fermi se trouvent en $\hbar\omega - E_1$ pour la bande de valence et en $E_{Fc} - E_c$ pour la bande de conduction.

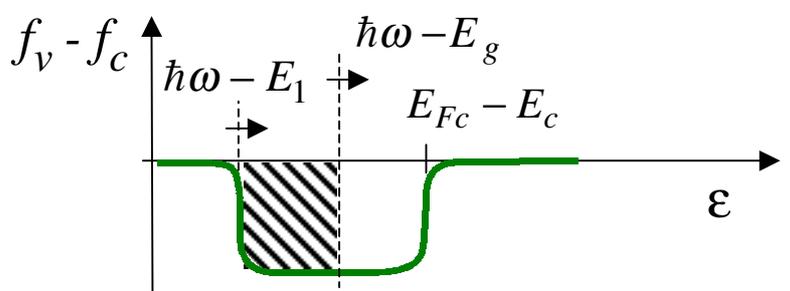


Lorsque $\hbar\omega < E_g$ il n'y a ni absorption ni gain puisque les photons ne peuvent pas relier deux états des bandes de valence et de conduction. A la figure de gauche ci-dessous nous représentons la situation rencontrée lorsque $E_g < \hbar\omega < E_1$. L'intégrale de $f_v - f_c$ étant prise de 0 à $\hbar\omega - E_g$ nous obtenons une croissance négative linéaire du coefficient d'absorption (le gain augmente linéairement en fonction de la fréquence). Une fois que $\hbar\omega > E_1$ l'intégrale devient constante puisque le quasi niveau de Fermi $\hbar\omega - E_1$ devient positif. On a donc apparition d'une région dans le spectre où le gain est constant. Ceci jusqu'à ce que $\hbar\omega - E_g$ atteigne le quasi niveau de Fermi de la bande de conduction $E_{Fc} - E_c$ soit $\hbar\omega = E_2$.

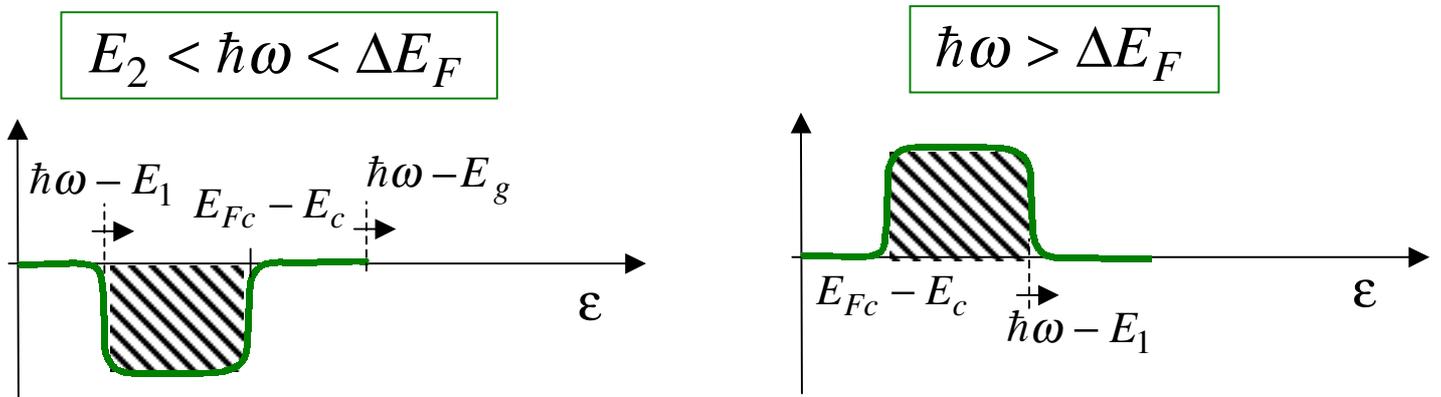
$$E_g < \hbar\omega < E_1$$



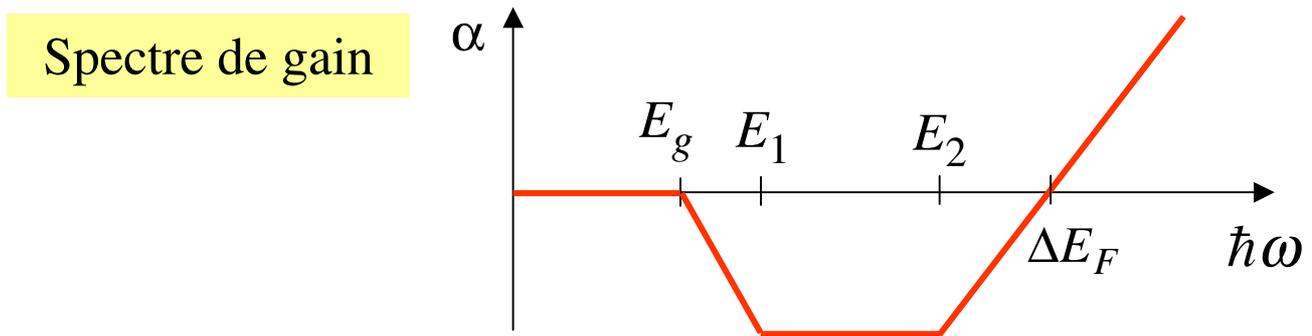
$$E_1 < \hbar\omega < E_g + E_{Fc} - E_c = E_2$$



Une fois la limite E_2 dépassée, la surface intégrée diminue avec une augmentation de fréquence. Le gain diminue donc linéairement jusqu'à son annulation en $\hbar\omega = E_{Fc} - E_{Fv} = \Delta E_F$. En effet, une fois que $\hbar\omega > \Delta E_F$ seules les transitions du bas vers le haut deviennent possibles et le système devient absorbant. L'absorption croît alors de manière linéaire avec la fréquence.



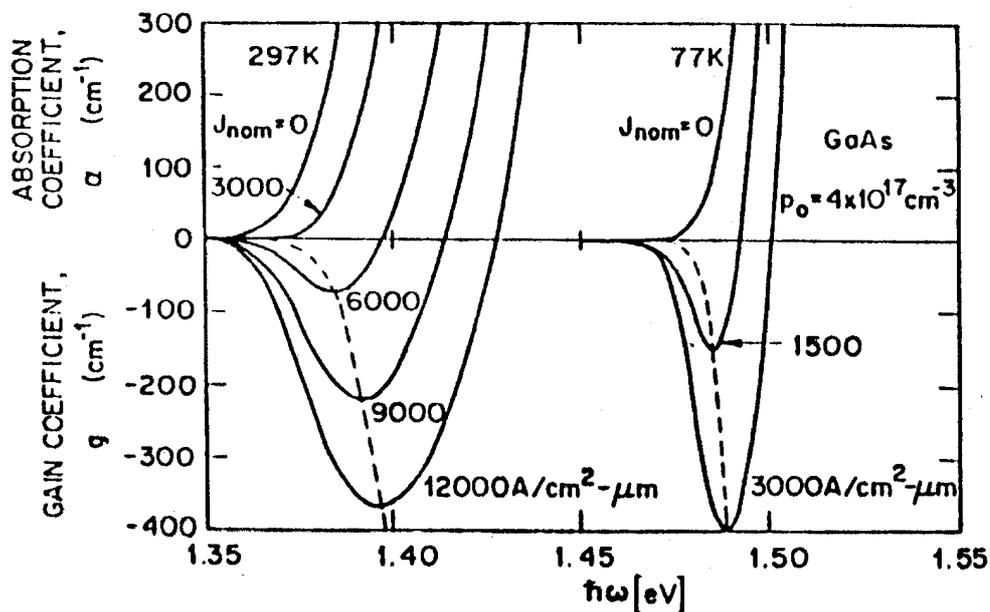
Ce calcul simplifié donne une bonne idée du spectre de gain d'un S-C. Celui-ci devient non-nul au gap et présente un maximum entre le gap et la différence des quasi-niveaux de Fermi ΔE_F . Au delà de ΔE_F il devient négatif.



Ces prédictions qualitatives sont bien vérifiées en pratique. Puisque les densités d'états et les distributions de Fermi ne présentent en réalité pas de sauts abrupts, les spectres réels sont bien entendu plus lisses.

Courbes de gain réelles

Nous donnons ci-dessous des exemples expérimentaux de spectres de gain à deux températures différentes.



Absorption/gain pour différents niveaux d'excitation mesurés en densité de courant traversant une zone active de 1 μm de large

Calcul des quasi-niveaux de Fermi

Nous faisons ici une petite parenthèse concernant le calcul des quasi-niveaux de Fermi. En pratique ces niveaux peuvent être calculés rigoureusement à partir de la mesure (ou du calcul) de la densité de porteurs minoritaires injectés. En effet, la densité d'électrons n_e est donnée par l'intégrale de la densité d'états multipliée par la distribution de Fermi (voir rappel). Connaissant n_e est ρ_e et la forme analytique de la distribution de Fermi on peut facilement remonter à la valeur de E_{Fc} .

Densité de porteurs :

$$n_e = 2 \int \rho_e(E) f(E) dE$$

$$\rho_e = \frac{1}{2\pi^2} \left(\frac{2m_c}{\hbar^2} \right)^{3/2} \sqrt{E - E_c}, \quad f(E) = \frac{1}{\exp[(E - E_{Fc})/k_B T] + 1}$$

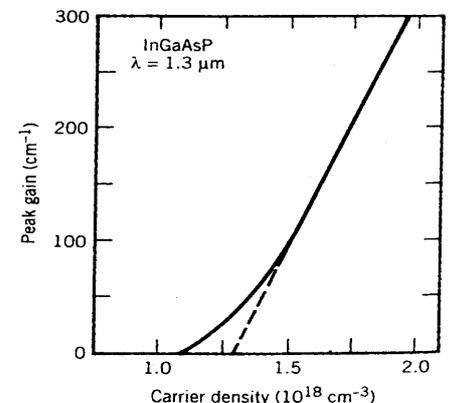
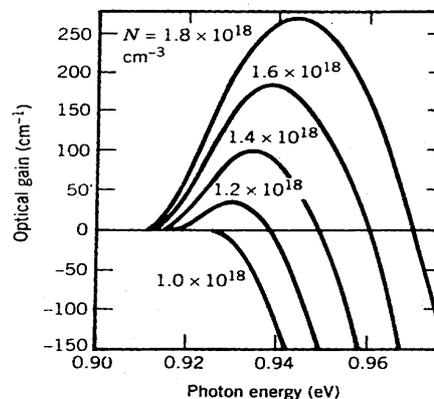


E_{Fc} obtenu à partir du calcul de n_e selon les lois du transport de charge dans les S-C.

(diffusion, mobilité, temps de recombinaison, etc...)

Enfin pour terminer ce chapitre sur le gain optique dans les S-C, il est intéressant pour la suite de voir comment varie le maximum du spectre de gain en fonction du degré d'excitation du S-C. Le degré d'excitation dans le cas de la jonction $p-n$ n'est rien d'autre que le courant d'injection du à la polarisation en sens passant de la jonction. Le courant d'injection est directement proportionnel à la densité de porteur. On voit ci-dessous à gauche différentes courbes de gain expérimentales en fonction de la densité de porteurs injectée. A droite est représenté le maximum du gain en fonction de la densité de porteurs. On constate que le gain maximum augmente presque linéairement après passage par un seuil. Comme nous le verrons, cette caractéristique est importante pour la modélisation des lasers à S-C.

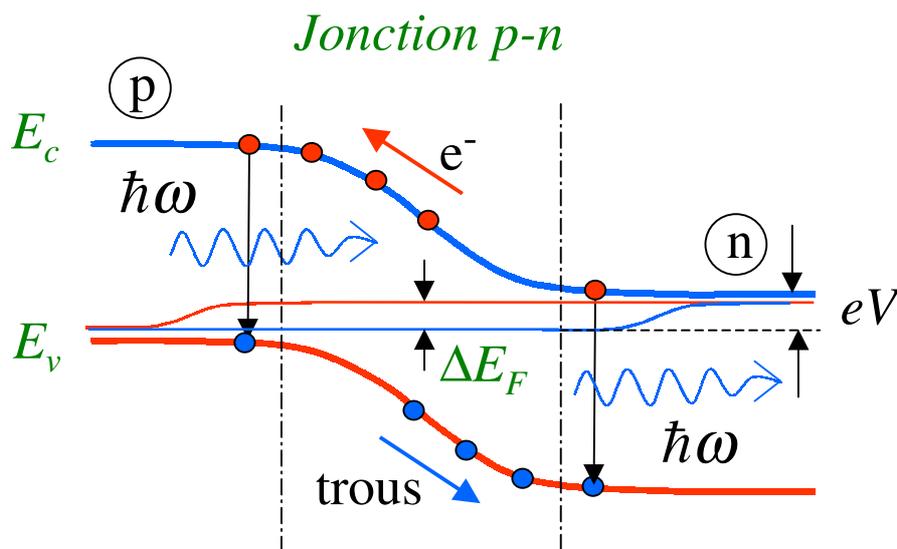
Gain optimal en fonction du degré d'excitation



La luminescence des semi-conducteurs

Jusqu'à présent nous avons vu que les mécanismes de base responsables de la génération de photons dans les semi-conducteurs sont l'émission spontanée et l'émission stimulée. Avec la notion de gain optique nous avons surtout insisté sur l'émission stimulée en considérant l'émission spontanée comme une simple source de bruit. Comme nous l'avons vu, l'émission stimulée exige des degrés d'excitation élevés induits par des polarisations de jonction $p-n$ supérieure à l'énergie du gap. De tels degrés d'excitation ne sont en réalité pas faciles à réaliser en pratique et ce n'est que grâce à des progrès significatifs dans la maîtrise des matériaux S-C que l'émission stimulée a pu être mise à profit (nous verrons cela plus en détail plus loin). Historiquement, c'est l'émission spontanée qui a été exploitée en premier lieu avec l'avènement des diodes électro-luminescentes. Les diodes électro-luminescentes (DEL ou LED en anglais). Les DEL fonctionnent sur le même principe que les amplificateurs : elles sont basées sur l'injection de porteurs minoritaires dans une jonction $p-n$ polarisée en sens passant. La différence essentielle avec les amplificateurs est que la jonction est faiblement polarisée et que les conditions de fonctionnement sont très éloignées des conditions d'inversion de population. Le rayonnement qu'elles émettent est donc essentiellement de nature spontanée.

Semi-conducteur hors de l'équilibre thermique : $\Delta E_F > 0$



Emission de photons non négligeable si :
temps de vie radiative < temps de vie non-radiative

Pour que les porteurs minoritaires injectés donnent lieu à de la génération de lumière il faut que leur temps de vie radiative soit plus court que leur temps de vie non-radiative. Le temps de vie non-radiative correspond à des mécanismes de désexcitation autres que la transition radiative. Des transitions de bande à bande peuvent survenir par exemple à partir de collisions avec le réseau cristallin (interactions photons-phonons), ou à partir de cascades de transitions de faible énergie rendues possibles par la présence de niveaux électroniques profonds dans la bande interdite (l'électron passe alors de niveau en niveau et aucun photon n'est émis du moins dans les longueurs d'onde optiques). Les niveaux électroniques profonds peuvent être dus à des impuretés chimiques, des défauts dans le réseau cristallin (dislocations) ainsi que les surfaces libres et interfaces avec d'autres matériaux. On voit donc que la qualité du matériau S-C est un paramètre essentiel pour les applications à la génération de lumière.

Caractérisation de la luminescence dans les S-C

Pour caractériser la luminescence d'un S-C il nous suffit de considérer le taux total de génération de photons dans les S-C R_{tot} , comme nous l'avons fait pour le calcul de l'amplification. Gardons néanmoins à l'esprit que cette fois le degré d'excitation est faible. En adoptant la même procédure que pour l'amplification nous arrivons à une expression de R_{tot} en fonction du facteur m .

$$R_{tot} = R_{st} - R_a + R_{sp}$$

$$R_{tot} = R_{coh} \left(1 + \frac{R_{sp}}{R_{coh}} \right)$$

$$R_{tot} = -\alpha v (m + n_\omega)$$

$$R_{coh} = Bn_\omega (f_c - f_v) \rho_t(\omega) = -\alpha n_\omega v$$

$$\frac{R_{sp}}{R_{coh}} = \frac{f_c (1 - f_v)}{n_\omega (f_c - f_v)} = \frac{m}{n_\omega}$$

$$\begin{cases} R_a = Bn_\omega f_v (1 - f_c) \rho_t(\omega) \\ R_{st} = Bn_\omega f_c (1 - f_v) \rho_t(\omega) \\ R_{sp} = Bf_c (1 - f_v) \rho_t(\omega) \end{cases}$$

où $m = \frac{1}{1 - e^{(E_g - \Delta E_F)/k_B T}}$

$$R_{tot} = \alpha v \left(\frac{1}{e^{(E_g - \Delta E_F)/k_B T} - 1} - n_\omega \right)$$

Cette expression permet d'étudier immédiatement les caractéristiques spectrales de la luminescence dans les S-C. Il suffit en effet de faire apparaître la fréquence ω dans les grandeurs où elle intervient telles que, en particulier, le coefficient d'absorption α .

Spectre de luminescence dans les S-C

$$R_{tot}(\hbar\omega) = \alpha(\hbar\omega) v \left(\frac{1}{e^{(\hbar\omega - \Delta E_F)/k_B T} - 1} - n_\omega \right)$$

Remarquons tout d'abord que si le S-C est à l'équilibre, c-à-d, $\Delta E_F = 0$, alors, par définition, n_ω est donné par la distribution de Bose-Einstein (voir début du cours) et dès lors le taux de génération de photons est nul. Ce qui est rassurant puisqu'on ne s'attend pas à ce qu'un S-C émette de la lumière en l'absence d'excitation.

Soit l'équilibre thermique : $\Delta E_F = 0$ et $n_\omega = \frac{1}{e^{\hbar\omega/k_B T} - 1}$

$$R_{tot}(\hbar\omega) = 0$$

Considérons maintenant un S-C faiblement excité. C'est à dire un S-C pour lequel ΔE_F est beaucoup plus faible que le gap (loin de la condition d'inversion de population). Dans ces conditions on peut supposer que la distribution spectrale de photons est toujours proche de celle à l'équilibre et on l'assimilera à la distribution de Bose-Einstein. De plus, comme l'énergie du gap dans les S-C est supérieure de deux ordres de grandeur à l'énergie thermique (à température ambiante, ~ 300 K) on peut négliger l'unité par rapport au facteur $\exp(\hbar\omega/k_B T)$. Nous obtenons de la sorte une expression particulièrement simple du taux de radiation par luminescence dans les S-C.

Soit un S-C faiblement excité : $\Delta E_F \ll \hbar\omega$

n_ω proche de l'équilibre

$$e^{(\hbar\omega - \Delta E_F)/k_B T} \gg 1 \quad \text{et} \quad e^{\hbar\omega/k_B T} \gg 1$$

$$R_{tot}(\hbar\omega) = \alpha(\hbar\omega) v \left(e^{-\hbar\omega/k_B T} e^{\Delta E_F/k_B T} - e^{-\hbar\omega/k_B T} \right)$$

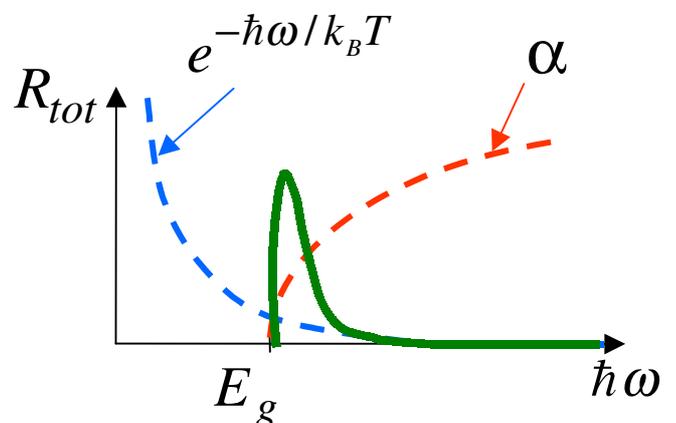
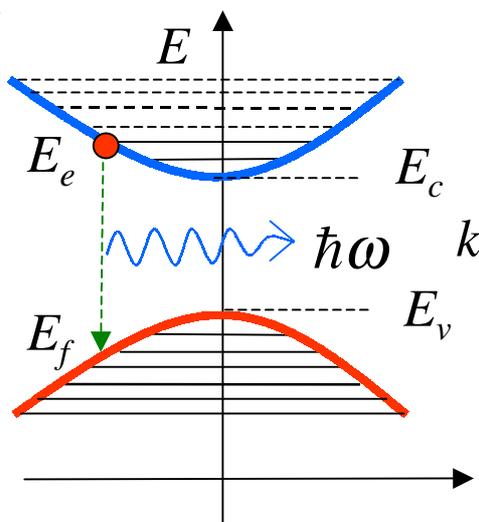
$$R_{tot}(\hbar\omega) = \alpha(\hbar\omega) v e^{-\hbar\omega/k_B T} \left(e^{\Delta E_F/k_B T} - 1 \right)$$

Spectre d'absorption

Thermalisation

Degré d'excitation

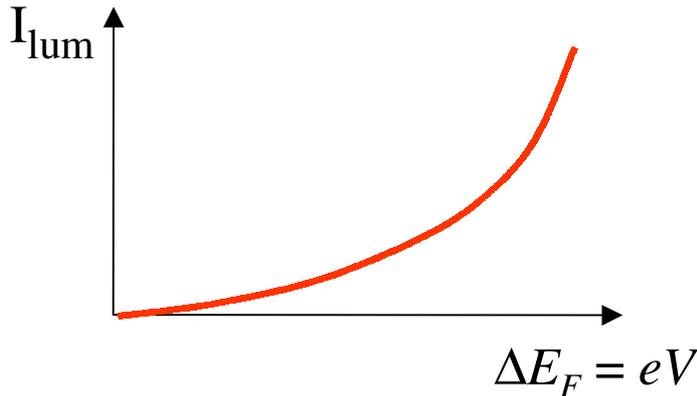
Cette expression montre que le spectre de luminescence est donné par le spectre d'absorption multiplié par le facteur de Boltzmann $\exp(-\hbar\omega/k_B T)$. Ce résultat s'interprète facilement : le coefficient d'absorption se retrouve ici parce qu'il représente en quelque sorte la quantité possible de transitions qui peuvent avoir lieu entre les bandes pour une fréquence donnée (cf. densité de transitions $\rho_i(\hbar\omega)$), et le facteur de Boltzmann représente le fait que les états pouvant donner lieu à ces transitions sont occupés par des électrons thermalisés en quasi-équilibre.



Les spectres de luminescence (courbe épaisse ci-dessus) présentent donc un seuil à l'énergie du gap et une décroissance exponentielle dépendante de la température.

Dans l'expression du taux de radiation par luminescence on trouve un terme représentant le rôle du degré d'excitation du S-C. Ce terme pouvait être attendu puisque le degré d'excitation est directement proportionnel au courant d'injection de la jonction $p-n$ polarisée en sens passant dont la valeur en fonction de la tension appliquée est donnée ci-dessous. Sachant que la tension appliquée détermine l'écart entre quasi-niveaux de Fermi $eV = \Delta E_F$, nous avons l'interprétation physique du terme représentant le degré d'excitation dans le taux de radiation par luminescence. Ce terme montre donc comment l'intensité lumineuse d'une DEL évolue en fonction de la tension appliquée à ses bornes.

$$\text{Degré d'excitation : } e^{\Delta E_F / k_B T} - 1$$



cf. réponse en intensité d'une diode polarisée :

$$I = I_s (e^{eV / k_B T} - 1)$$

Remarquons que le spectre mesurable en pratique n'est pas donné par R_{tot} . Un spectre expérimental est en réalité la mesure de la densité spectrale d'énergie émise. C'est à dire le taux d'apparition des photons $R_{tot}(\omega)$ à la fréquence ω multiplié par l'énergie de ces photons, le tout encore multiplié par la densité de modes électromagnétiques à la fréquence ω (à ne surtout pas oublier, cf. calcul du rayonnement thermique).

Spectre observé \equiv densité spectrale d'énergie émise

(mesuré par spectromètre optique)

$$I(\hbar\omega) = \hbar\omega \cdot R_{tot}(\hbar\omega) \cdot \rho(\hbar\omega)$$

Énergie des photons émis

Taux d'apparition des photons dans chaque mode ω

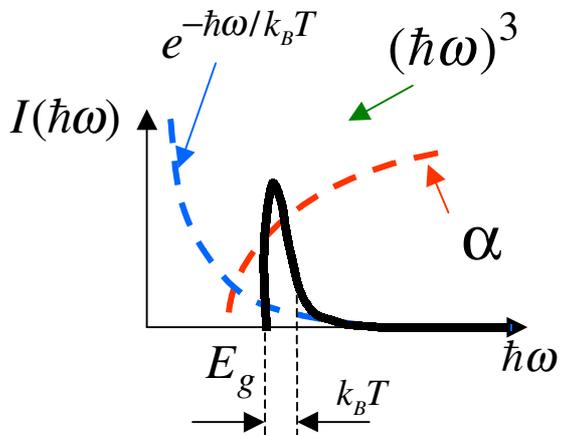
Densité de modes du champ E-M

$$\rho(\omega) = \frac{4}{\pi^2 c^3} \omega^2$$

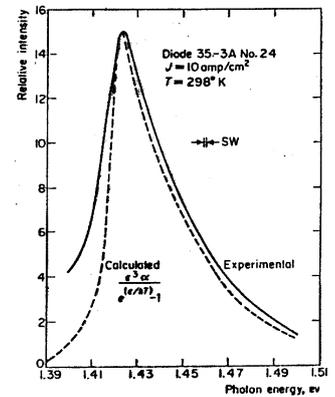
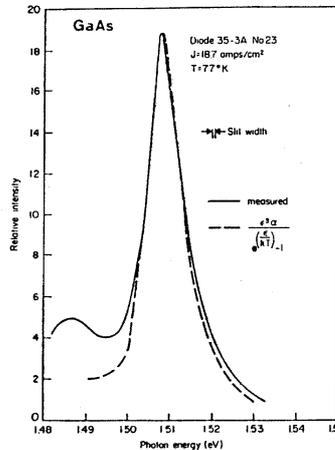
$$I(\hbar\omega) \div (\hbar\omega)^3 \alpha(\hbar\omega) e^{-\hbar\omega / k_B T} \left(e^{\Delta E_F / k_B T} - 1 \right)$$

Avec ces éléments nous sommes maintenant à même de faire la comparaison des prédictions théoriques et les spectres de luminescence mesurés en pratique. Sur les deux figures ci-dessous nous constatons un très bon accord quantitatif si ce n'est aux faibles énergies où l'on voit que les spectres expérimentaux présentent une plus grande densité spectrale (ce qui peut être facilement expliqué, comme nous le verrons par la suite).

Comparaison théorie-expérience



Largeur spectrale \approx énergie thermique

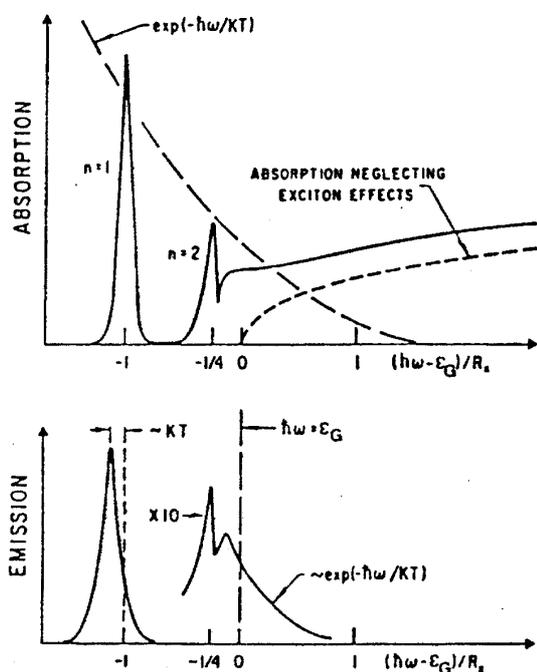


Une conclusion importante à tirer de cette étude est que les spectres de luminescence ont une largeur de l'ordre de l'énergie thermique. Ce résultat est similaire à celui trouvé pour le rayonnement thermique à la grande différence près que maintenant il s'agit de l'énergie thermique à température ambiante. Pour générer de la lumière visible à partir du rayonnement thermique il faut atteindre des températures de plusieurs milliers de degrés ce qui correspond à des largeurs spectrales énormes par rapport à celles obtenues avec les DEL. En ce qui concerne les applications aux télécommunications optiques, l'avantage des sources lumineuses à S-C apparaît donc très clairement. On est néanmoins encore très loin de la situation idéale pour les télécoms optiques que serait la génération de photons dans un mode électromagnétique spatio-temporel unique (une fréquence unique pour réduire la dispersion à son minimum et une direction unique pour l'injection dans les fibres optiques monomodes). Nous verrons plus loin que les lasers à S-C permettent de se rapprocher de cette situation idéale.

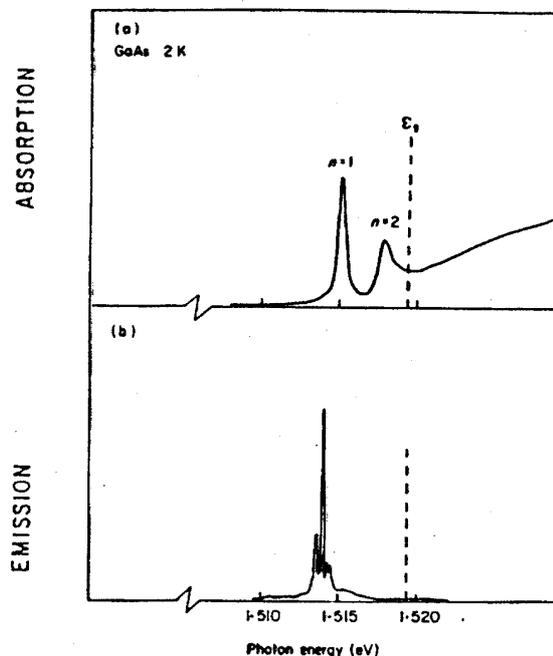
Il est intéressant d'étudier les spectres de luminescence à basse température car ils donnent des renseignements sur les états électroniques qui existent éventuellement sous le gap. En effet, à très basse température la thermalisation résulte en des spectres très fins situés sur les niveaux électroniques les plus bas (cf. rôle du facteur de Boltzmann). Comme le montrent les figures suivantes on peut ainsi étudier les excitons d'un S-C où les impuretés chimiques responsables de niveaux supplémentaires en dessous du gap dans la bande interdite. Pour chacun des cas, les figures donnent les spectres d'absorption et de luminescence. On constate que les spectres de luminescence mettent en évidence les structures fines de faible énergie. La figure de gauche est schématique alors que celle de droite donne des résultats expérimentaux pour un échantillon d'Arseniure de Gallium présentant des impuretés de faible énergie de liaison (niveaux proches du gap).

Spectres de luminescence à basse température

rôle des excitons



rôle des impuretés



Rendement des diodes électroluminescentes

Nous allons ici discuter brièvement les mécanismes affectant le rendement des DEL. Comme nous l'avons déjà mentionné plus haut, le rendement est conditionné par les transitions non radiatives. Un exemple important de transition non radiative est la transition par transfert d'énergie avec les phonons du réseau cristallin. Puisque l'énergie des phonons est généralement plus faible que celle du gap, il faut des échanges d'énergie avec plusieurs phonons pour une seule transition. On peut estimer que pour descendre tout le gap soit une énergie de l'ordre de 1eV, l'électron doit céder son énergie à 30-40 phonons. Ceci ne peut bien entendu pas se faire en une seule fois car la probabilité d'interaction entre un tel nombre de particules est infime. En pratique l'échange a lieu par étape grâce à l'existence de niveaux électroniques intermédiaires comme indiqué sur le schéma suivant. Si de tels niveaux intermédiaires n'existaient pas, une telle émission de phonons en cascade ne pourrait pas se produire et les transitions radiatives seraient alors privilégiées. On comprend ainsi l'importance des défauts et des impuretés dans les S-C destinés aux applications optiques. En pratique, on cherchera à éviter les dislocations, les lacunes, les impuretés chimiques et les interactions avec les surfaces du matériau S-C.

Rendement conditionné par les transitions non radiatives

Exemple 1 : transfert d'énergie avec les phonons

$$E_g \approx 1 \text{ eV} \approx 30 \text{ à } 40 \text{ phonons}$$

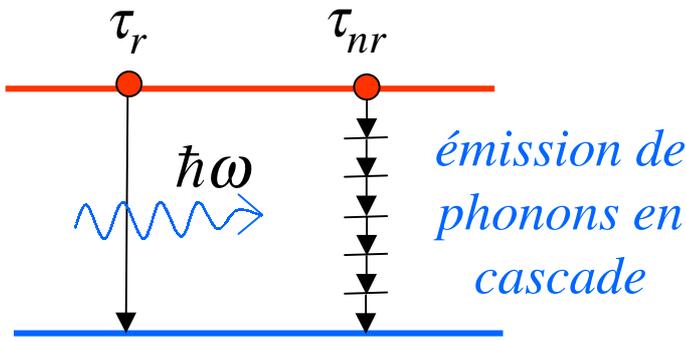
(collisions réseau cristallin)



niveaux intermédiaires nécessaires



rôle essentiel des impuretés et défauts

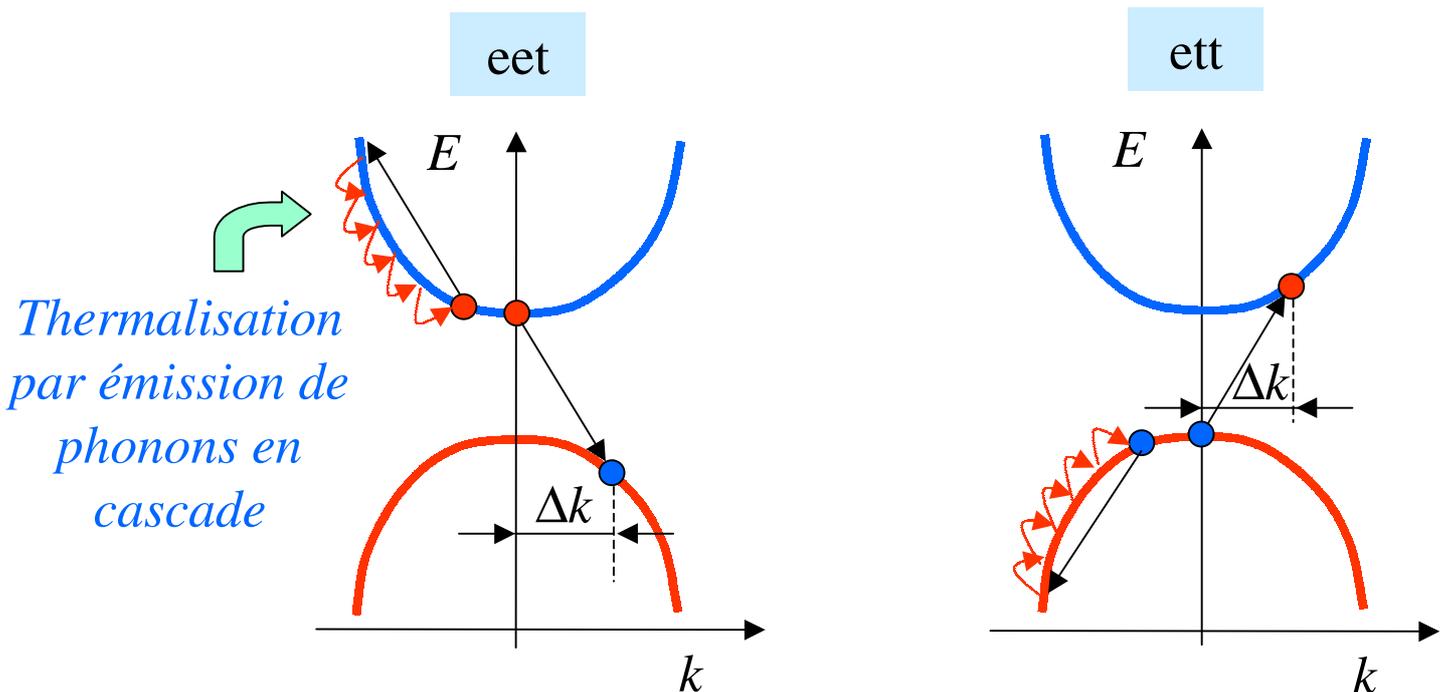


- lacunes
- surfaces
- dislocations
- ...

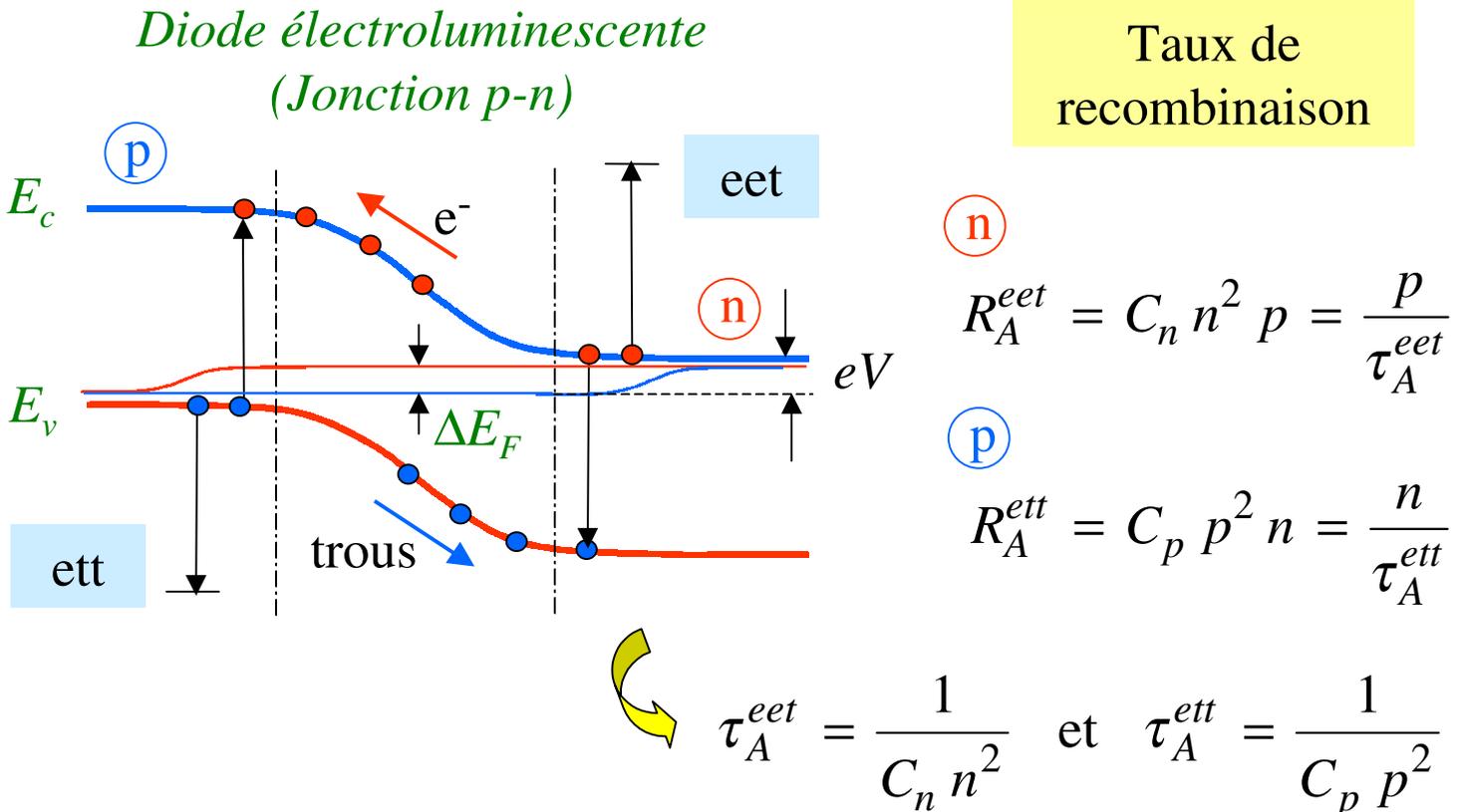
Exemple 2 : recombinaison Auger

Un autre mécanisme très important de recombinaisons non radiatives responsables de limitation de rendement des DEL est la recombinaison Auger. Lors d'une recombinaison Auger, l'électron cède son énergie à un autre électron de la bande de conduction. Ce dernier monte donc dans la bande mais il est très vite thermalisé (10^{-13} sec) par interaction avec le réseau cristallin (cascade de phonons). L'énergie est donc perdue sous forme de phonons (échauffement) et aucun photon n'est émis. Ce mécanisme de recombinaison non radiative est représenté schématiquement à la figure de gauche ci-dessous. On notera que l'interaction entre l'électron qui descend et l'électron Auger doit respecter la conservation de la quantité de mouvement.

La figure de droite illustre le même phénomène avec un électron Auger dans la bande de valence : l'électron qui descend donne son énergie et sa quantité de mouvement à un électron profond de la bande de valence, celui-ci est thermalisé par cascade de phonons. Il est habituel d'interpréter ce mécanisme en terme de transition de trous. L'électron qui descend est vu comme un trou qui monte grâce à un échange d'énergie avec un trou de la bande de valence. Ce trou « Auger » est alors thermalisé et remonte dans la bande par cascade de phonons. Pour différencier les deux mécanismes, on appelle le premier effet Auger *eet* et le second effet Auger *ett*. Ces lettres signifient recombinaison électron-trou par l'intermédiaire d'un électron (ou d'un trou) Auger (la particule Auger étant représentée par les lettres du milieu *e* et *t* dans *eet* et *ett*, respectivement).



A la figure ci-dessous nous illustrons les deux mécanismes de recombinaison Auger dans une jonction $p-n$ polarisée. Puisque le mécanisme eet fait intervenir deux électrons simultanément sa probabilité est proportionnelle au carré de la densité électronique, soit n^2 . Ce mécanisme affectera donc plus particulièrement les recombinaisons qui se font en région de type n . De même le mécanisme Auger ett , faisant intervenir deux trous simultanément, est proportionnel au carré de la densité de trous, soit p^2 . Il affecte donc les recombinaisons du côté p et est négligeable du côté n . Puisque le mécanisme eet , exige la présence d'un état électronique libre, c'est à dire un trou, dans la bande de valence sa probabilité est également proportionnelle à p , soit une probabilité globale proportionnelle à n^2p . Il en est de même pour le mécanisme ett qui met en jeu un électron de la bande de conduction (il faut un électron dans la bande de conduction pour que le trou puisse y monter). La probabilité associée à l'effet Auger ett est donc proportionnelle à p^2n .



Les formules ci-dessus définissent les taux de recombinaison Auger R_A pour les deux mécanismes eet et ett . Les constantes de proportionnalité C_n et C_p sont appelées les coefficients Auger. Comme le taux R_A^{eet} représente le nombre de trous qui disparaissent par unité de temps du côté n , soit $R_A^{eet} = -dp/dt$, on peut définir la durée de vie Auger eet comme étant $\tau_A^{eet} = 1/C_n n^2$. De même pour l'effet Auger ett on a $\tau_A^{ett} = 1/C_p p^2$. Ce sont ces durées de vie qui conditionneront le rendement des DEL. Notons que pour les applications optiques on utilise des semi-conducteurs fortement dopés afin d'atteindre aisément des degrés d'excitation élevés (le degré d'excitation est proportionnel à la densité de porteurs injectés qui est d'autant plus élevée que le dopage de la jonction $p-n$ est élevé). Comme les durées de vie Auger sont inversement proportionnelles aux concentrations de porteurs, on voit qu'il faudra être attentif à l'effet Auger en ce qui concerne le rendement de la génération de lumière et un compromis devra être établi entre rendement et excitabilité.

Exemple : GaAs

Nous illustrons ici cette brève discussion avec l'exemple du GaAs fortement dopé.

$$C_n = 10^{-31} \text{ cm}^{-6} \text{ s}^{-1}, \quad \text{soit} \quad n = 10^{18} \text{ cm}^{-3} \quad \rightarrow \quad \tau_A^{eet} = 10^{-5} \text{ s}$$

Si dopage élevé : compétition entre les recombinaisons Auger et radiative

Les deux formules ci-dessous font la synthèse des grandeurs introduites dans la discussion.

$$\text{Taux de recombinaison Auger : } \begin{cases} R_A^{eet} = C_n n^2 p = \frac{p}{\tau_A^{eet}} = -\frac{dp}{dt} \\ R_A^{ett} = C_p p^2 n = \frac{n}{\tau_A^{ett}} = -\frac{dn}{dt} \end{cases}$$

Dans les S-C fortement dopés tels que ceux utilisés dans les applications optiques, ce sont les recombinaisons Auger qui conditionnent le rendement (les autres mécanismes de recombinaison non radiative sont en général négligeables). Pour estimer le rendement il faut donc comparer les temps de vie radiative τ_r et Auger τ_A . Les formules ci-dessous expriment la compétition entre les deux effets. Les deuxièmes termes des membres de droite représentent les contribution à la génération de photons.

$$\text{Taux de recombinaison total : } \begin{cases} \textcircled{n} \quad \frac{dp}{dt} = -\frac{p}{\tau_A^{eet}} - \frac{p}{\tau_r} \\ \textcircled{p} \quad \frac{dn}{dt} = -\frac{n}{\tau_A^{ett}} - \frac{n}{\tau_r} \end{cases}$$

Les tableaux ci-dessous donnent les durées de vie Auger (inversement proportionnel au coefficient Auger C) et les durées de vie radiative pour différents S-C. Le Phosphure de Gallium, le Silicium et le Germanium ont tout trois un gap indirect, les temps de vie radiative sont donc relativement grands. Ils sont en tout cas supérieurs (ou comparable pour le GaP) aux durées de vie Auger. Il apparaît donc évident que ces S-C ne présenteront pas des bons rendements. Le cas de l'Antimoniure de Gallium est particulier car bien qu'étant à gap direct sa durée de vie Auger est plus courte que sa durée de vie radiative. Ce n'est donc pas un bon candidat pour la fabrication des DEL non plus. Le cas du GaAs apparaît par contre idéal.

Valeurs expérimentales des coefficients Auger (à 77 ou 300 K) pour la recombinaison Auger intrinsèque (eet ou ett): (1) GaAs: ett, 77 K. (2) GaSb: ett, 77 K. (3) GaP: ett. (4) Si: ett, 300 K. (5) Ge: eet, ett, 300 K. Les durées de vie Auger sont déterminées à partir des valeurs expérimentales des coefficients Auger ($\tau_A = 1/Cn^2$ ou $\tau_A = 1/Cp^2$).

Materiau	GaAs	GaSb	GaP	Si	Ge
C ($\text{cm}^6 \text{s}^{-1}$)	10^{-31}	10^{-25}	8×10^{-31}	3×10^{-31}	1×10^{-31}
processus Auger	ett	ett	ett	eet	eet, ett
τ_A (s) à 1×10^{18} (cm^{-3})	10^{-5}	10^{-11}	1.3×10^{-6}	3.3×10^{-6}	10^{-5}

Temps de vie radiative

Materiau	GaAs	GaSb	InSb	GaP	Si	Ge
r ($\text{cm}^3 \text{s}^{-1}$)	7.2×10^{-10}	2.4×10^{-10}	4.6×10^{-11}	5.4×10^{-14}	1.8×10^{-15}	5.3×10^{-14}
τ_r (s) à 1×10^{18} (cm^{-3})	1.3×10^{-9}	4.2×10^{-9}	2.2×10^{-8}	1.9×10^{-5}	5.6×10^{-4}	1.9×10^{-5}
ϵ_g [eV] à 300 K	1.435	0.72	0.18	2.26	1.12	0.66

Cette discussion sur le choix des matériaux est résumée ci-dessous.

Exemples :

Gap indirect : Si, Ge, GaP a priori pas utilisé pour les DEL

Gap direct : GaSb trop d'effet Auger, GaAs très bon matériau

Rendement quantique interne de luminescence d'un S-C

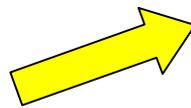
Nous allons ici préciser la notion de rendement de luminescence des S-C. La formule ci-dessous à gauche donne la variation du nombre de porteurs injectés (électrons ou trous) en la décomposant en ses contributions radiatives R_r et non radiative R_{nr} . Le rendement quantique interne est défini de manière naturelle comme étant le rapport du nombre de recombinaisons radiatives sur le nombre total de recombinaisons $R_r/(R_r + R_{nr})$. Cette expression peut facilement être exprimée en fonction des durées de vie. On arrive alors à la formule encadrée ci-dessous.

Soit δn l'excès de porteur ($\delta n = p$ pour type n, $\delta n = n$ pour type p)

$$\frac{d\delta n}{dt} = -\frac{\delta n}{\tau_{nr}} - \frac{\delta n}{\tau_r}$$

\downarrow
 R_{nr}

\downarrow
 R_r



$$\eta_{\text{int}} \equiv \frac{R_r}{R_r + R_{nr}}$$

$$\eta_{\text{int}} = \frac{1/\tau_r}{1/\tau_r + 1/\tau_{nr}}$$

Exemple :

$$\text{GaAs} : \eta_{\text{int}} \approx 1$$

$$\text{Si} : \eta_{\text{int}} \approx 0$$

$$\eta_{\text{int}} = \frac{1}{1 + \tau_r / \tau_{nr}}$$

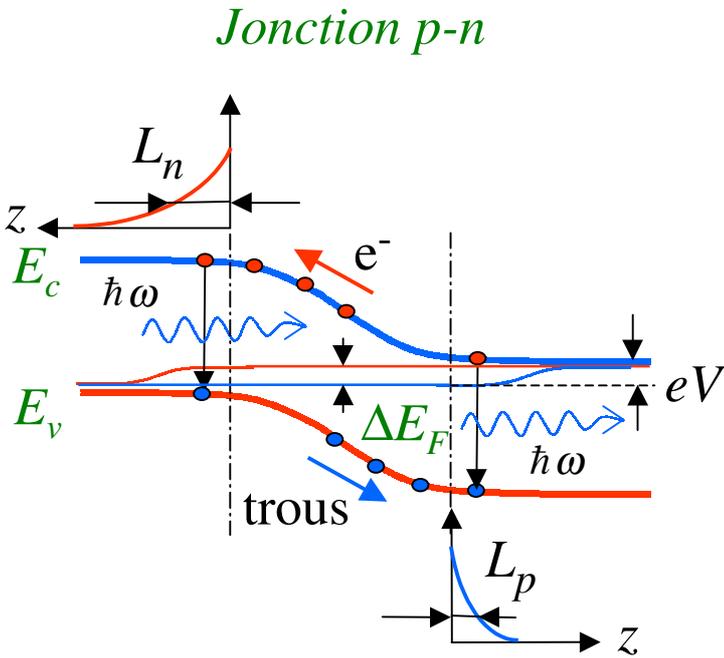
D'après les tableaux de la page précédente on peut facilement évaluer les rendements quantiques internes du GaAs et du Si donnés ci-dessus.

Le rendement quantique interne n'est proche de l'unité que si les conditions rappelées ci-dessous sont rencontrées.

$\eta_{\text{int}} \rightarrow 1$ si : gap direct, dopage limité (Auger), peu d'impuretés et défauts (cascades de phonons)

Les diodes électroluminescentes

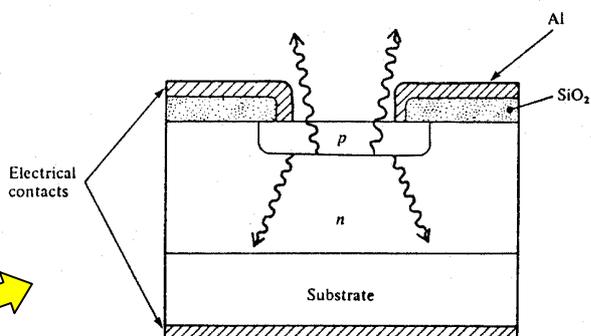
Les notions de base vues plus haut sur l'électroluminescence des S-C, nous permettent de préciser le fonctionnement des LED. La figure ci-dessous illustre ce fonctionnement. En raison de la polarisation de la jonction, des électrons sont injectés du côté *p* et des trous sont injectés du côté *n*. Ces porteurs excédentaires diffusent et se recombinent par transitions radiatives. Leur densité décroît donc de manière exponentielle en s'éloignant de la région de charge d'espace. Les longueurs de diffusion L_n et L_p sont données ci-dessous en fonction des coefficients de diffusion D_n et D_p eux-mêmes fonction de la mobilité des porteurs μ_n et μ_p . Puisque la mobilité de trous est significativement plus faible que celle des électrons, la longueur de diffusion des électrons L_n est plus grande que la longueur de diffusion des trous du côté *n*, L_p . L_n est typiquement de l'ordre de quelques microns. C'est donc du côté *p* que la zone sur laquelle les photons seront émis est la plus large. C'est la raison pour laquelle les DEL sont fabriquées avec la zone *p* proche de la surface émettrice comme indiqué sur le schéma de structure d'une DEL ci-dessous.



$$\begin{cases} L_n = \sqrt{D_n \tau_n} \\ L_p = \sqrt{D_p \tau_p} \end{cases} \quad D = \mu \frac{k_B T}{e}$$

$$\mu_n \approx 20 \mu_p, \quad L_n = 10 \mu_m$$

Structure d'une DEL (LED)



Limitation de la réabsorption : surface émettrice en zone *p*

La zone de génération de photons la plus large (soit la zone *p*) est placée près de la surface émettrice afin de limiter de problème de la réabsorption. En effet, les photons émis pourraient toujours être réabsorbés par transition d'un électron vers la bande de conduction. Cet électron pourra alors se désexciter de manière soit radiative soit non radiative. S'il est désexcité de manière radiative, on retrouve le photon initial et rien n'est perdu quant au rendement de la DEL. Mais il faut garder à l'esprit que l'électron peut se désexciter de manière non radiative avec une probabilité non négligeable surtout si l'électron excité se trouve près d'une surface ou tout autre défaut de la structure cristalline car il sera alors principalement désexcité de manière non radiative par cascade de phonons. Du point de vue du rendement, il y a donc un intérêt évident à faire en sorte que les photons soient générés dans une région la plus proche possible de la surface d'émission de la DEL. De cette façon les photons ont beaucoup moins de chance d'être réabsorbés. Si la région *n* était mise près de la surface émettrice, cette condition ne serait pas satisfaite car tous les photons seraient générés très près de la zone de charge d'espace et devraient encore traverser toute la région *n* où ils auraient de bonne chance d'être réabsorbés.

Rendement des diodes électroluminescentes

Nous nous intéressons maintenant au rendement global des DEL. Nous devons pour cela considérer la puissance électrique que requiert la polarisation de la jonction, $P_{elec} = I.V$ où I est fourni par la réponse en courant de la jonction $p-n$ donnée ci-dessous. On a vu plus haut que la puissance lumineuse émise est proportionnelle au courant I (voir taux de génération de photons total, R_{tot}). Mais comme nous venons de le voir, tous les photons créés ne sont pas émis en raison du phénomène de réabsorption. Ces considérations sont résumées ci-dessous.

$$\text{Puissance électrique : } P_{elec} = I \cdot V \quad \text{où} \quad I = I_s (e^{eV/k_B T} - 1)$$



$$\text{Puissance lumineuse émise : } P_{lum} \propto (e^{eV/k_B T} - 1)$$

Jonction $p-n$: zone d'émission réduite par $L_n = \text{qlqs } \mu\text{m}$



Perte de photons par réabsorption :

Création de paire e^- -trou et recombinaison non radiative

(impuretés, surfaces, ...)

Rendement de la diode

Le rendement global d'une DEL est par définition le rapport de la puissance lumineuse émise sur la puissance électrique consommée. C'est donc le produit de la tension appliquée V_a par le courant traversant la jonction; le courant étant le nombre d'électrons passant la jonction par seconde multiplié par la charge de l'électron.

$$\eta = \frac{\text{Puiss. lumineuse}}{\text{Puiss. électrique}} = \frac{\hbar\omega \cdot \text{Nbre de photons émis/sec}}{eV_a \cdot \text{Nbre d}'e^- \text{ passant la jonction/sec}}$$

En introduisant le nombre de photons créés en le multipliant au numérateur et au dénominateur, nous faisons apparaître le rendement interne qui n'est autre que le rapport du nombre de photons créés sur le nombre d'électrons passant dans la jonction (ce dernier étant le nombre de paires électron-trou créées). Le quotient du nombre de photons émis sur le nombre de photons créés définit le rendement optique.

$$\eta = \frac{\hbar\omega}{eV_a} \underbrace{\frac{\text{Nbre de photons émis/sec}}{\text{Nbre de photons créés/sec}}}_{\eta_{optique}} \underbrace{\frac{\text{Nbre de photons créés/sec}}{\text{Nbre d}'e^- \text{ passant/sec}}}_{\eta_{interne}}$$

Nous rappelons ci-dessous la définition du rendement interne de luminescence pour expliquer son expression en terme du rapport du nombre d'électrons passant sur le nombre de photons créés.

$$\eta_{\text{int}} = \frac{1}{1 + \tau_r / \tau_{nr}}, \quad \frac{d\delta n}{dt} = - \frac{\delta n}{\tau_{nr}} - \frac{\delta n}{\tau_r}$$

Nbre d'e⁻ passant/sec
Nbre de ph. créés/sec

L'expression du rendement d'une DEL est donc la suivante :

$$\eta = \frac{\hbar\omega}{eV_a} \eta_{\text{opt}} \eta_{\text{int}}$$

η_{int} déterminé par le choix du matériau (GaAs : $\eta_{\text{int}} \approx 1$)
 C'est le rendement optique qui limite l'efficacité d'une LED

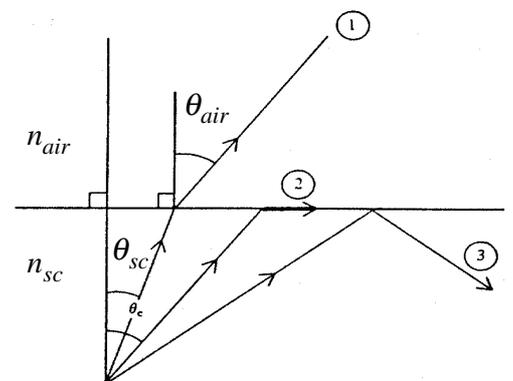
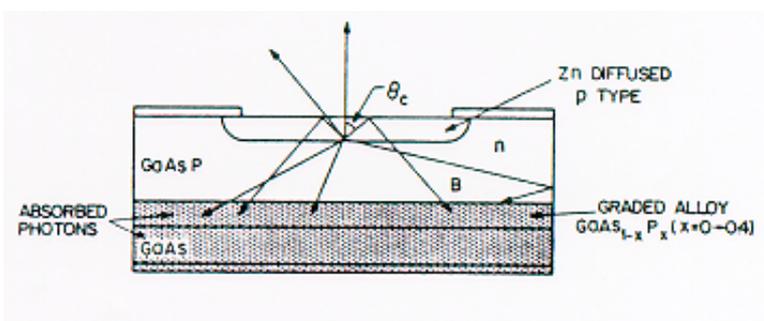
Rendement optique des LED

Nous discutons ici le rendement des DEL en fonction de leur structure (en particulier, leur géométrie). La structure des DEL est en effet déterminante pour leur rendement. Ceci provient du fait que l'indice de réfraction des S-C est élevé et occasionne, dès lors de nombreuses réflexions (voir plus bas). Ces réflexions ont pour effet d'augmenter le trajet des photons dans le S-C et donc d'augmenter le risque de réabsorption et diminuer le rendement. Cet effet est d'autant plus marqué que les réflexions renvoient les photons loin de la surface émettrice et en particulier vers le substrat. Le substrat contient plusieurs interfaces inévitablement associées à des défauts et donc à une augmentation de la réabsorption (voir schéma ci-dessous).

Réabsorption renforcée par les réflexions à l'interface S-C/air

De plus, il est à noter que l'émission spontanée étant isotrope (pas de direction privilégiée), les photons sont émis dans toutes les directions. Beaucoup seront donc perdus par réabsorption directe sur les parois et interface de la DEL. En ce qui concerne les réflexions, il y en a de deux types : les réflexions de Fresnel dues au passage d'un changement abrupt d'indice et les réflexions internes totales (voir schéma de droite ci-dessous).

Rayonnement de nature spontanée donc isotrope : $0 < \theta_{sc} < 2\pi$



Les réflexions de Fresnel sont partielles et le coefficient de réflexion associé R est donné ci-dessous en fonction des indices de réfraction du S-C ($n_{sc} = 3$ à 4 , typiquement) et de l'air ($n_{air} = 1$). A incidence normale ce coefficient de réflexion vaut à peu près 30% et il augmente quand l'angle d'incidence θ_{sc} augmente. Si l'angle d'incidence dépasse la valeur critique de l'angle de réflexion totale θ_c dont l'expression est donnée ci-dessous, la réflexion est de 100%. L'angle de réflexion totale est typiquement de l'ordre de 15° , ce qui signifie que tous les photons émis dans une direction faisant un angle supérieur à 15° par rapport à la normale à la surface émettrice restent à l'intérieur du S-C et finissent par être réabsorbés. On voit ainsi immédiatement que le rapport entre le nombre de photons émis et le nombre de photons générés dans la jonction (autrement dit, le rendement optique) sera très faible si aucune précaution n'est prise. Il est en général de l'ordre de 1%.

Réflexions partielles ($\theta_{sc} \approx 0$)

$$\begin{cases} n_{sc} = 3.5 \\ n_{air} = 1 \end{cases} \Rightarrow R = \left(\frac{n_{sc} - n_{air}}{n_{sc} + n_{air}} \right)^2 \approx 30\%$$

Réflexion totale ($\theta_{sc} > \theta_c$)

$$\begin{cases} n_{sc} \sin \theta_{sc} = n_{air} \sin \theta_{air} \\ \theta_c = \arcsin(n_{air} / n_{sc}) \approx 16^\circ \end{cases}$$

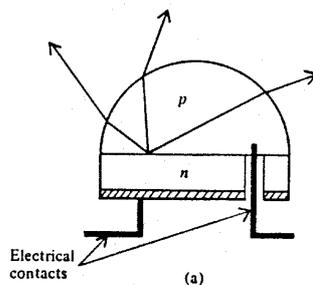
Conclusion : $\eta_{opt} \approx 1\%$

Augmentation du rendement optique :

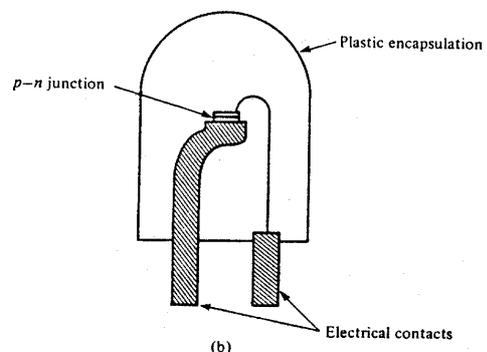
Afin d'augmenter le rendement optique des DEL on a recours à plusieurs astuces. L'une d'entre elles consiste à donner à la zone émettrice (région de type p) une forme de dôme sphérique. Comme indiqué sur le schéma ci-dessous à gauche, l'augmentation de rendement est simplement due au fait qu'un plus grand nombre de directions correspondent maintenant à un angle d'incidence inférieur à l'angle de réflexion totale. Une autre technique analogue consiste à encapsuler la DEL dans un dôme de matière transparente d'indice élevé (le plus près possible de l'indice du S-C) de façon à éliminer le plus possible les réflexions à la surface émettrice de la jonction. La forme de dôme de la capsule assure alors un minimum de réflexions totales (voir schéma de droite).

Le rendement peut être ainsi augmenté d'un facteur 4 à 5 (soit 4 à 5 % de rendement optique).

Région p en dôme hémisphérique



Encapsulation



$\eta_{opt} \rightarrow 4 - 5\%$

Choix des matériaux pour les diodes électroluminescentes

Le choix d'un matériau S-C pour une DEL dépend bien entendu de la longueur d'onde d'émission désirée puisque celle-ci est déterminée par le gap du matériau. En second lieu intervient le rendement quantique interne du matériau. Pour avoir un rendement élevé il faut bien entendu un matériau à gap direct, de grande pureté chimique et de bonne qualité cristalline (on utilisera donc préférentiellement des matériaux fabriqués par des techniques avancées d'épithaxie).

Critères

1. *Rendement quantique interne : gap direct, grande pureté chimique, pas de défauts (épithaxie)*
2. *Longueur d'onde d'émission : énergie du gap*

Pour faire le choix du matériau par rapport à sa longueur d'émission il est commode d'utiliser la formule ci dessous liant l'énergie du gap à la longueur d'onde. On peut tout de suite en déduire que les matériaux de gap proche de 1 eV émettent dans l'infrarouge proche et que pour avoir une émission dans le domaine visible il faut des gaps au moins deux fois plus élevés ; ce qui pose un problème pour les applications des DEL à l'affichage car les matériaux de gap élevé et direct ne sont en général pas ceux utilisés couramment dans les applications de l'électronique (matériaux moins connus, techniques de fabrication moins efficaces, difficultés technologiques pour le dopage, incompatibilité avec la technologie électronique usuelle, etc...)

$$\hbar\omega = E_g \quad \text{où} \quad \hbar = h/2\pi, \quad \omega = 2\pi c/\lambda$$

$$\left[h = 6.626 \cdot 10^{-34} \text{ J}\cdot\text{sec}, \quad c = 3 \cdot 10^8 \text{ m/sec}, \quad 1 \text{ eV} = 1.602 \cdot 10^{-19} \text{ J} \right]$$

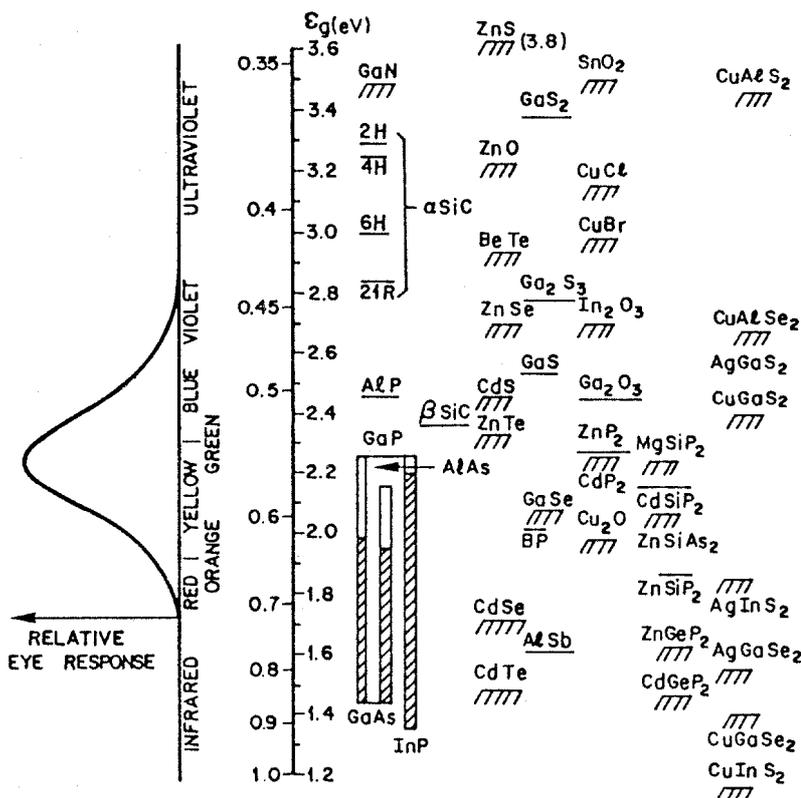

$$\lambda[\mu\text{m}] \cdot E_g[\text{eV}] = 1.24$$

$$\left\{ \begin{array}{l} E_g = 1 \text{ eV} \implies \lambda = 1.24 \mu\text{m} \text{ IR} \\ E_g = 2 \text{ eV} \implies \lambda = 0.62 \mu\text{m} \text{ Vis.} \end{array} \right.$$

Le tableau suivant donne les différentes caractéristiques intéressantes pour les application optiques des S-C. On y voit qu'aucun matériau à gap direct ne permet la fabrication de DEL dans le visible, en particulier, en raison de difficultés liées au dopage (pour les matériaux tels que ZnS, Zn Se, ZnTe).

Group(s)	Element/ Compound	Direct/ Indirect	E_g (eV)	Readily doped <i>n</i> or <i>p</i> type	r ($m^3 s^{-1}$)	λ_g (nm)
IV	C	i	5.47			227
	Si	i	1.12	yes	$1.79 \cdot 10^{-21}$	1106
	Ge	i	0.67	yes	$5.25 \cdot 10^{-20}$	1880
IV-VI	SiC (hex. α)	i	3.00	yes		413
III-V	AlP	i	2.45			506
	AlN	i	5.90	no		210
	AlSb	i	1.50			826
	AlAs	i	2.16			574
	GaN	d	3.40	uncertain		365
	GaP	i	2.26	yes	$5.37 \cdot 10^{-20}$	549
	GaAs	d	1.44	yes	$7.21 \cdot 10^{-16}$	861
	InN	d	2.40			516
	InP	d	1.35	yes	$1.26 \cdot 10^{-15}$	918
	InAs	d	0.35		$8.50 \cdot 10^{-17}$	354
	InSb	d	0.18		$4.58 \cdot 10^{-17}$	687
II-VI	ZnO	d	3.20	no		387
	ZnS(α)	d	3.80	no		326
	ZnS(β)	d	3.60	no		344
	ZnSe	d	2.28	no		480
	ZnTe	d	2.58	no		544
	CdS	d	2.53	no		490
	CdSe	d	1.74	no		712
	CdTe	d	1.50	yes		826

Application à l'affichage (LED visibles)



La figure ci-contre montre un classement des composés S-C selon leur groupe (III-V et II-VI pour les deux premières colonnes et leur énergie de gap. Sur la gauche est indiquée la courbe de sensibilité de l'œil humain. Les matériaux à gap direct sont soulignés en hachuré.

En pratique ce sont les composés III-V qui sont choisis en raison de leur facilité de fabrication. Le problème est que les seuls composés III-V à avoir un gap dans le visible sont le GaP et le AlAs qui ont tout deux un gap indirect. Le GaAs et l'InP ont un gap direct mais ils émettent dans le proche infrarouge. L'idée est alors de fabriquer des alliages ternaires. Les alliages GaAsP, AlGaAs et GaInP sont indiqués par les bandes verticales (celles-ci sont hachurées quand l'alliage a un gap direct).

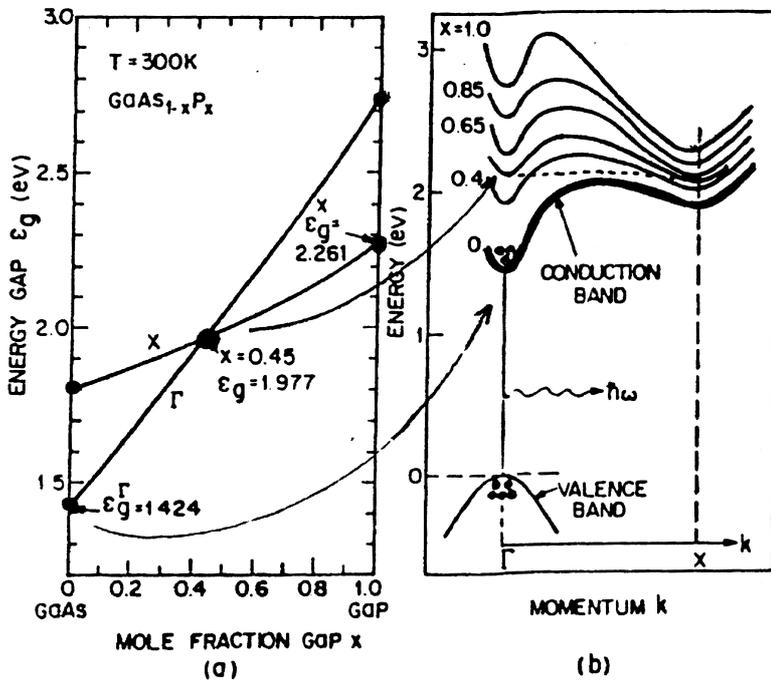
GaP : $E_g = 2.26 \text{ eV} \rightarrow \lambda = 0.549 \text{ } \mu\text{m}$ visible mais gap indirect

GaAs : $E_g = 1.44 \text{ eV} \rightarrow \lambda = 0.861 \text{ } \mu\text{m}$ IR proche mais gap direct



Solution : alliage ternaire $\text{GaAs}_{1-x}\text{P}_x$

L'idée de l'alliage ternaire est de réaliser un composé dans lequel deux éléments de même valence se trouvent ensemble, en une proportion donnée x , liés à l'élément de la valence complémentaire. Dans l'exemple ci-dessus les éléments As et P de valence V sont mélangés et combinés au Gallium de valence III pour donner un nouveau matériau S-C. Le pas du réseau du S-C résultant de cet alliage se voit modifié linéairement en fonction de la proportion de l'alliage, comme indiqué par la formule ci-dessous à droite. Il importe néanmoins que les constantes de réseau des deux composés purs correspondants (de proportions 0 et 1) soient suffisamment proches pour éviter des problèmes de dislocations (cf. pureté cristalline). De même l'énergie du gap de l'alliage et une fonction linéaire de la proportion x des éléments.



Pas du réseau

GaP : a , GaAs : A

$$a' = x a + (1-x) A$$

Energie du gap

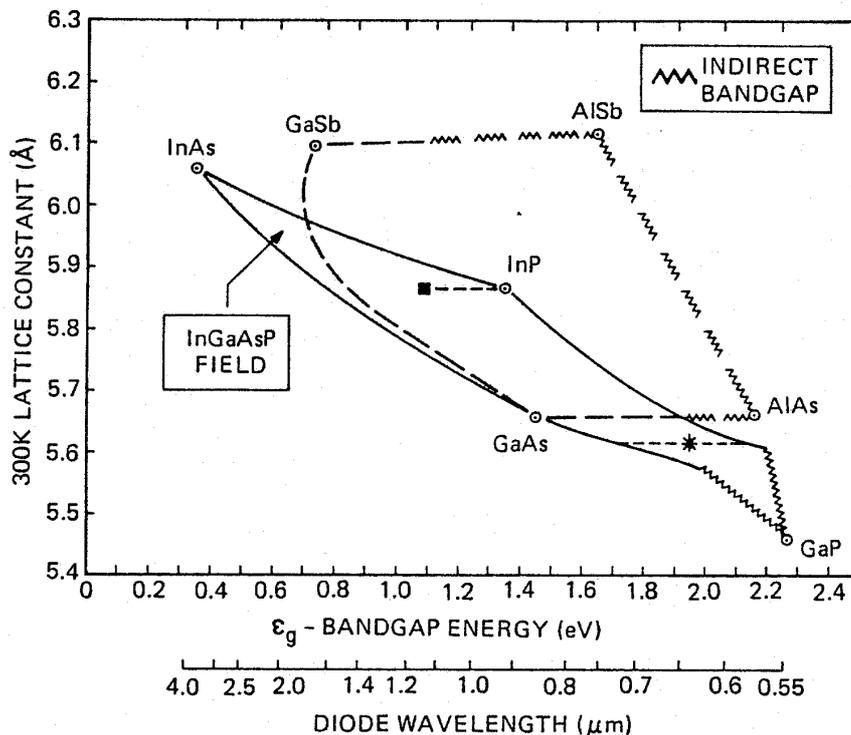
$$E_g' = x E_{g\text{GaP}} + (1-x) E_{g\text{GaAs}}$$

La figure ci-dessus montre les courbes de variation de l'énergie des gaps direct et indirect de l'alliage ternaire $\text{GaAs}_{1-x}\text{P}_x$ en fonction de la proportion de Phosphore x . On notera que la variation du gap avec x n'est pas strictement linéaire contrairement à la formule approchée ci-dessus. Les deux gaps augmentent néanmoins de manière monotone et quasi-linéaire avec x ; la structure de bande correspondante aux différentes valeurs de x est représentée sur le schéma de droite. Pour les applications à l'affichage dans le visible on utilisera donc des valeurs de x élevées. Néanmoins, comme on le voit sur la figure, au-delà de $x = 0.45$ le gap indirect (direction cristalline X) devient inférieur au gap direct (direction cristalline Γ). Il faudra donc, en principe, se limiter à des valeurs de x inférieures à 0.45 pour des raisons de limitation du rendement quantique interne de la DEL. Le gap limite correspondant est légèrement inférieur à 2 eV ce qui correspond à l'émission de lumière rouge-orangé (proche de la limite inférieure de la courbe de sensibilité de l'oeil).

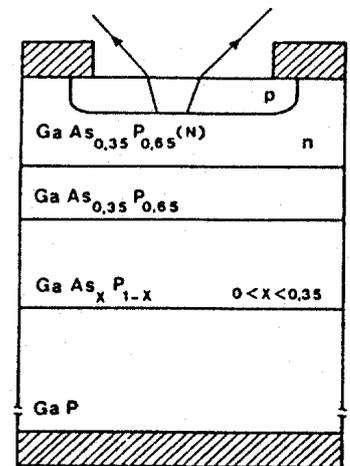
Alliages ternaires et quaternaires

Le tableau ci-dessous montre les courbes de constante de réseau en fonction du gap des alliages ternaires. Les courbes de pente proche de zéro (AlGaSb, AlGaAs) représentent les alliages ternaires les plus facilement réalisables puisqu'ils ne présentent pas de problème de contrainte induite par les différences de maille de réseau. Les cassures dans les courbes des alliages ternaires GaAsP et InGaP s'expliquent par le passage du gap direct au gap indirect (ce dernier étant indiqué par les lignes en zig-zag).

Pour certaines applications nécessitant des cristaux de bonne qualité (ceux-ci sont obtenus avec les alliages de composés de même constante de réseau tels que AlGaSb, AlGaAs) on a recours à des alliages quaternaires. Ceux-ci consistent en des alliages dans lesquels les deux éléments de valences complémentaires sont doublés. Parmi les alliages S-C quaternaires l'InGaAsP est particulièrement important car il est utilisé pour les applications aux télécommunications (sources lumineuses aux longueurs d'onde de 1.3 et 1.5 microns inaccessibles avec l'alliage AlGaAs).



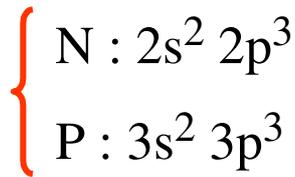
Exemple de Diode au GaAsP



Le schéma de droite montre une structure classique de DEL au GaAsP. Les zones hachurées représentent les électrodes métalliques. La première couche de S-C est du GaP pur qui sert de substrat de départ pour la croissance cristalline de l'alliage. La deuxième couche est une couche tampon dans laquelle la proportion d'Arsenic est augmentée progressivement de manière à éviter la présence d'un trop grand nombre de défauts du à la grande différence de constante de réseau du GaP et du GaAs (voir schéma ci-dessous). Une deuxième couche tampon à x constant (sur le schéma ci-dessous $x = 0.35$ correspond à $x = 0.65$ du schéma de la page précédente, ce qui correspond à un gap grand mais indirect) est également introduite de manière à éliminer au maximum les défauts dans la zone utile de la DEL. La zone à $x = 0.35$ est alors dopée n et ensuite dopée p de manière à obtenir la jonction $p-n$ avec le gap élevé voulu pour l'affichage dans le visible jaune-orangé.

Pour remédier au problème du gap indirect obtenu avec $x = 0.35$ on a recours à un dopage à l'azote de la zone utile de la DEL (jonction $p-n$). L'azote a la même valence que le phosphore mais il est plus petit (les couches périphériques correspondent au nombre quantique principal de $n=2$ au lieu de $n=3$ pour le phosphore). De ce fait l'azote forme un piège à électron dans le cristal et il se charge négativement. Les trous se combinent à ces charges négatives localisées et forment donc des excitons. Ces excitons étant liés, ne doivent pas être vus comme des états électroniques du S-C, en particulier, la transition électronique de recombinaison électron-trou n'obéit pas à la loi de conservation de la quantité de mouvement. En effet, si il s'agit d'un état confiné, les fonctions d'onde ont une extension Δx faible correspondant à une grande valeur de l'incertitude sur la quantité de mouvement, soit Δk . Dès lors, malgré le gap indirect, la présence de tels centres de recombinaison permet d'obtenir des rendements quantiques internes raisonnablement élevés. Ceci est résumé à la page suivante.

Pour des valeurs plus faibles de λ (gap indirect) : dopage à l'azote



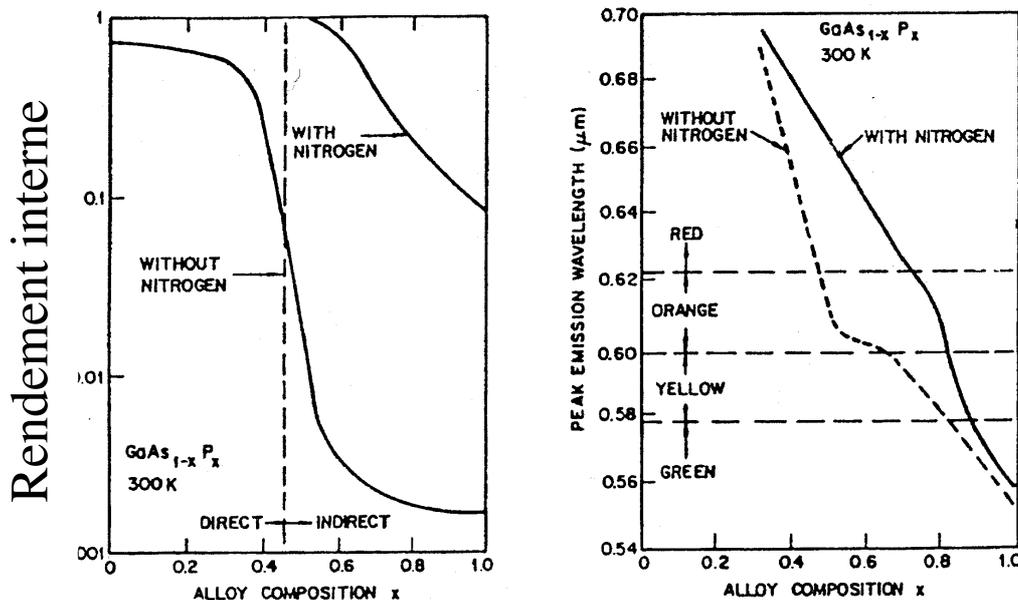
N capte un e^- qui attire un trou : exciton lié

↪ système confiné : $\Delta k \Delta x = 2\pi$

où $\Delta x < a \longrightarrow \Delta k > 2\pi/a$

↪ plus de règle de sélection

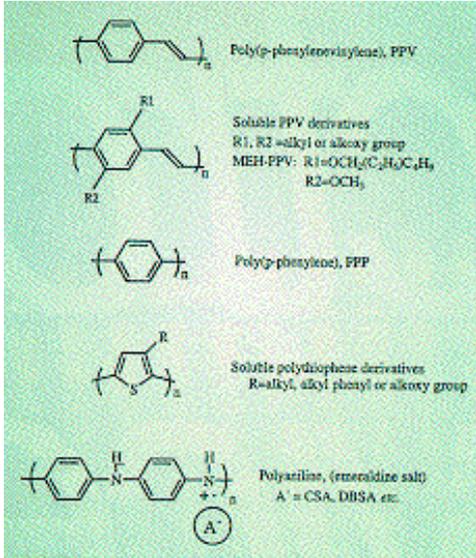
La figure de gauche montre la variation du rendement quantique interne de la DEL de GaAsP en fonction de la proportion de phosphore. On voit que ce dernier chute fortement au-delà de $x = 0.45$ (notez que l'échelle verticale est logarithmique). La figure montre aussi le rendement quantique interne en présence d'azote ; on constate que la suppression de la règle de transition $\Delta k = 0$, permet d'augmenter significativement l'efficacité d'une DEL à gap indirect. Le graphe de droite donne la longueur d'onde d'émission en fonction de la proportion d'azote. On notera la cassure dans la courbe correspondant au S-C non-dopé ; celle-ci est bien entendu causée par le changement du gap direct au gap indirect. On remarque également que la longueur d'onde augmente en présence d'azote, ce qui s'explique par le fait que les recombinaisons se font maintenant entre électrons et trous liés auxquels correspond une énergie de liaison inférieure à celle du gap naturel du matériau.



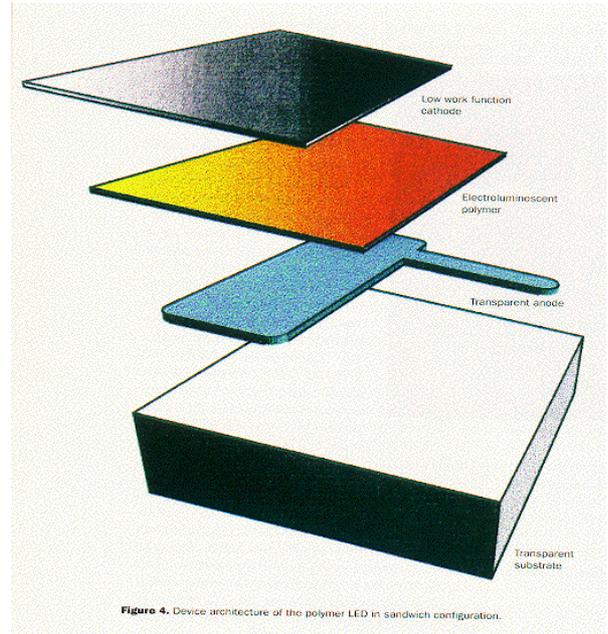
Cette brève discussion concernant le choix des matériaux pour les DEL ne fait que nous donner une petite indication du type de techniques auxquelles on peut recourir pour surmonter des obstacles a priori infranchissables avec les S-C. On voit que le secret du progrès dans la maîtrise de la lumière par les S-C vient principalement de l'ingénierie des matériaux (p.ex. la possibilité de fabriquer des alliages ternaires ou quaternaires, la possibilité de doper les S-C, etc...). Avec les progrès réalisés récemment les DEL permettent de couvrir un large spectre de longueur d'émission comprenant la quasi totalité du visible et l'infrarouge proche (pour les applications télécoms entre autre). Pour donner une idée de l'avancement des recherches dans ce domaine, la page suivante est consacrée à la présentation d'une DEL en matériau polymère. Ces matériaux ne forment pas de cristaux mais ont néanmoins une structure périodique (uni-dimensionnelle) leur conférant des propriétés semi-conductrices (présence de bandes interdites de largeur de l'ordre de l'eV). Ces matériaux sont donc photo-luminescents. Le simple passage d'un courant y crée une excitation (ils ne nécessitent pas de jonction p-n), c'est pourquoi leur structure est si simple (cf page suivante). De plus, la flexibilité (élastique) de ces matériaux rend leur utilisation très souple (pas de substrat rigide encombrant, etc..).

Applications futures : les polymères semi-conducteurs

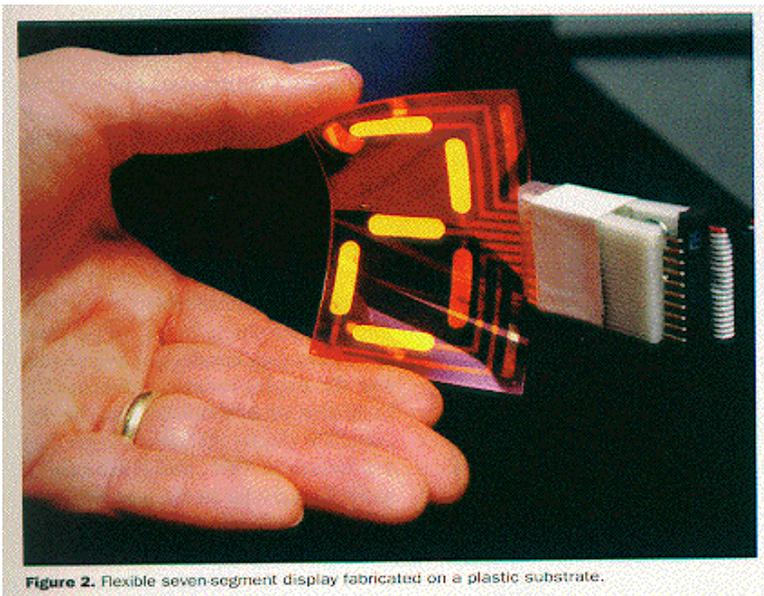
Exemples de matériaux



Structure des DEL en polymère



Application à l'affichage



Rendement quantique et luminescence

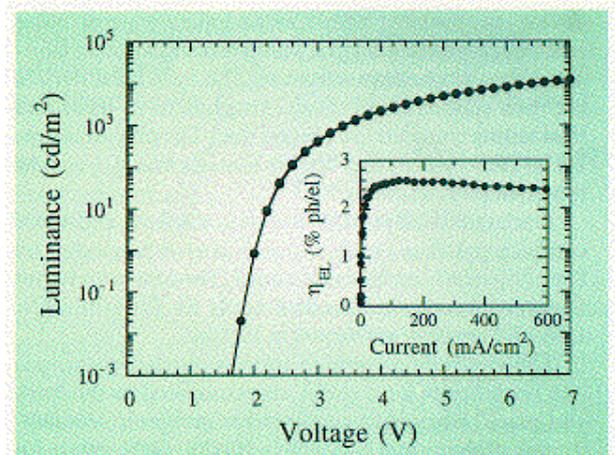


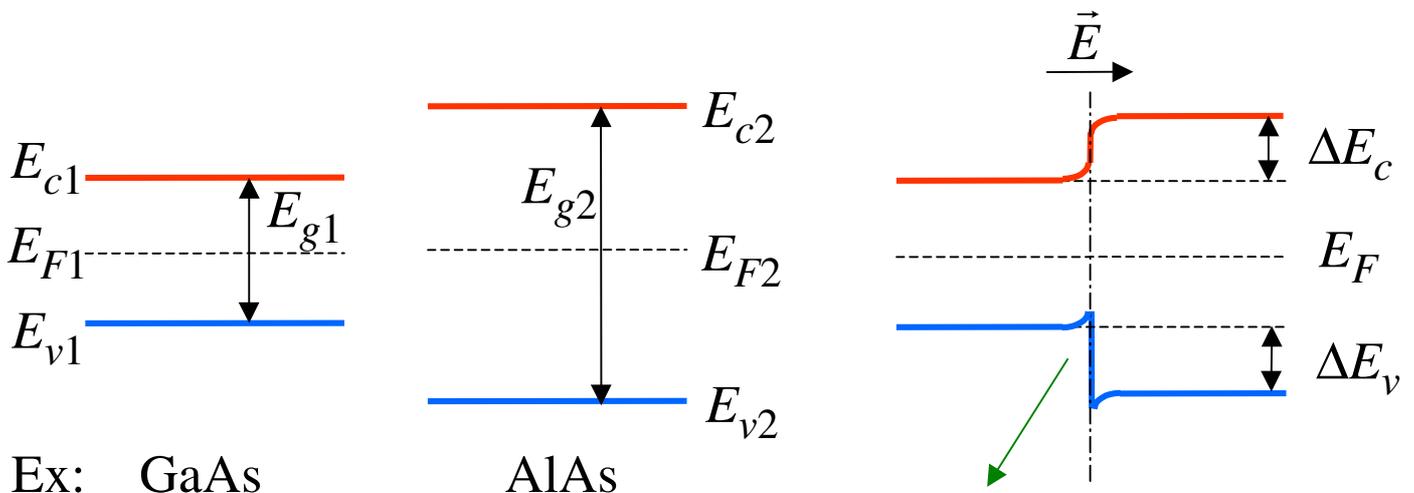
Figure 3. Typical data of luminance and quantum efficiency of MEH-PPV LEDs.

La figure ci-dessus à droite montre le désavantage principal des matériaux polymères : leur rendement quantique est relativement faible par rapport à celui des jonctions *p-n* des cristaux S-C classiques.

Les LED à hétérojonction double

L'hétérojonction

Le terme hétérojonction désigne la jonction entre deux semi-conducteurs de gaps différents. Il s'agit d'une jonction épitaxiée, c'est à dire, obtenue par croissance directe d'un cristal semi-conducteur sur un plan réticulaire de surface d'un autre cristal. Une telle croissance ne peut se faire que si les constantes de réseau des deux S-C sont suffisamment proches, de manière à éviter la présence de défauts (p. ex. dislocations) à la jonction. Une hétérojonction peut, par exemple, être réalisée en faisant croître un cristal de AlAs sur du GaAs car leur constantes de réseau sont toutes les deux de l'ordre de 5.65 Å. Le diagramme de bande de l'hétérojonction résultante est représenté schématiquement ci-dessous. Par définition, les niveaux de Fermi des deux composés s'alignent à l'équilibre (mise à niveau du potentiel électrochimique). Dès lors, puisque le gap de l'AlAs est plus grand que celui du GaAs, les bandes de conduction et de valence se raccordent en formant deux marches de potentiel. Puisque les constantes de réseau des deux cristaux ne sont pas exactement les mêmes, la densité de porteurs n'est pas la même de part et d'autre de l'hétérojonction (il y a plus de porteurs du côté de la constante de réseau la plus faible, le rapport des densité étant donné par le rapport de constantes de réseau au cube). La diffusion naturelle des porteurs tend alors à homogénéiser la concentration de porteur (le porteurs vont là où ils sont le moins nombreux au départ; vers l'AlAs dans l'exemple ci-dessous) et crée ainsi une charge d'espace qui courbe les bandes de conduction et de valence au voisinage de la jonction. Les marches de potentiel ainsi créés ont une hauteur donnée approximativement par la moitié de la différence des gaps des deux S-C (le calcul rigoureux des différences de potentiel dépasse de loin le cadre de ce cours).



Région de charge d'espace due à la différence de constante de réseau (cf. différence de densité)

$$\Delta E_c \approx \Delta E_v \approx (E_{g2} - E_{g1})/2$$

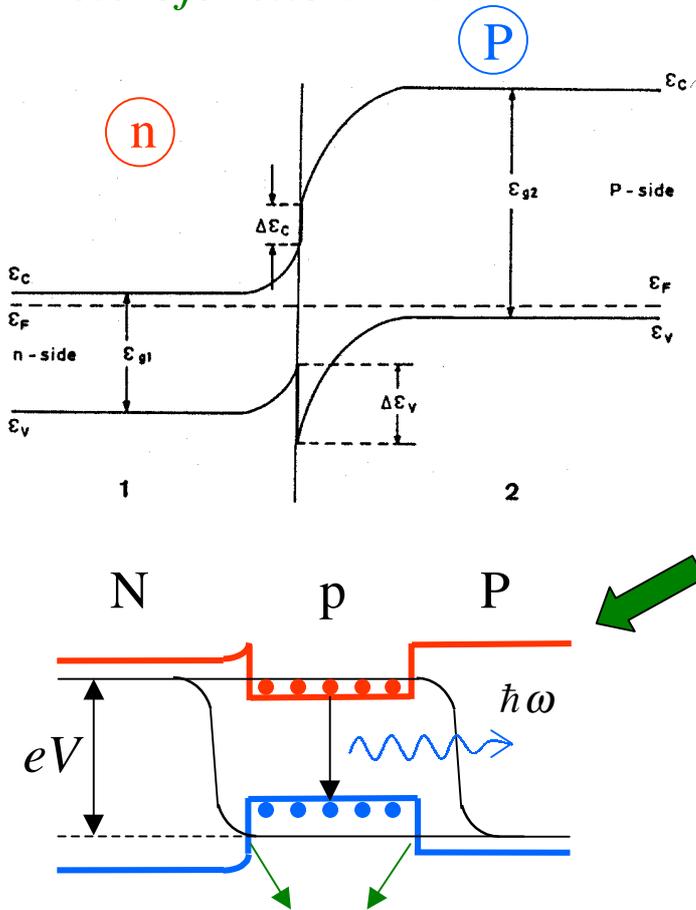
Lorsque, l'hétérojonction est réalisée entre deux semi-conducteurs dopés n et p , les marches de potentiel se rajoutent à la courbure des bandes due à la charge d'espace de la jonction p - n classique. Par convention, la région de grand gap sera notée par une lettre majuscule (N ou P) tandis que pour la région de petit gap la notation usuelle est conservée (n ou p).

L'hétérojonction est habituellement utilisée en paire appelée double hétérojonction. Dans l'exemple ci-dessous, nous donnons un exemple de double hétérojonction N-p-P en AlAs-GaAs-AlAs. Le diagramme de bande est donné pour la double hétérojonction non polarisée et polarisée. On constate dans le cas de la jonction polarisée que les bandes de conduction et de valence forment toutes deux un puits de potentiel pour leurs porteurs respectifs. Ces puits de potentiel créent un confinement des porteurs minoritaires injectés par polarisation de la jonction. La concentration de porteurs libres (e^- et trous) est donc élevée à cet endroit et le taux de recombinaison radiative en sera d'autant plus grand.

Notons que dans une jonction $p-n$ classique, les porteurs minoritaires, par exemple les électrons, sont injectés du côté p où peu de trous libres sont présents pour la recombinaison. Dans l'hétérojonction double, on met en présence les électrons et les trous en grand nombre. En plus de cet avantage, qui comme on le verra est surtout intéressant pour le laser semi-conducteur, les photons émis par recombinaison ne sont pas réabsorbés en dehors de leur zone de génération puisque le gap est plus grand que leur énergie dans ces régions (bien entendu, les photons ont l'énergie du petit gap). Il en résulte un très bon rendement optique pour les DEL à double hétérojonction.

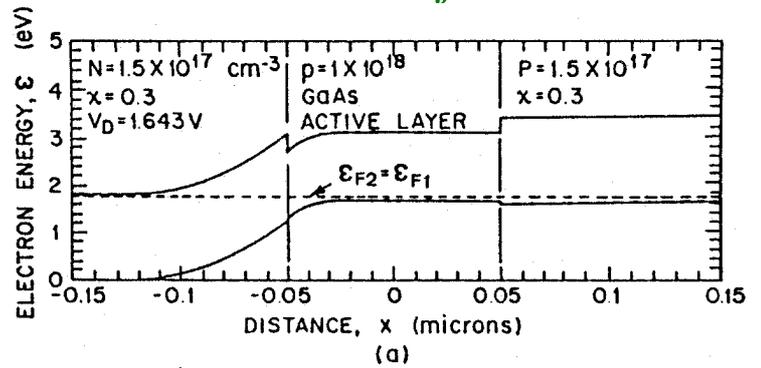
Semiconducteurs dopés

Hétérojonction P-n

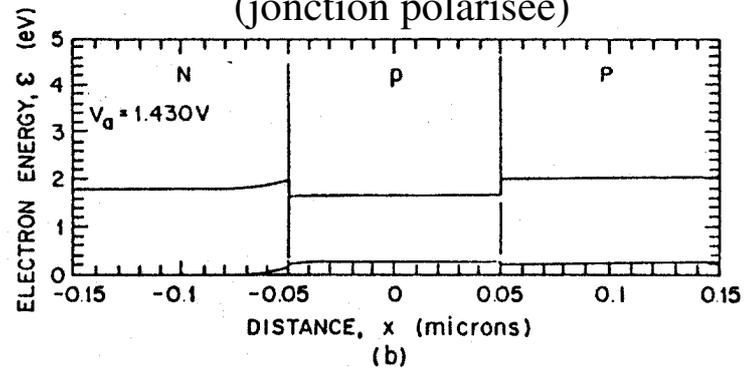


Barrières de potentiel : grand confinement de porteurs

Double hétérojonction



(jonction polarisée)



$$\hbar\omega \approx E_{gp} < E_{gN}, E_{gP}$$

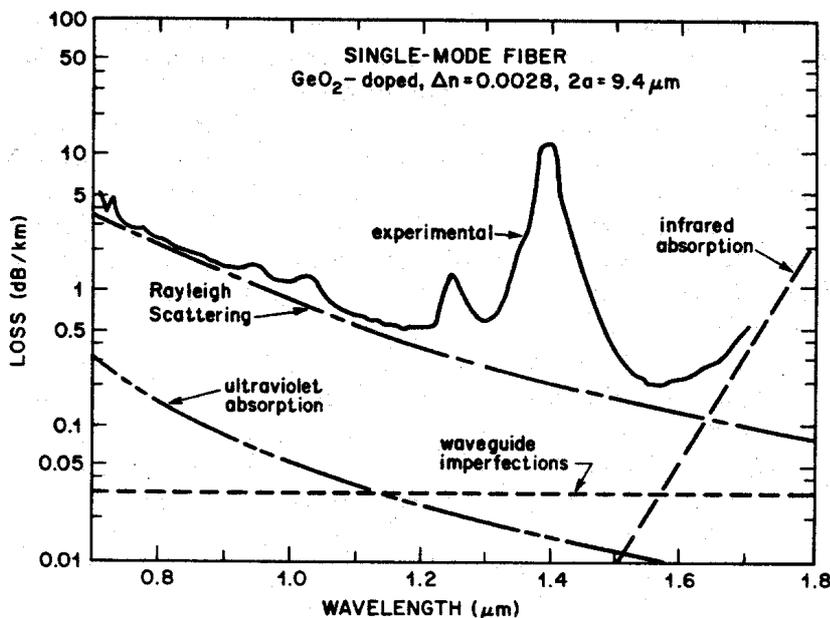
pas de réabsorption $\Rightarrow \eta_{opt} \nearrow \nearrow$

Nous présentons ci-dessous une brève discussion relative au choix des matériaux S-C pour les DEL à double hétérojonction en fonction de leurs applications. Nous nous concentrerons plus particulièrement sur les applications des DEL aux télécommunications optiques.

LED pour les télécommunications à fibre

Nous donnons ci-dessous le spectre d'absorption des fibres optiques de silice utilisées en télécommunication. Comme nous pouvons le voir le minimum d'absorption se situe aux environs de 1.55 microns juste après le pic d'absorption correspondant à l'absorption par les ions OH^- (dus à la présence inévitable de résidu aqueux lors de la fabrication des fibres). Pour des raisons technologiques que nous aborderons plus bas, les sources à 1.55 microns n'existaient pas au début de l'avènement des communications par fibre optique. Les semi-conducteurs permettaient tout au plus d'atteindre l'infrarouge très proche (λ inférieur au micron). Les premières applications des fibres optiques utilisaient donc le minimum relatif du spectre d'absorption se situant à 850 nm (ce minimum n'est pas bien visible sur le spectre ci-dessous car il s'agit du spectre d'absorption d'une fibre moderne). La région autour de ce minimum s'appelle la première fenêtre des télécommunications optiques (1ère fenêtre télécom). Les LED en GaAs permettaient d'atteindre cette fenêtre.

Spectre d'absorption des fibres optiques

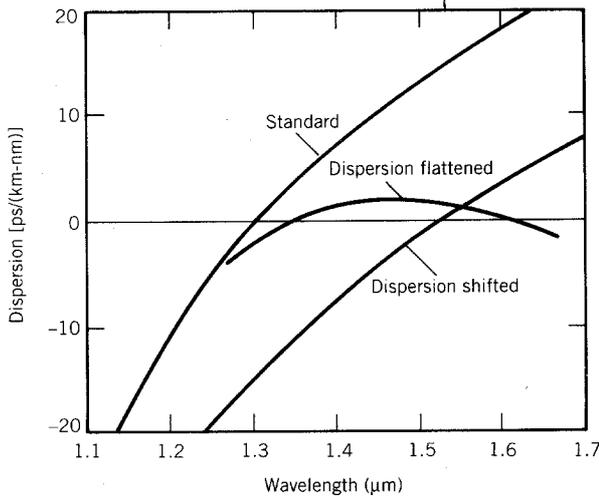


Fenêtres télécom

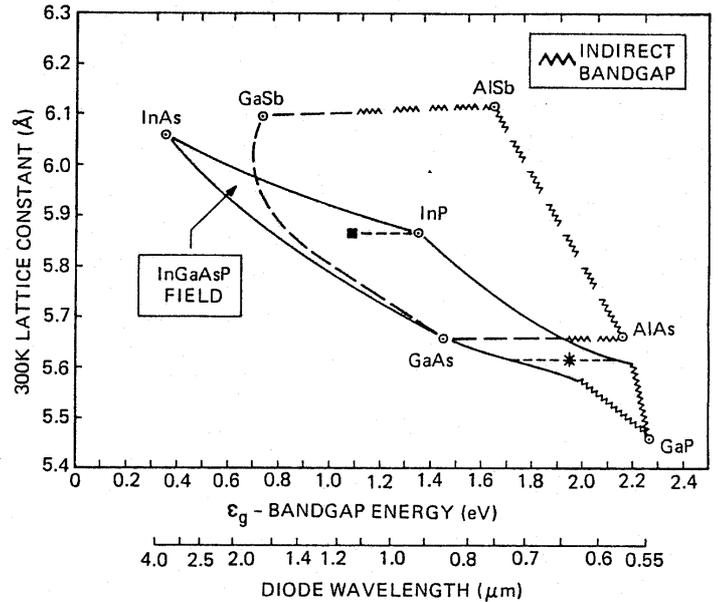
$$\left. \begin{array}{l} \lambda_1 = 0.85 \mu\text{m} \\ \lambda_2 = 1.3 \mu\text{m} \\ \lambda_3 = 1.55 \mu\text{m} \end{array} \right\}$$

Nous donnons ci-dessous, la dispersion chromatique des fibres optiques de silice en fonction de la longueur d'onde. Nous voyons que les fibres standard présentent un zéro de dispersion aux alentours de 1.3 microns. A cette longueur d'onde le spectre d'absorption présente un minimum relatif (avant le pic OH^-), cette zone s'appelle la seconde fenêtre télécom. Comme nous l'avons vu dans l'introduction du cours, il y a intérêt à minimiser la dispersion pour transmettre des données. Dans ce but il a fallu réaliser des sources lumineuses émettant à 1.3 microns. Ceci a pu être réalisé grâce aux progrès de l'ingénierie des matériaux semi-conducteurs et en particulier grâce à la maîtrise des composés ternaires. Nous voyons dans le diagramme donnant la constante de réseau en fonction du gap des S-C que l'utilisation de composés ternaires tels que le In GaAs ou le InAsP permet d'atteindre la seconde fenêtre télécom. La troisième fenêtre télécom a vu le jour lorsque la technologie des fibres a permis de déplacer le zéro de dispersion des fibres à 1.55 microns où l'absorption est à son minimum absolu. Les sources semi-conductrices pour cette longueur d'onde sont actuellement réalisées à partir d'alliages quaternaires (voir la zone InGaAsP du diagramme ci-dessous) permettant d'obtenir une large gamme de dispositifs basés sur les hétérojonctions et autres micro-structures telles que les puits quantiques multiples (voir fin du cours).

Dispersion des fibres optiques



Choix des matériaux



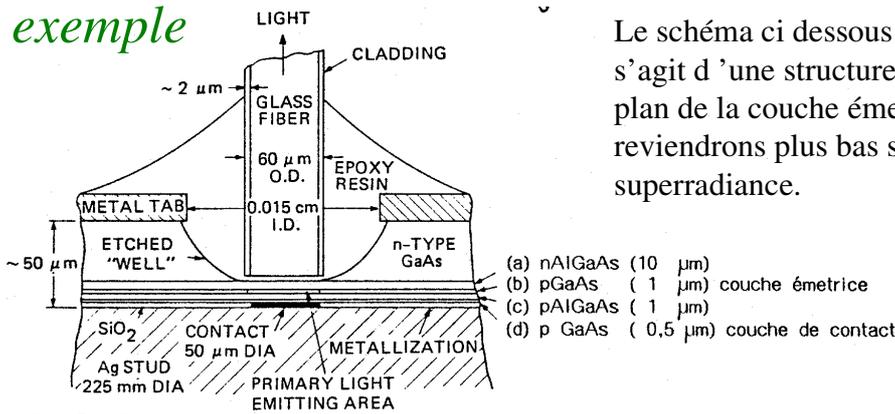
- 1^{ère} fenêtre : $\text{Al}_{1-x}\text{Ga}_x\text{As}$
- 2^{ème} et 3^{ème} fenêtres : $\text{In}_{1-x}\text{Ga}_x\text{As}_y\text{P}_{1-y}$ ($x = 0.45 y$, substrat d'InP)

Pour la première fenêtre télécom, soit 850 nm, le matériau de choix était l'alliage ternaire AlGaAs.

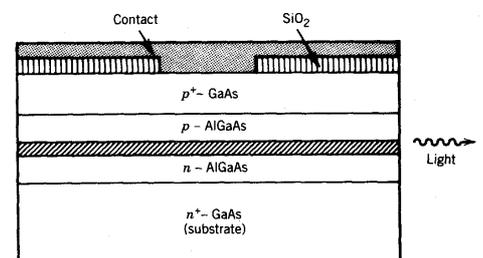
Pour les 2^{ème} et 3^{ème} fenêtres (1.32 et 1.55 microns), un matériau de choix est l'alliage quaternaire $\text{In}_{1-x}\text{Ga}_x\text{As}_y\text{P}_{1-y}$ dans lequel la relation $x = 0.45 y$ est vérifiée de manière à conserver une constante de réseau identique dans toute les couches de la structure à double hétérojonction en partant d'un substrat d'InP pur pour la croissance cristalline (épitaxie).

Nous montrons ci-dessous la structure d'une DEL en AlGaAs pour la première fenêtre télécom. L'épitaxie est réalisée de manière à ce que la couche émettrice de GaAs dopé p d'épaisseur de 1 micron soit placée entre les deux couches de gap supérieur en AlGaAs (double hétérojonction) le tout est épitaxié sur un substrat de silice sur lequel un contact métallique (cathode) est appliqué à l'endroit où la fibre optique est déposée de manière à avoir un maximum d'émission lumineuse dans la fibre (la fibre est une fibre multimode capable de recevoir les rayons de forte inclinaison (cf. DEL = rayonnement isotrope). Afin de récupérer le plus de lumière possible, la fibre est approchée de la zone émettrice grâce à une dépression créée par gravure dans la couche de consolidation supérieure (couche sur laquelle est appliquée l'anode). Le tout est scellé dans de la résine époxy pour assurer la stabilité mécanique de l'ensemble.

exemple



Le schéma ci dessous à droite montre un autre type de structure. Il s'agit d'une structure qui favorise l'émission de photons dans le plan de la couche émettrice (émission longitudinale). Nous reviendrons plus bas sur ce type de structure avec les DEL à superradiance.

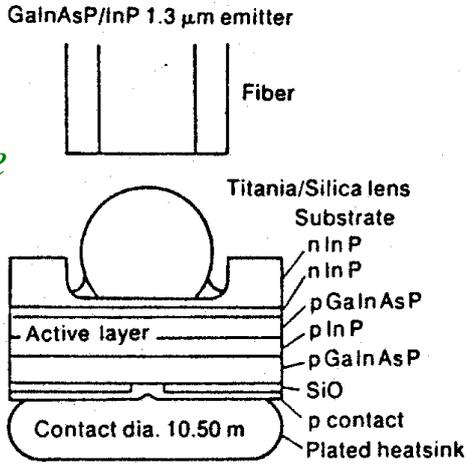


(LED à émission longitudinale)

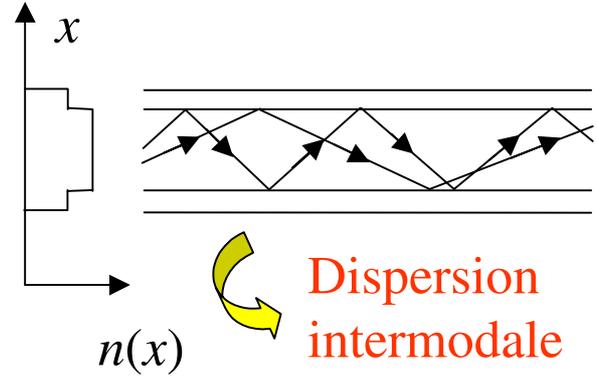
Problèmes liés à l'ouverture angulaire d'émission :

Ci-dessous à gauche nous voyons la structure d'une DEL basée sur l'utilisation d'une lentille sphérique de silice pour améliorer le rendement optique du système (il s'agit d'une DEL pour la 2^{ème} fenêtre télécom). Comme mentionné plus haut, un gros désavantage des DEL est l'isotropie de leur émission qui nécessite l'utilisation de fibres multimodales en raison de leur grande capacité angulaire de réception des rayons (une fibre monomode ne recevrait qu'une infime partie du rayonnement émis).

*Lentille
sphérique*



Utilisation de fibres multimodes

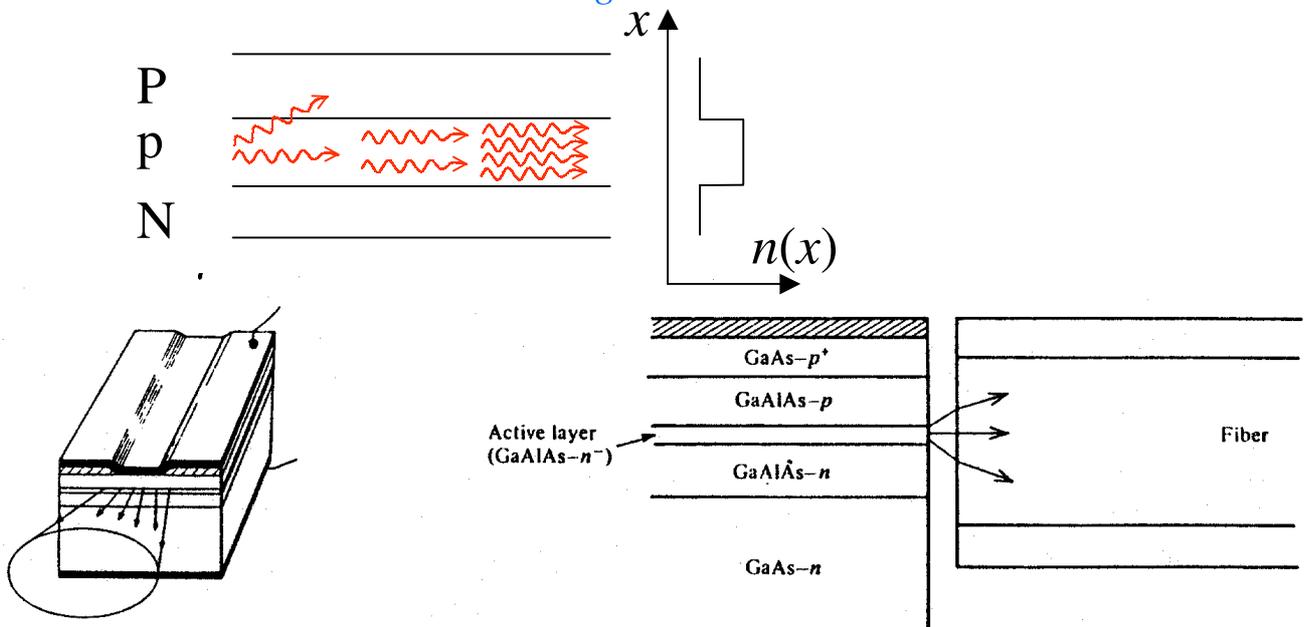


Le problème des fibres multimodales est que chaque angle de rayon guidé correspond à une vitesse de propagation différente de la lumière. L'information codée par la lumière sera donc facilement distordue car une impulsion lumineuse verra son énergie distribuée selon des rayons de vitesses différentes et sera donc élargie indépendamment de la dispersion chromatique de la silice, il s'agit de la dispersion intermodale (celle-ci est en général beaucoup plus élevée que la dispersion chromatique).

Amélioration : LED à superradiance

La situation peut être fortement améliorée en utilisant une DEL à superradiance. Le principe en est le suivant : si l'on polarise la jonction avec une $d\phi$ supérieure au gap, on aura amplification optique dans la zone émettrice de petit gap. Les photons ont donc tendance à se multiplier dans l'axe de la couche émettrice. Si l'on dispose alors les électrodes en bande, les photons seront multipliés sur une grande distance et le rayonnement deviendra intense tout en ayant une bonne directivité car seuls les photons se propageant dans l'axe des électrodes sont multipliés. On peut alors utiliser une fibre faiblement multimode (voire monomode). Ce principe de fonctionnement est très proche de celui du laser.

Polarisation élevée : $\Delta E_F > E_g \rightarrow$ Gain



Les quelques lignes ci-dessous résument les désavantages des DEL : en raison du rayonnement isotrope lié au principe même de la photoluminescence, le rendement optique des DEL sont faibles et les puissances injectées dans les fibres ne dépassent pas quelques μW . De plus la photoluminescence a le désavantage d'avoir une grande largeur spectrale fixée par l'énergie thermique (25 nm à température ambiante). Tout cela fait que les DEL ne sont utilisées que pour les communications à courtes distances (réseaux locaux) pour lesquels les problèmes de puissance lumineuse des sources et de dispersion ne se posent pas.

Problème : Puissance lumineuse injectée limitée à quelques μW



Grande largeur spectrale (cf. énergie thermique $\approx 25 \text{ nm}$)

LED pas utilisées pour communication à longues distances

Seulement pour réseaux locaux $L_{max} \approx \text{qlqs km}$

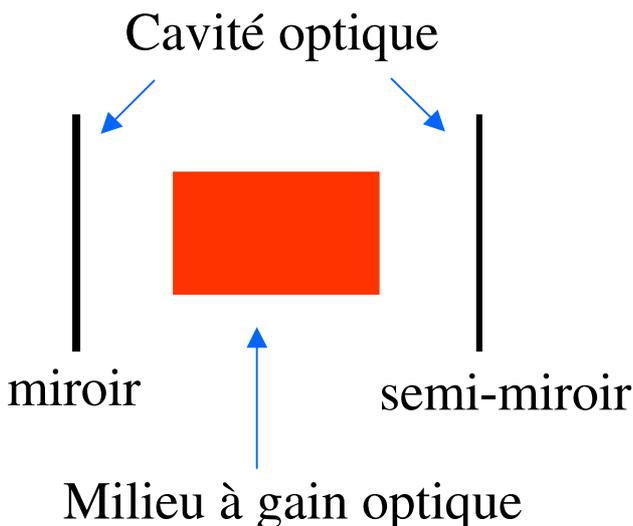
Nous verrons dans le chapitre suivant que le laser à S-C permet d'éviter ces problèmes.

Les Lasers à semiconducteur

Avant d'étudier le laser à S-C proprement dit, voyons tout d'abord le principe de fonctionnement du laser. Un laser est une cavité optique dans lequel se trouve un milieu à gain optique. La cavité peut être, par exemple, formée d'un semi-miroir et d'un miroir parallèles se faisant face de manière à ce que les photons puissent faire des aller-retour en traversant le milieu à gain. Comme nous l'avons vu plus haut, le gain optique peut être réalisé par émission stimulée sur base d'une inversion de population.

Le principe de fonctionnement du laser est le suivant : imaginons un premier photon spontané émis dans le milieu à gain (comme on l'a vu plus haut, l'émission spontanée y est inévitable). Si ce photon est émis dans l'axe de la cavité (perpendiculairement aux miroirs) il sera réfléchi par un des miroirs et reviendra dans le milieu à gain où il sera multiplié, disons N fois. Le groupe de N photons résultants sera à nouveau amplifié lors de son passage suivant dans le milieu à gain après réflexion sur un des miroirs. Il en résultera N^2 photons dans la cavité. En se répétant ce processus conduit à une croissance exponentielle du nombre de photons. Bien entendu la croissance est limitée par le phénomène de saturation de gain vu plus haut. Si l'intensité optique devient trop forte, un équilibre s'établit entre émission stimulée et absorption (annulation de l'inversion de population). Aussi, dans ce raisonnement nous n'avons pas tenu compte du fait que des photons quittaient la cavité à chaque aller-retour au travers du semi-miroir. Ce sont ces photons qui constituent le faisceau laser. Puisque, tous les photons sont identiques au premier photon spontané (cf. Propriétés de l'émission stimulée) ils sont bien entendu tous en phase et le rayonnement laser est parfaitement cohérent (du moins en principe). Naturellement, ce raisonnement est idéalisé. En pratique il n'y a pas qu'un seul photon spontané émis. Cependant, même si l'émission de photons spontanés a lieu de manière continue, ce sont toujours les premiers photons apparaissant dans la cavité qui seront amplifiés préférentiellement (ces premiers photons prendront toute l'énergie contenue dans l'inversion de population en raison du principe de croissance exponentielle des photons). Le faisceau sera donc bien cohérent et l'émission spontanée ne constitue jamais qu'une source de bruit comme pour l'amplificateur optique (l'émission spontanée est source de bruit puisqu'elle génère des photons qui n'ont pas de relation de phase avec les autres). Un faisceau laser n'est donc jamais parfaitement cohérent.

Principe de base du laser



{ Cavité = miroirs parallèles
Gain = émission stimulée

Principe : création d'une onde cohérente par multiplication des photons dans la cavité à partir d'un photon spontané

Historique :

Nous reprenons ci-dessous les différentes étapes qui ont été nécessaires dans le progrès de la physique pour aboutir à l'avènement du laser à S-C. Le premier ingrédient nécessaire a bien entendu été la découverte du photon par Planck et Einstein. Est venu ensuite le concept d'émission stimulée décrit pour la première fois par Einstein. La notion de gain optique n'est apparue que en 1950 avec les travaux de Kastler sur le pompage optique (technique d'inversion de population d'un système d'atomes). En 1954 Townes réalisa le premier MASER (Microwave Amplification by Stimulated Emission of Radiation) sur base de l'excitation des niveaux de vibration de molécules d'ammoniac se trouvant dans une cavité. En 1958 Townes suggère la possibilité de réaliser un Maser optique (le Laser) basé sur le même principe mais utilisant des transitions atomiques. En 1960 le premier laser fut construit utilisant des transitions atomiques de l'ion de Chrome dans une matrice d'alumine (rubis). En 1961 Bernard et Duraffourg (France) propose, sur base du concept de quasi-niveau de Fermi, l'utilisation d'une jonction $p-n$ semi-conductrice pour la réalisation de l'inversion de population. Il n'a fallu attendre qu'un an après cette proposition pour voir fonctionner le premier laser à S-C en 1962 dans plusieurs laboratoires de recherche industriels. Il s'agissait d'une diode (jonction $p-n$) en GaAs fonctionnant à très basse température (77K, nous verrons plus loin pourquoi une telle température). Il a fallu attendre 5 ans pour voir fonctionner la première diode laser à température ambiante. Celle-ci était basée sur le concept essentiel de double hétérojonction.

- 1900 : Rayonnement du corps noir, photon (Planck)
- 1917 : Emission stimulée (Einstein)
- 1950 : Pompage optique, gain optique (Kastler)
- 1954 : Maser à ammoniac, microonde $\lambda = 1.25$ cm (Townes)
- 1958 : Idée du Maser optique, cavité + gain optique (Townes)
- 1960 : 1^{er} Laser, rubis (Al_2O_3 + Chrome) $\lambda = 0.69$ μm
- 1961 : Idée du gain optique dans les S-C, quasi-niveau de Fermi $\Delta E_F > E_g$
(Bernard et Duraffourg)
- 1962 : 1^{er} Laser à S-C, GaAs $\lambda = 0.85$ μm , 77 K (IBM, MIT, GE)
- 1967 : Le laser à hétérojonction, λ variable, 300 K (Kroemer)

Le grand avantage des S-C par rapport aux autres matériaux à gain optique est qu'ils présentent une densité de dipôles électriques (ou porteurs excités) extrêmement élevée (les autres matériaux à gain sont essentiellement des gaz ou des verres ou des cristaux dopés dans lesquels la densité de dipôles électriques est très faible devant les 10^{18} cm^{-3} des S-C). Cela conduit naturellement à des gains optiques élevés sur des longueurs réduites. De plus le pompage optique dans les S-C peut se faire par simple polarisation d'une jonction $p-n$ ce qui est particulièrement facile et efficace par rapport aux techniques de pompage optique classiques.

Avantages des S-C : Densité de dipôles élevée (10^{18} cm^{-3})

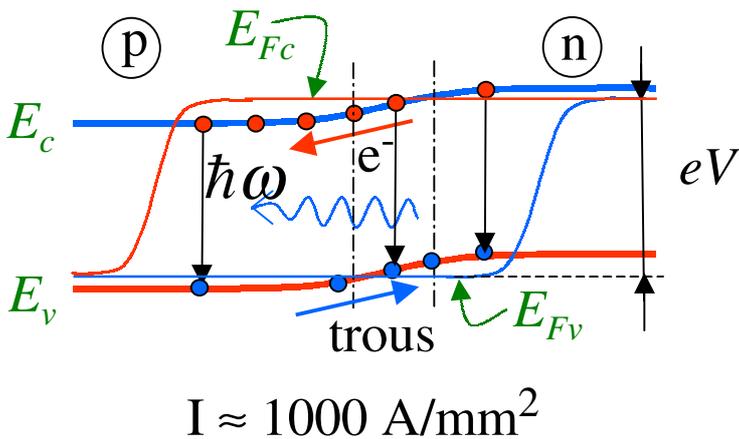


gain élevé \Rightarrow dimensions réduites

Pompage par courant (η élevé, modulable)

Le laser à homojonction

Comme nous l'avons vu plus haut, pour obtenir du gain optique dans une homojonction, il faut polariser la jonction avec une ddp supérieure au gap du S-C. L'inversion de population n'est alors réalisée que dans la zone de charge d'espace, là où à la fois les quasi-niveaux de Fermi E_{Fc} et E_{Fv} sont respectivement supérieur et inférieur au bas de la bande de conduction et au sommet de la bande de valence (entre les lignes pointillées du schéma ci-dessous). Cette zone est étroite et sa largeur dépend fortement de la ddp appliquée. En fait elle n'existe que si la ddp appliquée est suffisamment élevée. En pratique l'obtention de cette zone d'inversion de population requiert une telle ddp que le courant associé est très grand ce qui a pour effet de chauffer fortement la jonction. L'échauffement est si fort que pour conserver les propriétés de la jonction intacte, il faut refroidir le système à l'azote liquide (77 K). Une telle solution n'est évidemment pas acceptable pour les applications du laser à S-C.



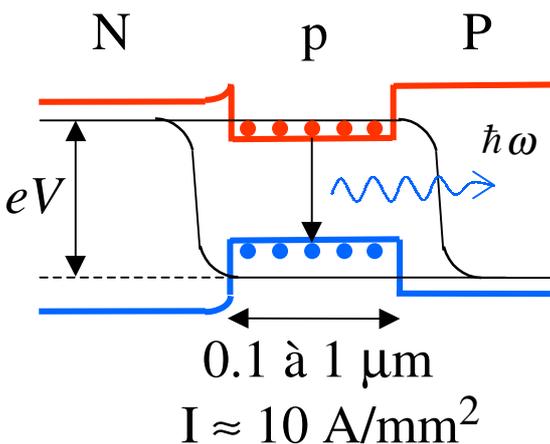
Problèmes :

- Zone active limitée (de taille variable)
- Courant élevé (échauffement)

↪ Refroidi à l'N liquide (77K)

Le laser à hétérojonctions

Le laser à S-C est devenu accessible pour les applications grâce à l'invention de la double hétérojonction. Comme nous l'avons déjà vu pour les DEL, la structure en double hétérojonction crée une double barrière de potentiel pour les porteurs minoritaires injectés. Le confinement de porteurs qui en résulte conduit à une grande facilité d'inversion de population. La ddp appliquée ne doit pas être très importante (elle doit juste excéder le petit gap de la structure) et la diode laser peut fonctionner à température ambiante. De plus la largeur de la zone d'inversion de population est maintenant fixée lors de la fabrication du dispositif et ne dépend donc plus de la ddp, le contrôle du fonctionnement de la diode laser en est donc fortement simplifié.



Avantages :

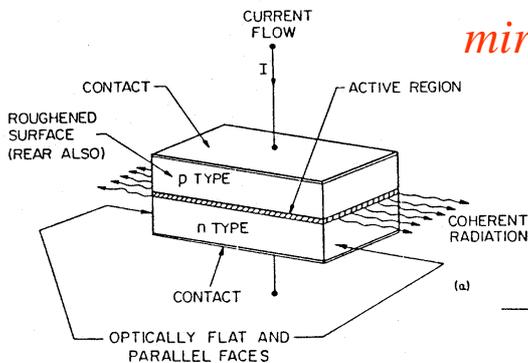
- Zone active déterminée par construction, confinement des porteurs minoritaires

↪ courant faible (300K)

Structure de base des lasers à S-C

Les schéma ci-dessous illustrent la structure de base de lasers à S-C à homojonction et double hétérojonction. La structure est particulièrement simple dans la mesure où la zone à gain est formée par la jonction p-n et que la cavité est simplement formée par les faces parallèles clivées du cristal se trouvant dans les plans perpendiculaires au plan de la jonction (les deux autres faces parallèles ne sont pas clivées ; au contraire elles sont rendues rugueuses de manière à minimiser leur réflectivité afin d'éviter d'avoir une deuxième direction de « lasage » (on désire en général un rayonnement uni-directionnel)). La réflectivité des miroirs ainsi obtenus est donnée par la loi de la réflexion de Fresnel et est de l'ordre de 30%. Un des grands avantages supplémentaires de la structure en double hétérojonction, est l'effet de guidage qu'elle offre naturellement. En effet, les indices de réfraction des couches à grand gap sont plus faibles que celui de la couche à petit gap. La zone à gain est donc également une couche guidante. Il s'en suit que même les photons émis en dehors de l'axe de la cavité peuvent être guidés et donc être également amplifiés pour participer au rayonnement laser (ce qui n'est pas le cas en l'absence de guidage puisque tout photon émis hors axe est évacué de la zone de gain).

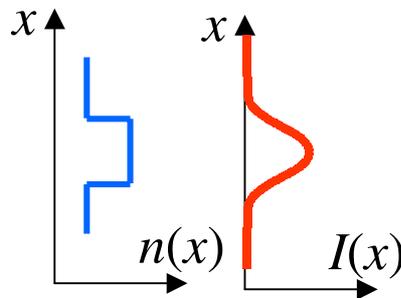
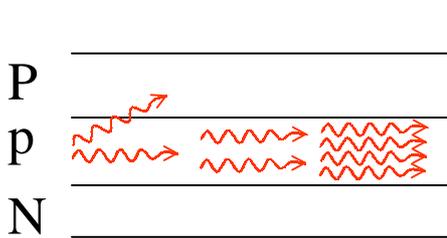
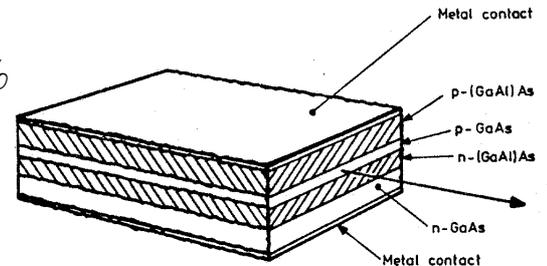
homojonction



miroirs = surfaces clivées

$$R = \left(\frac{n-1}{n+1} \right)^2 \approx 30\%$$

hétérojonctions

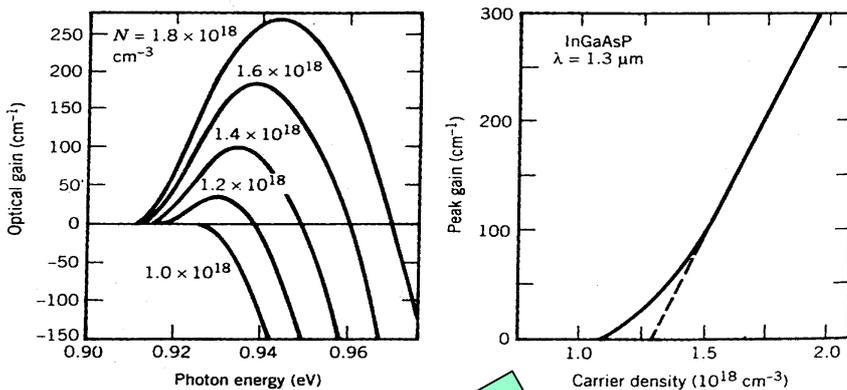


*Avantage supplémentaire :
effet de guidage, $\Delta n \approx 5\%$*

Modèle phénoménologique simplifié du laser

Dans le but d'établir un modèle simplifié d'une diode laser à S-C, nous commençons par analyser la manière de modéliser le gain optique en fonction du degré d'excitation de la jonction. A cette fin, nous avons représenté à la figure de gauche ci-dessous les courbes de spectre de gain en fonction de la densité de porteurs injectés. La densité de porteurs injectés est bien entendu fonction du courant traversant la jonction. Pour établir le lien entre ces deux quantités nous avons écrit ci-dessous à droite l'équation de bilan des porteurs. Celle-ci exprime simplement que le nombre d'électrons de la bande de conduction diminue avec une constante de temps τ_c (temps de vie des électrons dans la bande de conduction) et augmente avec le courant injecté (I étant le courant c'est à dire le nombre d'électrons passant la jonction par unité de temps, on divise ce nombre par la charge de l'électron e et par le volume de la zone de gain pour avoir la contribution du courant à la variation de densité de porteurs dans celle-ci). On arrive donc à établir la relation de proportionnalité existant entre le courant injecté et la densité de porteurs. La figure de droite montre comment le gain maximal (maximum des courbes spectrales de gain) varie en fonction de la densité de porteurs. Nous voyons que la variation comporte un seuil et que peu après ce seuil le gain est une fonction linéaire de la densité. Dans notre modèle simplifié du laser nous négligerons la courbure de la fonction autour du seuil et nous supposerons que le lien entre gain g et densité de porteurs n_e est linéaire selon la loi écrite ci-dessous, où σ est appelée la section efficace de gain optique et n_o la densité de seuil. Des valeurs typiques de ces deux quantités sont données ci-dessous.

Gain dans les S-C



Bilan des porteurs

$$\frac{dn_e}{dt} = -\frac{n_e}{\tau_c} + \frac{I}{eVol} = 0$$

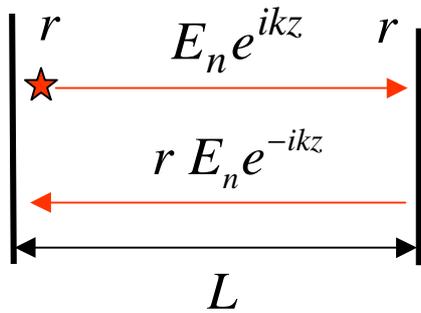
$$n_e = \frac{I\tau_c}{eVol}$$

$$g = \sigma(n_e - n_o)$$

$$\left\{ \begin{array}{l} \sigma \approx 10^{-16} \text{ cm}^2, \text{ section efficace de gain} \\ n_o \approx 10^{18} \text{ cm}^{-3}, \text{ densité de seuil} \end{array} \right.$$

Cavité : résonances du champ E-M

Nous devons ensuite modéliser le rôle de la cavité. Pour cela nous passons à une description classique du champ électromagnétique de la lumière et nous décrivons la variation temporelle de ce champ en terme de son évolution passage après passage dans la cavité. Le champ sur un passage dans la cavité peut s'exprimer comme $E_n \exp(ikz)$ où E est l'amplitude et k le vecteur d'onde. Pour le trajet dans l'autre sens nous avons donc la forme $r E_n \exp(-ikz)$ où r est la réflectivité du miroir (face clivée $r = 30\%$). En prenant en compte les réflexions sur les deux faces clivées, les pertes par absorption dans le matériau S-C (soit α_{int} le coefficient de perte par unité de longueur en terme d'intensité lumineuse) et le déphasage total de l'onde sur un aller-retour, nous pouvons exprimer le champ en un point de la cavité (ici l'étoile indiquée sur la figure ci-dessous) en fonction du champ au même point un aller-retour plus tôt, soit la formule donnant E_{n+1} en fonction de E_n ci-dessous (n représentant le nombre d'aller-retours effectués par l'onde électromagnétique). Le facteur rr représente l'effet des réflexions, $\exp(-\alpha_{int}2L)$ représente les pertes internes exprimées en intensité sur un aller-retour dans la cavité ; exprimées en amplitude les pertes sont donc bien $\exp(-(\alpha_{int}/2)2L)$, de même pour l'effet du gain nous trouvons $\exp(-(g/2)2L)$ où g est le gain en intensité (ou, de manière équivalente, en nombre de photons).



$$R = \left(\frac{n-1}{n+1} \right)^2 = r^2$$

Évolution du champ E-M sur un tour

$$E_{n+1} = E_n r r e^{ik2L} e^{-\frac{\alpha_{\text{int}}}{2}(2L)} e^{\frac{g}{2}(2L)}$$

Nous imposons ensuite que le régime du laser soit stationnaire (régime d'équilibre).

Régime stationnaire : $E_{n+1} = E_n$

Pour que le laser fonctionne en régime stationnaire il faut donc satisfaire deux conditions. L'une provient du terme de phase de la relation ci-dessus. Elle exprime simplement le fait que l'onde lumineuse dans la cavité est à la résonance. C'est à dire qu'après un aller-retour dans la cavité les photons se retrouvent avec la même phase de façon à ne pas se détruire par interférences destructives tour après tour dans la cavité. Le vecteur d'onde étant fonction de la fréquence de la lumière, cette condition de phase est en fait une condition sur la fréquence de la lumière émise par le laser. On voit que le laser ne peut « laser » que sur des fréquences bien déterminées fonction de la longueur de la cavité L et de son indice de réfraction n . Ces fréquences sont appelées les modes du lasers. Les expressions $\Delta\omega$ et $\Delta\nu$ donnent la distance entre les modes du laser en termes de fréquence angulaire et fréquence, respectivement.

1. Condition de phase : $e^{2ikL} = 1 \rightarrow 2kL = 2m\pi$

$\frac{n\omega}{c} L = m\pi \rightarrow \omega_m = \frac{m\pi c}{nL} \quad \Delta\omega = \frac{\pi c}{nL}, \quad \Delta\nu = \frac{c}{2nL}$

La deuxième condition de stationnarité ou d'équilibre du laser est fournie par l'égalisation des amplitudes (ou la conservation du nombre de photons d'un tour à l'autre dans la cavité). Cette condition exprime donc un équilibre entre les pertes internes de la cavité, les pertes par les miroirs et le gain optique. Les deux premières étant fixées par le dispositif, cette condition est une condition sur le gain : pour qu'il y ait équilibre il faut que le gain soit égal aux pertes totales de la cavité. Les pertes totale de la cavité sont exprimées ci-dessous au travers du coefficient α_{cav} qui globalise les pertes sous forme de pertes distribuées par unité de longueur (α_{mir} représente les pertes dues au miroirs comme si celles-ci étaient distribuées sur toute la longueur de la cavité comme les pertes internes d'absorption).

2. Condition d'amplitude : $R e^{-\alpha_{\text{int}}L} e^{gL} = 1$

$\ln(R) - \alpha_{\text{int}}L + gL = 0 \Rightarrow g = \alpha_{\text{int}} + \underbrace{\ln(1/R)/L}_{\alpha_{\text{mir}}} = \alpha_{\text{cav}}$

Cette condition sur le gain fait immédiatement entrevoir le fait que le laser présentera un seuil d'excitation avant de se mettre à « laser ». Il faudra en effet que n_e (c'est à dire le courant I) soit suffisamment élevé que pour que g puisse atteindre α_{cav} . Pour cette raison la condition $g = \alpha_{\text{cav}}$ porte le nom de condition d'oscillation du laser (oscillation signifie ici oscillation du champ électromagnétique de la lumière, c'est à dire « lasage »).

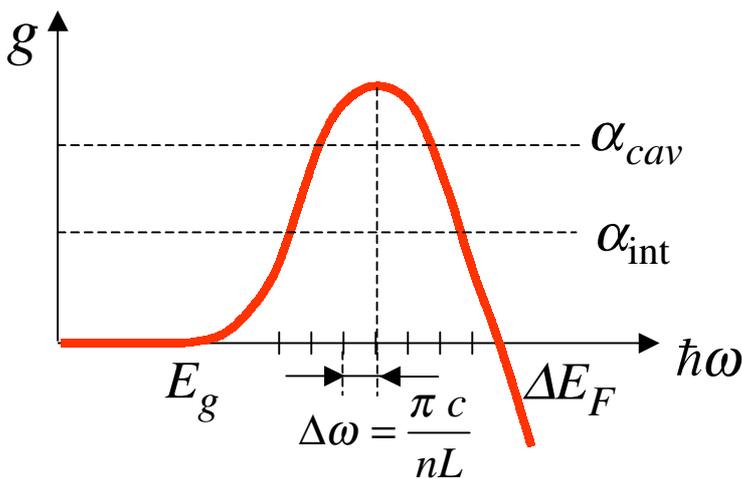
Effet de seuil :
$$\left\{ \begin{array}{l} g = \sigma(n_e - n_o) \quad \text{où} \quad n_e = \frac{I\tau_c}{eVol} \\ g = \alpha_{cav} \end{array} \right. \quad (\text{condition d'oscillation})$$

A l'aide des relations établies précédemment il est facile d'exprimer la condition d'oscillation en terme du courant d'excitation de la jonction. Soit I_s le courant de seuil du laser.

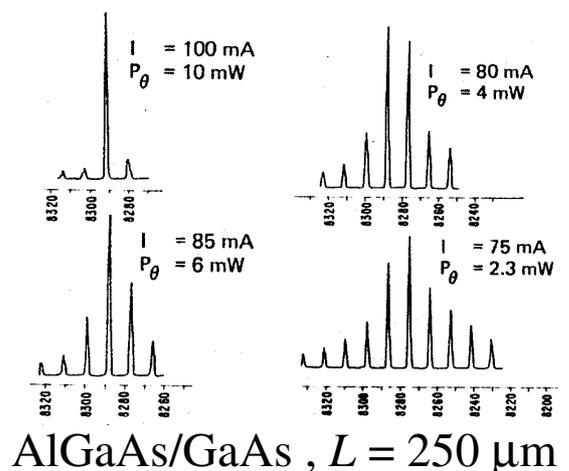
$$\sigma\left(\frac{I_s\tau_c}{eVol} - n_o\right) = \alpha_{cav} \quad \longrightarrow \quad I_s = \frac{eVol}{\tau_c} \left(\frac{\alpha_{cav}}{\sigma} + n_o\right)$$

La figure ci-dessous donne le spectre de gain ainsi que les pertes internes d'absorption α_{int} et les pertes globales de la cavité α_{cav} . Les modes du lasers sont également représentés par des tirets sur l'axe des absisses. On voit que plusieurs modes peuvent a priori remplir la condition d'oscillation, c'est à dire peuvent « laser ».

Pour comprendre comment les choses se passent en pratique il faut faire le raisonnement suivant. Imaginons que le courant augmente progressivement à partir de zéro. Tant que le gain reste inférieur aux pertes internes, les photons ne sont pas amplifiés dans la cavité et la diode laser fonctionne en régime de rayonnement spontané essentiellement. Une fois que le gain devient supérieur au pertes internes, les photons des modes de gain le plus élevés commencent à être amplifiés dans la cavité, mais pas suffisamment que pour compenser les pertes de photons par les miroirs. La diode laser fonctionne alors comme une DEL à superradiance ; on commence à avoir une meilleure directivité et le nombre de modes émettants est plus réduit qu'en régime spontané. Lorsque le courant est tel que le mode le plus amplifié remplit la condition d'oscillation, le système fonctionne en mode laser. Dans ce régime le gain est supérieur aux pertes et a priori le nombre de photons pourraient croître à l'infini. Bien entendu il n'en est rien puisque plus l'intensité lumineuse dans la cavité augmente, plus le gain devient faible (cf. Saturation de gain). Il y a donc un équilibre naturel qui s'établit de telle sorte que le gain compense tout juste les pertes totales (la courbe spectrales ci-dessous n'est donc qu'une vue de l'esprit). Donc, a priori, un seul mode peut « laser » dans le système : celui qui correspond au gain le plus élevé.



Spectre d'émission des lasers



Les spectres ci-dessus à droite, illustrent le raisonnement que nous venons d'établir. Pour une faible excitation (75 et 80 mA) le système fonctionne en régime spontané et ne sélectionne donc pas les modes. Pour une excitation plus élevée (85 mW) on entre en régime de superradiance où les modes commencent à être sélectionnés selon leur gain relatif. Au courant de 100mW le mode de gain maximal est en régime de lasage et est donc clairement sélectionné par rapport aux autres. Toute fois pour des raisons dont l'explication dépasse le cadre de ce cours, les autres modes sont toujours présents et sont même renforcés si le courant augmente encore, ce qui peut poser un problème en pratique puisque l'on cherche en général à avoir un rayonnement monochromatique (c'est à dire de largeur spectrale aussi faible que possible). Nous verrons par la suite comment remédier à ce problème en pratique.

Calcul de la puissance lumineuse émise :

On peut facilement calculer la puissance émise par un laser à S-C à partir des équations de bilan du nombre de photons dans la cavité optique et du nombre d'électrons dans la jonction.

La variation du nombre de photons est donnée par le taux de radiation cohérente et spontanée ainsi que par le nombre de photons perdus dans la cavité par unité de temps. Pour exprimer cette dernière quantité, on introduit le concept de durée de vie moyenne des photons dans la cavité. Il s'agit du temps $\tau_p = v \alpha_{cav}$ calculé ci-dessous à droite sur base de la décroissance temporelle exponentielle du nombre de photons dans la cavité en l'absence d'excitation (les 4 lignes de calculs indiquent clairement le lien entre τ_p , les pertes internes et les pertes aux miroirs : on assimile simplement la décroissance du nombre de photons tour après tour à une décroissance exponentielle de constante τ_p).

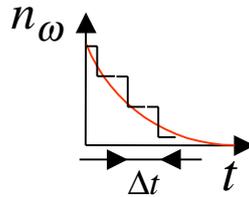
L'équation de bilan des photons dans la cavité est donc donnée par la formule ci-dessous encadrée en ligne pointillée. Pour la simplicité nous supposons que l'émission spontanée est négligeable, soit $R_{sp} = 0$. On exprime alors le taux de radiation cohérente à l'aide de la définition du coefficient d'absorption α ou du gain g associé à la jonction. On obtient alors l'équation encadrée en ligne continue dans laquelle il ne faut pas perdre de vue que g est une fonction de la densité d'électrons par l'intermédiaire de la section efficace de gain σ et de la densité de seuil n_o .

- Bilan du nombre de photons

$$\frac{dn_\omega}{dt} = R_c + R_{sp} - \frac{n_\omega}{\tau_p}$$

$$R_c = -\alpha v n_\omega = g v n_\omega, \quad R_{sp} \approx 0$$

$$\frac{dn_\omega}{dt} = g v n_\omega - \frac{n_\omega}{\tau_p} \quad \left(g = \sigma (n_e - n_o) \right)$$



Perte à chaque tour

$$R^2 e^{-2\alpha_{int} L} = e^{-\Delta t / \tau_p}$$

$$\text{où } \Delta t = \frac{2L}{c/n} = \frac{2L}{v}$$

$$\ln(1/R) + \alpha_{int} L = L / (v \tau_p)$$

$$\frac{1}{\tau_p} = v \left(\frac{1}{L} \ln \frac{1}{R} + \alpha_{int} \right) = v \alpha_{cav}$$

Le bilan du nombre d'électrons est également très simple à établir : la variation du nombre d'électrons dans la bande de conduction (porteurs minoritaires) est donnée par le nombre d'électrons qui se recombinent de manière non radiative et spontanée par unité de temps. Cette variation est donnée par la décroissance exponentielle naturelle de constante τ_c définissant le temps de vie des électrons dans la bande de conduction en l'absence de photons (ou ce qui revient au même en l'absence d'excitation). En présence d'excitation il faut tenir compte du fait que le nombre d'électrons croît de manière proportionnelle au courant d'excitation et du fait que chaque émission stimulée de photon correspond à la disparition d'un électron de la bande de conduction. Cette dernière contribution explique la présence du terme $-g v n_\omega$ dans l'équation de bilan.

- Bilan du nombre d'électrons

$$\frac{dn_e}{dt} = -\frac{n_e}{\tau_c} + \frac{I}{e Vol} - g v n_\omega$$

perte de porteurs due aux recombinaisons non-radiatives (+ spontané)

nbre d' e^- entrants

Recombinaisons stimulées

Modèle simplifié de la dynamique du laser

Les équations de bilan établies ci-dessus fournissent un modèle simplifié de la dynamique des lasers à S-C. Elles sont répétées ci-dessous en explicitant le rôle de la densité de porteurs dans le gain g de la jonction. Sous cette forme on voit que les équations forment un système d'équations différentielles non linéaires couplées. La non-linéarité provient des produits $n_e n_\omega$ (non-linéarité du 2^{ème} ordre). Il s'agit donc d'un système complexe pour lequel il n'y a, par exemple, pas unicité de la solution stationnaire. En effet on peut facilement vérifier que le système admet une solution pour laquelle $n_\omega = 0$. Cette solution (appelée solution triviale) correspond à l'absence d'émission de photons. Dans ce cas la densité électronique n_e est donnée par sa valeur à l'équilibre (déjà calculée plus haut pour établir le lien entre celle-ci et le courant I pour trouver la formule approchée du gain, $g = \sigma v (n_e - n_o)$). On trouve bien entendu également une solution pour laquelle n_ω n'est pas nul, c'est à dire une solution décrivant l'émission de photons par le laser.

$$\begin{cases} \frac{dn_\omega}{dt} = \sigma v (n_e - n_o) n_\omega - \frac{n_\omega}{\tau_p} \\ \frac{dn_e}{dt} = -\frac{n_e}{\tau_c} + \frac{I}{eVol} - \sigma v (n_e - n_o) n_\omega \end{cases}$$

Solutions stationnaires

$$\frac{dn_\omega}{dt} = 0, \quad \frac{dn_e}{dt} = 0$$

Solution triviale

$$\begin{cases} n_\omega = 0 \\ n_e = \frac{I\tau_c}{eVol} \end{cases}$$

Cette deuxième solution stationnaire appelée non-triviale peut être facilement calculée en annulant tout d'abord le terme de droite de la première équation de bilan (en supposant que n_ω n'est pas nul). Comme indiqué ci-dessous on en retire immédiatement la condition d'oscillation $g = \alpha_{cav}$ ce qui était bien sûr attendu puisque, à l'équilibre stationnaire le gain ne peut faire que compenser exactement les pertes.

Solution non-triviale

$$\textcircled{1} \quad \sigma v (n_e - n_o) = gv = \frac{1}{\tau_p}$$

$$\text{où } \frac{1}{\tau_p} = v\alpha_{cav} \quad \longrightarrow \quad g = \alpha_{cav}$$

La condition d'oscillation impose donc une valeur précise du gain c'est à dire une valeur précise de la densité de porteurs soit n_e^* cette densité à l'équilibre.



on retrouve la cond. d'oscillation

$$\sigma v (n_e - n_o) = \frac{1}{\tau_p} \quad \longrightarrow \quad n_e = n_o + \frac{1}{\sigma v \tau_p} \equiv n_e^*$$

Cette solution non-triviale correspond à l'état oscillant (ou « lasant ») du système. Il y correspond donc un flux de photons non-nul dont la densité n_ω peut être facilement calculée à partir de la deuxième équation de bilan. Ce calcul est immédiat, il suffit de remplacer dans l'équation n_e par sa valeur à l'équilibre n_e^* .

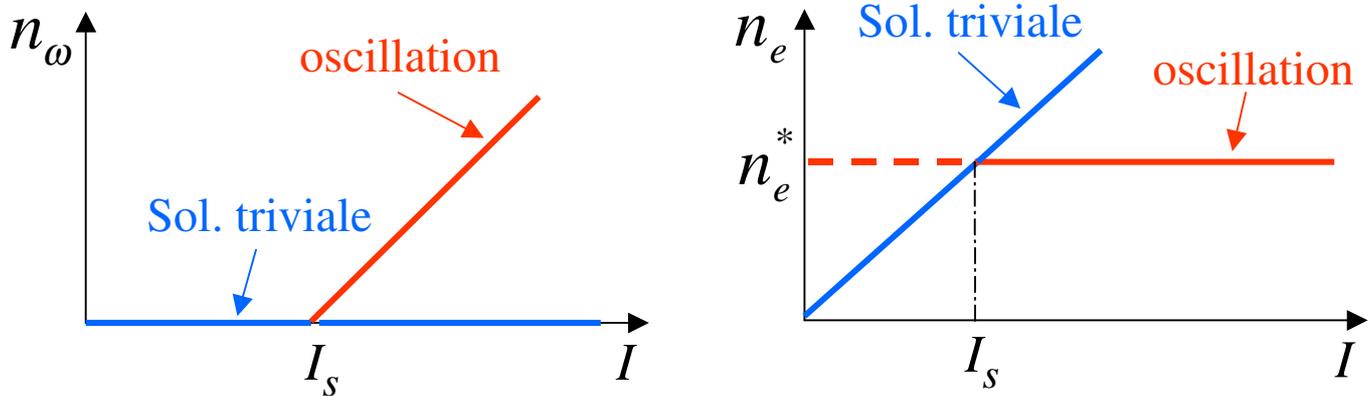
②

$$n_{\omega} = \tau_p \left(\frac{I}{eVol} - \frac{n_e^*}{\tau_c} \right) \rightarrow n_{\omega} = \frac{\tau_p}{eVol} (I - I_s) \quad \text{où} \quad I_s = \frac{eVol n_e^*}{\tau_c}$$

On constate que la densité de photons émis n_{ω} est une simple fonction linéaire du courant d'excitation I . On voit également que n_{ω} n'est non nul qu'au delà du courant de seuil I_s dont la valeur avait déjà été trouvée plus haut lors de la formulation de la condition d'oscillation (condition d'amplitude $g = \alpha_{cav}$).

Pour résumer la situation en ce qui concerne les deux états stationnaires possibles du laser nous illustrons graphiquement les solutions trouvées ci-dessus pour n_{ω} et n_e . Les solutions triviale et « oscillante » se croisent au point $I = I_s$. Comme n_{ω} ne peut être négatif, la solution oscillante commence en cette valeur du courant d'excitation. Un tel diagramme de représentation des solutions porte le nom de diagramme de bifurcation ; le terme bifurcation correspondant au fait que la solution triviale bifurque en $I = I_s$, comme si au delà de ce point le système avait le choix entre les deux solutions.

Représentation graphique (diagramme de bifurcation)



Calcul de stabilité de la solution triviale

Pour montrer quelle solution le système choisit en $I = I_s$, nous faisons ci-dessous le calcul de la stabilité de la solution triviale. A cette fin nous considérons les équations de bilan dynamiques (c'est à dire avec les dérivées temporelles) et nous supposons que n_{ω} et n_e évoluent mais restent très proches de leur valeur stationnaire de manière à pouvoir considérer leurs variations comme des quantités du premier ordre et négliger toutes les quantités d'ordre supérieur. Cette procédure est résumée dans les lignes qui suivent : $\varepsilon(t)$ représente physiquement l'émission spontanée qui fait qu'en réalité n_{ω} n'est jamais tout à fait nul. L'émission spontanée est une perturbation de la solution triviale. $\delta(t)$ représente l'écart de n_e par rapport à sa valeur stationnaire en raison de cette perturbation.

$$\begin{cases} n_{\omega}^{st} = 0 \\ n_e^{st} = \frac{I\tau_c}{eVol} \end{cases}$$

Émission spontanée = perturbation, soit :

$$\begin{cases} n_{\omega} = \varepsilon(t) \\ n_e = n_e^{st} + \delta(t) \end{cases}$$

$$\left[\varepsilon \text{ et } \delta \text{ sont d'ordre 1 : } \varepsilon \ll 1, \delta \ll n_e^{st} \rightarrow \varepsilon^2, \delta^2, \varepsilon\delta \approx 0 \right]$$

En substituant ces expressions de n_ω et n_e en fonction de $\varepsilon(t)$ et $\delta(t)$ et en négligeant tous les termes de puissance supérieure à 1 en ε et δ , nous trouvons un système d'équation différentielle ordinaire linéaire. Cette procédure de linéarisation est standard pour les calculs de stabilité des systèmes non-linéaires.

Linéarisation des équations

La solution générale de l'équation en ε est une simple exponentielle dont le coefficient λ est proportionnel à $(I - I_s)$. Connaissant la solution $\varepsilon(t)$ on peut facilement calculer la solution générale de l'équation en $\delta(t)$. On constate que $\delta(t)$ évolue selon deux exponentielles, l'une étant toujours décroissante et l'autre étant croissante pour $I > I_s$ comme pour l'évolution de $\varepsilon(t)$.

$$\frac{d\varepsilon}{dt} = \sigma v (n_e^{st} + \delta - n_o) \varepsilon - \frac{\varepsilon}{\tau_p}, \quad \frac{d\delta}{dt} = -\frac{n_e^{st}}{\tau_c} \delta + \frac{I}{eVol} - gv\varepsilon$$

$\varepsilon(t) = \varepsilon(0) e^{\lambda t}$
 $\delta(t) = \delta(0) e^{-t/\tau_c} - \frac{\varepsilon(0) gv}{\lambda + 1/\tau_c} (e^{\lambda t} - e^{-t/\tau_c})$

où $\lambda = \sigma v (n_e^{st} - n_o) - \frac{1}{\tau_p}$

$$\lambda = \sigma v \left(\frac{I\tau_c}{eVol} - n_o - \frac{1}{\sigma v \tau_p} \right)$$

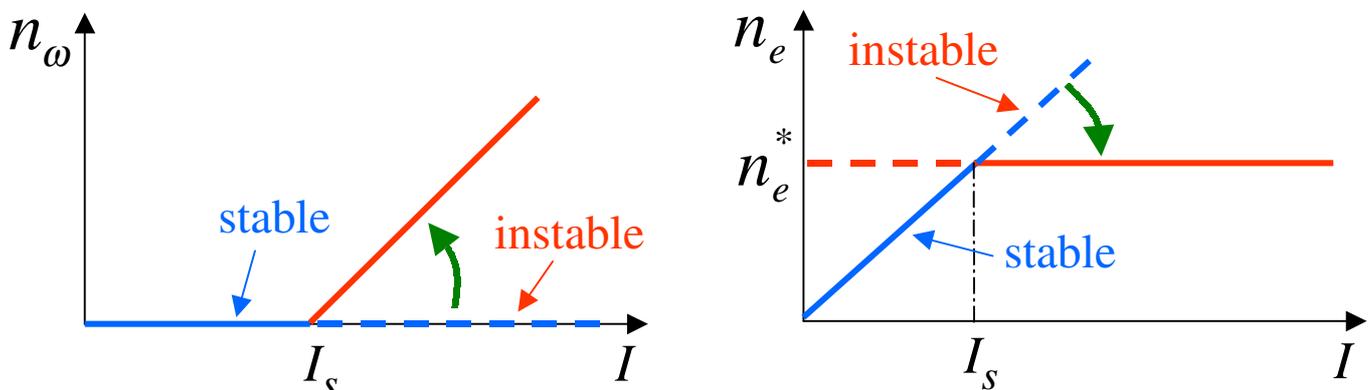
$\left[I_s = \frac{eVol n_e^*}{\tau_c} \right]$
 $\lambda = \sigma v \frac{\tau_c}{eVol} (I - I_s)$

On déduit de ce petit calcul que la perturbation $\varepsilon(t)$ décroît de manière exponentielle en dessous du courant de seuil du laser et qu'elle croît exponentiellement au delà de ce seuil. Physiquement ceci signifie simplement que la solution triviale est stable en dessous du seuil (l'émission spontanée est présente mais le nombre de photons reste limité) et est instable au dessus du seuil : l'émission spontanée est responsable d'une croissance exponentielle du nombre de photons et le système choisi spontanément (c'est le cas de le dire) d'osciller (ou « laser ») c'est à dire de se brancher sur la solution oscillante correspondant à l'émission de photons cohérents.

Conclusion :

{	si $I < I_s$	$\lambda < 0$	système stable
	si $I > I_s$	$\lambda > 0$	système instable

Le fonctionnement du laser tel que décrit à partir de ce raisonnement est représenté schématiquement ci-dessous.



Puissance émise et rendement

La puissance émise par le laser est donnée par l'énergie de l'ensemble des photons de la cavité (c'est à dire $\hbar\omega n_\omega Vol$) divisé par le temps de vie des photons dans la cavité en l'absence de milieu à gain, soit τ_{mir} . En effet, l'inverse de ce temps de vie représente bien le taux auquel les photons quittent la cavité par transmission au travers des miroirs. Ce taux multiplié par l'énergie des photons représente donc bien la puissance émise par le laser. Tout comme nous l'avons établi pour τ_p et α_{cav} , τ_{mir} est lié à α_{mir} par la relation $\tau_{mir} = 1/(\nu \alpha_{mir})$.

Puissance lumineuse émise

$$P_{lum} = \hbar\omega n_\omega Vol \alpha_{mir} \nu$$

$$\left(\begin{array}{l} R = \exp(-\alpha_{mir} L) \\ \alpha_{mir} = \ln(1/R) / L \end{array} \right)$$

$$\left(P_{lum} = \frac{\hbar\omega n_\omega Vol}{\tau_{mir}} \quad \text{où} \quad \tau_{mir} = \frac{1}{\alpha_{mir} \nu} \right)$$

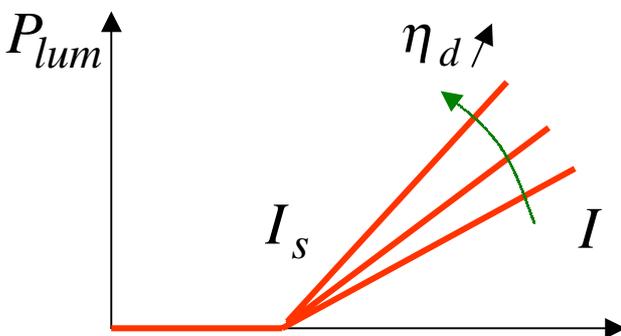
En adoptant pour n_ω sa valeur correspondant à la solution oscillante et en exprimant τ_p en fonction de $\alpha_{cav} = \alpha_{mir} + \alpha_{int}$, nous pouvons écrire P_{lum} comme indiqué ci-dessous :

$$n_\omega = \frac{\tau_p}{e Vol} (I - I_s) \quad \longrightarrow \quad P_{lum} = \frac{\hbar\omega}{e} \frac{\alpha_{mir}}{\alpha_{mir} + \alpha_{int}} (I - I_s)$$

$$\left(\tau_p = \frac{1}{\alpha_{cav} \nu} = \frac{1}{(\alpha_{mir} + \alpha_{int}) \nu} \right)$$



Le graphe ci-dessous, illustre la variation de la puissance émise en fonction du courant d'excitation de la jonction. Bien entendu, la puissance lumineuse ne devient non nulle qu'au delà du courant de seuil. La pente de la droite donnant la croissance linéaire de la puissance en fonction du courant définit le rendement différentiel du laser. C'est une quantité essentielle dans la mesure où l'on a toujours intérêt à avoir un rendement différentiel le plus élevé possible pour atteindre une puissance lumineuse donnée avec un minimum de courant.



Rendement différentiel

$$\eta_d \equiv \frac{\alpha_{mir}}{\alpha_{mir} + \alpha_{int}}$$

On constate, d'après la formule ci dessus, que le rendement différentiel est conditionné par le rapport entre les pertes internes du milieu à gain et les pertes par les miroirs. Pour avoir un bon rendement il faut que les pertes internes soient toujours bien inférieures aux pertes par les miroirs. En pratique cela signifie que les cristaux utilisés pour la fabrication des lasers à S-C doivent être de bonne qualité optique (les pertes internes principales se faisant par diffusion sur les défauts du cristal S-C). De même on aura intérêt à avoir une structure guidante telle que la double hétérojonction pour éviter au maximum la perte de photons par réabsorption et réémission hors axe.

$$\eta_d \rightarrow 1 \quad \text{si} \quad \alpha_{int} \ll \alpha_{mir} :$$



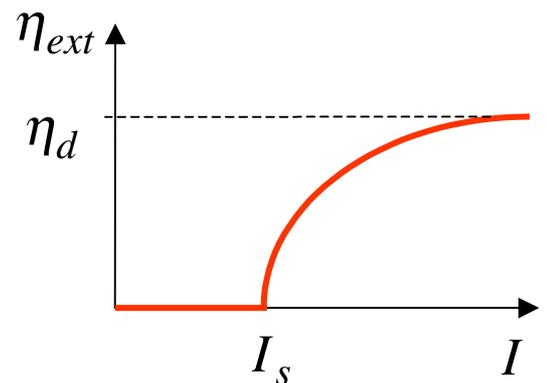
matériau S-C de qualité (absorption, diffusion)
guidage optique

Rendement externe

On définit le rendement externe comme étant le rapport du nombre de photons émis sur le nombre d'électrons utilisés. Au delà du seuil, le rendement externe tend de manière asymptotique vers le rendement différentiel.

$$\eta_{ext} = \frac{P_{lum}}{\hbar \omega} \frac{e}{I} = \eta_d \left(1 - \frac{I_s}{I}\right)$$

nbre de photons émis (pointing to $\frac{P_{lum}}{\hbar \omega}$)
1/nbre d'e⁻ utilisés (pointing to $\frac{e}{I}$)



Rendement total

On introduit finalement le rendement total comme étant le rapport entre la puissance lumineuse émise et la puissance électrique consommée. On voit qu'au delà du seuil le rendement total est essentiellement conditionné par le rendement différentiel.

$$\eta_{tot} = \frac{P_{lum}}{P_{elec}} = \frac{P_{lum}}{I V_a} \rightarrow \eta_{tot} = \frac{\hbar \omega}{e V_a} \eta_d \left(1 - \frac{I_s}{I}\right)$$

On retrouve dans cette formule l'expression du rendement externe.

$$\eta_{\text{tot}} = \frac{E_g}{eV_a} \eta_{\text{ext}}$$

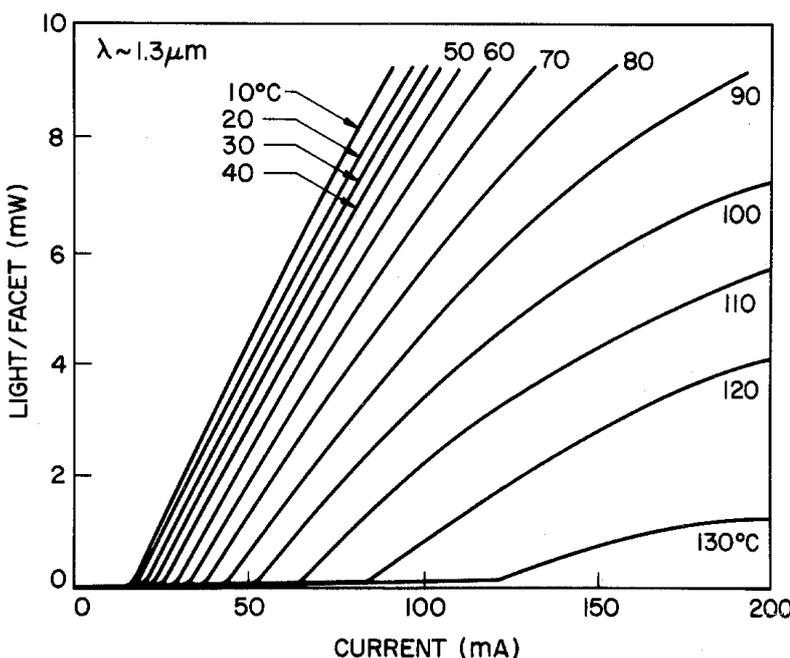
$$\left(eV_a \approx \Delta E_F > E_g \right)$$

Une chose importante à mentionner pour clore cette discussion sur le rendement des lasers à S-C est que ces lasers ont des rendements très élevés par rapport aux lasers classiques tels que les lasers à gaz (argon, hélium-néon, etc...) ou les lasers à cristaux ou verres dopés (cristaux YAG dopés au Néodyme, laser à fibre dopée à l'erbium, etc...). Les rendements des lasers à S-C sont de l'ordre de plusieurs dizaines de pour-cents alors que les lasers classiques ont des rendement typiquement de l'ordre du pour-cent. La raison de cette différence est, d'une part, la grande densité de dipôles électriques et d'autre part, la méthode d'inversion de la population par simple polarisation de la jonction. Dans les lasers classiques l'inversion de population se fait de manière optique (le pompage optique) par des lampes dont le spectre est très large et pour lequel beaucoup de photons restent dès lors inutilisés car ils sont en dehors de la condition d'excitation des transitions atomiques utiles au pompage optique. Il y a de plus une grande perte de photons due au fait que les lampes émettent un rayonnement incohérent isotrope qui est difficile à maîtriser du point de vue de la géométrie de l'éclairage du matériau à gain (par exemple, il faut utiliser des réflecteurs pour récupérer les photons émis dans la direction opposée au cristal amplificateur etc...).

Exemple : diode laser au GaAs $\eta_d = 80\%$, $\eta_{\text{tot}} = 50\%$

(Laser à gaz (Ar, He-Ne) : 1%)

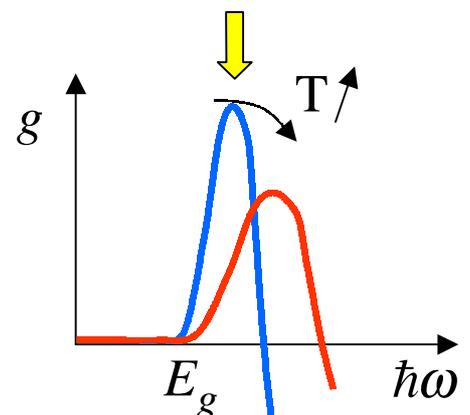
La figure suivante montre des courbes expérimentales de puissance émise en fonction du courant d'excitation. On constate que les prédictions qualitatives du modèle simplifié que nous avons discuté se vérifient bien. On voit en particulier l'existence d'un courant de seuil. On constate aussi que le rendement différentiel diminue avec une augmentation de la température. Ceci est dû à une augmentation des pertes internes liée à la température (diffusion des photons par les phonons du cristal). De même le seuil augmente avec la température. Ceci est dû au fait que l'augmentation de la température diminue le temps de vie non-radiative des porteurs minoritaires (interaction des électrons avec les phonons du cristal) de laquelle découle une augmentation de la densité de seuil n_o et de la densité de porteurs à l'équilibre n_o^* (comme suggéré par le schéma en bas à droite).



$\eta_d \downarrow$ si $T \uparrow$ car $\alpha_{\text{int}} \uparrow$

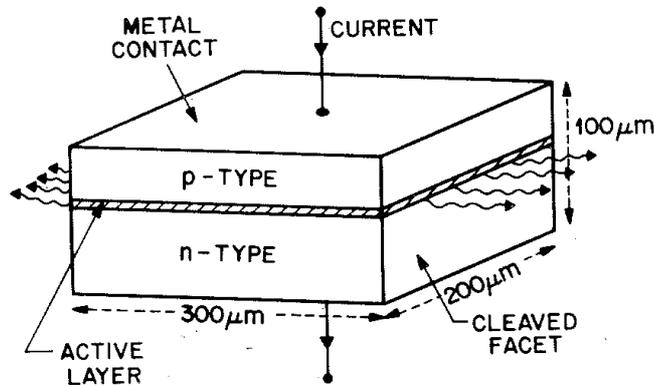
$$I_s = \frac{eV_o n_e^*}{\tau_c} \uparrow \quad \text{si } T \uparrow$$

car $\tau_c \downarrow$ $n_o \uparrow$ et $n_e^* \uparrow$



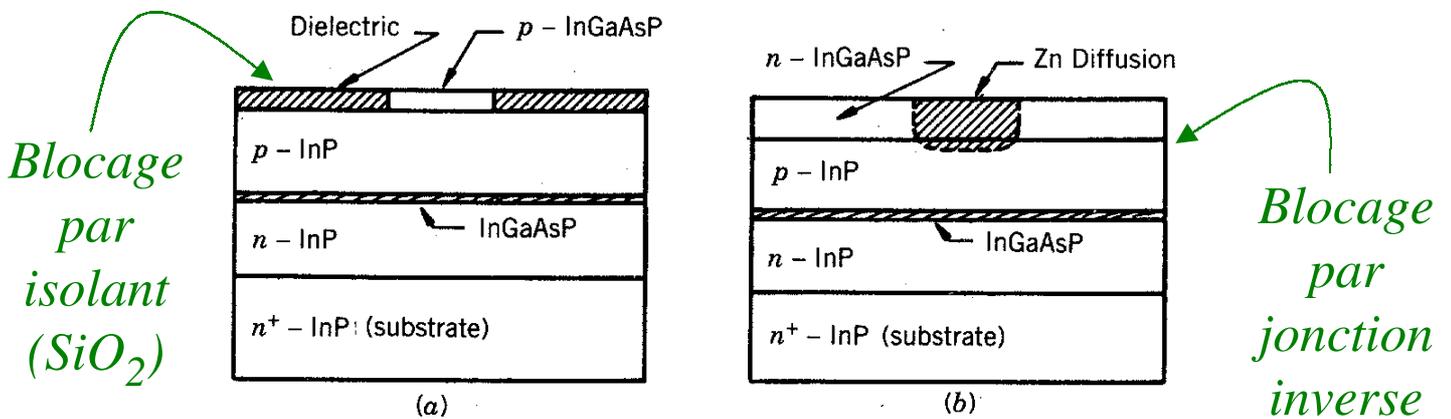
Structures des Lasers à semi-conducteurs

Rayonnement non confiné (broad area lasers)



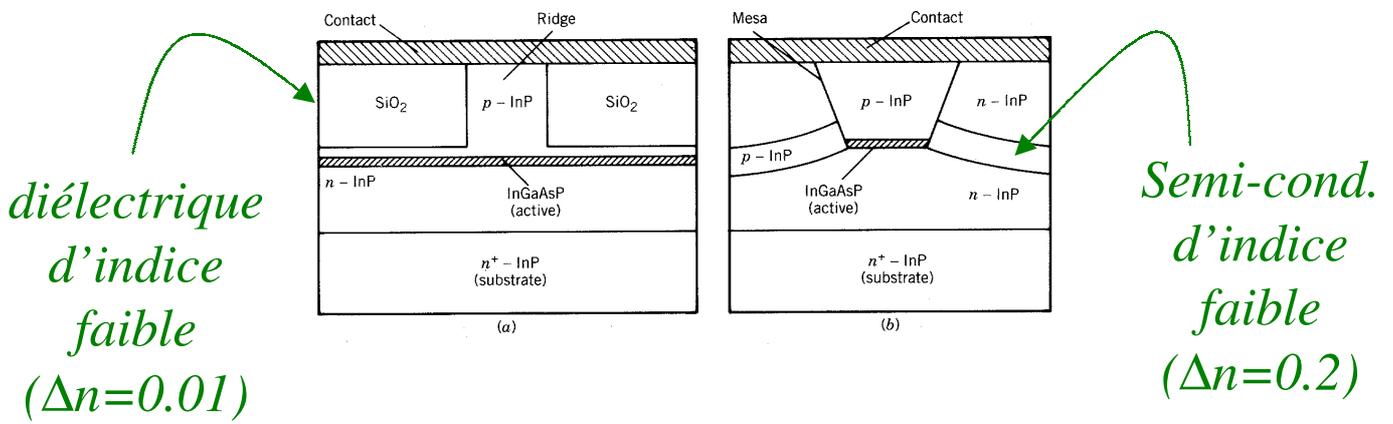
Le schéma ci-dessus illustre la structure la plus simple que l'on puisse imaginer pour le laser à S-C. En pratique, pour éviter que le rayonnement ne s'étale trop dans la dimension transverse (ce qui est recherché, en particulier pour l'injection dans les fibres), on limite la zone émettrice transversalement. La manière la plus simple de limiter la zone émettrice est de limiter spatialement le gain en fabricant des électrodes en forme de bande. De cette manière l'amplification optique a lieu seulement en dessous de la bande et les photons sont générés seulement à cet endroit comme s'ils étaient guidés ; c'est pourquoi les lasers basés sur ce principe s'appellent des lasers à « guidage par gain ». Cette solution est illustrée ci-dessous à l'aide de deux exemples basés sur des techniques différentes de guidage par gain. L'une consiste simplement à bloquer le courant par une couche d'isolant disposée de part et d'autre d'une bande conductrice (dans l'exemple de gauche ci-dessous, il s'agit d'une couche de S-C de type p sur laquelle on vient déposer la cathode). Dans l'autre exemple on bloque le passage du courant par déposition d'une couche de S-C de type n et on ne laisse passer le courant que sur une bande étroite en dopant cette couche localement avec des atomes de Zinc qui sont des accepteurs (valence 2, dopage de type p) et rendent la couche conductrice.

Guidage par gain (gain-guided lasers)



Pour améliorer l'efficacité du guidage par gain on peut avoir recours au guidage optique. Dans les exemples ci-dessous un guidage optique est obtenu en insérant des couches de silice d'indice de réfraction faible par rapport à celui des S-C. Ces zones d'indices faibles étant proches de la couche émettrice, les photons générés ont tendance à fuir ces zones et à se concentrer au centre de la zone de guidage par gain. Le schéma de droite illustre la solution la plus efficace pour le confinement transverse du rayonnement laser: la zone émettrice est complètement entourée par des couches de matériaux S-C d'indice de réfraction plus faible (cf. ce que l'on a vu pour le principe de guidage dans la double hétérojonction). Par cette technique le saut d'indice réellement vu par les photons générés peut être très élevé et le guidage est extrêmement efficace (dans la configuration de gauche, les photons ne voient pas directement l'indice plus faible, il ne font qu'en subir l'influence, d'où le guidage moins efficace: le saut d'indice effectif vu par les photons est d'un ordre de grandeur inférieur).

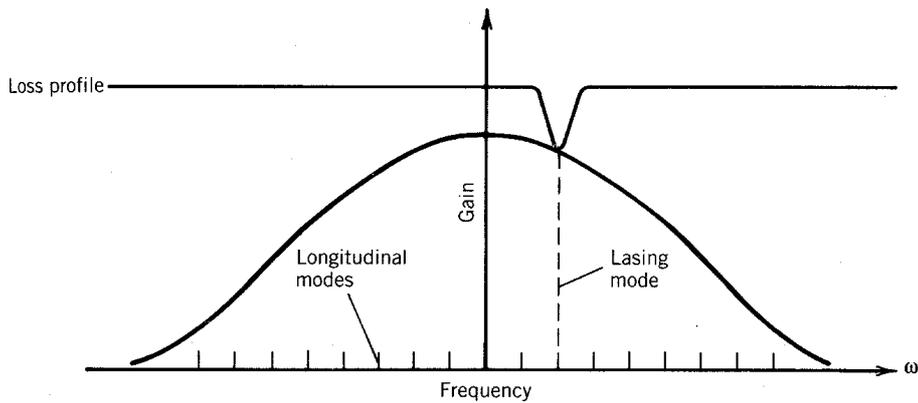
Guidage par saut d'indice (index-guided lasers)



Lasers monomodes

Comme on l'a déjà discuté à plusieurs reprises, pour les applications aux télécommunications optiques on a intérêt à avoir une source laser monochromatique. En pratique cela signifie un laser qui n'oscille que sur un mode. On a vu plus haut que, même si il y a un mode dominant correspondant au maximum de la courbe de gain, le laser émet toujours sur plusieurs modes. Pour éliminer complètement les modes non-dominants on crée artificiellement des pertes additionnelles pour ces modes. C'est ce qui est représenté à la figure suivante où l'on voit la courbe de gain, les modes du laser et une courbe de perte sélective en fréquence. C'est le mode correspondant au minimum de cette courbe de perte qui « lasera », les autres seront négligeables (pour autant que le contraste de perte soit suffisant bien entendu).

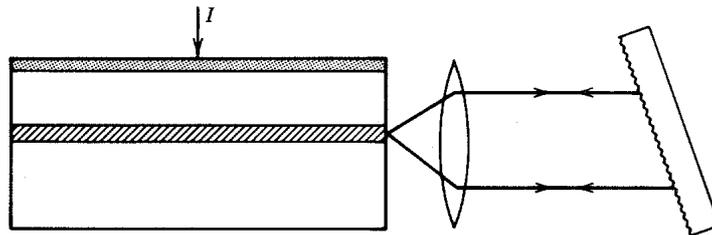
Principe : introduction de pertes localisées spectralement



Exemples

Le schéma suivant est celui du « laser à cavité externe ». Dans cette configuration un des miroirs du cristal S-C a été supprimé en appliquant sur la surface clivée une couche anti-reflet (multi-couche diélectrique). La cavité est alors formée à l'extérieur à l'aide d'un réseau de diffraction utilisé en monture de Littrow. Dans cette configuration, les différentes fréquences (modes du laser) sont diffractées par le réseau (c'est à dire que chaque fréquence est réfléchiée avec un angle différent) et une seule fréquence est réfléchiée dans l'axe de la cavité. En d'autre terme, un seul mode est sélectionné et le laser sera idéalement monochromatique (les autres modes étant supprimés complètement par diffraction en dehors de l'axe de la cavité). Cette configuration a le désavantage d'être encombrante et instable mécaniquement (elle requiert un système d'asservissement mécanique complexe) et ne peut être exploitée dans les applications industrielles (c'est un laser de laboratoire).

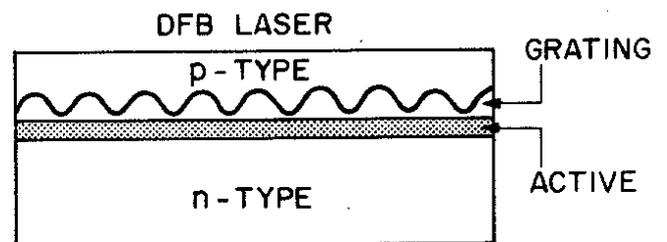
Laser à cavité externe (et réseau)



Une solution beaucoup plus pratique pour les applications est celle qui consiste à supprimer les deux miroirs de la cavité (par traitement anti-reflet des faces clivées) et les remplacer par un réflecteur de Bragg distribué. Le réflecteur de Bragg est créé en gravant un réseau à la surface de la dernière couche de la double hétérojonction. Ce réseau étant proche des photons générés, ces derniers subissent la réflexion de Bragg pour la longueur d'onde de Bragg (soit la longueur d'onde optique valant le double de la période du réseau). De la sorte, les photons générés à la longueur d'onde de Bragg sont réfléchis en tout point du système et effectue des aller-retours sous le réseau. Ces aller-retour sont analogues à ceux que l'on avait dans la cavité, à part que maintenant ils ne concernent qu'une et une seule fréquence. On a donc bien un laser monochromatique qui a en plus l'avantage d'être compact.

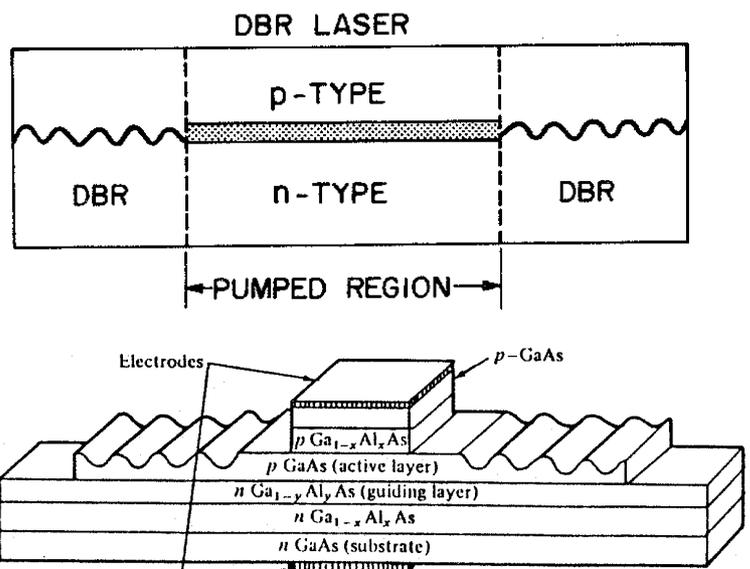
Laser à réflecteurs de Bragg

Réflecteur distribué
(distributed feedback)



Miroirs à
réflecteurs de Bragg

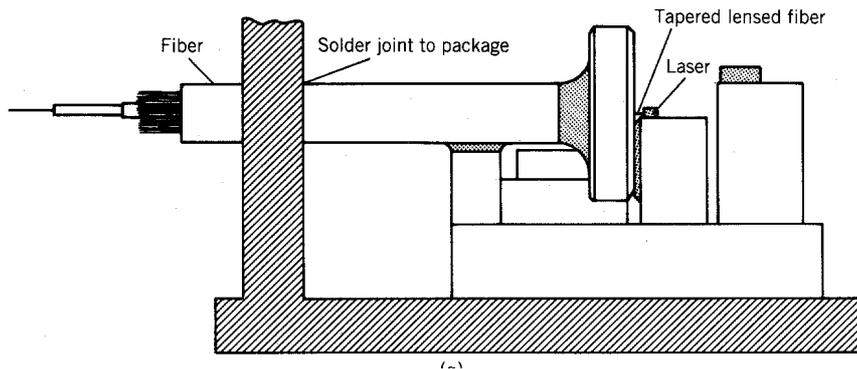
Dans cette dernière configuration, les miroirs de la cavité sont simplement remplacés par des réflecteurs de Bragg (réseaux optiques) gravés près de la zone émettrice. Les réflexions se font donc de manière sélective en longueur d'onde puisque seule la longueur d'onde de Bragg de réflecteurs fait des aller-retours dans la cavité. On obtient donc également ainsi un laser monomode.



Applications aux télécommunications

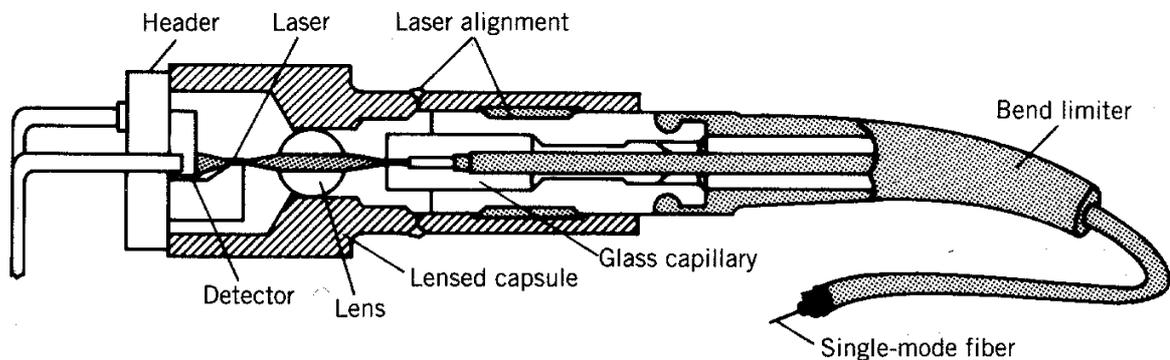
Les deux figures suivantes illustrent la manière dont les lasers sont montés pour injecter la lumière dans une fibre optique pour les applications à la transmission d'information. Dans le premier cas la fibre est approchée de la face de sortie du laser de manière à y injecter directement le faisceau laser. L'ensemble fibre-laser est stabilisé mécaniquement à l'aide d'un montage de pièces métalliques de grande précision.

Couplage laser-fibre direct



Couplage par micro-lentille

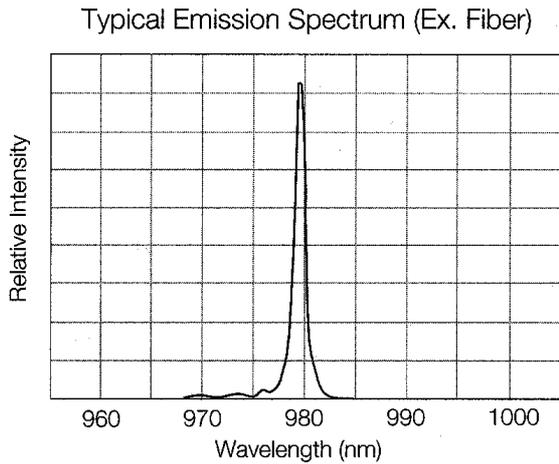
Dans cette deuxième configuration le faisceau laser est injecté dans la fibre par l'intermédiaire d'une lentille sphérique.



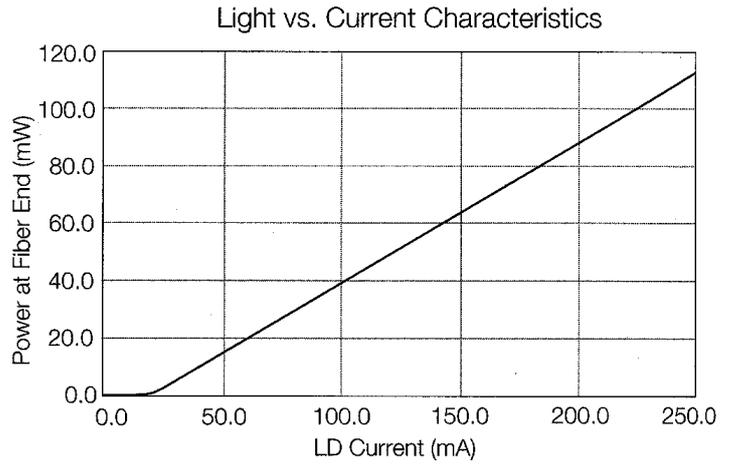
Exemple

Les figures ci-dessous montrent un exemple de spectre d'émission de laser à réseau de Bragg distribué (DFB laser) ainsi que la caractéristique puissance lumineuse-courant de ce laser. On voit que la largeur spectrale est de l'ordre du nanomètre. Ceci ne fait qu'illustrer les potentialités des lasers par rapport aux DEL, en pratique des largeurs spectrales aussi basses qu'une fraction d'Angström peuvent être obtenues avec des lasers du même type.

Spectre d'émission



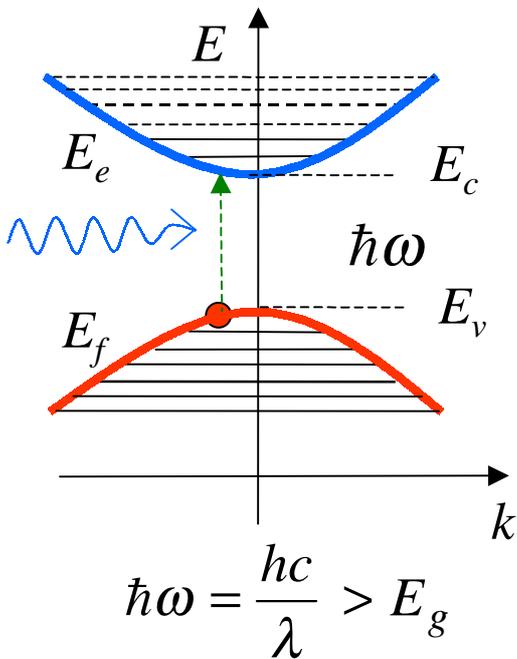
Caractéristique Puissance-courant



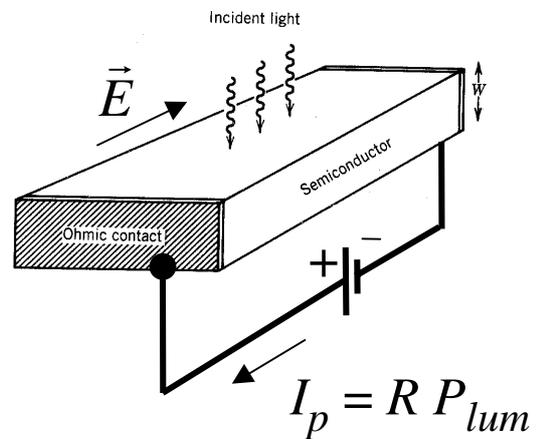
Les photodiodes

Les semi-conducteurs peuvent également être utilisés pour détecter la lumière. Le principe en est très simple. Imaginons que l'on place un S-C entre deux plaques d'un condensateur aux bornes duquel est appliqué une ddp. Si le S-C est éclairé avec une lumière dont la fréquence est supérieure à celle du gap, il y aura création de paires électron-trou. Les électrons et les trous vont dériver en sens inverse dans le champ électrostatique et vont donc créer un courant. La mesure de ce courant sera une mesure de la puissance lumineuse incidente sur le S-C.

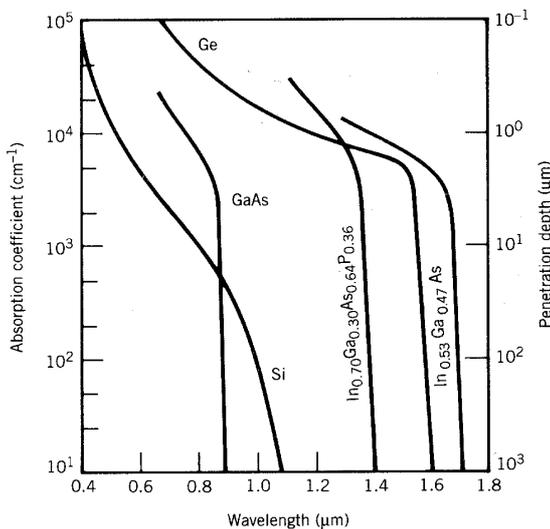
Principe de la photodétection dans les semi-conducteurs :



Création d'une paire e⁻-trou par absorption d'un photon et dérive des porteurs libres dans un champ électrique statique



La constante de proportionnalité entre le courant mesuré et la puissance lumineuse incidente s'appelle la sensibilité photoélectrique du S-C, elle est notée R (pour responsivity en anglais). R est donc donnée par le facteur $e/h\omega$ multiplié par le rapport du nombre de paire électron-trou générées sur le nombre de photons incidents. Ce rapport représente le rendement photoélectrique η du S-C. Si chaque photon incident conduit à la génération d'un électron contribuant au courant alors le rendement η vaut l'unité. Ceci n'est toute fois vrai que si l'épaisseur du S-C est suffisante que pour absorber tous les photons incidents, ce qui n'est pas nécessairement le cas puisque la profondeur de pénétration des photons dépend du coefficient d'absorption qui lui dépend de la longueur d'onde et du gap du S-C (voir schéma ci-dessous à gauche). A priori donc le rendement maximum est donné par le rapport de la puissance absorbée sur la puissance incidente et vaut $1 - \exp(-\alpha W)$ où α est le coefficient d'absorption et W l'épaisseur du S-C.



Sensibilité : $R = \frac{I_p}{P_{lum}}$

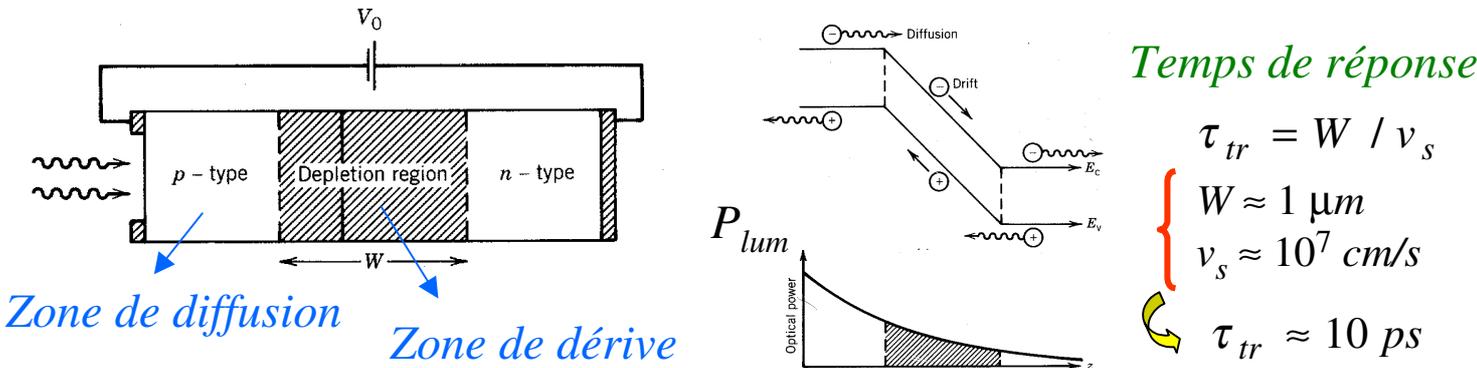
$$R = \frac{e \times \text{nbre d'e}^- \text{ générés / sec}}{\hbar\omega \times \text{nbre de photons / sec}}$$

$$R = \eta \frac{e}{\hbar\omega} = \eta \frac{\lambda [\mu m]}{1.24}$$

$$\eta_{max} = P_{abs} / P_{lum} = (P_{lum} - P_{trans}) / P_{lum} = 1 - \exp(-\alpha W)$$

Cette relation indique l'importance de la valeur du coefficient d'absorption (et donc de la relation entre gap et longueur d'onde) pour la photodétection, en particulier, il faut adapter le choix du S-C à la longueur d'onde de la lumière à détecter. Mais bien entendu, même si α est très grand et que $P_{abs} = P_{lum}$, le rendement n'atteint jamais l'unité car des paires électron-trou peuvent se recombiner sur le temps de la dérive au sein du S-C. Toute paire recombinée, ne contribuant plus au courant, diminue le rendement. Il y a donc intérêt à appliquer un champ électrostatique aussi élevé que possible de sorte que le temps de dérive soit court par rapport au temps de vie des porteurs libres. A cette fin on utilise la jonction $p-n$ (elle, une fois de plus) polarisée en inverse de manière à ce que, dans la région de charge d'espace, un champ électrostatique important soit créé. Il est en effet plus facile de créer à partir d'une ddp donnée, un champ élevé dans un jonction $p-n$ car la région de charge d'espace est d'extension limitée (typiquement de l'ordre du micron). La vitesse de dérive y est donc facilement très élevée ; elle atteint en pratique la vitesse de dérive maximale (limite due aux collisions avec les autres porteurs et le réseau cristallin) de l'ordre de 10^7 cm/s. Les temps de dérive typiques sont donc aussi faibles qu'une dizaine de picosecondes. Ce temps permet d'obtenir non seulement de bons rendements mais aussi des temps de réponse très courts (ce qui est important pour certaines applications comme les télécommunication où les impulsions codant l'information doivent être détectées à très grands débits).

Jonction p-n : champ de dérive élevé, $\eta \uparrow$ et réponse rapide



Le schéma ci-dessus à gauche montre la structure de base d'une photodiode. Les photons sont absorbés dans tout le S-C. Ceux absorbés avant la région de charge d'espace créent des paires électron-trou auxquelles n'est pas appliqué de champ. Ces paires diffusent donc simplement et seulement quelques-unes d'entre elles contribueront au courant. Les photons absorbés dans la région de charge d'espace créent des porteurs libres qui dérivent instantanément vers les régions neutres où ils diffusent et peuvent également se recombiner. Cette brève discussion fait immédiatement apparaître la nécessité de rendre la région de charge d'espace aussi grande que possible par rapport aux régions neutres. En effet, d'une part, la région neutre sur laquelle les photons sont incidents (région p sur le schéma ci-dessus) absorbe une grande partie des photons (voir schéma inférieur droit ci-dessus) avec un très mauvais rendement puisque les porteurs créés n'y dérivent pas. D'autre part, le courant provenant de la région de charge d'espace y est diminué en raison d'un ralentissement de la vitesse des porteurs (diffusion). En pratique on s'arrange donc pour que les zones neutres soient aussi peu épaisses que possible et en particulier la zone proche de la surface éclairée.

Cette discussion sur la physique de base et la géométrie de la photodiode est résumée à la figure suivante. Notez que c'est la zone p qui forme la surface éclairée car les trous ayant une mobilité moins bonne que les électrons il est important de faire cette zone aussi peu épaisse que possible.

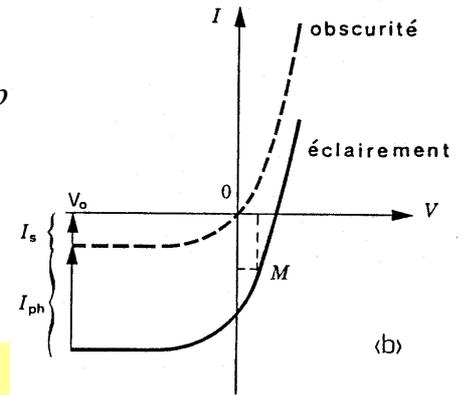
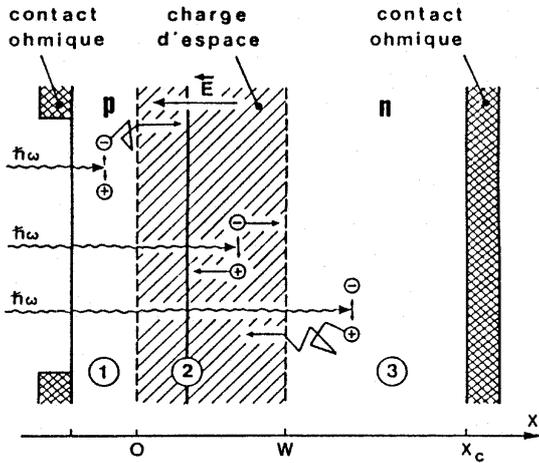
Mesure du flux lumineux : le photocourant

La puissance lumineuse incidente se mesure à l'aide d'une simple mesure du courant passant dans la photodiode. Comme celle-ci est polarisée en inverse le courant aura deux composantes : l'une est le courant de saturation naturel de la jonction I_s , l'autre est le courant généré par absorption de photons, soit le photocourant I_p .

$$I = I_s \left(e^{eV / k_B T} - 1 \right) - I_p$$

$$V = -V_o \text{ où } V_o \gg k_B T / e$$

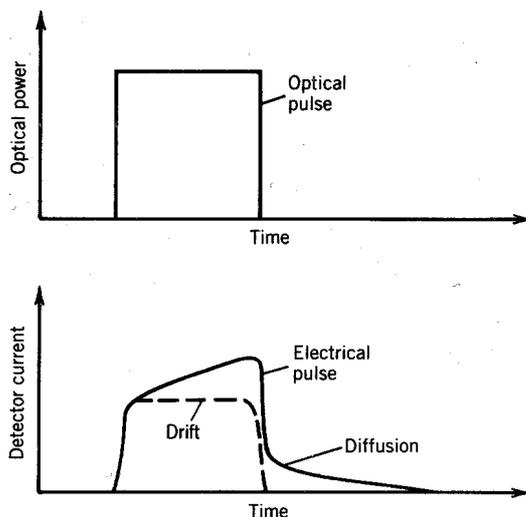
$$I \approx -I_s - I_p$$



Problèmes liés à la génération en zone neutre

Le problème de la lenteur des vitesses caractéristiques de diffusion dans les zones neutres n'est pas seulement qu'il conduit à une diminution du rendement, il est en effet responsable également d'un temps de réponse lent. En prenant 10^5 cm/s comme vitesse de diffusion (c'est une valeur typique, mais comme la diffusion est due au gradient de concentration des porteurs majoritaires, elle dépend en réalité des conditions de fonctionnement de la photodiode) et en prenant une zone neutre de 1 micron d'épaisseur nous obtenons un temps de réponse typique de l'ordre de la nanoseconde. Ce temps est beaucoup plus long que la dizaine de picosecondes correspondant au temps de réponse lié au courant de dérive (notez que le fait que le courant de dérive passe par la zone neutre ne change en rien sa dynamique car, même si la vitesse des porteurs y est réduite, le flanc montant de la concentration de porteurs correspondant à une augmentation du courant de dérive est conservé au cours de la traversée de la zone neutre). Il y aura donc deux contributions de dynamiques différentes au photocourant. Il en résultera une distorsion de mesure des signaux optiques. Ceci est illustré à la figure ci-dessous où l'on voit une impulsion lumineuse carrée et sa mesure sous forme de photocourant. Le courant de dérive conduit à une réponse fidèle alors que le courant de diffusion en raison de sa dynamique lente conduit à une forte distorsion (ceci n'est bien entendu valable que pour des impulsions subnanoseconde)

- Courant par diffusion des porteurs



1 nsec pour $1 \mu\text{m}$

$$v_d \approx 10^5 \text{ cm/s}$$

$$\tau_d \approx 1 \text{ ns}$$

Distortion des signaux

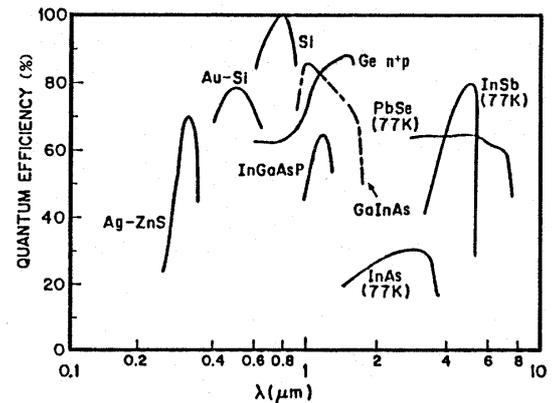
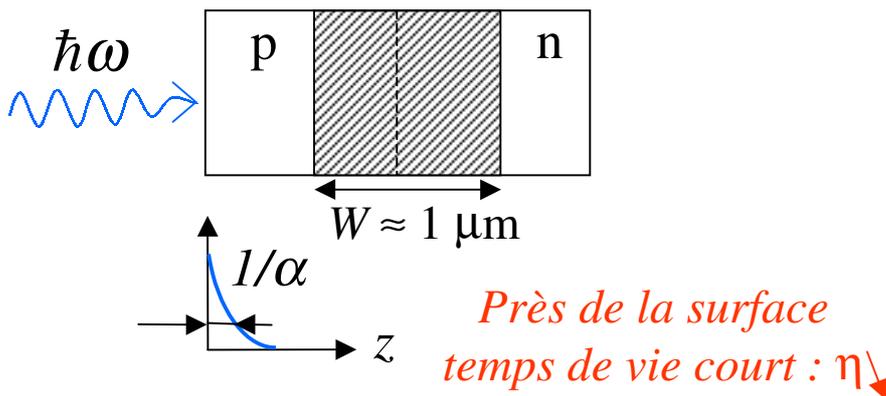
Limitation de bande passante

$$\Delta f_{\text{max}} \approx 1 \text{ GHz}$$

- Limitation du rendement à faible longueur d'onde

Le rendement des photodiodes diminue avec une diminution de la longueur d'onde de la lumière détectée contrairement à ce que pourraient laisser croire les courbes de coefficient d'absorption qui elle ne font qu'augmenter quand la longueur d'onde diminue. Le rendement diminue simplement parce que lorsque l'absorption devient trop forte, celle-ci a lieu principalement dans la zone neutre où le rendement est faible (en grande partie en raison de la proximité de la surface où les nombreux défauts du cristal forment des centres de recombinaison rapide). Ceci explique la différence entre les courbes d'absorption et les courbes de rendement des photodiodes. Ce phénomène limite donc fortement le domaine spectral utile d'une photodiode et rend le choix du matériau S-C crucial pour une application bien définie.

Si $\lambda \downarrow$ alors $\alpha \uparrow$ et absorption en dehors de la zone utile



Les photodiodes p-i-n

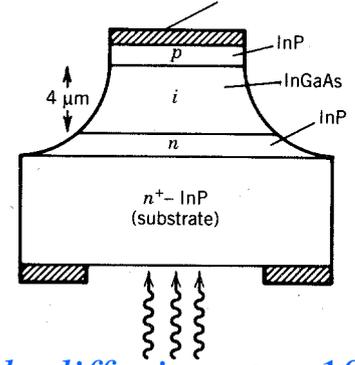
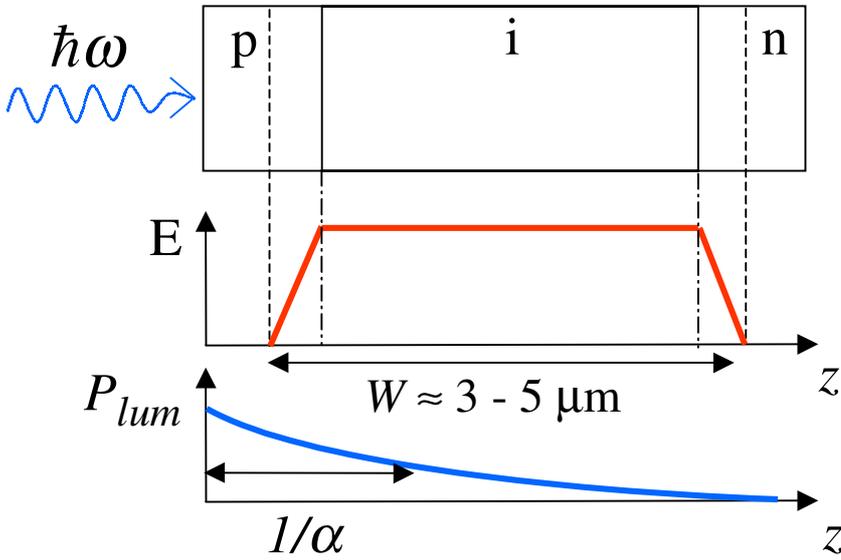
Afin d'augmenter la composante de dérive du courant par rapport à la composante de diffusion, on peut augmenter simplement la longueur de la zone de dérive W . Ceci peut être fait par l'introduction d'une couche de S-C intrinsèque entre les régions de type p et n . Avec une plus grande région de courant de dérive, la photodiode aura un meilleur rendement et la distorsion des signaux due au courant de diffusion se trouvera diminuée. De plus, si W est plus grande la capacité de la jonction est diminuée ce qui a pour avantage de donner au dispositif de mesure une dynamique plus rapide puisque celle-ci est conditionnée par la constante de temps RC du circuit équivalent. Cette configuration s'appelle la photodiode $p-i-n$ (ou PIN).

Pourquoi ? - Augmenter W (dérive \gg diffusion, $\eta \uparrow$)

- Diminuer la capacité ($\tau = RC$)

Le schéma de gauche ci-dessous montre la structure de la photodiode $p-i-n$. On voit que l'introduction de la couche de S-C intrinsèque conduit à une valeur élevée du champ électrostatique sur une grande longueur. En adaptant cette longueur au coefficient d'absorption, on peut s'arranger pour que la quasi-totalité de l'absorption se fasse dans la région de courant de dérive. La solution idéale pour la diode $p-i-n$ est de fabriquer la zone intrinsèque dans un matériau S-C de gap inférieur à celui des zones p et n . De la sorte, il n'y a plus de photons absorbés dans les zones de diffusion et le rendement de la photodiode peut approcher l'unité ; de même sa dynamique ne sera plus affectée par la lenteur du courant de diffusion. La figure de droite ci-dessous représente la structure d'une telle photodiode ; elle est appelée photodiode PIN à double hétérojonction ($P-i-N$).

Double hétérojonction P-i-N



Plus de diffusion, $\eta \approx 100\%$
(Télécoms : 1.3 et 1.55 μm)

En considérant une longueur typique de 3 à 5 microns pour la zone de dérive, on arrive à des temps de réponse de l'ordre de 30 à 50 picosecondes ce qui correspond à des bandes passantes limites de l'ordre de 20 GHz.

$$\left\{ \begin{array}{l} W \approx 3-5 \mu\text{m} \\ v_s \approx 10^7 \text{ cm/s} \end{array} \right. \Rightarrow \tau_{tr} \approx 30 - 50 \text{ ps} \Rightarrow \Delta f \approx 3-5 \text{ GHz (voire 20 GHz)}$$

Pour certaines applications à la détection dans le visible et infra-rouge très proche le silicium est choisi comme matériau. Son coefficient d'absorption étant faible (gap indirect) la longueur de la région intrinsèque doit être plus grande, la photodiode sera donc plus lente mais de capacité moindre.

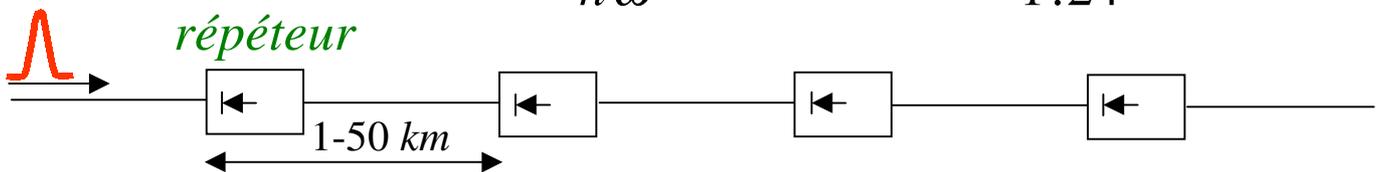
[Remarque : si α faible (ex. Si, gap indirect) $W \approx 50 \mu\text{m}$]

Pour augmenter la sensibilité R de la photodiode (ce qui en télécom optique a un intérêt immédiat puisque cela conduit à un plus grand espacement des répéteurs) on ajoute entre les zone i et n une couche p légèrement moins dopée de façon à créer une zone de champ de dérive plus élevé. Les électrons qui arrivent dans cette zone sont accélérés au point de pouvoir envoyer, par collision, un électron de valence dans la bande de conduction. C'est le processus d'ionisation d'impact. Ce processus conduit à une multiplication du nombre de porteurs et donc à une augmentation de la sensibilité. Des facteurs d'amplification de l'ordre de 10 à 100 peuvent être obtenus.

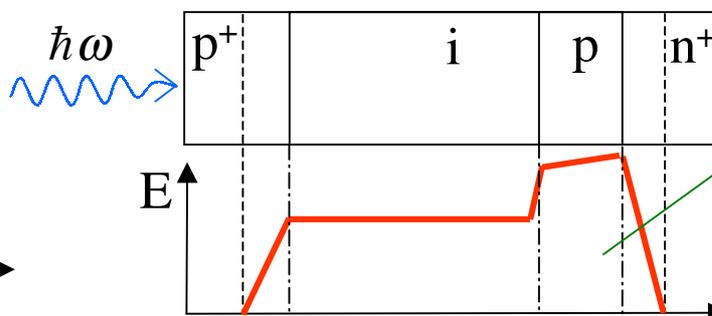
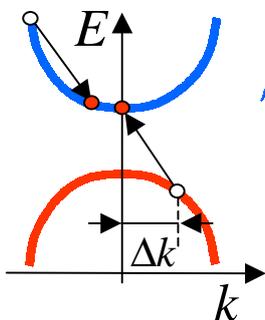
Les photodiodes à avalanche

$$I_p = R P_{lum}$$

Limite théorique de la sensibilité : $R = \eta \frac{e}{\hbar\omega} \rightarrow R_{max} = \frac{\lambda [\mu\text{m}]}{1.24} \quad (\eta = 1)$



Augmentation de R par effet d'avalanche (ionisation d'impact)



Zone de multiplication

$$I_p = M R P_{lum}$$

$$M = 10 - 100$$

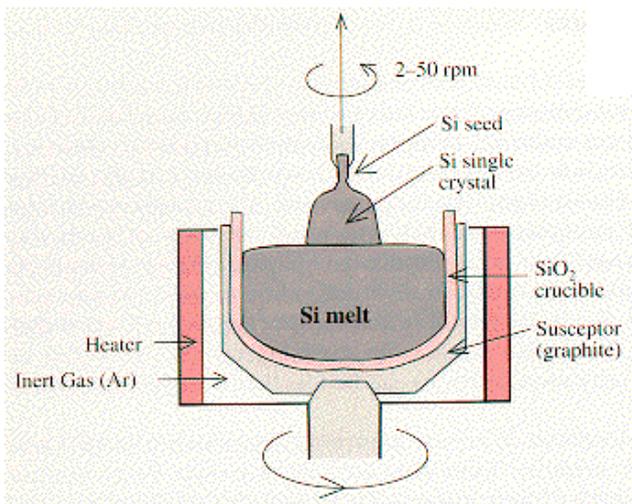
Techniques de fabrication des éléments semi-conducteurs

La technique utilisée pour fabriquer des cristaux S-C bruts sans structure particulière est la technique de Czochralski. Elle consiste simplement à tremper un germe de cristal dans un bain de S-C en fusion. Le germe est immédiatement élevé au dessus de la surface de sorte que le S-C emporté par tension de surface se cristallise. En plus du mouvement montant le germe est également mis en mouvement de rotation. Cela confère au cristal une forme de cylindre. En pratique des tailles de monocristaux de l'ordre de 20 cm de diamètre sur 1 à 2 mètres de long peuvent être obtenues. Les défauts de la méthode sont les contaminations possibles par la surface libre et le contact avec le creuset. Pour les alliages S-C à composante volatile (tels que le GaAs où l'As est volatile) on ajoute un bain de B_2O_3 liquide en surface (ce bain limite également les contaminations par la surface).

Le monocristal obtenu avec cette méthode est scié en tranche de quelques millimètre d'épaisseur. Ces tranches sont polies pour être ensuite traitées pour y inscrire les structures voulues par gravure, diffusion d'atome (dopage), lithographie, etc...

1. Croissance des cristaux S-C

Méthode de Czochralski

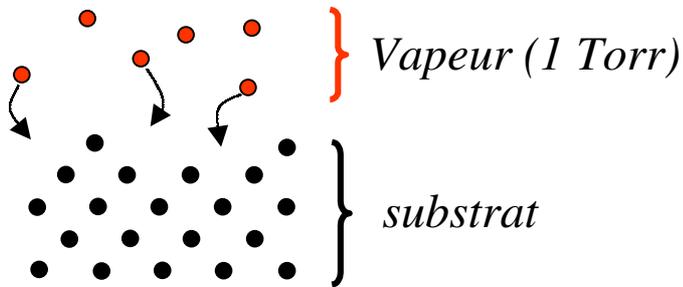


Croissance de cristaux bruts (non structurés) : Si et Ge

- *diamètre ~ 20 cm*
- *structure cristalline de bonne qualité, monocristal*
- *impuretés chimiques (creuset, air)*
- *pour S-C III-V (GaAs, InP), version améliorée avec B_2O_3 liquide en surface (As volatile)*

Épitaxie en phase vapeur (Chemical vapor deposition)

Pour fabriquer des S-C structurés de petites tailles on peut utiliser la technique de déposition appelée épitaxie en phase vapeur. Une vapeur du composé S-C est créée dans un vide poussé et les atomes viennent se déposer sur un substrat servant de germe à la croissance cristalline. Suivant que le S-C déposé est le même que celui du substrat ou non on parle d'homo-épitaxie ou hétéro-épitaxie.



Croissance de structures de petite taille

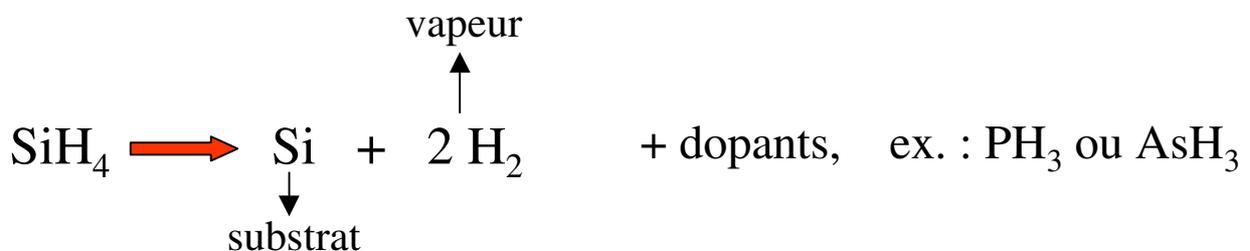
- épaisseur $\sim 1 \mu\text{m}$

- homo-épitaxie, ex. : Si sur Si

- hétéro-épitaxie, ex. : GaAs sur Si
(structure tétraédrale)

Exemples :

Un exemple typique de fabrication par épitaxie en phase vapeur est la fabrication de couches minces de silicium. Le silicium est mis en vapeur sous forme de silane. Au contact du substrat le Si se dépose et l'hydrogène reste dans la vapeur. On peut facilement faire croître des S-C dopés par cette technique en ajoutant les vapeurs adéquates.

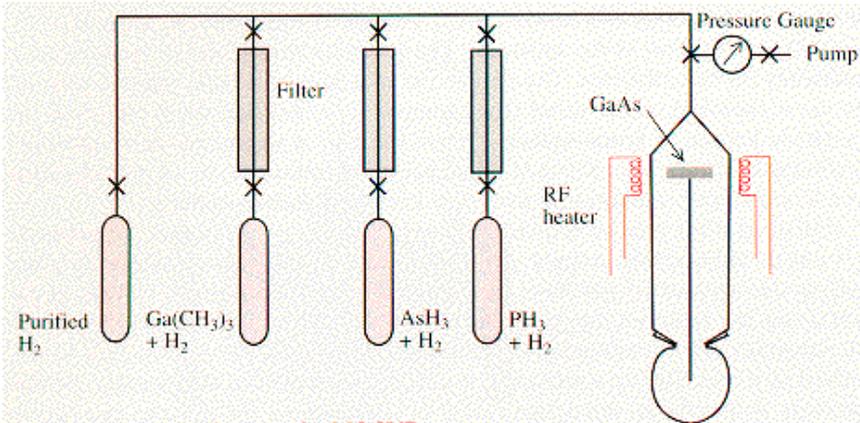


Pour les composés S-C III-V on utilise des vapeurs organiques (ou, plus précisément métallo-organiques). On parle alors de « Metal-organic chemical vapor deposition » ou MOCVD (c'est une technique très répandue).



Metal-Organic Chemical Vapor Deposition (MOCVD)

Schéma d'une installation de MOCVD



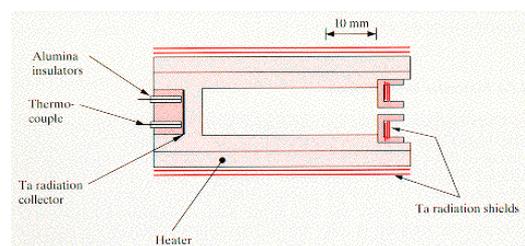
*Problème :
contamination par
résidus gazeux*

Epitaxie par faisceaux moléculaires (molecular beam epitaxy)

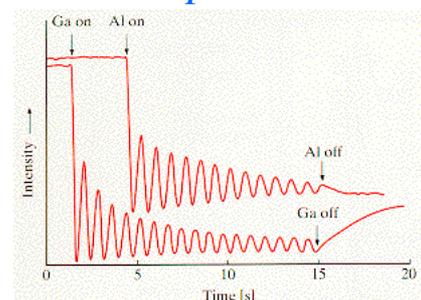
Pour les structures complexes de petites tailles la technique d'épitaxie par faisceaux moléculaires (molecular beam epitaxy) est utilisée. Elle consiste simplement à bombarder un substrat avec les atomes du composé S-C voulu. A l'aide d'un contrôle de la croissance en temps réel on peut faire croître des cristaux S-C avec une précision d'une couche atomique. Pour de telles précisions, le contrôle de la croissance se fait par l'intermédiaire des variations de réflexion d'un faisceau d'électrons rapides (voir schéma ci-dessous). La cellule de Knudsen également montrée ci-dessous constitue un canon à atomes. Il s'agit d'un four dans lequel est chauffé l'élément jusqu'à évaporation. Les atomes sont alors éjectés à grande vitesse par un orifice de petit diamètre, ceci en raison du vide poussé se trouvant dans l'enceinte de croissance.

*Croissance de structures complexes de haute qualité (en laboratoire)
ex. : puits quantiques multiples GaAs/AlAs (e ~ 10 nm)*

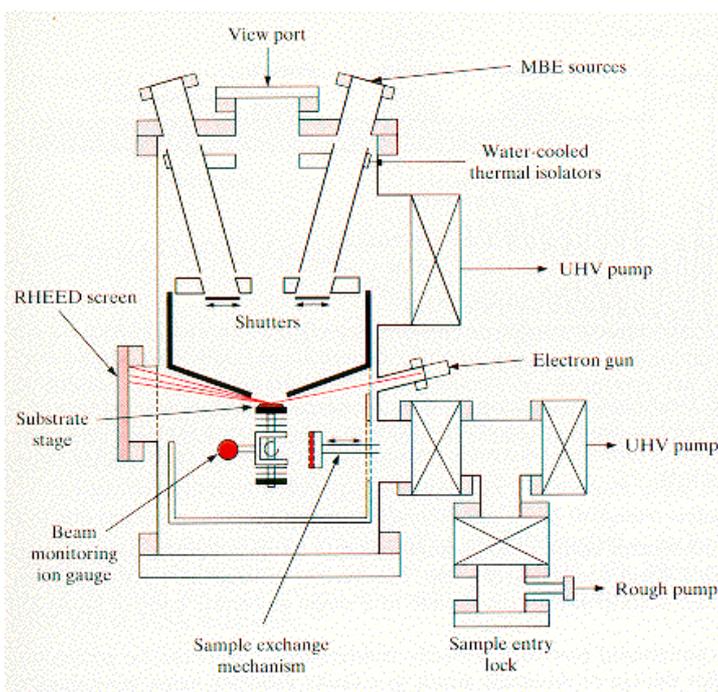
Cellule de Knudsen



*Contrôle du processus en
temps réel*

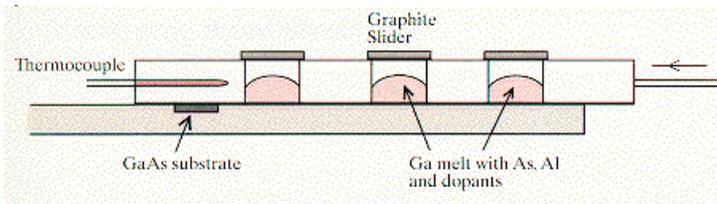


Réflexion d'électrons rapides



Epitaxie en phase liquide (Liquid Phase Epitaxy, LPE)

L'épitaxie en phase liquide est également utilisée pour les dispositifs ne demandant pas trop de précision. Il s'agit de faire cristalliser successivement des composés S-C fondus dans des creusets séparés. Les différentes couches sont déposées par simple déplacement des creusets sur la surface du cristal. Cette technique à l'avantage d'être très simple et donc d'être peu coûteuse. La fabrication à grande échelle de diodes laser en GaAs est un exemple typique d'application de cette technique.



- coût réduit

- grande production possible

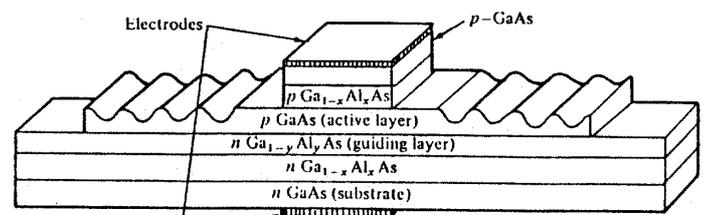
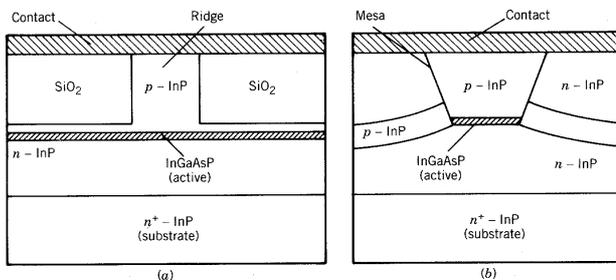
Ex. : GaAs laser diodes

2. Gravure des semiconducteurs

Pour réaliser des structures complexes faisant intervenir une géométrie tri-dimensionnelle, on doit souvent avoir recours à des techniques de gravure. Pour rappeler le type de structures complexes rencontrés dans les applications optiques nous avons repris ci-dessous les schémas des structures de certaines diodes lasers. Dans les applications électroniques usuelles la gravure est beaucoup utilisée pour la réalisation de circuits intégrés.

Création des structures tri-dimensionnelles

ex. : double hétérojonction confinée, réflecteurs à réseau

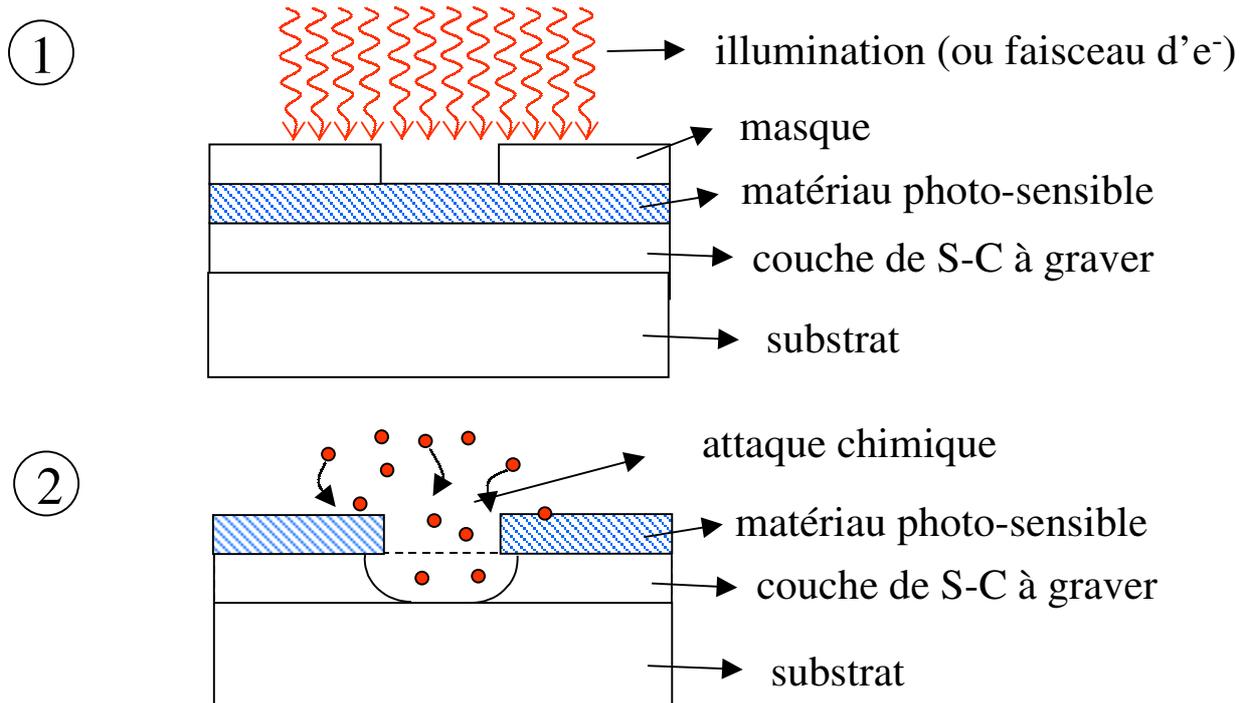


(application usuelle de la gravure : circuits intégrés)

La lithographie

La lithographie consiste à réaliser une gravure par attaque chimique. Son principe est décrit ci-dessous. Un film photosensible est déposé sur le S-C à graver. Un masque déterminant la forme de la gravure est déposé sur le film avant exposition. Après exposition du film à la lumière (ou un faisceau d'électrons) le matériau photosensible est enlevé chimiquement. On attaque ensuite chimiquement le S-C découvert.

Principe: gravure chimique par technique de masquage



« Wet Chemical Etching » : peu précis (0.5 μm)

Le « wet chemical etching » désigne la technique basée sur l'attaque chimique en phase liquide. Cette technique est peu précise car l'attaque chimique est isotrope et des effets de bord inévitables (représentés schématiquement ci-dessus) rendent la précision limitée à 0.5 micron. Une grande amélioration de la technique de lithographie a été réalisée avec le remplacement de l'attaque chimique par le bombardement d'ions (appelé par contraste avec la première technique « dry etching »). Le bombardement est fortement anisotrope (directionnel) et les gravures sont donc beaucoup plus précises.

Améliorations: bombardement d'ions (Dry Etching)

attaque directionnelle \Rightarrow précision