

Physique générale

Chapitre 6 - Physique moderne

Marc HAELTERMAN

Professeur

D/2011/0098/010

1e édition – Tirage 2011-12/2-S

PHYS-H-100_F

En application de l'article 23 du Décret du 31 mars 2004, l'auteur consent à mettre son support de cours obligatoire à la disposition des étudiants régulièrement inscrits. Toute reproduction, communication au public, commercialisation ou autre forme d'exploitation de ce fascicule est interdite et pourra faire l'objet de poursuites disciplinaires, civiles et/ou pénales, conformément à la loi du 30 juin 1994 relative au droit d'auteur et aux droits voisins, à moins qu'il s'agisse d'une reproduction effectuée par un étudiant régulièrement inscrit, dans le cadre de son usage strictement privé d'étudiant, qui ne porte pas préjudice à l'exploitation normale de l'oeuvre.

Toute citation doit mentionner le nom de l'auteur et la source.

Tél. : 02-649 97 80 – Fax : 02-647 79 62 – [Http://www.ulb.ac.be/pub](http://www.ulb.ac.be/pub) – E-mail : mpardoen@ulb.ac.be

« Un peuple qui discute librement est un peuple qui vit et qui marche, un peuple qui ne discute pas est un peuple qui meurt. »

Pierre-Théodore Verhaegen (1796-1862)
Avocat, homme politique libéral, fondateur de l'ULB.

Le label FSC : la garantie d'une gestion responsable des forêts

Les Presses Universitaires de Bruxelles s'engagent !

Les P.U.B. impriment depuis de nombreuses années les syllabus sur du papier recyclé. Les différences de qualité constatées au niveau des papiers recyclés ont cependant poussé les P.U.B. à se tourner vers un papier de meilleure qualité et surtout porteur du label FSC.

Sensibles aux objectifs du FSC et soucieuses d'adopter une démarche responsable, les P.U.B. se sont conformé aux exigences du FSC et ont obtenu en avril 2010 la certification FSC (n° de certificat COC spécifique aux P.U.B. : CU-COC-809718-HA).

Seule l'obtention de ce certificat autorise les P.U.B. à utiliser le label FSC selon des règles strictes. Fortes de leur engagement en faveur de la gestion durable des forêts, les P.U.B. souhaitent dorénavant imprimer tous les syllabus sur du papier certifié FSC. Le label FSC repris sur les syllabus vous en donnera la garantie.

Qu'est-ce que le FSC ?

FSC signifie "Forest Stewardship Council" ou "Conseil de bonne gestion forestière". Il s'agit d'une organisation internationale, non gouvernementale, à but non lucratif qui a pour mission de promouvoir dans le monde une gestion responsable et durable des forêts.

Se basant sur dix principes et critères généraux, le FSC veille à travers la certification des forêts au respect des exigences sociales, écologiques et économiques très poussées sur le plan de la gestion forestière.

Quelles garanties ?

Le système FSC repose également sur la traçabilité du produit depuis la forêt certifiée dont il est issu jusqu'au consommateur final. Cette traçabilité est assurée par le contrôle de chaque maillon de la chaîne de commercialisation/transformation du produit (Chaîne de Contrôle : Chain of Custody – COC). Dans le cas du papier et afin de garantir cette traçabilité, aussi bien le producteur de pâte à papier que le fabricant de papier, le grossiste et l'imprimeur doivent être contrôlés. Ces contrôles sont effectués par des organismes de certification indépendants.

Les 10 principes et critères du FSC

1. L'aménagement forestier doit respecter les lois nationales, les traités internationaux et les principes et critères du FSC.
2. La sécurité foncière et les droits d'usage à long terme sur les terres et les ressources forestières doivent être clairement définis, documentés et légalement établis.
3. Les droits légaux et coutumiers des peuples indigènes à la propriété, à l'usage et à la gestion de leurs territoires et de leurs ressources doivent être reconnus et respectés.
4. La gestion forestière doit maintenir ou améliorer le bien-être social et économique à long terme des travailleurs forestiers et des communautés locales.
5. La gestion forestière doit encourager l'utilisation efficace des multiples produits et services de la forêt pour en garantir la viabilité économique ainsi qu'une large variété de prestations environnementales et sociales.
6. Les fonctions écologiques et la diversité biologique de la forêt doivent être protégées.
7. Un plan d'aménagement doit être écrit et mis en œuvre. Il doit clairement indiquer les objectifs poursuivis et les moyens d'y parvenir.
8. Un suivi doit être effectué afin d'évaluer les impacts de la gestion forestière.
9. Les forêts à haute valeur pour la conservation doivent être maintenues (par ex : les forêts dont la richesse biologique est exceptionnelle ou qui présentent un intérêt culturel ou religieux important). La gestion de ces forêts doit toujours être fondée sur un principe de précaution.
10. Les plantations doivent compléter les forêts naturelles, mais ne peuvent pas les remplacer. Elles doivent réduire la pression exercée sur les forêts naturelles et promouvoir leur restauration et leur conservation. Les principes de 1 à 9 s'appliquent également aux plantations.



Le label FSC apposé sur des produits en papier ou en bois apporte la garantie que ceux-ci proviennent de forêts gérées selon les principes et critères FSC.

® FSC A.C. FSC-SECR-0045

FSC, le label du bois et du papier responsable

Plus d'informations ?

www.fsc.be

A la recherche de produits FSC ?

www.jecherchedufsc.be

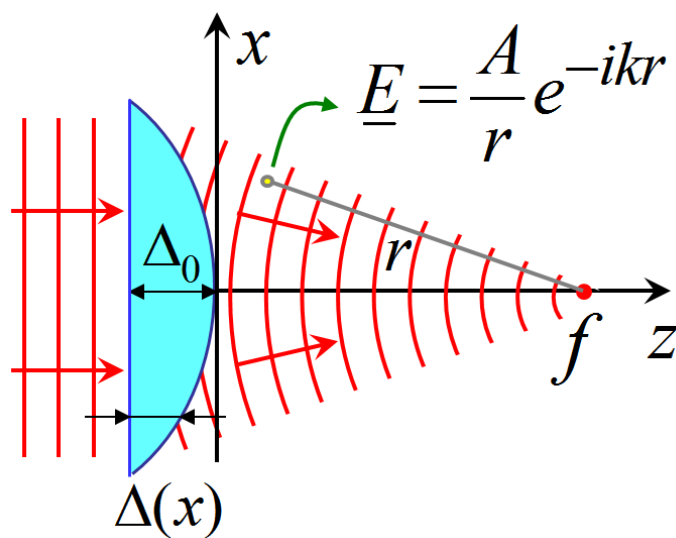


Cours de Physique Générale

Marc Haelterman

Chapitre 6

Physique moderne



Physique moderne

Le présent chapitre est principalement consacré à une introduction aux deux branches de la physique moderne que constituent la mécanique quantique et la relativité restreinte. Bien que ces deux théories soient apparues dès le début du 20^{ème} siècle, elles sont considérées comme des « *sciences modernes* » dans la mesure où elles sont encore actuellement en développement. A titre d'exemple, la théorie de la mécanique quantique a depuis peu été développée dans le cadre de la théorie de l'information en vue d'applications telles que l'ordinateur quantique ou la cryptographie quantique. De même, la théorie de la relativité d'Einstein fait encore l'objet de nombreux développements dans le domaine de la cosmologie dont le but est de comprendre la structure de notre univers. Nous allons commencer le chapitre avec le thème de la mécanique quantique. Pour bien comprendre les bases de la mécanique quantique il est essentiel de maîtriser la notion d'ondes dispersives, c'est pourquoi je propose dans la première section d'étudier les ondes électromagnétiques dispersives. Il s'agit d'ondes qui se propagent dans la matière et non plus simplement dans le vide comme on l'a considéré jusqu'à présent.

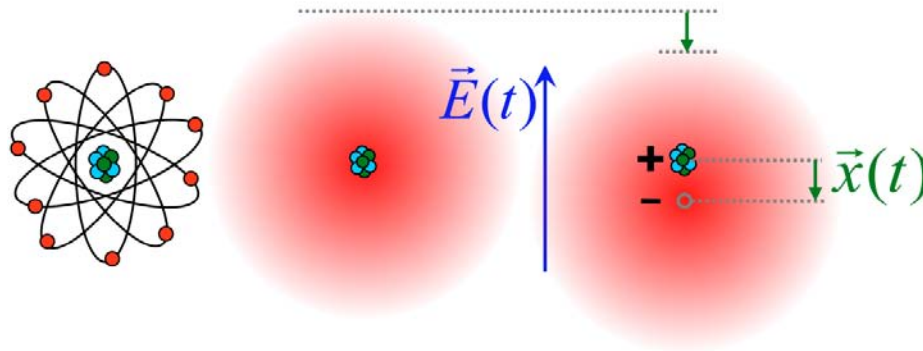
1. Ondes dispersives

Pour présenter la notion d'ondes dispersives nous allons considérer ici les ondes électromagnétiques (EM) se propageant dans les milieux diélectriques. La description de la propagation des ondes EM en milieu diélectrique exige un développement mathématique relativement long mais il a l'avantage d'être d'une grande valeur pédagogique car, d'une part, il utilise des concepts et des outils mathématiques dont vous aurez besoin dans la suite de vos études (en particulier, les phaseurs) et, d'autre part, il est d'une grande simplicité conceptuelle dans la mesure où il se base sur la physique classique du 19^{ème} siècle. De ce point de vue il a également une grande valeur historique car il s'agit d'un développement proposé par Lorentz, avant l'avènement de la mécanique quantique, pour expliquer de façon remarquablement simple les propriétés de la propagation lumineuse dans la matière.

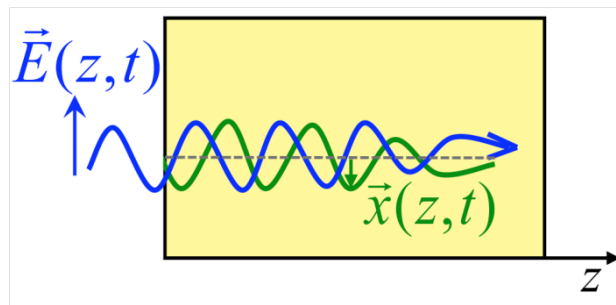
1.1 Densité de courant de charge liée

Comme on l'a vu dans le chapitre d'électrostatique, un milieu diélectrique est caractérisé par l'absence de charges libres. Les électrons sont liés aux noyaux des atomes qui constituent le milieu et forment autour de ceux-ci ce que l'on appelle un « *nuage électronique* ». Si un champ électrique est appliqué à un tel milieu, les nuages électroniques des atomes se déplacent dans le champ et une polarisation du milieu apparaît (création de dipôles électriques microscopiques). Ceci est représenté

schématiquement ci-dessous. Sur ce schéma, je considère le cas général d'un champ électrique variable $\vec{E}(t)$ provoquant une variation de position des centres de charge des nuages électroniques. J'introduis pour cela le vecteur $\vec{x}(t)$ représentant le déplacement du nuage électronique par rapport à sa position de repos.



Si le champ extérieur \vec{E} provient d'une onde harmonique incidente sur le milieu diélectrique, soit $\vec{E}(z,t) = E_0 \cos(kz - \omega t) \vec{1}_x$, la polarisation du milieu subit des variations harmoniques. C'est-à-dire que les positions \vec{x} des charges liées subissent des oscillations harmoniques. Remarquez que l'on suppose une onde plane de direction parallèle à l'axe z , de sorte que les vecteurs \vec{x} de déplacement des nuages électroniques constituent un champ vectoriel dépendant de z et du temps, soit $\vec{x}(z,t)$.



A la section 6 du chapitre 2 consacrée au condensateur nous avons vu qu'un champ électrique provoque une polarisation du milieu diélectrique, ce qui se manifeste dans le condensateur par l'apparition de deux fines pellicules de charges électriques opposées. Ces pellicules de charge sont responsables d'une diminution du champ électrique dans le milieu diélectrique. Dans la situation présente, nous ne pouvons pas nous servir de cette approche car le milieu diélectrique n'est pas limité comme il l'était par les plaques du condensateur. En effet, pour des raisons évidentes de simplicité mathématique, l'onde que l'on étudie est supposée être une onde plane d'extension illimitée se propageant dans un milieu également supposé illimité. Il n'y a dès lors pas de zones d'accumulation de charge contrairement au cas du condensateur. Autrement dit, le milieu reste parfaitement neutre et en conséquence il n'y a pas de champ électrique induit par la polarisation des atomes du milieu. Néanmoins, comme on le verra ci-dessous, la polarisation des atomes décrite par la fonction $\vec{x}(z,t)$ est responsable d'une modification de l'onde par rapport à ce qu'elle est dans le vide. Nous verrons que cette modification est due au courant induit par la polarisation variable des atomes c'est-à-dire, par le courant dû au mouvement des charges liées.

S'il y a un mouvement oscillant des charges, cela signifie effectivement qu'il y a un courant oscillant dans le milieu. Le courant peut être décrit par la densité de courant $\vec{J} = \eta_c q_e \vec{v}$ où η_c est le nombre d'électrons par unité de volume, q_e est la charge électronique (dont nous n'explicitons pas le signe) et \vec{v} est la vitesse des électrons. Mais dans le cas présent c'est le mouvement des nuages électroniques qui est responsable du courant. C'est donc la charge totale des nuages électroniques qui bouge et il est dès lors plus approprié de considérer que la densité de courant est déterminée par la densité d'atomes η_a , la charge totale q des nuages électroniques et par leur vitesse de déplacement \vec{v} , soit $\vec{J} = \eta_a q \vec{v}$. La vitesse des nuages électroniques est bien sûr donnée par la dérivée de leur position, soit $\vec{v} = \partial \vec{x} / \partial t$, ce qui nous conduit à l'expression suivante de la densité de courant :

densité de courant des charges liées : $\vec{J} = \eta_a q \frac{\partial \vec{x}}{\partial t}$

1.2 Équation d'onde

Pour obtenir l'équation d'onde EM dans le milieu diélectrique, on a bien entendu recours aux équations de Maxwell. Mais cette fois, contrairement à ce que l'on a fait pour les ondes EM dans le vide, on ne pourra plus considérer que la densité de courant \vec{J} y est nulle.

$$\left\{ \begin{array}{l} \text{rot } \vec{E} = -\frac{\partial \vec{B}}{\partial t} \\ \text{rot } \vec{B} = \mu_0 \vec{J} + \mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t} \end{array} \right.$$

L'équation d'onde des milieux diélectriques s'obtient en prenant le rotationnel de la loi de Faraday et en y substituant l'expression du rotationnel de \vec{B} fournie par la loi d'Ampère-Maxwell.

$$\text{rot}(\text{rot } \vec{E}) = -\frac{\partial \text{rot } \vec{B}}{\partial t} = -\mu_0 \frac{\partial \vec{J}}{\partial t} - \mu_0 \epsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2}$$

On a vu précédemment que le rotationnel du rotationnel est égal à l'opposé du laplacien lorsque la divergence de \vec{E} est nulle. Or, ici comme dans le vide la divergence de \vec{E} est nulle car le milieu diélectrique est neutre, ce qui correspond à une densité de charge électrique ρ nulle, soit $\text{div } \vec{E} = \rho / \epsilon_0 = 0$. On peut donc écrire :

$$\Delta \vec{E} = \mu_0 \frac{\partial \vec{J}}{\partial t} + \mu_0 \epsilon_0 \frac{\partial^2 \vec{E}}{\partial t^2}$$

Pour simplifier le développement, nous choisissons l'axe z parallèle à la direction de propagation de l'onde plane. De cette manière le champ électrique de l'onde est uniforme en x et y , ce qui se traduit par des dérivées nulles par rapport aux variables x et y . La variable z constitue donc l'unique dimension spatiale de l'équation d'onde. En regroupant dans l'expression ci-dessous le terme de dérivée seconde temporelle avec le terme de dérivée seconde spatiale, on donne à l'équation la forme de l'équation d'onde standard avec un terme supplémentaire dans le membre de droite. Le membre de gauche représente les évolutions spatiales et temporelles du champ électrique. En l'absence de courant cette expression est égale à zéro, indiquant que les variations spatiales sont identiques aux variations temporelles (c'est ce qui conduit à la notion d'onde). En présence de courant, le membre de droite est non nul. Ce terme doit être interprété comme un terme de « source » de champ électromagnétique puisqu'on sait que le courant génère effectivement un champ magnétique et comme il s'agit d'un courant variable, on a affaire à un champ magnétique variable qui génère par induction un champ électrique variable. On doit donc s'attendre à ce que l'onde dans le milieu diélectrique soit différente d'une onde se propageant dans le vide.

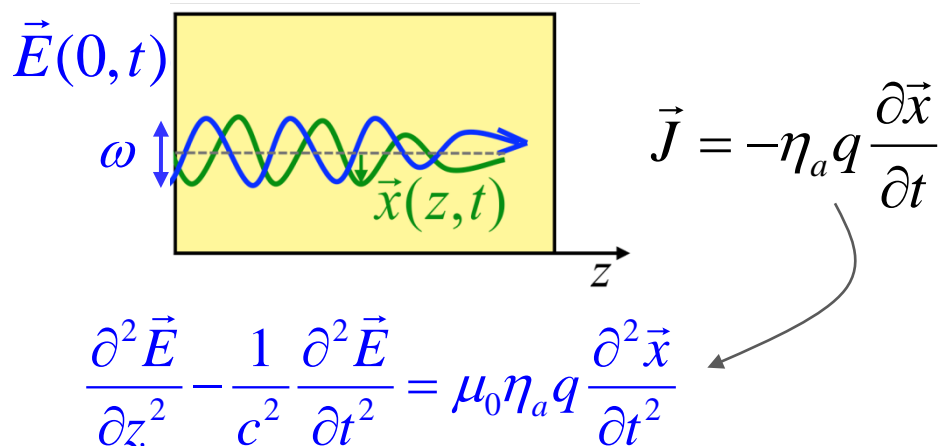
onde plane (1 dimension) :

$$\frac{\partial^2 \vec{E}}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} = \mu_0 \frac{\partial \vec{J}}{\partial t} \quad \text{"source de champ"}$$

Pour terminer, nous substituons la densité de courant \vec{J} par son expression donnée plus haut, soit $\vec{J} = \eta_a q \partial \vec{x} / \partial t$. Le terme de source devient ainsi proportionnel à la dérivée seconde temporelle de la coordonnée de position des nuages électroniques. Autrement dit, c'est un terme proportionnel à l'accélération des charges. Je rappelle ici que \vec{x} est un champ vectoriel représentant en toute valeur de z les positions des nuages électroniques par rapport à leur position de repos.

$$\frac{\partial^2 \vec{E}}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} = \mu_0 \eta_a q \frac{\partial^2 \vec{x}}{\partial t^2}$$

Le schéma et les expressions mathématiques ci-dessous résument la situation : le champ électrique de l'onde est modifié par rapport à ce qu'il serait dans le vide en raison de la présence des courants de charges liées dans le diélectrique, ce qui s'exprime ici au travers de l'équation d'onde modifiée. La double flèche dessinée à la gauche du milieu diélectrique avec la pulsation ω illustre la condition aux limites que nous considérerons dans la suite : nous supposons que l'onde dans le milieu diélectrique est créée par l'arrivée d'une onde plane harmonique se propageant dans le vide à gauche du milieu. Le champ en $z=0$ a donc une simple évolution harmonique à la pulsation ω et il en sera bien sûr de même de l'onde qui se propage dans le diélectrique.



$$\vec{J} = -\eta_a q \frac{\partial \vec{x}}{\partial t}$$

$$\frac{\partial^2 \vec{E}}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} = \mu_0 \eta_a q \frac{\partial^2 \vec{x}}{\partial t^2}$$

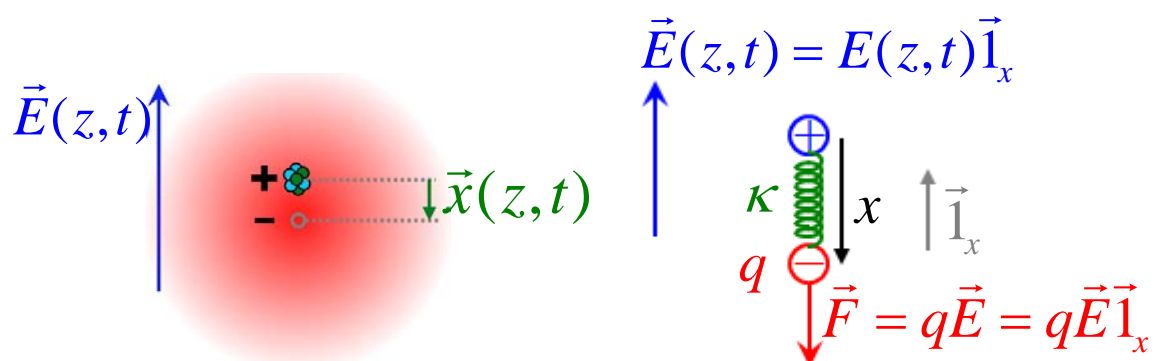
1.3 Relation constitutive

L'équation d'onde modifiée obtenue ci-dessus montre que l'évolution du champ électrique est déterminée par l'accélération des charges liées. Ce modèle de propagation d'onde n'est bien sûr pas complet car nous avons deux inconnues, les champs \vec{x} et \vec{E} , pour une seule équation. Mais ces deux inconnues sont naturellement liées puisque c'est le champ électrique \vec{E} qui est responsable du déplacement des nuages électroniques représenté par le champ vectoriel \vec{x} . Du point de vue mathématique il apparaît dès lors évident que les positions \vec{x} des nuages électroniques sont fonctions du champ électrique \vec{E} . Notre but est de trouver le lien mathématique entre \vec{x} et \vec{E} afin de pouvoir résoudre l'équation d'onde EM des milieux diélectriques. Si nous avons ce lien sous la forme d'une fonction $\vec{x}(\vec{E})$ alors l'équation d'onde modifiée deviendrait effectivement une équation à une seule inconnue, ce qui nous permettrait de la résoudre.

$$\frac{\partial^2 \vec{E}}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} = \mu_0 \eta_a q \frac{\partial^2 \vec{x}(\vec{E})}{\partial t^2}$$

Le lien entre \vec{x} et \vec{E} , s'appelle la « *relation constitutive* » du matériau diélectrique. Chaque matériau possède une relation constitutive qui lui est propre car la déformation des nuages électroniques sous l'effet du champ électrique dépend de la configuration atomique qui caractérise le matériau. Pour trouver la relation constitutive des milieux diélectriques il nous faut donc avant tout choisir un modèle mathématique du matériau permettant d'établir un lien entre le champ électrique et la déformation des nuages électroniques. Un tel modèle d'interaction champ-matière devrait en principe se faire dans le cadre de la mécanique quantique, ce qui nous est bien sûr impossible à ce stade, c'est pourquoi je propose ici le modèle de Lorentz qui a l'avantage d'être à la fois simple et suffisamment fidèle à la réalité. A l'image de ce que l'on a fait pour modéliser les mouvements des atomes dans les cristaux au chapitre 5, ce modèle suppose que les nuages électroniques sont liés aux noyaux atomiques par l'intermédiaire d'un potentiel parabolique ou harmonique (si on suppose des déplacements \vec{x} de petite amplitude, il est effectivement légitime de considérer un potentiel harmonique car, dans ces conditions, celui-ci est une approximation valable de tout potentiel de liaison en « U »).

En faisant ce choix de potentiel la dynamique des nuages électroniques est assimilée à celle d'un oscillateur linéaire correspondant à une force de rappel élastique $\vec{f} = -\kappa\vec{x}$. Dès lors, la force $\vec{F} = q\vec{E}$ exercée par le champ électrique sur les nuages électroniques apparaît comme une force extérieure harmonique s'exerçant sur un oscillateur linéaire. La force \vec{F} a la direction du champ \vec{E} et donc les déplacements \vec{x} ont également la direction du champ \vec{E} . Nous choisissons ici l'axe x dans la direction du champ \vec{E} de sorte que les déplacements \vec{x} se fassent aussi parallèlement à l'axe x . Ceci est schématisé sur la figure ci-dessous (notez que la charge q sur ce schéma est négative puisqu'il s'agit de la charge du nuage électronique).



En tenant compte d'un amortissement des vibrations des nuages électroniques au travers d'un coefficient d'amortissement λ , on obtient la force totale exercée sur un nuage électronique situé en z . La composante en x de cette force est $F = -\kappa x - \lambda \dot{x} + qE$. Sa substitution dans la deuxième loi de Newton nous donne l'équation de l'oscillateur linéaire amorti forcé.

$$\ddot{x} + \frac{\kappa}{m} x + \frac{\lambda}{m} \dot{x} - \frac{qE(z,t)}{m} = 0$$

La solution de cette équation nous donne a priori la relation constitutive recherchée, c'est-à-dire la fonction $\bar{x}(\vec{E})$. Mais la situation n'est pas aussi simple car \bar{x} est donné ici comme la solution d'une équation différentielle temporelle, ce qui implique que sa valeur dépend de toute l'évolution du champ \vec{E} et pas seulement de sa valeur à un instant donné. La relation constitutive ne prend donc pas la forme d'une simple fonction $\bar{x}(\vec{E})$.

Une relation simple entre \bar{x} et \vec{E} peut toutefois être obtenue si on fait la supposition que le champ électrique de l'onde dans le diélectrique a la forme d'une onde plane harmonique, soit $\vec{E}(z,t) = E_\omega \cos(kz - \omega t) \vec{1}_x$, comme dans le vide. Cette démarche peut paraître étonnante car cette supposition revient à donner la solution de l'équation d'onde modifiée sans se prendre la peine de résoudre cette équation. Mais ceci n'est pas tout à fait vrai car on ne fait que supposer que l'onde dans le diélectrique est plane et harmonique et on ne dit rien quant aux paramètres de cette onde, en particulier, en ce qui concerne son amplitude E_ω et son nombre d'onde k . On supposera donc dans la suite que E_ω et k sont des inconnues du problème qui nous seront données par la résolution de l'équation d'onde modifiée. La pulsation ω peut quant à elle être considérée comme une constante imposée ; c'est la pulsation de l'onde incidente sur le milieu diélectrique. Il n'y a effectivement aucune raison pour que la pulsation de l'onde soit modifiée au passage du vide vers le milieu diélectrique.

En résumé, on suppose que la solution de l'équation d'onde modifiée est de la forme $\vec{E}(z,t) = E_\omega \cos(kz - \omega t) \vec{1}_x$. La substitution de cette expression dans l'équation d'onde modifiée nous permettra, d'une part, de confirmer a posteriori qu'il s'agit bien d'une de ses solutions et, d'autre part, de calculer l'amplitude E_ω et le nombre d'onde k .

Avec l'expression du champ électrique $\vec{E}(z,t) = E_\omega \cos(kz - \omega t) \vec{1}_x$, on peut dire que la composante en x de la force extérieure \vec{F} est égale à $F = qE_\omega \cos(kz - \omega t)$ et l'équation du mouvement devient celle de l'oscillateur linéaire amorti forcé que nous avons étudiée en détail au chapitre 5.

$$\ddot{x} + \frac{\kappa}{m} x + \frac{\lambda}{m} \dot{x} - \frac{qE_\omega}{m} \cos(kz - \omega t) = 0$$

En introduisant les paramètres habituels de la description mathématique de l'oscillateur linéaire amorti et forcé, soit $\omega_0 = (\kappa/m)^{1/2}$, $\alpha = \lambda/(2m)$, et $a = qE_\omega/m$, nous obtenons l'équation sous sa forme canonique.

$$\ddot{x} + \omega_0^2 x + 2\alpha \dot{x} - a \cos(kz - \omega t) = 0$$

En introduisant la notion de phaseur, nous allons transformer cette équation du mouvement en une simple équation algébrique.

$$x \equiv \text{Re} \left[\underline{X} e^{-i\omega t} \right]$$

"convention de signe des ondes"

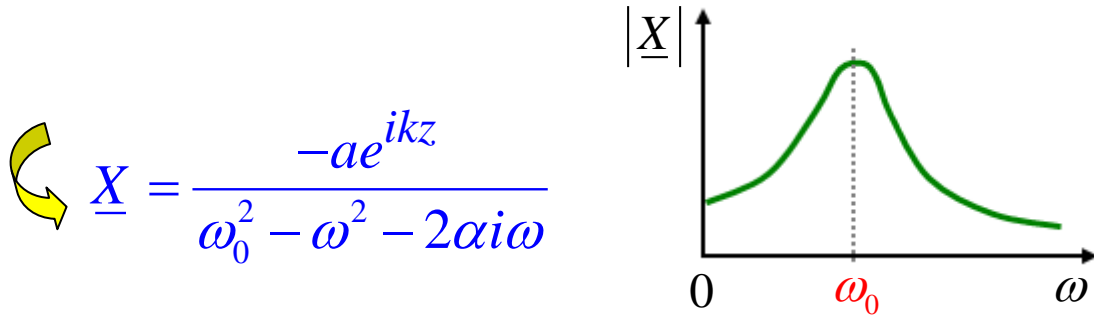
Remarque : Notez que dans la définition du phaseur ci-dessus, on a changé le sens de rotation du nombre complexe tournant par rapport à la définition du phaseur introduite lors de l'étude des systèmes oscillants. Au lieu de $\exp(i\omega t)$, on a $\exp(-i\omega t)$. Cette différence de signe n'a aucune importance en pratique puisque le mouvement harmonique réel obtenu en prenant la partie réelle du phaseur multiplié par $\exp(-i\omega t)$ ou par $\exp(i\omega t)$ sera exactement le même [la partie réelle de $\exp(-i\omega t)$ ou $\exp(i\omega t)$ est toujours $\cos(\omega t)$]. Le signe moins est introduit ici car c'est la convention habituelle utilisée pour la description des ondes en physique. C'est notamment celle que vous utiliserez en deuxième année au cours de mécanique quantique et statistique. Nous comprendrons plus loin l'intérêt de cette convention.

L'équation algébrique du phaseur \underline{X} est dérivée ci-dessous de la façon habituelle. La seule subtilité pour l'obtenir est de tenir compte du fait que le terme de forçage $a \cos(kz - \omega t)$ a pour phaseur $a \exp(ikz)$. Je vous rappelle que l'argument du phaseur (ici kz) représente le déphasage de l'oscillation harmonique correspondante. Le déphasage kz rend compte du fait que la force extérieure provient d'une onde harmonique de champ électrique ; la force extérieure oscille en tout point z de façon harmonique avec un déphasage valant kz .

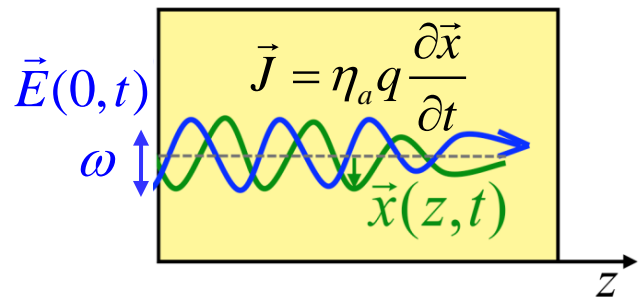
$$\cancel{\text{Re}}[-\omega^2 \cancel{\underline{X} e^{-i\omega t}} + \omega_0^2 \underline{X} e^{-i\omega t} - 2\alpha i \omega \cancel{\underline{X} e^{-i\omega t}} + a e^{ikz} \cancel{e^{-i\omega t}}] = 0$$

$$\curvearrowright -\omega^2 \underline{X} + \omega_0^2 \underline{X} - 2\alpha i \omega \underline{X} + a e^{ikz} = 0$$

En isolant \underline{X} dans le membre de gauche, on obtient une expression du phaseur analogue à celle que l'on avait pour l'OLA forcé. La seule nuance est qu'ici le phaseur a une valeur dépendante de z (le phaseur est plus précisément une fonction complexe de z). Mais ceci n'a que peu d'importance sur le résultat du forçage harmonique dans la mesure où le module de $\exp(ikz)$ vaut l'unité. Ainsi, comme pour l'OLA forcé, on voit que l'amplitude des oscillations des nuages électroniques (donnée par le module du phaseur) est résonante lorsque la pulsation de l'onde de champ électrique ω est égale à la fréquence de résonance des nuages électroniques ω_0 .



Notre démarche est résumée ci-contre : nous avons postulé l'existence d'une solution harmonique pour l'équation d'onde modifiée du milieu diélectrique (notez que l'équation d'onde est ici écrite en notation scalaire E représentant la composante en x du champ \vec{E}). Si cette solution existe, elle met les nuages électroniques en mouvement harmonique tel que décrit par le phaseur \underline{X} obtenu ci-dessus. Le mouvement réel $x(z,t)$ des nuages électroniques est donné en prenant la partie réelle de ce phaseur multiplié par $\exp(-i\omega t)$. Si on substitue cette valeur de la fonction $x(z,t)$ dans l'équation on pourra vérifier si la solution d'onde harmonique postulée au départ est bien solution du problème et on pourra en calculer l'amplitude et le nombre d'onde. Mais avant de procéder à ce calcul nous allons ici analyser l'interprétation physique du phaseur \underline{X} .



$$\left\{ \begin{array}{l} \frac{\partial^2 E}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2 E}{\partial t^2} = -\mu_0 \eta_a q \frac{\partial^2 x}{\partial t^2} \\ E(z,t) = E_\omega \cos(kz - \omega t) \end{array} \right.$$

$$a = qE_\omega / m$$

$$\left\{ \begin{array}{l} \underline{X} = \frac{-ae^{ikz}}{\omega_0^2 - \omega^2 - 2\alpha i\omega} \\ x = \text{Re} \left[\underline{X} e^{-i\omega t} \right] \end{array} \right.$$

En remplaçant le coefficient de forçage « a » par sa valeur qE_ω / m , on trouve l'expression suivante du phaseur :

$$a = qE_\omega / m \quad \Rightarrow \quad \underline{X} = -\frac{q}{m} \frac{E_\omega e^{ikz}}{\omega_0^2 - \omega^2 - 2\alpha i\omega}$$

Cette expression est la relation constitutive du milieu diélectrique exprimée en termes de phaseur. On y reconnaît en particulier le phaseur du champ électrique $\underline{E} = E_\omega \exp(ikz)$, ce qui nous permet de constater que la relation constitutive en termes de phaseur consiste en une simple relation de proportionnalité entre le champ électrique et le déplacement des nuages électroniques. Son caractère complexe nous indique par contre que les grandeurs réelles correspondantes, soit \vec{x} et \vec{E} , ne sont pas simplement proportionnelles entre elles car leur phaseurs n'ont pas le même argument (ou la même « phase »), ce qui révèle effectivement un décalage temporel entre les variations harmoniques de \vec{x} et \vec{E} (voir chapitre 5).

1.3.1 Limite Électrostatique

Pour l'interprétation physique du phaseur des déplacements de charge \underline{X} il est intéressant d'analyser la situation d'un champ électrique statique correspondant dans notre développement à une pulsation ω nulle. Si on fait tendre ω vers zéro, le champ électrique devient simplement $\vec{E}(z, t) = E_\omega \vec{1}_x$ ce qui physiquement signifie que l'on se retrouve dans la situation d'un champ électrostatique appliqué au diélectrique. On doit donc s'attendre à retrouver le résultat obtenu en électrostatique lors de la modélisation du condensateur diélectrique (voir chapitre 2). On vérifie cela aisément en constatant que lorsque $\omega = 0$ le phaseur \underline{X} prend la valeur réelle $\underline{X} = qE_{\omega=0} / (m\omega_0^2)$ qui, avec la définition de la pulsation propre $\omega_0 = (\kappa / m)^{1/2}$, devient $\underline{X} = qE_{\omega=0} / \kappa$.

$$\underline{X} = \frac{q}{m\omega_0^2} E_{\omega=0} = \frac{q}{\kappa} E_{\omega=0}$$

Dans cette expression le facteur qE_ω représente la force F appliquée au nuage électronique. Il est donc normal de trouver comme valeur de \underline{X} le rapport de cette force et de la constante de rappel κ , soit $\underline{X} = F / \kappa$. En effet, d'une part, dans la limite électrostatique $\omega = 0$ le phaseur devient identique à la grandeur réelle correspondante, soit $x = \underline{X}$ et, d'autre part, la force de rappel élastique impose la relation $F = \kappa x$.

La limite électrostatique nous permet ainsi d'introduire naturellement la notion de susceptibilité électrique. Nous avons vu en électrostatique que la susceptibilité χ est donnée par l'expression $\eta_a q \alpha / \varepsilon_0$ où α est par définition le coefficient de déformabilité de l'atome. Plus précisément ce coefficient lie le champ électrique à la longueur « d » du dipôle résultant du déplacement des nuages électroniques (à ne pas confondre avec le coefficient d'amortissement α ci-dessus), soit $d = \alpha \|\vec{E}\|$ (voir chapitre 2, section 6.9). Dans le présent contexte, cette longueur d n'est rien d'autre que le module du phaseur des déplacements, soit $d = x = \underline{X}$, ce qui, avec la relation $\underline{X} = q E_{\omega=0} / \kappa$ trouvée ci-dessus, nous montre que le coefficient de déformabilité α vaut q / κ . La susceptibilité à fréquence nulle $\chi = \eta_a q \alpha / \varepsilon_0$ peut donc être réécrite comme suit : $\chi = \eta_a q^2 / (\varepsilon_0 \kappa)$.

$$\left\{ \begin{array}{l} \chi = \frac{1}{\varepsilon_0} \eta_a q \alpha \\ d \equiv \alpha \|\vec{E}\| \end{array} \right. \quad \Rightarrow \quad \alpha = \frac{q}{\kappa} \quad \text{et} \quad \chi(\omega = 0) = \frac{\eta_a q^2}{\varepsilon_0 \kappa}$$


1.3.2 Susceptibilité électrique complexe

Revenons au cas général des ondes EM pour lesquelles la fréquence angulaire ω est différente de zéro. L'expression du phaseur \underline{X} (reprise ci-dessous) montre que ce dernier est proportionnel au phaseur de champ électrique $E_\omega \exp(ikz)$. Ceci nous permet d'interpréter cette expression comme une généralisation de la relation $d = \alpha \|\vec{E}\|$ de l'électrostatique. Cette généralisation est effectivement assez naturelle puisque la norme du phaseur \underline{X} représente l'amplitude des déformations des nuages électroniques alors que la norme du phaseur de champ électrique représente l'amplitude du champ électrique responsable de ces déformations. L'expression de \underline{X} montre en l'occurrence que l'amplitude de ces oscillations est proportionnelle à celles du champ électrique E_ω , à l'image de la relation $d = \alpha \|\vec{E}\|$ de l'électrostatique. Le coefficient de proportionnalité entre \underline{X} et le phaseur de champ électrique $E_\omega \exp(ikz)$ est donc la généralisation du coefficient de déformabilité α , comme le suggère la ligne pointillée dans l'expression de \underline{X} ci-dessous.

$$\underline{X} = - \frac{q}{m} \frac{E_\omega e^{ikz}}{\omega_0^2 - \omega^2 - 2\alpha i \omega} \quad \rightarrow \quad \equiv \alpha$$

Puisque la déformabilité α des atomes entre dans la définition de la susceptibilité nous pouvons également généraliser au domaine des phaseurs la notion de susceptibilité. La susceptibilité valant par définition $\chi = \eta_a q \alpha / \varepsilon_0$, il suffit effectivement de multiplier le facteur α généralisé donné ci-dessus par $\eta_a q / \varepsilon_0$, pour obtenir une généralisation de la susceptibilité électrique pour les ondes :

$$\alpha = \frac{q}{m} \frac{1}{\omega_0^2 - \omega^2 - 2\alpha i \omega} \quad \text{et} \quad \chi = \frac{\eta_a q \alpha}{\varepsilon_0}$$

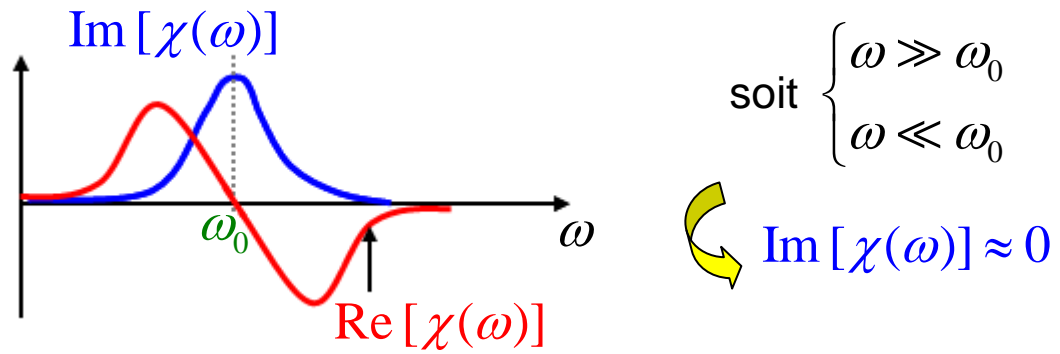


$$\chi(\omega) \equiv \frac{\eta_a q^2}{\varepsilon_0 m} \frac{1}{\omega_0^2 - \omega^2 - 2\alpha i \omega}$$

Remarquez que cette généralisation fait de la susceptibilité une fonction complexe de la pulsation ω . En exprimant la déformabilité α en fonction de la susceptibilité χ , soit $\alpha = \varepsilon_0 \chi / (\eta_a q)$, le lien entre le phaseur \underline{X} des mouvements électroniques et le phaseur de champ électrique devient le suivant :

$$\underline{X} = -\frac{\varepsilon_0}{\eta_a q} \chi(\omega) E_\omega e^{ikz}$$

Il est facile de voir dans l'expression de χ ci-dessus que sa partie réelle présente un zéro en $\omega = \omega_0$ alors que sa partie imaginaire présente un maximum en ce point (voir graphe ci-dessous). On peut montrer que la partie imaginaire de la susceptibilité est responsable d'une absorption des ondes EM dans le milieu diélectrique (l'onde EM fournit de l'énergie aux oscillateurs linéaires amortis qui représentent ici les nuages électroniques). Comme le maximum de la partie imaginaire se situe en $\omega = \omega_0$ on constate que l'absorption de l'énergie EM est d'autant plus forte que l'on se trouve proche de la résonance des nuages électroniques. Pour simplifier les choses nous éliminerons cette situation dans la suite du développement en nous imposant la restriction d'être très loin de la résonance. Nous supposerons donc que la pulsation est, soit beaucoup plus petite, soit beaucoup plus grande que la pulsation propre des nuages électroniques. Un grand nombre de situations pratiques répondent à cette restriction. C'est le cas, par exemple, des ondes lumineuses dans les milieux transparents tels que le verre : s'il y a transparence l'absorption y est effectivement négligeable et donc la partie imaginaire de la susceptibilité peut être considérée comme nulle. Nous considérerons donc dans la suite que la susceptibilité $\chi(\omega)$ est une fonction réelle.



La situation à laquelle nous sommes arrivés est résumée ci-dessous : après avoir postulé l'existence d'une solution d'onde harmonique plane, nous avons obtenu sous forme d'un phaseur la fonction $x(z, t)$ donnant les déplacements des nuages électroniques en fonction du champ électrique.

$$\frac{\partial^2 E}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2 E}{\partial t^2} = -\mu_0 \eta_a q \frac{\partial^2 x}{\partial t^2} \quad \vec{E}(0, t)$$

$$\begin{cases} E(z, t) = E_\omega \cos(kz - \omega t) \\ \underline{X} = -\frac{\epsilon_0}{\eta_a q} \chi(\omega) E_\omega e^{ikz} \end{cases} \Rightarrow x(z, t) = \text{Re} \left[\underline{X} e^{-i\omega t} \right]$$

1.4 Solution de l'équation d'onde modifiée

Notez que pour le moment, nous ne connaissons pas l'onde de champ électrique, nous en avons seulement la forme mathématique : il s'agit d'une onde plane d'amplitude E_ω et de nombre d'onde k encore indéterminés. Pour connaître ces paramètres de l'onde pour une fréquence angulaire ω donnée, on doit en substituer la forme mathématique dans l'équation d'onde modifiée des milieux diélectriques. Pour la commodité du calcul nous ferons cette substitution après avoir exprimé l'équation dans le domaine des phaseurs. A cette fin nous introduisons la notation complexe de l'onde de champ électrique dans l'équation d'onde modifiée.


$$E(z, t) = \text{Re} \left[E_\omega e^{i(kz - \omega t)} \right]$$

La substitution de cette expression du champ dans l'équation d'onde EM conduit à une relation algébrique car la dérivée seconde en z devient le simple facteur $-k^2$ et la dérivée seconde en temps devient $-\omega^2$.

$$\text{Re}[-k^2 E_\omega e^{i(kz-\omega t)} + \frac{\omega^2}{c^2} E_\omega e^{i(kz-\omega t)}] = \text{Re}[\omega^2 \mu_0 \eta_a q \underline{X} e^{-i\omega t}]$$


En substituant le phaseur $\underline{X} = -\varepsilon_0 \chi / (\eta_a q) E_\omega \exp(ikz)$ et en regroupant tous les termes dans le membre de gauche on obtient l'expression ci-dessous dans laquelle on voit un nombre complexe tournant dont la partie réelle est nulle pour tout temps t et toute position z , ce qui n'est possible que si ce nombre est nul.

$$\text{Re}[-k^2 E_\omega e^{i(kz-\omega t)} + \frac{\omega^2}{c^2} E_\omega e^{i(kz-\omega t)} + \omega^2 \mu_0 \eta_a q \frac{\varepsilon_0}{\eta_a q} \chi(\omega) E_\omega e^{ikz} e^{-i\omega t}] = 0$$




$$-k^2 \cancel{E_\omega} + \frac{\omega^2}{c^2} \cancel{E_\omega} + \omega^2 \mu_0 \eta_a q \frac{\varepsilon_0}{\eta_a q} \chi(\omega) \cancel{E_\omega} = 0$$

On constate que l'amplitude E_ω du champ électrique apparaît dans tous les termes dont la somme est nulle, ce qui révèle que cette amplitude est indéterminée. En conclusion, comme pour une onde se propageant dans le vide, l'amplitude de l'onde peut avoir n'importe quelle valeur. En pratique cette valeur est bien sûr imposée par les conditions aux limites (voir chapitre 5). L'équation des phaseurs nous donne donc uniquement une information sur le nombre d'onde k . On trouve facilement l'expression suivante :



$$k^2 = \frac{\omega^2}{c^2} + \omega^2 \mu_0 \varepsilon_0 \chi(\omega)$$


 $c^2 = 1 / \mu_0 \varepsilon_0$

En se rappelant que la vitesse de la lumière c est donnée par $c = 1/\sqrt{\mu_0 \epsilon_0}$ et que la permittivité du milieu diélectrique est, par définition, donnée par $\epsilon = \epsilon_0(1 + \chi)$, on trouve encore :

$$\begin{aligned} \hookrightarrow k^2 &= \omega^2 \mu_0 \epsilon_0 [1 + \chi(\omega)] = \omega^2 \mu_0 \epsilon(\omega) \\ &\quad \downarrow \\ \epsilon(\omega) &\equiv \epsilon_0 [1 + \chi(\omega)] \end{aligned}$$

Remarquez que, à l'image de la susceptibilité, la permittivité apparaît comme étant une fonction de la pulsation ω de l'onde. Nous trouvons finalement l'expression suivante du nombre d'onde :

$$\hookrightarrow \boxed{k(\omega) = \omega \sqrt{\mu_0 \epsilon(\omega)}}$$

Notre description de l'onde EM dans le milieu diélectrique est maintenant complète, on peut l'écrire comme suit :

$$\begin{cases} E(z, t) = E_0 \cos[k(\omega)z - \omega t] \\ k(\omega) = \omega \sqrt{\mu_0 \epsilon(\omega)} \end{cases}$$

Ce résultat est remarquablement simple dans la mesure où la seule différence avec l'onde se propageant dans le vide est l'expression du nombre d'onde. Cette expression $k(\omega)$ contient toute la complexité du lien dynamique existant entre le champ électrique et le mouvement des nuages électriques. En effet, $k(\omega)$ dépend de la permittivité $\epsilon(\omega)$ qui elle-même dépend de la susceptibilité $\chi(\omega)$ qui, à son tour, rend compte de la déformabilité des atomes. La dynamique des nuages électroniques est incluse dans cette description dans la mesure où la dépendance de la susceptibilité en la pulsation traduit le fait que le comportement des nuages électroniques dépend de la pulsation.

La différence essentielle par rapport à l'onde EM se propageant dans le vide est que le nombre d'onde n'est plus simplement proportionnel à la pulsation. Je rappelle que pour une onde dans le vide le nombre d'onde est défini par la relation $k = \omega/c$ que nous avons introduite de façon à ce que la fonction $\cos(kz - \omega t)$ représente une onde se propageant à la vitesse de la lumière dans le vide c . Ceci se voit facilement si on y

met le nombre d'onde en évidence, soit $\cos[k(z - (\omega/k)t)]$ où $\omega/k = c$ apparaît bien comme étant la vitesse de l'onde.

1.5 Vitesse de propagation et relation de dispersion

Puisque le nombre d'onde dans le milieu diélectrique $k(\omega) = \omega\sqrt{\mu_0\epsilon(\omega)}$ est a priori une fonction plus compliquée de ω que le simple rapport ω/c on doit s'attendre à avoir une modification de la vitesse de propagation des ondes. La vitesse apparaît en mettant $k(\omega)$ en évidence dans l'expression de l'onde, comme ci-dessous (je rappelle que pour comprendre cela il suffit de remarquer que la valeur du cosinus est maintenue constante tant que son argument ne varie pas, c'est-à-dire, tant que $z - (\omega/k)t = \text{cste}$, ce qui en différenciant donne $dz/dt = \omega/k$).

$$E(z, t) = E_\omega \cos\left[k(\omega)\left(z - \frac{\omega}{k(\omega)}t\right)\right]$$

En introduisant la vitesse $v = \omega/k(\omega)$, on peut écrire le champ électrique comme suit :

$$v(\omega) \equiv \frac{\omega}{k(\omega)} \quad \Rightarrow \quad E(z, t) = E_\omega \cos[k(\omega)(z - vt)]$$

En conclusion, la fonction $v(\omega) = \omega/k(\omega)$ donne la vitesse des ondes harmoniques dans le milieu diélectrique en fonction de la pulsation ω de ces ondes. En toute généralité, cette vitesse est différente de la vitesse de la lumière dans le vide c .

$$\text{vitesse des ondes harmoniques : } v(\omega) = \frac{\omega}{k(\omega)} \neq c$$

En reprenant l'expression du nombre d'onde donnée à la section précédente, soit $k(\omega) = \omega\sqrt{\mu_0\epsilon(\omega)}$, on voit que la vitesse est inversement proportionnelle à la racine de la permittivité caractéristique du milieu diélectrique.

$$v(\omega) = \frac{\cancel{\omega} 1}{\cancel{\omega} \sqrt{\mu_0 \epsilon(\omega)}}$$

Comme chaque milieu diélectrique est caractérisé par une déformabilité différente des molécules et des atomes qui le constituent, la susceptibilité et la permittivité varient d'un milieu à l'autre et il en est donc de même de la vitesse de propagation $v = \omega/k(\omega) = 1/\sqrt{\mu_0\epsilon(\omega)}$.

Pour caractériser les propriétés électromagnétiques d'un milieu diélectrique on introduit la notion de permittivité relative définie par la relation suivante : $\varepsilon_r \equiv \varepsilon / \varepsilon_0 = 1 + \chi$. Il faut avoir à l'esprit que cette grandeur est une fonction de la pulsation de la lumière. Chaque milieu diélectrique est caractérisé par sa fonction de permittivité relative $\varepsilon_r(\omega)$. En introduisant la permittivité relative dans l'expression de la vitesse des ondes harmoniques, on y fait apparaître la vitesse de la lumière dans le vide $c = 1/\sqrt{\mu_0\varepsilon_0}$:

$$v(\omega) = \frac{1}{\sqrt{\mu_0\varepsilon(\omega)}} = \frac{1}{\sqrt{\mu_0\varepsilon_0\varepsilon_r(\omega)}} = \frac{c}{\sqrt{\varepsilon_r(\omega)}}$$

$\varepsilon \equiv \varepsilon_0\varepsilon_r$

On définit alors l'indice de réfraction « n » comme étant la racine de la permittivité relative, ce qui conduit à l'expression suivante de la vitesse de propagation de l'onde harmonique de pulsation ω :

$$n(\omega) \equiv \sqrt{\varepsilon_r(\omega)} \quad \Rightarrow \quad v(\omega) = \frac{c}{n(\omega)}$$

Cette relation montre que la vitesse de propagation d'une onde harmonique est inversement proportionnelle à l'indice de réfraction du milieu dans lequel elle se propage. Puisque l'indice de réfraction est fonction de la pulsation ω , la vitesse de propagation d'une onde harmonique dépend bien sûr de sa pulsation. Comme la permittivité relative est, d'après sa définition $\varepsilon_r = 1 + \chi$, une grandeur toujours supérieure à l'unité, il en est de même pour l'indice de réfraction $n = \sqrt{1 + \chi}$ et on peut en conclure que la vitesse de propagation dans un milieu diélectrique est toujours plus faible que dans le vide.

Sachant que $k = \omega/v$ et $v = c/n$, on constate que le nombre d'onde est proportionnel à l'indice de réfraction, soit $k = n\omega/c$. On verra plus loin que cette relation générale entre le nombre d'onde et la fréquence angulaire joue un rôle essentiel dans la description de la propagation des ondes. Pour une raison qui apparaîtra plus loin, on donne à cette relation le nom spécifique de « *relation de dispersion* ».

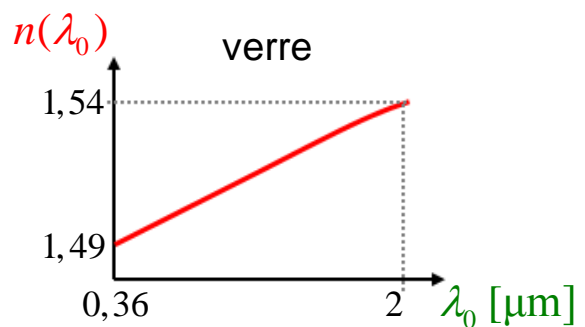
$$k(\omega) = \frac{\omega}{c} n(\omega)$$

Remarque : Notez que la notion d'indice de réfraction est surtout utilisée dans le domaine de l'optique. C'est une notion qui a été introduite bien avant la découverte des ondes électromagnétique pour décrire les propriétés de la propagation lumineuse dans les diélectriques transparents. Dans les autres domaines de longueur d'onde on utilise plutôt la notion de permittivité relative pour décrire la propagation d'onde EM (par exemple aux longueurs d'ondes des applications de télécommunications).

La longueur d'onde de l'onde harmonique dans le milieu diélectrique est donnée par $2\pi/k$ et puisque $k = \omega n(\omega)/c$ on constate que la longueur d'onde dans un milieu diélectrique est plus petite que dans le vide ($\lambda_0 = 2\pi c/\omega$ est la longueur d'onde de l'onde harmonique de fréquence angulaire ω se propageant dans le vide).

$$\lambda = \frac{2\pi}{k(\omega)} = \frac{2\pi c}{\omega} \frac{1}{n} = \frac{\lambda_0}{n}$$

On utilise souvent la longueur d'onde dans le vide plutôt que la fréquence angulaire pour caractériser une onde harmonique (la longueur d'onde étant connue, la fréquence angulaire s'obtient avec la relation $\omega = 2\pi c/\lambda_0$). Le graphe ci-contre montre une variation typique de l'indice de réfraction du verre en fonction de la longueur d'onde dans le vide. Ce graphe montre, en particulier, que l'indice de réfraction varie avec la longueur d'onde (ou la fréquence) et que la longueur d'onde dans le verre est à peu près 1,5 fois plus petite que dans le vide.



1.6 Propagation d'un battement et vitesse de groupe

Pour comprendre l'importance de la relation de dispersion $k(\omega) = \omega n(\omega)/c$ dans la description des phénomènes ondulatoires, nous allons ici étudier la propagation simultanée de deux ondes de fréquences différentes. Nous avons déjà considéré ce problème dans le cas de la propagation dans le vide (voir chapitre 5, section 10). Nous avons alors vu que deux ondes de fréquences légèrement différentes provoquent des interférences constructives et destructives alternées associées à la présence de ventres et de nœuds dans l'enveloppe de l'onde. C'est ce que nous avons appelé le phénomène de battement. Nous allons voir ici que les ondes de battement se comportent de façon légèrement différente quand elles se propagent dans les milieux diélectriques.


Pour la simplicité, je considérerai seulement ici la superposition de deux ondes de même amplitude A et dont les vecteurs champ électrique ont la même direction. Si les deux ondes ont un déphasage nul en $z=0$ et ont des fréquences ω et $\omega + \Delta\omega$, on peut écrire le champ total en $z=0$ comme l'indique la relation suivante. Pour être précis, il s'agit d'une composante de champ puisque la notation est scalaire (admettons par exemple qu'il s'agisse de la composante en x d'un champ parallèle à l'axe x et se propageant dans la direction z).

$$E(0, t) = A \cos(\omega t) + A \cos[(\omega + \Delta\omega)t]$$

On utilise la formule de Simpson pour transformer la somme des cosinus en un produit.

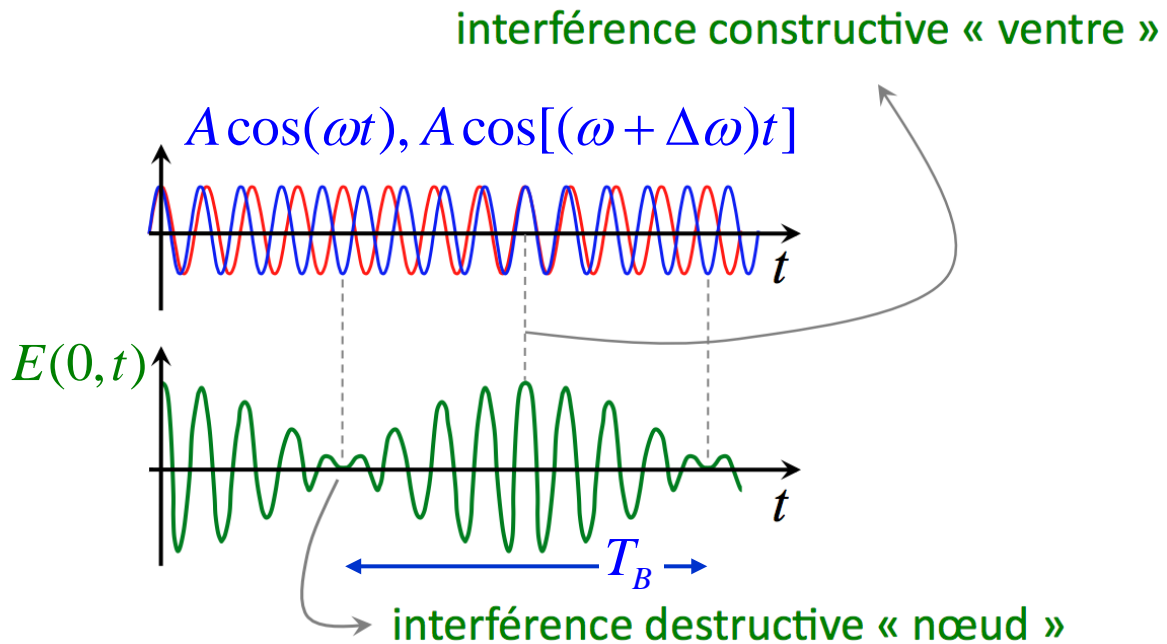
$$\cos \alpha + \cos \beta = 2 \cos\left(\frac{\alpha + \beta}{2}\right) \cos\left(\frac{\beta - \alpha}{2}\right)$$

Le résultat est un signal de fréquence angulaire $\omega + \Delta\omega/2$ dont l'amplitude est modulée de façon harmonique à une fréquence angulaire de petite valeur $\Delta\omega/2$. C'est le phénomène de battement (voir chapitre 5, section 10).



$$E(0, t) = 2A \cos\left(\frac{2\omega + \Delta\omega}{2}t\right) \cos\left(\frac{\Delta\omega}{2}t\right)$$

Ce battement est représenté schématiquement ci-dessous. Sa période de battement est $T_B = 2\pi/\Delta\omega$.



Étudions maintenant la propagation de ce signal de battement dans un milieu diélectrique dont la relation de dispersion est donnée au travers de l'indice de réfraction, soit $k(\omega) = n(\omega)\omega / c$. Pour décrire la propagation des deux ondes il nous suffit de tenir compte du fait que le nombre d'onde dépend de la pulsation, ce qui donne :


$$E(z, t) = A \cos[k(\omega)z - \omega t] + A \cos[k(\omega + \Delta\omega)z - (\omega + \Delta\omega)t]$$

En utilisant la formule de Simpson, on obtient une expression dans laquelle on voit apparaître le phénomène de battement aux niveaux temporel et spatial. Ce phénomène se traduit par la génération d'une onde de pulsation élevée $\omega + \Delta\omega/2$ et de nombre d'onde $[k(\omega) + k(\omega + \Delta\omega)]/2$ dont l'amplitude est modulée selon une variation harmonique lente de pulsation $\Delta\omega/2$ et de nombre d'onde $[k(\omega + \Delta\omega) - k(\omega)]/2$. Cette fonction de modulation harmonique lente s'appelle l'« *enveloppe* » de l'onde de battement. Le nombre d'onde $[k(\omega) + k(\omega + \Delta\omega)]/2$ peut être vu comme la moyenne des deux nombres d'onde des ondes individuelles, on l'appelle pour cette raison \bar{k} . De même, on appelle $\omega + \Delta\omega/2$ la fréquence moyenne des deux ondes et on la note $\bar{\omega}$.

$$E(z, t) = 2A \cos\left[\underbrace{\frac{k(\omega) + k(\omega + \Delta\omega)}{2}}_{\bar{k}} z - \underbrace{\left(\frac{2\omega + \Delta\omega}{2}\right)}_{\bar{\omega}} t\right] \times \cos\left(\frac{k(\omega + \Delta\omega) - k(\omega)}{2} z - \frac{\Delta\omega}{2} t\right)$$

$k(\omega + \Delta\omega) \approx k(\omega) + k'(\omega)\Delta\omega$

Comme le suggère la flèche verte sur la deuxième ligne de cette dernière expression, le nombre d'onde de l'enveloppe $[k(\omega + \Delta\omega) - k(\omega)]/2$ peut être exprimé en première approximation à l'aide de la dérivée de la fonction $k(\omega)$ par rapport à ω . On trouve $[k(\omega + \Delta\omega) - k(\omega)]/2 = k'(\omega)\Delta\omega/2$. Ceci est une approximation valable si la fréquence différence $\Delta\omega$ est beaucoup plus faible que la fréquence ω elle-même. Avec ces nouvelles notations, on peut écrire le champ total de l'onde de battement comme suit :



$$E(z, t) = 2A \cos[\bar{k}(\omega)z - \bar{\omega}t] \cos\left(\frac{k'(\omega)\Delta\omega}{2}z - \frac{\Delta\omega}{2}t\right)$$

Finalement, nous faisons apparaître les vitesses de l'onde et de son enveloppe en mettant les nombres d'onde en évidence dans les arguments des fonctions cosinus.

$$E(z, t) = 2A \cos\left[\bar{k}(\omega)\left(z - \frac{\bar{\omega}}{\bar{k}(\omega)}t\right)\right] \cos\left[\frac{k'(\omega)\Delta\omega}{2}\left(z - \frac{1}{k'(\omega)}t\right)\right]$$

On voit ainsi aisément que l'onde de fréquence élevée se propage à la vitesse $v = \bar{\omega}/\bar{k}$, c'est ce que l'on appelle la « *vitesse de phase* » de l'onde, elle est notée v_ϕ . Comme la différence de pulsation $\Delta\omega$ est supposée beaucoup plus petite que la pulsation ω de l'onde elle-même, il est évident que les valeurs moyennes $\bar{\omega}$ et \bar{k} peuvent être assimilées en première approximation à ω et k dans l'expression de v_ϕ .

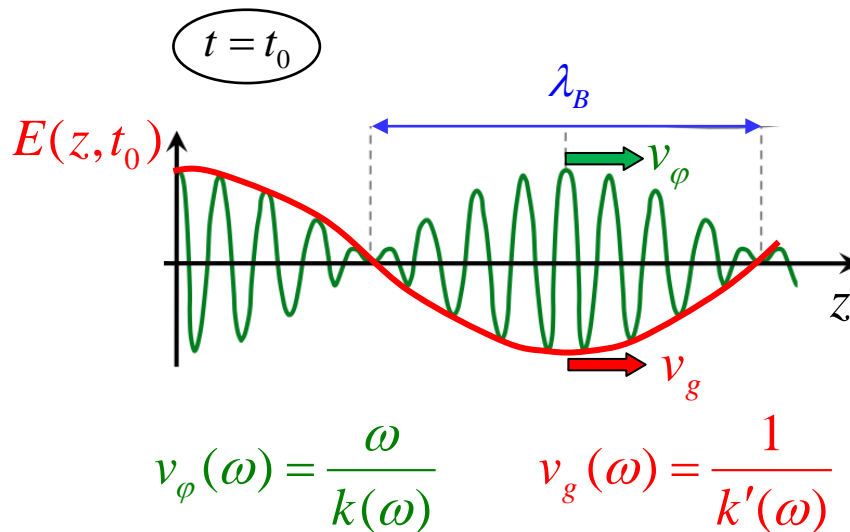
$$\text{vitesse de "phase": } v_\phi(\omega) = \frac{\bar{\omega}}{\bar{k}(\omega)} \approx \frac{\omega}{k(\omega)}$$

On voit également que la vitesse de l'enveloppe est $1/k'(\omega)$. Cette vitesse est appelée la « *vitesse de groupe* » et est notée v_g . Elle est en toute généralité différente de la vitesse de phase.

$$\text{vitesse de "groupe": } v_g(\omega) = \frac{1}{k'(\omega)}$$

La situation est représentée sur le schéma ci-dessous où est dessinée en vert l'onde de battement, c'est-à-dire, le champ total. L'enveloppe de modulation du champ total est représentée par la courbe rouge, c'est la fonction cosinus de fréquence faible. Cette dernière avance à la vitesse de groupe v_g alors que l'onde verte de fréquence élevée avance à la vitesse de phase v_ϕ . Comme ces deux vitesses sont en général différentes, on voit qu'il y a un « *glissement* » entre l'onde de battement et son

enveloppe. Sur le dessin ci-dessous les maxima de l'onde et de son enveloppe coïncident mais ceci est une situation particulière observée seulement à des instants précis car les maxima de l'onde de battement se déplacent au sein de l'enveloppe. On peut facilement voir que ce glissement est dû au fait que les ondes harmoniques de fréquences ω et $\omega + \Delta\omega$ ont des vitesses de propagation différentes. En effet, si ces ondes avaient la même vitesse de propagation cela signifierait que les rapports $\omega/k(\omega)$ et $(\omega + \Delta\omega)/k(\omega + \Delta\omega)$ sont égaux et donc que la relation de dispersion $k(\omega)$ est une simple relation de proportionnalité $k(\omega) = \alpha\omega$ auquel cas on a bien l'égalité $\omega/k(\omega) = 1/k'(\omega) = 1/\alpha$ correspondant à l'absence de glissement.



J'attire votre attention sur le fait que la fonction $k(\omega)$, c'est-à-dire, la relation de dispersion apparaît dans les expressions des vitesses de phase et de groupe. Ceci donne une indication de l'importance de la relation de dispersion dans la description des phénomènes de propagation d'ondes.

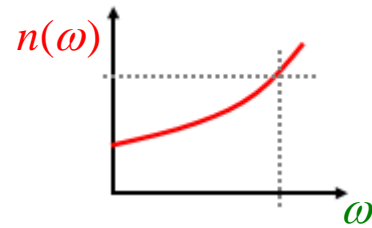
$$k(\omega) = \frac{\omega}{c} n(\omega)$$

Pour illustrer cela, reprenons l'exemple de la propagation de l'onde de battement dans le vide. La relation de dispersion du vide est $k(\omega) = \omega/c$. La vitesse de phase $v_{\phi} = \omega/k(\omega)$ est donc la vitesse de la lumière dans le vide, c . La vitesse de groupe est donnée par $v_g = 1/k'(\omega)$, ce qui de manière évidente donne aussi la vitesse c .

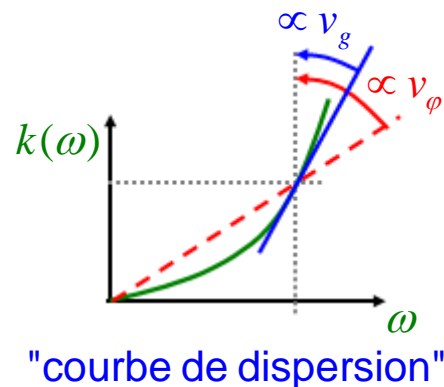
Vide: $k(\omega) = \frac{\omega}{c} \quad \Rightarrow \quad k'(\omega) = \frac{1}{c} \quad \Rightarrow \quad v_g = v_{\phi} = c$

Ce résultat était, bien sûr, attendu puisque l'on sait que toutes les ondes EM, quelque soit leur fréquence, se propagent dans le vide à la vitesse c . On dit pour cela que le vide est un milieu « *non dispersif* ». Un battement s'y propage sans distorsion (pas de glissement de l'onde dans son enveloppe). Ceci est dû au fait que l'indice de réfraction du vide vaut l'unité quelque soit la fréquence (ce qui revient à dire que $k = \omega/c$).

Par contre, tout milieu diélectrique présente un indice de réfraction dépendant de la fréquence, comme le suggère le graphe ci-contre.



La relation de dispersion $k(\omega) = n(\omega)\omega/c$ n'est donc plus simplement linéaire, comme c'était le cas pour le vide. La pente de la courbe $k'(\omega)$ est donc différente du rapport k/ω , ce qui signifie que les vitesses de phase et de groupe sont différentes. Le graphe ci-dessous illustre cette situation, il s'agit du graphe de la « *courbe de dispersion* » $k(\omega) = n(\omega)\omega/c$. La vitesse de phase est représentée par le rapport ω/k qui peut être vu sur le graphe comme la pente de la droite pointillée rouge par rapport à la verticale. Plus $v_\phi = \omega/k = c/n(\omega)$ est grande, plus la droite pointillée se rapproche de l'axe horizontal des fréquences. De même, la vitesse de groupe $v_g = 1/k'(\omega)$ est la pente de la courbe de dispersion par rapport à la verticale. Plus la pente de la courbe de dispersion se rapproche de la verticale, plus petite est la vitesse de groupe.



Quand des ondes sont caractérisées par une relation de dispersion non linéaire (c'est-à-dire quand la vitesse de groupe est différente de la vitesse de phase ce qui est équivalent à dire que les ondes de fréquences différentes se propagent à des vitesses différentes), on dit que les ondes sont « *dispersives* » ce qui signifie qu'elles sont sujettes à la « *dispersion* ». Le terme « *dispersion* » traduit le fait que les ondes harmoniques se propagent à des vitesses différentes et ont donc tendance à se séparer, ou se « *disperser* », par propagation (dans le cas de l'onde de battement cela se traduit par le glissement entre l'onde de fréquence élevée et son enveloppe).

Comme nous l'avons vu ici, le caractère dispersif des ondes provient des caractéristiques du milieu qui véhicule ces ondes. C'est donc le milieu qui impose la relation de dispersion par ses propriétés microscopiques. Un milieu qui est caractérisé par une relation de dispersion non linéaire est appelé « *milieu dispersif* ». La théorie est développée ici pour les ondes EM dans les milieux diélectriques mais on rencontre des milieux dispersifs pour tous les types d'ondes (ondes acoustiques, ondes hydrodynamiques de surface, etc...).

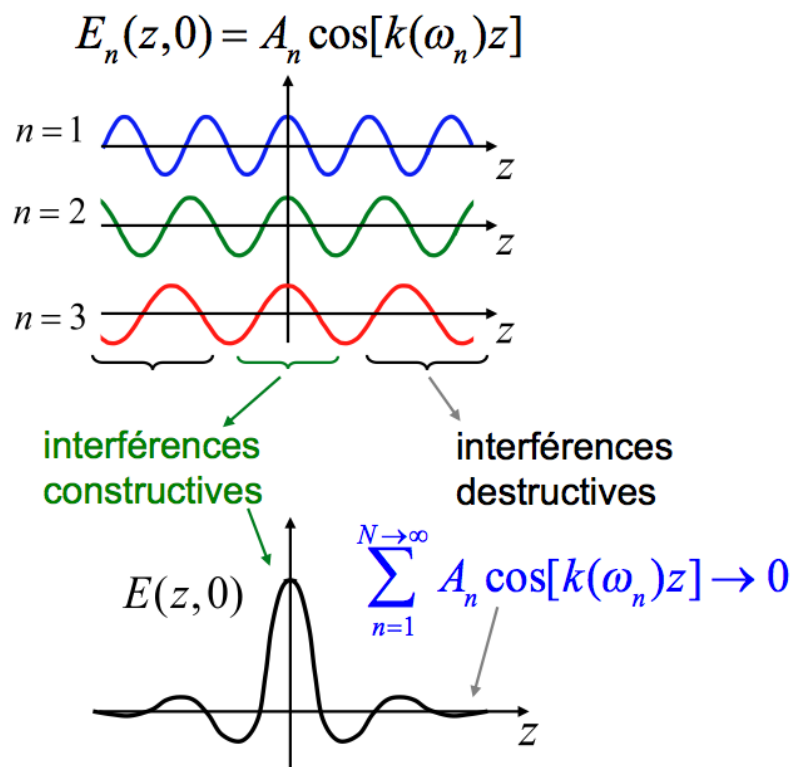
1.7 Paquet d'ondes et dispersion

Un paquet d'ondes est, comme son nom le suggère, une somme d'ondes harmoniques telle que celle décrite mathématiquement ci-dessous. Les ondes ont chacune une fréquence donnée ω_n correspondant à un nombre d'onde donné par la relation de dispersion $k(\omega_n)$.

$$E(z, t) = \sum_{n=1}^N E_n(z, t) = \sum_{n=1}^N A_n \cos[k(\omega_n)z - \omega_n t]$$

Si toutes les ondes harmoniques de la somme ont des déphasages nuls en $z=0$, comme c'est le cas de l'expression ci-dessus, leur somme donnera une amplitude importante en $z=0$ car tous les cosinus s'ajoutent avec le même signe (le champ total en $z=0$ est donné par la somme des amplitudes A_n). Le schéma ci-dessous montre trois ondes harmoniques de la somme au temps $t=0$ (photographie en $t=0$ des trois ondes). La somme en $z=0$ donne une valeur importante du champ E . Par contre, lorsqu'on s'éloigne de l'origine, la situation est toute différente : puisque les ondes harmoniques ont des nombres d'onde différents, elles prennent progressivement des valeurs de plus en plus diverses au fur et à mesure qu'on s'éloigne de $z=0$. On voit clairement sur le graphe qu'en des valeurs de z suffisamment grandes, on finit par sommer de façon aléatoire des valeurs positives et négatives. On se rend facilement compte que si le nombre d'harmoniques de la somme est très grand (tend vers l'infini), la somme en des valeurs de z éloignées de l'origine se fera sur des valeurs positives autant que sur des valeurs négatives des fonctions cosinus. Il en résultera donc une somme nulle.

On décrit souvent les sommes d'ondes harmoniques en termes d'interférences. Lorsque la somme se fait avec les mêmes valeurs de phase (ici la phase nulle en $z=0$), on dit que l'on a des interférences constructives. Dans l'autre cas, quand la somme se fait sur des champs de signes opposés en raison de valeurs de phase différentes (aléatoires), on dit qu'il y a interférences destructives.



On voit de cette façon qu'un paquet d'ondes harmoniques constitue une distribution de champ localisée dans l'espace pour un temps t donné (ici on a choisi le temps $t = 0$). On appelle cette distribution localisée une « *impulsion* » de champ électrique. De telles impulsions sont souvent utilisées en pratique et c'est là que réside tout l'intérêt de la notion de paquet d'ondes. On peut voir les choses dans l'autre sens en se disant que toute impulsion de champ électrique constitue une superposition d'ondes harmoniques de même phase à un instant donné mais de fréquences différentes.

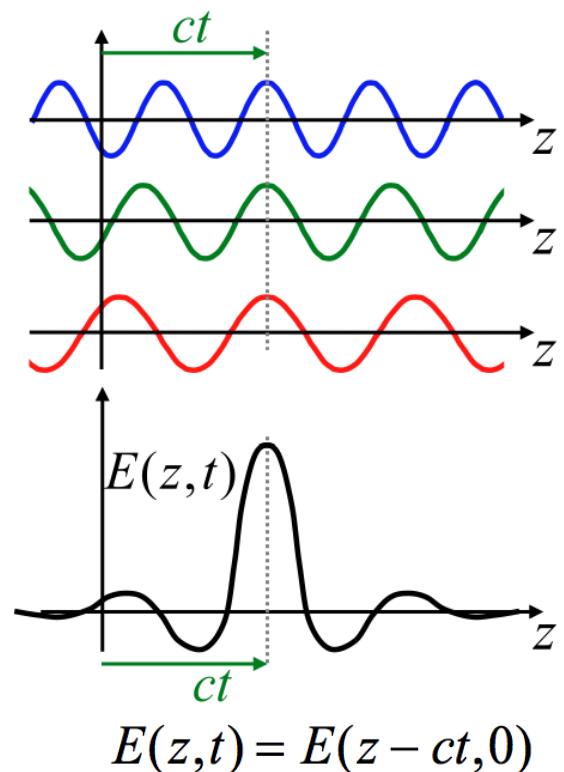
Cette décomposition des impulsions de champ électrique en somme d'ondes harmoniques est très utile en pratique car elle permet de décrire facilement leur propagation. Considérons, par exemple, la propagation du paquet d'onde EM dans le vide. Pour avoir la valeur du champ en tout point z et en tout temps t , il suffit de faire la somme des fonctions cosinus en y remplaçant $k(\omega_n)$ par sa valeur donnée par la relation de dispersion. Dans le cas du vide la relation de dispersion est $k = \omega/c$.

$$\text{propagation dans le vide : } k(\omega) = \frac{\omega}{c}$$

En mettant $k(\omega_n) = \omega_n/c$ en évidence dans l'argument des cosinus, on fait apparaître la vitesse de la lumière en facteur de t , et on voit ainsi que toutes les ondes harmoniques se propagent à la même vitesse.

$$E(z, t) = \sum_{n=1}^N A_n \cos\left[\frac{\omega_n}{c}(z - ct)\right]$$

En conclusion, l'impulsion de champ électrique (le paquet d'onde) avance à la vitesse c comme chacune de ses composantes harmoniques. Ceci peut se traduire mathématiquement par l'égalité $E(z, t) = E(z - ct, 0)$ qui exprime que la distribution spatiale du champ électrique en $t = 0$, soit $E(z, 0)$, est simplement translatée en $z = ct$ au temps t . Ce raisonnement est illustré ci-contre où l'on voit que la condition d'interférence constructive est remplie en $z = ct$, ce qui signifie que l'impulsion a son sommet en permanence en ce point bougeant à la vitesse c . En conclusion, l'impulsion se propage à la vitesse c par simple translation.



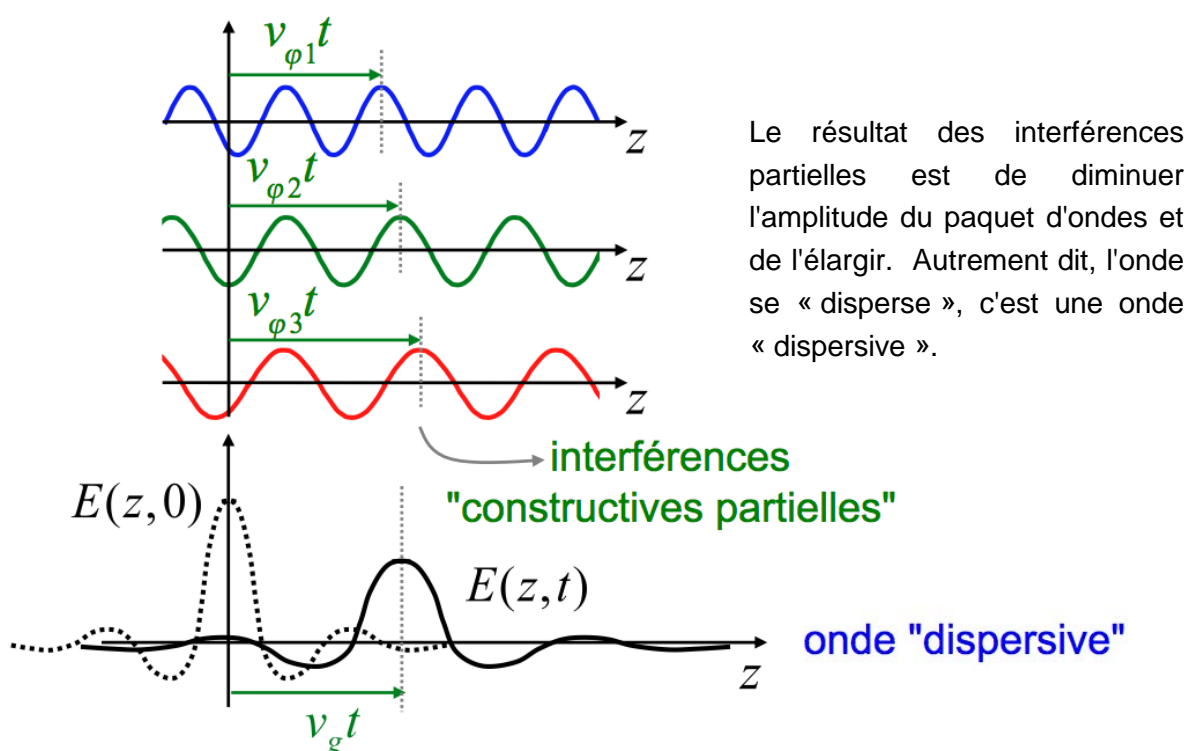
Voyons maintenant ce qui se produit quand une impulsion de champ électrique se propage dans un milieu diélectrique. Pour décrire cette propagation, on exprime la somme des ondes harmoniques en tenant compte de leur relation de dispersion $k(\omega) = \omega n(\omega) / c$.

$$E(z, t) = \sum_{n=1}^N A_n \cos[k(\omega_n)z - \omega_n t] \quad k(\omega) = \frac{\omega}{c} n(\omega)$$

En mettant les nombres d'onde en évidence dans les arguments des cosinus, on fait apparaître les vitesses de phase.

$$E(z, t) = \sum_{n=1}^N A_n \cos\left[k(\omega_n)\left(z - \frac{\omega_n}{k(\omega_n)} t\right)\right] \quad \nearrow v_\phi(\omega_n)$$

Comme le milieu est dispersif, chaque vitesse de phase est différente. Cette situation est représentée sur le graphe ci-dessous où on voit que la condition d'interférences constructives correspondant à la phase nulle de toutes les harmoniques n'est tout simplement plus possible car les points de phase nulle des harmoniques « n » se déplacent à des vitesses différentes $v_\phi(\omega_n)$. Même si l'on avait interférence constructive parfaite au temps $t=0$ pour former l'impulsion $E(z, 0)$, aux temps t successifs les points de phase nulles [soit les points dont la coordonnée z satisfait la relation $k_n(z - v_{\phi n} t) = 0$] ont glissé les uns par rapport aux autres et la somme des ondes harmoniques ne peut plus se faire avec la même phase pour toutes. La somme des harmoniques dans la zone des phases nulles, soit $z \approx v_{\phi n} t$, n'est donc pas aussi grande qu'elle ne l'était en $t=0$. On dit qu'on a interférences constructives partielles (ni tout à fait constructives, ni tout à fait destructives).

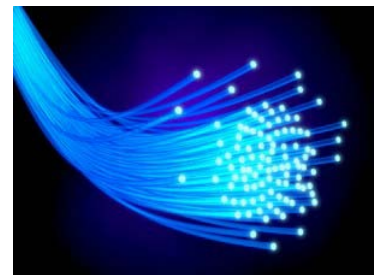


On peut montrer que le maximum du paquet d'ondes se déplace à la vitesse de groupe $v_g = 1/k'(\omega)$. Ceci n'est pas facile à démontrer rigoureusement mais pour le comprendre vous pouvez vous référer au résultat obtenu pour l'onde de battement qui constitue un cas particulier de paquet d'ondes à deux composantes. Dans ce cas nous avons vu que l'enveloppe se déplace à la vitesse de groupe $v_g = 1/k'(\omega)$ (voir section 1.6). Dans le cas présent d'une somme de plus de deux ondes harmoniques, l'équivalent de l'enveloppe est le profil de l'impulsion. Il n'est donc pas difficile d'admettre que ce profil se déplace à la vitesse $v_g = 1/k'(\omega)$ comme l'enveloppe de l'onde de battement.

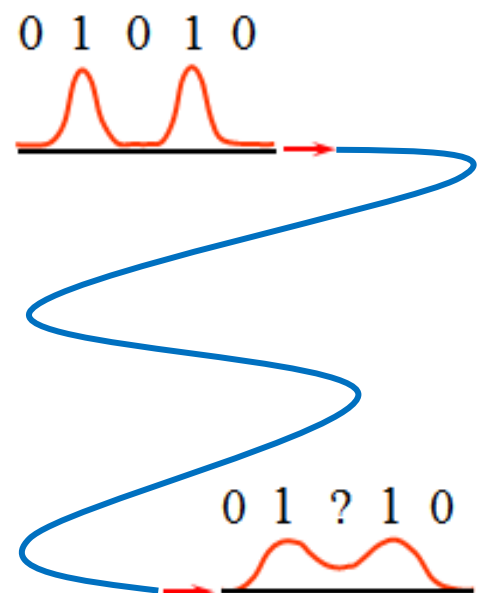
On retiendra surtout de ce développement que la dispersion (c'est-à-dire la relation de dispersion non linéaire) est responsable d'un étalement du paquet d'ondes et d'une diminution de son amplitude. C'est le phénomène de dispersion.

Illustration : les fibres optiques

La dispersion des ondes est un phénomène qui se rencontre dans beaucoup de systèmes pratiques et il peut avoir une grande importance dans certaines applications des ondes. Ceci est illustré ici avec la propagation d'impulsions lumineuses dans les fibres optiques (l'image ci-contre montre une fibre optique enroulée sur un cylindre). Comme vous le savez probablement, l'information est véhiculée dans les fibres optiques sous forme d'impulsions lumineuses très courtes, c'est-à-dire, sous forme de paquets d'ondes.



Le principe de la transmission de données par fibre optique est schématisé ci-contre. L'information est mise sous forme binaire, c'est-à-dire qu'elle est discrétisée sous forme d'une suite de 1 et de 0. Les 1 et les 0 sont disposés dans des « fenêtres » temporelles bien précises dont la durée détermine la cadence de la transmission. Chaque fenêtre temporelle contient un 1 ou un 0. Pour transmettre un 1 dans une fenêtre temporelle donnée, on y envoie une impulsion lumineuse. Les 0, eux, sont codés par l'absence d'impulsion dans les fenêtres temporelles correspondantes, comme l'indique le schéma. Les impulsions lumineuses sont des impulsions EM qui, en cours de propagation dans la ligne, subissent la

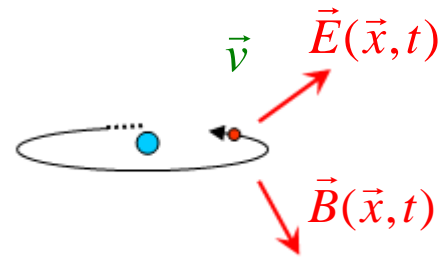


dispersion du verre de la fibre optique [on a vu à la fin de la section 1.6 que l'indice de réfraction « n » du verre dépend de la longueur d'onde, c'est-à-dire, de la pulsation, $n = n(\omega)$, il y a donc bien dispersion]. Les impulsions subissent dès lors un élargissement dispersif. Bien entendu, cet effet est d'autant plus important que la longueur de la ligne de transmission est grande.

L'élargissement dispersif des impulsions lumineuses est la cause majeure de limitation de la cadence des transmissions optiques. En effet, pour augmenter la cadence des transmissions, il faut diminuer la durée des fenêtres contenant les 0 et les 1 du signal binaire. Les impulsions EM sont en conséquence plus proches l'une de l'autre et lors de leur élargissement dispersif, elles finissent par se recouvrir l'une l'autre. Le schéma ci-dessus illustre cette situation d'élargissement excessif. Il suggère que le 0 se trouvant entre deux 1 pourrait être mal lu puisqu'en raison de l'élargissement, il y a maintenant une intensité lumineuse non nulle qui pourrait être interprétée comme codant un 1 au lieu d'un zéro. Cet exemple illustre l'importance des phénomènes dispersifs en pratique.

2. Introduction à la mécanique quantique

La mécanique quantique est née de l'idée que la matière est constituée d'ondes et non de particules massiques occupant une position et un volume précis dans l'espace. A titre d'exemple, les électrons qui tournent autour des noyaux atomiques forment une onde diffuse que l'on appelle usuellement le nuage électronique de l'atome. Les « *ondes de matière* » ont été introduites en physique au début du 20^{ème} siècle pour répondre à un certain nombre d'interrogations et d'observations expérimentales relatives à la structure de la matière. Un des problèmes qui se posait à l'époque est illustré ci-contre. Il s'agit du schéma élémentaire d'un atome (l'atome d'hydrogène pour la simplicité) tel qu'il était vu vers 1913. Niels Bohr, un physicien danois, avait donné cette image « *de système planétaire* » de l'atome pour rendre compte d'observations expérimentales.



Ce modèle posait problème (y compris à Niels Bohr) car en gravitant autour du noyau sur une orbite circulaire l'électron provoque des variations périodiques de la charge électrique et du courant électrique dans les régions de l'espace par lesquelles il passe (en un point donné de sa trajectoire la charge électrique est nulle tout le temps sauf au moment où l'électron s'y trouve et il en est de même du courant). Ces variations de charge et de courant provoquent inévitablement des perturbations locales des champs électriques et magnétiques. Ces perturbations locales se propagent à la vitesse de la lumière dans tout l'espace sous forme d'ondes électromagnétiques. Or, les ondes électromagnétiques sont constituées de distributions de champs électrique et magnétique possédant des densités d'énergie données par $(\epsilon_0/2) \|\vec{E}\|^2$ et $1/(2\mu_0) \|\vec{B}\|^2$ (voir le chapitre d'électrostatique, section 6.11 et le chapitre d'électromagnétisme, section 8.7). Cette énergie diffusée dans l'espace ne peut provenir que du mouvement de l'électron, c'est-à-dire de son énergie cinétique. L'électron perd donc de l'énergie cinétique dans son mouvement orbital et se rapproche progressivement du noyau. En effet, lorsqu'il perd de l'énergie cinétique, la force centrifuge qui le maintient en orbite ne compense plus l'attraction coulombienne du proton et il finit donc par entrer en collision avec ce dernier. Ce raisonnement simple montre que le modèle planétaire de Bohr n'est physiquement pas acceptable. Il fallait trouver un modèle dans lequel les électrons ne bougent pas, ce qui était tout aussi difficile à imaginer car si l'électron ne bouge pas, il n'y a rien qui le retient loin du noyau et on ne peut expliquer pourquoi les atomes sont de taille beaucoup plus grande que les noyaux.

Alors que pendant de longues années toute la communauté scientifique était devant ce problème concernant la structure de la matière, en 1923 le jeune physicien français Louis de Broglie fit, dans le cadre de sa thèse de doctorat, une suggestion très

audacieuse pour le résoudre. Cette suggestion constitue un des fondements de la mécanique quantique.

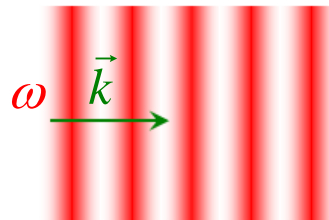
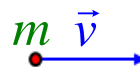
2.1 La conjecture de de Broglie

En simplifiant fortement les aspects scientifiques et historiques, on peut résumer la proposition de Louis de Broglie comme suit : plutôt que de considérer l'électron (ou toute autre particule élémentaire) comme une petite bille solide possédant une masse localisée m se déplaçant à une certaine vitesse \vec{v} , il le considéra comme une onde. En d'autres termes, il conjectura que les particules élémentaires sont en réalité des ondes possédant une certaine fréquence angulaire ω et un certain vecteur d'onde \vec{k} , à l'image des ondes électromagnétiques. On appelle les ondes de de Broglie des « ondes de matière » puisqu'elles constituent véritablement la matière.

Louis de Broglie



1892 - 1987



Pour des raisons qu'il n'est pas possible de présenter ici, car elles impliquent un grand nombre de travaux antérieurs à ceux de de Broglie, la fréquence de l'onde qui représente la particule fut associée par de Broglie à l'énergie de la particule. De même, de Broglie associa l'impulsion (ou la quantité de mouvement) de la particule au vecteur d'onde de cette onde. Les liens entre ces grandeurs mettent en jeu la constante de Planck $h = 6,6 \times 10^{-34}$ J.s, qui constitue une des constantes fondamentales de la nature (cette constante était connue en 1923).

constante de Planck : $h = 6,610^{-34}$ J.s

Pour la simplicité, nous n'allons considérer ici que le cas de la particule libre, c'est-à-dire, le cas d'une particule se déplaçant dans l'espace libre en l'absence de potentiel d'interaction. L'énergie de la particule se réduit dans cette situation à son énergie cinétique. La pulsation de l'onde de matière associée à la particule de masse m et de vitesse v est donnée par son énergie cinétique divisée par $h/2\pi$, ce qui revient à dire que l'énergie de la particule est donnée par la constante de Planck divisée par la période de l'onde, comme indiqué ci-dessous à droite.

$$\omega = \frac{2\pi}{h} \frac{1}{2} m v^2 = \frac{2\pi}{h} E \quad \Rightarrow \quad E = \frac{h}{2\pi} \omega = \frac{h}{T}$$

Le nombre vecteur d'onde \vec{k} quant à lui est donné par la quantité de mouvement $\vec{p} = m\vec{v}$ divisée par $h/2\pi$. Le vecteur d'onde est donc un vecteur qui pointe dans la direction du mouvement de la particule. En module, la relation de de Broglie revient à dire que la quantité de mouvement est égale à la constante de Planck divisée par la longueur d'onde λ .

$$k = \frac{2\pi}{h} mv = \frac{2\pi}{h} p \quad \Rightarrow \quad p = \frac{h}{2\pi} k = \frac{h}{\lambda}$$

Pour la simplicité des notations, on introduit la constante $\hbar = h/(2\pi)$ (prononcée hache-barre) de sorte que les relations de de Broglie s'écrivent $E = \hbar\omega$ et $p = \hbar k$.

$$\hbar = h / (2\pi) \quad \Rightarrow \quad \begin{cases} E = \hbar\omega \\ p = \hbar k \end{cases}$$

Pour décrire les ondes de matière, de Broglie a proposé d'utiliser des fonctions d'onde complexes. Ainsi une particule de masse m se déplaçant parallèlement à l'axe z à la vitesse v est représentée par la fonction d'onde suivante :

$$\text{fonction d'onde : } \psi = Ae^{i(kz - \omega t)}$$

Où $\omega = mv^2 / (2\hbar)$ et $k = mv / \hbar$. Il est important de réaliser que cette fonction d'onde complexe n'est pas un phaseur (c'est pourquoi il n'y a pas de barre sous la lettre ψ) dans la mesure où il ne lui correspond pas une grandeur réelle oscillante. Les fonctions d'onde de la mécanique quantique sont par essence même des ondes complexes. Nous comprendrons plus loin pourquoi de Broglie a proposé cette modélisation pour le moins abstraite de la matière.

Comme son écriture l'indique, cette onde avance vers les z positifs à la vitesse de phase $v_\phi = \omega/k$.

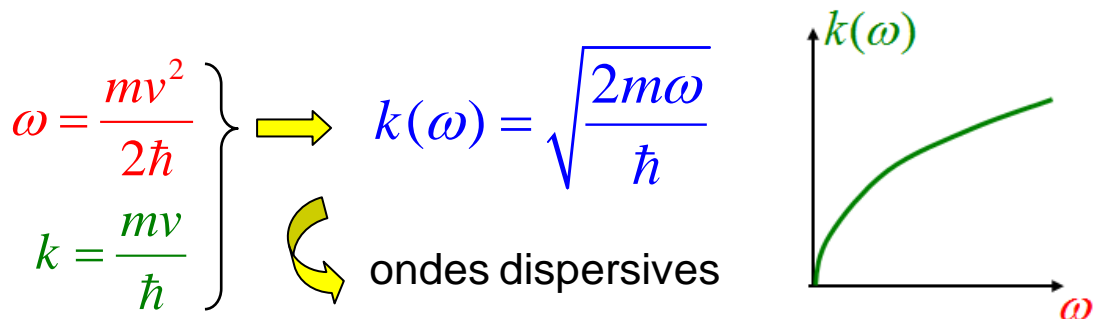
$$\psi = Ae^{i[k(\omega)z - \omega t]} \quad \Rightarrow \quad v_\phi = \frac{\omega}{k(\omega)}$$

Notez que cette proposition de représentation mathématique des particules élémentaires de de Broglie constitue ce que l'on peut appeler une conjecture. Il s'agit effectivement d'une construction purement intellectuelle qui n'a reçu de premières vérifications expérimentales que deux années plus tard (en 1925). Ce qui explique d'ailleurs pourquoi de Broglie ne reçut le prix Nobel qu'en 1929, six ans après sa conjecture « *audacieuse* ». En ce qui concerne le caractère abstrait de la proposition

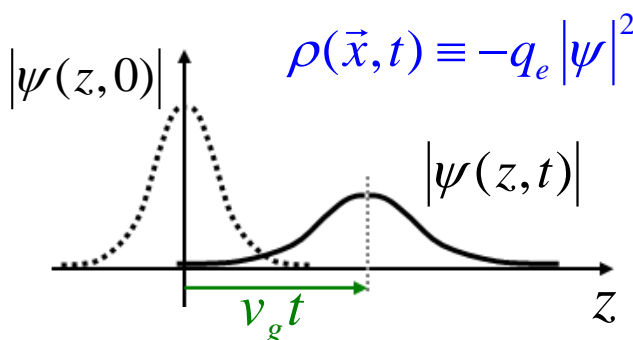
de de Broglie, il est important d'avoir à l'esprit que ce n'est pas parce que l'on a réalisé de nombreuses vérifications expérimentales de la validité de cette conjecture que l'on peut mieux la comprendre aujourd'hui. En réalité, personne n'est tenu de « *comprendre* » le sens physique profond de la proposition de de Broglie. Il s'agit d'un « *postulat* » constituant un modèle mathématique fidèle de la réalité expérimentale à l'échelle microscopique. Essayer de comprendre la conjecture de de Broglie est en quelque sorte comparable à essayer de comprendre l'origine de la loi de Coulomb : tout ce qu'on a pu faire c'est décrire cette loi mathématiquement mais personne n'a jamais pu expliquer son origine.

2.2 Relation de dispersion des ondes de matière

Le lien entre le nombre d'onde et la pulsation fournit la relation de dispersion qui caractérise les ondes. Pour les ondes électromagnétiques nous avons vu dans la section 1 du présent chapitre que la relation de dispersion conditionne la vitesse de propagation ainsi que la dispersion des paquets d'ondes (ou impulsions). Pour voir si les ondes de matière sont sujettes à la dispersion, il suffit de vérifier si la relation de dispersion est linéaire ou non. La relation entre le nombre d'onde et la pulsation s'obtient en éliminant la vitesse dans leurs expressions, soit $\omega = mv^2 / (2\hbar)$ et $k = mv / \hbar$. En écrivant cette dernière relation $v = \hbar k / m$ on trouve $\omega = \hbar k^2 / (2m)$, soit encore $k(\omega) = \sqrt{2m\omega / \hbar}$, ce qui indique que le lien entre la fréquence angulaire et le nombre d'onde est de type parabolique (voir graphe ci-dessous). Il apparaît donc évident que les ondes de matière sont dispersives.



Si, dans son état initial, une particule est localisée dans l'espace sous forme d'un paquet d'ondes étroit, comme celui représenté schématiquement ci-dessous, celui-ci va inévitablement subir un élargissement dispersif lors de sa propagation (même si celle-ci a lieu dans le vide, ce qui est implicite depuis le début). La particule perd progressivement sa localisation en se déplaçant ; elle se « *disperse* » en quelque sorte dans l'espace.



Notez que s'il s'agit d'une particule chargée telle qu'un électron, la densité de charge est donnée par le module carré de la fonction d'onde, soit $\rho(\vec{x}) = -q_e |\psi(\vec{x})|^2$ (la fonction d'onde est supposée ici normalisée dans le sens où l'intégrale de

volume de $|\psi(\vec{x})|^2$ sur tout l'espace vaut l'unité, de sorte que la charge totale du nuage électronique soit égale à celle de l'électron). L'interprétation de ce que j'appelle ici localisation de la particule vous sera donnée au cours de physique quantique et statistique de 2^{ème} année (je peux juste mentionner que cette problématique est liée à ce que l'on appelle la « *dualité onde-corpuscule* » et l'interprétation « *probabiliste* » de la mécanique quantique).

Remarquez que le schéma ci-dessus rappelle que le déplacement du paquet d'ondes se fait avec la vitesse de groupe v_g . La vitesse de groupe étant donnée par $1/k'(\omega)$, on trouve facilement, à l'aide de la relation de dispersion $\omega = \hbar k^2 / (2m)$, que la vitesse de groupe du paquet d'ondes électronique n'est rien d'autre que la vitesse v de la particule « *classique* », en effet, $v_g = 1/k'(\omega) = d\omega / dk = \hbar k / m = v$ (pour la dernière égalité je rappelle que $p = \hbar k$ et $p/m = v$). Ce résultat nous permet en quelque sorte que concilier les points de vue de la mécanique classique et de la mécanique quantique en voyant la particule classique comme un paquet d'onde localisé.

2.3 Équation de Schrödinger

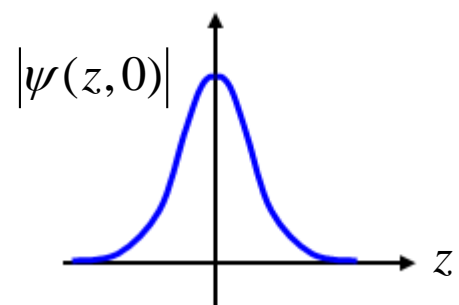
L'intérêt de la relation de dispersion ne réside pas seulement dans la capacité qu'elle offre de décrire les phénomènes de dispersion. Elle est également très utile pour dériver les équations d'ondes des ondes qu'elle régit. Nous allons l'utiliser ici pour construire l'équation de base de la mécanique quantique, l'équation de Schrödinger, du nom d'Erwin Schrödinger, le physicien autrichien qui l'a dérivée pour la première fois vers 1925 sur base des travaux de de Broglie. Le raisonnement de Schrödinger fut très simple. Pour le comprendre, considérons la relation de dispersion en la mettant au carré dans les deux membres, comme ci-dessous :

$$\text{relation de dispersion : } k(\omega) = \sqrt{\frac{2m\omega}{\hbar}} \quad \Rightarrow \quad k^2 = \frac{2m\omega}{\hbar}$$

Considérons maintenant un paquet d'ondes de matière. Celui-ci est, par définition, sujet à la relation de dispersion de de Broglie. Nous pouvons donc écrire ce paquet d'ondes comme suit $\psi(z, t) = \sum A_n \exp\{i[k(\omega_n)z - \omega_n t]\}$ où $k(\omega_n)$ est le nombre d'onde de la composante harmonique « n » de pulsation ω_n du paquet d'ondes. Le paquet d'ondes peut, par exemple, représenter un état plus ou moins localisé de la particule, comme suggéré sur le schéma ci-contre. L'idée de Schrödinger était de décrire l'évolution du paquet d'ondes en termes d'une

Paquet d'ondes :

$$\psi(z, t) = \sum_n A_n e^{i[k(\omega_n)z - \omega_n t]}$$



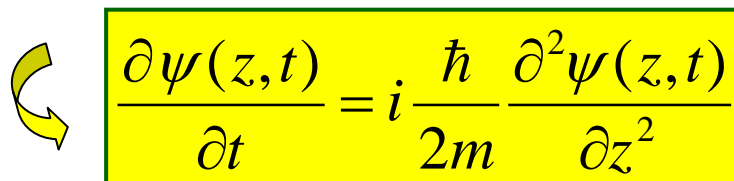
équation aux dérivées partielles. Il lui a suffi pour cela de calculer la dérivée seconde du paquet d'ondes par rapport à la coordonnée d'espace z pour voir que celle-ci est proportionnelle à sa dérivée temporelle.

$$\frac{\partial^2 \psi(z,t)}{\partial z^2} = -\sum_n k_n^2 A_n e^{i[k(\omega_n)z - \omega_n t]} = -\frac{2m}{\hbar} \sum_n \omega_n A_n e^{i[k(\omega_n)z - \omega_n t]}$$

On voit, en effet, que la dérivée par rapport au temps du paquet d'ondes fait apparaître les mêmes termes dans la somme que ceux de la dérivée seconde par rapport à z .

$$\frac{\partial \psi(z,t)}{\partial t} = -i \sum_n \omega_n A_n e^{i[k(\omega_n)z - \omega_n t]}$$

En comparant les deux dernières égalités obtenues, on voit qu'il suffit de multiplier la dérivée seconde spatiale par le facteur $i\hbar/(2m)$ pour obtenir la dérivée simple temporelle. Ce qui signifie que le lien entre ces dérivées est $\partial_t \psi = i\hbar/(2m) \partial_{zz}^2 \psi$, c'est l'équation d'onde de matière de Schrödinger.



$$\frac{\partial \psi(z,t)}{\partial t} = i \frac{\hbar}{2m} \frac{\partial^2 \psi(z,t)}{\partial z^2}$$

« équation de Schrödinger »
(particule libre)

La découverte de cette équation a permis de résoudre une quantité considérable de problèmes liés à la structure de la matière. Schrödinger reçut le prix Nobel en 1933 pour cette contribution (simple mais d'importance majeure) à la physique moderne.

Remarque : Notez que cette équation n'est valable que pour la particule libre sans potentiel de liaison. Pour la particule plongée dans un potentiel (par exemple le cas de l'électron plongé dans le potentiel coulombien du proton dans l'atome d'hydrogène), l'équation de Schrödinger prend la forme suivante, où $V(z)$ est proportionnel à l'énergie potentielle de la particule (l'équation est ici écrite pour une seule dimension spatiale, elle peut facilement être généralisée à trois dimensions en remplaçant la dérivée seconde spatiale par l'opérateur laplacien).

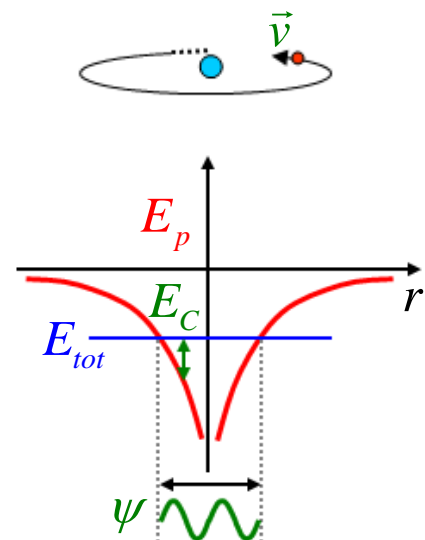
particule liée :

$$\frac{\partial \psi}{\partial t} = i \frac{\hbar}{2m} \frac{\partial^2 \psi}{\partial z^2} + iV(z)\psi$$

2.4 Quantification de l'énergie

L'équation de Schrödinger est un modèle mathématique qui décrit le comportement dynamique des particules élémentaires vues comme des ondes. Elle remplace en quelque sorte la deuxième loi de Newton de la mécanique classique que Bohr avait utilisé dans son modèle planétaire de l'atome. Ce modèle ne permettait pas d'expliquer la stabilité des atomes car il implique des variations périodiques de champs électrique et magnétique (voir introduction). De même, il ne permettait pas d'expliquer de façon satisfaisante l'existence de niveaux d'énergie discrets des électrons en orbite autour des noyaux (la discrétisation des niveaux d'énergie était connue depuis l'observation des spectres de la lumière émise par les atomes). Nous allons voir ici que la proposition de de Broglie de considérer les particules élémentaires comme des ondes a permis d'expliquer de façon très simple et naturelle ces deux propriétés physiques des atomes (stabilité et quantification de l'énergie).

Le schéma ci-contre montre un électron gravitant autour d'un proton, le tout constituant un atome d'hydrogène. Si l'électron gravite autour du noyau, c'est parce qu'il est capté par le potentiel électrique attractif du proton. Il s'agit du potentiel coulombien représenté ci-contre par la courbe rouge reproduisant la dépendance en $1/r$ de part et d'autre du noyau situé en l'origine du système de coordonnées. En principe, on ne considère pas les valeurs de r de part et d'autre de l'origine puisqu'il s'agit de la variable radiale supposée toujours positive. Je représente néanmoins le potentiel de cette manière pour que l'on comprenne bien qu'il s'agit d'un potentiel symétrique par rapport au noyau (symétrie centrale).



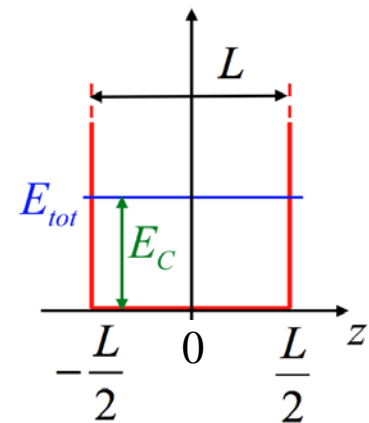
Pour la simplicité, nous allons considérer ici un mouvement linéaire d'aller-retour de l'électron autour du noyau et non pas une orbite comme dans le modèle planétaire de Bohr représenté ci-dessus. L'état atomique correspondant à une telle trajectoire linéaire d'aller-retour existe bien en pratique, c'est d'ailleurs celle qui correspond à l'énergie la plus faible de l'électron et qui est donc la plus souvent rencontrée dans l'atome d'hydrogène. La raison pour laquelle l'électron ne rentre pas en collision avec le noyau pendant son trajet d'aller-retour apparaîtra plus tard.

Comme suggéré sur le schéma ci-dessus, le mouvement d'aller-retour se comprend facilement à partir de la notion de conservation de l'énergie : l'énergie totale est décomposée en énergies cinétique et potentielle et l'électron alterne entre ces deux formes d'énergie tout en maintenant son énergie totale constante. Si le mouvement de

l'électron autour du noyau est périodique c'est donc parce que le potentiel coulombien présente un minimum. On dit en physique que l'électron est confiné dans un « *puits de potentiel* ».

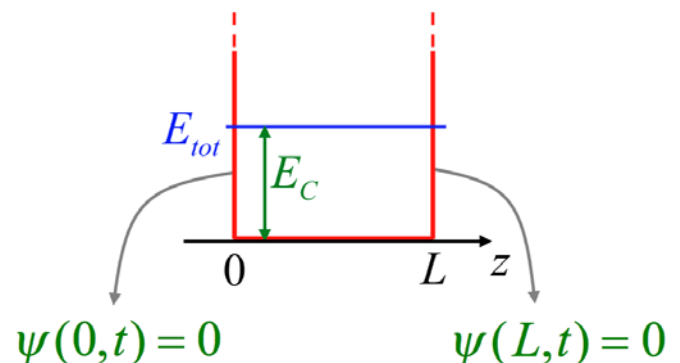
Pour passer de cette description « *classique* » de l'atome à sa description quantique, nous allons considérer un puits de potentiel plus simple que le potentiel coulombien. La résolution de l'équation de Schrödinger pour le puits de potentiel coulombien exige en effet des connaissances de mathématiques qui dépassent le cadre de ce cours. Pour simplifier les mathématiques nous allons nous contenter de résoudre le problème d'un électron confiné dans un puits de potentiel de forme carrée. Il est représenté sur le schéma ci-contre.

puits de potentiel carré

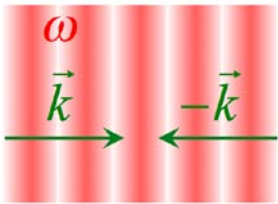


De cette manière, plutôt que de décrire les pentes du potentiel en $1/r$, on décrit le confinement de l'électron par un potentiel qui varie abruptement au-delà d'une certaine distance de l'origine (cette distance vaut $L/2$ sur le schéma ci-dessus) correspondant au rayon de l'atome (c'est-à-dire, le rayon du nuage électronique). Le potentiel passe brusquement de 0 à l'infini au-delà de $\pm L/2$. Cette façon de modéliser l'atome nous permet d'éviter le recours à l'équation de Schrödinger de l'électron lié à un potentiel. En effet, ici l'électron est libre dans tout le puits de potentiel puisque son énergie potentielle ne varie pas (c'est comme s'il était dans le vide en l'absence de toute charge électrique). Mais l'électron ne peut pas non plus dépasser les limites données par le rayon $L/2$ de l'atome car l'énergie potentielle y est infinie, ce qui veut dire obligatoirement que son énergie cinétique s'annule quand il arrive à ces limites $z = \pm L/2$. Selon le principe de conservation de l'énergie l'électron ne peut donc que rebondir sur les bords du puits. Le mouvement de l'électron est donc un mouvement d'aller-retour autour du centre du puits, ce qui est tout à fait comparable au mouvement qu'il a en réalité dans le puits de potentiel coulombien du proton. Plutôt que de faire demi-tour de façon progressive sur les pentes du potentiel en $1/r$, il fait demi-tour abruptement en $z = \pm L/2$. En termes de fonction d'onde, on traduit l'absence de l'électron au-delà des limites du puits par l'annulation de sa fonction d'onde.

Pour la simplicité, nous décrivons ce système avec un axe de coordonnée z dont l'origine est situé sur un bord du puits, comme indiqué sur le schéma ci-contre. On peut ainsi exprimer le confinement de l'électron par le puits de potentiel carré au travers de l'annulation de la fonction d'onde en $z = 0$ et $z = L$, soit $\psi(0,t) = \psi(L,t) = 0$.



Considérons maintenant la résolution de l'équation de Schrödinger pour le puits de potentiel carré. Dans la zone du puits $0 < z < L$, l'équation à considérer est l'équation de Schrödinger de la particule libre, soit $\partial_t \psi = i\hbar / (2m) \partial_{zz}^2 \psi$. Comme on l'a vu dans l'introduction (section 2.1), pour une énergie donnée la fonction d'onde de la particule libre est une onde plane de type $A \exp[i(kz - \omega t)]$. On peut effectivement facilement vérifier que cette fonction est solution de l'équation de Schrödinger pour autant que le nombre d'onde k satisfasse la relation de dispersion de de Broglie, soit $k(\omega) = \sqrt{2m\omega} / \hbar$. Mais cette fonction d'onde ne peut pas être annulée aux bords du puits comme on voudrait l'imposer (le module de la fonction $A \exp[i(kz - \omega t)]$ vaut effectivement toujours A). Cette incohérence apparente est due au fait que la fonction d'onde proposée n'est pas la solution générale de l'équation de Schrödinger. Pour être générale il faut lui rajouter l'onde régressive, comme nous l'avons fait au chapitre 5 pour les ondes de corde. La solution générale est donc de la forme $A \exp[i(kz - \omega t)] + B \exp[i(-kz - \omega t)]$. La fonction d'onde est maintenant annulable en $z = 0$ et $z = L$. Comme pour les ondes de cordes stationnaires, il suffit effectivement que $B = -A$ de sorte que la fonction d'onde devienne $\psi(z, t) = 2iA \sin(kz) \exp(-i\omega t)$.

$$k = 2\sqrt{\frac{\pi m \omega}{h}}$$


$$\psi(z, t) = A e^{i(kz - \omega t)} - A e^{i(-kz - \omega t)}$$

$$\psi(z, t) = [A e^{ikz} - A e^{-ikz}] e^{-i\omega t}$$

$$\psi = A e^{i(kz - \omega t)} - A e^{i(-kz - \omega t)} \quad \psi(z, t) = 2iA \sin(kz) e^{-i\omega t}$$

On constate ainsi l'apparition d'une onde stationnaire comparable à l'onde stationnaire de corde étudiée au chapitre 5. L'interprétation physique de ce résultat est simple. Les discontinuités du potentiel en $z = 0$ et $z = L$ imposent la réflexion des ondes en ces points. Et la réflexion engendre la superposition d'une onde progressive et d'une onde régressive de même amplitude, ce qui conduit à la formation d'une onde stationnaire.

Le problème est tout à fait comparable à celui de l'onde stationnaire de corde résonante mais il convient néanmoins de souligner deux différences de grande importance : d'une part la fonction d'onde est complexe et d'autre part la valeur de k à prendre en compte dans la fonction d'onde est donnée par la relation de dispersion de de Broglie pour la particule libre, soit $k(\omega) = \sqrt{2m\omega} / \hbar$ (et non pas la relation de dispersion des ondes de corde $k(\omega) = \omega / v$, ni celle des ondes électromagnétiques $k(\omega) = \omega / c$!).

Le calcul de la fonction d'onde n'est bien entendu pas terminé car il faut encore imposer les conditions aux limites de l'onde. Pour que la fonction d'onde soit physiquement acceptable dans le cas présent, il faut qu'elle s'annule à la fois en $z=0$, [ce qui est automatique puisque $\sin(0)=0$] et en $z=L$, soit $\psi(L,t) = 2iA \sin(kL) \exp(-i\omega t) = 0$, ce qui implique que le produit kL soit égal à un multiple de π .

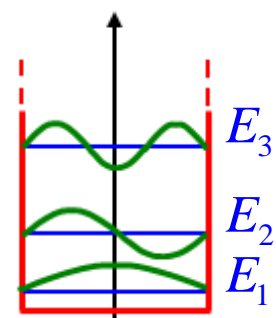
$$\psi(0,t) = 0 \quad \text{et} \quad \psi(L,t) = 0 \quad \text{OK} \quad \text{si} \quad kL = n\pi$$

Autrement dit, le nombre d'onde doit satisfaire l'égalité $k = n\pi/L$. En remplaçant k par sa valeur en fonction de ω , cette relation devient une contrainte sur les valeurs de la fréquence angulaire de l'onde. Autrement dit, la pulsation ne peut prendre que des valeurs discrètes données par la relation ci-dessous où n est un entier :

$$k(\omega) = \frac{n\pi}{L} \Rightarrow \sqrt{\frac{2m\omega_n}{\hbar}} = \frac{n\pi}{L} \Rightarrow \omega_n = \frac{n^2 \hbar \pi^2}{2mL^2}$$

La physique à laquelle nous sommes confrontés ici est tout à fait analogue à celle qui nous avait conduits à la notion de mode propre de vibration d'une corde : une corde fixée à ses deux extrémités ne peut vibrer qu'à des fréquences bien déterminées.

Sachant, d'après les postulats de la mécanique quantique, que l'énergie de l'électron est donnée par sa pulsation multipliée par la constante \hbar , soit $E = \hbar\omega$, on voit que les valeurs discrètes de la fréquence signifient que l'énergie de l'électron dans le puits de potentiel ne peut prendre qu'un ensemble de valeurs discrètes. On dit que l'énergie est « *quantifiée* ». La figure ci-contre montre de façon symbolique les « *modes* » de vibration électronique (les fonctions d'onde) associés aux trois premières valeurs de l'énergie. Les niveaux d'énergie correspondants y sont représentés par des traits horizontaux bleus sur l'axe de l'énergie de la représentation du puits de potentiel.

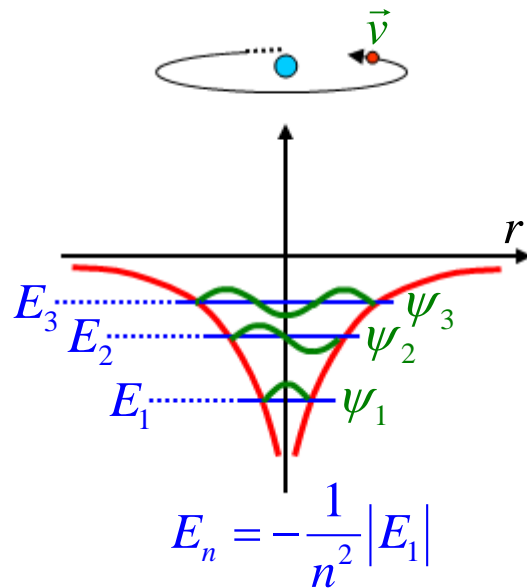


"quantification"

$$E_n = \hbar\omega_n = \frac{n^2 \hbar^2}{8mL^2}$$

De manière évidente, le spectre discret d'énergie en « n^2 » que l'on obtient avec le puits de potentiel carré ne correspond pas à celui qui est observé en pratique sur l'atome d'hydrogène car l'assimilation du potentiel coulombien à un potentiel carré est bien trop grossière. Néanmoins, la phénoménologie reste la même. Ce n'est que du point de vue quantitatif que les choses changent quand on passe du puits carré au potentiel coulombien. La résolution de l'équation de Schrödinger montre plus précisément que les niveaux d'énergie forment une série en $-1/n^2$ plutôt qu'en n^2 .

Ceci est dû au fait que plus on monte en énergie, plus l'électron voit les limites de son mouvement s'éloigner (dans l'analogie à la corde vibrante, tout se passe comme si la corde devenait de plus en plus longue), les fréquences des modes propres de vibration se rapprochent donc l'une de l'autre au fur et à mesure que l'on monte en énergie [ceci peut se comprendre à l'aide du résultat obtenu ci-dessus, soit $\omega_n = n^2 \hbar \pi^2 / (2mL^2)$, qui montre que les fréquences de vibration propres du puits carré sont d'autant plus serrées que la largeur L du puits est grande]. Le signe négatif dans l'expression des énergies de l'atome d'hydrogène provient simplement de la convention de potentiel coulombien nul à l'infini. Avec cette convention, l'électron lié a effectivement toujours une énergie totale négative.



Remarque : Notez que le modèle simplifié du puits carré proposé ici est un modèle unidimensionnel du mouvement de l'électron. Dans la pratique l'électron se déplace dans un espace à trois dimensions et la fonction d'onde associée doit également être décrite à trois dimensions.

2.5 Stabilité de l'atome

Un aspect très important des travaux de de Broglie et de Schrödinger est qu'ils ont apporté une explication naturelle de la stabilité des atomes qui restait jusqu'alors un véritable mystère. En effet, la fonction d'onde d'un état lié de l'électron autour du noyau prend toujours la forme que nous venons de voir pour le puits carré, c'est-à-dire, une fonction réelle des coordonnées de position $f_n(\vec{x})$ multipliée par l'exponentielle imaginaire $\exp(-i\omega_n t)$. La fonction des coordonnées de position $f_n(\vec{x})$ représente physiquement une onde stationnaire (c'est l'onde stationnaire en $\sin(kz)$ du puits carré).

$$\psi_n(\vec{x}, t) = f_n(\vec{x}) e^{-i\omega_n t}$$

$\omega_n = \frac{1}{\hbar} E_n$

onde stationnaire

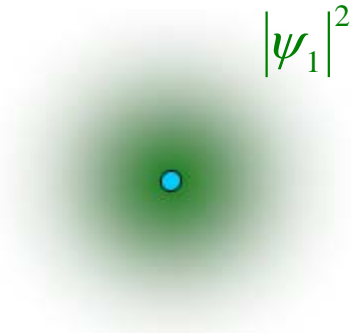
Le fait qu'un état électronique d'énergie donnée corresponde à une onde stationnaire implique que la répartition de charge $\rho(\vec{x})$ du nuage électronique est stationnaire. En effet, selon la proposition de de Broglie, la distribution de charge du nuage électronique est donnée par le module carré de la fonction d'onde. Or, comme la variation temporelle de la fonction d'onde est donnée par une exponentielle imaginaire dont le module vaut toujours l'unité, le module carré de la fonction d'onde ne présente pas de dépendance temporelle, ce qui signifie que le nuage électronique reste rigoureusement immobile.

répartition de charge stationnaire :

$$\rho(\vec{x}, t) \equiv q_e |\psi_n|^2 = q_e f_n^2(\vec{x})$$

En suivant la proposition de de Broglie nous sommes passés du mouvement classique de l'électron ponctuel qui fait des allers-retours autour du noyau, provoquant de la sorte l'émission d'une onde électromagnétique, à une distribution de charge délocalisée stationnaire qui ne peut engendrer la moindre onde électromagnétique. Cet état de l'électron ne peut donc pas perdre d'énergie et il peut dès lors subsister indéfiniment, comme les atomes de la matière qui nous entourent (et dont nous sommes constitués) qui ont été formés il y a des milliards d'années...

Ce schéma ci-contre représente la fonction d'onde d'énergie la plus basse de l'atome d'hydrogène. Elle possède une symétrie sphérique. Le noyau est si petit par rapport au nuage électronique que l'électron n'a qu'une probabilité infime de le rencontrer (la taille des noyaux est de l'ordre du femtomètre, soit 10^{-15} m, alors que les nuages électroniques sont à l'échelle de l'Angström 10^{-10} m, ce qui, en termes de volume, fait une différence de 15 ordres de grandeur.



2.6 Le photon

Le formalisme de la mécanique quantique de Schrödinger permet de comprendre pourquoi les atomes peuvent dans certaines conditions émettre des ondes électromagnétiques (bien souvent sous forme de lumière). Pour aborder cette problématique, considérons ce qui se produit lorsque l'état de l'électron est constitué de la superposition de deux fonctions d'onde stationnaires. L'équation de Schrödinger étant linéaire, on sait que la somme de deux de ses solutions en est également une solution. La somme de deux fonctions d'onde stationnaires peut donc, en principe, représenter en état physiquement acceptable de l'électron. L'expression ci-dessous représente une telle superposition, ω_1 et ω_2 sont les pulsations propres a priori différentes des deux états superposés et les coefficients α et β sont des nombres arbitraires (que nous supposons réels) représentant les poids relatifs de chacune des fonctions d'onde dans l'état superposé.

$$\psi = \alpha\psi_1 + \beta\psi_2 = \alpha f_1 e^{-i\omega_1 t} + \beta f_2 e^{-i\omega_2 t}$$

Pour avoir la répartition de charge de cet état électronique particulier nous avons besoin du module carré de sa fonction d'onde, soit $|\psi|^2$, qui s'obtient en multipliant la fonction d'onde par son complexe conjugué :

$$|\psi|^2 = (\alpha f_1 e^{-i\omega_1 t} + \beta f_2 e^{-i\omega_2 t}) \times (\alpha f_1 e^{i\omega_1 t} + \beta f_2 e^{i\omega_2 t})$$

$$\hookrightarrow |\psi|^2 = \alpha^2 f_1^2 + \beta^2 f_2^2 + \alpha\beta f_1 f_2 e^{-i\omega_1 t} e^{i\omega_2 t} + \alpha\beta f_1 f_2 e^{i\omega_1 t} e^{-i\omega_2 t}$$

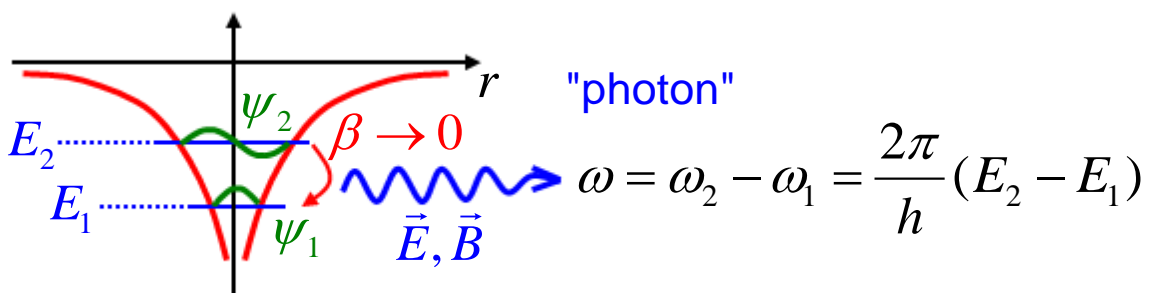
$$\hookrightarrow |\psi|^2 = \alpha^2 f_1^2 + \beta^2 f_2^2 + \underbrace{2\alpha\beta f_1 f_2 \cos[(\omega_1 - \omega_2)t]}_{\text{(battement)}}$$

On constate ainsi que le nuage électronique présente cette fois des variations temporelles à la fréquence angulaire $\omega_2 - \omega_1$. C'est le phénomène de battement entre deux signaux harmoniques que nous avons rencontré au chapitre 5. Le fait que les deux fonctions d'onde n'ont pas la même énergie (pas la même pulsation) fait que leur superposition « bat » à la fréquence différence $\omega_2 - \omega_1$. Les fonctions f_1 et f_2 étant réelles, on peut facilement combiner les deux exponentielles imaginaires pour obtenir une fonction cosinus (oscillation harmonique). Ceci signifie que la distribution de charge n'est plus stationnaire mais oscille à la fréquence différence $\omega_2 - \omega_1$. Une telle distribution de charge variable génère naturellement une onde électromagnétique.

$$|\psi|^2 = \alpha^2 f_1^2 + \beta^2 f_2^2 + 2\alpha\beta f_1 f_2 \cos[(\omega_2 - \omega_1)t]$$

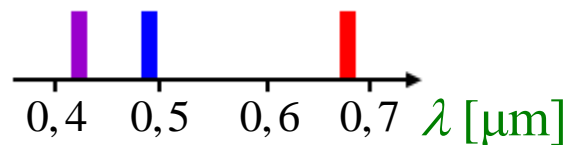
$$\rho(\vec{x}, t) \equiv q_e |\psi|^2 \quad \longrightarrow \quad \frac{\partial \rho}{\partial t} \neq 0 \quad \longrightarrow \quad \text{onde EM}$$

Le schéma ci-dessous explique comment ce phénomène de battement a lieu en pratique. Supposons que, suite à un phénomène de transfert d'énergie quelconque (une collision avec un autre atome, par exemple), l'électron ait été mis dans un état d'énergie E_2 supérieure à celle de l'état d'énergie la plus basse E_1 (appelé état fondamental). Supposons également que l'état 2 n'ait pas été généré parfaitement et qu'il y a une légère superposition avec l'état d'énergie E_1 (le coefficient α ci-dessus est petit mais non nul). La fonction d'onde de l'électron se mettra donc à battre à la fréquence différence $\omega_2 - \omega_1$. L'électron émettra donc une onde EM. En émettant une onde EM, l'électron donne son énergie au champ EM. Son énergie qui était initialement proche de E_2 va donc diminuer, ce qui correspond à une augmentation du poids de la fonction d'onde d'énergie E_1 autrement dit, α grandit et β diminue. Le produit $\alpha\beta$ apparaissant dans la composante oscillante de la distribution de charge (voir la formule ci-dessus) passe donc par un maximum et puis diminue vers zéro jusqu'à ce que toute la différence d'énergie $E_2 - E_1$ ait été transférée vers l'onde EM. Le paquet d'ondes EM résultant est un photon. L'énergie du photon est l'énergie donnée par l'électron c'est-à-dire $E_2 - E_1 = \hbar(\omega_2 - \omega_1)$. Or, $\omega_2 - \omega_1$ est précisément la pulsation de l'onde EM (c'est-à-dire, la pulsation du photon), on arrive donc à la conclusion que la fréquence du photon est donnée par la différence des niveaux d'énergie impliqués dans la dynamique du nuage électronique de l'atome.



Ceci explique pourquoi la lumière émise par les atomes présente des spectres discrets. Les raies spectrales représentées ci-dessous correspondent aux différences d'énergie des premiers états électroniques de l'atome d'hydrogène. C'est une partie du spectre d'émission de l'hydrogène. L'étude de ce type de spectre a permis de connaître les niveaux d'énergie des états électroniques de la plupart des atomes existants (c'est le principe de la spectroscopie atomique).

spectres atomiques



Remarque 1 : Il est intéressant de noter que le raisonnement ci-dessus nous amène à la conclusion que l'énergie E du photon est liée à sa pulsation ω par la relation $E = \hbar\omega$. En effet, nous avons vu d'une part que l'énergie de l'onde électromagnétique générée par l'atome est l'énergie perdue par l'électron lors de la transition de l'état 2 à l'état 1, soit $E = E_2 - E_1 = \hbar(\omega_2 - \omega_1)$. Et d'autre part, nous avons vu que la pulsation de l'onde EM générée est la pulsation de battement des deux fonctions d'onde correspondantes, soit $\omega = \omega_2 - \omega_1$, on trouve donc bien la relation $E = \hbar\omega$ entre l'énergie E du photon et sa pulsation ω . Cette relation avait déjà été découverte par Planck dès le début du 20^{ème} siècle lors de l'observation des spectres d'émission de corps chauffés à haute température. C'est pourquoi la constante h porte le nom de constante de Planck. C'est le raisonnement que nous venons de faire sur l'émission du photon qui a poussé Louis de Broglie à proposer $E = \hbar\omega$ comme relation entre l'énergie des électrons orbitaux et la pulsation de leur fonction d'onde.

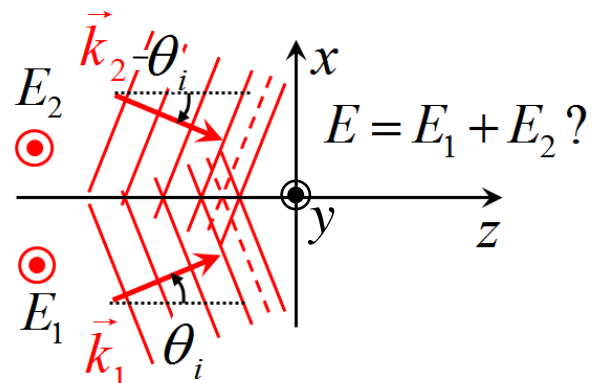
Remarque 2 : Notez que cette vision des choses est extrêmement simplifiée. En particulier, dans la version moderne de la mécanique quantique le champ électromagnétique lui-même est quantifié et le photon n'est pas considéré comme un paquet d'ondes classique, comme je le suggère ci-dessus. Néanmoins, cette vision simple des choses n'est pas entièrement fausse et elle a l'avantage d'expliquer la phénoménologie de façon imagée et intuitive correspondant de plus aux aspects historiques de la démarche scientifique qui a mené à la compréhension du monde microscopique. Une approche beaucoup plus rigoureuse de la physique atomique vous sera donnée dans le cours de Physique Quantique et Statistique de deuxième année.

3. Phénomènes d'interférence

Les ondes sont omniprésentes dans le monde qui nous entoure, ceci tant à l'échelle macroscopique qu'à l'échelle microscopique. Comme nous l'avons vu à plusieurs reprises, ces ondes sont pour la plupart régies par l'équation d'onde standard dont le caractère linéaire reflète le principe de superposition. Ce principe dit que les ondes peuvent se superposer sans interagir l'une avec l'autre (sans se modifier mutuellement), ce qui implique du point de vue mathématique qu'elles s'additionnent simplement pour donner lieu à des phénomènes d'interférences. Si deux ondes sont sommées en phase il en résultera une onde dont l'amplitude vaut la somme des amplitudes des ondes individuelles ; on a alors ce que l'on appelle une interférence constructive. Par contre si les deux ondes sont en opposition de phase, on a affaire à une interférence destructive. A titre d'exemple, à l'échelle macroscopique nous sommes plongés dans un monde traversé en permanence par une multitude d'ondes électromagnétiques utilisées à des fins de communication avec des applications telles que la radio, la télévision, le téléphone, le GPS, etc.... La superposition de ces ondes donne lieu à des phénomènes d'interférence qui peuvent avoir des répercussions très importantes sur ces applications. A l'échelle microscopique, en raison de la nature ondulatoire des particules élémentaires, les phénomènes d'interférence jouent un rôle déterminant dans de nombreuses propriétés de la matière. Par exemple, c'est le phénomène d'interférence qui est responsable de la quantification des niveaux d'énergie atomiques dans la mesure où la formation d'ondes stationnaires est précisément le résultat d'une interférence entre deux ondes contra-propagatives. On comprend ainsi que l'étude des phénomènes d'interférence est indispensable à la compréhension d'une multitude de phénomènes naturels et de dispositifs techniques. Nous allons dans cette section découvrir certains aspects des interférences. Nous verrons, en particulier, le phénomène de diffraction qui est le résultat d'interférences entre un grand nombre d'ondes. Les illustrations sont proposées dans le domaine des ondes électromagnétiques mais les conclusions que nous allons pouvoir en tirer sont valables pour tout type d'onde, qu'il s'agisse d'ondes acoustiques, d'ondes de matière ou encore d'ondes de surface (vagues à la surface de l'eau par exemple).

3.1 Interférences d'ondes planes

Pour commencer avec un exemple simple, je propose d'étudier le phénomène d'interférence entre deux ondes planes telles que des faisceaux laser collimatés. Les deux ondes planes dont on va considérer la superposition sont représentées schématiquement ci-contre. Elles ont pour champ électrique \vec{E}_1 et \vec{E}_2 de vecteurs d'onde \vec{k}_1 et \vec{k}_2 . Les ondes ont la même pulsation ω et la même polarisation



parallèle à l'axe y , telle que symbolisée par le vecteur sortant du plan de la feuille sur la gauche du schéma ci-dessus (le même symbole est utilisé pour l'axe y). Le système d'axe de coordonnées choisi pour décrire ces ondes a son axe z orienté parallèlement à la bissectrice des deux vecteurs d'onde \vec{k}_1 et \vec{k}_2 . Les deux ondes font donc avec l'axe z des angles θ_i identiques en valeur absolue.

Vu la symétrie du système, il apparaît évident que les vecteurs d'onde sont la forme suivante :

$$\begin{cases} \vec{k}_1 = k_x \vec{1}_x + k_z \vec{1}_z \\ \vec{k}_2 = -k_x \vec{1}_x + k_z \vec{1}_z \end{cases} \quad \text{où} \quad \begin{cases} k_x = k \sin \theta_i \\ k_z = k \cos \theta_i \end{cases}$$

θ_i est l'angle que font les deux ondes planes avec l'axe z (voir schéma ci-dessus).

Les composantes en y des champs ont la forme mathématique suivante dans laquelle nous avons introduit la notation des phaseurs :

$$\begin{cases} E_1(\vec{x}, t) = A \cos(\vec{k}_1 \cdot \vec{x} - \omega t) = \text{Re}[A e^{i(\vec{k}_1 \cdot \vec{x} - \omega t)}] \\ E_2(\vec{x}, t) = A \cos(\vec{k}_2 \cdot \vec{x} - \omega t) = \text{Re}[A e^{i(\vec{k}_2 \cdot \vec{x} - \omega t)}] \end{cases}$$

Nous identifions facilement dans ces expressions les phaseurs des champs :

$$\begin{cases} \underline{E}_1 = A e^{i\vec{k}_1 \cdot \vec{x}} = A e^{ik_x x} e^{ik_z z} \\ \underline{E}_2 = A e^{i\vec{k}_2 \cdot \vec{x}} = A e^{-ik_x x} e^{ik_z z} \end{cases}$$


La somme des champs en termes de phaseur donne l'expression suivante :

$$\underline{E}_1 + \underline{E}_2 = A e^{ik_x x} e^{ik_z z} + A e^{-ik_x x} e^{ik_z z}$$

On peut mettre l'exponentielle imaginaire en $k_z z$ en évidence :


$$\underline{E}_1 + \underline{E}_2 = A (e^{ik_x x} + e^{-ik_x x}) e^{ik_z z}$$

Ce qui fait apparaître la fonction cosinus :




$$\underline{E}_1 + \underline{E}_2 = 2A \cos(k_x x) e^{ik_z z}$$

Le champ réel est donc donné par l'expression suivante :



$$E = \text{Re}[2A \cos(k_x x) e^{ik_z z} e^{-i\omega t}]$$

Ce qui donne après extraction de la partie réelle :



$$E = 2A \underbrace{\cos(k_x x)}_{\text{modulation}} \underbrace{\cos(k_z z - \omega t)}_{\text{onde plane}}$$

Le deuxième facteur de cette expression montre qu'il s'agit d'une onde se propageant dans la direction des z positifs à la vitesse de phase ω/k_z et dont les fronts d'onde sont plans et perpendiculaires à l'axe z (le lieu des points de phase constante de l'onde sont effectivement des plans $z = cste$). Le premier facteur quant à lui montre que cette onde plane présente une modulation stationnaire d'amplitude en x selon la fonction $\cos(k_x x)$. Cette modulation représente le phénomène d'interférence. Ce phénomène peut être comparé au phénomène de battement de deux ondes de fréquences différentes vu au chapitre 5. Dans le cas du battement, l'amplitude de l'onde plane est modulée dans le temps tandis que dans le cas présent l'amplitude est modulée selon la variable spatiale transverse x .

Remarque : Notez que ce résultat aurait pu être obtenu directement avec l'expression réelle des champs mais nous aurions eu besoin pour cela de la formule de Simpson pour la somme des cosinus. Cet exemple montre que l'utilisation des phaseurs permet d'obtenir très simplement les formules de Simpson, ce qui illustre l'utilité des phaseurs en tant qu'outils de simplification des calculs.

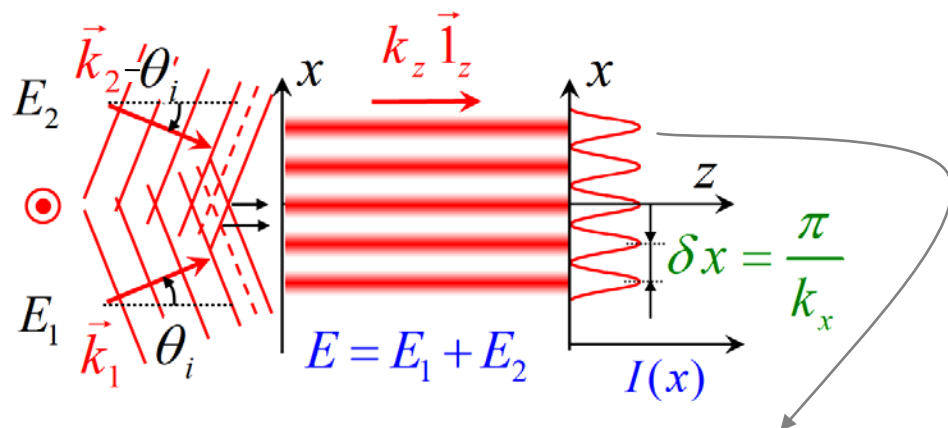
3.2 Figure d'interférence, intensité lumineuse

Comme c'est l'intensité lumineuse que l'on observe dans toute expérience d'optique, il est intéressant de voir quelle est la distribution d'intensité de l'onde résultant des interférences. L'intensité de l'onde est la puissance électromagnétique transportée par l'onde par unité de surface. Elle s'exprime en W/m^2 . En d'autres termes, l'intensité d'une onde électromagnétique en un point de l'espace représente l'énergie passant en ce point par unité de temps et de surface, c'est une densité de flux d'énergie. Pour

comprendre qu'une onde électromagnétique transporte de l'énergie, il faut se rappeler que le champ électrique et le champ magnétique renferment une densité d'énergie proportionnelle à leur amplitude au carré, soit $w_e = \epsilon_0 \|\vec{E}\|^2 / 2$ et $w_m = \|\vec{B}\|^2 / (2\mu_0)$. Dans une onde électromagnétique se propageant dans le vide, le champ magnétique et le champ électrique sont liés par la vitesse de la lumière selon la relation de proportionnalité $\|\vec{B}\| = \|\vec{E}\| / c$. On en déduit que la somme des énergies électrique et magnétique $w_e + w_m = \epsilon_0 \|\vec{E}\|^2 / 2 + \|\vec{B}\|^2 / (2\mu_0)$ est proportionnelle au carré du module du champ électrique. La théorie de l'électromagnétisme montre plus précisément que l'intensité est proportionnelle au module carré du phaseur du champ électrique (voir le théorème de Poynting dans le cours d'électromagnétisme de 3^{ème} année). Pour simplifier nous allons ici considérer que la constante de proportionnalité est unitaire et que l'intensité est donc donnée par le module carré du phaseur de champ électrique. Or, le phaseur du champ électrique est ici $2A \cos(k_x x) e^{ik_z z}$, ce qui donne :

$$I(x, z) \equiv |\underline{E}(x, z)|^2 = 4A^2 \cos^2(k_x x)$$

Cette distribution d'intensité est représentée sur le schéma ci-dessous sous forme de niveaux de gris. On remarque que la répartition d'intensité est indépendante de z , conformément à l'expression mathématique ci-dessus. On peut voir sur le schéma que les maxima de la fonction $\cos^2(k_x x)$ se situent là où les fronts d'onde se croisent. En ces points de croisement, les champs sont maximum pour les deux ondes et leur superposition donne donc un maximum d'intensité (les fronts d'onde sont définis comme étant les lieux des points \vec{x} pour lesquels la phase du champ vaut un multiple de 2π , soit $\vec{k}_1 \cdot \vec{x} - \omega t = 2m\pi$). Ce sont les points qui correspondent aux interférences « constructives » des ondes E_1 et E_2 .



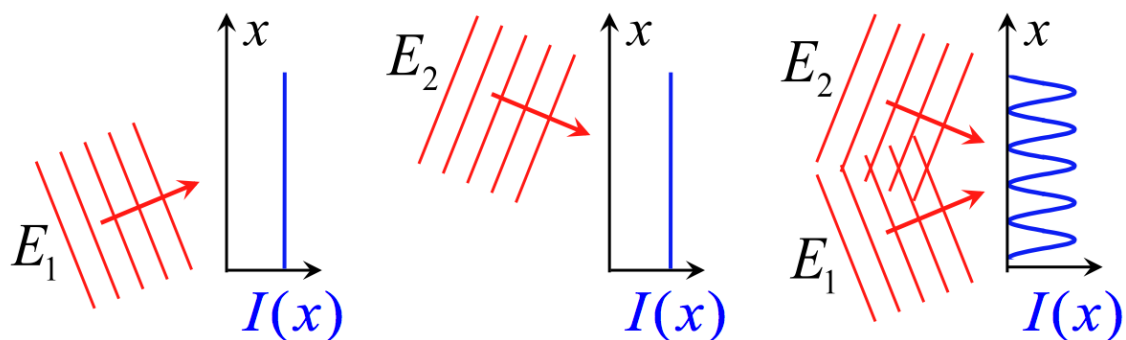
interférences constructives et destructives

On voit également que les zéros de la fonction $\cos^2(k_x x)$ se situent là où les fronts d'onde d'une onde (soit E_1) se superposent avec les plans d'amplitude minimale (négative) de l'autre onde (soit E_2). Ces plans d'amplitude minimale sont les plans qui se trouvent à mi-distance entre les fronts d'onde ; ils sont représentés en coupe par les lignes pointillées sur le schéma. La phase correspondante vaut un multiple impair de

π soit, $\vec{k}_1 \cdot \vec{x} - \omega t = (2m+1)\pi$ et correspond donc bien à une amplitude négative $-A$ opposée à l'amplitude maximale A de l'onde. La superposition des champs en ces points donne donc bien un champ nul, ce sont les zéros de la fonction $\cos^2(k_x x)$ correspondant aux interférences « destructives » des deux ondes.

La structure périodique d'intensité variable correspondant à la fonction $\cos^2(k_x x)$, peut s'observer facilement en plaçant un écran perpendiculairement à l'axe z . Un tel écran diffuserait effectivement la lumière avec une intensité proportionnelle à la fonction $\cos^2(k_x x)$. Ce profil d'intensité est représenté avec à l'aide de la courbe $I(x)$ sur la droite du schéma ci-dessus. Les bandes lumineuses brillantes et noires observées s'appellent des « franges d'interférence ». Comme indiqué sur le schéma ci-dessus, les franges sont séparées de la distance π/k_x [c'est la période de la fonction $\cos^2(k_x x)$].

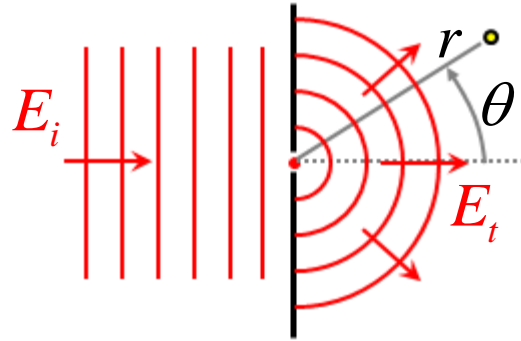
Ce phénomène peut paraître étonnant si on raisonne en termes de flux de photons. Le faisceau lumineux E_1 constitue un flux de photons uniforme dans l'espace et dans le temps (onde plane) et il en est de même pour le faisceau lumineux E_2 (voir les deux schémas de gauche ci-dessous). Il semble donc légitime de penser que lorsque les deux flux de photons sont superposés, la répartition des photons reste uniforme. Mais on constate mathématiquement et expérimentalement qu'il n'en est rien (voir schéma de droite ci-dessous). En particulier, la probabilité de rencontrer un photon à une distance $\pi/(2k_x)$ de l'axe z est nulle [zéro de la fonction $\cos^2(k_x x)$] même si les deux flux de photons sont uniformes. Ceci est dû au fait que ce point est le lieu d'une interférence destructive entre les deux flux de photons.



Ce résultat est applicable aux ondes de matière auquel cas les ondes planes représentent des faisceaux de particules élémentaires telles que, par exemple, des électrons. Si l'on bombarde une cible à l'aide de deux faisceaux uniformes d'électrons il y aura des zones (répétées périodiquement) de la cible qui ne seront jamais touchées par les électrons. De telles expériences sont simples à réaliser ; elles constituent autant de preuves irréfutables de la nature ondulatoire des particules élémentaires qui constituent notre monde.

3.3 Expérience de Young

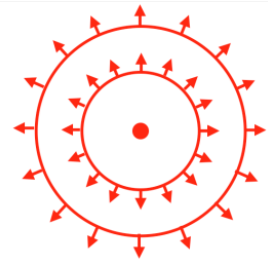
Le phénomène d'interférence est ce qui a permis de découvrir la nature ondulatoire de la lumière dans le courant du 17^{ème} siècle. C'est un scientifique du nom de Grimaldi qui fit les premières descriptions du phénomène vers 1650 mais il a fallu attendre 1800 pour que ces expériences soient reproductibles et prennent ainsi la valeur de preuves scientifiques. C'est Thomas Young qui réussit cette prouesse grâce à une expérience qui porte maintenant son nom. L'« *expérience de Young* » est toute simple mais pour la réussir à l'époque il a fallu faire preuve de beaucoup d'adresse et de perspicacité (ceci principalement en raison de la nature polychromatique et incohérente des sources lumineuses existant à l'époque). Pour améliorer la cohérence de la lumière du soleil utilisée dans son expérience, Young faisait passer les faisceaux lumineux par un trou microscopique percé dans un écran, à l'image du schéma ci-contre.



En faisant cela, Young créait un « *point source* » d'où la lumière émergeait dans toutes les directions de façon isotrope. En d'autres termes, il a créé de cette manière ce que l'on appelle une onde « *sphérique* ». Sur le schéma ci-dessus le champ de cette onde est noté E_t (t pour onde transmise) alors que l'onde incidente est notée E_i . Ce résultat peut être compris intuitivement en se disant que, si le trou est suffisamment petit, la vibration du champ incident constitue une perturbation homogène du champ électromagnétique (le champ prend la même valeur sur toute la surface du trou). En vertu de l'équation d'onde EM, cette perturbation se transmettra dans tout l'espace et, puisqu'elle est quasi-ponctuelle, elle se propage de façon isotrope dans toutes les directions (il faut pour cela que le trou ait des dimensions inférieures ou de l'ordre de grandeur de la longueur d'onde de la lumière). On peut facilement expérimenter l'émission d'ondes dans de telles conditions avec les ondes de surface se propageant à la surface de l'eau : en agitant la surface en un point on génère une onde qui se propage dans toutes les directions autour du point en question (l'onde est dans ce cas circulaire).

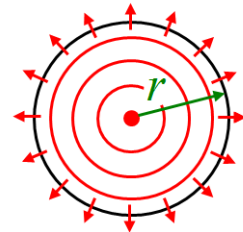
Une autre manière d'interpréter ce qui se produit est de dire que tous les points situés sur une sphère centrée sur la source ponctuelle « *voient* » la même perturbation de champ électromagnétique et subissent donc la même variation de champ en raison de la propagation de cette perturbation. L'onde ainsi générée possède de façon évidente une géométrie sphérique. La forme mathématique de cette onde est donc du type $A \cos(kr - \omega t)$ où r est la coordonnée radiale centrée sur le point source. En effet, on voit facilement dans cette expression que les fronts d'onde, c'est-à-dire les lieux des points de phase constante à un temps t donné, sont des sphères $r = cste$.

Puisque la surface des fronts d'onde de l'onde sphérique grandit lors de leur propagation, la densité de flux d'énergie qu'ils transportent diminue avec la distance, comme le suggère le schéma ci-contre. On en déduit que l'amplitude d'une onde sphérique varie avec la distance r , à la différence de l'onde plane qui a une amplitude constante. Ceci est facile à comprendre si on considère la conservation de l'énergie. Pour



cela il convient d'introduire la notation des phaseurs, puisque l'intensité est égale au module carré du phaseur de champ électrique. En termes de phaseurs l'onde sphérique a la forme $\underline{E}_t = A(r) \exp(ikr)$ où $A(r)$ est son amplitude. L'intensité de l'onde est donc donnée par $I(r) = |\underline{E}_t|^2 = A^2(r)$ (je rappelle que l'intensité lumineuse est une densité de flux d'énergie donnée en W/m^2). Si on intègre l'intensité $I = |\underline{E}_t|^2$ de l'onde sur une sphère de rayon r , on obtient la puissance totale radiée par le point source, soit $4\pi r^2 A^2$. Par conservation de l'énergie, cette puissance est bien entendu indépendante du rayon r de la sphère sur laquelle on intègre (on suppose qu'il n'y a pas d'absorption de l'onde), ce qui mathématiquement s'exprime en écrivant $4\pi r^2 A^2 = \text{cste}$. Cette relation nous indique que l'amplitude A de l'onde sphérique est inversement proportionnelle à la distance de propagation, soit $A = a/r$. On peut donc écrire le champ électrique de l'onde sphérique de la manière suivante :

Onde sphérique : $\underline{E}_t = \frac{a}{r} e^{ikr}$



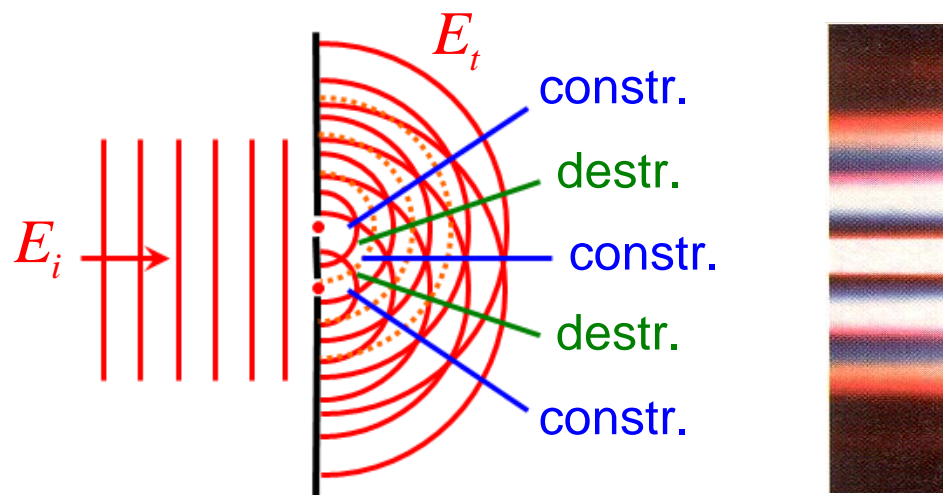
Le raisonnement que l'on vient de faire est résumé à l'aide du schéma et des formules ci-contre. Notez que l'on retrouve la dépendance en $1/r^2$ de la loi de Coulomb que l'on avait formulée au chapitre 2 en termes de densité de flux de photons virtuels en appliquant un raisonnement tout à fait semblable.

$$I = |\underline{E}_t|^2 = \frac{a^2}{r^2}$$

$$\hookrightarrow 4\pi r^2 I = \text{cste}$$

Cette brève introduction à la notion d'onde sphérique nous permet d'aborder l'expérience de Young proprement dite. Cette expérience réalisée vers 1800 a provoqué un véritable choc dans la communauté scientifique car elle a permis de mettre en évidence de manière irréfutable la nature ondulatoire de la lumière. Malgré les travaux de Grimaldi, la nature ondulatoire de la lumière n'était pas reconnue jusqu'alors. En particulier, Newton avait formulé un modèle de la lumière dans lequel celle-ci apparaissait comme un flux de particules élémentaires. L'expérience de Young a une grande importance historique car elle constitue le point de départ d'une réflexion et d'une démarche scientifique qui a conduit à la théorie ondulatoire de la lumière qui s'est trouvée confirmée par les travaux de Maxwell sur les ondes électromagnétiques. Et c'est ensuite sur la théorie des ondes électromagnétiques que s'est calquée la théorie de la mécanique quantique.

L'expérience de Young consiste simplement à faire passer une onde sphérique par deux ouvertures de petite dimension percées dans un écran opaque, comme l'indique le schéma ci-dessous (l'onde sphérique incidente est représentée par des fronts d'onde plans car la source est en général très éloignée de l'écran relativement à l'écart entre les deux trous). La représentation des fronts d'onde sur le schéma montre qu'en se superposant les deux ondes sphériques générées interfèrent constructivement dans certaines directions. Il s'agit des directions dans lesquelles les fronts d'onde se superposent (ceux-ci sont représentés par les lignes circulaires pleines sur le schéma). On voit également que les ondes interfèrent également destructivement dans les directions où les fronts d'onde d'une onde se superposent aux lieux des points d'amplitude minimale (lignes circulaires pointillées sur le schéma).

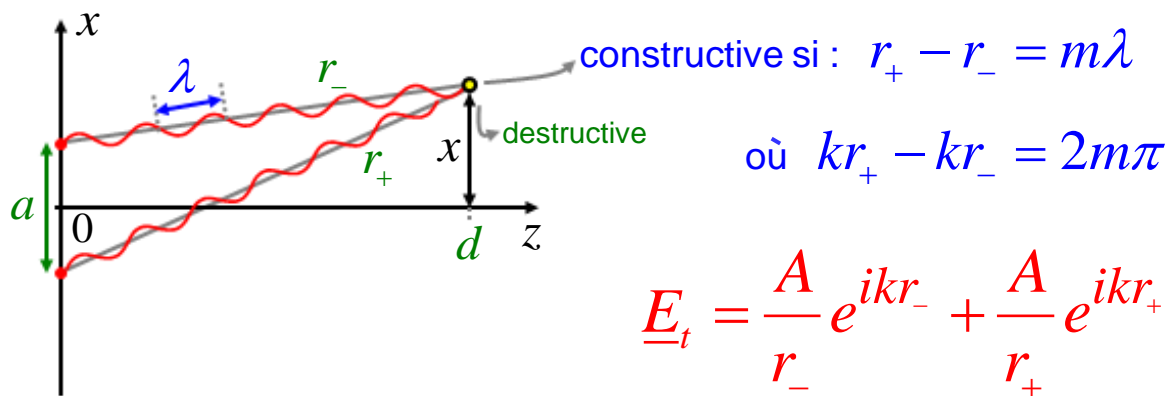


En conclusion, si on place un deuxième écran à une certaine distance de l'écran percé, on observe une série de franges d'interférences, comme dans la situation vue plus haut avec la superposition des deux ondes planes. On en voit une photographie sur la droite du schéma ci-dessus. C'est ce que l'on appelle les « *franges de Young* ».

Remarque : Notez que si l'on fait cette expérience avec un faisceau de particules élémentaires telles que des électrons, on obtient le même phénomène de formation de franges périodiques. C'est une des expériences qui a permis de mettre en évidence la nature ondulatoire des électrons postulée par de Broglie, ce qui illustre l'importance historique de l'expérience de Young une grande importance dans l'histoire des sciences.

Les franges d'interférences de la photo à droite du schéma ci-dessus ont été obtenues en utilisant une source de lumière « *blanche* », c'est-à-dire, de la lumière qui contient toutes les couleurs du spectre visible. Ces couleurs sont visibles sur la photo sous forme de bandes rouges et bleues de part et d'autre des franges brillantes. Ce phénomène de décomposition de la lumière blanche peut s'expliquer facilement à partir de la description mathématique de l'expérience de Young.

Pour introduire la description mathématique, l'expérience de Young est schématisée sur le dessin ci-dessous où sont représentés deux rayons des ondes sphériques émanant des deux ouvertures et arrivant en un point donné de l'écran d'observation. Sur les rayons sont représentées schématiquement les variations périodiques du champ électrique des deux ondes. On voit que le point de l'écran choisi sur le schéma correspond à une condition d'interférence destructive (le maximum d'une onde est superposé à un minimum de l'autre). De manière évidente, les ondes ne se somment constructivement que si les distances r_- et r_+ (qui séparent le point de l'écran des deux ouvertures) diffèrent d'un nombre entier de longueurs d'onde, soit $r_+ - r_- = m\lambda$, ce que l'on peut encore écrire, en introduisant le nombre d'onde $k = 2\pi / \lambda$, comme suit $k(r_+ - r_-) = 2m\pi$.



Nous allons ici calculer la position des franges de Young sur un écran d'observation situé à une distance d des deux trous. Il suffit pour cela d'exprimer l'onde transmise par les deux trous comme étant la somme des deux ondes sphériques, comme indiqué ci-dessus en termes de phaseurs. Les distances r_- et r_+ peuvent être exprimées facilement en termes de la hauteur x dans le plan d'observation. Dans les expressions ci-dessous, la distance entre les ouvertures de l'écran est notée a .

Ces expressions de r_- et r_+ permettent de calculer la position des franges brillantes données par la condition $r_+ - r_- = m\lambda$.

$$r_{\pm} = \sqrt{d^2 + \left(x \pm \frac{1}{2}a\right)^2}$$

Si la distance d est beaucoup plus grande que la distance a de séparation des trous, et si l'on se contente de décrire les premières franges de Young caractérisées par des petites valeurs de x , on peut considérer l'inégalité $d \gg a, x$ de façon à pouvoir approcher l'expression mathématique de r_- et r_+ au premier ordre en les rapports $(a/d)^2$, $(x/d)^2$ et ax/d^2 . C'est ce qui est fait ci-dessous :

$$r_{\pm} = \sqrt{d^2 + x^2 \pm xa + a^2 / 4}$$

$$d \gg a, x$$

$$\hookrightarrow r_{\pm} = d \sqrt{1 + (x/d)^2 \pm xa/d^2}$$

$$\hookrightarrow r_{\pm} = d \left[1 + \frac{1}{2} (x/d)^2 \pm \frac{1}{2} xa/d^2 \right]$$

La différence entre r_- et r_+ donne donc simplement xa/d . L'expression de la condition de formation des interférences constructives $r_+ - r_- = m\lambda$ devient donc $xa/d = m\lambda$. Ceci nous indique que les franges d'interférences constructives (franges brillantes) sont situées aux distances $x_m = m\lambda d/a$. Les franges apparaissent donc équidistantes dans notre approximation du 1^{er} ordre, la distance entre franges étant $\lambda d/a$.

$$\hookrightarrow r_+ - r_- = x \frac{a}{d} = m\lambda \quad \Rightarrow \quad x_m = \frac{m\lambda d}{a}$$

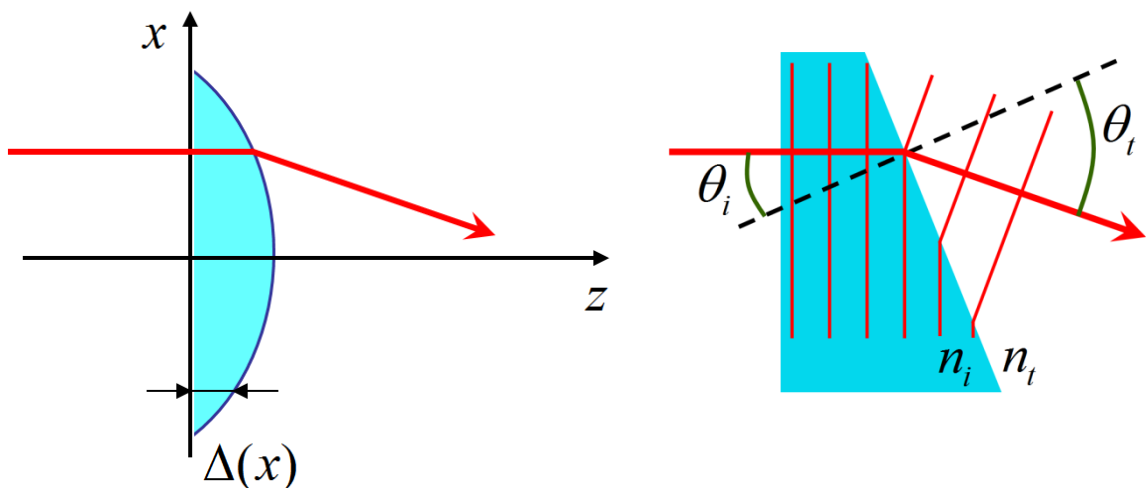
Cette dernière relation montre que plus la longueur d'onde est grande, plus la frange d'un ordre m donné est éloignée de l'origine. Ceci explique pourquoi la lumière blanche est décomposée dans le processus d'interférence. Si l'expérience de Young se fait avec de la lumière blanche, les différentes longueurs d'onde λ qui la composent forment des interférences constructives à des distances différentes données par $x_m = m\lambda d/a$. Les franges rouges se trouvent donc plus éloignées de l'origine que les franges bleues (sur la photo ci-dessus on voit de la lumière blanche entre les franges de couleurs bleues et rouges car dans ces zones centrales toutes les couleurs se superposent). Remarquons encore que pour la frange d'ordre zéro ($m=0$), toutes les couleurs forment leur interférence constructive en $x=0$, c'est la raison pour laquelle il n'y a pas de décomposition lumineuse pour cette frange particulière (voir la photo ci-dessus où cette frange apparaît effectivement blanche).

4. Les lentilles

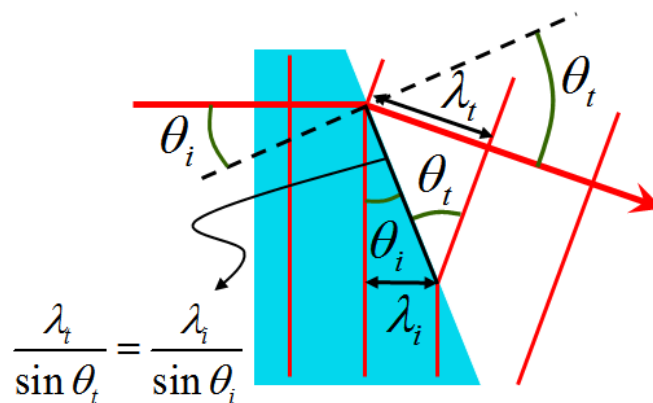
Une lentille est une lame de verre d'épaisseur variable qui a pour effet de déformer les fronts d'onde de la lumière qui passe au travers, le plus souvent dans le but de former des images. Les lentilles sont souvent décrites dans le contexte de l'optique géométrique. Dans ce contexte c'est la trajectoire des rayons lumineux qui est considérée et non les fronts d'onde. C'est le phénomène de réfraction régi par la loi de Snell-Descartes qui est responsable de la déviation des rayons lumineux sur les surfaces de la lentille. A titre d'illustration une lentille convergente plan-convexe est représentée schématiquement ci-dessous. La fonction $\Delta(x)$ donne l'épaisseur de la lentille en fonction de la coordonnée transverse x (pour la simplicité nous nous limitons ici à une description à 2 dimensions, il nous sera facile plus loin de généraliser nos résultats à 3 dimensions). La déflexion du rayon lumineux est due au fait que l'angle d'incidence qu'il fait avec la surface de sortie de la lentille n'est pas nul. La loi de Snell-Descartes, permet de quantifier la déflexion en donnant l'angle θ_t du rayon transmis (angle donné par rapport à la normale à la surface de verre).

4.1 Loi de la réfraction de Snell-Descartes

La déflexion des rayons à la sortie de la lentille est régie par la loi de Snell-Descartes $n_i \sin \theta_i = n_t \sin \theta_t$ où les grandeurs n_i et n_t sont les indices de réfraction du verre et de l'air, respectivement. Cette loi est illustrée ci-dessous à droite avec un schéma montrant une vue rapprochée du rayon émergent de la lentille. En vue très rapprochée la courbure de la lentille n'apparaît plus et l'interface peut être considérée comme plane, c'est-à-dire que l'on peut effectivement lui appliquer la loi de Snell-Descartes développée pour les dioptries plans. On constate que le rayon est défléchi vers l'axe optique (axe z de symétrie de la lentille) puisque l'indice du verre est plus grand que l'indice de l'air, soit $n_i > n_t$, et donc, en vertu de la loi de Snell-Descartes, on a bien $\theta_i < \theta_t$.



Afin de comprendre l'origine de la loi de Snell-Descartes je propose de l'interpréter dans le domaine de l'optique ondulatoire. Dans ce but le schéma ci-dessous représente les fronts d'onde des ondes incidente et réfractée ainsi que les rayons qui leur sont associés. Les fronts d'onde sont les lieux des points pour lesquels la phase de l'onde vaut un multiple pair de π , c'est-à-dire les points pour lesquels le champ électromagnétique de l'onde passe par un maximum. Il est évident que l'onde traverse l'interface sans provoquer de rupture des fronts d'onde (on comprendrait effectivement mal pourquoi le champ ne resterait pas à son maximum d'un côté de l'interface s'il est à son maximum de l'autre côté). Les fronts d'onde doivent donc être représentés sur le schéma de sorte qu'ils se recoupent sur l'interface verres-air de la lentille. On constate, par simple inspection des triangles rectangles apparaissant dans ce schéma, que cette condition de « *continuité* » des fronts d'onde impose l'égalité $\lambda_t / \sin \theta_t = \lambda_i / \sin \theta_i$. Sachant que la longueur d'onde dans les milieux diélectriques est donnée par la longueur d'onde dans le vide divisée par l'indice de réfraction, $\lambda_i = \lambda_0 / n_i$ et $\lambda_t = \lambda_0 / n_t$, la condition de continuité des fronts d'onde devient $\lambda_0 / n_t \sin \theta_t = \lambda_0 / n_i \sin \theta_i$, ce qui conduit immédiatement à la loi de Snell-Descartes.



Remarque : Notez qu'en passant d'un milieu d'indice élevé vers un milieu d'indice plus faible l'angle de transmission θ_t est toujours plus grand que l'angle incident θ_i . On voit donc qu'il existe un angle incident critique θ_{ic} pour lequel l'angle de transmission est égal à 90° , soit $n_i \sin \theta_{ic} = n_t$. Bien sûr l'angle de transmission ne pourra pas être plus grand que 90° alors que rien n'empêche de choisir un angle incident plus grand que l'angle critique $\theta_{ic} = \arcsin(n_t / n_i)$. On peut montrer dans ce cas que l'onde n'est plus transmise par l'interface mais qu'elle est totalement réfléchie. C'est ce que l'on appelle le phénomène de « *réflexion interne totale* ».

Une lentille convergente est une lentille pour laquelle les rayons lumineux parallèles à son axe optique convergent tous vers un même point appelé le foyer de la lentille. Pour cela il faut bien sûr que la lentille ait une forme bien précise. L'optique géométrique ne permet pas facilement de savoir quelle forme donner à la lentille pour obtenir un foyer, c'est pourquoi nous allons ici faire appel à l'optique ondulatoire.

4.2 Lame de verre d'épaisseur constante

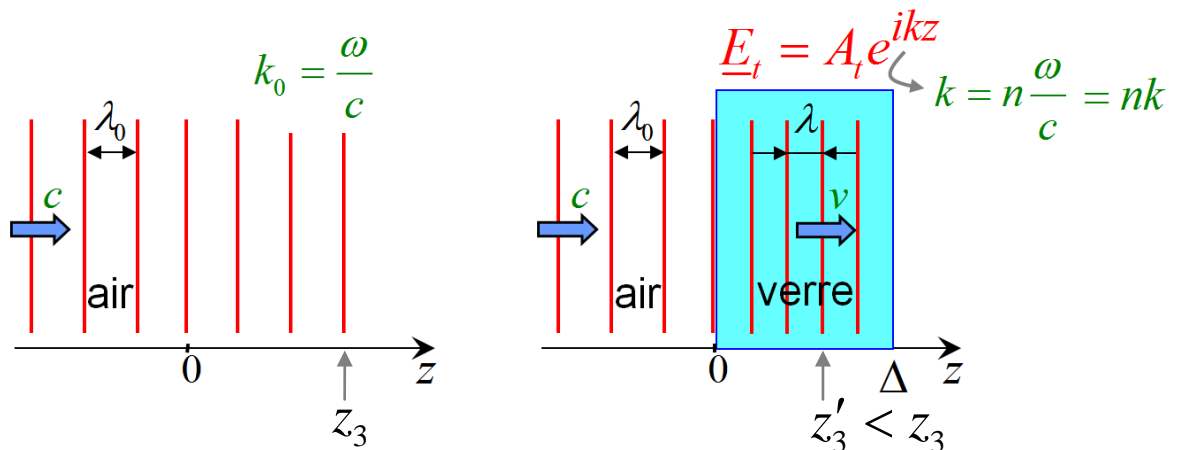
Pour décrire les lentilles, commençons d'abord par voir l'effet d'une lame de verre d'épaisseur constante Δ sur la propagation d'une onde plane. Pour cela je définis tout d'abord ce que l'on appellera ici la phase de l'onde : la phase d'une onde sera simplement considérée comme étant l'argument de son phaseur. Ainsi, pour une onde plane harmonique se propageant le long de l'axe z la phase de l'onde de champ électrique est simplement $\varphi = kz$.

$$E(x, z, t) = A \cos[kz - \omega t] = \text{Re}[A e^{ikz} e^{-i\omega t}]$$

"phase"

$$\underline{E} = A e^{ikz} = A_i e^{i\varphi}$$

Les deux schémas ci-dessous illustrent l'effet de la lame de verre. Sans lame, à gauche, la longueur d'onde reste constante et la phase en tout point z vaut simplement $\varphi_0 = k_0 z$ où k_0 est le nombre d'onde dans l'air (ou dans le vide), soit $k_0 = 2\pi / \lambda_0 = \omega / c$.



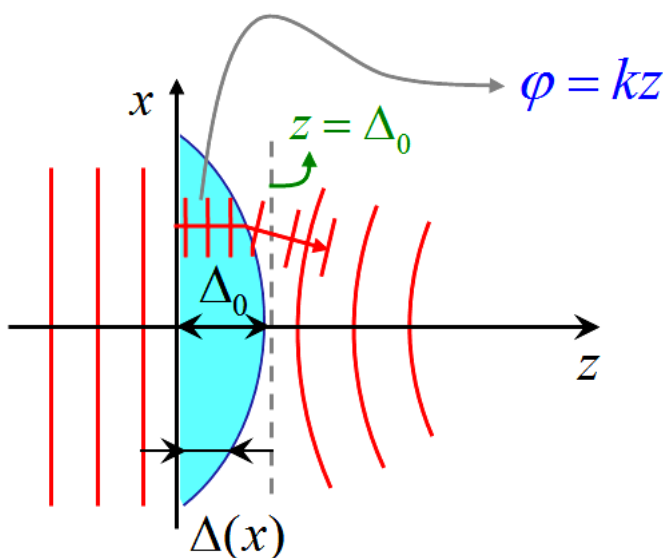
Considérons maintenant le schéma de droite ci-dessus. On y voit une lame de verre et les fronts d'onde de l'onde lumineuse s'y propageant. Dans le verre le nombre d'onde est plus grand que le nombre d'onde $k_0 = \omega / c$ du vide (ou de l'air) car il vaut $k = n\omega / c$ où l'indice de réfraction n est toujours supérieur à l'unité. Le fait que k est plus grand que k_0 se manifeste par une longueur d'onde plus petite dans le verre que dans l'air (les fronts d'onde sont plus rapprochés dans le verre que dans l'air, comme on le voit sur le schéma de droite ci-dessus). Si les fronts d'onde sont plus rapprochés dans le verre cela signifie que l'onde est ralentie, ce que l'on exprime en disant que les fronts d'onde sont retardés. En effet, en présence de la lame de verre un front d'onde donné à un instant donné se situera en une valeur de z plus petite qu'en l'absence de la lame (ceci est illustré sur les deux schémas ci-dessus par les positions z_3 et z'_3 du troisième front d'onde après l'origine).

Ce retard des fronts d'onde dans le verre s'exprime facilement au travers de la notion de phase accumulée dans la lame de verre. Par définition, la phase accumulée sur une distance donnée vaut le produit de cette distance par le nombre d'onde. La phase accumulée sur l'épaisseur Δ de la lame vaut donc $\varphi_l = k\Delta = nk_0\Delta$ alors que dans le vide, sur la même épaisseur, on avait $\varphi_0 = k_0\Delta$. Le retard des fronts d'onde se reflète donc dans une valeur plus grande de la phase accumulée sur une distance donnée. Lors de la mise en place de la lame, la phase de l'onde en $z = \Delta$ a donc changé de la valeur $\varphi_l - \varphi_0 = k_0\Delta(n-1)$. On dit que la lame a provoqué un « déphasage ». On peut donc voir une lame de verre plane comme étant un dispositif qui provoque un déphasage sur une onde lumineuse. Ce déphasage est proportionnel à l'épaisseur de verre Δ et à $(n-1)$.

$$\text{déphasage : } \varphi_l - \varphi_0 = k_0\Delta(n-1)$$

4.3 Lame de verre d'épaisseur variable

Considérons maintenant la lentille plan-convexe analysée plus haut et représentée schématiquement ci-dessous. C'est pour la simplicité du développement que nous considérons ici une lame dont la face d'entrée est plane mais les résultats obtenus peuvent facilement se généraliser à toute forme de lentille. De même nous ne considérerons que deux dimensions de l'espace x et z , en considérant que la lame est d'épaisseur constante en y .



Dans l'épaisseur du verre la phase de l'onde est proportionnelle à z et à k , c'est à dire que les fronts d'onde sont plus rapprochés que dans le vide. En d'autres termes, ils sont retardés par rapport à ceux qui se propagent dans le vide.

Comme l'épaisseur $\Delta(x)$ est variable en x , le retard accumulé sur l'épaisseur de la lentille est aussi variable. Là où l'épaisseur de verre est grande le retard du front d'onde est grand et vice-versa ; il s'en suit une courbure des fronts d'onde. Comme suggéré sur le schéma ci-dessus, on constate que ceci donne la possibilité de courber des fronts d'onde, en particulier, on voit la possibilité de transformer une onde plane en une onde sphérique. C'est le principe de fonctionnement d'une lentille.

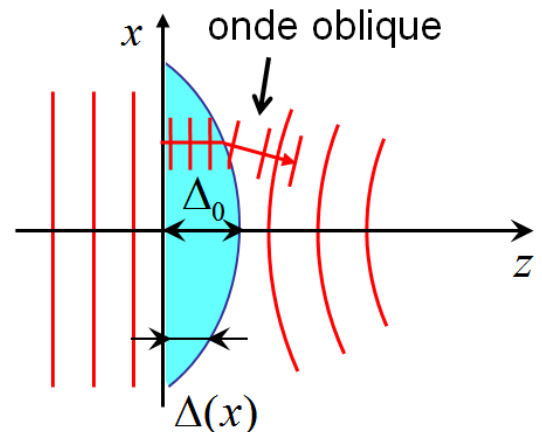
Pour pouvoir calculer la forme à donner à une lentille pour générer une onde sphérique convergente en un point, il nous faut exprimer le retard de phase accumulé par l'onde en toute valeur de la coordonnée transverse x . A cette fin, pour le calcul du retard de phase nous prenons comme ordonnée de référence $z = \Delta_0$ [Δ_0 étant l'épaisseur de la lentille en $x = 0$, soit $\Delta_0 = \Delta(0)$]. Le plan $z = \Delta_0$ s'appelle le plan de sortie de la lentille. En l'absence de lame le déphasage en $z = \Delta_0$ vaut $k_0 \Delta_0$.

sans lame : $\varphi_0 = k_0 \Delta_0$

En présence de la lame, le déphasage accumulé en $z = \Delta_0$ résulte de la somme du déphasage accumulé sur l'épaisseur $\Delta(x)$ de verre traversé $\Delta(x)$, et du déphasage accumulé sur la distance qu'il reste à parcourir dans le vide (ou l'air) pour arriver en $z = \Delta_0$, soit sur la distance $\Delta_0 - \Delta(x)$. Comme l'épaisseur $\Delta(x)$ correspond à une propagation dans un milieu d'indice de réfraction n , on trouve pour la phase accumulée dans le verre l'expression $nk_0 \Delta(x)$ alors que pour la distance parcourue dans le vide, on a la phase accumulée $k_0 [\Delta_0 - \Delta(x)]$ car l'indice du vide (ou de l'air) vaut 1.

avec lame : $\varphi_l = nk_0 \Delta(x) + k_0 [\Delta_0 - \Delta(x)]$


Il faut noter toutefois que le déphasage dans l'air $k_0 [\Delta_0 - \Delta(x)]$ est celui d'une onde d'incidence oblique (voir schéma ci-contre), on pourrait donc croire a priori que le résultat ci-dessus n'est pas correct car il ne considère pas l'inclinaison de l'onde, le calcul étant fait comme si l'onde se propageait parallèlement à l'axe z . Pour montrer que ce résultat est néanmoins valable, décrivons l'onde en la décomposant localement en portions d'onde plane d'angle variable en x (voir schéma), ce qui permet d'écrire le phaseur de l'onde (en une valeur de x donnée) comme suit :



$$A_x e^{ik \sin(\theta)x} e^{ik \cos(\theta)z}$$

Où θ est l'angle que fait l'onde par rapport à l'axe z . Si on adopte l'approximation des petits angles on peut écrire $\cos \theta \approx 1$, ce qui conduit immédiatement à la conclusion que la phase accumulée sur une longueur z parallèlement à l'axe optique (axe de symétrie de la lame) vaut kz et est donc bien indépendante de l'angle θ . La phase accumulée dans l'air à la sortie de la lentille vaut donc bien $k_0 [\Delta_0 - \Delta(x)]$, conformément à l'expression de la phase φ_l donnée ci-dessus.

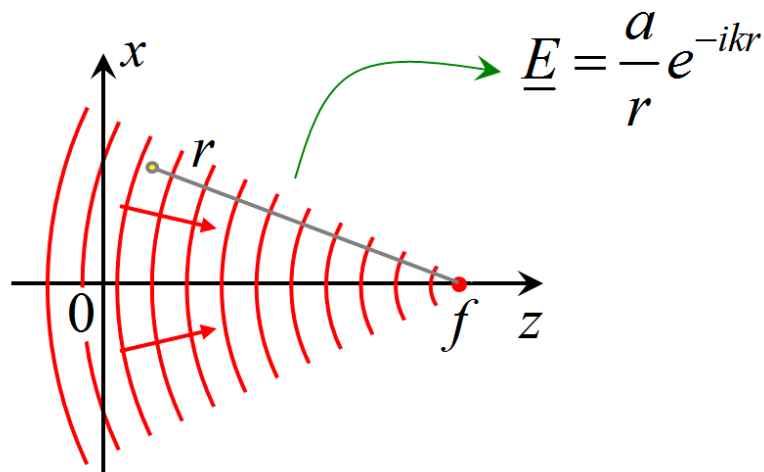
En conclusion, le déphasage introduit par une lame d'épaisseur variable à une distance donnée x de l'axe optique vaut $nk_0\Delta(x) + k_0[\Delta_0 - \Delta(x)] - k_0\Delta_0$, soit l'expression reprise ci-dessous :

 **déphasage total en x : $\varphi(x) = \varphi_l - \varphi_0 = k_0\Delta(x)(n-1)$**

Notez que cette expression est identique à celle obtenue plus haut pour une lame d'épaisseur constante, à la différence près qu'ici Δ est une fonction de x .

4.4 Onde sphérique convergente

Nous allons ici calculer la forme $\Delta(x)$ à donner à une lentille pour qu'elle focalise la lumière d'une onde plane incidente en un point donné de l'axe de la lentille. Si l'onde générée par passage au travers de la lentille se réduit à un point sur l'axe de la lentille, c'est qu'il s'agit bien sûr d'une onde sphérique convergente. Pour calculer la forme de la lentille qui génère une telle onde il est utile d'étudier tout d'abord la phase d'une onde sphérique convergente. Nous avons vu précédemment que le phaseur de l'onde sphérique est $(a/r)\exp(ikr)$ mais cette expression correspondait à une onde divergente. En effet, si on fait apparaître la dépendance temporelle en écrivant le champ à partir de son phaseur, soit $E = \text{Re}[(a/r)\exp(ikr - i\omega t)]$, on trouve l'expression $(a/r)\cos(kr - \omega t)$ qui correspond bien à une onde se propageant dans la direction des valeurs croissantes de r . Pour avoir une onde se propageant vers les valeurs décroissantes de r il convient d'inverser le signe du nombre d'onde, comme pour une onde plane. En effet, le phaseur $(a/r)\exp(-ikr)$ correspond au champ $(a/r)\cos(kr + \omega t)$ qui représente bien une onde se propageant dans la direction des valeurs décroissantes de r puisque le lieu des points de phase constante est donné par l'équation $kr + \omega t = \text{cste}$. Il s'agit donc bien du phaseur d'une onde sphérique convergente telle qu'illustrée sur le schéma ci-dessous.



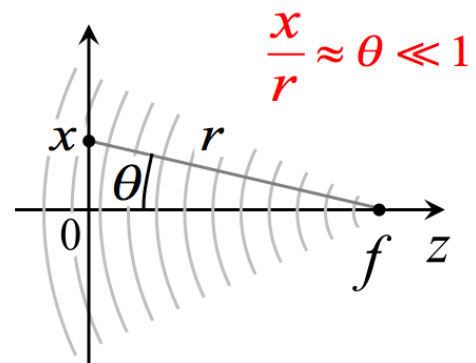
Étudions maintenant la distribution de phase $\varphi = -kr$ de l'onde sphérique convergente centrée sur l'axe z en $z = f$ (voir schéma ci-dessus). Pour obtenir la distribution de phase $\varphi = -kr$ dans l'espace il suffit de remplacer r par son expression en fonction des coordonnées x, z (je rappelle que pour la simplicité, on ne prend pas en compte la coordonnée y , nous décrivons donc en réalité une onde « cylindrique » invariante par translation en y et c'est donc implicitement une lentille dont l'épaisseur est constante en y que nous considérerons). On trouve :

$$r^2 = (z - f)^2 + x^2$$

Intéressons nous maintenant à la distribution de phase $\varphi = -kr$ dans le plan $z = 0$. Pour cela il suffit de considérer les valeurs de r obtenue pour $z = 0$, ce qui donne :

En $z = 0$ on a :
$$r = \sqrt{f^2 + x^2} = f \sqrt{1 + x^2 / f^2}$$

Pour simplifier cette expression nous considérons l'hypothèse des petits angles introduites plus haut pour pouvoir négliger l'effet de l'angle de réfraction non nul à la sortie de la lentille. Dans la limites des petits angles il est évident que la description de l'onde sphérique convergente peut se limiter à des valeurs de x très petites par rapport à la distance f (le schéma ci-contre illustre bien cette situation).



Nous pouvons donc approcher la racine ci-dessus au premier ordre en x^2 / f^2 , ce qui donne :

$$r \approx f \left(1 + \frac{x^2}{2f^2} \right) = f + \frac{x^2}{2f}$$

En remplaçant cette expression de r dans le phaseur $\underline{E} = (a / r) \exp(-ikr)$ nous obtenons l'expression du phaseur de l'onde lumineuse en tout point du plan $z = 0$. On constate que la phase évolue en x de façon parabolique (phase proportionnelle au carré de x).

$$\underline{E}(x, z = 0) = \frac{a}{r} e^{-ik_0 \left(f + \frac{x^2}{2f} \right)}$$


Cette dernière expression indique que la distribution de phase dans le plan $z=0$ de l'onde sphérique convergente centrée en $z=f$ est donnée par la fonction $-k_0[f + x^2 / (2f)]$. Or, cette onde est précisément celle que l'on cherche à générer avec la lentille. En conséquence, si l'on donne à la lentille un profil $\Delta(x)$ tel que le déphasage $\varphi(x) = k_0\Delta(x)[n-1]$ introduit par cette lentille reproduit la fonction $-k_0[f + x^2 / (2f)]$ alors cette lentille générera une onde sphérique convergente centrée en $z=f$. En effet, si la lentille génère dans son plan de sortie la distribution de phase de l'onde sphérique alors l'onde se propagera après la lentille comme cette onde sphérique, c'est-à-dire qu'elle sera convergente en un point de l'axe situé à une distance f du plan de sortie de la lentille. De la sorte, la distance f apparaîtra naturellement comme la distance focale de la lentille.

4.5 Lentille convergente

En résumé, si on égale la fonction $\varphi(x) = k_0\Delta(x)[n-1]$ à la distribution de phase de l'onde sphérique, soit $-k_0[f + x^2 / (2f)]$, on trouve la fonction $\Delta(x)$ donnant l'épaisseur de la lentille qui génère la même onde sphérique. Ceci revient à dire que l'on aura la forme à donner à la lentille pour qu'elle génère une onde sphérique convergente. En pratique les choses sont un peu plus compliquées que cela car l'égalité des deux fonctions conduit à une difficulté d'interprétation physique. En effet, considérons le déphasage sur l'axe de la lentille $\varphi(0) = k_0\Delta_0[n-1]$. Cette valeur du déphasage est incompatible avec la distribution de phase de l'onde sphérique qui elle vaut $\varphi(0) = -k_0f$. On constate en effet que pour avoir égalité entre ces deux phases il faudrait que l'épaisseur de la lentille soit négative en $x=0$, soit $\Delta_0 < 0$. En réalité, cette difficulté n'est qu'apparente car une distribution de phase est une fonction qui peut toujours être définie à une constante près. En effet, si on reprend l'onde sphérique convergente considérée plus haut, soit $(a/r)\cos(kr + \omega t)$, on voit facilement qu'on peut la réécrire avec un déphasage quelconque sans pour autant changer la forme de l'onde. En d'autres termes l'expression $(a/r)\cos(kr + \omega t + \varphi_0)$ représente la même onde sphérique que l'expression $(a/r)\cos(kr + \omega t)$ (la seule différence entre les deux ondes est que leurs fronts d'onde sont translatés en r de la distance φ_0/k). Autrement dit, l'égalité entre le déphasage provoqué par la lentille et la distribution de phase de l'onde sphérique peut être exprimée à une constante près, comme écrit ci-dessous.

$$\varphi(x) = k\Delta(x)(n-1) = -k\left(f + \frac{x^2}{2f}\right) + \varphi_0$$

Cette égalité nous donne immédiatement la fonction recherchée $\Delta(x)$:



$$\Delta(x) = \frac{1}{n-1} \left[-f - \frac{x^2}{2f} + \frac{\varphi_0}{k} \right]$$

La constante φ_0 qu'il convient de choisir dans l'expression ci-dessus est fixée par la valeur de l'épaisseur en $x = 0$, soit :

$$\Delta_0 = \frac{1}{n-1} \left[-f + \frac{\varphi_0}{k} \right]$$

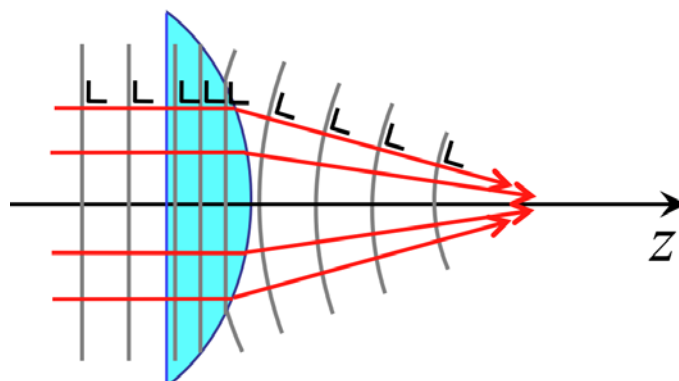
Ce qui conduit à la forme finale du profil de la lentille $\Delta(x)$.

$$\Delta(x) = \Delta_0 - \frac{1}{(n-1)} \frac{x^2}{2f}$$

surface parabolique

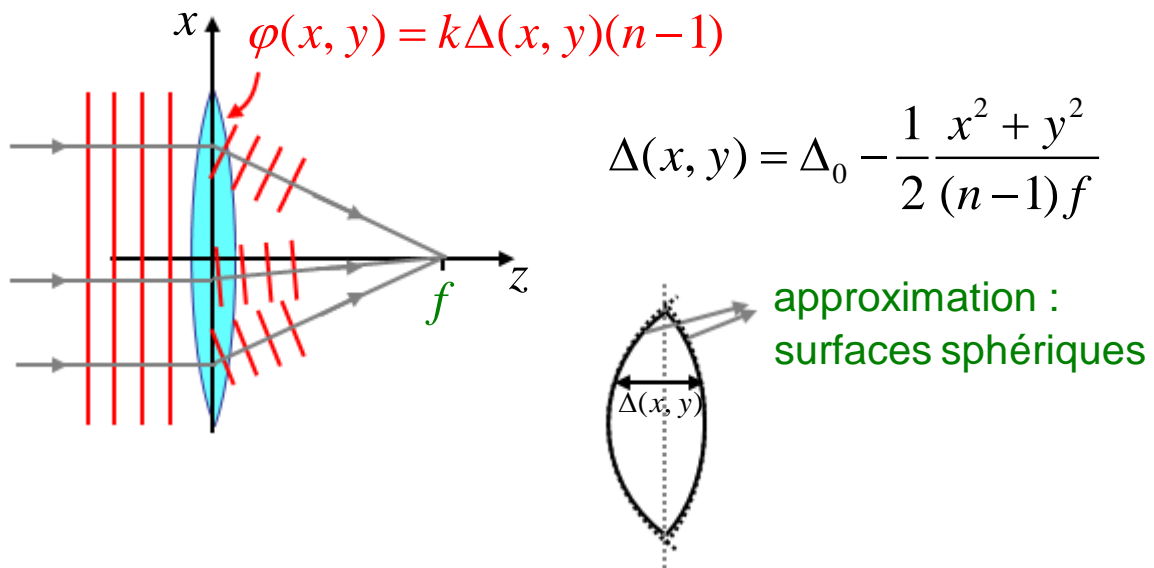
On constate de cette façon que le profil d'une lentille convergente est parabolique. On voit également que plus la distance focale f est petite, plus forte doit être la courbure du profil parabolique. De même, on voit que, pour une distance focale donnée, la courbure est d'autant plus faible que l'indice de réfraction est grand. Ces constations se comprennent aisément en considérant le principe de fonctionnement de la lentille vu plus haut (courbure des fronts d'onde d'une onde plane par déphasage variable en x).

Remarque : Ce résultat peut être obtenu dans le cadre de l'optique géométrique sur base d'une analyse de la géométrie de la trajectoire des rayons. Notez à ce propos que les rayons de l'optique géométrique sont définis comme étant les lignes partout perpendiculaires aux fronts d'onde de la théorie de l'optique ondulatoire développée ici (voir schéma ci-dessous). Une onde sphérique apparaît donc comme un ensemble de rayons convergents vers un point unique.



4.6 Généralisation à trois dimensions

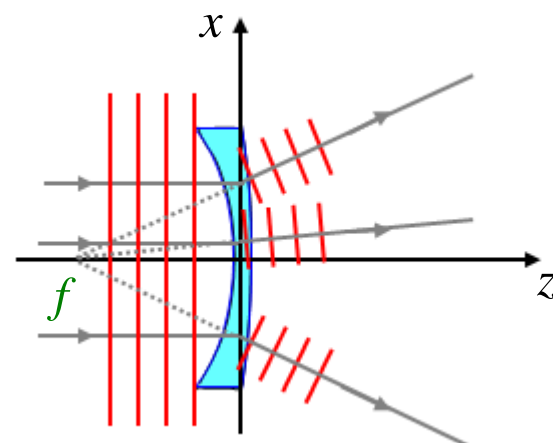
Le raisonnement fait ci-dessus peut facilement être généralisé à 3 dimensions, c'est-à-dire, en y incluant la variation d'épaisseur de la lame dans la direction y . Par symétrie, on peut directement conclure que si l'on veut la même distance de focalisation f pour les rayons passant sur l'axe y , l'épaisseur en y doit varier de manière parabolique avec les mêmes coefficients que la variation en x . L'épaisseur de la lentille est donc la suivante : $\Delta(x, y) = \Delta_0 - (x^2 + y^2) / [2(n-1)f]$. On remarquera qu'elle présente une symétrie de révolution par rapport à l'axe z . La surface de la lentille est donc un paraboloïde à symétrie de révolution.



En pratique, pour des raisons technologiques, les surfaces des lentilles sont la plupart du temps sphériques. Mais pour autant que les lentilles soient minces, les surfaces sphériques constituent une excellente approximation des surfaces paraboliques.

4.7 Lentilles divergentes

L'approche théorique proposée plus haut peut être appliquée au cas où la coordonnée $z = f$ est négative. Les calculs sont identiques, il suffit d'y considérer que f est



négative. L'interprétation de la distance focale négative est immédiate, elle correspond à une lentille dont l'épaisseur augmente paraboliquement en s'éloignant du centre et les rayons ne croisent l'axe z que dans leur prolongement vers les z négatifs (voir lignes pointillées). Il s'agit donc d'une lentille divergente.

$$\Delta(x, y) = \Delta_0 - \frac{1}{2} \frac{x^2 + y^2}{(n-1)f}$$

où $f < 0$

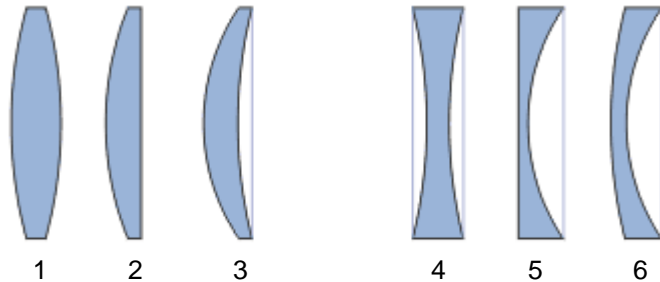
Pour des raisons techniques, les lentilles convergentes et divergentes peuvent avoir des formes différentes selon la concavité de leurs faces paraboliques, comme le montre la figure ci-dessous. Mais dans tous les cas, les lentilles convergentes ont une épaisseur qui diminue en s'éloignant de leur centre et les lentilles divergentes ont une épaisseur qui augmente en s'éloignant de leur centre.

Lentilles convergentes

1. lentille biconvexe
2. lentille plan-convexe
3. ménisque convergent

Lentilles divergentes

4. lentille biconcave
5. lentille plan-concave
6. ménisque divergent



4.8 Formation d'images

4.8.1 L'appareil photographique

Une des applications les plus répandues des lentilles est la formation d'images (mais elles ont beaucoup d'autres applications également : voir cours d'Optique Physique de 3^{ème} année section Physique). Une seule lentille convergente peut, par exemple, être utilisée pour réaliser l'image réelle d'un objet, comme l'indique le schéma ci-dessous. Le terme « *image réelle* » peut être expliqué comme suit : un objet visible est un objet qui diffuse les rayons de la lumière qui lui arrive d'une source lumineuse (elle-même diffuse ou non). Chaque point de l'objet (tel que la pointe de la flèche « *objet* » sur le schéma ci-contre où la lentille est représentée de façon conventionnelle par une double flèche) diffuse les rayons dans toutes les directions. Le rayon diffusé, qui est parallèle à l'axe optique z de la lentille, passe par le foyer $z = f$. Le rayon qui passe par le centre de la lentille n'est pas défléchi car la lentille est supposée mince et la réfraction est négligeable. Ces deux rayons se croisent donc après la lentille en un point situé au-delà de la distance focale f . On peut montrer par simple développement de géométrie élémentaire qu'en ce même point au-delà de la lentille se croisent tous les rayons diffusés par le point de l'objet considéré (ici la pointe de la flèche « *objet* »). A titre d'exemple, le rayon qui passe par $z = -f$, est parallèle à l'axe z après passage par la lentille et passe donc bien par le même point que les deux premiers rayons considérés.

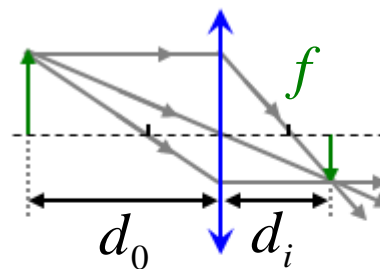


image "réelle"

$$\frac{1}{d_0} + \frac{1}{d_i} = \frac{1}{f}$$

$$M = \frac{d_i}{d_0}$$



De la sorte, la configuration de rayons diffusés par l'objet est reproduite (à un facteur d'agrandissement et un renversement près) après la lentille, exactement comme si l'objet y était reproduit en plus grand ou plus petit et renversé. C'est pourquoi on parle dans ce cas d'image « réelle ». Du point de vue des ondes électromagnétiques, tout se passe comme si l'objet était « réellement » placé après la lentille. Un calcul élémentaire de géométrie montre que la distance d_i à laquelle se reforme la configuration de rayons diffusés de l'objet est déterminée par la distance d_0 entre l'objet et la lentille et la distance focale f de la lentille, au travers de la relation $1/d_0 + 1/d_i = 1/f$. Cette relation s'appelle la « *formule de conjugaison* » des lentilles. Remarquez que le facteur d'agrandissement de l'image est donné de façon évidente par le rapport $M = d_i/d_0$ (voir schéma ci-dessus).

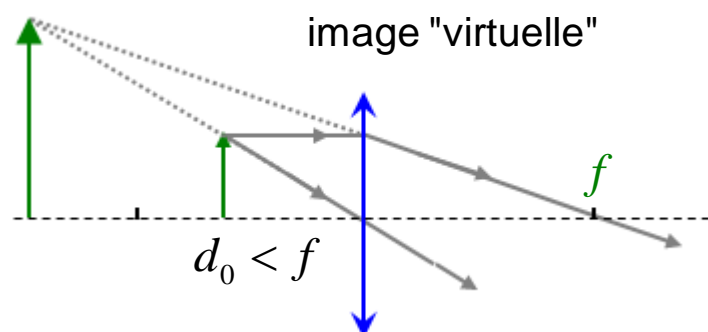
Remarque : la formule de conjugaison des lentilles $1/d_0 + 1/d_i = 1/f$ est donnée ici en considérant que toutes les distances sont positives. Il existe d'autres formes de cette formule qui sont basées sur des conventions sur les signes des distances image, objet et focale. Il est souvent considéré qu'une distance mesurée avant la lentille est négative et les distances mesurées après la lentilles sont positives. Il en résulte la forme suivante de la formule de conjugaison $-1/d_0 + 1/d_i = 1/f$ pour une lentille convergente (voir séances de laboratoire).

Ce principe de formation d'images réelles est utilisé, par exemple, en photographie. La pellicule photographique est placée à la distance d_i , de façon à ce que l'objet vienne s'y reproduire. Les variations d'intensité lumineuse de l'objet sont alors enregistrées sur la cellule photosensible pour être reproduite ensuite sur un écran d'ordinateur ou sur papier par des procédés divers.

4.8.2 La loupe

Un autre exemple d'application des lentilles convergentes est la loupe. Pour utiliser une lentille convergente en tant que loupe, il faut placer l'objet par rapport à la lentille à une distance plus petite que sa distance focale ($d_0 < f$). Le rayon diffusé par l'objet qui est

parallèle à l'axe optique passe par le foyer au-delà de la lentille (en $z = f$) et celui qui passe par le centre de la lentille n'est pas dévié. Au-delà de la lentille, les rayons sont donc divergents et l'observateur les « verra » comme s'ils provenaient d'un objet plus lointain mais surtout plus grand, comme indiqué par le prolongement pointillé des rayons sur le schéma ci-contre.



la loupe

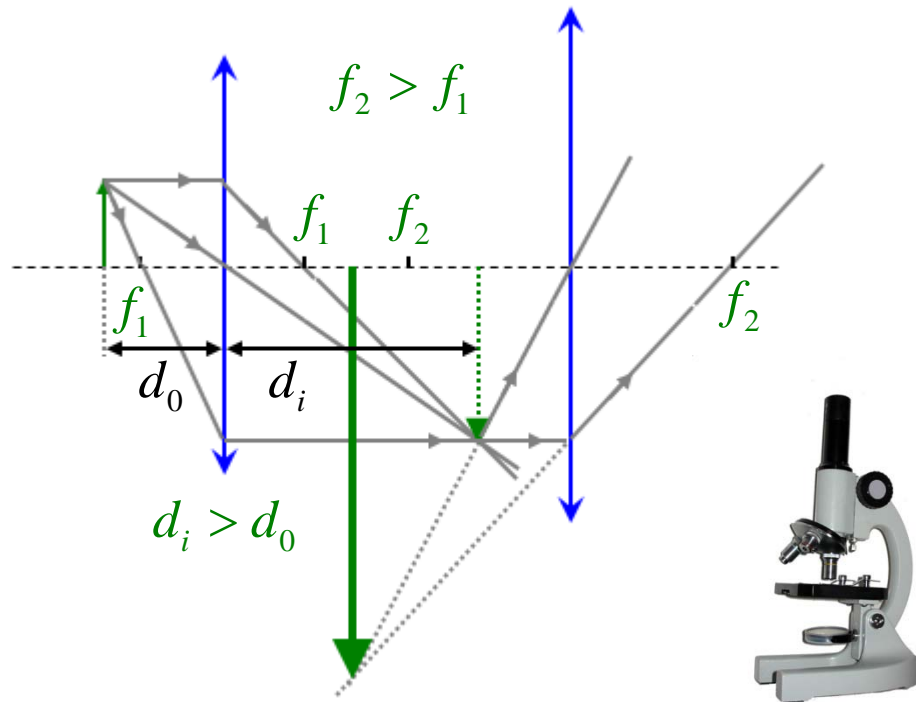
Dans cette situation, la configuration des rayons diffusés par l'objet n'est reproduite nulle part dans l'espace. Si on voit l'objet plus grand, c'est par pur effet d'optique lié au prolongement des rayons vers l'arrière (exactement comme l'effet de cassure provoqué par la réfraction sur des objets à moitié immergés dans l'eau). C'est pourquoi, on dit qu'il s'agit d'une image « virtuelle ».



Remarque : Notez que si l'objet est placé exactement à la distance focale de la loupe, l'image virtuelle se forme à l'infini (voir schéma ci-dessus). Les rayons qui apparaissent à l'observateur sont dans ce cas parallèles. C'est le mode d'utilisation le plus naturel de la lentille car si l'image virtuelle est à l'infini, l'œil de l'observateur ne doit en principe faire aucun effort d'accommodation. Remarquez que même si l'image est à l'infini, elle apparaîtra plus grande à l'observateur car les rayons proviennent d'un angle plus grand qu'en l'absence de loupe. En pratique le facteur de grossissement d'une loupe s'exprime en grossissement angulaire. Ce concept sera vu en séance d'exercices.

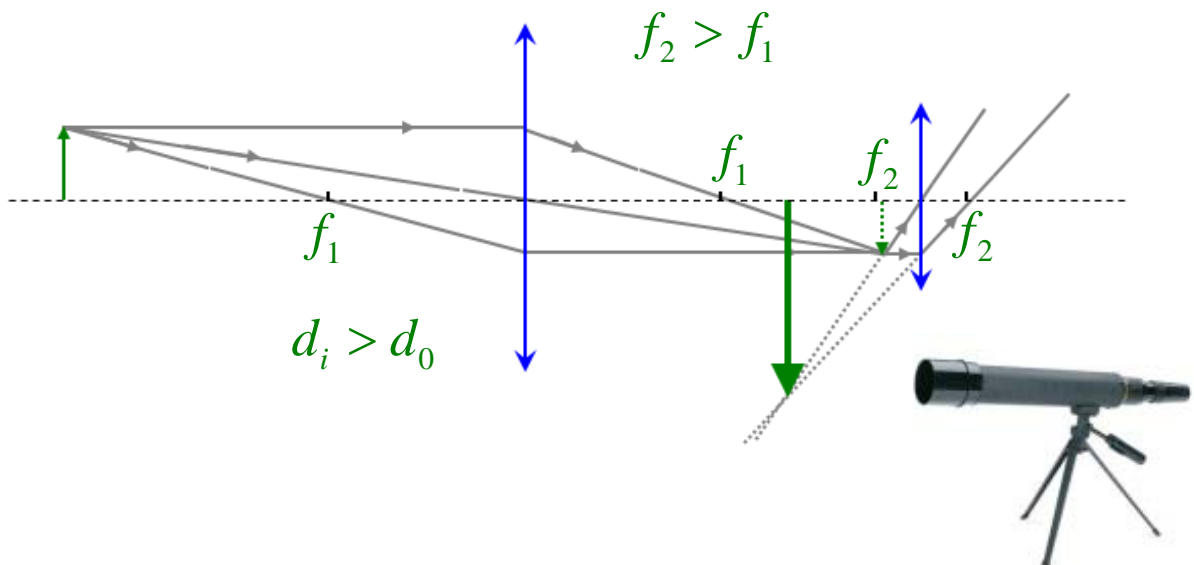
4.8.3 Le microscope

Les lentilles peuvent également être combinées pour obtenir d'autres effets intéressants du point de vue des applications. A titre d'exemple, deux lentilles convergentes de distances focales différentes peuvent être combinées pour obtenir l'image agrandie d'objets de petites tailles. C'est le principe du microscope représenté schématiquement ci-dessous. Une première lentille de petite distance focale f_1 est placée à une distance $d_0 > f_1$ de l'objet à observer. Comme la distance d_0 est plus grande que la distance focale, la formule de conjugaison des lentilles indique que la distance « image » $d_i = 1/(1/f_1 - 1/d_0)$ est positive ce qui signifie que l'image est réelle (voir l'appareil photographique ci-dessus). Si d_0 est néanmoins proche de f_1 , la distance $d_i = 1/(1/f_1 - 1/d_0)$ est grande et le facteur d'agrandissement $M = d_i/d_0$ est grand. La seconde lentille de plus grande distance focale f_2 est utilisée en mode « loupe » pour encore agrandir cette image réelle. Pour cela, il faut que la nouvelle distance objet d_0 soit plus petite que la distance focale f_2 , comme indiqué sur le schéma.



4.8.4 La lunette, les jumelles

Un autre exemple de système de lentilles convergentes est la lunette (ou les jumelles) qui sert à agrandir des images lointaines. Le principe est à peu près le même que celui du microscope. La différence est que la première lentille est de grande distance focale f_1 et la seconde est de courte distance focale f_2 . Comme l'objet est très éloigné, la distance d_0 est toujours beaucoup plus grande que f_1 et, dès lors, la première lentille forme une image réelle de petite taille car $d_i = 1/(1/f_1 - 1/d_0)$ est à peu près égale à f_1 et le facteur d'agrandissement $M = f_1/d_0$ est donc très petit. L'image réelle est formée près de la lentille de courte focale f_2 , de sorte que l'effet loupe en donne une image virtuelle plus grande.



5. La relativité restreinte

Nous terminerons ce chapitre consacré à la physique moderne avec une introduction à la relativité restreinte. Comme nous allons le voir, la théorie de la relativité restreinte bouleverse l'image que nous nous forçons depuis la naissance des concepts les plus fondamentaux de notre cadre de vie, à savoir, l'espace et le temps. Au-delà de son impact philosophique considérable, cette théorie a conduit à une amélioration profonde de la compréhension de nombreux phénomènes physiques ainsi qu'à une multitude d'avancées scientifiques et technologiques. C'est la raison pour laquelle il est essentiel de l'aborder dans les études d'ingénieurs. L'approche pédagogique poursuivie ici vise essentiellement à donner, de façon aussi efficace que possible, une image mentale de ce qu'est notre univers selon la théorie de la relativité restreinte. Grâce à cette approche imagée il est possible d'avoir une compréhension intuitive de cette théorie malgré le très haut niveau d'abstraction auquel elle nous confronte.

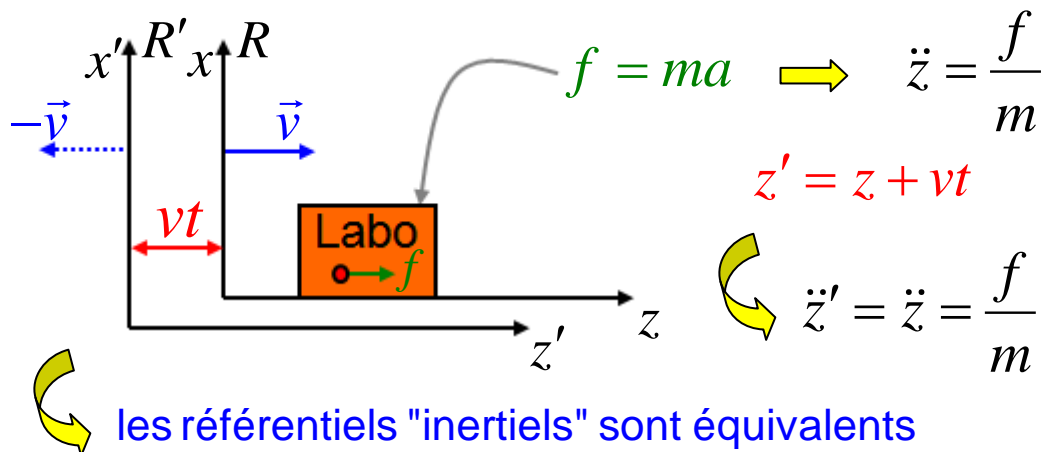
Nous commencerons par étudier la démarche intellectuelle poursuivie par les scientifiques de la fin du 19^{ème} siècle et du début du 20^{ème} siècle pour aboutir à la notion de relativité restreinte. A cette fin, je commence par rappeler le principe de relativité de Galilée (il serait sans doute plus juste de dire qu'il s'agit du principe de Galilée et Newton car si Galilée a formulé les premières notions de relativité, c'est Newton qui les a concrétisées et précisées au travers de ses lois fondamentales de la mécanique).

5.1 Principe de relativité de Galilée

Le principe de relativité de Galilée stipule que tous les référentiels inertiels sont équivalents. Ceci est illustré sur le schéma ci-dessous où on voit un laboratoire dans lequel des expériences sont effectuées. Comme tout phénomène physique a pour origine les interactions fondamentales de la nature (l'interaction gravitationnelle, l'interaction électrofaible, et l'interaction forte), je représente schématiquement l'expérience par une particule élémentaire subissant une force f . Les systèmes réels sont en général composés d'un très grand nombre de telles particules en interaction. L'expérience de physique mènera à des observations qui sont déterminées par les forces f sur chacune des particules. Le mouvement des particules est régi par la loi de Newton $f = ma$. Si pour simplifier la discussion on ne considère que des déplacements parallèles à l'axe z , on peut écrire cette loi comme suit : $\ddot{z} = f/m$. Toutes les particules impliquées dans l'expérience de physique obéissent à cette loi pour donner lieu au phénomène étudié. Les positions et les mouvements des particules sont repérés par l'axe z du référentiel R dans lequel le laboratoire est fixe (au repos).

Supposons maintenant que l'expérience est observée d'un référentiel R' dans lequel le référentiel R est en mouvement à la vitesse \vec{v} constante parallèle à l'axe z . Sur un temps t , le référentiel R avance de la distance vt dans le référentiel R' . Les positions des particules de l'expérience dans le référentiel R' sont donc $z' = z + vt$. La loi du mouvement des particules dans R' peut être obtenue en dérivant deux fois z' . On trouve $\ddot{z}' = \ddot{z} = f/m$, c'est-à-dire la même loi du mouvement que dans le référentiel R .

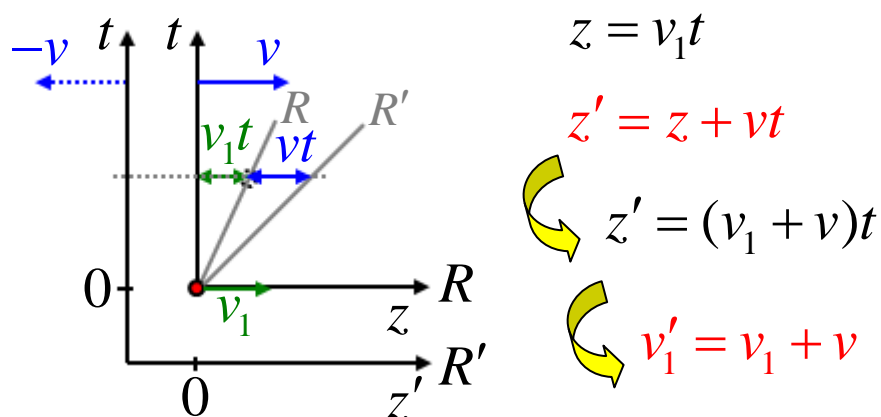
En conclusion, les lois du mouvement sont les mêmes dans les deux référentiels en mouvement uniforme l'un par rapport à l'autre. On dit que les référentiels « *inertiels* » sont équivalents.



Notez que l'on peut voir la situation schématisée ci-dessus en considérant que c'est le référentiel R qui est au repos et le référentiel R' qui est en mouvement à la vitesse $-\vec{v}$. Il y a « *relativité* » du mouvement uniforme. C'est la raison de la présence de la flèche pointillée de vitesse $-\vec{v}$ dirigée vers la gauche sur le schéma ci-dessus.

5.2 Formule d'addition des vitesses

Au principe de relativité galiléenne est associée la formule d'addition des vitesses. Cette formule exprime simplement le fait que les vitesses des particules dans le référentiel R' sont données par les vitesses des particules dans le référentiel R augmentées de la vitesse relative v des deux référentiels. Pour une particule de vitesse v_1 dans R , on a donc la vitesse $v'_1 = v_1 + v$ dans R' .



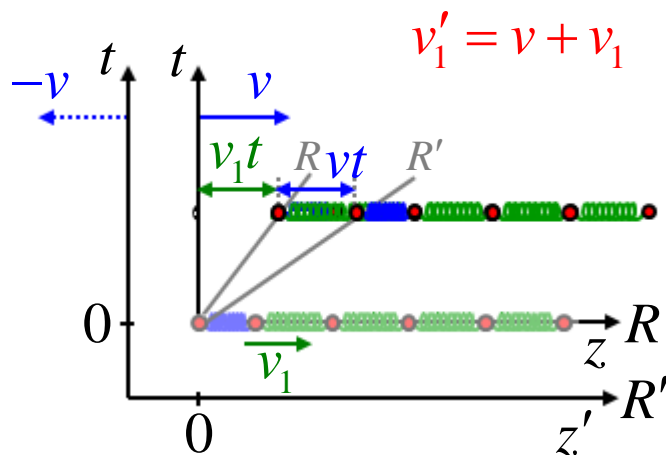
La particule est représentée sur le schéma ci-dessus dans des diagrammes d'espace-temps. L'évolution de la particule dans le référentiel R est représentée graphiquement par sa coordonnée $z = v_1 t$ en fonction du temps t . L'évolution dans ce diagramme est donc une simple droite de pente v_1 si la vitesse v_1 est constante. Cette droite est représentée en gris et est identifiée par la lettre R sur le schéma ci-dessus.

On peut faire la même représentation de l'évolution de la particule dans R' . On a alors une droite de pente $v_1 + v$ dans le diagramme d'espace-temps. C'est la ligne droite grise identifiée par la lettre R' sur le schéma ci-dessus.

5.3 Ondes et relativité

Jusqu'à présent il n'y a pas de difficulté conceptuelle liée à ce principe de relativité et tout pourrait même sembler trivial mais nous allons voir que tout se complique lorsque l'on considère les ondes électromagnétiques. Voyons tout d'abord le cas d'une onde non électromagnétique telle que l'onde de compression représentée schématiquement ci-dessous. Cette onde se propage à une certaine vitesse v_1 , déterminée par les propriétés mécaniques du milieu matériel qui la véhicule. Cette vitesse v_1 est, de façon évidente, définie dans le référentiel R dans lequel le milieu matériel est au repos. Dès lors, si l'onde est observée d'un autre référentiel R' dans lequel R a une vitesse v (ou du référentiel R' bougeant à la vitesse $-v$ par rapport à R qui est au repos), celle-ci aura une vitesse $v'_1 = v + v_1$.

Les évolutions de l'onde de compression vues des référentiels R et R' sont indiquées par les lignes grises sur le schéma ci-contre.



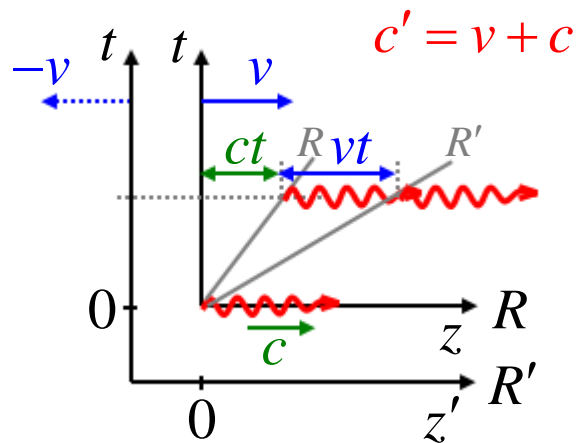
En conclusion, rien ne change par rapport au mouvement d'une particule et on ne voit toujours pas pourquoi les choses se compliqueraient. En particulier, la loi d'addition des vitesses est d'application et semble pour le moins logique et naturelle : si l'onde avance à la vitesse v_1 dans le milieu matériel et que ce milieu matériel avance lui-même à la vitesse v par rapport à l'observateur, on s'attend bien à ce que la vitesse de l'onde soit $v'_1 = v + v_1$ pour cet observateur.

Mais avec la découverte des ondes électromagnétiques par Maxwell dans la deuxième moitié du 19^{ème} siècle, les choses sont apparues nettement plus délicates. La raison de cette difficulté provient du fait que les ondes électromagnétiques n'ont pas besoin de milieu matériel pour se propager, elles peuvent se propager dans le vide. De même, leur vitesse « c » est déterminée par des grandeurs physiques qui ne caractérisent aucun matériau mais qui caractérise le vide, il s'agit, bien sûr, de la permittivité et de la perméabilité du vide. Ceci provient directement des équations de Maxwell qui combinées entre elles donnent l'équation d'onde standard. Elle est donnée ci-dessous à une dimension :

Equations de Maxwell :

$$\frac{\partial^2 \vec{E}}{\partial t^2} = c^2 \frac{\partial^2 \vec{E}}{\partial z^2}$$

Considérons l'onde EM vue des deux référentiels R et R' , comme on l'a fait pour l'onde de compression. Vue du référentiel R , l'onde a une vitesse c . Ceci est illustré sur le schéma ci-contre avec un paquet d'ondes (par exemple un photon) dont l'évolution est représentée par une droite de pente c repérée par la lettre R sur le schéma. Le même paquet d'ondes vu du référentiel R' aura logiquement une vitesse $c' = v + c$, ce qui est représenté par la droite de pente plus élevée repérée par la lettre R' .



On tombe alors sur une difficulté car la lumière dans R' est supposée, d'après les équations de Maxwell, se propager à la vitesse c et non à la vitesse $c' = v + c$. En effet, puisque le paquet d'ondes se propage dans le vide, il n'y a plus de repère absolu pour déterminer sa vitesse. Dans le référentiel R l'onde est dans le vide et sa vitesse est c , de même que dans le référentiel R' l'onde est dans le vide et sa vitesse est c . Si les équations de Maxwell sont correctes, dans tous les référentiels inertiels la vitesse de la lumière doit être c . Mais alors, on peut se demander ce que devient la loi d'addition des vitesses : si l'onde avance à la vitesse c dans R et que R avance lui-même à la vitesse v dans R' , on s'attendrait bien à ce que la vitesse de l'onde dans R' soit $c' = v + c$. Et pourtant, les équations de Maxwell disent le contraire car dans R et R' on a le « même vide », rien ne permet de différencier ces deux référentiels. Puisqu'ils sont équivalents du point de vue des lois physiques on s'attend à ce que les équations de Maxwell soient les mêmes pour ces deux référentiels et donc que la vitesse de l'onde EM soit c dans les deux. Il y a clairement un problème majeur à résoudre. J'insiste ici sur le fait que ce problème n'existe pas avec les ondes de mouvement de matière telles que les ondes sonores car celles-ci se propagent dans

un milieu matériel et ont une vitesse fixée dans le référentiel de ce milieu (vitesse déterminée par les caractéristiques du milieu) et non pas une même vitesse dans plusieurs référentiels. Ceci peut être résumé en disant que, bien que les ondes sonores et les ondes électromagnétiques sont régies par la même équation d'onde, soit $\partial^2 f / \partial t^2 = v^2 \partial^2 f / \partial z^2$, ces deux types d'ondes diffèrent fondamentalement dans la mesure où pour les premières la vitesse v qui apparaît dans l'équation est attachée au milieu matériel qui supporte les ondes tandis que pour les deuxièmes la vitesse vaut c et elle n'est rattachée à aucun référentiel particulier puisque ces ondes se propagent dans le vide et n'ont donc aucun support matériel.

La solution adoptée pour résoudre ce problème était de considérer que les équations de Maxwell étaient fausses ou, en tout cas, qu'elles n'étaient pas complètes. Comme la difficulté venait du fait que les ondes se propagent dans le vide, on se disait que les ondes EM se propagent en réalité dans un milieu impalpable pour les sens humains et que l'on a appelé pour cette raison « *éther* ». L'éther est l'équivalent du milieu matériel supportant l'onde de compression dans l'exemple précédent. Dès ce moment, toute la difficulté tombait car on pouvait à

nouveau appliquer la loi d'addition des vitesses de la relativité galiléenne. Mais, bien sûr, en adoptant cette solution on mettait en défaut les équations de Maxwell. Celles-ci devaient être modifiées pour tenir compte de l'existence de l'éther en ce sens que si le milieu pour lequel on formule les équations de Maxwell est en mouvement à la vitesse v par rapport à l'éther, alors ces équations doivent conduire à l'équation d'onde avec un facteur de vitesse égal à $c' = v + c$.

$$\frac{\partial^2 \vec{E}}{\partial t^2} = c^2 \frac{\partial^2 \vec{E}}{\partial z^2}$$

↓
 $c' \neq c$???

- équations de Maxwell fausses
- l'éther est le support des ondes EM

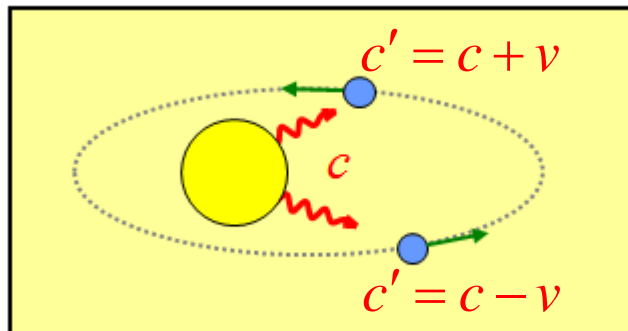
???

Ether = référentiel absolu (repos absolu)

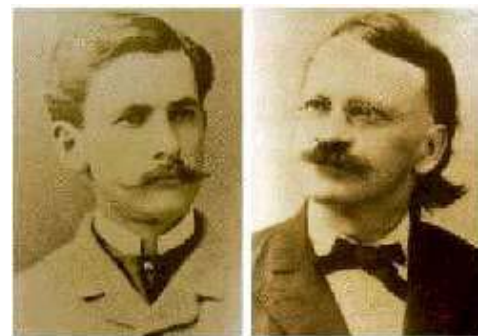
Cette solution semblait naturelle et pratique mais elle présentait deux difficultés importantes. L'une était d'ordre pratique et était due au fait qu'il s'est avéré impossible de modifier les équations de Maxwell de façon cohérente pour tenir compte de l'éther. L'autre était d'ordre philosophique car, en adoptant l'idée de l'éther, on adoptait l'idée de l'existence d'un référentiel absolu dans la nature. Tous les référentiels inertiels n'étaient plus équivalents contrairement à ce qu'avaient établi Galilée et Newton. Après avoir pensé que la terre était le centre de l'univers, l'homme est passé à l'idée que le soleil était le centre de l'univers (l'héliocentrisme), et avec Galilée et Newton, l'homme s'est « *libéré* » de la notion de référentiel absolu en disant que tous les référentiels inertiels sont équivalents et qu'il n'y a donc pas de repère absolu dans

l'univers. L'hypothèse de l'existence de l'éther était de ce point de vue une importante régression.

Pour vérifier l'existence de l'éther expérimentalement, il fallait recourir à des méthodes expérimentales de mesure de la vitesse de la lumière. Ceci est illustré sur le schéma ci-contre avec la lumière solaire observée de la terre. La terre a une orbite elliptique autour du soleil et elle a donc, à une période de l'année, une composante de vitesse dirigée vers le soleil, soit v . Si la lumière solaire se déplace dans l'éther elle s'y déplace avec la vitesse c et vue de la terre cette vitesse devient supérieure à c , soit $c' = c + v$. De même, six mois plus tard dans l'année, la terre s'éloigne du soleil avec la même vitesse v et la vitesse de la lumière devrait donc être inférieure à c , soit $c' = c - v$.



Mais aucune des tentatives de mesure de la vitesse de la lumière solaire n'a pas pu montrer de variation par rapport à c . Tout se passe comme si la terre était immobile par rapport au soleil ($v = 0$). Quelque soit la vitesse de la terre par rapport au soleil, la vitesse de la lumière est toujours la même et elle vaut rigoureusement c , comme prévu par les équations de Maxwell. Ce sont les américains Michelson et Morley qui en 1881 firent la première preuve expérimentale convaincante de la constance de la vitesse de la lumière.



Michelson

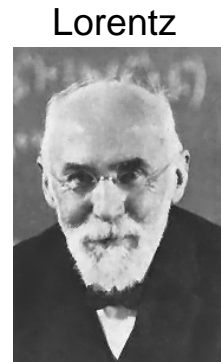
Morley

expérience de Michelson et Morley (1881) : $c' = c$

Ce résultat mettait en difficulté la théorie de l'éther même si pour la sauver des explications saugrenues étaient avancées telles que l'idée que la terre entraîne l'éther dans son sillage et l'immobilise autour d'elle sur une certaine épaisseur à l'image d'un objet se déplaçant dans un fluide visqueux, ce qui permettait d'expliquer pourquoi la mesure donnait toujours c .

5.4 Invariance des équations de Maxwell

Le physicien et mathématicien hollandais Hendrik Lorentz s'intéressait aux propriétés mathématiques des équations de Maxwell. Il étudia en particulier des transformations des variables d'espace et de temps (z et t) qui laissent inchangées ces équations. Pour comprendre son travail, voyons d'abord le problème simple de la transformation de l'équation d'onde pour un changement de référentiel galiléen. A cette fin, il est commode d'écrire l'équation d'onde à partir de la multiplication des deux opérateurs différentiels $\partial/\partial t + c\partial/\partial z$ et $\partial/\partial t - c\partial/\partial z$, comme indiqué ci-dessous.



Lorentz
1853-1928

$$\frac{\partial^2 \vec{E}(z,t)}{\partial t^2} - c^2 \frac{\partial^2 \vec{E}(z,t)}{\partial z^2} = 0 \quad \Rightarrow \quad \left[\frac{\partial}{\partial t} + c \frac{\partial}{\partial z} \right] \left[\frac{\partial}{\partial t} - c \frac{\partial}{\partial z} \right] \vec{E} = 0$$

Cet artifice mathématique nous permet de voir que si la multiplication du champ \vec{E} par un des opérateurs $\partial/\partial t + c\partial/\partial z$ ou $\partial/\partial t - c\partial/\partial z$ donne zéro, alors \vec{E} est solution de l'équation d'onde.

$$\left[\frac{\partial}{\partial t} \pm c \frac{\partial}{\partial z} \right] \vec{E} = 0$$

Cette dernière relation montre que les variations temporelles et spatiales du champ sont identiques (au signe près pour l'opérateur contenant le signe $+$, soit $\partial/\partial t + c\partial/\partial z$). C'est pour cette raison que la solution générale de l'équation d'onde a la forme d'une somme d'une onde progressive et d'une onde régressive.

$$\vec{E} = \vec{f}(z - ct) + \vec{g}(z + ct)$$

Il est, en effet, très facile de vérifier que cette forme du champ est solution de l'équation $(\partial/\partial t + c\partial/\partial z)(\partial/\partial t - c\partial/\partial z)\vec{E} = 0$

5.4.1 Changement de référentiel galiléen

Effectuons maintenant le changement de variables correspondant au passage d'un référentiel inertiel R à un autre référentiel inertiel R' dans lequel le référentiel R a une vitesse v . Si z est la coordonnée spatiale du référentiel R et z' la coordonnée spatiale du référentiel R' , on a $z' = z + vt$. C'est ce que l'on appelle un changement de variable galiléen car il est construit sur la loi galiléenne d'addition des vitesses. Comme nous cherchons à voir comment se transforme l'équation d'onde avec ce changement de variable nous devons appliquer les règles de changement de variables pour les dérivées partielles. Ces règles, pour être cohérentes sur le plan mathématique, nous obligent à introduire une variable temporelle t' représentant le temps dans le nouveau repère. Comme le temps est absolu en relativité galiléenne, on a bien sûr $t' = t$ (le temps s'écoule de la même façon quelque soit le référentiel dans lequel on se trouve).

$$\begin{cases} z' = z + vt \\ t' = t \end{cases}$$

Avec le changement de variables $z' = z + vt$ et $t' = t$, les dérivées partielles se transforment de la façon qui est détaillée ci-dessous :

$$\left\{ \begin{array}{l} \frac{\partial}{\partial z} = \frac{\partial}{\partial z'} \frac{dz'}{dz} + \frac{\partial}{\partial t'} \frac{dt'}{dz} \\ \frac{\partial}{\partial t} = \frac{\partial}{\partial z'} \frac{dz'}{dt} + \frac{\partial}{\partial t'} \frac{dt'}{dt} \end{array} \right. \Rightarrow \left\{ \begin{array}{l} \frac{\partial}{\partial z} = \frac{\partial}{\partial z'} \\ \frac{\partial}{\partial t} = \frac{\partial}{\partial t'} + v \frac{\partial}{\partial z'} \end{array} \right.$$

Les deux opérateurs de l'équation d'onde se transforment donc comme indiqué ci-dessous :

$$\left\{ \begin{array}{l} \left[\frac{\partial}{\partial t} + c \frac{\partial}{\partial z} \right] \vec{E} \rightarrow \left[\frac{\partial}{\partial t'} + (c + v) \frac{\partial}{\partial z'} \right] \vec{E} \\ \left[\frac{\partial}{\partial t} - c \frac{\partial}{\partial z} \right] \vec{E} \rightarrow \left[\frac{\partial}{\partial t'} - (c - v) \frac{\partial}{\partial z'} \right] \vec{E} \end{array} \right.$$

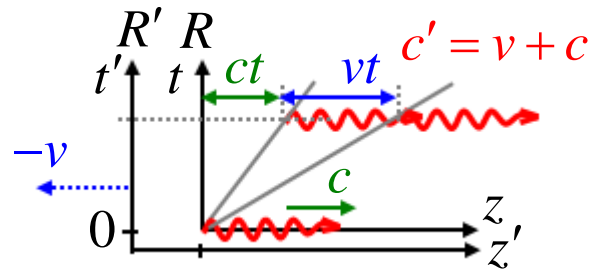
La solution de l'équation d'onde est donnée par le champ qui satisfait soit l'équation $[\partial/\partial t' + (c+v)\partial/\partial z']\vec{E}$ soit la suivante $[\partial/\partial t' - (c-v)\partial/\partial z']\vec{E}$. On voit ainsi que le champ \vec{E} dans le référentiel R' a pour forme générale :

$$\vec{E} = \vec{f}[z - (c + v)t] + \vec{g}[z + (c - v)t]$$

Comme attendu, ce résultat est en accord avec la relativité galiléenne. En particulier, la vitesse de l'onde progressive dans le référentiel R' vaut $c' = c + v$ alors que l'onde régressive s'y propage avec la vitesse $c' = c - v$.

Considérer ce résultat comme physiquement acceptable signifierait que les équations de Maxwell ne sont pas « *invariantes* » par rapport à une transformation de référentiel galiléen. La multiplication de l'opérateur transformé avec \vec{E} donne, en effet, $(\partial/\partial t' + (c + v)\partial/\partial z')(\partial/\partial t' - (c - v)\partial/\partial z')\vec{E}$ qui est une expression différente de l'équation d'onde $(\partial^2/\partial t'^2 - c^2 \partial^2/\partial z'^2)\vec{E} = 0$ obtenue à partir des équations de Maxwell exprimées dans le référentiel R .

Clairement, avec les variables t' et z' l'équation d'onde standard n'est pas satisfaite, comme indiqué sous le graphe ci-contre. Autrement dit, cette équation n'est pas invariante par rapport à un changement de repère galiléen.



$$\frac{\partial^2 \vec{E}}{\partial t'^2} - c^2 \frac{\partial^2 \vec{E}}{\partial z'^2} \neq 0$$

équations de Maxwell

non - invariantes

Ce résultat est impossible à accepter si on adopte l'idée que tous les référentiels inertiels sont équivalents. En effet l'équation d'onde EM écrite dans le référentiel R' vient des équations de Maxwell dans le vide et elle devrait donc être la même que dans le référentiel R . Dans le référentiel R' de variables z' et t' , les équations de Maxwell donnent effectivement $(\partial^2/\partial t'^2 - c^2 \partial^2/\partial z'^2)\vec{E} = 0$, ce qui implique que la vitesse de l'onde dans R' doit être c et non $c' = c + v$.

Sur base de ce constat Hendrik Lorentz chercha un changement de variables différents qui conserve l'équation d'onde (ou les équations de Maxwell). En d'autres termes, son but était de trouver un changement de variables $(z, t) \rightarrow (z', t')$ tel que dans les deux référentiels l'équation d'onde soit la même, soit :

$$\frac{\partial^2 E}{\partial z^2} = \frac{1}{c^2} \frac{\partial^2 E}{\partial t^2} \rightarrow \frac{\partial^2 E}{\partial z'^2} = \frac{1}{c^2} \frac{\partial^2 E}{\partial t'^2}$$

Si un tel changement de variable existe, cela signifie que la vitesse de la lumière est la même dans les deux référentiels correspondants, ce qui permettrait d'expliquer pourquoi la vitesse de la lumière apparaît comme une constante indépendante du référentiel dans laquelle on la mesure.

5.4.2 Invariance de rotation

Avant de décrire le développement de Lorentz, je propose ici de traiter un problème permettant d'établir l'invariance d'une équation pour une certaine transformation. Pour cela je prends l'exemple du changement de référentiel correspondant à une rotation autour d'un axe, le but étant de montrer que l'équation d'onde EM est invariante par rapport à cette transformation. Notez que cette invariance de rotation est une évidence puisque notre univers ne présente pas de direction privilégiée on s'attend effectivement à ce que l'équation d'onde EM soit indépendante de l'orientation du repère que l'on considère pour l'écrire.

Considérons l'équation d'onde EM à trois dimensions : $\Delta E = \frac{1}{c^2} \frac{\partial^2 E}{\partial t^2}$

En explicitant le laplacien en termes des coordonnées cartésiennes, elle s'écrit :

$$\frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial y^2} + \frac{\partial^2 E}{\partial z^2} = \frac{1}{c^2} \frac{\partial^2 E}{\partial t^2}$$

Sa solution d'onde plane est : $E(\vec{x}, t) = \text{Re} \left[A e^{i\vec{k} \cdot \vec{x}} e^{-i\omega t} \right]$

Si l'on explicite le vecteur d'onde $\vec{k} = k_x \vec{1}_x + k_y \vec{1}_y + k_z \vec{1}_z$, le champ complexe s'écrit comme suit :

$$\underline{E} = A e^{ik_x x} e^{ik_y y} e^{ik_z z} e^{-i\omega t}$$


Notez qu'il ne s'agit pas d'un phaseur dans la mesure où la dépendance temporelle y est incluse. Il s'agit d'une généralisation à 4 dimensions de la notion de phaseur. On l'appellera par la suite le « *champ complexe* » et l'on retiendra qu'il suffit d'en prendre la partie réelle pour obtenir le champ réel.

Si l'on s'intéresse à cette expression c'est parce que Lorentz s'est basé sur la notion de relation de dispersion pour étudier l'invariance de l'équation d'onde EM. Or la relation de dispersion se trouve en substituant le champ complexe dans l'équation d'onde (voir l'introduction à la mécanique quantique, section 2.3). Cette substitution donne effectivement :

$$k_x^2 + k_y^2 + k_z^2 = \frac{1}{c^2} \omega^2$$

Cette relation est la généralisation à 3 dimensions spatiales de la relation de dispersion des ondes EM dans le vide (à une dimension cette relation était simplement $k = \omega / c$).

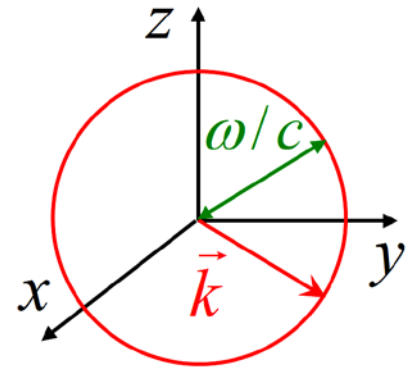
Nous avons vu dans l'introduction à la mécanique quantique qu'il existe un lien direct entre la relation de dispersion des ondes et l'équation qui régit ces ondes. Ce lien apparaît ici évident en constatant qu'il suffit, pour obtenir l'équation d'onde, de remplacer dans la relation de dispersion les composantes du vecteur d'onde et la pulsation par les opérateurs différentiels suivants : $k_x \rightarrow -i\partial / \partial x$, $k_y \rightarrow -i\partial / \partial y$, $k_z \rightarrow -i\partial / \partial z$, et $\omega \rightarrow i\partial / \partial t$. On vérifie effectivement que l'application de l'opérateur $-i\partial / \partial x$ au champ complexe \underline{E} donne bien $k_x \underline{E}$ (et de même pour les autres dérivées).

$$\frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial y^2} + \frac{\partial^2 E}{\partial z^2} = \frac{1}{c^2} \frac{\partial^2 E}{\partial t^2}$$


$$k_x^2 + k_y^2 + k_z^2 = \frac{\omega^2}{c^2}$$

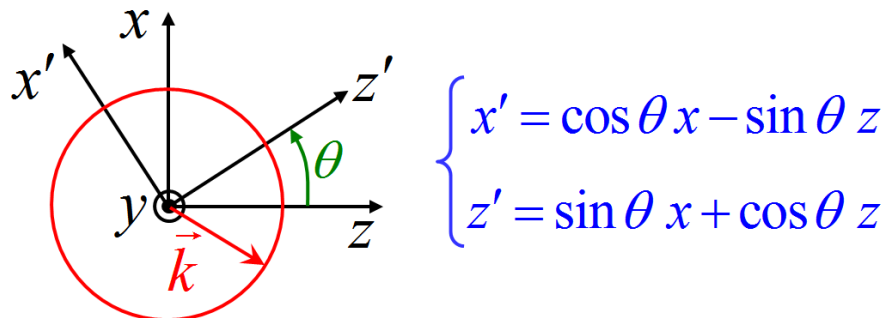
C'est sur cette équivalence entre relation de dispersion et équation d'onde que Lorentz s'est basé pour faire son étude de l'invariance des équations de Maxwell (ou équation d'onde EM). L'idée est la suivante : puisqu'il y a une correspondance biunivoque entre la relation de dispersion et l'équation d'onde associée, il n'y a qu'à démontrer l'invariance de la relation de dispersion pour prouver l'invariance de l'équation.

Le schéma ci-contre est une représentation graphique de la relation de dispersion des ondes EM, soit $k_x^2 + k_y^2 + k_z^2 = \omega^2 / c^2$. Cette relation est d'une interprétation évidente : elle montre que le vecteur à une norme donnée par ω / c . Autrement dit, le vecteur d'onde d'une onde EM a toujours son extrémité sur une sphère de rayon ω / c . L'invariance de rotation de l'équation d'onde EM apparaît donc comme une évidence puisque la sphère est effectivement invariante par rotation.



Malgré l'évidence de ce résultat je propose de le traiter ici en détail avec l'exemple de la rotation d'un angle θ autour de l'axe y . Le changement des variables de coordonnées correspondant $(x, z) \rightarrow (x', z')$ est indiqué ci-dessous.

invariance par rotation



Pour vérifier l'invariance par rotation, le changement de variables $(x, z) \rightarrow (x', z')$ peut être appliqué à l'équation d'onde. Pour cela les règles de transformation des dérivées partielles doivent être appliquées. Les nouvelles expressions des dérivées partielles qui en résultent sont indiquées ci-contre. Mais nous ne ferons pas ce calcul ici car notre but est de travailler sur l'invariance de la relation de dispersion.

$$\left\{ \begin{array}{l} \frac{\partial}{\partial x} = \frac{\partial}{\partial x'} \frac{dx'}{dx} + \frac{\partial}{\partial z'} \frac{dz'}{dx} \\ \frac{\partial}{\partial z} = \frac{\partial}{\partial x'} \frac{dx'}{dz} + \frac{\partial}{\partial z'} \frac{dz'}{dz} \end{array} \right.$$

$\begin{matrix} \nearrow \cos \theta & \nearrow \sin \theta \\ \searrow -\sin \theta & \searrow \cos \theta \end{matrix}$

Dans le nouveau repère (x', y', z') le vecteur d'onde s'écrit $\vec{k} = k_{x'} \vec{1}_{x'} + k_{y'} \vec{1}_{y'} + k_{z'} \vec{1}_{z'}$. Il est évident que celui-ci a toujours la même norme ω / c puisque la rotation conserve la norme, ce qui signifie que la nouvelle relation de dispersion s'écrit :

$$k_{x'}^2 + k_{y'}^2 + k_{z'}^2 = \frac{\omega^2}{c^2}$$

Cette relation est identique à la relation de dispersion dans le référentiel (x, y, z) . Si on lui applique les règles de transposition pour obtenir l'équation d'onde correspondante, soit $k_{x'} \rightarrow -i\partial/\partial x'$, $k_{y'} \rightarrow -i\partial/\partial y'$, $k_{z'} \rightarrow -i\partial/\partial z'$, et $\omega \rightarrow i\partial/\partial t$, nous constatons que l'on obtient, comme prévu la même équation d'onde. Nous avons ainsi vérifié l'invariance de rotation des équations de Maxwell (ou de l'équation d'onde EM).

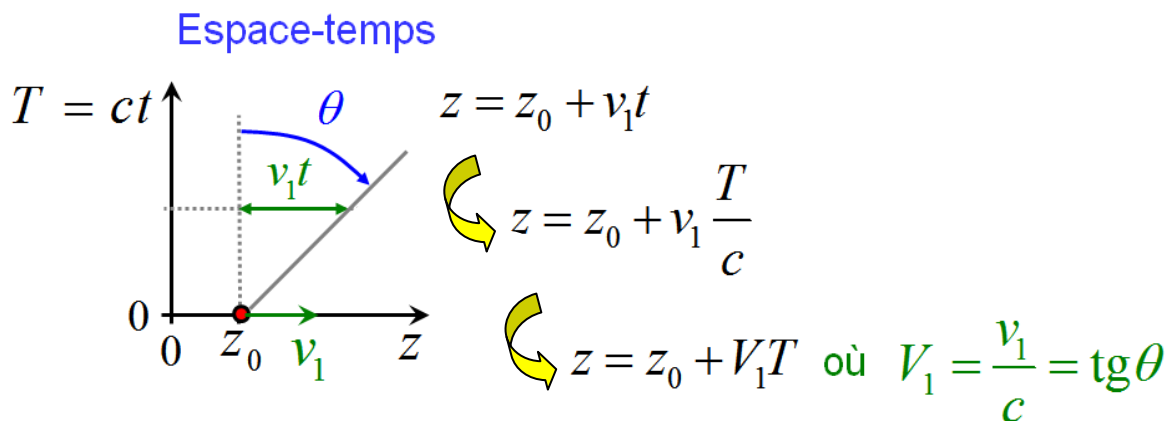
$$\frac{\partial^2 E}{\partial x'^2} + \frac{\partial^2 E}{\partial y'^2} + \frac{\partial^2 E}{\partial z'^2} = \frac{1}{c^2} \frac{\partial^2 E}{\partial t^2}$$

5.4.3 Transformations de Lorentz

Nous allons ici considérer la même procédure mais pour une transformation impliquant l'espace et le temps. Pour la simplicité des calculs nous ne considérerons que l'équation d'onde EM à une dimension.

$$\frac{\partial^2 E}{\partial z^2} = \frac{1}{c^2} \frac{\partial^2 E}{\partial t^2}$$

Nous commençons par symétriser cette équation en appliquant le changement de variable temporelle suivant : $T = ct$. La nouvelle variable temporelle $T = ct$ a les dimensions d'une longueur. Nous l'appellerons pour cette raison le temps « espace ». Le schéma suivant montre l'évolution d'une masse ponctuelle se déplaçant à vitesse constante v_1 le long de l'axe z . L'évolution $z = z_0 + v_1 t$ est indiquée par la ligne droite grise inclinée sur le graphe. Plus la vitesse est grande, plus cette droite se rapproche de l'horizontale. Pour exprimer la vitesse dans la nouvelle variable il suffit de remplacer $t = T/c$ dans la relation $z = z_0 + v_1 t$, ce qui donne $z = z_0 + V_1 T$ où $V_1 = v_1/c$ est une vitesse sans dimension. Cette vitesse est la pente de la droite donnant le mouvement de la masse dans le référentiel (z, T) , soit $dz/dT = V_1$. Notez qu'il s'agit de la pente par rapport à la verticale de sorte que $\tan \theta = V_1$ où θ est l'angle que la droite fait avec la verticale.



Le nouvel ensemble de variables d'espace et de temps (z, T) constituent les coordonnées de ce que l'on appelle en physique l'« *espace-temps* » (l'espace-temps a en pratique 4 dimensions). De même la droite donnant l'évolution de la position de la masse au cours du temps s'appelle une trajectoire d'espace-temps ou une « *ligne d'univers* ».

Avec cette nouvelle variable temporelle l'équation d'onde EM s'écrit :

$$\frac{\partial^2 E}{\partial z^2} = \frac{\partial^2 E}{\partial T^2}$$

Pour étudier la solution harmonique de l'équation d'onde exprimée avec cette nouvelle variable temporelle nous repassons un instant à l'équation de départ en (z, t) , soit $\partial^2 E / \partial t^2 = c^2 \partial^2 E / \partial z^2$. La solution harmonique de cette équation peut s'écrire sous la forme suivante :

$$E(z, t) = \text{Re} \left[A e^{ik_z z} e^{-i\omega t} \right]$$

Où le nombre d'onde est noté k_z pour une raison qui apparaîtra plus loin. Si l'on applique le changement de variable $t = T / c$ on trouve :

$$E(z, T) = \text{Re} \left[A e^{ik_z z} e^{-i\frac{\omega}{c} T} \right]$$

Si on introduit alors la constante $k_T = \omega / c$ on peut écrire :

$$E(z, T) = \text{Re} \left[A e^{ik_z z} e^{-ik_T T} \right]$$

La substitution de cette solution dans l'équation d'onde fournit la relation de dispersion associée, soit :

$$k_z^2 = k_T^2$$

Remarquez que cette égalité n'est rien d'autre que l'expression de la relation de dispersion habituelle des ondes EM dans le vide, soit $k = \omega / c$. On peut encore l'écrire comme suit :

$$k_z^2 - k_T^2 = 0$$

Pour trouver la transformation d'espace-temps $(z, T) \rightarrow (z', T')$ qui conserve cette relation de dispersion (je rappelle que c'est le but de ce développement) il est utile de faire un pas supplémentaire dans l'abstraction en introduisant une variable de temps complexe au travers de la relation :

$$T_i = iT$$

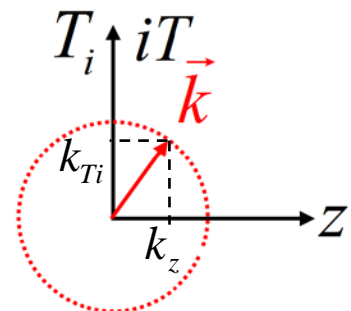
De même, nous introduisons le nombre d'onde complexe associé à cette nouvelle variable temporelle :

$$k_{Ti} = ik_T$$

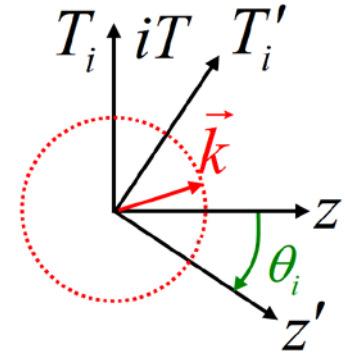
De sorte que la solution d'onde harmonique et la relation de dispersion deviennent :

$$\begin{cases} E(z, T_i) = \text{Re} \left[A e^{ik_z z} e^{ik_{Ti} T_i} \right] \\ k_z^2 + k_{Ti}^2 = 0 \end{cases}$$

Le fait d'être passé au domaine complexe a permis de symétriser ces deux expressions. En particulier, la relation de dispersion a maintenant la forme d'une simple relation liant les composantes d'un vecteur d'onde, plus exactement, cette relation donne la norme au carré d'un vecteur d'onde (notez que cette norme est nulle mais, vu que k_{Ti} est une grandeur imaginaire, ceci ne signifie pas que les composantes du vecteur d'onde sont nulles). Dans la représentation graphique donnée ci-contre k_z et k_{Ti} apparaissent comme les composantes d'un vecteur d'onde \vec{k} dans les directions $\vec{1}_z$ et $\vec{1}_{Ti}$. Notez que cette représentation est purement abstraite car une des dimensions est imaginaire (en particulier, la norme nulle du vecteur \vec{k} ne peut pas être représentée graphiquement avec des composantes non-nulles). Ce schéma nous guide néanmoins dans notre recherche d'une transformation qui laisse l'équation d'onde EM invariante.



Par analogie au développement de la section précédente sur l'invariance par rotation dans l'espace, on peut facilement anticiper que l'équation d'onde EM sera invariante également par rotation dans l'espace temps complexe (z, T_i) . En effet, puisque la rotation conserve la norme, la relation de dispersion $k_z^2 + k_{T_i}^2 = 0$ restera inchangée. La transformation d'angle θ_i représentée sur le schéma ci-contre et dont les formules de transformation des variables sont écrites ci-dessous, conserve donc la relation de dispersion :



$$\begin{cases} z' = \cos \theta_i z - \sin \theta_i T_i \\ T_i' = \cos \theta_i T_i + \sin \theta_i z \end{cases} \quad \Rightarrow \quad k_{z'}^2 + k_{T_i'}^2 = 0$$

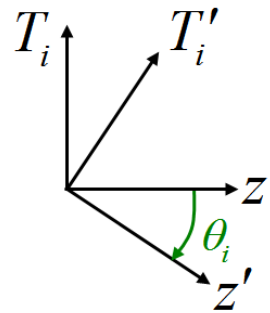
En introduisant la variable de « temps espace » réelle $T' = -iT_i'$ et le nombre d'onde réel associé $k_{T'} = -ik_{T_i'}$, cette relation de dispersion s'écrit $k_z^2 = k_{T'}^2$, ce qui montre l'invariance de la relation de dispersion initiale $k_z^2 = k_T^2$ par rapport à la transformation $(z, T) \rightarrow (z', T')$ proposée. Il en est bien sûr de même pour l'équation d'onde EM associée qui, d'après les règles de correspondance $k_z \rightarrow -i\partial/\partial z'$ et $k_{T'} \rightarrow i\partial/\partial T'$, s'écrit comme indiqué ci-dessous.

$$k_{z'}^2 = k_{T'}^2 \quad \Rightarrow \quad \frac{\partial^2 E}{\partial z'^2} = \frac{\partial^2 E}{\partial T'^2}$$

En réintroduisant la variable temporelle t , on obtient finalement l'équation d'onde EM dans le nouveau référentiel (z', t') :

$$\frac{\partial^2 E}{\partial z'^2} = c^2 \frac{\partial^2 E}{\partial t'^2}$$

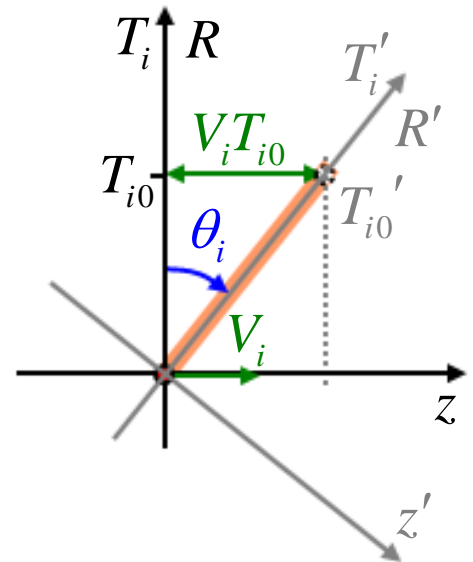
Cette équation est identique à l'équation d'onde EM dans le référentiel (z, t) . Nous pouvons donc en conclure que nous avons trouvé le changement de variables $(z, t) \rightarrow (z', t')$ qui laisse invariante l'équation d'onde EM. Les formules de transformation des variables correspondantes s'appellent les transformations de Lorentz. Mais nous n'avons pas encore les transformations de Lorentz $(z, t) \rightarrow (z', t')$ de façon explicite puisqu'elles n'ont été formulées jusqu'à présent que dans les variables d'espace-temps complexe (z, T_i) , comme illustré ci-dessous.



$$\begin{cases} z' = \cos \theta_i z - \sin \theta_i T_i \\ T'_i = \cos \theta_i T_i + \sin \theta_i z \end{cases}$$

Avant de passer aux transformations de Lorentz en variables réelles (z, t) et (z', t') , il est intéressant d'analyser la transformation en termes des variables (z, T_i) . Nous allons à cette fin considérer que l'espace-temps complexe (z, T_i) constitue le cadre spatio-temporel dans lequel nous vivons.

Le graphe ci-contre montre la ligne d'univers dans l'espace-temps complexe (z, T_i) d'un point matériel en mouvement à vitesse constante dans un référentiel R . Puisque T_i représente le temps, la ligne d'univers est bien sûr inclinée par rapport à la verticale pour représenter le mouvement du point. La vitesse du point est plus exactement donnée par la pente de la ligne d'univers, soit $V_i = dz / dT_i = \tan(\theta_i)$. Considérons maintenant le référentiel R' donné par les transformations de rotation de l'angle θ_i considérées ci-dessus (comparez les deux schémas). Le point est clairement au repos dans ce référentiel dans la mesure où sa ligne d'univers est le long de l'axe T'_i .



Ce qui frappe dans ce modèle d'univers c'est que le temps et l'espace d'un référentiel ne sont pas les mêmes que ceux d'un autre référentiel. Par exemple, un laps de temps T'_{i0} dans le référentiel R' apparaîtra de durée différente dans le référentiel R . Sur le schéma on voit effectivement qu'un évènement ayant eu lieu au temps T'_{i0} dans le référentiel R' apparaît au temps $T_{i0} = T'_{i0} \cos \theta_i$ dans le référentiel R . Ceci est dû au fait que le temps d'un référentiel a une projection non nulle dans l'espace de l'autre et vice-versa, l'espace d'un référentiel a une projection non nulle dans le temps de l'autre. Ceci se retrouve sur plan mathématique dans la structure des transformations de Lorentz qui implique précisément des projections de l'espace dans le temps et inversement. En résumé, on constate qu'espace et temps forment un tout indissociable dans lequel l'espace peut se transformer en temps et le temps peut se transformer en espace au gré des mouvements qui y sont effectués. L'invariance de la vitesse de la lumière (c'est-à-dire de l'équation d'onde EM) suggère que nous vivons dans un tel univers.

Mais bien sûr, il s'agit de préciser les choses car notre temps (celui que l'on mesure avec une horloge) n'est pas le temps imaginaire T_i , et si le changement d'un référentiel à l'autre se fait par une simple rotation dans l'espace-temps complexe (z, T_i) , il n'en est pas de même pour les variables réelles (z, t) . Pour obtenir les transformations de Lorentz en termes des variables (z, t) et (z', t') il nous suffit d'appliquer en sens inverse les changements de variables introduits successivement pour arriver aux variables (z, T_i) . Commençons par introduire les temps « espace » réels T et T' à partir de la définition du temps imaginaire : $T_i = iT$ et $T'_i = iT'$. On obtient les relations données ci-dessous à droite.

$$\begin{cases} z' = \cos \theta_i z - \sin \theta_i T_i \\ T'_i = \cos \theta_i T_i + \sin \theta_i z \end{cases} \longrightarrow \begin{cases} z' = \cos \theta_i z - i \sin \theta_i T \\ iT'_i = i \cos \theta_i T + \sin \theta_i z \end{cases}$$

En multipliant la deuxième ligne par $-i$, on obtient :

$$\begin{cases} z' = \cos \theta_i z - i \sin \theta_i T \\ T' = \cos \theta_i T + i \sin \theta_i z \end{cases}$$

Il convient pour continuer de se débarrasser des coefficients imaginaires qui apparaissent dans ces expressions. Pour cela il faut avoir à l'esprit que l'angle θ_i définit la vitesse du mobile dans le référentiel R . Pour le voir clairement, considérons le point $z' = 0$ dans le référentiel R' , d'après la transformation $z' = \cos \theta_i z - \sin \theta_i T_i$ on constate que pour ce point les coordonnées z et T_i sont liées par la relation $z = \tan \theta_i T_i$, ce qui montre que la vitesse dans le référentiel R du point $z' = 0$ est donnée par $V_i = \tan \theta_i$. En termes du temps réel on trouve $z = i \tan \theta_i T$, ce qui montre que la vitesse du point $z' = 0$ est donnée par $i \tan \theta_i$. La vitesse dans l'espace-temps réel étant une grandeur réelle, il apparaît clair que θ_i ne peut être qu'un angle imaginaire. Ceci nous amène donc à poser $\theta_i = -i\theta$ où θ est une grandeur réelle. Pour voir ce que deviennent les transformations de Lorentz il convient de voir ce que deviennent les fonctions $\sin \theta_i$ et $\cos \theta_i$.

$$\sin \theta_i = \frac{1}{2i}(e^{i\theta_i} - e^{-i\theta_i}) = \sin(-i\theta) = \frac{1}{2i}(e^{\theta} - e^{-\theta}) = -i \operatorname{sh}(\theta)$$

$$i \sin \theta_i = \operatorname{sh}(\theta)$$

$$\cos \theta_i = \frac{1}{2}(e^{i\theta_i} + e^{-i\theta_i}) = \cos(i\theta) = \frac{1}{2}(e^{-\theta} + e^{\theta}) = \operatorname{ch}(\theta)$$

$$\cos(\theta_i) = \operatorname{ch}(\theta)$$

Les transformations de Lorentz prennent donc la forme suivante :

$$\begin{aligned} z' &= \text{ch } \theta \, z - \text{sh } \theta \, T \\ T' &= \text{ch } \theta \, T - \text{sh } \theta \, z \end{aligned}$$

On notera que la vitesse V du point $z' = 0$ (c'est-à-dire, la vitesse du référentiel R') est cette fois donnée par le rapport $\text{sh } \theta / \text{ch } \theta$. Cette vitesse n'est autre que la vitesse du référentiel R' incliné d'un angle $\theta_i = -i\theta$ dans l'espace-temps complexe (z, T_i) du référentiel R , elle est donnée par la tangente hyperbolique de θ , soit $V = \text{sh } \theta / \text{ch } \theta = \text{th } \theta$. Ceci donne la signification de la variable θ : cette variable détermine la vitesse V d'un référentiel par rapport à l'autre au travers de la relation $V = \text{th } \theta$.

Il est facile d'écrire les transformations de Lorentz en fonction de la vitesse adimensionnelle V ; il suffit effectivement pour cela de connaître l'expression du cosinus et du sinus hyperboliques en fonction de la tangente hyperbolique. Sachant que $\text{ch}^2 \theta - \text{sh}^2 \theta = 1$, ces expressions sont évidentes :

$$\begin{aligned} \text{sh } \theta &= \frac{\text{th } \theta}{\sqrt{1 - \text{th}^2 \theta}} = \frac{V}{\sqrt{1 - V^2}} \\ \text{ch } \theta &= \frac{1}{\sqrt{1 - \text{th}^2 \theta}} = \frac{1}{\sqrt{1 - V^2}} \end{aligned}$$

Ceci nous mène à la forme suivante des transformations de Lorentz :

$$\begin{cases} z' = \frac{1}{\sqrt{1 - V^2}} z - \frac{V}{\sqrt{1 - V^2}} T \\ T' = \frac{1}{\sqrt{1 - V^2}} T - \frac{V}{\sqrt{1 - V^2}} z \end{cases}$$

Sachant que le « temps espace » T est lié au temps réel t par la relation $T = ct$ et que la vitesse adimensionnelle V vaut v/c on arrive finalement à l'expression des transformations de Lorentz en variables (z, t) :

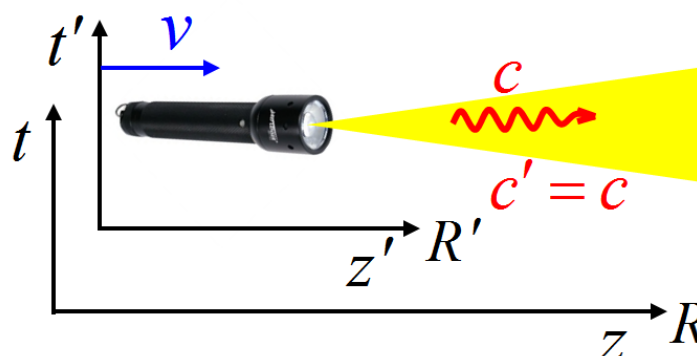
$$\begin{aligned} z' &= \frac{c}{\sqrt{c^2 - v^2}} z - \frac{v}{\sqrt{c^2 - v^2}} ct \\ ct' &= \frac{c}{\sqrt{c^2 - v^2}} ct - \frac{v}{\sqrt{c^2 - v^2}} z \end{aligned}$$

En divisant la deuxième ligne par c on trouve finalement les transformations de Lorentz sous leur forme standard :

Transformations de Lorentz

$$\begin{aligned} z' &= \frac{1}{\sqrt{1 - v^2/c^2}} z - \frac{v}{\sqrt{1 - v^2/c^2}} t \\ t' &= \frac{1}{\sqrt{1 - v^2/c^2}} t - \frac{v/c^2}{\sqrt{1 - v^2/c^2}} z \end{aligned}$$

A la différence des transformations de Galilée, ces transformations de coordonnées spatio-temporelles conservent l'équation d'onde EM ou, en d'autres termes, elles conservent les équations de Maxwell. Ceci signifie qu'une onde électromagnétique se déplace à la même vitesse dans les référentiels R et R' qui ont entre eux une vitesse relative v . Ceci est illustré sur le schéma suivant où l'on voit un faisceau lumineux émis par une lampe au repos dans le référentiel R' . Si les variables (z, t) et (z', t') sont liées entre elles par les transformations de Lorentz alors l'onde lumineuse se propage à la vitesse c dans les deux référentiels, conformément à l'expérience de Michelson et Morley.



Lorentz a ainsi trouvé une solution mathématique au problème de la constance de la vitesse de la lumière observée expérimentalement. Le problème du point de vue de Lorentz sur ce problème conceptuel est qu'il n'a jamais considéré que ces transformations traduisent la réalité de notre cadre d'espace-temps. Selon lui ces transformations n'avaient pas plus de valeur qu'un simple jeu mathématique. C'est Albert Einstein à l'âge de 26 ans qui formula pour la première fois l'idée que les transformations de Lorentz traduisent la réalité de notre cadre d'espace-temps. Il est arrivé à cette conviction sur base d'un travail théorique qui va bien au-delà des transformations de Lorentz. Dans un article de 1905 Albert Einstein va effectivement jusqu'à remettre en question toute la mécanique Newtonienne qui était à l'époque considérée comme d'une rigueur et d'une exactitude absolues. Ces réflexions l'ont amené à dire que l'éther n'existe pas et que les équations de Maxwell sont malgré tout correctes dans tout référentiel, ce qui allait à l'encontre de la pensée dominante de l'époque.

Albert Einstein



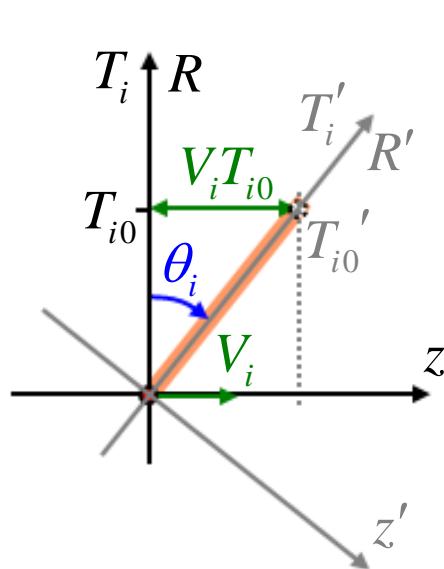
(1879-1955)

5.5 Transformations de Lorentz dans l'espace-temps complexe

Dans le but d'acquérir une image mentale et une intuition de la nature étrange de notre cadre d'espace-temps, nous allons souvent par la suite travailler dans la représentation d'espace-temps complexe z, T_i pour laquelle les transformations de Lorentz se limitent à une simple rotation. Comme la rotation d'un référentiel par rapport à un autre est facile à représenter graphiquement, la représentation dans l'espace-temps complexe z, T_i pourra être utilisée comme outil efficace pour appréhender et résoudre des problèmes de relativité. Le seul problème de cet outil est que les représentations graphiques qu'il fournit ne sont pas interprétables directement car elles impliquent des grandeurs complexes de façon implicite. Il faudra donc être très prudent lors de son utilisation.

5.5.1 Ligne d'univers et vitesse

Pour nous familiariser avec cette représentation, je suggère de revoir en détail la notion de ligne d'univers et de vitesse. Pour cela je reprends ci-dessous le graphe de la ligne d'univers d'un mobile se déplaçant à vitesse constante le long de l'axe z du référentiel R . Les transformations de Lorentz dans l'espace-temps z, T_i sont reprises ci-dessous. Le référentiel R' de coordonnées (z', T'_i) est le référentiel dans lequel le mobile est au repos, ce dernier fait un angle θ_i avec le référentiel R .



$$\begin{cases} z' = z \cos \theta_i - T_i \sin \theta_i \\ T_i' = T_i \cos \theta_i + z \sin \theta_i \end{cases}$$

Pour bien comprendre les représentations graphiques dans l'espace-temps complexe (z, T_i) il est important de réaliser qu'un évènement donné est représenté par un point dont les coordonnées (z, T_i) représentent respectivement la position et le moment de l'évènement par rapport à l'origine.

L'angle θ_i entre les deux référentiels représente la différence de vitesse v qu'il y a entre ces deux référentiels. En effet, l'angle θ_i est défini par $\theta_i = -i\theta$ où $\text{th}\theta = V = v/c$. L'angle θ_i est donc un angle imaginaire dont le module est d'autant plus grand que la vitesse v est grande.

Dans l'espace-temps complexe la vitesse du mobile vaut : $V_i = \frac{dz}{dT_i} = \text{tg} \theta_i$

Il s'agit d'une vitesse imaginaire qui n'a donc pas de signification physique directe en tant que pente d'une droite, c'est pourquoi il faudra être prudent avec l'interprétation des graphes dans l'espace-temps complexe. Dans la vitesse imaginaire V_i on retrouve facilement la vitesse adimensionnelle réelle V . Il suffit pour cela de remplacer l'angle θ_i par son expression $\theta_i = -i\theta$, ce qui d'après le calcul ci-dessous donne $V_i = -iV$.

$$V_i = \frac{dz}{dT_i} = \text{tg} \theta_i = \text{tg}(-i\theta) = -\frac{\sin(i\theta)}{\cos(i\theta)}$$

$$\hookrightarrow V_i = -i \frac{\text{sh}\theta}{\text{ch}\theta} = -i \text{th}\theta \quad \Rightarrow \quad V_i = -iV$$

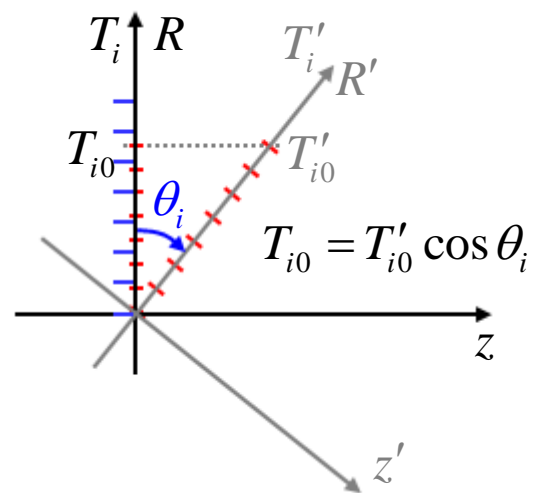
Il est intéressant de comparer ce résultat à celui que l'on obtient à partir du changement de variable $T_i = iT$ dans la dérivée de z par rapport au temps. Cela donne les relations ci-dessous à gauche. La comparaison avec l'expression de V_i obtenue ci-dessus $V_i = -iV$ montre que la dérivée de z par rapport au « temps espace » T donne, comme attendu, la vitesse adimensionnelle réelle V .

$$V_i = \frac{dz}{iT} = -i \frac{dz}{dT} \quad \Rightarrow \quad \frac{dz}{dT} = V$$

Le fait que les transformations de Lorentz représentent une rotation dans l'espace-temps complexe fait que le temps d'un référentiel a une projection non nulle dans l'espace de l'autre référentiel et inversement. En d'autres termes, l'espace et le temps apparaissent comme deux dimensions complètement indissociables. Deux référentiels inertiels ayant une vitesse relative non nulle entre eux mêlent leurs dimensions d'espace et de temps. Ceci conduit aux phénomènes maintenant bien connus de contraction des longueurs et de dilatation du temps.

5.5.2 Dilatation du temps

Le diagramme ci-contre illustre l'écoulement du temps dans les deux référentiels. Il montre, sous forme de traits équidistants, le défilement des unités de temps (disons des secondes, ce qui dans le « temps espace » correspond à une distance de 300.000 km). Les traits de l'axe du temps T'_i représentent le passage des secondes sur une horloge au repos dans R' . Admettons que l'on mesure le temps T'_{i0} pour la durée d'un phénomène avec cette horloge dans R' . Comme les coordonnées du référentiel R sont obtenues par les transformations correspondant à la rotation d'angle θ_i , la durée T_{i0} dans le référentiel R correspondant à la durée T'_{i0}



"relativité" du temps

dans R' est donnée par la projection d'angle θ_i de la longueur T'_{i0} , ce qui donne $T_{i0} = T'_{i0} \cos \theta_i$. On voit donc que la durée d'un même phénomène apparaît différente dans les deux référentiels. Pour mettre en évidence cette différence, le diagramme montre également les traits correspondant au passage des secondes s'écoulant dans R' mais mesurées à l'aide d'une horloge au repos dans le référentiel R . On constate que les secondes de R' apparaissent plus courtes que les secondes de R . On

exprime cette différence de mesure du temps en disant qu'il y a « *relativité* » du temps. Le temps n'est plus une grandeur absolue qui s'écoule partout de la même manière dans l'univers.

Voyons maintenant ce que donnent en grandeurs réelles les résultats obtenus en variables z, T_i . Nous avons vu que les durées imaginaires dans les deux référentiels sont liées par la relation :


$$T_{i0} = T'_{i0} \cos \theta_i$$

Si on remplace le temps imaginaire T_i par iT , on obtient : ~~$iT_0 = iT'_0 \cos \theta_i$~~

On exprime ensuite le cosinus en fonction de la tangente : $T_0 = T'_0 \frac{1}{\sqrt{1 + \operatorname{tg}^2 \theta_i}}$

Et en se rappelant que $\operatorname{tg} \theta_i = V_i = -iV$ on trouve : $T_0 = T'_0 \frac{1}{\sqrt{1 - V^2}}$

Finalement en introduisant la vitesse dimensionnelle $v = cV$ et le temps $t = T/c$ on obtient la formule de transformation des durées observées lors du passage d'un référentiel à l'autre.



$$t_0 = t'_0 \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} > t'_0$$

Dans cette formule, le temps t'_0 représente la durée d'un évènement ayant lieu au repos dans le référentiel R' qui lui-même est en mouvement rectiligne uniforme à la vitesse v dans le référentiel R dans lequel on mesure un temps t_0 pour le même évènement. La formule indique que la durée mesurée dans le référentiel en mouvement est toujours plus longue que la durée mesurée dans le référentiel au repos. C'est ce que l'on appelle le phénomène de « *dilatation du temps* ».

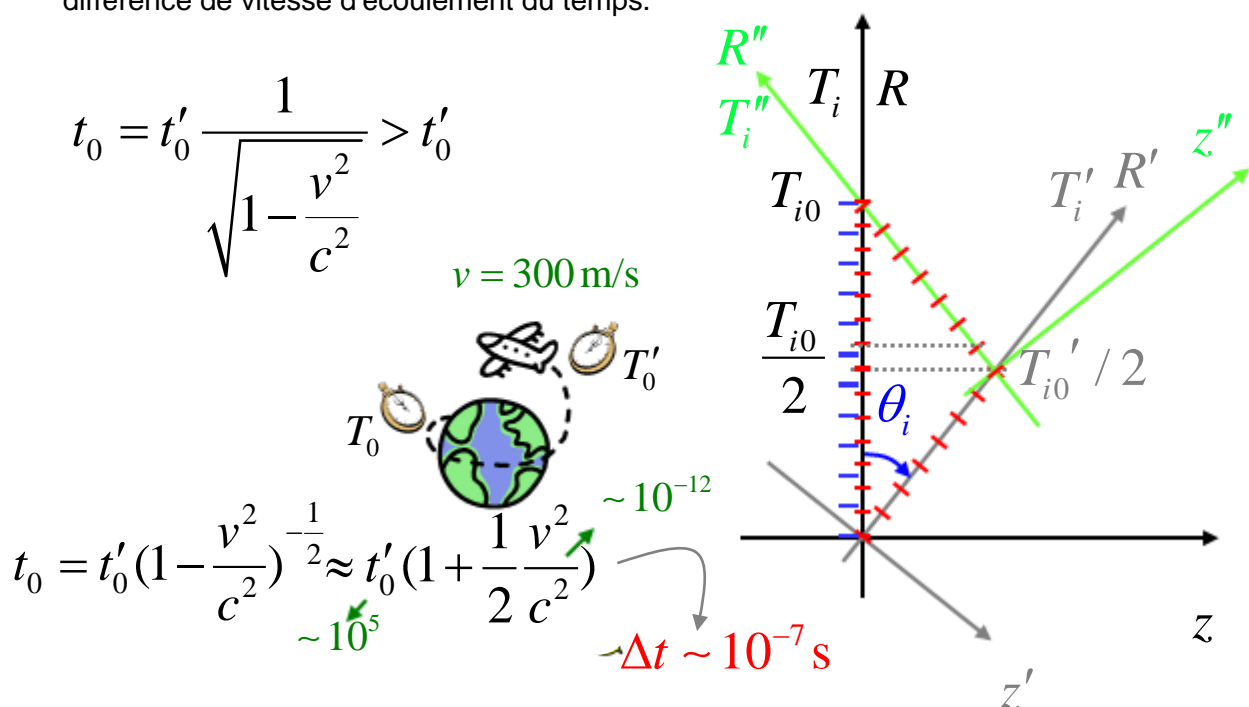
Remarque : Notez que dans l'espace-temps complexe la formule indiquait une contraction de la durée au travers de la formule $T_{i0} = T'_{i0} \cos \theta_i$. Mais il ne faut bien sûr pas oublier que l'angle θ_i est imaginaire et que, dès lors, le cosinus impliqué dans la projection de la durée T'_0 est en fait un cosinus hyperbolique plus grand que l'unité, d'où le fait que l'on a bien affaire à une dilatation et non à une contraction. La représentation graphique est donc à considérer avec prudence mais son grand avantage ici est que le phénomène de dilatation du temps y apparaît clairement

comme étant dû au fait que les dimensions temporelles et spatiales de deux référentiels en mouvement relatif ont des projections non-nulles entre elles : le temps de R' se projette sur l'espace R (le temps devient de l'espace) et l'espace de R' se projette sur le temps de R (l'espace devient du temps) et vice versa.

Bien sûr, il s'agit d'un phénomène réciproque car les deux référentiels sont équivalents et jouent des rôles parfaitement symétriques. Ainsi, si on mesurait, pour un phénomène donné, une durée t_0 avec une horloge au repos dans R , on mesurerait un temps plus long $t'_0 = t_0 / [1 - (v/c)^2]^{1/2}$ pour le même phénomène mesuré à partir de R' . Ceci apparaît évident sur la représentation graphique dans l'espace-temps complexe (z, T_i) donnée plus haut.

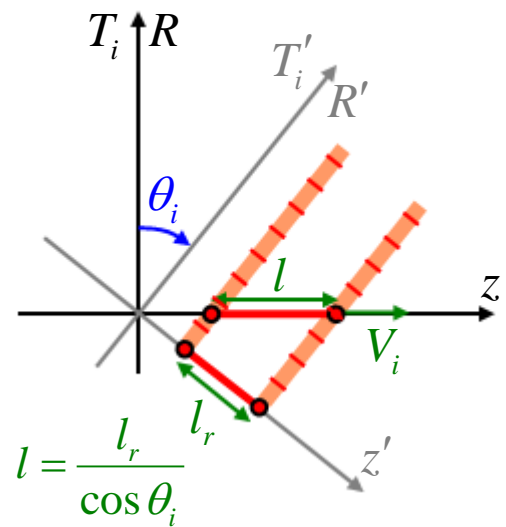
Ce phénomène de dilatation du temps a pu être vérifié expérimentalement de plusieurs manières. Une des expériences de vérification a consisté à mesurer la différence d'indication de deux horloges de haute précision disposées dans des référentiels en mouvement relatif. Une horloge était embarquée dans un avion et l'autre restait au repos dans le référentiel de la terre. L'avion a voyagé pendant une bonne vingtaine d'heures (soit à peu près 10^5 secondes) à une vitesse de l'ordre de 300 m/s. Cette vitesse étant beaucoup plus faible que celle de la lumière, le calcul de la dilatation du temps peut se faire dans une approximation du premier ordre en v^2/c^2 , ce qui donne une différence entre t_0 et t'_0 d'à peu près 10^{-7} s. Cette différence de temps correspondait bien à la différence des temps affichés par les deux horloges à l'arrivée de l'avion.

Le diagramme ci-dessous schématise, en variables z, T_i , un aller-retour de l'avion. Le référentiel R' correspond à l'aller et le référentiel R'' correspond au retour de l'avion à la vitesse opposée à celle de l'aller. On visualise très bien sur ce diagramme la différence de vitesse d'écoulement du temps.



5.5.3 Contraction des longueurs

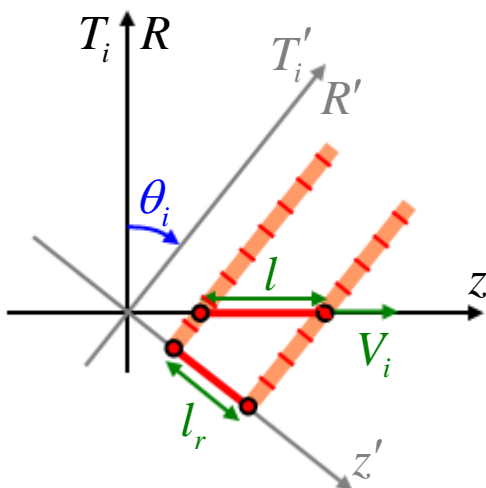
Le diagramme ci-contre montre une tige de longueur l_r au repos dans le référentiel R' . Les lignes d'univers des deux extrémités de la tige sont représentées avec le défilement des secondes comme il apparaîtrait sur deux horloges synchronisées fixées aux extrémités de cette tige. Pour voir comment sera vue la tige à partir du référentiel R , il suffit de considérer les intersections de ses lignes d'univers avec l'axe z . Ces intersections donnent l'état de la tige au temps $T_i = 0$ dans le référentiel R . On constate que la tige a une longueur différente donnée par $l = l_r / \cos \theta_i$. On en conclut que les longueurs apparaissent différentes d'un référentiel à l'autre. Il y a « *relativité* » des longueurs.



"relativité" des longueurs
et de la simultanéité

De même, on voit que les horloges n'apparaissent pas synchronisées dans le référentiel R . Le temps à une extrémité est en avance (ou en retard) par rapport au temps à l'autre extrémité. Des événements qui étaient simultanés dans R' (p. ex., les clics des horloges) ne le sont plus dans R . Il y a « *relativité* » de la simultanéité.

La formule $l = l_r / \cos \theta_i$ n'est pas très utile en tant que telle pour quantifier les choses en pratique puisque l'angle θ_i est imaginaire. Mais en exprimant le cosinus en fonction de la tangente et en introduisant la relation $\operatorname{tg} \theta_i = -iv / c$, on obtient la formule de modification de la longueur en variables réelles : $l = l_r [1 - (v/c)^2]^{1/2}$. C'est la formule de « *contraction des longueurs* » ; elle montre que la longueur l d'un objet mesurée dans un référentiel dans lequel cet objet est en mouvement est plus courte que sa longueur l_r mesurée dans le référentiel dans lequel il est au repos.

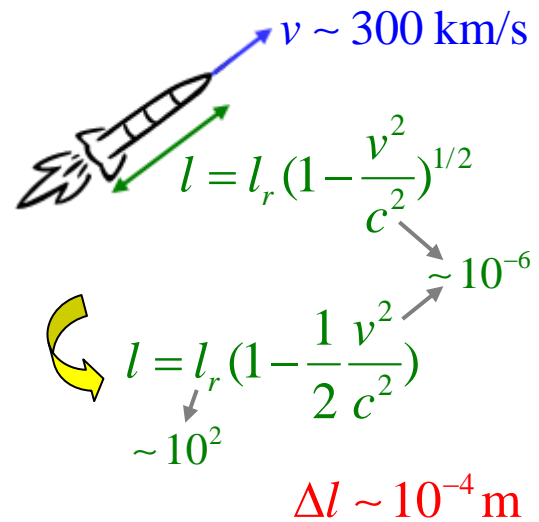


$$l = \frac{l_r}{\cos \theta_i} \quad \Rightarrow \quad l = l_r \sqrt{1 + \operatorname{tg}^2 \theta_i}$$

$$\operatorname{tg} \theta_i = iV \quad \Rightarrow \quad l = l_r \sqrt{1 - V^2} < l_r$$

$$l = l_r \sqrt{1 - \frac{v^2}{c^2}}$$

Pour illustrer la formule de contraction des longueurs, considérons une fusée de 100m de long voyageant à 300 km/s dans le référentiel de la terre. La mesure de longueur de cette fusée à partir de la terre (si elle était faisable) donnerait, d'après un calcul au premier ordre en v^2/c^2 , une longueur plus courte de 10^{-4} m (cette variation de longueur est très petite malgré la grande vitesse de la fusée et il n'est pas réaliste de vouloir en faire la vérification expérimentale).



$v \sim 300 \text{ km/s}$

$l = l_r \left(1 - \frac{v^2}{c^2}\right)^{1/2}$

$\sim 10^{-6}$

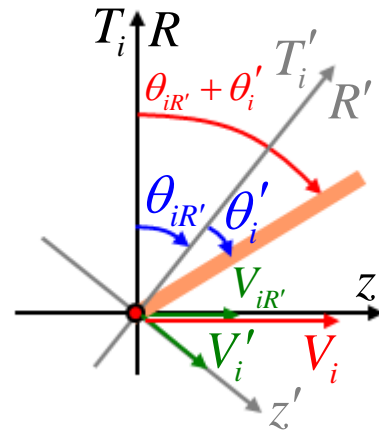
$l = l_r \left(1 - \frac{1}{2} \frac{v^2}{c^2}\right)$

$\sim 10^2$

$\Delta l \sim 10^{-4} \text{ m}$

5.5.4 Loi de composition des vitesses

Comme nous l'avons déjà longuement discuté, la loi d'addition des vitesses obtenue dans le cadre de la relativité galiléenne, n'est pas compatible avec les équations de Maxwell qui indiquent que la vitesse de la lumière est la même dans tout référentiel. Il est donc évident que dans le cadre de la relativité restreinte d'Einstein, on ne retrouve pas cette formule d'addition des vitesses.



Le diagramme ci-dessus montre la ligne d'univers d'une particule en mouvement uniforme dans le référentiel R' qui est lui-même en mouvement uniforme dans le référentiel R . Si la vitesse de la particule dans R' est v' et que la vitesse de R' dans R est dirigée dans le même sens et vaut $v_{R'}$, la loi d'addition des vitesses de la relativité galiléenne donnerait dans le référentiel R la vitesse $v = v' + v_{R'}$. Mais le diagramme montre clairement que ce n'est pas cette vitesse qui est obtenue dans R en réalité car le mouvement uniforme ne correspond pas à une simple translation de l'axe z mais bien à une rotation. Quand on somme des vitesses dans le cadre de la relativité restreinte, on somme des angles.

Dans les variables z, T_i , la vitesse v' de la particule dans R' correspond à une ligne d'univers inclinée d'un angle θ'_i . La relation entre cet angle et la vitesse réelle normalisée $V' = v'/c$ est donnée par la relation $\text{tg} \theta'_i = V'_i = -iV'$. De même, la vitesse relative entre les deux référentiels $v_{R'}$ correspond à une inclinaison relative d'un angle $\theta_{iR'}$ entre ces deux référentiels. Cet angle est lié à la vitesse relative adimensionnelle $V_{R'} = v_{R'}/c$ par la relation $\text{tg} \theta_{iR'} = V_{iR'} = -iV_{R'}$.

Sur cette base, nous pouvons calculer la vitesse v de la particule dans R car celle-ci est donnée par l'inclinaison de sa ligne d'univers dans R , soit $\theta_{iR'} + \theta'_i$ (voir le schéma ci-dessus). En se servant de la formule d'addition des angles pour la tangente, on obtient facilement l'expression de la vitesse v en fonction des vitesses v' et $v_{R'}$.

$$V_i = \operatorname{tg}(\theta_{iR'} + \theta'_i) = \frac{\operatorname{tg} \theta_{iR'} + \operatorname{tg} \theta'_i}{1 - \operatorname{tg} \theta_{iR'} \operatorname{tg} \theta'_i} = \frac{V_{iR'} + V'_i}{1 - V_{iR'} V'_i}$$

$$V_i = \frac{-iV_{R'} - iV'}{1 + V_{R'} V'} = -iV$$

$$V = \frac{V_{R'} + V'}{1 + V_{R'} V'} \quad \Rightarrow \quad v = \frac{v_{R'} + v'}{1 + v_{R'} v' / c^2}$$

C'est la formule de « *composition* » des vitesses de la relativité restreinte. Cette formule montre, comme attendu, que les vitesses ne s'additionnent pas simplement. En particulier, la composition des vitesses donnent une vitesse plus faible que l'addition des vitesses.

Il est intéressant de voir ce qui se passe si la particule est un photon. Dans ce cas, la vitesse v' est la vitesse de la lumière, soit $v' = c$. Le calcul de la vitesse du photon dans R peut s'obtenir en appliquant la formule de composition des vitesses :

$$\text{Soit } v' = c \quad \Rightarrow \quad v = \frac{v_{R'} + c}{1 + v_{R'} / c} = c$$

Comme il fallait s'y attendre, on trouve la vitesse c . Ceci est un résultat essentiel car il montre que les équations de Maxwell (ou l'équation d'onde EM) sont effectivement valables dans tous les référentiels inertiels. La notion d'éther n'a plus de raison d'être.

Ce résultat montre également que la vitesse de la lumière apparaît comme une vitesse limite que l'on ne peut dépasser. Ceci est cohérent avec l'expression de la vitesse en termes de la variable θ , nous avons vu en effet que la vitesse adimensionnelle est donnée par la tangente hyperbolique de cette variable, soit $V = \text{th}\theta$. Sachant que la vitesse réelle est donnée par la relation $v = cV$, on peut écrire $v = c \text{th}\theta$. Or la fonction tangente hyperbolique n'a pas de valeur supérieure à l'unité, ce qui implique que la vitesse d'une particule ne peut être supérieure à la vitesse de la lumière.

Remarque 1 : Notez que cette notion de vitesse limite ne peut être appréhendée avec la représentation graphique dans l'espace-temps complexe. Ceci est dû au fait que l'angle θ_i est imaginaire et que sa représentation graphique ne peut pas être interprétée directement. En l'occurrence, on ne voit pas comment l'angle θ_i peut être limité à une valeur finie (rien n'empêche a priori de prendre des angles aussi grands que l'on veut). On trouve là clairement les limites de la représentation graphique dans l'espace-temps complexe.

Remarque 2 : Notez finalement que si les vitesses impliquées dans la loi de composition des vitesses sont faibles alors on retrouve en première approximation la loi d'addition galiléenne des vitesses. Comme les vitesses impliquées dans notre vie de tous les jours sont en général beaucoup plus faibles que celle de la lumière, cette dernière loi peut être appliquée sans problème (mais il est bon d'avoir à l'esprit qu'il s'agit d'une approximation).

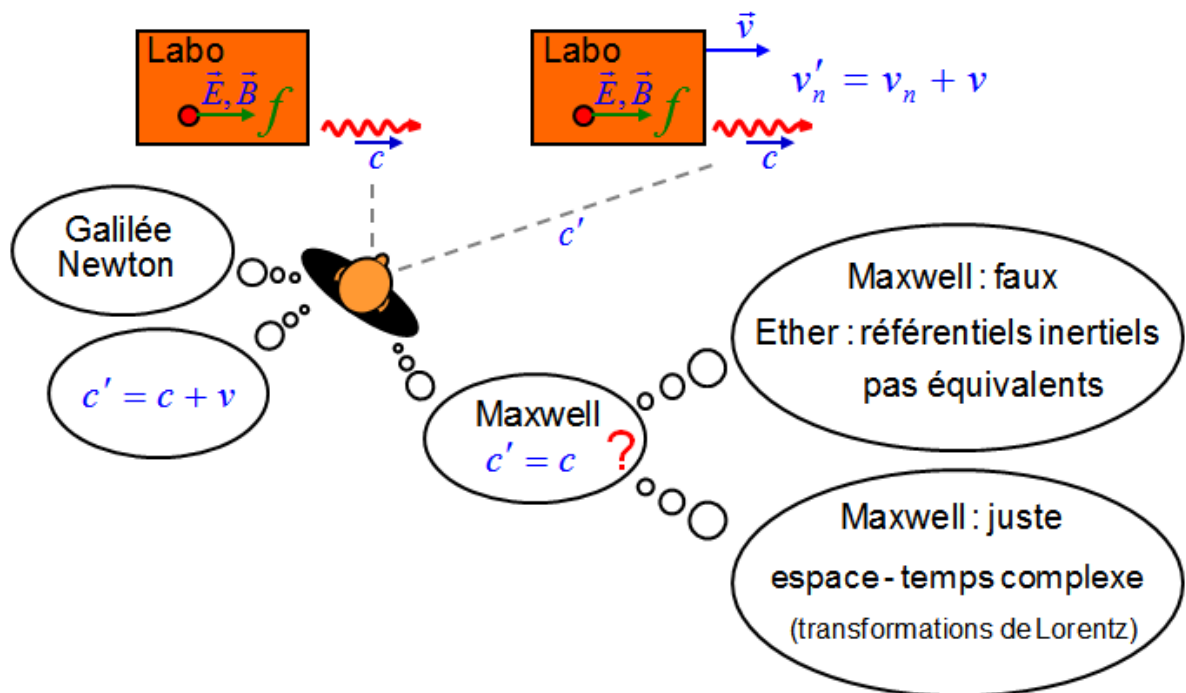
$$\text{Soit } v_{R'}, v' \ll c \quad \Rightarrow \quad v = \frac{v_{R'} + v'}{1 + \cancel{v_{R'} v' / c^2}} \approx v_{R'} + v'$$

5.6 Synthèse de cinématique relativiste

Le schéma ci-dessous récapitule le cheminement qui a mené la communauté scientifique au concept de relativité restreinte. Le cheminement part du principe de relativité de Galilée et Newton : tous les référentiels inertiels sont équivalents, les lois physiques y ont les mêmes expressions. On s'attend donc à ce que, à l'image des lois de Newton, les équations de Maxwell soient les mêmes dans tous les référentiels inertiels. Et c'est bien là que le problème se pose. Pour schématiser la situation, j'ai représenté ci-dessous un observateur observant deux laboratoires en mouvement relatif uniforme à la vitesse v . Dans ces deux laboratoires, on effectue des expériences identiques avec des charges électriques en mouvement qui émettent des ondes électromagnétiques (symbolisées par un photon sur le schéma). Puisqu'on s'attend à ce que les référentiels soient équivalents, la vitesse des photons émis est la même dans les deux laboratoires, elle vaut c comme l'imposent les équations de Maxwell. L'observateur qui est au repos dans le référentiel du laboratoire de gauche observe, à gauche donc, un photon qui avance à la vitesse c . Puisque la vitesse du photon de droite est c dans le référentiel de droite, l'observateur se dit que, d'après la

loi d'addition des vitesses de la relativité de Galilée, la vitesse de ce photon doit être $c' = c + v$. Il est alors confronté à une difficulté puisque s'il revient aux équations de Maxwell, il se dit que la vitesse du photon de droite ne peut être que c puisqu'il se propage dans les mêmes conditions que le photon de gauche.

Pour sortir de cette difficulté, l'observateur a un choix. Ou bien il se dit que les équations de Maxwell sont fausses et qu'elles doivent être complétées ou modifiées pour tenir compte d'un éther supportant les ondes électromagnétiques de façon à retomber sur la loi d'addition des vitesses. Dans l'expérience ci-dessous, l'éther pourrait, par exemple, être au repos avec l'observateur et la vitesse qu'il mesure serait c sans discussion pour les deux photons, mais les équations de Maxwell dans le référentiel de droite devraient avoir une autre forme que celles du référentiel de gauche puisque la vitesse de la lumière y serait plus petite, elle vaudrait $c - v$. En faisant ce choix, l'observateur perd la notion de relativité du mouvement uniforme. Tous les référentiels inertiels ne sont plus équivalents. Il existe un seul référentiel absolu dans la nature, celui attaché à l'éther.

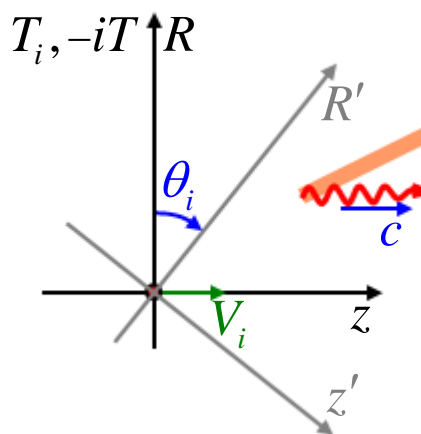


Son deuxième choix est de se dire, comme Einstein l'a fait, que les équations de Maxwell sont justes et que tous les référentiels inertiels sont équivalents. Ce qui n'est pas un choix très rationnel, a priori, puisque cela signifie que les équations de Maxwell sont effectivement identiques dans les référentiels de droite et de gauche et que donc les photons sont émis avec la même vitesse c dans ces deux référentiels et qu'en même temps ces deux vitesses devraient encore être égales vue d'un seul référentiel, c'est-à-dire $c' = c$ pour l'observateur. Ce choix a néanmoins été fait par Einstein en interprétant les transformations de Lorentz comme traduisant la réalité physique de notre univers : l'espace et le temps forment un espace-temps complexe dans lequel le passage d'un référentiel inertiel à un autre est décrit en termes de rotation des axes de temps et d'espace. L'expérience a montré que c'est ce choix qui est le bon, même s'il

mène à des phénomènes pour le moins contre-intuitifs tels que la contraction des longueurs, la dilatation du temps, et la relativité de la simultanéité...

L'espace-temps complexe dans lequel on vit est schématisé ci-dessous à deux dimensions, comme on l'a considéré partout dans ce chapitre. En réalité, il y a bien sûr trois dimensions spatiales et l'espace-temps complexe a donc quatre dimensions.

Ce qui est important à retenir c'est que le temps constitue une dimension imaginaire par rapport à l'espace et que la rotation correspondant au mouvement relatif rectiligne uniforme de deux référentiels inertiels est à considérer avec des angles imaginaires, ce qui en rend la représentation graphique impossible (ou tout au moins incorrecte). Cet espace-temps complexe a été formalisé peu après les travaux d'Einstein par le physicien allemand Herman Minkowski (qui fut pendant un temps le professeur de physique d'Einstein).



Espace - temps complexe de Minkowski

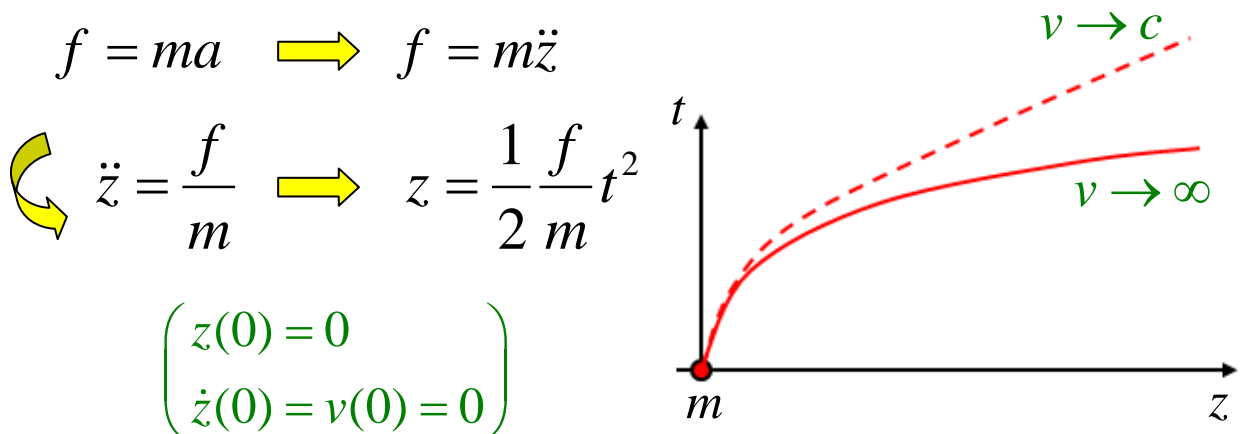
Notons pour terminer que dans l'espace-temps de Minkowski, les choses apparaissent simples en ce qui concerne l'électromagnétisme (les équations de Maxwell sont invariantes pour tout changement de référentiels inertiels et dans tous les référentiels elles ont donc la forme que l'on a étudié dans ce cours) mais en revanche les équations de la mécanique prennent une allure toute différente de celles de la mécanique Newtonienne. Les notions de masse, de force, d'impulsion (quantité de mouvement) et d'énergie ont du être complètement revues. Ce sont les thèmes que nous allons aborder brièvement dans la section suivante consacrée à la dynamique relativiste.

5.7 Dynamique relativiste

Nous allons maintenant voir comment le mouvement uniformément accéléré est décrit dans le cadre de la relativité restreinte. Les quelques développements qui suivent constituent la base de ce que l'on appelle la dynamique relativiste. Ils nous permettront, en particulier, de comprendre l'origine de la fameuse formule $E = mc^2$. Pour cela je vous propose de commencer par une brève analyse du mouvement rectiligne uniformément accéléré (MRUA) en dynamique classique Newtonienne.

5.7.1 Rappel de dynamique Newtonienne

Considérons le mouvement d'une particule assimilable à une masse ponctuelle m subissant une force constante f et se déplaçant le long de l'axe z d'un repère cartésien. En intégrant deux fois par rapport au temps la deuxième loi de Newton $f = ma$ pour un MRUA le long de l'axe z , soit $\ddot{z} = f/m$, on constate que la position de la masse z est proportionnelle au carré du temps écoulé depuis l'origine $z = 0$ (on suppose que la particule est partie de l'origine à vitesse nulle). Ce résultat est représenté graphiquement ci-dessous par la courbe en trait rouge plein (avec l'axe du temps absolu en position verticale, comme précédemment).



Cette évolution parabolique de la position, $z = ft^2 / (2m)$, est bien entendu associée à une évolution linéaire de la vitesse, $v = ft / m$, ce qui signifie que la vitesse tend vers l'infini lorsque le temps tend vers l'infini ou, en d'autres termes, la vitesse de la masse m n'a pas de valeur limite.

$$v = \dot{z} = \frac{f}{m} t$$

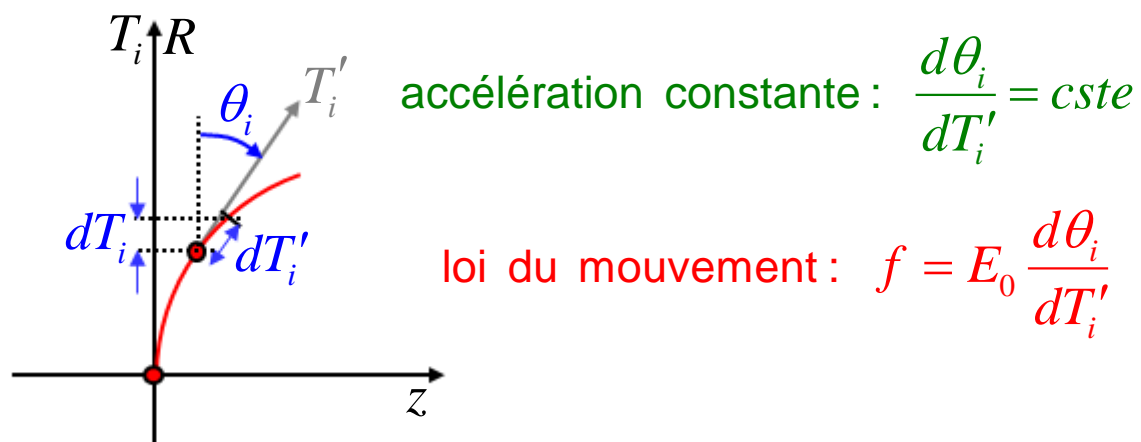
Or, on a vu plus haut que dans le cadre de la relativité restreinte la vitesse peut, au plus, atteindre la valeur c (une vitesse supérieure à c n'a tout simplement pas de sens physique ou plus exactement pas de sens « géométrique » si on prend le point de vue de l'espace-temps de Minkowski). Sur la représentation graphique du MRUA ci-dessus, la pente de la courbe ($v = dz / dt$) devrait donc avoir une valeur limite égale à c , comme suggéré par la courbe rouge pointillée. Cette contradiction entre la loi de

Newton et la relativité restreinte nous indique qu'il faut s'attendre à une différence fondamentale entre les dynamiques relativiste et Newtonienne ; en d'autres termes, il faut s'attendre à ce que la deuxième loi de Newton ne soit plus valable.

5.7.2 Accélération relativiste

Appréhender le MRUA en dynamique relativiste n'est pas chose facile car la particule massique est en mouvement à une vitesse variable et elle change donc continûment de référentiel inertiel, ce qui a pour conséquence que l'axe du temps propre de la masse T'_i s'incline de façon continue par rapport à l'axe du temps propre du laboratoire T_i . Autrement dit, l'angle θ_i que font les référentiels du laboratoire et de la particule (représenté sur le schéma ci-dessous) évolue continuellement.

Que devient donc la notion d'accélération de la loi de Newton $f = ma$ (et donc également la notion de force) si le temps n'est plus absolu ? Doit-on dériver deux fois la position par rapport à T_i ou par rapport à T'_i . Les réflexions d'Einstein sur le sujet ont été longues et complexes mais on peut néanmoins exprimer l'équivalent relativiste de l'accélération très simplement sur base du raisonnement suivant. Dans la représentation (z, T_i) une particule accélérée a une trajectoire courbe car elle change continuellement de référentiel. La courbure de la trajectoire est d'autant plus marquée que l'accélération est grande. Si la particule subit dans son référentiel une force constante elle va tout simplement tracer une trajectoire à courbure constante, c'est-à-dire un cercle. Pour traduire cela mathématiquement il suffit d'exprimer que la vitesse de variation de l'angle θ_i exprimée dans le temps propre de la particule T'_i est une constante. Une accélération constante correspond donc à une variation de l'angle θ_i à taux constant en termes du temps T'_i , soit $d\theta_i/dT'_i = cste$. On peut alors généraliser la notion de force en disant que la force « relativiste » est donnée par le produit d'une constante et de l'accélération relativiste, soit $f = E_0 d\theta_i/dT'_i$.



La constante E_0 caractérise la particule en mouvement, elle représente en quelque sorte la « résistance » de la particule au changement de référentiel. À une force f donnée, plus E_0 est grande, plus le changement continu de référentiel (représenté par

la dérivée $d\theta_i/dT_i'$ est lent. La constante E_0 est donc l'équivalent relativiste de la masse de la particule en dynamique Newtonienne. Il est important de remarquer que les dimensions de E_0 sont des énergies puisque f est en Newton, que le temps T_i a les dimensions d'une longueur et que θ_i est bien entendu sans dimension.

Si la vitesse de variation de l'angle θ_i est une constante dans le temps propre de la particule, il est évident que la trajectoire d'espace-temps (ou ligne d'univers) de la particule forme un cercle dans la représentation graphique ci-dessus (temps complexe T_i , angle complexe θ_i et espace réel z'). On peut facilement s'assurer de cela en exprimant la loi de la dynamique relativiste en fonction des coordonnées du référentiel du laboratoire (z, T_i) . Ce changement de référentiel est aisé car seul le temps intervient dans la loi et, comme l'indique le graphique ci-dessus, un accroissement du temps propre de la particule dT_i' correspond à un accroissement $dT_i = dT_i' \cos \theta_i$ du temps dans le référentiel au repos R . En remplaçant dT_i' par $dT_i/\cos \theta_i$, on obtient donc la loi du MRUA relativiste dans le référentiel au repos :

référentiel R :

$$dT_i = dT_i' \cos \theta_i \implies dT_i' = \frac{dT_i}{\cos \theta_i} \implies \boxed{f = E_0 \cos \theta_i \frac{d\theta_i}{dT_i}}$$

Pour obtenir l'équation de la ligne d'univers du MRUA, il suffit d'écrire la loi du MRUA ci-dessus sous la forme suivante :

$$\curvearrowright dT_i = (E_0 / f) \cos \theta_i d\theta_i$$

On introduit ensuite la définition de l'angle θ_i , soit $dz/dT_i = \text{tg}(\theta_i)$, dans laquelle on substitue l'expression de dT_i ci-dessus :

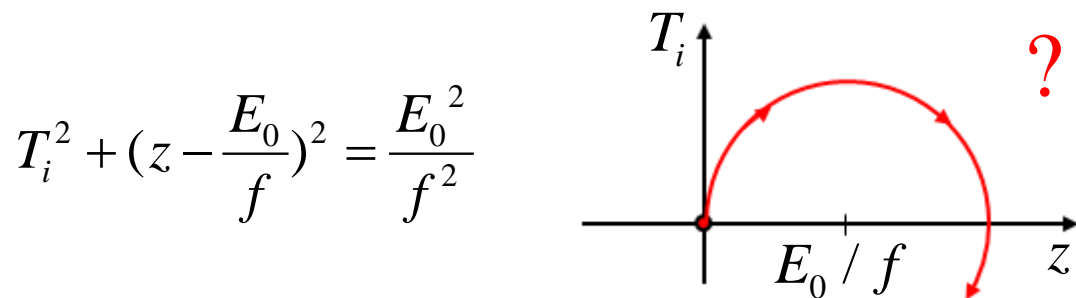
$$dz = \text{tg} \theta_i dT_i = \text{tg} \theta_i (E_0 / f) \cos \theta_i d\theta_i$$

$$\curvearrowright dz = (E_0 / f) \sin \theta_i d\theta_i$$

En intégrant les expressions finales de dT_i et dz avec la condition initiale de position et de vitesse nulle en $T_i = 0$, soit $z(T_i = 0) = 0$ et $\theta_i(T_i = 0) = 0$, on trouve :

$$\curvearrowright \begin{cases} T_i = (E_0 / f) \sin \theta_i \\ z = -(E_0 / f) \cos \theta_i + E_0 / f \end{cases}$$

Grâce à la relation $\sin^2 \theta_i + \cos^2 \theta_i = 1$, on élimine facilement θ_i dans ces deux dernières expressions et on constate que la ligne d'univers est un cercle de rayon E_0/f centré en $(z = E_0/f, T_i = 0)$, comme représenté graphiquement ci-dessous.

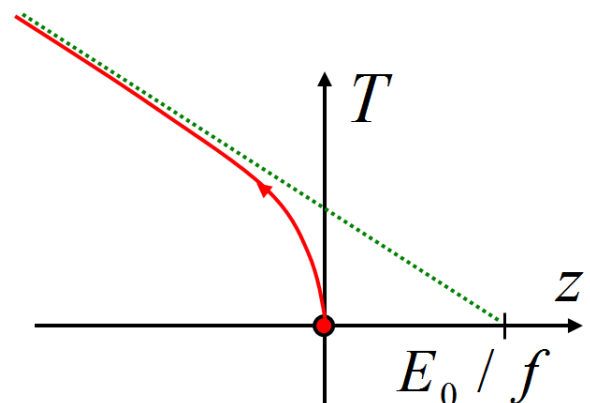


Le point d'interrogation sur le graphe de la ligne d'univers rappelle qu'il faut être attentif à l'interprétation de la représentation dans l'espace-temps complexe (z, T_i) . En effet, si notre temps réel était comme le temps T_i , la ligne d'univers circulaire ci-dessus montre qu'à partir d'un certain temps dans le référentiel au repos, soit le temps $T_i = E_0/f$ (point de pente horizontale du cercle ci-dessus) la particule disparaîtrait après s'être collisionnée avec une autre particule qui n'est en réalité qu'elle-même pour des temps propres plus lointains... De même, l'existence d'une pente négative montre la possibilité de remonter le temps, ce qui n'a jamais pu être observé dans notre univers et ceci est plutôt heureux car, comme on s'en rend compte ici, il serait bien chaotique...

Pour retrouver la véritable ligne d'univers il faut bien entendu introduire le temps réel à partir de la définition du temps complexe, ce qui conduit à l'équation d'une branche d'hyperbole dont la pente dz/dT_i , cette fois ne passe plus par l'infini, interdisant de la sorte le recul de la particule dans le temps (voir figure ci-dessous). On constate ainsi que le principe de causalité n'est pas remis en question par la relativité restreinte.

$$T_i = -iT \quad \Rightarrow \quad -T^2 + \left(z - \frac{E_0}{f}\right)^2 = \frac{E_0^2}{f^2}$$

De même, on constate que la vitesse (pente) maximale de la trajectoire, dz/dT , vaut 1 (pente de l'asymptote de l'hyperbole), ce qui correspond en grandeurs réelles à la vitesse c qui apparaît donc bien comme la vitesse limite que peut atteindre tout objet dans notre univers (même après une accélération arbitrairement longue, la vitesse n'atteint jamais tout à fait c). La ligne d'univers hyperbolique du MRUA est représentée sur le schéma ci-contre.



On constate sur le schéma ci-dessus que la ligne d'univers hyperbolique correspond à un déplacement du point vers les z négatifs, c'est-à-dire qu'elle correspond à une vitesse négative. Ceci est dû au fait que la représentation (z, T_i) ne peut être interprétée directement sur base de l'analyse des images qu'elle produit. Mathématiquement, l'inversion du signe des vitesses vient du fait que l'accélération positive dans la représentation (z, T_i) se traduit par une accélération négative dans la représentation (z, T) ; en effet, sachant que $\theta_i = -i\theta$ et $T_i = iT$, on vérifie facilement la relation $d\theta_i / dT_i' = -d\theta / dT'$. Par la suite, nous n'attacherons que peu d'importance à ce problème de signe.

5.8 Énergie et quantité de mouvement

Avant d'aborder les notions d'énergie et de quantité de mouvement relativistes, je propose ici de revoir ces notions dans leur version classique Newtonienne. Pour une particule libre de masse m subissant une force f , l'expression de l'énergie cinétique Newtonienne (non relativiste) s'obtient simplement en intégrant la seconde loi de Newton $f = ma$ par rapport à la coordonnée d'espace (ici z).

Dynamique Newtonienne : $f = ma$

L'intégrale en z de la force fournit le travail donné à la particule. Puisque la particule est libre (pas de potentiel liant) ce travail se transforme intégralement en énergie cinétique. Le bref développement ci-dessous nous montre donc que l'énergie cinétique a pour expression $E_C = (1/2)mv^2$.

$$E_C = \int_0^z f dz = m \int_0^z \frac{dv}{dt} dz = m \int_0^v v dv = \frac{1}{2}mv^2$$

De même, l'intégrale de la force par rapport au temps (ici t) donne l'expression de l'impulsion (ou la quantité de mouvement) $p = mv$.

$$p = \int_0^t f dt = m \int_0^t \frac{dv}{dt} dt = m \int_0^v dv = mv$$

Nous allons utiliser ces définitions de l'énergie cinétique et de la quantité de mouvement pour en obtenir les expressions relativistes. Pour cela nous partons, de la loi de la dynamique relativiste exprimée dans le référentiel au repos (je rappelle ici que cette loi est l'équivalent relativiste de la seconde loi de Newton).

Dynamique relativiste : $f = E_0 \cos \theta_i \frac{d\theta_i}{dT_i}$

Pour faciliter l'interprétation physique, il est plus commode de passer aux coordonnées réelles d'espace-temps. En appliquant le changement de variables définissant les variables complexes θ_i et T_i , nous pouvons facilement exprimer la dérivée première de la loi de la dynamique relativiste en variables réelles :

$$\text{où } \begin{cases} \theta_i = i\theta \\ T_i = -iT \end{cases} \quad \Rightarrow \quad \frac{d\theta_i}{dT_i} = -\frac{d\theta}{dT}$$

Et en se rappelant que le cosinus de l'angle imaginaire $\theta_i = i\theta$ donne le cosinus hyperbolique de θ , on peut écrire toute la loi en termes des coordonnées d'espace-temps réelles.

$$\cos \theta_i = \frac{1}{2}(e^{i\theta_i} + e^{-i\theta_i}) = \frac{1}{2}(e^{-\theta} + e^{\theta}) = \text{ch}(\theta)$$

Cela nous donne l'expression de la force relativiste en variables réelles :

$$f = E_0 \text{ch} \theta \frac{d\theta}{dT}$$

Notez qu'un changement de signe a été opéré artificiellement dans ce changement de variable. Celui-ci n'a aucune importance pratique ou conceptuelle car il provient du fait qu'une accélération négative dans l'espace-temps (z, T_i) correspond à une accélération positive dans l'espace-temps réel (z, T) (voir schémas des lignes d'univers circulaire et hyperbolique vues plus haut). La représentation (z, T_i) n'ayant pas de sens physique nous aurions pu éviter cette difficulté en disant au départ qu'une accélération réelle positive se traduit par une valeur négative de $d\theta_i / dT_i$ mais ce choix n'a pas été fait pour ne pas rentrer dans ces subtilités dès l'introduction à la dynamique relativiste. Pour la suite du développement nous considérerons donc, pour la simplicité, une accélération positive dans l'espace-temps réel (z, T) , conformément à l'expression de la force relativiste ci-dessus.

5.8.1 Énergie relativiste

A partir de la définition de la force relativiste nous pouvons aisément trouver l'expression de l'énergie cinétique d'une particule libre en exprimant simplement la conservation de l'énergie : le travail de la force f sur une distance z donnée, c'est-à-dire, l'intégrale de la force f de 0 à z est égal à l'énergie cinétique accumulée E_C .

$$f = E_0 \operatorname{ch} \theta \frac{d\theta}{dT}$$

$$E_C = \int_0^z f dz = E_0 \int_0^z \operatorname{ch} \theta \frac{d\theta}{dT} dz = E_0 \int_0^\theta \operatorname{ch} \theta \frac{dz}{dT} d\theta$$

Ici, en déplaçant le dénominateur dT d'un facteur à l'autre, on peut naturellement effectuer le changement de variables d'intégration $z \rightarrow \theta$.

Dans cette dernière intégrale on voit apparaître la vitesse dz/dT de la particule. Cette vitesse est bien entendu liée à la variable θ et c'est ce lien qui nous permettra de calculer l'intégrale. Pour retrouver la relation entre la vitesse dz/dT et θ , il suffit de se rappeler que $\theta = i\theta_i$ où θ_i est l'angle de la ligne d'univers dans l'espace-temps complexe (z, iT_i) , c'est-à-dire, $\operatorname{tg} \theta_i = dz/dT_i$. Sachant que $T_i = iT$, et que $\operatorname{tg} \theta_i = -i \operatorname{th} \theta$, on voit que la vitesse dz/dT n'est autre que la tangente hyperbolique de θ , soit $V = dz/dT = \operatorname{th} \theta$, comme on l'a déjà montré à la section 1.4.3.

$$\frac{dz}{dT} = V = \operatorname{th}(\theta)$$

Ce constat nous permet de calculer aisément l'intégrale de l'énergie :

$$\hookrightarrow E_C = E_0 \int_0^\theta \operatorname{ch} \theta \operatorname{th} \theta d\theta = E_0 \int_0^\theta \operatorname{sh} \theta d\theta$$

$$\hookrightarrow E_C = E_0 [\operatorname{ch} \theta]_0^\theta = E_0 (\operatorname{ch} \theta - 1)$$

Finalement, sur base de la relation suivante entre le cosinus et la tangente hyperbolique,

$$\text{ch}(\theta) = \frac{1}{\sqrt{1 - \text{th}^2(\theta)}}$$

Nous voyons que l'énergie cinétique est une fonction relativement simple de la vitesse $V = \text{th} \theta$.

$$\hookrightarrow E_C = E_0 \frac{1}{\sqrt{1 - \text{th}^2(\theta)}} - E_0 = E_0 \frac{1}{\sqrt{1 - V^2}} - E_0$$

Cette expression est néanmoins plus complexe que l'expression non relativiste de l'énergie cinétique, soit $E_C = (1/2)mv^2$. Bien entendu, la validité de la relativité restreinte impose que, dans la limite des vitesses faibles par rapport à la vitesse de la lumière ($V \ll 1$), cette expression plus complexe se rapproche de l'expression non relativiste de l'énergie cinétique. Si ce n'était pas le cas, cela signifierait que la relativité restreinte n'est pas compatible avec la dynamique Newtonienne, ce qui ne serait pas acceptable puisque la validité de la dynamique Newtonienne à faible vitesse ne fait aucun doute. L'équivalence des expressions relativiste et non relativiste de l'énergie cinétique à basse vitesse ($V \ll 1$) peut se vérifier aisément à l'aide du bref développement suivant basé sur l'approximation du premier ordre de la racine, soit $(1+x)^n = 1 + nx$ (si $x \ll 1$):

$$\text{soit } V = \frac{v}{c} \ll 1 \quad \Rightarrow \quad E_C \approx E_0 \left(1 + \frac{1}{2} \frac{v^2}{c^2}\right) - E_0$$

Pour que cette expression de l'énergie cinétique relativiste soit identique à celle de l'énergie cinétique non relativiste, il faut évaluer la constante E_0 à mc^2 .


$$\hookrightarrow E_C \approx \frac{1}{2} \overset{E_0 = mc^2}{E_0} \frac{v^2}{c^2} = \frac{1}{2} m v^2$$

En conclusion, pour assurer la validité de la relativité restreinte (c.-à-d., pour que la dynamique relativiste devienne identique à la dynamique Newtonienne dans la limite des vitesses arbitrairement faibles), la constante E_0 (qui au départ n'était qu'une simple constante représentant l'inertie de la particule en dynamique relativiste) doit être égalée à la masse de la particule multipliée par le carré de la vitesse de la lumière. La vitesse de la lumière étant une constante universelle, on peut dire que E_0 représente la masse de la particule exprimée dans d'autres unités, en l'occurrence, en unités d'énergie. Nous verrons par la suite que cette énergie E_0 représente une véritable énergie physiquement accessible (alors que la masse m en dynamique Newtonienne ne représente qu'un nombre abstrait caractérisant l'inertie de la particule).

En remplaçant E_0 par mc^2 et V par v/c dans l'expression de l'énergie cinétique relativiste, nous en obtenons l'expression suivante.

$$E_C = E_0 \frac{1}{\sqrt{1-V^2}} - E_0 = mc^2 \underbrace{\frac{1}{\sqrt{1-v^2/c^2}}}_{E} - mc^2$$

Nous verrons par la suite que la seule façon cohérente d'interpréter ce résultat est de considérer que le premier terme du membre de droite est l'énergie totale de la particule ; on la notera E . De la sorte l'énergie totale E de la particule apparaît comme étant la somme de son énergie cinétique et de l'énergie mc^2 , soit :




$$E = E_C + mc^2$$

Pour comprendre comment Einstein a été amené à adopter cette expression de l'énergie totale de la particule libre, il est intéressant de revenir en arrière, à l'expression obtenue pour l'intégrale de l'énergie :

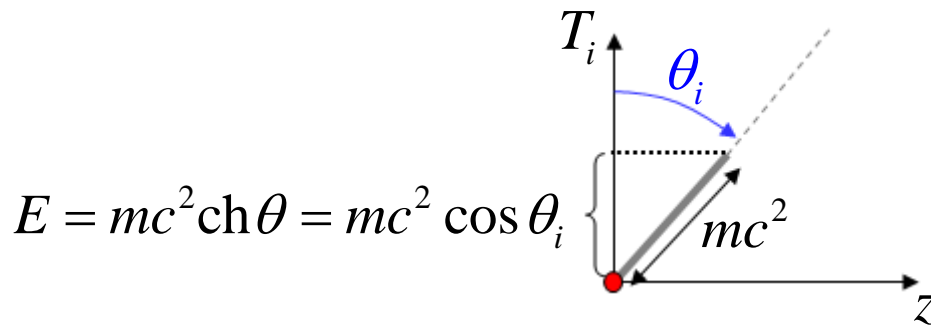
$$E_C = E_0 (\text{ch} \theta - 1) = mc^2 (\text{ch} \theta - 1) = mc^2 \text{ch} \theta - mc^2$$

Dans cette expression on voit que l'énergie totale n'est rien d'autre que :

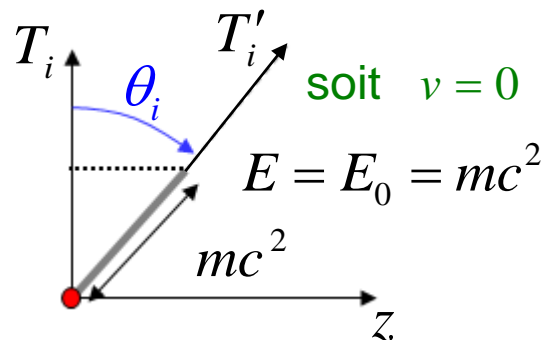


$$E = mc^2 \text{ch}(\theta)$$

Dans notre interprétation géométrique de la relativité restreinte cette expression représente la projection sur l'axe des temps d'un vecteur de longueur $E_0 = mc^2$ et dont la direction est donnée par la ligne d'univers de la particule. Ceci est représenté ci-dessous dans la représentation géométrique d'espace-temps complexe (z, T_i) dans laquelle le cosinus hyperbolique de θ devient le cosinus de θ_i .



Dans le référentiel de la particule (voir schéma ci-contre), la ligne d'univers est par définition le long de l'axe des temps (ici T'_i) et la projection devient donc la norme du vecteur, soit la valeur E_0 elle-même, ce qui correspond à une énergie totale $E = E_0 = mc^2$; c'est ce que l'on appelle l'énergie au repos de la particule. L'énergie de la particule dans son référentiel ($v=0$) est illustrée ci-contre. L'énergie au repos d'une particule caractérise cette particule au même titre que la masse caractérise une particule en dynamique Newtonienne.



Pour une particule en mouvement [ici dans le repère (z, T_i)], la formule $E = E_C + mc^2$ indique que l'énergie totale de la particule est la somme de son énergie cinétique E_C et de son énergie au repos E_0 .

Remarque : Notez que sur les graphiques ci-dessus l'énergie totale apparaît plus petite que l'énergie au repos. Ceci est évidemment dû à la représentation d'espace-temps complexe : θ_i est un angle imaginaire et le cosinus (toujours inférieur à l'unité) de la projection sur l'axe des temps est donc en fait un cosinus hyperbolique toujours supérieur à l'unité.

5.8.2 Impulsion relativiste

Voyons maintenant l'expression de l'impulsion (ou quantité de mouvement) relativiste. Pour cela reprenons l'expression de la force relativiste donnée ci-contre. Je rappelle la signification de cette expression : une telle force constante agissant sur une particule engendre, vue de son propre repère, une accélération constante.

$$f = E_0 \operatorname{ch} \theta \frac{d\theta}{dT}$$

Comme en dynamique Newtonienne, l'impulsion est donnée par l'intégrale de la force dans le temps. En y remplaçant le temps t par le temps $T = ct$, cette intégrale peut être aisément calculée en variable θ .

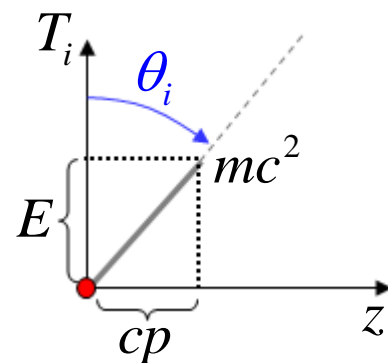
$$p = \int_0^t f dt = E_0 \int_0^t \operatorname{ch} \theta \frac{d\theta}{dT} dt = E_0 \int_0^\theta \operatorname{ch} \theta \frac{d\theta}{dT} \frac{dT}{c}$$

$T = ct \longrightarrow dt = \frac{dT}{c}$

On constate ainsi que l'impulsion est proportionnelle au sinus hyperbolique de θ .

$$p = \frac{E_0}{c} \int_0^\theta \operatorname{ch} \theta d\theta = \frac{E_0}{c} \operatorname{sh} \theta = mc \operatorname{sh} \theta$$

Cette expression montre que l'impulsion multipliée par la vitesse de la lumière c , soit $cp = mc^2 \operatorname{sh} \theta$, est égale à la projection sur l'axe d'espace z , du vecteur de module $E_0 = mc^2$ dont la direction est celle de la ligne d'univers de la particule. Ainsi, en particulier, dans son propre référentiel une particule a une impulsion nulle, $p = 0$, comme en dynamique Newtonienne.



En exprimant le sinus hyperbolique en fonction de la tangente hyperbolique qui est elle-même égale à la vitesse adimensionnelle $V = \operatorname{th} \theta = v/c$, on obtient l'expression de l'impulsion p en fonction de la vitesse v de la particule.

$$\operatorname{sh}(\theta) = \frac{\operatorname{th}(\theta)}{\sqrt{1 - \operatorname{th}^2(\theta)}}$$

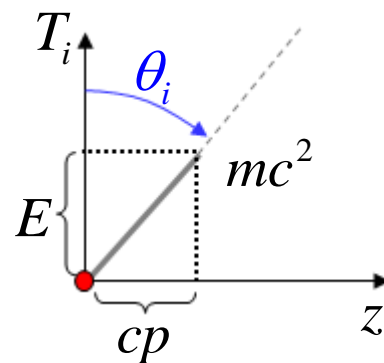
$$\curvearrowright p = mc \frac{v/c}{\sqrt{1-v^2/c^2}} = \frac{mv}{\sqrt{1-v^2/c^2}}$$

On constate bien sûr qu'à faible vitesse ($v \ll c$), l'impulsion prend sa forme « Newtonienne » $p = mv$ (au premier ordre en le rapport v/c).

$$\text{soit } V = \frac{v}{c} \ll 1 \quad \Rightarrow \quad p \approx mv \left(1 + \cancel{\frac{1}{2} \frac{v^2}{c^2}}\right) \approx mv$$

Nous verrons en séance d'exercices que ces expressions relativistes de l'impulsion et de l'énergie ont d'énormes conséquences sur le comportement des particules se déplaçant à des vitesses proches de celle de la lumière.

Pour terminer, il est intéressant de calculer le module du vecteur E_0 dans l'espace-temps réel. Dans la représentation d'espace-temps complexe ce module s'obtient en prenant la somme du carré des composantes du vecteur, soit : $E_0^2 = E_0^2 \cos^2 \theta_i + E_0^2 \sin^2 \theta_i$. En exprimant cette relation en variable $\theta = i\theta_i$, on obtient : $E_0^2 = E_0^2 \text{ch}^2 \theta - E_0^2 \text{sh}^2 \theta$ (notez que, sachant que $\text{ch}^2 \theta - \text{sh}^2 \theta = 1$, cette relation est triviale). On retrouve dans cette égalité l'expression de l'énergie et l'impulsion qui dès lors apparaissent liées par la relation $E^2 - c^2 p^2 = m^2 c^4$.



$$\text{ch}^2(\theta) - \text{sh}^2(\theta) = 1$$

$$\left. \begin{array}{l} E = mc^2 \text{ch}(\theta) \\ p = mc \text{sh} \theta \end{array} \right\} \quad \Downarrow \quad E^2 - c^2 p^2 = m^2 c^4$$

Cette relation est intéressante car elle permet d'introduire la notion de particule sans masse, qui est une des conséquences majeures de la théorie de la relativité restreinte. En effet, dans le cadre de la relativité restreinte une particule peut avoir une masse nulle, et en même temps une énergie $E = mc^2 \text{ch} \theta$, non nulle et une impulsion $p = mc \text{sh} \theta$ non nulle. Il suffit effectivement que les fonctions $\text{ch} \theta$ et $\text{sh} \theta$ prennent des valeurs infinies, ce qui est obtenu pour une vitesse égale à la vitesse de la lumière. En effet, si $v = c$ alors $V = \text{th} \theta = 1$, et donc $\theta = \infty$. Un exemple bien connu de particule sans masse est le photon.

5.8.3 Le photon, particule sans masse

D'après la relation $E^2 - c^2 p^2 = m^2 c^4$, si une particule est sans masse, $m = 0$, alors $E^2 = c^2 p^2$ ou autrement dit : $E/c = p$.

$$\begin{array}{l} E^2 - c^2 p^2 = 0 \\ \swarrow \\ E / c = p \end{array} \quad \left[\begin{array}{l} \theta = \infty \\ v = c \end{array} \right]$$

Il est intéressant de comparer cette dernière relation aux relations de de Broglie appliquées au photon.

$$\text{de Broglie : } \begin{cases} E = \hbar \omega \\ p = \hbar k \end{cases}$$

Pour le photon nous savons que $k = \omega/c$, ce qui permet tout de suite de vérifier la relation ci-dessus : $p = E/c$.

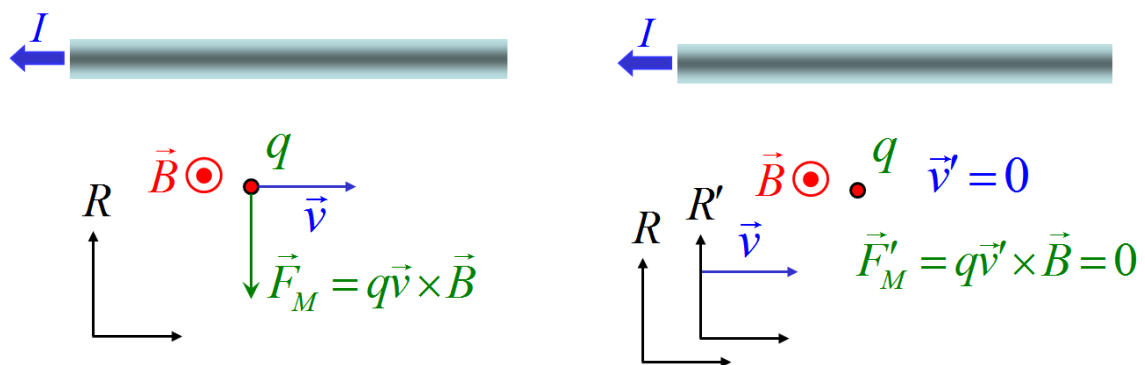
$$\text{le photon : } p = \hbar \frac{\omega}{c} = \frac{E}{c}$$

On voit donc que, dans le cadre de la relativité restreinte, le photon peut être considéré comme une particule au sens classique du terme dans la mesure où on peut lui définir une impulsion. Le photon est une particule qui transporte une quantité de mouvement qu'il est capable de transmettre à d'autres particules lors de collisions avec ces dernières. C'est ce concept qui a conduit Einstein à formuler sa description de l'effet photoélectrique pour laquelle il reçut le prix Nobel en 1921 (selon Einstein, l'effet photoélectrique consiste en l'éjection d'un électron du nuage électronique d'un atome sous l'effet d'une « *collision* » avec un photon). Ainsi, Einstein a redonné au photon le statut de particule qu'il avait « *définitivement* » perdu avec la découverte des ondes électromagnétiques par Maxwell et Hertz. C'était le point de départ de la véritable révolution scientifique qu'a représenté l'avènement de la mécanique quantique. Vous découvrirez cette très élégante théorie en deuxième année dans le cours de Physique quantique et statistique.

Nous approfondirons en séance d'exercices les conséquences de la théorie de la relativité restreinte. Nous verrons, en particulier, que la masse d'un système est liée à son énergie au travers de la formule $E = mc^2$, de sorte qu'une variation de son énergie ΔE conduit à une variation Δm de sa masse selon la relation $\Delta E = \Delta mc^2$.

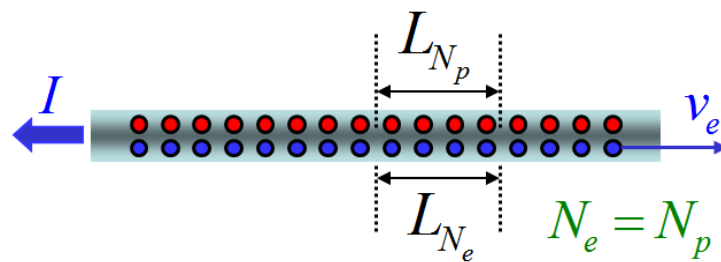
5.9 Force magnétique: manifestation relativiste

Pour terminer cette section consacrée à la relativité restreinte, je propose ici d'illustrer le rôle fondamental qu'a eu cette théorie dans le développement des sciences. A cette fin j'ai choisi de discuter l'origine physique de la force magnétique et du phénomène d'induction électromagnétique de Faraday. La théorie de la relativité restreinte a montré que la force magnétique constitue en fait une manifestation relativiste de la charge électrique. Pour comprendre cette assertion, considérons un fil conducteur parcouru par un courant I , comme représenté schématiquement ci-dessous. Dans le référentiel du laboratoire R dans lequel le fil est au repos, le courant est responsable de la génération d'un champ magnétique \vec{B} donné par la formule de Biot et Savart. Si une charge q est disposée dans ce champ et qu'elle est animée d'une vitesse \vec{v} non nulle, celle-ci subit une force magnétique \vec{F}_M donnée par $q\vec{v} \times \vec{B}$. Jusque là, il n'y a rien de nouveau mais supposons maintenant que nous observions la charge dans son propre référentiel R' , c'est-à-dire, le référentiel dans lequel la charge est au repos (sa vitesse \vec{v}' est nulle). La force magnétique exercée sur la charge dans ce référentiel est bien sûr nulle, soit $\vec{F}'_M = q\vec{0} \times \vec{B} = \vec{0}$. Ce résultat est troublant car dans l'un des référentiels on a une force qui tire la charge vers le bas (si q est positive et \vec{v} est dirigée vers la droite sur le schéma) et dans l'autre on n'a pas de force et la charge reste immobile. Nous avons déjà rencontré cette difficulté lors de l'introduction du phénomène d'induction électromagnétique. Dans le référentiel R' , les lignes de champ magnétique balayent l'espace à la vitesse $-\vec{v}$ et sont donc responsables d'un champ électrique induit $\vec{E}_i = \vec{v} \times \vec{B}$. Nous avons introduit ce champ induit de façon artificielle sans le justifier et donc sans le comprendre en profondeur, à la manière dont Faraday l'avait introduit ne se doutant pas qu'il s'agissait d'une manifestation relativiste de la charge électrique.

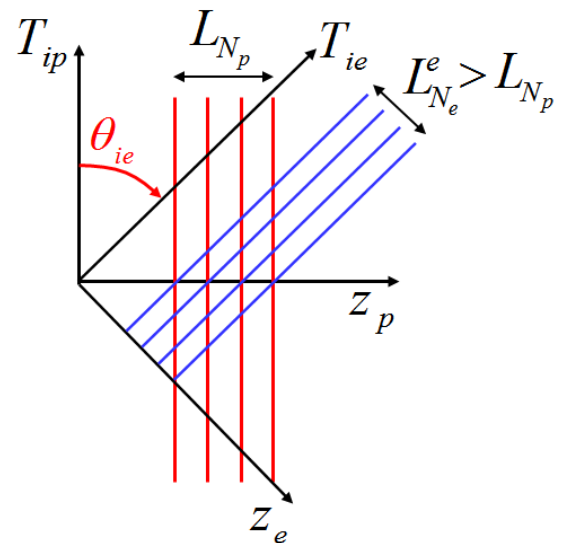


Notre but ici est de comprendre que le champ induit provient du phénomène relativiste de contraction des longueurs. A cette fin, analysons le système en détail à l'aide du schéma suivant sur lequel le courant I apparaît comme étant dû au mouvement des électrons animés d'une vitesse d'ensemble v_e (ils sont représentés sur le schéma ci-dessous par les boules bleues se déplaçant vers la droite). Les atomes du métal constituant le fil portent une charge positive qui compense exactement la charge négative des électrons (ils sont représentés par les boules rouges fixes sur le schéma).

Si cette compensation n'était pas exacte, le fil serait le siège de phénomènes électrostatiques (on constate bien dans la pratique que ce n'est effectivement pas le cas). La neutralité du fil peut s'exprimer à partir de la notion de densité de charges linéiques. Si N_p est le nombre d'atomes compris sur le segment de fil de longueur L_{Np} et si q_e est la charge des atomes (on suppose que les atomes du métal du fil perdent un seul électron), la densité linéique de charge positive est $\lambda_+ = q_e N_p / L_{Np}$. La neutralité du fil impose que la densité de charge linéique négative due aux électrons vaut exactement l'opposé, soit $\lambda_- = -q_e N_e / L_{Ne} = -\lambda_+$, ce qui signifie que sur la longueur L_{Np} on trouve exactement le même nombre d'électrons que de protons $N_e = N_p$.



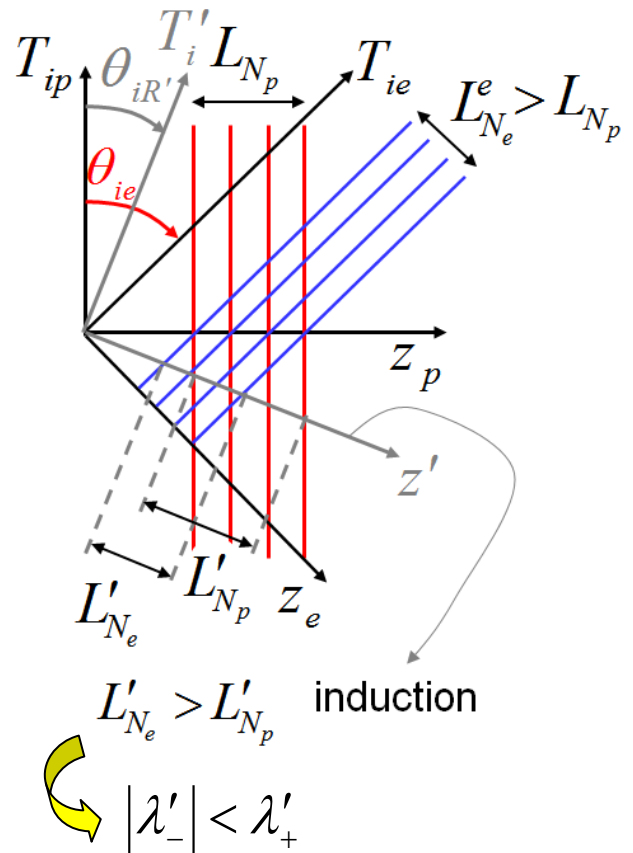
Les lignes d'univers des protons et des électrons sont représentées schématiquement sur le schéma ci-contre. Les référentiels des protons et des électrons sont également représentés à l'aide des axes (z_p, T_{ip}) et (z_e, T_{ie}) , respectivement. On voit sur ce schéma que si la densité linéique de charge est la même dans le référentiel du laboratoire (ou des protons), il faut que les électrons aient une densité plus faible dans leur propre référentiel, puisqu'en vertu du phénomène de contraction des longueurs, ils occupent un segment de plus grande longueur dans leur référentiel (remarquez que sur le schéma la longueur L_{Ne}^e du segment occupé par les N_e électrons dans leur propre référentiel apparaît plus petite que la longueur L_{Np} de ce segment dans le référentiel du laboratoire, ceci doit bien entendu être interprété à l'envers, soit $L_{Ne}^e > L_{Np}$). On peut conclure que la densité de charge négative est plus petite que la densité de charge positive dans le référentiel des électrons. Le fil apparaît donc chargé positivement dans ce référentiel.



Analysons maintenant le système dans le référentiel R' dans lequel la charge d'essai q a une vitesse nulle. Ce référentiel avance à la vitesse v vers la droite dans le référentiel R du laboratoire, il est représenté sur le schéma suivant à l'aide des axes (z', T'_i) . Comme les protons ont une vitesse nulle dans le référentiel R du laboratoire, leur vitesse est $v'_p = -v$ dans le référentiel R' . D'après la formule de contraction des longueurs, la longueur L_{Np} qu'occupent N_p atomes dans le référentiel R apparaît plus

petite dans le référentiel R' , soit $L'_{Np} < L_{Np}$ (ce qui sur le schéma se voit à l'inverse). Comme, selon la théorie de la relativité restreinte, la charge électrique est conservée lors d'un changement de référentiel, cette longueur est toujours occupée par la même quantité de charge soit $q_e N_p$, ce qui implique que la densité linéique de charge positive a augmenté en passant du référentiel R au référentiel R' , soit $\lambda'_+ = q_e N_p / L'_{Np} > \lambda_+$.

Pour la densité linéique de charge négative, la situation est opposée. En effet, puisque la vitesse des électrons est plus faible dans le référentiel R' que dans le référentiel R (notez que la vitesse v de la charge q a été choisie dans le même sens que la vitesse v_e des électrons), la longueur L'_{Ne} qu'occupent les électrons (qui occupaient la longueur L_{Np} dans le référentiel R) dans R' est plus grande que celle occupée par les protons dans R' , soit $L'_{Ne} > L'_{Np}$ (ceci se voit à l'inverse sur le schéma). On peut donc en conclure que la densité linéique de charge négative est, en valeur absolue, plus faible que la densité linéique de charge positive, soit $|\lambda'_-| < \lambda'_+$.



En conclusion, la densité de charge linéique totale du fil $\lambda'_+ + \lambda'_- = \lambda'_+ - |\lambda'_-|$ vue du référentiel R' de la charge d'essai est non nulle et positive. Le fil émet donc un champ électrostatique dans le référentiel R' . Ce champ repousse la charge q vers le bas car il est dirigé vers l'extérieur du fil. De cette façon, on voit que dans les deux référentiels la charge q subit une force vers le bas. Dans le référentiel R , cette force est décrite comme une force magnétique et dans le référentiel R' cette force est décrite comme une force électrique. Un calcul complet et rigoureux montre que les forces sont effectivement égales dans les deux référentiels, elles ont donc bien le même effet sur la charge. Ceci montre que le phénomène d'induction électromagnétique a pour origine une manifestation relativiste de la charge électrique.

Ce résultat est général dans la mesure où la théorie de la relativité restreinte montre que les champs \vec{E} et \vec{B} sont effectivement transformés l'un en l'autre en passant d'un référentiel à un autre, de la même manière que le temps et l'espace s'échangent leurs rôles.

Table des matières

Physique moderne

1. Ondes dispersives	1
1.1 Densité de courant de charge liée	1
1.2 Équation d'onde	3
1.3 Relation constitutive	5
1.3.1 Limite Électrostatique	10
1.3.2 Susceptibilité électrique complexe	11
1.4 Solution de l'équation d'onde modifiée	13
1.5 Vitesse de propagation et relation de dispersion	16
1.6 Propagation d'un battement et vitesse de groupe	19
1.7 Paquet d'ondes et dispersion	24
 2. Introduction à la mécanique quantique	 29
2.1 La conjecture de de Broglie	30
2.2 Relation de dispersion des ondes de matière	32
2.3 Équation de Schrödinger	33
2.4 Quantification de l'énergie	35
2.5 Stabilité de l'atome	40
2.6 Le photon	41

3. Phénomènes d'interférence	44
3.1 Interférences d'ondes planes	44
3.2 Figure d'interférence, intensité lumineuse	46
3.3 Expérience de Young	49
4. Les lentilles	54
4.1 Loi de la réfraction de Snell-Descartes	54
4.2 Lambe de verre d'épaisseur constante	56
4.3 Lambe de verre d'épaisseur variable	57
4.4 Onde sphérique convergente	59
4.5 Lentille convergente	61
4.6 Généralisation à trois dimensions	63
4.7 Lentilles divergentes	63
4.8 Formation d'images	64
4.8.1 L'appareil photographique	64
4.8.2 La loupe	65
4.8.3 Le microscope	66
4.8.4 La lunette, les jumelles	67

5. La relativité restreinte	68
5.1 Principe de relativité de Galilée	68
5.2 Formule d'addition des vitesses	69
5.3 Ondes et relativité	70
5.4 Invariance des équations de Maxwell	74
5.4.1 Changement de référentiel galiléen	75
5.4.3 Transformations de Lorentz	80
5.5 Transformations de Lorentz dans l'espace-temps complexe	88
5.5.1 Ligne d'univers et vitesse	88
5.5.2 Dilatation du temps	90
5.5.3 Contraction des longueurs	93
5.5.4 Loi de composition des vitesses	94
5.6 Synthèse de cinématique relativiste	96
5.7 Dynamique relativiste	99
5.7.1 Rappel de dynamique Newtonienne	99
5.7.2 Accélération relativiste	100
5.8 Énergie et quantité de mouvement	103
5.8.1 Énergie relativiste	105
5.8.2 Impulsion relativiste	109
5.8.3 Le photon, particule sans masse	111
5.9 Force magnétique: manifestation relativiste	112
