

Lesson 4 : Théorie classique des champs

C'est l'époque de la Série mondiale de baseball, et la Taverne d'Hermann est bondée avec les fans que viennent regarder les matchs. Art, qui arrive en retard, approche un tabouret à côté de Lenny.

Art : « Quels sont les noms des types dans le grand champ ? »

Lenny : « Qui est dans un champ classique ; et qu'est-ce qu'il y a dans un champ quantique ? »

Art : « Arrête de me chambrer, Lenny, je pose simplement la question. Qui est dans un champ quantique ? »

Lenny : « Qui est dans un champ classique. »

Art : « C'est ce que je demande. OK. Que sait-on aussi du champ électrique ? »

Lenny : « Tu veux connaître le nom du gars dans le champ électrique ? »

Art : « Naturellement. »

Lenny : « Naturellement ? Ah non. Naturellement est dans un champ magnétique. »

Jusqu'ici nous nous sommes concentrés sur le mouvement relativiste des particules. Dans cette leçon, nous allons aborder la théorie des champs – pas la théorie quantique des champs, mais la théorie classique relativiste des champs. Il pourra y avoir des points de contacts occasionnels avec la mécanique quantique ; je les signalerai alors. Mais nous nous occuperons avant tout de la théorie *classique* des champs.

La théorie des champs que vous connaissez probablement le mieux est l'électromagnétisme : les ondes électromagnétiques et leur interaction avec les particules chargées électriquement. L'électromagnétisme fait intervenir un champ électrique et un champ magnétique.

Mais tout d'abord qu'est-ce qu'un champ ?

4.1 Champs et espace-temps

Un champ n'est rien d'autre qu'une *fonction* d'un ensemble vers un autre

$$f : E \rightarrow F$$

Toutefois comme l'ensemble de départ E qu'on considère a généralement des caractéristiques de nature spatiale – c'est-à-dire une géométrie intéressante –, l'usage est de parler non pas de fonction mais de *champ*. Nous appellerons aussi l'ensemble de départ *l'espace sous-jacent*. Si l'ensemble d'arrivée F est un espace vectoriel, on parle de champ vectoriel. Si l'ensemble d'arrivée est l'ensemble des nombres réels on parle de champ scalaire.

La lettre consacrée pour un champ scalaire est ϕ (la lettre grecque *phi*). Et pour un champ vectoriel on utilisera une lettre majuscule latine comme P ou grecque comme Ξ (la lettre grecque majuscule *kxi*).

En physique l'espace sous-jacent est le plus souvent un

espace-temps, parce qu'on s'intéresse généralement à la dynamique des systèmes et pas seulement à leur statique. Il aura plusieurs dimensions spatiales et une dimension temporelle. Pour commencer à nourrir notre intuition, on peut penser à l'espace-temps newtonien. Mais bien sûr ce sera bientôt l'espace-temps de Minkowski. Les mathématiciens et les physiciens théoriciens sont capables d'imaginer des espaces-temps avec un nombre quelconque de dimensions temporelles et un nombre quelconque de dimensions spatiales. Mais dans le monde physique, même les théories qui ont 10, 11 ou 26 dimensions spatiales¹, ont toujours une seule dimension temporelle. Personne ne sait donner un sens à plus d'une dimension temporelle. Donc nous aurons une dimension temporelle t et un certain nombre de dimensions spatiales X^i .

Un champ est ainsi une fonction qui à chaque point de l'espace-temps (t, X^i) associe un élément d'un espace vectoriel ou de l'ensemble des réels. En d'autres termes, à chaque point spatial correspond un vecteur ou un scalaire, qui lui-même dépend du temps.

La météorologie fournit des exemples naturels. La vitesse du vent mesurée en chaque point X^i d'une région et à chaque temps t est un champ vectoriel. La température est un champ scalaire.

Quand on considère les coordonnées (t, X^i) des points de l'espace-temps, il faut garder à l'esprit que les X^i ne sont pas des degrés de liberté du système que l'on regarde ; ce sont simplement des étiquettes attachées aux points de l'espace *sous-jacent*. De même les points de l'espace-temps – ce qu'on a appelé des *événements* – sont étiquetés par

1. Des esprits facétieux ont dit qu'on ne va pas plus loin que 26 dimensions spatiales en théorie des cordes, parce qu'au delà il n'y a plus de lettres disponibles.

(t, X^i) , où l'indice i va de 1 au nombre de dimensions spatiales. Les degrés de liberté dans la théorie des champs sont dans l'*espace d'arrivée*. En d'autres termes ce sont eux-mêmes des champs.

Nous allons commencer par considérer un champ scalaire ϕ . À chaque point de l'espace-temps (t, X^i) , le champ associe le *scalaire* $\phi(t, X^i)$. Les mathématiciens notent cette correspondance ainsi

$$\phi : (t, X^i) \rightarrow \phi(t, X^i)$$

Quand on travaille avec des champs il est possible d'être parfois un peu perdu entre l'ensemble de départ, dit encore sous-jacent, et l'ensemble d'arrivée. Il est important de ne pas les confondre. L'espace d'arrivée est les réels ou un espace vectoriel. L'espace sous-jacent est en général un espace vectoriel. La figure 4.1 montre un *champ scalaire* de (t, x) vers les nombres réels.

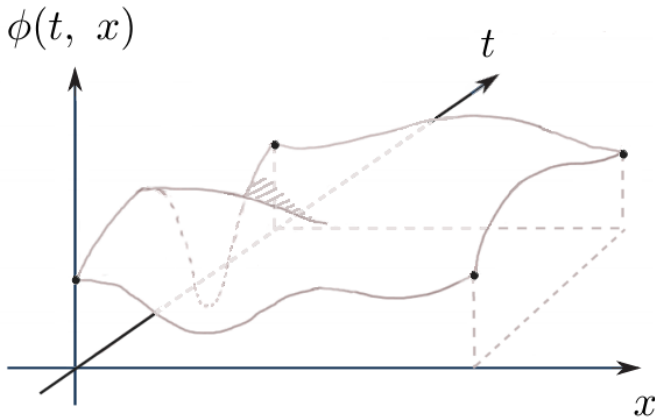


Figure 4.1 : Champ scalaire $\phi(t, x)$.

L'ensemble de départ est le plan horizontal (t, x) , et l'ensemble d'arrivée est l'axe vertical unidimensionnel représentant les réels. Le champ est donc représenté par une nappe ou une surface. À chaque point (t, x) correspond une hauteur sur la nappe.

On peut aussi représenter graphiquement des champs vectoriels, mais le dessin est un peu plus compliqué (cf. *Volume 1*, figure 2 du chapitre 9, sur le théorème de Gibbs-Liouville, où est représenté un champ de vitesses).

Il est habituel en théorie des champs de dire que l'espace sous-jacent est de dimension $(3 + 1)$, voulant dire par là qu'il a trois dimensions spatiales et une dimension temporelle. Plus généralement nous pourrions être intéressés par l'étude de champs sur des espaces-temps avec un nombre plus grand de dimensions spatiales. Si l'espace-temps a d dimensions spatiales, on parlera alors d'un espace-temps de dimension $(d + 1)$. Mais en théorie de la relativité l'ensemble de départ sera habituellement un espace-temps $(3 + 1)$, et l'ensemble d'arrivée les nombres réels ou bien un espace vectoriel.

4.2 Champs et action

Comme je l'ai dit, le principe de moindre action est l'un des principes les plus fondamentaux de la physique. Il gouverne toutes les lois connues de la physique. Sans lui nous n'aurions pas de raison de croire en la conservation de l'énergie ni même en l'existence de solutions aux équations que nous écrivons. Nous allons aussi fonder notre étude des champs sur le principe d'action. Le principe d'action qui gouverne le mouvement des champs est une généralisation de celui qui gouverne le mouvement des particules. Notre plan est

d'examiner en parallèle le principe d'action appliqué aux particules et celui appliqué aux champs, et de les comparer. Pour simplifier cette comparaison, nous allons tout d'abord reformuler le principe d'action pour les particules non-relativistes dans le langage des champs².

4.2.1 La trajectoire d'une particule vue comme un champ

Je souhaite retourner brièvement à la théorie des particules non-relativistes, non pas que je sois particulièrement intéressé par les particules qui vont lentement, mais parce que les mathématiques ont des similarités avec la théorie des champs. En fait, en un certain sens la trajectoire d'une particule non-relativiste est un champ d'une variété très simple – un champ où l'espace sous-jacent a seulement sa dimension temporelle et n'a plus de dimensions spatiales.

La figure 4.2 montre la trajectoire d'une particule comme on la représente habituellement, avec en abscisse le *temps*, et en ordonnée la *position* de la particule sur un axe.

La figure 4.2 peut aussi être vue comme une partie de la figure 4.1 : l'intersection de la nappe et du plan vertical formé par l'axe horizontal t et l'axe vertical sur lequel est repéré ϕ .

En d'autres termes, la trajectoire d'une particule $\phi(t)$ est simplement un champ dont l'espace sous-jacent a été privé

2. Le professeur Susskind emploie parfois le terme traditionnel de *principe de moindre action* et parfois celui de *principe d'action*. Il n'y a pas de différence entre les deux. Mais le deuxième est plus correct car le principe est en réalité de trouver une solution qui rende l'action non pas minimale mais stationnaire.

de tous ses axes spatiaux X^i , pour ne conserver que l'axe temporel t . Quant à l'espace d'arrivée c'est ici l'ensemble des réels sur lequel la particule se déplace en fonction du temps. Il s'agit donc d'un *champ scalaire*.

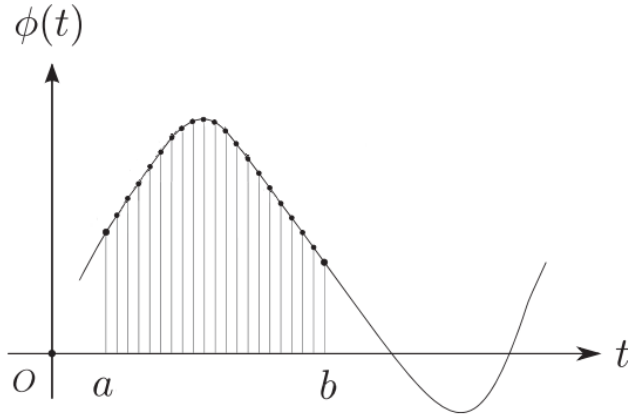


Figure 4.2 : Trajectoire d'une particule vue comme un champ, où l'espace sous-jacent n'a plus de dimensions spatiales, seulement la dimension temporelle t .

Les physiciens parlent parfois pour la théorie du mouvement d'une particule de théorie des champs de dimension $(0 + 1)$.

La figure 4.2 montre donc la trajectoire d'une particule non-relativiste. Pour trouver cette trajectoire nous allons utiliser le principe de moindre action.

Le lecteur ou la lectrice se souviennent que l'action est définie comme l'intégrale d'un certain lagrangien \mathcal{L} en fonction du temps, depuis une date a jusqu'à une date b :

$$A = \int_a^b \mathcal{L} dt$$

Notez pour plus tard – quand nous généraliserons à un espace sous-jacent (t, X^i) – que c’est l’intégrale sur l’espace sous-jacent du champ qu’est la trajectoire de la particule.

Pour des particules non-relativistes, le lagrangien est simple : c’est l’énergie cinétique moins l’énergie potentielle. L’énergie cinétique est ordinairement exprimée sous la forme $\frac{1}{2}mv^2$, mais avec notre nouvelle notation utilisant un champ $\phi(t)$, et non plus l’habituel $x(t)$ pour la position de la particule, nous allons utiliser $\dot{\phi}$, c’est-à-dire $\frac{d\phi}{dt}$, au lieu de v pour la vélocité. Avec cette notation, l’énergie cinétique devient $\frac{1}{2}m\dot{\phi}^2$. Nous allons simplifier davantage en prenant une masse m égale à 1. Comme pour la vitesse de la lumière, c’est juste une question de choix d’unités, et ça permet de faire disparaître des équations un facteur qui – en tout cas à certaines étapes – ne joue pas de rôle déterminant. Ainsi nous avons

$$\text{Énergie cinétique} = \frac{1}{2} \dot{\phi}^2$$

Tournons-nous à présent vers l’énergie potentielle. Dans notre exemple, l’énergie potentielle est juste une fonction de la position de la particule sur son axe – en d’autres termes c’est une fonction de ϕ . Nous allons la noter $V(\phi)$. Le lagrangien étant dans les cas simples, comme ici, l’énergie cinétique moins l’énergie potentielle, on arrive à

$$\mathcal{L}(\phi, \dot{\phi}) = \frac{\dot{\phi}^2}{2} - V(\phi) \tag{4.1}$$

Et l’action est

$$A = \int_a^b \left(\frac{\dot{\phi}^2}{2} - V(\phi) \right) dt \tag{4.2}$$

Nous avons appris dans le *Volume 1* à utiliser les équations d'Euler³-Lagrange pour trouver la fonction $\phi(t)$, c'est-à-dire la trajectoire, qui minimise l'action. L'action est en effet une fonctionnelle dont la variable indépendante est la fonction ϕ . Et trouver la fonction qui minimise la fonctionnelle est un problème de calcul des variations qui généralise la minimisation des fonctions avec laquelle on est familier en calcul différentiel ordinaire.

Quand il n'y a comme ici qu'une seule fonction en variable indépendante, la fonction ϕ , il n'y a qu'une seule équation d'Euler-Lagrange

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \frac{\partial \mathcal{L}}{\partial \phi}$$

Notre tâche est de résoudre cette équation avec le lagrangien donné par l'équation 4.1. Attaquons-nous au terme de gauche, et commençons par dériver le lagrangien par rapport à $\dot{\phi}$. On voit immédiatement que c'est simplement

$$\frac{\partial \mathcal{L}}{\partial \dot{\phi}} = \dot{\phi}$$

Du côté gauche de l'équation d'Euler-Lagrange, il faut encore dériver par rapport au temps. Cela donne $\ddot{\phi}$.

Du côté droit nous devons dériver \mathcal{L} par rapport à ϕ . C'est simplement

$$\frac{\partial \mathcal{L}}{\partial \phi} = - \frac{\partial V(\phi)}{\partial \phi}$$

3. Leonhard Euler (1707-1783), mathématicien suisse qui vécut principalement en Russie et en Allemagne. Il reconnut en Lagrange (1736-1813) son fils spirituel.

En rassemblant nos calculs nous parvenons à l'équation d'Euler-Lagrange sous la forme plus explicite de l'équation différentielle suivante

$$\frac{d^2\phi}{dt^2} = -\frac{\partial V(\phi)}{\partial\phi} \quad (4.3)$$

Pour ceux qui ont déjà lu le *Volume 1*, ce calcul n'était qu'un petit exercice pour se rafraîchir la mémoire. Et l'équation 4.3 à laquelle nous sommes arrivés est familière. C'est tout simplement l'équation de Newton $F = ma$ où les deux côtés sont inversés, et $m = 1$. En effet le terme de droite de l'équation 4.3 est l'opposé du gradient de l'énergie potentielle. C'est par définition la force exercée sur la particule. Et le terme de droite est la dérivée seconde de la position, c'est-à-dire l'accélération de la particule.

Pour trouver la trajectoire de la particule, il faut connaître l'énergie potentielle en fonction de la position afin de pouvoir calculer son gradient. Quand il n'y a, comme dans l'exemple, qu'un seul axe pour le déplacement de la particule c'est simplement la dérivée de V par rapport à ϕ . Nous voulons dériver ou calculer le gradient de V car cela conduit au champ de forces. Par exemple si le gradient est uniforme comme pour le champ de gravitation terrestre (localement à la surface de la Terre), on obtient une chute uniformément accélérée, verticale ou parabolique (si on rajoute des degrés de liberté à la particule) selon les conditions initiales de position et de vitesse de la particule.

Pour le lecteur ou la lectrice qui ne sont pas familiers avec l'utilisation des équations d'Euler-Lagrange, voici comment nourrir son intuition à son sujet. L'action A est l'intégrale depuis le point a jusqu'au point b du lagrangien en fonction du temps. Imaginons qu'on agite comme une ficelle en

quelque sorte la trajectoire reliant le point $[a, \phi(a)]$ au point $[b, \phi(b)]$ sur la figure 4.3, ces deux points étant maintenus fixes. Cela produit toute une collection de trajectoires, pour chacune desquelles on peut calculer une action avec l'intégrale 4.2.

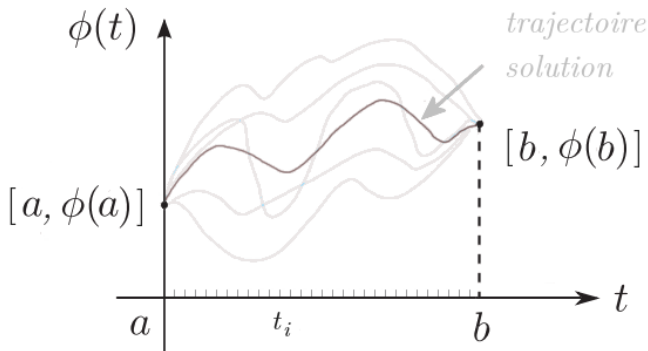


Figure 4.3 : Détermination de la trajectoire qui minimise l'action (ou plus précisément la rend stationnaire).

En outre pour une trajectoire donnée cette intégrale est la limite d'une somme finie d'éléments, calculés en chacune des positions montrées sur la figure 4.3 entre a et b . Quand les intervalles sont choisis de plus en plus étroits et de plus en plus nombreux la somme converge vers l'intégrale.

Rentrons encore un peu plus dans le détail de la façon de penser à la trajectoire solution. Il y a deux façons de le faire, qui inversent l'ordre d'une optimisation et d'un passage à la limite, mais qui sont équivalentes. On peut

- a) soit penser à chaque trajectoire possible et à son action, calculée par une intégrale, elle-même définie comme limite d'une suite de sommes, puis chercher la trajectoire optimale,

- b) soit penser à une collection finie et fixe de temps intermédiaires t_i couvrant le segment de a à b , comme sur les figures 4.2 et 4.3. On approxime alors l'équation 4.2 par une somme finie, qu'on peut aussi noter A , de termes faisant intervenir la collection des $\phi(t_i)$ et a priori aussi des $\dot{\phi}(t_i)$. Mais pour chaque $\dot{\phi}(t_i)$ on prend l'approximation

$$\dot{\phi}(t_i) \approx \frac{\phi(t_{i+1}) - \phi(t_i)}{t_{i+1} - t_i}$$

Donc la somme finie approximant une action est

$$A = \sum_a^b \left\{ \frac{1}{2} \left(\frac{\phi(t_{i+1}) - \phi(t_i)}{t_{i+1} - t_i} \right)^2 - V(\phi(t_i)) \right\} \Delta t$$

où la collection des $\phi(t_i)$ sont les *variables indépendantes*. Pour trouver la collection de $\phi(t_i)$ qui minimise A , on est ramené à un simple problème d'algèbre et de calcul différentiel ordinaire. La collection de points correspondants $[t_i, \phi(t_i)]$ approxime une trajectoire continue. Ensuite on augmente le nombre de temps dans la collection des t_i tout en réduisant leurs intervalles. À la limite on obtient la trajectoire optimale.

Les équations d'Euler-Lagrange (ou l'équation d'Euler-Lagrange quand elle est unique) permettent de trouver quelle est la trajectoire dont l'action est minimale, ou pour être plus exact stationnaire. Elles transforment en effet le problème "trouver la trajectoire rendant l'action stationnaire" en une équation différentielle sur $\phi(t)$, comme on l'a vu dans l'exemple plus haut où elle nous a conduits à l'équation 4.3.

4.3 Principes de la théorie des champs

Nous venons d'étudier un exemple de champ où il n'y avait pas de dimensions spatiales dans l'espace sous-jacent. Basée sur cet exemple, essayons de développer notre intuition pour une théorie des champs dans un monde ressemblant davantage au nôtre – un monde avec une ou plusieurs dimensions spatiales.

Nous partons du postulat que la théorie des champs – en fait toute la physique – est gouvernée par un principe d'action. Rendre l'action stationnaire est un principe très puissant qui encode et synthétise un très grand nombre de lois de la physique.

4.3.1 Le principe d'action

Définissons le principe d'action pour les champs. On vient de voir que pour une particule se mouvant sur une dimension dans l'espace d'arrivée, et pour laquelle l'espace sous-jacent était seulement le temps, figure 4.2, nous avons considéré les deux extrémités fixes $[a, \phi(a)]$ et $[b, \phi(b)]$ de sa trajectoire sur le diagramme t, ϕ , et envisagé toutes les trajectoires possibles entre ces deux points, figure 4.3. Puis nous avons cherché celle qui rend l'action stationnaire. Cette procédure est très similaire à celle de trouver la distance la plus courte entre deux points⁴. Le principe de moindre action nous dit comment calculer la valeur de $\phi(t)$ à toutes les dates intermédiaires entre les deux dates frontières a et b .

4. Dans ce cas, au lieu de chercher à minimiser $\int \mathcal{L} dt$, on cherche à minimiser $\int \sqrt{dt^2 + d\phi^2}$. C'est généralement le premier exemple traité dans les ouvrages sur le calcul des variations. Le deuxième exemple est la détermination de la courbe brachistochrone.

Le problème de la théorie des champs est une généralisation de cette idée : on va calculer la valeur de $\phi(t, X^i)$ en tous les points à l'intérieur d'une *région* de l'espace sous-jacent quand on connaît la valeur du champ en tous les points de sa *frontière*. Nous écrivons $\phi(t, X^i)$ au lieu de $\phi(t, x)$ pour suggérer qu'il y a plusieurs dimensions spatiales. Et le champ $\phi(t, X^i)$ que l'on calculera est celui qui minimise ou rend stationnaire une certaine intégrale généralisant $\int \mathcal{L} dt$.

La figure 4.1 représentait un champ $\phi(t, x)$ avec un espace sous-jacent de dimension $(1+1)$. Dans ce cas l'intégrale d'action que l'on va chercher à minimiser aura la forme

$$A = \iint \mathcal{L} dt dx$$

La région d'intégration sera par exemple le rectangle représenté en perspective dans le plan x, t . Et les conditions aux limites seront que l'on connaît la valeur du champ $\phi(t, x)$ sur les bords du rectangle.

La représentation du champ ou même seulement de son espace sous-jacent devient malaisée dès qu'il y a plusieurs dimensions spatiales. Considérons un champ dont l'espace sous-jacent a les dimensions t, x et y . On peut alors décider de ne pas représenter l'axe des valeurs de ϕ , et représenter une région de l'espace sous-jacent comme sur la figure 4.4. Il s'agit sur la figure d'une région parallélépipédique mais on peut très bien imaginer une région de forme quelconque. Le problème général de la théorie des champs peut être énoncé de la manière suivante :

Ses valeurs sur la frontière d'une région de l'espace-temps étant données, déterminer en tous les points de la région le champ ϕ qui rend une certaine action stationnaire.

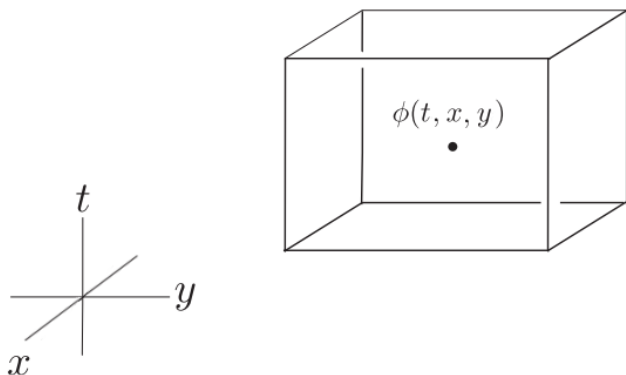


Figure 4.4 : Espace sous-jacent d'un champ $\phi(t, x, y)$. L'axe sur lequel ϕ prend ses valeurs n'est pas représenté.

Quel champ minimise l'action ? Intuitivement de même que sur la figure 4.3 on a agité la trajectoire pour suggérer toutes les trajectoires possibles afin de choisir celle optimale, sur la figure 4.1 on peut imaginer d'agiter la nappe – comme quand on plie un drap avant de le ranger – pour suggérer tous les champs possibles, *mais dont les valeurs sur la frontière sont données*, avant de déterminer celui qui minimise l'action.

L'action correspondant à un champ est maintenant une intégrale multiple sur t et les dimensions spatiales comme on l'a vu plus haut. Pour ne pas surcharger les notations on va laisser tomber le signe intégrale multiple. Pour un champ sur l'espace-temps (3+1) on écrira l'action sous la forme

$$A = \int \mathcal{L} dt dx dy dz$$

où le lagrangien n'a pas encore été spécifié. Mais en théorie

de la relativité nous avons appris à traiter ces quatre dimensions sur un pied d'égalité, chacune d'entre elle participant à la construction de l'espace-temps. Nous avons gommé la différence entre l'axe temporel et les axes spatiaux en leur donnant des noms similaires : X^0 , X^1 , X^2 et X^3 , que l'on résume sous la forme X^μ . Il est habituel d'écrire l'intégrale ci-dessus sous la forme plus ramassée suivante

$$A = \int \mathcal{L} d^4x$$

Tournons-nous à présent vers le lagrangien sous le signe intégrale. Et regardons comment on va minimiser l'action.

4.3.2 Champ ϕ rendant l'action stationnaire

Étant donné que le lagrangien est intégré sur le temps mais aussi l'espace, il est souvent maintenant appelé *densité lagrangienne*. C'est juste une question de cohérence dimensionnelle. Comme on veut que l'action ait les mêmes unités dans le cas d'un champ et dans le cas d'une particule, et que le lagrangien est multiplié par un petit élément de volume sous le signe intégrale, ses propres unités ont changé. Ce doit être une énergie par unité de volume, c'est-à-dire une densité.

Quelles sont les variables (ou fonctions) indépendantes qui gouvernent le lagrangien ? Retournons un instant à une particule non-relativiste. Le lagrangien dépendait des coordonnées de la particule, c'est-à-dire du champ $\phi(t)$, et de sa vitesse, c'est-à-dire $\partial\phi/\partial t$. La généralisation naturelle, inspirée par l'idée d'espace-temps de Minkowski, est que la densité lagrangienne soit maintenant fonction de ϕ et de

toutes les dérivées partielles de ϕ par rapport au temps et à l'espace. C'est-à-dire, \mathcal{L} va maintenant dépendre de

$$\phi, \frac{\partial \phi}{\partial t}, \frac{\partial \phi}{\partial x}, \frac{\partial \phi}{\partial y}, \frac{\partial \phi}{\partial z}$$

Alors, avec nos notations ramassées, nous allons écrire l'action sous la forme

$$A = \int \mathcal{L} \left(\phi, \frac{\partial \phi}{\partial X^\mu} \right) d^4x$$

où l'indice μ parcourt les quatre dimensions temporelle et spatiales. Dans cette intégrale d'action, nous n'avons pas écrit de dépendance explicite de \mathcal{L} vis-à-vis de t, x, y ou z . Mais de même que dans certains problèmes de mouvement de particules le lagrangien peut dépendre explicitement du temps (en plus de sa dépendance implicite, car la position et la vitesse dépendent de toute façon du temps), la densité lagrangienne peut avoir une dépendance explicite à X^μ . Néanmoins, pour un système fermé – dans lequel l'énergie et l'impulsion sont conservées – \mathcal{L} ne dépend pas explicitement du temps ni de la position dans l'espace.

Comme en mécanique classique ordinaire, les équations du mouvement sont obtenues en "agitant" le champ ϕ jusqu'à trouver celui qui minimise l'action. C'est bien sûr seulement une image. La procédure pratique est d'utiliser les équations d'Euler-Lagrange qui, après qu'on a un peu travaillé dessus, conduisent à des équations différentielles sur le champ ϕ . Dans le cas d'une particule avec un degré de liberté, on a vu dans le *Volume 1*, et partiellement révisé ici, que l'équation d'Euler-Lagrange

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \frac{\partial \phi}{\partial t}} - \frac{\partial \mathcal{L}}{\partial \phi} = 0 \tag{4.4}$$

conduit à l'équation de Newton, qui est une équation différentielle d'ordre 2, ou bien de manière équivalente aux équations de Hamilton, qui sont deux équations différentielles d'ordre 1 (*Vol 1* chapitre 8).

Comment les équations d'Euler-Lagrange se transforment-elles dans le cas d'un espace sous-jacent multidimensionnel ? Nous avons démontré dans le *Volume 1* que l'équation d'Euler-Lagrange 4.4 était satisfaite par la trajectoire optimale, c'est-à-dire celle effectivement empruntée par la particule. Nous n'allons pas démontrer quelle équation généralisant l'équation 4.4 est satisfaite par le champ $\phi(X^\mu)$, mais nous allons donner sa forme. Le premier terme dans l'équation 4.4 est une dérivée par rapport au temps, c'est-à-dire la première des quatre composantes. Il va être remplacé par une somme de termes similaires pour chaque composante. En d'autres termes,

$$\frac{d}{dt} \rightarrow \sum_{\mu} \frac{\partial}{\partial X^\mu}$$

et ce sur quoi porte chacune de ces quatre dérivées partielles est modifié respectivement de manière analogue. L'équation 4.4 devient

$$\sum_{\mu} \frac{\partial}{\partial X^\mu} \frac{\partial \mathcal{L}}{\partial \frac{\partial \phi}{\partial X^\mu}} - \frac{\partial \mathcal{L}}{\partial \phi} = 0 \quad (4.5)$$

Par exemple le premier terme de la somme de gauche dans cette équation est simplement le premier terme dans l'équation 4.4 :

$$\frac{\partial}{\partial t} \frac{\partial \mathcal{L}}{\partial \frac{\partial \phi}{\partial t}}$$

Le d droit est devenu un ∂ partiel, pour insister sur le fait qu'il y a maintenant trois autres variables sur le même plan que t .

L'équation 4.5 est l'équation d'Euler-Lagrange pour un champ scalaire dont l'espace sous-jacent est l'espace-temps à quatre dimensions de la relativité. Nous allons voir dans la section 4.3.4 qu'elle est étroitement liée à une équation décrivant des ondulations de ϕ .

Mais d'abord signalons que, comme en mécanique des particules, il peut y avoir plus d'un degré de liberté. Exprimé dans le langage de la théorie des champs, cela veut dire que ϕ peut être vectoriel, ou si l'on préfère qu'il peut y avoir plusieurs champs scalaires sur l'espace-temps. Soyons explicites et supposons que l'on considère deux champs scalaires ϕ et χ (la lettre grecque *chi*, prononcée qui). L'action dépendra de chaque champ et ses dérivées

$$\phi, \frac{\partial \phi}{\partial X^\mu}$$

$$\chi, \frac{\partial \chi}{\partial X^\mu}$$

Il doit alors y avoir deux équations d'Euler-Lagrange du même type que l'équation 4.5

$$\sum_{\mu} \frac{\partial}{\partial X^\mu} \frac{\partial \mathcal{L}}{\partial \frac{\partial \phi}{\partial X^\mu}} - \frac{\partial \mathcal{L}}{\partial \phi} = 0$$

$$\sum_{\mu} \frac{\partial}{\partial X^\mu} \frac{\partial \mathcal{L}}{\partial \frac{\partial \chi}{\partial X^\mu}} - \frac{\partial \mathcal{L}}{\partial \chi} = 0$$
(4.6)

Et s'il y a encore plus de degrés de liberté, c'est-à-dire de champs scalaires, il y aura encore plus d'équations comme

ci-dessus. On pourrait exprimer tout ceci en termes d'un seul champ vectoriel, mais il n'y en aurait pas moins une équation par composante dans l'espace d'arrivée.

4.3.3 Plus sur les équations d'Euler-Lagrange

Le lagrangien le plus simple pour une particule non-relativiste est donné par l'équation 4.1. Il contient une énergie cinétique

$$\frac{1}{2} \left(\frac{d\phi}{dt} \right)^2$$

(on a pris $m = 1$) et une énergie potentielle

$$-V(\phi)$$

Nous pourrions penser que la généralisation en théorie des champs aura une forme similaire, mais où l'énergie cinétique contiendra aussi des dérivées spatiales :

$$\mathcal{L} = \frac{1}{2} \left[\left(\frac{\partial \phi}{\partial t} \right)^2 + \left(\frac{\partial \phi}{\partial x} \right)^2 + \left(\frac{\partial \phi}{\partial y} \right)^2 + \left(\frac{\partial \phi}{\partial z} \right)^2 \right] - V(\phi)$$

Cependant il y a quelque chose de manifestement faux dans cette équation : le temps et l'espace y jouent exactement le même rôle. Certes la relativité les traite plus ou moins sur un pied d'égalité, mais il ne faut pas être plus royaliste que le roi. Même en théorie de la relativité, le temps et l'espace ne sont pas complètement symétriques. L'expérience

ordinaire nous dit bien que ce n'est pas la même chose. L'expression de l'invariant de base en géométrie de Minkowski nous met la puce à l'oreille. On a en effet

$$d\tau^2 = dt^2 - dx^2 - dy^2 - dz^2$$

D'autre part nous verrons plus tard que les dérivées $\partial\phi/\partial X^\mu$ forment un quadrivecteur et que la quantité

$$\left(\frac{\partial\phi}{\partial t}\right)^2 - \left(\frac{\partial\phi}{\partial x}\right)^2 - \left(\frac{\partial\phi}{\partial y}\right)^2 - \left(\frac{\partial\phi}{\partial z}\right)^2$$

est invariante par transformation de Lorentz. Cela suggère d'utiliser ce terme dans l'expression du lagrangien. Plus tard, quand nous nous occuperons de l'invariance de Lorentz, nous verrons pourquoi c'est la bonne façon de faire. Donc, dans la théorie des champs, l'expression

$$\mathcal{L} = \frac{1}{2} \left[\left(\frac{\partial\phi}{\partial t}\right)^2 - \left(\frac{\partial\phi}{\partial x}\right)^2 - \left(\frac{\partial\phi}{\partial y}\right)^2 - \left(\frac{\partial\phi}{\partial z}\right)^2 \right] - V(\phi) \quad (4.7)$$

sera notre lagrangien. Nous pouvons aussi le voir comme un prototype. Par la suite nous construirons d'autres théories fondées sur des lagrangiens similaires à 4.7.

La fonction $V(\phi)$ est appelée le *potentiel de champ*. Elle est analogue à l'énergie potentielle d'une particule. Plus précisément, c'est une *densité d'énergie* (une énergie par unité de volume) en chaque point de l'espace sous-jacent mais qui dépend seulement de la valeur du champ en ce point. La fonction $V(\phi)$ dépend du contexte, et dans au moins un cas important elle peut être déduite de l'expérience. Nous discuterons de ce cas dans la section 4.5.1.

Pour le moment, $V(\phi)$ peut être n'importe quelle fonction de ϕ .

Avant de voir comment utiliser le lagrangien donné par l'équation 4.7 dans les équations d'Euler-Lagrange, faisons une brève digression.

Mécanique des milieux continus

Dans ce chapitre, ce que nous étudions, que j'ai appelé la théorie classique des champs, porte aussi un autre nom : c'est la *mécanique des milieux continus*.

La mécanique des milieux continus embrasse de nombreux domaines de la physique, depuis la mécanique des fluides, jusqu'à l'élasticité, en passant par la résistance des matériaux, l'électromagnétisme et la magnétohydrodynamique des fluides. Si nous parlons d'élasticité, les degrés de liberté peuvent être les déplacements des points dans le matériau élastique. En mécanique des fluides, ils peuvent être les vitesses des petits volumes de fluide, et ainsi de suite. Ce que nous faisons est donc aussi appelé mécanique des milieux continus.

Un autre point qui mérite d'être souligné concerne la présence ou l'absence de la dimension temporelle dans l'espace sous-jacent. Si nous regardons un espace sous-jacent, en chaque point duquel on mesure un ou plusieurs champs, et qui n'a pas de dimension temporelle, les équations d'Euler-Lagrange 4.6 restent valides. Simplement dans ce cas elles nous permettront de trouver la solution de *problèmes de statique*.

Par exemple, nous pourrions considérer un champ électrique ou un champ magnétique créés par un courant électrique circulant dans un solénoïde. Si le problème n'avait

pas de dépendance au temps – si le courant était continu et stable –, dans les équations 4.6 nous éliminerions toutes les dérivées temporelles. Alors nous aurions les mêmes équations mais avec seulement les composantes spatiales.

Exemple d'utilisation de la densité lagrangienne

Prenons la densité lagrangienne donnée par l'équation 4.7 et déduisons pas à pas les équations de mouvement du champ – pensez à la nappe représentée sur la figure 4.1 mais imaginez-la avec trois dimensions spatiales au lieu d'une seule à côté de la dimension temporelle, et le champ quant à lui prend ses valeurs scalaires dans autre dimension dans ce qu'on appelé dans ce chapitre l'espace d'arrivée. L'équation d'Euler-Lagrange 4.5 nous dit exactement comment procéder. L'indice μ prend les valeurs 0, 1, 2 et 3. Chacune des valeurs de cette variable muette engendre un terme séparé dans l'équation pour le champ. Voici comment cela se passe :

- 1) Nous prenons d'abord pour l'indice μ la valeur 0. En d'autres termes le premier X^μ est la coordonnée temporelle X^0 , dénotée aussi t . Et nous commençons par calculer la dérivée partielle de \mathcal{L} par rapport à $\partial\phi/\partial t$. Avec la formule 4.7 cela donne

$$\frac{\partial\mathcal{L}}{\partial\frac{\partial\phi}{\partial t}} = \frac{\partial\phi}{\partial t}$$

Puis nous devons prendre la dérivée partielle par rapport à X^0 (c'est-à-dire t) de ce résultat inter-

médiaire. On obtient

$$\frac{\partial}{\partial X^0} \frac{\partial \mathcal{L}}{\partial \frac{\partial \phi}{\partial t}} = \frac{\partial^2 \phi}{\partial t^2}$$

Ce terme $\partial^2 \phi / \partial t^2$ est analogue à l'accélération d'une particule.

- 2) Ensuite on passe à $\mu = 1$. La coordonnée X^μ devient la première coordonnée spatiale X^1 , dénotée aussi x . Nous devons maintenant calculer la dérivée partielle de \mathcal{L} par rapport à $\partial \phi / \partial x$. Cela donne, en prêtant attention au signe,

$$\frac{\partial \mathcal{L}}{\partial \frac{\partial \phi}{\partial x}} = - \frac{\partial \phi}{\partial x}$$

Puis on calcule la dérivée partielle par rapport x de ce deuxième résultat intermédiaire.

$$\frac{\partial}{\partial x} \frac{\partial \mathcal{L}}{\partial \frac{\partial \phi}{\partial x}} = \frac{\partial}{\partial x} \left(- \frac{\partial \phi}{\partial x} \right) = - \frac{\partial^2 \phi}{\partial x^2}$$

Et nous faisons de même avec $\mu = 2$ et $\mu = 3$, c'est-à-dire avec y et z .

- 3) Rassemblant nos calculs, l'équation d'Euler-Lagrange 4.5 devient

$$\frac{\partial^2 \phi}{\partial t^2} - \frac{\partial^2 \phi}{\partial x^2} - \frac{\partial^2 \phi}{\partial y^2} - \frac{\partial^2 \phi}{\partial z^2} + \frac{\partial V}{\partial \phi} = 0 \quad (4.8)$$

Comme d'habitude, si nous voulons utiliser des unités conventionnelles nous devons réintroduire le facteur c explicitement. (Il est déjà présent dans l'équation 4.8, mais sous la forme d'un 1 qui ne se voit pas.) Le résultat est

$$\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \frac{\partial^2 \phi}{\partial x^2} - \frac{\partial^2 \phi}{\partial y^2} - \frac{\partial^2 \phi}{\partial z^2} + \frac{\partial V}{\partial \phi} = 0 \quad (4.9)$$

4.3.4 Ondes et équations d'onde

De tous les phénomènes décrits par la théorie classique des champs, le plus courant et plus simple à comprendre est la propagation des ondes. Les ondes sonores, les ondes lumineuses, les ondes à la surface ou dans l'eau, les ondes le long d'une corde vibrante : toutes sont décrites par des équations similaires, qui sans surprise sont appelées des équations d'onde. On a déjà vu dans les cours précédents l'importance de l'oscillateur harmonique. La relation entre la théorie des champs et les mouvements ondulatoires est l'un des faits les plus importants de la physique⁵. Il est temps de l'explorer.

L'équation 4.9 est une généralisation de l'équation de Newton du mouvement, qui était l'équation 4.3 reproduite

5. Les cours de physique générale de premier cycle sont parfois découpés en trois parties : 1) mécanique classique newtonienne, 2) électricité et magnétisme, 3) vibrations et ondes. Voir par exemple les cours de Walter Lewin au Massachusetts Institute of Technology <https://tinyurl.com/hnpye3z> qui sont une bonne préparation à ceux plus avancés de Leonard Susskind <http://theoreticalminimum.com/courses/> dont la collection Le Minimum Théorique est la version livre.

ci-dessous,

$$\frac{d^2\phi}{dt^2} = -\frac{\partial V(\phi)}{\partial\phi}$$

Dans cette équation de Newton le terme $-\partial V(\phi)/\partial\phi$ est une force agissant sur la particule et déterminant son accélération à chaque instant. Dans l'équation 4.9 décrivant le mouvement d'un champ, c'est aussi une sorte de force agissant sur le champ, et l'éloignant de son mouvement naturel stable⁶ quand il n'y a pas de force.

Pour une particule le mouvement sans force externe est un mouvement uniforme, à vitesse constante. Pour un champ, le mouvement sans force externe est une onde se propageant comme un son ou une radiation électromagnétique stables.

Pour voir cela, simplifions l'équation 4.9. D'une part omettons le terme lié à une force en prenant $V(\phi) = 0$. D'autre part au lieu de trois dimensions spatiales, considérons-en seulement une. Nous sommes maintenant dans un espace-temps sous-jacent plus simple de dimension $(1 + 1)$. Et l'équation 4.9 devient

$$\frac{1}{c^2} \frac{\partial^2\phi}{\partial t^2} - \frac{\partial^2\phi}{\partial x^2} = 0 \tag{4.10}$$

Je vais vous montrer tout une famille de solutions. Considérons n'importe quelle fonction de la quantité $(x + ct)$. Dénotons-la $F(x + ct)$. D'après notre définition très générale d'un champ, donnée au début de ce chapitre, F est un champ scalaire sur l'espace sous-jacent de dimension $(1+1)$

6. Un mouvement stable est a priori un oxymore. On veut dire par là un mouvement périodique.

(t, x) . On peut le voir comme une nappe, figure 4.1. Mais ce champ a une forme intéressante un peu comme une chevelure crantée.

Considérons les dérivées de F par rapport à x et par rapport à t . Il est facile de voir qu'elles satisfont

$$\frac{\partial F(x+ct)}{\partial t} = c \frac{\partial F(x+ct)}{\partial x}$$

En appliquant la méthode une seconde fois, et en se rappelant qu'on peut inverser l'ordre des dérivations partielles, on a même

$$\frac{\partial^2 F(x+ct)}{\partial t^2} = c^2 \frac{\partial^2 F(x+ct)}{\partial x^2}$$

soit

$$\frac{1}{c^2} \frac{\partial^2 F(x+ct)}{\partial t^2} - \frac{\partial^2 F(x+ct)}{\partial x^2} = 0 \quad (4.11)$$

L'équation 4.11 n'est autre que l'équation d'onde 4.10 appliquée à la fonction F . Nous avons donc trouvé une vaste classe de solutions de l'équation d'onde. Toute fonction de x et t seulement à travers la quantité $(x+ct)$ est une solution.

Quelles sont les propriétés des fonctions de la forme $F(x+ct)$? Au temps $t=0$ c'est juste la fonction $F(x)$. À mesure que le temps passe, la valeur de la fonction au point x de l'espace sous-jacent change, car elle dépend de $(x+ct)$. Son graphe dans un plan x, z se déplace rigidement vers la gauche avec la vitesse c (pour compenser le terme $+ct$ dans son argument). Prenons un exemple de fonction F (que l'on identifie maintenant à notre champ ϕ). Soit une fonction sinusoïdale de période spatiale $2\pi/k$

$$\phi(t, x) = \sin k(x+ct)$$

C'est une onde sinusoïdale se déplaçant vers la gauche avec la vitesse c . Il y a aussi des solutions en cosinus, et même des paquets d'onde et des impulsions. Tant qu'elles se déplacent vers la gauche *avec la vitesse c* ce sont des solutions de l'équation 4.10.

Qu'en est-il des ondes se déplaçant vers la droite ? Sont-elles aussi décrites par l'équation 4.10 ? La réponse est oui. Tout ce qu'on a à faire pour que l'onde se déplace vers la droite est de changer $(x + ct)$ en $(x - ct)$. Il est laissé au lecteur ou à la lectrice le soin de montrer que toute fonction dépendant de x et t seulement à travers $(x - ct)$ est solution de l'équation 4.10 et a un graphe dans le plan x, z qui se déplace vers la droite à mesure que le temps passe.

Avant de passer aux champs relativistes concluons cette section sur les ondes avec quelques observations.

Une onde sinusoïdale $\phi(t, X^i)$ sur l'espace sous-jacent (t, X^i) est le mouvement le plus simple que peut prendre un champ *quand il n'y a pas de potentiel de champ $V(\phi)$* . C'est la généralisation à un champ avec des dimensions spatiales du mouvement uniforme $\phi(t)$ d'une particule sur l'espace réduit au simple axe temporel t , quand il n'y a pas de champ de forces externe. Dans les deux cas il ne faut pas perdre de vue que l'espace dans lequel ϕ prend ses valeurs est l'*espace d'arrivée* du champ. *C'est un autre axe spatial qui ne fait pas partie des axes de l'espace sous-jacent.*

Pour penser à un champ stable dont l'espace sous-jacent a plusieurs dimensions spatiales, on peut penser à la *houle sur la mer*. Ici l'espace sous-jacent a deux dimensions spatiales et bien sûr l'axe temporel. En un point donné de la mer, son niveau monte et descend périodiquement en fonction du passage de la houle. Le champ est le niveau de la mer. Il prend ses valeurs sur un autre axe que les deux

axes spatiaux des cartes marines. Quand on veut dessiner la houle, en général on représente les trois dimensions x , y et l'axe d'arrivée z (qui n'est pas un des axes de l'espace sous-jacent). Et on ne représente pas l'axe temporel, contrairement à ce qu'on a fait sur la figure 4.1, où l'existence d'un seul axe spatial x le permettait.

Quand on dit que la houle se déplace, il faut faire attention à une erreur classique dans laquelle peut nous faire trébucher le bon sens. Les points (x, y) sur la mer par définition ne se déplacent pas. Alors qu'est-ce qui se déplace ? D'une part, si on met un bouchon au point (x, y) son altitude va changer en fonction de la houle ; il montera et descendra alternativement, mais en restant toujours au même point (x, y) . D'autre part on peut dire si l'on veut qu'une crête de vague de la houle se déplace, ce n'est pas inexact, mais rien de tangible ne se déplace sur le plan (x, y) .

Quand le champ est dans une dimension perpendiculaire à l'espace sous-jacent on parle d'onde transverse. Il existe aussi des ondes longitudinales, par exemple quand on fait circuler une impulsion dans un long ressort, ou tout simplement les ondes sonores. L'axe d'arrivée est alors physiquement confondu avec un des axes de l'espace sous-jacent, mais du point de vue de l'analyse il reste distinct.

Les vibrations sont à première vue simples, mais elles conduisent à des phénomènes complexes. Il existe aussi des vibrations stables sur l'espace sous-jacent \mathbb{R}^2 qui ne se déplacent pas (au sens de la houle). Penser aux vibrations complexes qu'on peut donner à une plaque de métal, sur laquelle on a versé du sable pour visualiser les nœuds de vibrations⁷ Un autre domaine est la recherche en anti-bruit actif. Elle bute encore sur des mathématiques insolubles par les méthodes analytiques, et encore formidables pour

7. Un film pédagogique est à <https://youtu.be/wvJAgUBF4w>

les méthodes numériques, et de ce fait progresse lentement. Imaginez, cependant, un jour pouvoir ouvrir votre fenêtre en ville et ne pas entendre le bruit des voitures, seulement le chant des oiseaux.

La maîtrise des vibrations dans les machines est un domaine important de la technologie. Les TGV les plus récents, dans lesquels on peut boire un café, lire et écrire sur la tablette devant soi alors qu'ils se déplacent à 300 km/h, en termes de contrôle des vibrations sont aux premiers TGV ce que ceux-ci étaient aux locomotives à vapeur.

Une machine, même statique, mal conçue peut se briser à cause des vibrations engendrées par son fonctionnement. Un exemple célèbre de dégâts causés par les vibrations est celui du pont du détroit de Tacoma, dans l'État de Washington aux États-Unis, qui quelques mois après son inauguration en 1940 entra en vibration sous l'effet du vent comme une anche de clarinette, et se désintégra.

4.4 Champs relativistes

Quand nous construisons une théorie des particules, il y a deux principes de base que nous devons respecter, en plus du principe de moindre action.

1. Une action doit être une intégrale. Dans le cas d'une particule, c'est une intégrale sur le temps le long de la trajectoire. Dans le cas d'un champ avec des dimensions spatiales, $\phi(t, X^i)$, c'est une intégrale sur l'espace-temps.
2. Une action doit avoir la même valeur dans tous les référentiels. Elle doit être construite à partir de quantités de telle sorte qu'elle conserve exactement la même forme.

Comment nous sommes-nous assurés de respecter ces principes dans le cas d'une particule ? Tout d'abord nous avons construit l'action totale A en sommant les petites actions de nombreux petits segments le long de la trajectoire de la particule. Ensuite, nous avons pris la limite de cette somme quand la taille des petits segments tendait vers zéro. Par définition, la somme devient l'intégrale. Le point crucial est ceci : quand nous avons défini l'action pour l'un de ces petits segments, nous avons choisi une quantité qui était invariante – le temps propre le long de la trajectoire. Étant donné que tous les observateurs s'accordent sur la valeur du temps propre de chaque petit segment, ils s'accordent forcément aussi sur la valeur de l'action sur chaque trajectoire possible. Par conséquent, les équations du mouvement que l'on déduit de l'ensemble des actions possibles – en appliquant le principe de moindre action – sont exactement les mêmes dans tous les référentiels. Les lois de la mécanique des particules sont invariantes. J'ai déjà fait allusion à cette façon de procéder pour les champs dans la section 4.3.3. Mais maintenant nous voulons aborder l'invariance de Lorentz dans la théorie relativiste des champs de manière plus formelle et détaillée.

Nous devons savoir comment former des quantités invariantes à partir des champs, et ensuite les utiliser pour construire des intégrales d'action qui sont invariantes elles aussi. À cette fin, nous devons avoir une vue claire du concept de transformation invariante pour un champ.

4.4.1 Propriétés des transformations sur les champs

Retournons à l'exemple où Art était dans une gare et Lenny dans un train traversant la gare avec une certaine vitesse v . Ils considéraient tous les deux le même événement, par exemple une lampe qui s'allume quelque part dans l'espace-temps. L'espace-temps, comme notre espace à trois dimensions spatiales ordinaires, est défini avant toute notion de repère ou de référentiel. Et un événement y est bien défini indépendamment de toute coordonnée⁸.

Dans le référentiel de la gare, arbitrairement dit au repos, et utilisé comme référentiel de base pour le diagramme de Minkowski, l'événement a les coordonnées (t, x, y, z) . Dans le référentiel du train, il a les coordonnées (t', x', y', z') . Et, puisque nous avons laissé nos considérations d'événements et de référentiels en mouvement relatif depuis un moment, rappelons qu'il s'agit d'un seul et même événement, simplement avec deux labels différents.

Art et Lenny ont aussi des détecteurs de champ. Et leurs appareils enregistrent des valeurs numériques pour un champ ϕ . La variété de champ la plus simple est un champ pour lequel ils obtiennent exactement la même mesure. Si nous appelons a priori le champ que mesure Art $\phi(t, x, y, z)$, et le champ que mesure Lenny $\phi'(t', x', y', z')$, alors la transformation la plus simple est

$$\phi'(t', x', y', z') = \phi(t, x, y, z) \quad (4.12)$$

8. C'est à vrai dire une question philosophique intéressante. Mais nous faisons ici simplement référence au fait géométrique élémentaire qu'un même espace peut avoir différents repères pour étiqueter ses points, ou un même espace-temps peut avoir différents référentiels pour étiqueter ses événements.

En d'autres termes, en n'importe quel événement de l'espace-temps Art et Lenny (et qui que ce soit d'autre) sont d'accord sur la valeur du champ ϕ en cet événement-là.

Cette propriété est une caractéristique d'un *champ scalaire*. En effet un champ scalaire n'est pas concerné par le référentiel; dès qu'un événement est choisi dans l'espace-temps, les observateurs seront forcément d'accord sur la valeur du champ en cet événement, figure 4.5.

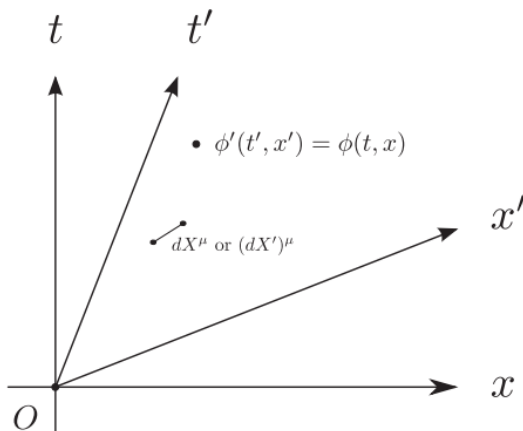


Figure 4.5 : Champ scalaire vu par deux référentiels en mouvement l'un par rapport à l'autre. Les valeurs du champ scalaire en chaque événement ne sont pas affectées par le choix de référentiel. En d'autres termes $\phi'(t', x') = \phi(t, x)$.

On a dit en introduction que la température est typiquement un champ scalaire. À vrai dire ce n'est pas un très bon exemple car la température est un concept un peu plus compliqué qu'il n'y paraît et que nous définirons de manière précise dans le *Volume 6* sur la thermodynamique statis-

tique⁹. Si vous mesurez la température atmosphérique avec deux thermomètres l'un au repos et l'autre en mouvement vous n'obtiendrez pas exactement les mêmes valeurs aux mêmes points de l'espace-temps. Néanmoins le concept de champ scalaire devrait être clair.

Tous les champs ne sont pas scalaires. Nous avons mentionné en introduction la vitesse du vent. Dans un espace à trois dimensions spatiales, plus une composante temporelle, la vitesse du vent en un point donné à une date donnée a trois composantes spatiales. Des anémomètres dans des référentiels différents ne mesureront pas les mêmes composantes. Par exemple, l'air peut être au repos par rapport à la gare, mais si Lenny passe la tête par la fenêtre il aura les cheveux, ou disons plutôt la barbe, dans le vent.

Venons-en maintenant au concept de champ quadrivecteuriel. La vitesse du vent reste un bon exemple. En un événement quelconque de l'espace-temps, dans le référentiel de la gare Art mesure la vitesse de l'air, c'est-à-dire les trois composantes dX^i/dt des molécules d'air en cet événement-là. Il obtient le tri-vecteur ordinaire

$$V^x(t, x, y, z), V^y(t, x, y, z), V^z(t, x, y, z)$$

Mais vous pouvez déjà deviner que dans la théorie de la relativité nous devrions représenter la vitesse des molécules par leur quadrivitesse, c'est-à-dire non pas avec $V^i = dX^i/dt$ mais avec $U^\mu = dX^\mu/dt$.

Connaissant les composantes de la quadrivitesse du vent dans le référentiel de la gare, quelles sont-elles dans le référentiel du train ? Comme la quadrivitesse est un quadrivec-

9. Les notes de cours en anglais sont disponibles à https://www.lapasserelle.com/statistical_mechanics/

teur (c'est-à-dire que ce sont quatre nombres qui se transforment selon la transformation de Lorentz), on sait automatiquement que

$$\begin{aligned}
 (U')^0 &= \frac{U^0 - vU^1}{\sqrt{1 - v^2}} \\
 (U')^1 &= \frac{U^1 - vU^0}{\sqrt{1 - v^2}} \\
 (U')^2 &= U^2 \\
 (U')^3 &= U^3
 \end{aligned}
 \tag{4.13}$$

Je n'ai pas inclus les dépendances à (t, x, y, z) ou (t', x', y', z') pour ne pas surcharger les notations, mais la règle est simple : du côté droit les composantes dépendent de (t, x, y, z) et du côté gauche de (t', x', y', z') . Mais ces deux jeux de coordonnées sont des labels sur le même événement.

Considérons un autre quadrivecteur intéressant que nous appellerons le gradient dans l'espace-temps du champ scalaire ϕ . Ses composantes sont les dérivées de ϕ . Nous utiliserons la notation allégée suivante pour ces dérivées

$$\partial_\mu \phi = \frac{\partial \phi}{\partial X^\mu}
 \tag{4.14}$$

Notez l'indice μ en position basse dans le terme de gauche. La justification pour cette façon de noter apparaîtra quand nous manipulerons des tenseurs et serons amenés à faire des sommes sur des indices.

Nous pourrions penser que $\partial_\mu \phi$ est un quadrivecteur et se transforme de même façon que U^μ . Nous ferions une erreur, mais pas une grosse erreur.

4.4.2 Interlude mathématique : composantes covariantes

Faisons une pause pour parler de divers aspects de la transformation de la *représentation* des objets mathématiques d'un référentiel vers un autre. Nous avons vu que les champs scalaires ont tout simplement la même représentation, c'est-à-dire la même mesure tant qu'on ne change pas d'unités, dans tous les référentiels. Regardons maintenant des objets mathématiques plus élaborés.

Supposons que nous ayons un espace décrit par deux jeux de coordonnées X^μ et $(X')^\mu$. Considérons un intervalle infinitésimal représenté respectivement par dX^μ et par $d(X')^\mu$. Le calcul différentiel à plusieurs variables ordinaire implique que les deux intervalles infinitésimaux satisfassent la relation

$$d(X')^\mu = \sum_{\nu} \frac{\partial (X')^\mu}{\partial X^\nu} dX^\nu \quad (4.15)$$

Nous expliquons pourquoi plus en détail un peu plus bas quand nous parlons de fonctions composées.

Einstein écrivit beaucoup d'équations de ce genre. Après quelques temps, il observa un motif : chaque fois qu'il y avait un indice répété dans une même expression – l'indice ν du côté droit de l'équation 4.15 est un tel indice – l'expression comportait une somme sur cet indice. Dans l'un de ses articles sur la relativité générale, au bout de quelques pages apparemment il se fatigua d'écrire le signe somme et déclara que désormais chaque fois qu'une expression comporterait deux fois le même indice il faudrait opérer une somme sur lui. Cette convention devint connue sous le nom de *convention de sommation d'Einstein*. Elle est mainte-

nant universellement employée au point qu'aucun physicien ne prend plus la peine de la mentionner quand il écrit des équations. Adoptons-la nous aussi et dispensons-nous du signe \sum_ν dans l'équation 4.15. On la réécrit plus simplement

$$d(X')^\mu = \frac{\partial(X')^\mu}{\partial X^\nu} dX^\nu \quad (4.16)$$

où la sommation sur l'indice muet μ du côté droit est implicite. Un autre garde-fou est que ν apparaît à droite mais pas à gauche, donc il y a clairement une opération implicite du côté droit : c'est la sommation sur ν .

Si les équations reliant X à X' sont linéaires, comme c'est le cas dans la transformation de Lorentz, alors les dérivées partielles $\frac{\partial(X')^\mu}{\partial X^\nu}$ sont nécessairement des coefficients constants, comme l'est la pente d'une droite. Considérons la transformation de Lorentz permettant de passer de X à X' , et limitons-nous aux deux premières coordonnées puisque dans un boost le long de X^1 , les deux autres composantes spatiales ne sont pas affectées.

Les équations reliant $(X')^0$ et $(X')^1$ à X^0 et X^1 sont

$$(X')^0 = \frac{X^0 - vX^1}{\sqrt{1 - v^2}}$$

$$(X')^1 = \frac{X^1 - vX^0}{\sqrt{1 - v^2}}$$

Voici la liste des quatre dérivées partielles $\frac{\partial(X')^\mu}{\partial X^\nu}$ en utilisant la notation γ pour $1/\sqrt{1 - v^2}$

$$\begin{aligned}
\frac{\partial(X')^0}{\partial X^0} &= \gamma \\
\frac{\partial(X')^0}{\partial X^1} &= -v\gamma \\
\frac{\partial(X')^1}{\partial X^0} &= -v\gamma \\
\frac{\partial(X')^1}{\partial X^1} &= \gamma
\end{aligned}
\tag{4.17}$$

Si nous introduisons ces coefficients dans les équations 4.16 nous obtenons comme il se doit

$$\begin{aligned}
d(X')^0 &= \gamma dX^0 - v\gamma dX^1 \\
d(X')^1 &= -v\gamma dX^0 + \gamma dX^1
\end{aligned}
\tag{4.18}$$

Cela a bien sûr aussi une notation matricielle avec la matrice de changement de base exprimant la nouvelle base (du référentiel en mouvement) dans l'ancienne (du référentiel au repos).

Tirons de cet exercice une règle générale pour la transformation des quadrivecteurs. Retournant à l'équation 4.16, nous pouvons remplacer les composantes infinitésimales par les composantes de n'importe quel quadrivecteur dans les référentiels au repos et en mouvement. L'équation 4.16 devient pour n'importe quel quadrivecteur

$$(A')^\mu = \frac{\partial(X')^\mu}{\partial X^\nu} A^\nu
\tag{4.19}$$

Dans le cas d'une transformation de Lorentz, les dérivées partielles sont les composantes de la matrice de change-

ment de base de l'équation 4.18. On peut donc écrire plus explicitement

$$\begin{aligned}(A')^0 &= \gamma A^0 - v\gamma A^1 \\ (A')^1 &= -v\gamma A^0 + \gamma A^1\end{aligned}\tag{4.20}$$

Mais l'équation 4.19, qui est une notation tensorielle, est plus générale. Elle est vraie pour des transformations de référentiels quelconques¹⁰ – auquel cas les dérivées partielles ne sont plus des constantes, mais dépendent de l'événement (X^0, X^1) où l'on se trouve.

Mon réel objectif en présentant tout ceci n'est pas d'expliquer comment dX^μ ou A^μ se transforment, mais de montrer comment $\partial_\mu\phi$ se transforme. Les quatre composantes de $\partial_\mu\phi$ forment aussi un quadrivecteur, mais d'une sorte un peu différente de ceux qu'on a vus jusqu'à présent, par exemple dX^μ ou U^μ .

Les composantes de $\partial_\mu\phi$, par leur définition même à l'aide de l'équation 4.14, font manifestement référence au système de coordonnées X , mais elles peuvent être transformées pour que l'objet qu'elles représentent – le gradient d'un champ scalaire – soit exprimé dans le système de coordonnées X' .

La règle basique de transformation provient du calcul différentiel élémentaire. C'est une généralisation de la règle de dérivation en chaîne – qui dans les occasions solennelles

10. On se rappelle qu'un référentiel n'est qu'une collection d'étiquettes sur les événements. Deux référentiels sont simplement deux collections d'étiquettes différentes. Quand elles sont raisonnables on peut différencier les coordonnées dans l'une par rapport aux coordonnées dans l'autre.

porte le nom de théorème de dérivation des fonctions composées. Rafraîchissons-nous la mémoire. Soit une fonction $f(x)$ de \mathbb{R} vers \mathbb{R} . Si l'on change de variable et que x soit lui-même une fonction $g(x')$, alors on peut s'intéresser à la fonction f en fonction de x' que l'on note aussi $f \circ g(x')$.

Concrètement, à chaque point x' de la droite \mathbb{R} correspond un point x par la fonction g , et à x correspond une valeur par la fonction f . On peut s'intéresser au taux de variation de f quand on fait varier x' .

La dérivée de f par rapport à x' est obtenue par la règle de dérivation des fonctions composées

$$\frac{df}{dx'} = \frac{df}{dx} \frac{dg}{dx'}$$

On s'en rappelle parfois en l'écrivant

$$\frac{df}{dx'} = \frac{df}{dx} \frac{dx}{dx'}$$

et en "simplifiant" par dx . Il se trouve que c'est formellement correct ici, mais les ratios d'infinitésimaux, inventés par Leibniz, ne se manipulent pas comme des ratios habituels. C'est parce qu'ils étaient vagues qu'au XIX^e siècle des mathématiciens comme Augustin Cauchy (1789 - 1857) et Karl Weierstrass (1815 - 1897) ont redéfini la notion de dérivée plus proprement à l'aide de limites de vrais ratios. Mais comme nous l'avons déjà mentionné, les infinitésimaux, quand ils sont manipulés avec soin, sont très intuitifs et commodes, et sont restés des instruments dans la boîte à outil des physiciens et des mathématiciens^{11 12}.

11. Ne serait-ce qu'en observant que la notation dy/dx a perduré, même si elle trouble – à juste titre – les néophytes.

12. De même Bernhard Riemann (1826 - 1866) a défini proprement la notion d'intégrale utilisée depuis déjà deux siècles avant lui.

La règle de dérivation en chaîne se généralise aisément au calcul différentiel à plusieurs variables. Pour un champ scalaire ϕ défini sur un espace sous-jacent à plusieurs dimensions elle devient

$$\frac{\partial \phi}{\partial (X')^\mu} = \sum_\nu \frac{\partial X^\nu}{\partial (X')^\mu} \frac{\partial \phi}{\partial X^\nu}$$

ou, en utilisant la convention de sommation,

$$\frac{\partial \phi}{\partial (X')^\mu} = \frac{\partial X^\nu}{\partial (X')^\mu} \frac{\partial \phi}{\partial X^\nu}$$

Simplifions encore à l'aide des notations abrégées $\partial_\nu \phi$ pour $\frac{\partial \phi}{\partial X^\nu}$ et $\partial'_\mu \phi$ pour $\frac{\partial \phi}{\partial (X')^\mu}$. Nous obtenons

$$\partial'_\mu \phi = \frac{\partial X^\nu}{\partial (X')^\mu} \partial_\nu \phi \quad (4.21)$$

Soyons plus général et remplaçons $\partial'_\mu \phi$ par B_μ . L'équation 4.21 devient

$$B'_\mu = \frac{\partial X^\nu}{\partial (X')^\mu} B_\nu \quad (4.22)$$

Examinons un instant les équations 4.19 et 4.22 qui spécifient la façon de passer respectivement de A^μ à $(A')^\mu$ et de B_μ à B'_μ . Réécrivons-les l'une au dessus de l'autre pour faciliter leur comparaison

$$\begin{aligned} (A')^\mu &= \frac{\partial (X')^\mu}{\partial X^\nu} A^\nu \\ B'_\mu &= \frac{\partial X^\nu}{\partial (X')^\mu} B_\nu \end{aligned}$$

On voit deux différences.

- a) La première différence est que dans la première équation les indices grecs sur A ou A' – qui listent leurs composantes – sont en position supérieure, tandis que dans la seconde équation ceux sur B ou B' sont en position inférieure. Cette différence de notation semble n'avoir aucune importance, mais en fait elle en a : elle est utilisée pour signifier quelque chose, qui va devenir clair dans un instant.
- b) La seconde différence concerne les coefficients dérivées partielles. Dans l'équation du haut on a les dérivées partielles de X' par rapport à X , tandis que dans celle du bas on a les dérivées partielles de X par rapport à X' .

Clairement les A^μ et les B_μ sont deux sortes différentes de quadrivecteurs qui se transforment de manière différente pour passer du repère sans prime au repère avec prime. En fait ceux qu'on avait considérés jusqu'à présent étaient du type A^μ .

Les quadrivecteurs qui se transforment comme A^μ , c'est-à-dire selon l'équation 4.19 sont dits *contravariants*. Et les quadrivecteurs qui se transforment comme B_μ , c'est-à-dire selon l'équation 4.22, sont dits *covariants*. Vous pouvez aussi dire informellement si vous préférez "quadrivecteurs avec indice en haut" et "quadrivecteurs avec indice en bas".

D'où viennent les termes contravariants et covariants ? Cela veut dire respectivement "qui se transforment contrairement à la base" et "qui se transforment comme la base". Pour le comprendre pensons à deux bases dans l'espace. Peu importe le nombre de dimensions, une seule suffit. Considérons donc un espace $(1 + 0)$, c'est-à-dire juste un axe spatial. Supposons que le vecteur de base prime soit le vec-

teur de base sans prime *divisé* par 10. Alors $dX'/dX = 10$. Et pour passer de A à A' il faut *multiplier* A par 10. On dit donc que A est contravariant, car il varie contrairement à la base.

Maintenant regardons comment B se transforme. Pour passer de B à B' il faut *diviser* B par 10. On dit donc que B est covariant, car il se transforme comme la base.

D'une manière générale tous les vecteurs habituels (position, vitesse, accélération...) sont contravariants. Tandis que le gradient est un exemple de vecteur covariant. Certains auteurs objectent d'appeler le gradient un vecteur, et l'appellent un covecteur.

Retournons à la transformation de Lorentz. Dans l'expression 4.17, j'avais listé les dérivées partielles des composantes avec prime par rapport aux composantes sans primes. Nous pouvons faire de même pour les dérivées partielles des composantes sans prime par rapport aux composantes avec prime. C'est particulièrement simple à établir. Il suffit d'observer que si le train va à la vitesse v dans le repère de la gare, alors la gare va à la vitesse $-v$ dans le repère du train. Donc il suffit de changer le signe de v dans l'expression 4.17. On obtient

$$\frac{\partial X^0}{\partial (X')^0} = \gamma$$

$$\frac{\partial X^0}{\partial (X')^1} = v\gamma$$

$$\frac{\partial X^1}{\partial (X')^0} = v\gamma$$

$$\frac{\partial X^1}{\partial (X')^1} = \gamma$$

Et voici les règles de transformations pour les composantes d'un quadrivecteur covariant :

$$\begin{aligned} B'_0 &= \gamma B_0 + v\gamma B_1 \\ B'_1 &= v\gamma B_0 + \gamma B_1 \end{aligned} \tag{4.23}$$

Jetons de nouveau un œil à l'espace vectoriel des quadrivecteurs contravariants. Nous avons déjà vu des exemples de quadrivecteurs contravariants comme les X^μ , c'est-à-dire les déplacement de O au point de coordonnées X^μ . Le déplacement différentiel entre deux points proches, dX^μ , est aussi un quadrivecteur contravariant. Si on multiplie un quadrivecteur contravariant par un scalaire (c'est-à-dire un *invariant*), le résultat est encore un quadrivecteur contravariant car les invariants sont complètement passifs dans une transformation de référentiel. Nous avons déjà vu que le temps propre $d\tau$ est un invariant, et donc la quantité $dX^\mu/d\tau$, qui est la quadrivitesse, est un quadrivecteur contravariant :

$$U^\mu = \frac{dX^\mu}{d\tau}$$

Généralement quand on parle d'un quadrivecteur, sans préciser s'il est contravariant ou covariant, on parle d'un quadrivecteur contravariant.

Le tableau 4.1 récapitule ce qu'on a appris sur les scalaires et les quadrivecteurs contravariants. Un exemple de *champ* A^μ avec ces propriétés est obtenu à partir d'un fluide qui remplirait tout l'espace-temps. En chaque point du fluide, il y a une quadrivitesse ainsi qu'une vitesse ordinaire à trois dimensions. La quadrivitesse est notée par exemple $U^\mu(t, X^i)$. Si le fluide s'écoule, la vitesse peut être

différente en différents endroits. La quadrivitesse du fluide est un champ vectoriel sur l'espace-temps. Puisque c'est une quadrivitesse (c'est-à-dire le quadrivecteur contravariant dX^μ divisé par le champ scalaire $d\tau$) c'est automatiquement un quadrivecteur contravariant. U^μ se transforme exactement comme A^μ dans la tableau 4.1.

Champ	Transformation	Exemple
Scalaire :	Même valeur $\phi'(t', x') = \phi(t, x)$	Température Temps propre
Vectoriel :	Transformation de Lorentz $(A')^0 = \gamma A^0 - v\gamma A^1$ $(A')^1 = -v\gamma A^0 + \gamma A^1$ $(A')^2 = A^2$ $(A')^3 = A^3$	Déplacements X^μ, dX^μ 4-vecteurs

Tableau 4.1 : Transformations de champs scalaires et quadrivecteurs contravariants.

Comme on l'a vu, pour les quadrivecteurs covariants, l'indice est noté en position basse. Et pour les équations de transformation du référentiel sans prime vers le référentiel avec prime, il suffit de changer v en $-v$ (équation 4.23) dans la transformation de Lorentz. Le premier exemple que l'on a rencontré est le gradient relativiste d'un champ scalaire.

À partir d'un quadrivecteur on peut faire un scalaire. On l'a déjà fait quand nous avons construit le scalaire $d\tau^2$ à partir du quadrivecteur dX^μ :

$$d\tau^2 = dt^2 - dx^2 - dy^2 - dz^2$$

Nous pouvons suivre la même procédure avec *n'importe quel*

quadrivecteur. Si A^μ est un quadrivecteur, alors la quantité

$$(A^0)^2 - (A^1)^2 - (A^2)^2 - (A^3)^2$$

est un scalaire exactement pour les mêmes raisons. C'est laissé comme exercice.

Nous venons de voir comment fabriquer un scalaire à partir d'un quadrivecteur, et donc aussi un champ scalaire à partir d'un champ quadrivectoriel. C'est simplement la même opération en chaque événement de l'espace-temps.

Voyons une dernière fois l'opération opposée dans un cadre général, c'est-à-dire construire un champ de quadrivecteurs à partir d'un champ scalaire. On fait cela en différenciant le champ scalaire par rapport aux quatre composantes de temps et d'espace. Ensemble, ces quatre dérivées partielles forment un quadrivecteur covariant. Si le champ scalaire est ϕ , son gradient est

$$\frac{\partial \phi}{\partial X^\mu} = \left(\frac{\partial \phi}{\partial X^0}, \frac{\partial \phi}{\partial X^1}, \frac{\partial \phi}{\partial X^2}, \frac{\partial \phi}{\partial X^3} \right)$$

Insistons : c'est un quadrivecteur *covariant*. Quand on le représente avec une lettre, on place l'indice en position basse, par exemple B_μ . Et il est plaisant de noter que c'est déjà le cas de l'indice μ dans les dérivées partielles, qui est en position haute mais *au dénominateur*, ce qu'on peut considérer comme en position basse au numérateur. Et ces notations s'avèrent cohérentes.

4.4.3 Construction d'un lagrangien relativiste

Nous avons maintenant tous les outils nécessaires pour construire des lagrangiens relativistes. Nous savons comment les choses se transforment¹³ quand on change de référentiel, et nous savons construire des scalaires invariants à partir de quadrivecteurs et d'autres objets. Comment construit-on un lagrangien ? C'est simple. Le lagrangien lui-même – la quantité que nous calculons en chaque point t d'une trajectoire possible et multiplions par dt , ou en chaque point (t, x, y, z) et multiplions par le petit volume $dt dx dy dz$ sur l'espace sous-jacent d'un champ possible, puis sommes pour obtenir une intégrale d'action – ce lagrangien doit rester invariant dans tous les référentiels. En d'autres termes, le lagrangien doit être un champ scalaire.

C'est la seule contrainte fondamentale qui existe sur la construction d'un lagrangien. On se rappelle que *l'existence du lagrangien est un des principes fondamentaux de la physique, donc il ne fait pas l'objet de discussions*. Ensuite le lagrangien doit simplement conduire à des résultats cohérents et significatifs. C'est ainsi qu'on avait trouvé que le lagrangien d'une particule était non pas son énergie cinétique K plus son énergie potentielle V mais $K - V$ (Volume 1, pp 119-123). Ce dernier conduisait aux équations de Newton, donc il était acceptable. Et du reste, nous ne l'avons pas montré mais c'était essentiellement le seul possible. Nous allons procéder de la même façon pour trouver le lagrangien d'un champ relativiste.

13. Plus précisément, si "les choses" dont on parle ont une existence intrinsèque, comme les événements ou les quadrivecteurs de l'espace-temps, nous savons comment leurs *représentations* se transforment d'un référentiel à l'autre.

Nous prenons un champ $\phi(t, X^i)$ et considérons tous les champs scalaires possibles qu'on peut construire à partir de ϕ . Ces champs scalaires sont des candidats pour le lagrangien dans le problème considéré.

Regardons quelques exemples. Bien sûr, le champ ϕ étant lui-même un scalaire dans cet exemple, c'est un candidat. Mais toute fonction de ϕ l'est aussi. Si tout le monde est d'accord sur la valeur en chaque événement de ϕ , tout le monde sera d'accord sur la valeur de ϕ^2 , ϕ^3 , $\sin \phi$, $\sinh \phi$, etc. Par exemple le potentiel de champ $V(\phi)$ est un champ scalaire fonction de ϕ et donc un candidat pour faire partie du lagrangien. En fait on a déjà vu souvent $V(\phi)$ apparaître dans le lagrangien.

Quels autres ingrédients pourrions-nous utiliser ? Certainement nous voulons inclure les dérivées du champ, sans quoi notre théorie des champs serait triviale et sans intérêt. Nous devons simplement nous assurer qu'elles entrent en jeu de telle sorte que le lagrangien reste bien un champ scalaire. C'est facile. Tout d'abord nous utilisons les dérivées partielles du champ pour construire le quadrivecteur

$$\frac{\partial \phi^\mu}{\partial X}$$

Ensuite, nous utilisons ces quatre composantes pour construire le champ scalaire que nous connaissons

$$\left(\frac{\partial \phi}{\partial t}\right)^2 - \left(\frac{\partial \phi}{\partial x}\right)^2 - \left(\frac{\partial \phi}{\partial y}\right)^2 - \left(\frac{\partial \phi}{\partial z}\right)^2$$

Voilà une expression non-triviale qui a sa place dans un lagrangien \mathcal{L} . Quoi d'autre utiliser ? On peut multiplier par une constante. On peut même multiplier par n'importe quelle fonction d'un champ scalaire. Multiplier deux choses

invariantes en produit une troisième qui est invariante aussi. Par exemple, l'expression

$$\left[\left(\frac{\partial \phi}{\partial t} \right)^2 - \left(\frac{\partial \phi}{\partial x} \right)^2 - \left(\frac{\partial \phi}{\partial y} \right)^2 - \left(\frac{\partial \phi}{\partial z} \right)^2 \right] F(\phi)$$

serait un lagrangien légal.

Ce lagrangien commence à être un peu compliqué, et nous n'allons pas le considérer pour l'instant. Mais il est invariant par transformation de Lorentz et forme donc un champ scalaire candidat.

Nous pourrions même faire quelque chose d'encore plus laid. En principe, nous pourrions utiliser des dérivées d'ordre supérieur si nous parvenions à bâtir avec elles un champ scalaire. Mais cela nous entraînerait au delà des limites de la mécanique classique. Dans le cadre de la mécanique classique, nous pouvons utiliser des fonctions des coordonnées et leurs dérivées partielles d'ordre un. Pour une particule sur l'espace sous-jacent consistant seulement en l'axe temporel c'est ce que l'on a noté

$$\mathcal{L}(\phi(t), \dot{\phi}(t))$$

Des *puissances* supérieures des dérivées premières sont acceptables, mais des dérivées d'*ordre* supérieur ne le sont pas.

Ainsi en théorie de la relativité, si nous voulons rester dans le cadre de la physique standard, nous aurons dans le lagrangien des dérivées partielles d'ordre un par rapport au temps et par rapport à l'espace, et des puissances de celles-ci, mais pas de dérivées d'ordre supérieur.

4.4.4 Emploi du lagrangien

Malgré les restrictions imposées par la mécanique classique et l'exigence que le lagrangien soit Lorentz-invariant, nous avons une liberté considérable dans le choix du lagrangien avec lequel travailler. Examinons plus attentivement celui-ci

$$\mathcal{L} = \frac{1}{2} \left[\frac{1}{c^2} \left(\frac{\partial \phi}{\partial t} \right)^2 - \left(\frac{\partial \phi}{\partial x} \right)^2 - \left(\frac{\partial \phi}{\partial y} \right)^2 - \left(\frac{\partial \phi}{\partial z} \right)^2 \right] - \frac{\mu^2}{2} \phi^2 \quad (4.24)$$

C'est essentiellement le même lagrangien que dans l'équation 4.7, où l'on a réintroduit c pour être en unités conventionnelles, et où $V(\phi)$ a une expression spécifique. Vous pouvez maintenant voir pourquoi j'avais choisi ce lagrangien pour notre exemple non-relativiste de la section 4.3.1.

Le facteur $\frac{1}{2}$ n'est rien de plus qu'une convention sans signification physique. C'est le même facteur $\frac{1}{2}$ qui apparaît dans l'énergie cinétique ordinaire $\frac{1}{2}mv^2$. Nous aurions pu prendre mv^2 à la place. Si nous faisons cela, cependant, notre masse serait celle de Newton divisée par 2.

Dans le chapitre 1, nous avons vu qu'une transformation de Lorentz générale est équivalente à une combinaison de rotations spatiales rigides et d'une transformation de Lorentz simple le long de l'axe des x . Nous avons montré que l'expression 4.24 est invariante dans une transformation de Lorentz simple. Est-elle invariante par une rotation rigide des axes spatiaux ? Oui, car la partie spatiale dans l'expression 4.24 est la somme des carrés des composantes d'un 3-vecteur ordinaire. C'est le carré d'une longueur qui ne change pas dans une rotation. Cela établit que le lagrangien 4.24 est invariant par n'importe quelle transformation

de Lorentz générale.

Le lagrangien 4.24 conduit à l'une des plus simples des théories classiques des champs.

L'équation d'Euler-Lagrange appliquée au lagrangien 4.24 produit une équation analogue à celle obtenue à partir de l'équation 4.7. Il n'est pas difficile, en suivant la même procédure que précédemment, d'arriver à

$$\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \frac{\partial^2 \phi}{\partial x^2} - \frac{\partial^2 \phi}{\partial y^2} - \frac{\partial^2 \phi}{\partial z^2} + \mu^2 \phi = 0 \quad (4.25)$$

C'est une équation d'onde particulièrement simple car elle est linéaire. Le champ et ses dérivées n'apparaissent qu'à la puissance un. Il n'y a pas de terme en ϕ^2 , ou ϕ multiplié par une dérivée de ϕ . Par conséquent si on a deux solutions ϕ_1 et ϕ_2 , toute combinaison linéaire de ces deux-là est encore une solution.

Nous pouvons obtenir un champ encore plus simple en fixant $\mu = 0$. Il satisfait alors l'équation d'onde

$$\frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} - \frac{\partial^2 \phi}{\partial x^2} - \frac{\partial^2 \phi}{\partial y^2} - \frac{\partial^2 \phi}{\partial z^2} = 0 \quad (4.26)$$

4.4.5 Résumé des champs classiques

Nous disposons désormais d'une procédure pour développer une théorie classique des champs. Notre premier exemple était un champ scalaire, mais la même technique s'applique aussi à un champ vectoriel ou même un champ tensoriel. À partir du champ se figurer tous les scalaires qu'on peut fabriquer avec lui-même et ses dérivées d'ordre un. Une fois

tous les scalaires listés et caractérisés, construire un lagrangien à partir de fonctions de ces scalaires, par exemple des sommes de termes. Ensuite, appliquer les équations d'Euler-Lagrange. Cela revient à écrire les équations de mouvement du champ, ou les équations de champ de propagation d'une onde, ou quoi que ce soit d'autre que la théorie de champ en considération est supposée décrire. L'étape suivante est d'étudier les équations de mouvement obtenues.

Les théories classiques des champs doivent être continues. Un champ qui ne serait pas continu aurait des dérivées infinies, et par conséquent une action infinie. Il aurait aussi une énergie infinie. Nous aurons plus à dire sur l'énergie dans le chapitre suivant.

4.5 Champs et particules – un avant goût

Pour clore ce chapitre, je voudrais faire quelques remarques sur la relation entre les particules et les champs – pas dans un cadre quantique. Je voudrais seulement dans un cadre classique parler de l'interaction entre les particules ordinaires et les champs classiques. Si nous avons travaillé sur les lois de l'électrodynamique au lieu d'étudier un simple champ scalaire, j'aurais pu vous montrer comment les particules chargées interagissent avec un champ électromagnétique. Nous ne l'avons pas encore fait. Mais comment une particule pourrait-elle interagir avec un champ scalaire ϕ ? En quoi la présence d'un champ scalaire pourrait-elle affecter le mouvement d'une particule?

Pensons au lagrangien d'une particule en mouvement en présence d'un champ préétabli. Supposons que quelqu'un ait résolu les équations du mouvement, et que nous sachions que le champ $\phi(t, x)$ est une certaine fonction spécifiée du

temps et de l'espace. La particule pourrait être couplée au champ, d'une manière analogue à une particule chargée dans un champ électromagnétique. Comment la particule se déplace-t-elle ? Pour répondre à cette question, retournons à la mécanique d'une particule en présence d'un champ. Nous avons déjà écrit un lagrangien pour une particule. C'était $-m d\tau$, car $d\tau$ est essentiellement la seule quantité invariante disponible. L'intégrale d'action était

$$A = -m \int d\tau \quad (4.27)$$

Pour obtenir les solutions non-relativistes correctes pour des vitesses faibles, nous avons trouvé qu'il fallait le signe moins devant l'intégrale, et que le paramètre m se comportait comme une masse non-relativiste. En utilisant la relation $d\tau^2 = dt^2 - dx^2$, nous avons réécrit cette intégrale sous la forme

$$A = -m \int \sqrt{dt^2 - dx^2}$$

où dx représente un infinitésimal dans toutes les directions spatiales. Ensuite nous avons factorisé dt de la manière suivante

$$A = -m \int dt \sqrt{1 - \left(\frac{dx}{dt}\right)^2}$$

Le ratio dx/dt sous la racine carrée est la vitesse v , donc l'action est encore

$$A = -m \int dt \sqrt{1 - v^2}$$

Nous avons ensuite développé la quantité $\sqrt{1-v^2}$ en série entière et utilisé l'approximation fournie par ses deux premiers termes 1 et $-v^2/2$ (équation 3.18) valable quand v est très petite¹⁴.

Nous avons trouvé que le lagrangien $-m + mv^2/2$, qui peut être simplifié en $mv^2/2$ car une constante additive dans un lagrangien ne joue aucun rôle, était celui avec lequel nous sommes familiers pour une particule ordinaire libre, c'est-à-dire quand il n'y a pas de champ de forces créant une énergie potentielle.

Mais le nouveau lagrangien est relativiste. Et nous voulons regarder la particule couplée à un champ.

Que pouvons-nous faire à ce lagrangien relativiste pour représenter le fait que la particule soit couplée avec un champ ? Pour que le champ affecte la particule, il faut que le champ lui-même apparaisse quelque part dans le lagrangien. Il faut que nous l'insérions d'une manière qui soit Lorentz-invariante. En d'autres termes, nous devons construire un scalaire à partir du champ. Comme nous l'avons vu précédemment, il y a de nombreuses façons d'y parvenir, mais une action très simple que nous pourrions essayer est celle-ci

$$A = - \int [m + \phi(t, x)] d\tau$$

soit

$$A = - \int [m + \phi(t, x)] \sqrt{1-v^2} dt \quad (4.28)$$

Cela correspond au lagrangien

$$\mathcal{L} = - [m + \phi(t, x)] \sqrt{1-v^2} \quad (4.29)$$

14. Et nous avons utilisé implicitement le fait que deux lagrangiens proches conduisent à deux actions proches.

Ceci est l'une des choses les plus simples que l'on puisse considérer, mais il y a beaucoup d'autres possibilités. Par exemple, dans le lagrangien ci-dessus, on pourrait remplacer $\phi(t, x)$ par son carré, ou n'importe quelle autre fonction de $\phi(t, x)$. Pour le moment, nous allons utiliser le lagrangien donné par l'expression 4.29, et l'intégrale d'action correspondante de l'équation 4.28.

L'équation 4.29 est donc un lagrangien possible pour une particule se déplaçant dans un champ préétabli. Maintenant nous pouvons nous demander : comment la particule se déplace-t-elle dans ce champ ? C'est similaire à se demander comment une particule chargée se déplace dans un champ électrique ou un champ magnétique. On écrit le lagrangien pour la particule dans le champ électrique ou magnétique ; on ne se soucie pas de savoir comment le champ est arrivé là. À la place, on écrit simplement le lagrangien puis les équations d'Euler-Lagrange. Nous allons creuser un peu cet exemple. Mais avant de la faire, je souhaiterais signaler une caractéristique intéressante de l'équation 4.29.

4.5.1 Le champ mystère

Supposons que, pour une raison ou pour une autre, le champ $\phi(t, x)$ ait tendance à migrer vers une constante spécifique autre que zéro. Mettons qu'il "aime" être à une valeur particulière non nulle. Dans ce cas, $\phi(t, x)$ serait constant ou approximativement constant, malgré sa dépendance formelle à t et x . Le mouvement de la particule aurait alors exactement la même allure que si la particule avait la masse $m + \phi$. Reformulons cela :

Dans le champ scalaire ϕ approximativement constant, la particule de masse m se comporterait comme si elle avait la masse $m + \phi$.

Ceci est l'exemple le plus simple d'un champ scalaire causant un décalage apparent de la masse d'une particule.

Il y a dans la nature un champ qui ressemble beaucoup à celui de cet exemple. Devinez-vous lequel ? Pour vous mettre sur la voie : il a fait beaucoup parler de lui en juillet 2012 au CERN à Genève.

Notre champ scalaire a beaucoup de ressemblance avec le champ de Higgs¹⁵. Dans notre exemple, un décalage dans la valeur du champ scalaire cause un décalage dans la masse des particules. Notre exemple n'est pas exactement le mécanisme de Higgs, mais il en est très proche. Si une particule démarre avec une masse nulle et est couplée à un champ de Higgs, ce couplage a pour effet de décaler la masse de la particule vers une valeur non nulle. Ce décalage est en gros ce que veulent dire les gens quand ils disent que le champ de Higgs confère à la particule sa masse.

4.5.2 Encore un peu de lagrangien

Jetons encore un œil aux équations d'Euler-Lagrange dans notre exemple de champ scalaire. Nous n'allons pas pousser les calculs jusqu'au bout car ils deviennent rapidement inextricables. Nous allons juste voir le cheminement général.

15. En 1964, Peter Higgs, François Englert et Robert Brout, ainsi que plusieurs autres physiciens, prédirent l'existence d'une particule appartenant à la famille des bosons, résultant d'un mécanisme similaire dans un cadre quantique. Au CERN, 48 ans plus tard, une expérience réalisée avec le LHC (*Large Hadron Collider*) confirma leur prédiction.

Pour simplifier nous considérerons que nous sommes dans un espace-temps sous-jacent de dimension (1+1), c'est-à-dire un seul axe spatial en plus de l'axe temporel. Reconsidérons notre lagrangien de l'équation 4.29 où v est remplacé par son expression \dot{x} .

$$\mathcal{L} = -[m + \phi(t, x)] \sqrt{1 - \dot{x}^2}$$

La première étape quand on applique les équations d'Euler-Lagrange, en vue d'arriver à une équation différentielle sur $\phi(t, x)$ qui déterminera le mouvement, est de calculer la dérivée partielle de \mathcal{L} par rapport à \dot{x} . Rappelez-vous que calculer la dérivée partielle veut dire regarder le taux de variation de \mathcal{L} quand \dot{x} varie alors que toutes les autres variables indépendantes déterminant \mathcal{L} sont maintenues fixes. Dans ce cas, nous considérons l'expression entre crochets comme constante car elle n'a pas de dépendance explicite à \dot{x} . En revanche l'expression $\sqrt{1 - \dot{x}^2}$ dépend bien explicitement de \dot{x} . Quand nous prenons la dérivée partielle de \mathcal{L} nous obtenons

$$\frac{\partial \mathcal{L}}{\partial \dot{x}} = \frac{[m + \phi(t, x)] \dot{x}}{\sqrt{1 - \dot{x}^2}} \quad (4.30)$$

Cette expression devrait avoir un air familier. Nous avons obtenu un résultat presque identique dans le chapitre 3 quand nous avons calculé l'impulsion d'une particule relativiste, équation 3.30. La seule différence ici est le terme supplémentaire $\phi(t, x)$ ajouté à l'intérieur des crochets au numérateur. Cela renforce l'idée que l'expression $[m + \phi(t, x)]$ se comporte comme une masse *dépendant de la position*.

Toujours avec les équations d'Euler-Lagrange, l'étape suivante est de différentier par rapport au temps l'expres-

sion obtenue dans l'équation 4.30. Nous allons juste indiquer cette opération de manière symbolique en écrivant

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}} = \frac{d}{dt} \frac{[m + \phi(t, x)] \dot{x}}{\sqrt{1 - \dot{x}^2}}$$

C'est le membre de gauche de l'équation d'Euler-Lagrange. Regardons à présent le membre de droite

$$\frac{\partial \mathcal{L}}{\partial x}$$

Puisque nous voulons différencier partiellement par rapport à x , nous nous demandons si \mathcal{L} dépend explicitement de x . La réponse est affirmative, car $\phi(t, x)$ dépend de x , et il se trouve que c'est le seul endroit où \mathcal{L} dépend de x . La dérivée partielle est donc

$$\frac{\partial \mathcal{L}}{\partial x} = -\frac{\partial \phi}{\partial x} \sqrt{1 - \dot{x}^2}$$

Et l'équation d'Euler-Lagrange devient

$$\frac{d}{dt} \frac{[m + \phi(t, x)] \dot{x}}{\sqrt{1 - \dot{x}^2}} = -\frac{\partial \phi}{\partial x} \sqrt{1 - \dot{x}^2} \quad (4.31)$$

C'est l'équation différentielle que satisfait le mouvement du champ. Si on essaie de calculer la dérivée temporelle du côté gauche on arrive à une expression repoussante. Donc nous allons nous arrêter ici.

Comme exercice le lecteur ou la lectrice peut réintroduire le facteur c dans l'équation 4.31, et même voir comment le champ se comporte à la limite non-relativiste quand la vitesse \dot{x} est petite par rapport à c . Nous aurons plus à dire sur la question dans le chapitre suivant.