

# Relativité restreinte et initiation à la physique mathématique

## 1. Introduction générale

### 1.1. Qu'est-ce que la science mathématique ?

Pour le dire brièvement, c'est l'étude des systèmes éventuellement infinis mais constitués de purs objets élémentaires, dont l'existence est abstraite, indépendante du monde extérieur. Chaque constituant (élément, relation...) de ces systèmes a pour seule nature le fait d'être exact, du type "tout ou rien" : deux objets sont égaux ou différents, sont reliés ou ne sont pas reliés, une opération donne un résultat exactement...

Le choix de tels systèmes à étudier est lui-même également un choix mathématique, en ce sens qu'il est donné par des conditions de complexité limitée (contrairement par exemple à la biologie qui dépend des choix arbitraires secrètement accumulés par la Nature pendant des milliards d'années). La science mathématique est ainsi autonome par rapport aux contingences de notre Univers, mais peut toujours s'appliquer à chacun des innombrables lieux ou aspects particuliers de ce dernier où de telles situations de complexité limitée peuvent se rencontrer.

Bien que ces objets soient de nature indépendante de toute forme de représentation sensible (ou apparence) particulière, on ne peut les comprendre qu'en se les représentant sous une certaine forme. Il peut y avoir plusieurs formulations ou manières de se représenter une même théorie, toutes étant rigoureusement équivalentes entre elles mais leur efficacité, leur pertinence peut être très différente. Le plus souvent, les idées sont d'abord développées sous forme d'intuitions, d'imaginations plus ou moins visuelles, puis cristallisées en des formules et des phrases littéraires pour être communiquées, écrites ou travaillées sous ces autres formes. Et on ne peut se libérer d'une forme de représentation particulière qu'en développant d'autres formes de représentation et en s'exerçant à traduire les objets et les concepts d'une forme à l'autre.

### 1.2. Qu'est-ce que la géométrie ?

Étymologiquement, le mot *géométrie* signifie "mesure de la terre". C'est donc, à l'origine, une science reliée à la réalité concrète : celle de l'espace qui nous entoure, ou celle des plans : celui du sol lorsqu'il est plat, ou de toute surface plane (papier, tableau...), que l'on peut voir et sur laquelle on peut dessiner facilement. Le géomètre grec Euclide en a écrit le premier une formulation axiomatique dans le cas du plan, d'où le nom donné à la géométrie plane usuelle.

La géométrie euclidienne est ainsi la première *théorie physique*, c'est-à-dire une théorie mathématique qui modélise la réalité physique, en l'idéalisant : tandis que deux points trop proches l'un de l'autre peuvent difficilement être distingués par des mesures physiques, en mathématique il faut trancher : ils sont soit égaux, soit distincts (ainsi, entre deux points distincts se trouvent d'autres points, et ainsi de suite : dans un segment de droite de longueur finie il existe une infinité de points). Et on tranche au plus simple : cette géométrie se définit comme étant la plus simple théorie du plan ou de l'espace en accord avec l'expérience physique habituelle. C'est aussi l'unique théorie dans laquelle est vrai exactement tout énoncé simple qui semble vrai, aux incertitudes de mesure près, pour cette même expérience.

Pour un mathématicien moderne, le mot "géométrie" désigne le nom de famille que portent un grand nombre de théories mathématiques qui ont entre elles plus ou moins de liens de parenté, notamment le fait qu'il est plus ou moins possible de les visualiser, d'en dessiner des figures. D'un point de vue mathématique, elles sont toutes également "vraies", se rapportant toutes à des réalités mathématiques (des ensembles de "points"...).

### 1.3. Comment peut-on visualiser d'autres géométries ?

Mais comment visualise-t-on la géométrie euclidienne, d'abord ? Elle n'est pas celle que l'on voit directement, car la géométrie du champ de vision est la *géométrie sphérique*.

En effet, la vue d'un point ne nous renseigne que sur la demi-droite issue de l'œil auquel ce point appartient; et cette demi-droite peut se représenter par son point d'intersection avec une sphère centrée sur l'œil. Lorsqu'on bouge l'œil et/ou la tête, la modification de ce qu'on voit correspond à une rotation de cette sphère. L'exemple le plus clair en ce sens est celui des objets astronomiques dont on ne perçoit que l'image sur la "sphère céleste" (qui tourne à cause de la rotation de la Terre sur elle-même).

Mais la géométrie euclidienne (du plan ou de l'espace) est seulement celle qu'on a le plus l'habitude de *concevoir*. En effet, en voyant ou en imaginant une image de dimension 2 (sur la sphère de vision), on attribue à chaque point une sensation de proximité ou d'éloignement, et on est familier avec la manière dont cette perception se transforme lorsqu'on tourne l'objet ou qu'on change de point de vue par rapport à lui. C'est cette faculté, fruit d'une construction mentale, qui constitue notre compréhension de la géométrie, et notre agilité à l'employer en rend l'usage tellement facile comme celui d'une sensation élémentaire, que son caractère indirect en est presque anecdotique.

Or, la perception intuitive des autres géométries relève du même principe : celui qui consiste à *voir* une chose pour en *concevoir* une autre. Et pour parvenir à concevoir ainsi d'autres géométries en s'appuyant sur une vision qui (hélas) demeure la même, un bon moyen est de travailler sur cette faculté de concevoir l'espace euclidien, en la particularisant, en la modifiant ou en la généralisant. (Mais ensuite, il est vrai que ce lien avec la vision ne se préserve pas facilement, et que le travail de conception devient plus intense et traduit en formules avec un rôle de la vision relativement moindre).

Par exemple, on peut concevoir des espaces (euclidiens ou non) de dimension supérieure à 3. Ils ne comportent mathématiquement aucune véritable nouveauté par rapport à la droite, au plan et à l'espace usuel, même si l'absence totale d'expérience par rapport à eux donne aux novices l'impression d'une difficulté insurmontable. Pour concevoir ces espaces, on peut par exemple se dire qu'on ne voit que deux ou trois dimensions en même temps sous forme de coupe ou de projection, mais qu'il y en a d'autres en réserve, on peut les changer...

#### 1.4 La géométrie euclidienne décrit-elle exactement l'espace physique ?

Non. Non seulement le temps s'y ajoute en relativité restreinte pour former un espace de dimension 4 qui n'est pas euclidien (l'espace-temps) et est plus réel (proche de la réalité physique) que notre espace de dimension 3 usuel, mais même si on se restreint à considérer son apparence d'espace de dimension 3, la théorie de la relativité générale nous enseigne qu'en réalité, cet espace, pour autant qu'on puisse encore le mesurer, n'obéit pas exactement à la géométrie euclidienne (indépendamment de toute question de matérialisation des objets ou de précision des mesures). Bien sûr, l'erreur de le supposer euclidien est extrêmement faible, plus encore que d'autres erreurs beaucoup plus compréhensibles et déjà couramment négligées : lorsqu'on cartographie une région de la Terre (de forme à peu près carrée par exemple) de taille  $L$ , en l'assimilant à une région d'un plan, l'erreur géométrique normale due à la rotondité de la Terre vaut (en distance sur le terrain)

$$\frac{L^3}{R^2}$$

où  $R = 6370$  km est le rayon de la Terre, ce qui fait par exemple 2,4 cm pour une région large de 10 km.

Celle due à la relativité générale, qui mesure donc le même type de décalage par rapport à la géométrie euclidienne plane mais concerne le plan tangent à la Terre et constitue un effet supplémentaire par rapport à celui qu'on calculerait en termes de la seule rotondité de la Terre, vaut

$$\frac{2L^3 g}{Rc^2}$$

où  $g = 9,8 \text{ m.s}^{-2}$  est l'accélération de pesanteur et  $c = 3.10^8 \text{ m.s}^{-1}$  est la vitesse de la lumière. Elle est donc 720 millions de fois plus faible (précisément, pour n'importe quelle région choisie);

dans le cas de la mesure de la Terre entière au lieu d'une région, l'erreur est de l'ordre de quelques centimètres. Ainsi, étant donné un choix d'"équateur" parfaitement circulaire, sa longueur est inférieure à sa distance au centre multipliée par  $2\pi$  (mais le calcul précis de sa distance au centre dépend de la répartition de la masse de la Terre entre les différentes profondeurs).

Si maintenant on considère un plan vertical, une autre erreur survient, elle aussi plus forte que celle issue de la relativité générale : le solide de longueur  $L$  que l'on tourne pour comparer des longueurs horizontales et verticales se déforme sous son propre poids (il s'étire s'il est pendu, il se contracte s'il est posé), d'une amplitude de l'ordre de

$$\frac{L^2 g}{c'^2}$$

où  $c'$  est la vitesse du son dans le solide, qui est forcément inférieure à  $c$  d'après la relativité restreinte.

De l'autre côté, la physique de l'infiniment petit pose actuellement tellement de problèmes que certains physiciens théoriciens envisagent sérieusement que les concepts usuels de géométrie ne seraient plus valables à l'échelle de Planck (qui est bien plus petite que la plus petite distance pour laquelle le comportement des particules est actuellement connu).

### 1.5. Qu'est-ce que la physique mathématique ?

#### *État des lieux*

Il y a deux types de travaux en physique mathématique: la physique fondamentale et la physique appliquée. La partie théorique de la physique appliquée consiste essentiellement à développer les conséquences de la physique fondamentale à l'aide d'approximations adaptées (en s'aidant parfois de l'expérience quand les approximations sont trop difficiles à réaliser sur le plan théorique).

La partie physique de cet ouvrage aura pour but de présenter des théories physiques fondamentales bien établies et largement vérifiées à ce jour, des plus simples (mathématiquement simples et proches de l'expérience courante) aux plus évoluées (mathématiquement sophistiquées et à la base de l'explication du plus grand nombre de phénomènes). Mais cet ordre de rangement n'est pas strict. Du point de vue de la "réalité physique", seules les dernières auraient le titre de théories physiques fondamentales puisqu'elles sont physiquement plus fondamentales et précisément conformes à la réalité que les premières qui n'en sont que des résultats et approximations; mais il se trouve que l'ordre pédagogique des théories fondamentales (l'ordre de la compréhension dans leur apprentissage et le sens intuitif donné aux notions) est plus ou moins l'inverse de leur ordre de dépendance en tant que théories physiques. (C'est finalement logique puisque si une théorie  $A$  précède une théorie  $B$  à la fois physiquement et pédagogiquement, on classerait  $B$  dans la physique appliquée).

D'abord il y a la mécanique classique, celle de Galilée et Newton, qui traite des forces et des mouvements; la loi de la gravitation s'y énonce simplement. A partir d'elle on peut avancer dans deux directions, qui sont les deux types de physique moderne qui s'opposent à la physique classique : d'une part la Relativité (restreinte puis générale), d'autre part, la physique quantique et la physique statistique, qui sont deux théories parentes et étroitement liées entre elles. Ces théories seront introduites plus en détails dans les chapitres correspondants.

Nous ne préciserons pas les justifications théoriques et expérimentales sur lesquelles repose leur reconnaissance, car ces justifications sont actuellement innombrables et parmi elles, celles à l'origine historique de leur découverte ont désormais perdu tout caractère privilégié.

Une telle démarche pédagogique débarrassée des questions de justification expérimentale est largement satisfaisante du fait que ces théories bien établies et vérifiées avec une extrême précision, englobent déjà dans leur domaine d'application la plupart des expériences de physique réalisables. Cependant, beaucoup de ces expériences échappent en pratique à cette réduction à cause de l'extrême complexité de leur expression théorique, ou des difficultés parfois insurmontables de la résolution des problèmes mathématiques correspondants.

Ceci ne fait pas de la physique fondamentale une science morte pour autant, car non seulement les théories établies sont manifestement incomplètes (la relativité générale et la physique

quantique sont difficilement conciliables pour des raisons impossibles à vulgariser), mais la théorie quantique sous sa forme complète compatible avec la relativité restreinte (théorie quantique des champs) s'avère, au niveau fondamental, d'une complexité quasi-insondable manquant d'un fondement mathématique bien clair sur lequel reposer, et le travail de reformulation à la recherche d'un meilleur fondement demeure actif.

### *Aspects philosophiques*

Même si on supposait que les objets physiques aient une nature, nous ne pourrions de toute manière pas la connaître par l'expérience car toutes nos perceptions passent par une *traduction* des phénomènes extérieurs en la conscience qu'on peut avoir d'eux. Donc, le mieux qu'on puisse faire est d'étudier le monde physique en tant que système mathématique, ce qui nous laisse libres de nous le représenter sous des formes sans rapport avec celles sous lesquelles nous le percevons habituellement. Expliquons cela en détails:

Au moins jusqu'à maintenant, il semble que toute prétention à dire qu'un élément de l'imagination serait plus qu'un autre ressemblant à la "vraie nature" d'un certain élément de la réalité physique est vaine, parce qu'il n'existe aucun moyen de comparer ces natures : toutes nos perceptions du monde extérieur sont indirectes, traduites par notre corps en impulsions nerveuses, que le cerveau traite pour en distinguer les formes globales. Donc, du monde physique extérieur, seules les structures, c'est-à-dire les relations entre les choses, sont accessibles à nos sens et peuvent faire l'objet de nos discours, non une quelconque nature ontologique. Et de toute manière, comment espérer qu'un objet physique puisse avoir une quelconque identité de nature avec un phénomène mental ?

Par exemple il n'y a aucune ressemblance de nature entre le sentiment d'une couleur et les qualités des objets vus, qui sont responsables de cette couleur. Cette qualité consiste en des coefficients d'absorption de différentes fréquences d'ondes électromagnétiques par ces objets, qui déterminent la composition en fréquences de la lumière que l'oeil reçoit. Puis, ce n'est qu'une faible partie de l'information sur cette composition que les cellules de l'oeil peuvent transmettre (suivant les propriétés des molécules dans la rétine qui font ce travail).

Également, nous pouvons à peine imaginer le temps, même juste conformément à la manière dont nous le vivons, qui montrerait la nature de cette expérience vivante du temps au-delà de la seule forme de la droite qui le représente mathématiquement : en aucun instant nous ne pouvons imaginer une durée en tant que telle, puisqu'un instant ne contient aucune durée; il n'y a que les durées futures qui n'existent pas encore, et les durées passées, que nous percevons dans notre souvenir. Mais déjà, un souvenir d'une durée n'est plus une durée, car il ne fait que représenter cette durée par une certaine autre sensation instantanée.

Ainsi, la prétention de décrire les choses "comme elles sont" serait vaine, ou alors, ce n'est plus de la physique mais de la métaphysique, ce qui peut éventuellement se défendre à condition de l'annoncer clairement. Eventuellement, cela peut être utile lorsqu'on travaille avec une théorie provisoire qu'on est en train de remettre en question pour tenter d'en construire une autre plus précise comme explicitation de structures sous-jacentes à une structure donnée (par exemple, cela est pertinent pour passer des moles aux molécules, et de la thermodynamique classique à la mécanique statistique). Mais cela n'a lieu que de manière relative et non absolue, et n'est pas le cas général.

L'objet de la Physique est donc de décrire les lois physiques, c'est-à-dire l'expression sous forme de théorie mathématique des relations constatées entre les résultats de nos observations du monde physique, puisque c'est une caractéristique générale des mathématiques que de s'intéresser aux structures, relations entre les objets, et non à leur nature. Alors, entre plusieurs présentations mathématiquement équivalentes (ou traductions) d'une théorie, faisant différents choix de formes d'imagination pour représenter tel ou tel aspect de la réalité, la meilleure est celle qui permet de saisir cette théorie mathématique de la manière la plus facile ou efficace.

L'intuition quotidienne du monde physique est adaptée à la compréhension des problèmes de la vie quotidienne. Mais lorsqu'on veut passer à l'étude des structures de la réalité qui apparaissent dans d'autres contextes, il peut être judicieux de modifier ce choix. Mais de toute manière, un discours s'appuyant ainsi judicieusement sur un système inhabituel de correspondances entre objets

réels et formes d'imagination employées à les représenter, est donc bien sûr pareillement vide de signification sur la vraie nature de ces aspects ou objets de la réalité physique que tout autre discours.

Faute d'avoir une théorie complète de la physique, on a des théories partielles, dont chacune concerne les propriétés des résultats des expériences d'un certain domaine expérimental de la réalité, c'est-à-dire un certain type d'expériences avec un certain degré de précision des mesures (dépendant des paramètres de l'expérience). Au cours de l'exploration de la physique mathématique, on est amené à considérer une succession de théories correspondant à différents domaines expérimentaux, de plus en plus larges et dont les conséquences théoriques sont de plus en plus précisément en accord avec l'expérience (ou donnent de plus en plus d'informations à lui confronter).

Chacune de ces théories est d'abord par nature une théorie mathématique; sa signification physique (son titre de théorie physique) consiste alors à attribuer à chaque expérience du domaine expérimental considéré une traduction dans la théorie, c'est-à-dire un certain sous-système d'objets de la théorie qui représente mathématiquement le dispositif expérimental utilisé.

Mais une théorie a aussi une signification intuitive, qui réside dans le choix de la forme de représentation la mieux adaptée à la compréhension de la théorie. Dans cette exploration, beaucoup de théories géométriques interviennent, non seulement pour rendre compte du comportement de la matière évoluant à l'intérieur de l'espace qui nous est familier, mais aussi pour remplacer cet espace lui-même par d'autres (ce dernier cas n'arrive pas souvent : principalement en Relativité restreinte, Relativité générale et ce qui inclut la gravitation quantique). Et, parmi nos intuitions ou capacités habituelles à percevoir le monde physique, la vision est celle qui est la plus claire mathématiquement, et la plus facile à exercer pour pouvoir, avec elle, concevoir beaucoup de ces théories géométriques.

La difficulté du naïf à remettre en question la forme de ses perceptions se présente aussi en termes du principe de réalité, suivant lequel "la réalité physique existe" que nous la percevons ou pas, au nom de quoi il voudrait continuer à penser les choses en termes d'"objets réels" ou concrets tels qu'il les perçoit habituellement. Or le problème n'est pas de savoir si la réalité comporte ou non d'autres objets que ceux de notre expérience ou perception, car de toute manière les théories physiques avec lesquelles nous travaillons en présentent. Le problème est de savoir lesquels. Le naïf imagine ces objets de la réalité sous-jacente, qu'il ne peut percevoir directement, comme ressemblant à ceux qui lui sont familiers. Il n'a pas toujours tort, d'ailleurs, dans la mesure où tel ou tel phénomène qu'il a besoin d'expliquer peut être dû à des mécanismes situés dans un domaine expérimental plus large que son expérience courante, mais pas encore assez pour que son intuition cesse d'être valide, son incapacité à observer directement ces mécanismes suivant cette même intuition étant seulement contingente et remédiable par des instruments adaptés (microscopes, télescopes, etc.).

Par contre, lorsqu'en poussant plus avant le domaine expérimental on se trouve confronté à une impossibilité plus fondamentale d'affiner les instruments de mesure pour observer le même type de mécanismes réels explicatifs, il est temps d'envisager qu'il y ait à cela une bonne raison: à savoir qu'il n'existe pas de réalité explicative ressemblant à ce que le naïf voudrait croire, parce que la réalité est d'une autre forme qui ne peut être comprise que par des moyens différents, en des termes différents.

## 1.6. Quelques mots de géométrie euclidienne

Évoquons ici les géométries euclidiennes de dimensions 1, 2 et 3. Celle de dimension 1 (de la droite) est la plus simple mathématiquement; les deux suivantes (du plan et de l'espace) sont presque de complexité équivalente, et seront traitées ici de la même manière.

Notre vision est habituée à comprendre le plan et l'espace, mais se trouve un peu large pour parler de la droite : quelle droite ? horizontale, verticale... ? On a tendance à vouloir placer cette droite quelque part. On a déjà moins tendance à vouloir disposer le plan dans l'espace, puisque les deux dimensions du plan remplissent les deux dimensions du champ de vision, ne laissant pas voir le vide qui se trouve de chaque côté. Mais quels que soient nos modes de représentations et d'éventuels rapports avec notre Univers physique, chacune de ces géométries, en tant que théorie mathématique, se limite à considérer les constructions pouvant être faites à l'intérieur de son

système (la droite, le plan, l'espace) indépendamment de toute considération d'un éventuel espace extérieur dans lequel ce système pourrait être disposé. Même si une telle disposition peut être utile pour des calculs et raisonnements intermédiaires ou comme support intuitif, elle n'est pas considérée comme faisant partie des réalités premières dans cette théorie, et doit donc disparaître dans les objets d'étude et les résultats finaux. Cette indépendance de chaque géométrie apparaît clairement dans l'exemple déjà évoqué du champ de vision, qui a la forme (ou géométrie) d'une sphère : cette sphère a certes un centre bien défini (l'œil) mais n'a pas de rayon, il serait faux de la confondre avec une sphère particulière disposée dans l'espace. (Elle est l'ensemble des demi-droites issues de l'œil.)

Pour cette raison, il est bon de fournir à disposition de notre intuition des formes de représentations autres que visuelle pour figurer la géométrie de la droite (ou géométrie de dimension 1). En voici deux. D'abord, la sensation du temps. Ensuite, la notion de fil (infiniment fin), en tant qu'objet mécanique dont la nature est préservée quelles que soient les dispositions dans l'espace (droites, courbes, lignes brisées...) qu'on lui donne. Présentons maintenant une introduction intuitive à la géométrie euclidienne.

Considérons des fils de longueur finie, arrêtés en deux extrémités (on les coupe là, ou bien on oublie les prolongements comme s'ils n'existaient pas). On appellera *courbe* une image d'un fil, lieu du plan ou de l'espace en lequel il peut être disposé. Chaque fil se caractérise par sa longueur, qui est une quantité positive. Pour toute courbe, les fils pouvant être disposés dessus sont tous ceux qui ont une même longueur, qu'on appelle alors la longueur de la courbe. Tout point d'un fil le divise en deux parties, dont la somme des longueurs est égale à la longueur du fil de départ. La longueur d'une courbe se mesure également par le temps qu'on met à parcourir cette courbe du regard à vitesse fixée (ainsi sont reliées ces 3 intuitions unidimensionnelles: les fils, le temps et les courbes).

Marquer le fil d'une graduation régulière repérant les longueurs depuis une extrémité (ou plus généralement depuis un point quelconque du fil), ou de manière équivalente chronométrer le temps de parcours à vitesse constante, fournit un marquage de chaque point de la courbe par un nombre ou une quantité appelé son *abscisse curviligne*.

Entre deux points distincts  $A$  et  $B$ , il y a une valeur minimale non nulle des longueurs de courbes qui les relient; c'est la seule longueur pour laquelle il n'existe qu'une seule courbe qui les relie. Cette courbe est notée  $[AB]$ , et est appelée un *segment*. Sa longueur s'appelle la *distance* entre  $A$  et  $B$ . Inversement, la longueur d'une courbe peut se définir à partir de la notion de distance (sans utiliser la mesure des longueurs des fils) : c'est la valeur vers laquelle s'approche la somme des distances entre des points de la courbe consécutifs pour le fil, lorsque ces points sont de plus en plus nombreux et rapprochés dans le fil (envahissant toutes les parties du fil). On pourrait définir à partir des segments les demi-droites et les droites, ou inversement les segments à partir des droites définies en termes de symétrie axiale qui sont des antidéplacements (voir plus bas), mais passons.

La notion de disposition d'un fil dans le plan ou dans l'espace se généralise à la disposition d'un morceau de plan dans l'espace (une feuille de papier), d'un volume dans l'espace (un solide), ou d'un morceau de plan dans un plan (une feuille de papier posée sur une table).

Exprimons-nous maintenant dans le cas de la géométrie plane.

L'ensemble des points à égale distance  $d$  d'un point donné  $O$  s'appelle le *cercle* de centre  $O$  et de rayon  $d$ . Une droite divise le plan en deux demi-plans. L'intersection ou la réunion de deux demi-plans s'appelle un *secteur angulaire*  $S$  qui se mesure par son *angle* :  $O$  étant à l'intersection des deux droites frontières, l'intersection de  $S$  avec un cercle de centre  $O$  est une courbe, dont la longueur divisée par le rayon est la mesure en radians de l'angle (qui ne dépend pas du rayon).

(Ce qui suit serait généralisable à la géométrie dans l'espace en remplaçant feuille par solide et table par espace.)

Chaque disposition d'une feuille sur une table fait voir une figure sur la feuille comme étant une figure sur la table.

On appelle *déplacement* la transformation d'une figure (dessinée sur la feuille mais vue comme étant sur la table), qui résulte d'un changement de disposition de la feuille sur la table réalisé par simple glissement, sans retournement. La notion d'*antidéplacement* se définit de la même manière

sauf que la feuille (de préférence un papier calque) est de plus retournée. (Dans le cas de l'espace, le "retournement" ne peut être que virtuel, et consiste à prendre l'image miroir). Les déplacements conservent l'orientation du plan (le sens des aiguilles d'une montre), les antidéplacements les renversent.

Tout déplacement dans le plan est soit une rotation, obtenue en tournant la feuille autour d'un point (appelé centre et unique point fixe) suivant un certain angle et dans un certain sens (défini par rapport à l'orientation du plan), soit une translation (qui avance la feuille sans la tourner, et peut être vue comme la limite d'une rotation d'angle infiniment petit autour d'un centre infiniment éloigné).

**Théorème et définition.** On appelle *isométrie* une transformation  $f$  du plan (ou l'espace) dans lui-même qui vérifie les trois conditions équivalentes suivantes :

- 1)  $f$  conserve les distances (ou les longueurs de courbes)
- 2)  $f$  conserve toutes les notions de géométrie euclidienne
- 3)  $f$  est un déplacement ou un antidéplacement.

### 1.7. Fondement expérimental de la géométrie affine plane

Nous définirons la géométrie affine du plan par l'expérience suivante. On observe de loin une figure tracée sur une vitre; on ne sait pas (et on ne saura jamais) à quelle distance cette vitre se trouve, ni comment elle est disposée (de face ou de biais; on ne voit pas son contour), ni si on la voit par devant ou par derrière (ou, ce qui revient au même, à travers un miroir). Différents observateurs seront toujours supposés ignorer même comment ils sont disposés dans l'espace les uns par rapport aux autres.

On sait seulement ce qu'on en voit, et la taille apparente en est petite (elle pourrait tenir dans le disque lunaire par exemple) : on peut la prendre en photo, et faire des mesures géométriques sur la photo. Le passage de la figure réelle à la figure ainsi vue s'appelle une *transformation affine*.

Dans ces conditions, quels sont les notions de géométrie (propriétés ou mesures d'une figure – on appellera cela plus loin les *structures*) dont l'apparence de tout point de vue de ce type est toujours fidèle à la réalité, aux incertitudes de mesure près (négligeant les effets de distorsion liés au quotient de la taille de l'objet à sa distance) ?

On appelle *géométrie affine du plan* la géométrie du plan\* réduite à l'étude des notions qui apparaissent fidèlement dans cette expérience. Elle diffère donc de la géométrie euclidienne à laquelle nous sommes habitués, par le fait que son langage est plus réduit : lorsque deux observateurs placés en des points de vue différents se communiquent leurs observations d'une même figure par téléphone, en ayant seulement le droit de s'exprimer dans ce langage, ils seront toujours d'accord. Et ils peuvent même introduire de nouveaux termes, tant qu'ils peuvent les définir entièrement au moyen de ce langage, ils resteront toujours d'accord.

Ainsi, toutes les notions de géométrie affine sont des notions qui existent aussi en géométrie euclidienne. Mais certaines notions habituelles de géométrie euclidienne n'étant plus valables, on a alors besoin, pour pouvoir encore travailler, de recourir à des notions moins habituelles qui subsistent en géométrie affine, pouvant correspondre aux anciennes notions pour des observateurs particuliers.

Un même objet de géométrie affine admet de multiples interprétations (ou mesures) possibles dans le langage de la géométrie euclidienne (chaque observateur a son interprétation définie par ce qu'il voit), et aucune mesure ou propriété qui utilise seulement le langage affine ne permet d'exprimer comment distinguer "quelle est la vraie". N'importe laquelle peut être considérée comme vraie du point de vue correspondant.

La notion de droite fait partie de la géométrie affine. Toutes les notions définissables à partir de notions de géométrie affine, font aussi partie de la géométrie affine. Par exemple, font partie de la géométrie affine les notions d'alignement des points, de demi-droites, segments, parallèles et

---

\* Voici pourquoi on se permettra de parler de plan (infini) bien que la figure soit supposée tenir dans le disque lunaire : il suffit de partir de figures beaucoup plus petites que le disque lunaire (on a donc besoin d'un télescope pour l'observer), à partir de quoi le disque lunaire semblera presque infiniment grand comme un plan: on peut y continuer la figure librement. . .

parallélogrammes, rapports de longueurs de segments parallèles, rapports des aires, barycentres, translations et homothéties, paraboles, ainsi que les ellipses et hyperboles avec leurs centres.

En fait, la notion de droite est suffisante pour définir toutes les autres notions de géométrie affine, c'est-à-dire qui ne dépendent pas de l'observateur dans l'expérience ci-dessus. Par suite, toute transformation qui transforme les droites en droites préserve aussi toutes ces notions et est une transformation affine.

Ne font pas partie de la géométrie affine les notions de distances, longueurs de courbe ou de segment, d'angles, d'orthogonalité, de cercles, de rotations, de produit scalaire, d'axes et de foyers de coniques, car des observateurs placés en des lieux différents peuvent ne pas être d'accord dessus.

Par exemple, si l'un croit voir un cercle, un autre ne verra pas un cercle mais seulement une ellipse. Ainsi peut-on définir la notion d'ellipse: une ellipse est une figure obtenue par transformation affine à partir d'un cercle. De même, les rotations et antidéplacements de la géométrie euclidienne sont des transformations affines du plan qui en géométrie affine ne sont plus distinguables dans un ensemble plus large de transformations affines.

On a vu en géométrie euclidienne les dispositions de fils dans le plan, qui sont certaines applications d'un fil dans le plan. Ici, il n'y a plus lieu de les distinguer à l'intérieur d'un certain ensemble plus vaste d'applications d'un fil dans le plan, par exemple celui des *courbes paramétrées* (le fil est remplacé par un élastique; si on garde l'intuition du temps pour imaginer le fil, la vitesse de parcours de la courbe peut alors varier d'une manière quelconque, différente d'un point de vue à l'autre suivant la transformation affine subie).

## 2. L'étrangeté de la théorie de la Relativité

Ce chapitre d'introduction à la relativité et à son étrangeté vise notamment à analyser la méthodologie sur laquelle reposent traditionnellement les cours d'initiation à la relativité restreinte, et à motiver l'adoption d'une nouvelle approche reflétant mieux l'usage de cette théorie dans un contexte plus large de physique théorique. Il s'adresse particulièrement à ceux ayant vu la présentation habituelle, mais devrait être au moins partiellement compréhensible par tous.

Il est possible d'aborder directement les choses sérieuses en passant au chapitre 3.

### 2.1. Son nom et quelques autres aspects extérieurs

La relativité est en deux étapes nommées "restreinte" et "générale". La théorie de la relativité restreinte est une théorie de l'espace et du temps qui remplace celle précédemment énoncée par Galilée et Newton, en réduisant cette dernière au rang d'approximation valable lorsque les vitesses relatives des objets étudiés sont négligeables devant celle de la lumière. (Nous sous-entendons parfois l'adjectif "restreinte", comme nous n'aborderons pas ici la relativité générale).

Une conséquence de la relativité est qu'aucun signal, aucune information ne peut être transmis plus vite que la vitesse de la lumière dans le vide (qui est aussi celle des ondes radio), pas plus qu'elle ne peut arriver avant d'être partie, pour des raisons de non-contradiction de la réalité. Cette vitesse est une constante universelle  $c \approx 3.10^8 \text{ m.s}^{-1}$ , ce qui signifie précisément que celui qui envoie un signal à cette rapidité indépassable, sous forme d'ondes radio par exemple, à un point situé à  $3.10^8 \text{ m}$  de là d'où il est aussitôt renvoyé, devra attendre deux secondes avant d'en recevoir le retour.

Il ne faut pas se laisser abuser par le nom de la "relativité" qui reflète mal le sens de cette théorie : certains auteurs de la théorie l'ont par la suite regretté, pensant que celui de "chronogéométrie" par exemple aurait mieux convenu, mais il était trop tard. Le slogan caricatural "tout est relatif" avait déjà pris possession de l'imagination populaire, mais son sens n'était ni nouveau ni correctement représentatif de la théorie. En effet, d'une part le Principe de relativité était déjà contenu dans la théorie galiléenne, la seule différence étant que le changement de référentiel affecte un plus grand nombre de paramètres (par exemple le temps) qui décrivent des objets physiques donnés, en théorie de la relativité par rapport à la théorie galiléenne (et ce principe est ainsi réhabilité après que l'électromagnétisme a semblé le contredire). D'autre part, ce slogan risque de provoquer par exemple des blocages psychologiques à l'acceptation du caractère objectif et mesurable du mouvement de rotation d'un objet sur lui-même (gyroscope, pendule de Foucault...).

La relativité restreinte a été découverte par différents chercheurs : Poincaré, Lorentz, Minkowski, Einstein. La relativité générale a été découverte par Einstein et Hilbert; elle englobe la relativité restreinte et la gravitation (qu'elle interprète comme n'étant pas une force mais l'effet de la courbure de l'espace-temps. . . ).

Entre les deux on peut placer l'électromagnétisme, pour son niveau de complexité (la relativité générale n'en dépend pas en toute rigueur mais utilise à peu près les mêmes outils mathématiques). La relativité générale cohabite naturellement avec l'électromagnétisme : le champ électromagnétique et le champ gravitationnel interagissent de manière cohérente tout en restant des objets distincts.

On peut remarquer un outil mathématique qui traverse un grand nombre de théories physiques : le calcul tensoriel. Pour résumer, ce formalisme consiste en assemblages d'opérations générales de "multiplications" entre des vecteurs pouvant appartenir à différents espaces vectoriels. Il s'exprime dans toute son étendue en théorie quantique des champs où n'importe quelle expression tensorielle qu'on peut écrire à partir de certains objets (donnés par les types de particules présents) se réalise par le diagramme de Feynman correspondant.

Son absence apparente des cours habituels d'électromagnétisme est en fait une maladresse : certains calculs (en mécanique classique et électromagnétisme) utilisent des formules vectorielles ou autres règles de transformations d'expressions qui semblent alors d'origine très mystérieuse ou nécessitent des démonstrations compliquées, mais deviendraient simples et naturelles une fois traduites en langage tensoriel. La raison pour laquelle cette situation perdure est qu'il ne semble pas exister actuellement dans la littérature une présentation suffisamment claire et pédagogique du calcul tensoriel pour qu'on puisse songer à l'introduire à ce niveau d'enseignement.

Historiquement, l'électromagnétisme a été développé avant la relativité restreinte, et actuellement on enseigne encore l'électromagnétisme d'abord en premières années d'université, pour en déduire qu'il oblige à construire la relativité restreinte d'après la loi de constance de la vitesse de la lumière qu'il induit. Mais cette démarche héritée de l'histoire, perpétuée par inertie, manque fortement d'élégance (sans relativité, le champ électromagnétique se présente sous forme de deux objets, champ électrique et champ magnétique, obéissant aux quatre équations de Maxwell qui semblent bien peu naturelles, alors qu'une ou deux équations tensorielles sur un seul objet suffit en partant de la relativité). Il serait plus élégant et efficace de commencer par la relativité restreinte et le calcul tensoriel, si seulement des approches convenables de ces théories étaient trouvables dans la littérature.

L'électromagnétisme dans le cadre relativiste s'appelle également l'*électrodynamique classique* (par opposition à l'*électrodynamique quantique* qui traite cette force en physique quantique). On peut concevoir cette théorie comme une extension de l'électrostatique, de la manière suivante.

Il y a une similitude formelle entre les théories de l'électrostatique et de l'induction magnétique (dans l'espace tridimensionnel) qui se correspondent en remplaçant les charges de dimension zéro (densités) par des charges de dimension 1 (courants). Une fois cela remarqué en termes de calcul tensoriel, il suffit de réappliquer cette même théorie, formellement indépendante du choix d'une dimension particulière, au cas de l'espace-temps de la relativité, pour obtenir l'électrodynamique classique. Le champ magnétique et le champ électrique ne sont que les deux composantes d'un même champ décomposé par rapport à un référentiel choisi.

Cette théorie présente l'étrange défaut de ne pas pouvoir considérer les électrons comme ponctuels en toute exactitude sans conséquences absurdes, mais seulement suivant des approximations, en renonçant à définir la position d'un électron avec une précision plus fine qu'une distance de l'ordre du *rayon classique*\* de l'électron  $r_e = 2,82 \cdot 10^{-15}$  m. Cependant cette taille est

---

\* Comparable à la taille du noyau atomique, c'est le double du rayon de la sphère à l'extérieur de laquelle l'énergie du champ électrique calculé classiquement atteint l'énergie de masse  $mc^2$  de l'électron : vu de plus près que cela, si l'électron était ponctuel, en lui soustrayant cette énergie et donc cette masse qui se trouve dans le champ qui l'entoure, il lui resterait une masse négative, défiant les lois de la mécanique. . . mais en fait, l'énergie de masse du champ magnétique dû au spin de l'électron est plus importante encore près de ce rayon classique, et donnerait donc un rayon plus grand.

bien inférieure à la fameuse “incertitude” de la physique quantique (dans le cadre de laquelle ce problème de l’énergie du champ se résoud mais de manière très compliquée).

## 2.2. L’origine de ses paradoxes : l’intuition galiléenne

Vue du haut d’une tour, une voiture peut s’éloigner tant qu’elle veut, indéfiniment; elle semble rapetisser et ralentir au fur et à mesure qu’elle s’approche de l’horizon, bien qu’il ne lui arrive en réalité rien de spécial à son approche. Elle ne pourra jamais atteindre cet horizon si ce n’est à cause de la rotondité de la terre, ni encore moins le dépasser. Ces faits n’étonnent personne. Pourquoi s’étonnerait-t-on alors des fait suivants qui leur sont pourtant analogues : un voyageur peut accélérer tant qu’il veut sans qu’il ne lui arrive quoi que ce soit, tandis que quelqu’un qui se prétendrait immobile le verrait ainsi atteindre des vitesses arbitrairement proches de celle de la lumière, sans jamais l’atteindre, mais en subissant à son approche quelques distorsions, dont un ralentissement du temps et une contraction des longueurs dans le sens de sa vitesse (qui n’apparaît d’ailleurs même pas visuellement comme une contraction mais comme une rotation; voir les détails plus loin) ?

Cet étonnement, qui traduit une légitime soif de comprendre, est dû à notre tendance naturelle à penser le problème par ce que (à moins que vous ayez une meilleure suggestion) nous appellerons “l’intuition galiléenne”, que nous allons maintenant expliquer.

Nous percevons le monde sous forme d’images reçues par les yeux et instantanément interprétées par le cerveau en images tridimensionnelles. On a ainsi une succession continue d’images, ou pour le dire autrement une image en évolution. De même dans l’activité théorique, on imagine des figures ou des images qui se succèdent ou évoluent au cours du temps. L’intuition galiléenne est alors le fait d’utiliser cette dimension temporelle de notre imagination, cette faculté à faire évoluer dans notre imagination les images au cours du temps, pour représenter la dimension temporelle de l’univers physique que l’on cherche à concevoir.

Ce mode de représentation était adapté à la théorie galiléenne, qui autorisait les interactions instantanées à distance et dans laquelle la vitesse de la lumière était infinie. En effet, l’image qu’un observateur reçoit du monde à un instant y correspondait à l’image qu’un autre observateur éloigné du premier (et pouvant être en mouvement par rapport à lui) en reçoit aussi à un certain instant (donc au “même” instant), par une simple transformation géométrique (un déplacement). C’est donc une image de dimension 3 que le théoricien peut imaginer à un certain instant, et à laquelle il peut attribuer une réalité objective.

Mais cela devient faux en relativité, du fait que la vitesse de la lumière est finie et indépassable par l’information. L’intuition galiléenne n’est donc pas adaptée pour concevoir cette théorie.

## 2.3. Le piège méthodologique de son enseignement actuel

La plupart des cours de relativité restreinte, même récents, que l’on trouve sur le marché, emploient encore plus ou moins la même démarche fidèle à l’histoire de sa découverte, loin de toute tentative de rénovation réputée impossible ou sans objet, tellement cette méthode a été répétée plus ou moins telle quelle depuis bientôt un siècle: au nom du sens des principes et réalités expérimentales, ils commencent par considérer comme réalité première le type d’apparences du monde perçues par les misérables points dans l’univers que nous sommes. S’enfermant ainsi dans l’intuition galiléenne, ils ne peuvent alors qu’entreprendre d’y faire entrer également coûte que coûte la réalité de l’espace-temps physique.

La correspondance entre réalité physique et intuition galiléenne que donne naturellement l’observation n’étant pas satisfaisante, ils en construisent une autre artificielle plus satisfaisante pour l’esprit, appelée “référentiel galiléen”, au moyen d’un dispositif expérimental fictif, complexe et pratiquement irréalisable. Cette notion de référentiel galiléen qui était héritée de la théorie galiléenne, sera finalement un outil mathématique équivalent à la notion de repère orthogonal en géométrie euclidienne (ou parfois à la donnée d’un seul vecteur de base de ces repères, à savoir le vecteur temps) : notion utile mais non essentielle à l’étude de la géométrie, et ici détournée de son sens au profit d’une intuition inadaptée. (En effet, elle déclare instantanés des événements indépendants, donnant à “voir” en imagination un événement “présent” éloigné dont la lumière n’arrivera que plus tard, comme si on pouvait voir venir un signal lumineux avant qu’il

ne nous parvienne. . . ) (Le qualificatif “galiléen” des référentiels était dû au fait que dans la théorie galiléenne on pouvait aussi considérer des référentiels “non-galiléens” sans trop de complications, ce qui n’est plus le cas : car les systèmes matériels accélérés aussi étudiables, qui sont sensés définir ces autres référentiels, sont alors de plus mauvais points de vue pour l’intuition galiléenne; un tel emploi de cet intuition correspondrait mathématiquement à l’emploi de systèmes de coordonnées curvilignes.)

Naturellement incapables d’articuler cette intuition conformément à la théorie à concevoir, ils déclarent que la théorie ne peut être appréhendée que par la rigueur des formules et des démonstrations, en rejetant toute inspiration intuitive. Mais en même temps, au nom du sens des réalités, ils continuent à s’accrocher à cette intuition désespérément. Car les formules ne peuvent exister à l’état pur sous peine de perdre toute signification, elles ne peuvent qu’être des relations entre des paramètres qui correspondent à la réalité, laquelle se définit comme objet de nos sens et donc de notre intuition, oui mais lesquels et suivant quelle correspondance ? C’est ce choix malheureux d’un mode de représentation fixé arbitrairement, ni fidèle à l’expérience dans sa relation à la réalité, ni adapté à la compréhension interne de la théorie à concevoir, qui détermine l’apparence des formules “fondamentales” et leur complexité.

Alors, faute de pouvoir donner un sens physique intuitif à une formule, ils la consacrent de la force d’autorité que lui donne la longueur des preuves du fait qu’il ne peut exister aucune théorie de l’espace-temps ou de la mécanique (compatible avec certains principes formulés expérimentalement), aussi tordue soit-elle, dans laquelle elle ne serait rigoureusement vraie. Ainsi, au nom du sens des réalités physiques, ils prouvent l’unicité de la théorie sans la donner à comprendre vraiment, en oubliant presque d’en justifier l’existence (c’est-à-dire la non-contradiction), satisfaits pour cela de constater que l’univers existe. (D’où la prolifération de pseudo physiciens révolutionnaires qui prétendent réfuter la relativité par de simples expériences de pensée, la présentant comme une interprétation absurde et médiocre des fameuses expériences).

#### **2.4. Pour une nouvelle approche de la théorie**

Nous aborderons la relativité restreinte par une présentation intuitive qui consiste à employer, à la place de l’intuition galiléenne, une intuition géométrique purement spatiale.

En voici le motif. Nous percevons les dimensions de la réalité physique sous forme de trois sensations : images visuelles (de dimension 2), profondeur et temps; elles fournissent naturellement trois formes d’imagination pour le théoricien. On reconnaît que les trois dimensions d’espace sont physiquement de même nature et en continuité entre elles malgré leur différence de forme sensible (vision et profondeur). Cette identité de nature et cette continuité apparaissent intuitives lorsque par une rotation de l’objet la profondeur s’échange avec une des deux dimensions visuelles. Alors, généralisons ce raisonnement en supposant que la dimension temporelle elle aussi est physiquement de même nature que les trois autres, et rendons ce fait intuitif en représentant le temps lui aussi par une dimension visuelle de l’imagination (en fait l’identité de nature n’est pas complète mais cette démarche n’en est pas moins pertinente, et la différence sera précisée plus loin).

On pourra être tenté de reprocher à cette méthode l’absence de démarche rigoureuse fondée sur la réalité (les énoncés de principes physiques issus de constatations expérimentales) pour construire et prouver la théorie (ou pour le dire plus clairement, le fait qu’elle ne se fonde pas sur l’intuition galiléenne). Mais son objectif est de donner une claire vision des choses sans s’embarrasser au départ de leurs apparences expérimentales. La cohérence et les conséquences de cette vision montreront naturellement qu’elle décrit un monde possible compatible avec les fameux principes physiques. Les lecteurs intéressés par la preuve qu’il n’y a pas d’autre monde possible, à part la relativité restreinte et la théorie galiléenne, sont donc invités à se référer pour cela aux cours traditionnels sur ce sujet. (Remarquons en passant que la relativité générale décrit justement un autre monde plus proche de la vérité. . . )

On pourra aussi lui reprocher de n’être pas plus simple, voire d’être plus difficile à comprendre que la méthode traditionnelle. Déjà, les gens habitués à cette dernière, ou même ceux qui ayant commencé à l’apprendre voudraient qu’on la leur explique mieux, risquent de mal accepter de reprendre les choses à zéro en changeant ainsi complètement d’approche. Mais peut-être même certains, suivant leur manière de penser, pouvaient réellement mieux s’y retrouver. En effet, elle avait

en quelque sorte l'avantage de permettre de formuler la théorie et résoudre des problèmes sans faire vraiment le travail de la comprendre et de rentrer dedans : "il suffit" d'écrire les démonstrations et d'appliquer les formules. Cependant, cet avantage qu'on peut trouver jusqu'à la formulation de la théorie risque de se transformer en handicap ensuite, lorsque face à un problème concret on se perd devant ces formules qu'on ne sait pas toujours comment appliquer ou manipuler, ou dont l'utilisation nécessite des pages de calcul.

Enfin, on pourrait être tenté de penser qu'elle ne serait pas complète, omettant les formules consacrées réputées nécessaires à la résolution exacte des problèmes pratiques (en confondant souvent le caractère "pratique" d'un problème avec l'emploi arbitraire des référentiels galiléens ainsi que du vocabulaire et des formes d'imagination associés dans sa formulation). Mais finalement, il n'est pas nécessaire de développer tellement de formules spécifiques comme si c'était une théorie totalement étrangère à nos connaissances courantes, car la partie essentielle de la relativité est une géométrie au fond très similaire à la géométrie euclidienne (comme on l'expliquera plus loin) : il suffit donc, en vue de faire face à la plupart des problèmes, de reprendre les divers outils d'étude de la géométrie euclidienne, portés par l'intuition qui les accompagne, et de préciser comment ils sont modifiés dans la nouvelle géométrie.

Et pourtant, le point de vue que la présentation qui suit vise à exprimer n'est pas réellement nouveau, comme en témoignent maints développements ultérieurs de la physique théorique ayant suivi historiquement la découverte de la relativité. On observe en effet que les lourdes formules ("transformation de Lorentz" et ce qui s'ensuit) prétendument constitutives de la relativité, tout en restant vraies, n'apparaissent plus telles quelles dans les chapitres de relativité générale, électrodynamique classique ou autres théories fondées sur la relativité restreinte. A leur place, intervient un simple objet (une "métrique" dite "pseudo-euclidienne"!) qui témoigne de l'assimilation réelle de cette nature géométrique de l'espace-temps dans l'esprit des physiciens. C'est pour eux si naturel qu'il est inutile se s'attarder sur ce fait, le plus intéressant étant d'aller de l'avant dans son utilisation. Donc, l'approche qui sera proposée ici serait nouvelle non pas en elle-même, mais seulement *relativement à l'univers habituel de l'enseignement universitaire*.

Alors, pourquoi donc une telle rénovation de son enseignement ne s'est-elle jamais produite depuis que tant de physiciens l'ont si bien acquise ? La réponse est simple, mais seulement étrangère à notre monde habitué à juger les connaissances des élèves et étudiants d'après leur capacité à rédiger des démonstrations dans les formes règlementaires : c'est qu'un abîme sépare parfois la claire assimilation personnelle d'une connaissance, de la capacité à trouver les mots justes pour la traduire en une forme verbale claire et transmissible.

### 3. Nouvelle présentation de la relativité restreinte

#### 3.1. Le schéma de la théorie

En vertu de ce qui précède, la meilleure façon de comprendre l'espace-temps est de commencer par renoncer à la tentative de le comprendre en tant que tel, en sorte de pouvoir assimiler tranquillement les notions qui le constituent comme si cela n'avait aucun rapport, pour ne pas dépendre d'arrière-pensées qui risqueraient d'être contre-productives. Le fait que la relativité restreinte a le titre de théorie physique équivaut à dire que le monde imaginaire que nous allons maintenant présenter correspond mathématiquement à l'espace-temps de notre univers dans le domaine d'approximation considéré. Mais ce fait n'a pour le moment aucune importance dans la mesure où il n'apporte rien à la compréhension de la théorie elle-même. Il suffira à la fin, quand tout sera clairement compris, de remarquer simplement : "Et pourquoi pas ?".

Passons donc au contenu conceptuel de la relativité.

Voici:

*Imaginons un monde où toutes choses matérielles sont fixes, en équilibre.*

Pouvez-vous imaginer cela ? Vous allez dire : bien sûr, c'est trop facile, ce genre de situation est un cas particulier de l'expérience de la vie de tous les jours, laquelle est plus compliquée. Mais la relativité restreinte doit être forcément quelque chose de plus compliqué encore, sinon on l'aurait comprise depuis longtemps, non ?

Certes, il n'est pas exactement suffisant d'imaginer un monde immobile, mais mathématiquement la différence est accessoire. L'essentiel de la complexité des choses se trouve contenu dans ce type particulier d'expérience quotidienne.

Cette expérience n'est mathématiquement pas une chose triviale du tout. En effet, pour décrire la disposition des différents objets en équilibre, il faut faire appel au langage de la géométrie. Or la géométrie euclidienne est en fait une théorie mathématiquement complexe, qu'on ne pourrait pas facilement expliquer de manière simple et purement logique sans s'appuyer sur l'expérience visuelle (par exemple à un aveugle qui n'aurait au départ aucune expérience de l'espace, s'il en existe). De même pour la notion d'équilibre. Mais nous avons de la chance, ces deux choses nous sont familières. Il ne reste qu'à y ajouter les trois points suivants, qui peuvent s'introduire chacun indépendamment des autres, ainsi qu'à préciser certaines interactions particulières qu'il y a entre eux.

- Le rapport de cette géométrie à la perception spatio-temporelle que nous en donne l'expérience.

- Le changement de géométrie: l'espace dans lequel les choses sont ainsi disposées obéit à une géométrie différente de celle qui nous est familière; il faut donc la définir.

- L'expression mathématique de la mécanique de l'équilibre et les propriétés mécaniques des particules.

Cette différence de géométrie est faible en soi, la nouvelle géométrie n'étant pas particulièrement plus compliquée que la géométrie euclidienne, mais nous sommes tellement marqués par notre habitude que cette différence peut nous sembler extraordinaire et insurmontable.

Plus précisément, il y a deux différences de géométrie. La première est la dimension: elle vaut 4, alors que nous ne sommes habitués qu'aux dimensions 2 et 3. La deuxième est qu'il ne s'agit pas là de la géométrie euclidienne de dimension 4 qui généralise automatiquement celles du plan et de l'espace usuels. Mais c'est une autre géométrie qui a de fortes parentés avec la géométrie euclidienne, et quelques différences. Suivant un certain usage, nous l'appellerons la géométrie pseudo-euclidienne. Comme la géométrie euclidienne, elle existe en dimension quelconque  $n \geq 2$ .

Il y a environ trois manières possibles de l'introduire. La première serait de partir de principes physiques dont l'invariance de la vitesse de la lumière. Elle masque son analogie profonde avec la géométrie euclidienne, en privilégiant les structures qui lui sont spécifiques, liées au cône de lumière (que nous verrons en dernier).

Une deuxième méthode consiste à en donner une construction mathématique à partir de la géométrie affine, de manière analogue à la construction de la géométrie euclidienne (le moyen le plus commode étant probablement l'introduction d'un produit scalaire sur l'espace vectoriel associé).

Comme ces deux premières méthodes se trouvent facilement par ailleurs, ce que nous présenterons ici sera une troisième méthode (non qu'elle soit meilleure dans l'absolu mais qu'elle a aussi des avantages qui lui sont propres), qui consiste en une sorte de transformation "magique" (non rigoureuse) qui fait la correspondance entre géométrie euclidienne et géométrie pseudo-euclidienne. Son avantage est qu'elle ne nécessite pas de remonter aux fondements mathématiques de la géométrie, mais qu'elle permet de réutiliser les connaissances naïves et les méthodes générales dont on dispose en géométrie euclidienne, en appliquant finalement une simple traduction des résultats pour obtenir ceux correspondants en géométrie pseudo-euclidienne. Elle a donc l'avantage de nous permettre de transporter directement dans la nouvelle géométrie des intuitions très diverses que nous avons pu hériter de l'expérience en rapport avec la géométrie euclidienne sans nécessiter leur reconstruction mathématique, avantage dont nous ne nous priverons pas par la suite.

Cette méthode s'exprime mathématiquement par la notion de *prolongement analytique*: on démontre des formules ou relations géométriques pour les valeurs positives d'une variable, avant d'appliquer finalement le résultat à une valeur négative. Dans cette correspondance, les égalités sont généralement préservées, mais non les inégalités.

Nous nous contenterons de l'introduire en dimension 2 (géométrie du plan pseudo-euclidien), car il est plus facile d'y exercer l'intuition visuelle de la géométrie, et cela suffit à réaliser l'essentiel du travail de présentation des nouveautés par rapport à la géométrie euclidienne. Ensuite, passer à la dimension 3 ou 4 serait un exercice d'intuition difficile mais secondairement utile, car d'un point

de vue abstrait cette généralisation est quasi automatique et ne comporte que peu de phénomènes nouveaux (qu'on abordera par la suite; mais on n'a pas souvent besoin en pratique de tenir compte de toutes les dimensions en même temps).

Les deux autres problèmes, le rapport à l'expérience et la mécanique, s'appuient sur les structures géométriques de l'espace, donc ils dépendent de cette géométrie. Cependant, ils peuvent s'exprimer de la même manière avec n'importe laquelle de toutes ces géométries (euclidienne ou pseudo-euclidienne, de dimension quelconque), grâce au fait que ces géométries ont toutes en commun le même langage fondamental (la même liste de structures; seules les propriétés de certaines de ces structures présentent des différences).

Nous les exprimerons donc en invoquant les structures géométriques suivant le langage de la géométrie euclidienne. Et il sera permis de les interpréter effectivement dans le cadre de la géométrie euclidienne, autrement dit celui de notre expérience quotidienne d'objets fixes, pour bien en comprendre la nature.

Le fait que l'expression de la mécanique relativiste soit la même que celle de l'équilibre s'explique simplement : l'équilibre est la situation de la mécanique relativiste où les systèmes matériels ne subissent aucune évolution dans une certaine direction temporelle; alors, tout peut encore se produire dans les trois dimensions restantes suivant les lois normales de la mécanique relativiste, à la seule différence près que la dimension temporelle (plus précisément la direction temporelle en question) n'intervient plus. La loi de l'équilibre est donc la même que la mécanique relativiste avec une dimension de moins, ce qu'il fallait démontrer.

### 3.2. Lien à l'expérience

La question du rapport de cette réalité à notre expérience de l'espace et du temps, est la seule question facile, qui ne nécessite pas de travail mathématique important; elle n'apporte aucune information sur les lois de la physique elles-mêmes. La question est celle-ci: de quelle manière visitons-nous ce monde de dimension 4 où les choses sont fixes, en sorte qu'il nous apparaisse sous forme d'un monde de dimension 3 où les choses évoluent ?

En général, cette question s'applique à la visite d'un monde de dimension  $n \geq 2$  où les choses sont fixes (et pour lesquelles le temps n'existe donc pas), le laissant percevoir sous la forme d'un monde de dimension  $n - 1$  où les choses évoluent.

Dans cette histoire, le temps au cours duquel les choses paraissent évoluer n'est pas un temps objectif, mais il n'a de sens que subjectif, étant vécu séparément pour chaque observateur, et n'ayant donc pas d'autres relations nécessaires a priori avec celui des autres que celles s'établissant par l'intermédiaire des choses matérielles qui sont fixes.

En fait, nous avons déjà effleuré la réponse en introduisant la géométrie euclidienne : physiquement, chaque observateur (ou son corps ou son vaisseau spatial ou même la Terre si on veut, vu qu'ici la taille de la Terre est finalement négligeable car elle vaut moins d'un dixième de seconde-lumière), est semblable à un fil disposé dans l'espace en une courbe appelée sa *ligne d'univers*. L'idée est que ce fil en tant qu'espace de dimension 1 joue le rôle de ligne du temps personnel de l'observateur. Ainsi, au cours de son temps à lui, il visite le monde en parcourant sa ligne d'univers à vitesse constante, comme un train suivrait une voie ferrée à vitesse constante ou comme une impulsion lumineuse suit une fibre optique. On postule que cette vitesse de parcours est une constante universelle notée  $v$ ; on peut lui donner numériquement la valeur qu'on veut puisque le temps étant ainsi une quantité subjective et non physique (les objets physiques ne peuvent pas le mesurer puisqu'ils sont fixes), on ne peut définir physiquement son unité qu'en termes de la longueur de courbe parcourue à cette vitesse.

Un observateur peut, de sa ligne d'univers, observer la ligne d'univers d'un autre observateur, non l'autre observateur lui-même (car seule la ligne d'univers est matérielle). La sensation d'accélération est définie par la courbure de cette ligne de la même manière que la force centrifuge dans une voiture roulant à vitesse constante est définie par la courbure de la route ou de manière équivalente par la position du volant.

Une horloge est physiquement semblable à un observateur: c'est aussi un fil disposé dans l'espace, mais gradué régulièrement. Donc si la ligne d'univers d'un observateur est superposée

à celle d'une horloge (il garde l'horloge avec lui), il voit au cours de son temps personnel les graduations de l'horloge défilier à un rythme régulier, comme une bonne horloge doit se comporter.

Les paradoxes de la relativité viennent des effets géométriques qui se présentent dans leur complexité mais coupés de l'intuition géométrique explicative.

On peut vérifier que la théorie galiléenne de l'espace-temps s'obtient comme limite de ce modèle lorsque la vitesse de parcours  $v$  des lignes d'univers tend vers l'infini; et par rapport à la théorie galiléenne, le présent modèle apporte des corrections dont le premier terme a une amplitude de l'ordre de  $v^{-2}$ . Or, le travail de changement de géométrie que nous ferons plus loin peut se résumer (si on oublie les expériences faisant intervenir la lumière elle-même, qui ne correspond à rien dans ce modèle euclidien) au fait de donner à ce  $v^{-2}$  une valeur négative, suivant la formule

$$v^2 = -c^2.$$

Donc les effets relativistes (exprimés à l'aide de  $c$ ) sont finalement inversés par rapport à ceux qu'on obtiendrait normalement dans la situation telle que nous venons de la décrire dans le cas de la géométrie euclidienne.

Les effets géométriques suivants, exactement inverses des effets relativistes correspondants, ne seraient-ils pas paradoxaux d'un point de vue purement logique en l'absence de support visuel explicatif :

– Quand on voit au loin un camion avancer de gauche à droite, pourquoi rétrécit-il subitement lorsque son chemin se tourne vers nous ou vers le lointain (dilatation relativiste du temps) ?

– Pourquoi ses deux roues avant qui étaient superposées se séparent-elles par rapport à nous (relativité de la simultanéité) ?

– Pourquoi cela prend-il des temps différents pour aller d'un point à un autre suivant des chemins différents, alors qu'on y va dans tous les cas à la même vitesse (paradoxe des jumeaux) ?

– Enfin, en définissant la grosseur d'un saucisson par la largeur de ses rondelles, comment se fait-il qu'il est plus gros lorsque son axe est incliné par rapport au plan de coupe que lorsqu'il lui est orthogonal (contraction relativiste des longueurs) ?

Voilà, l'essentiel est dit. Ajoutons quelques remarques.

Un *référentiel galiléen* est la donnée d'une direction temporelle (i.e. qu'on peut parcourir comme temps, ce qui en géométrie pseudo-euclidienne ne peut pas être une direction "spatiale") pour être la direction de droite d'univers d'un observateur ou des observateurs choisi(s) comme référence.

**Vocabulaire.** Par commodité pour la suite de cet exposé, nous emploierons le qualificatif de "théorique" comme désignant l'emploi du mode de représentation géométrique de l'espace-temps, où les choses apparaissent fixes (pouvant se comprendre d'abord comme si la géométrie était euclidienne, puis en tenant compte de sa forme pseudo-euclidienne). A cela s'opposera le qualificatif d'"expérimental" pour désigner la description donnée par les apparences expérimentales ordinaires de fixité ou de non fixité dans l'espace à 3 dimensions que peuvent prendre les objets de la mécanique relativiste par rapport à un référentiel donné (la fixité qualifiant les systèmes invariants par translation temporelle suivant le référentiel).

Cet usage quelque peu détourné des termes opposés de "théorie" et "expérience" ne devrait pas prêter à confusion puisque leur usage normal auquel il fait concurrence n'apparaîtra pas dans cet exposé (où tout est théorique au sens usuel).

Par exemple, la notion théorique de point est désignée expérimentalement par le terme d'*événement*, c'est-à-dire un point qui pour un observateur n'apparaît qu'à un instant précis.

**Remarque 1.** A la description ci-dessus il faut ajouter deux informations nécessitant l'usage de notions spécifiques à la géométrie pseudo-euclidienne.

La première information est la condition pour qu'une courbe théorique puisse être le lieu d'une ligne d'univers (d'un observateur ou d'une particule). En effet, en géométrie euclidienne toute courbe pouvait être le lieu de disposition d'un fil (qui permet de définir l'abscisse curviligne), ce qui n'est plus vrai en géométrie pseudo-euclidienne. Notamment, une courbe refermée sur elle-même

n'est pas une ligne d'univers possible, de sorte qu'on ne puisse pas remonter le temps. Précisément, la succession des points d'une ligne d'univers doit respecter l'*ordre temporel* (ou ordre de causalité): c'est une relation d'ordre partiel qui se conserve par le groupe des rotations-translations théoriques. Il est évident qu'une telle relation d'ordre ne peut pas exister dans le cas de la géométrie euclidienne, puisque la symétrie centrale y étant une rotation de  $\pi$  radians la renverserait. Mais la symétrie centrale du plan pseudo-euclidien n'est pas une rotation.

La deuxième chose est de définir ce que l'observateur observe effectivement : l'image observée en un évènement d'observation, est donnée par la lumière, qui arrive "à la vitesse de la lumière", donc issue d'un ensemble d'évènements formant un cône de sommet l'évènement d'observation, appelé le cône de lumière passée. Nous décrirons plus loin ses propriétés géométriques.

**Remarque 2.** Un défaut superficiel de cette vision est qu'elle semble porter une conception fataliste du monde (philosophie d'après laquelle toutes choses à venir seraient écrites avant qu'elles ne se produisent), ainsi qu'une distinction entre matière et esprit. Mais ce n'est là qu'une impression liée à l'usage d'une forme d'intuition particulière choisie pour ses avantages pratiques de compréhension de l'espace physique, et qui n'exprime aucun véritable argument en faveur de l'une ou l'autre de ces philosophies.

### 3.3. Etude déformée de la géométrie euclidienne plane

Nous allons présenter ici la géométrie euclidienne exprimée à travers des modes de représentation déformés. Puis les moyens ici nécessaires pour gérer une telle approche nous permettront "magiquement" de passer à la géométrie pseudo-euclidienne.

Cette déformation s'applique, de manière équivalente, aux deux modes de représentation habituels de la géométrie: les dessins (dans le plan habituel), et les systèmes de coordonnées. Les dessins ne représenteront pas fidèlement les propriétés des figures géométriques que l'on veut étudier, car par rapport à la réalité, elles apparaîtront écrasées dans une certaine direction, comme lorsqu'on les voit de biais, d'un point de vue éloigné. On ne peut alors se fier aux apparences du dessin qu'en ce qui concerne les notions de géométrie affine du plan. Les autres notions doivent être reconstruites, d'une manière non conforme à ce qui se voit sur le dessin.

#### *Les cercles*

Un cercle a sur le dessin l'aspect d'une ellipse, avec un grand axe et un petit axe, qui sont des droites orthogonales (aussi bien en réalité que sur le dessin; leurs directions sont la seule paire de directions ayant cette propriété). La direction du petit axe et le rapport mesuré sur le dessin

$$k = \frac{\text{longueur du petit axe}}{\text{longueur du grand axe}}$$

dépendent du point de vue employé pour faire le dessin, mais ne dépendent pas du cercle.

Remarque: à partir d'une ellipse quelconque dans un dessin, on peut supposer qu'elle représente un cercle. Puis, ayant choisi une ellipse pour jouer le rôle de cercle, les autres cercles se définissent comme étant les ellipses obtenues à partir de celle-ci par homothéties et translations.

#### *Les repères orthogonaux*

Nous allons utiliser ces axes pour former un système de coordonnées, orthonormé sur le dessin (en effet, à l'aide de la géométrie euclidienne du dessin on peut copier l'unité de longueur apparente d'un axe sur l'autre), mais seulement orthogonal en réalité. Les coordonnées seront notées  $(t, x)$ , où l'*axe du temps* (d'équation  $x = 0$ ) est le petit axe de l'ellipse, et l'*axe d'espace* (d'équation  $t = 0$ ) est le grand axe. Notons  $(O, \vec{t}, \vec{x})$  ce repère. Les vecteurs  $\vec{t}$  et  $\vec{x}$  semblent tous deux de norme 1 sur le dessin, mais en fait  $\|\vec{x}\| = k\|\vec{t}\| = 1$ .

Cet emploi purement artificiel du vocabulaire d'espace et de temps pour désigner les coordonnées est seulement motivé ici par le fait qu'un point parcourant l'"axe du temps" à la vitesse  $v$  aura une vitesse apparente égale à 1 sur le dessin, en sorte que la "coordonnée de temps" sur cet axe s'identifie au temps de ce parcours. On aura donc  $v = k^{-1}$ .

Ensuite, à partir d'un tel repère principal, les autres repères qu'on se permettra d'utiliser seront ceux qui sont réellement orthogonaux et ont les mêmes valeurs des normes des vecteurs de base (et

donc la même valeur de  $k$ ) sans tenir compte des apparences du dessin, autrement dit des repères obtenus par de vraies rotations à partir d'un repère orthogonal fixé construit comme ci-dessus.

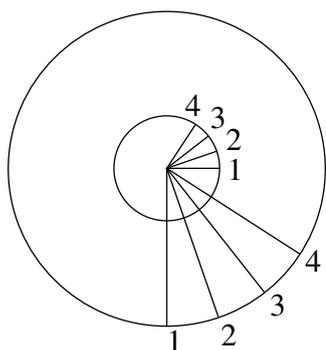
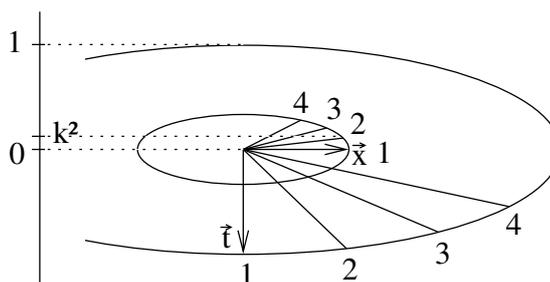


Figure réelle étudiée



Dessin utilisé (agrandi ici)

### Les distances

Il ne sera pas question de mentionner les distances apparentes sur le dessin, car elles n'ont pas de rapport naturel avec la réalité étudiée. Nous allons considérer les mesures des distances réelles, mais suivant deux unités de longueur possibles, qui sont les longueurs réelles des deux vecteurs de base principaux. Précisément, l'unité d'espace vaut  $k$  fois l'unité de temps. Ainsi, la mesure en unité de temps de la norme d'un vecteur  $u$  de coordonnées  $(t, x)$  est donnée par

$$\|u\|_t = \sqrt{t^2 + k^2 x^2} = \sqrt{t^2 + \frac{x^2}{v^2}}$$

et est la longueur apparente sur le dessin du vecteur obtenu par une vraie rotation de  $u$  pour l'amener sur l'axe du temps, et de manière analogue sa mesure en unité d'espace est donnée par

$$\|u\|_x = k^{-1} \|u\|_t = \sqrt{v^2 t^2 + x^2}.$$

À chacune de ces unités correspond de la même manière une mesure du produit scalaire :  $u \cdot u' = tt' + k^2 xx'$  ou  $v^2 tt' + xx'$ .

Ainsi,  $k$  sera traité comme un nombre fixe. Cependant, sans le traiter d'une manière différente, et puisque sa valeur effective n'intervient pas dans les calculs formels, on peut plus judicieusement l'interpréter comme étant non un nombre réel mais une quantité dans le cas d'une étude en coordonnées : les mesures en unités de temps et d'espace sont alors interprétées non comme des nombres réels mais des quantités indépendantes (des temps et des longueurs), et la quantité  $v$  ou son inverse  $k$  est alors la constante universelle qui fait le lien physique naturel entre les deux.

### Les angles

Les angles sont définis par le rapport de la longueur d'un arc de cercle à la mesure du rayon. Or, pour cela on mesurera le rayon en unité de temps, et la longueur de l'arc en unité d'espace. L'unité d'angle utilisée sera donc égale à  $k$  radians. Les fonctions cos et sin utilisant comme il se doit la mesure des angles en radians, la rotation d'un angle  $\alpha$  exprimé dans cette nouvelle unité enverra donc le vecteur temps  $\vec{t} = (1, 0)$  sur le vecteur de composantes  $(\cos(k\alpha), k^{-1} \sin(k\alpha))$ , et le vecteur espace  $\vec{x} = (0, 1)$  sur  $(-k \sin(k\alpha), \cos(k\alpha))$ .

On remarque que tant qu'on ne s'intéresse jamais aux intersections d'un même cercle avec les deux axes d'un repère, ni aux rotations d'un angle droit, le coefficient  $k$  et les fonctions trigonométriques n'apparaissent pas de manière indépendante, mais on peut toujours regrouper leurs apparitions en blocs de  $\cos(k\alpha)$ ,  $k^{-1} \sin(k\alpha)$  et  $k^2$ . (Ou plus généralement, les fonctions de  $k$  qui interviennent sont des fonctions analytiques de  $k^2$ ). Nous excluons dans la suite toute notion géométrique dans laquelle ils n'apparaissent pas ainsi groupés.

### 3.4. Passage à la géométrie pseudo-euclidienne

Pour passer de l'étude déformée de la géométrie euclidienne à celle de la géométrie pseudo-euclidienne (dont toute représentation est nécessairement déformée), il suffit de supposer  $k^2$  négatif. Précisément, la géométrie pseudo-euclidienne, où la correspondance naturelle entre les mesures d'espace et de temps est exprimée par la vitesse de la lumière  $c$ , s'obtient à partir du modèle précédent par la formule

$$v^2 = -c^2 \quad (= k^{-2})$$

d'où on peut déduire que

$$\begin{aligned} \cos(k\alpha) &= \cos\left(\frac{\alpha}{v}\right) = \operatorname{ch}\left(\frac{\alpha}{c}\right) \\ k^{-1} \sin(k\alpha) &= v \sin\left(\frac{\alpha}{v}\right) = c \operatorname{sh}\left(\frac{\alpha}{c}\right) \end{aligned}$$

où les fonctions *cosinus hyperbolique*  $\operatorname{ch}$  et *sinus hyperbolique*  $\operatorname{sh}$  sont définies par

$$\operatorname{ch} u = \frac{e^u + e^{-u}}{2}, \quad \operatorname{sh} u = \frac{e^u - e^{-u}}{2}.$$

(Ces expressions peuvent aussi se déduire des calculs suivant les repères de lumière présentés plus loin).

Les inégalités qui étaient vraies en géométrie euclidienne ne le sont généralement plus ici. En particulier, le produit scalaire continue à exister mais le carré scalaire n'est plus toujours positif.

Contrairement au cas précédent, pour mesurer la norme d'un vecteur donné  $\vec{u}(t, x)$  on ne peut plus choisir arbitrairement entre l'unité de temps et celle d'espace, mais ce choix est déterminé par le *genre* du vecteur, c'est-à-dire par le signe de son carré scalaire, disons ici (celui en unité de temps au carré)  $t^2 - c^{-2}x^2$  : s'il est positif, le vecteur  $\vec{u}$  est du genre temps et  $\|\vec{u}\|$  peut seulement se mesurer en unité de temps; s'il est négatif,  $\vec{u}$  est du genre espace et  $\|\vec{u}\|$  peut seulement se mesurer en unité d'espace; s'il est nul bien que  $\vec{u} \neq 0$ ,  $\vec{u}$  est un vecteur *isotrope* ou vecteur de lumière (sa direction est une direction possible d'un photon dans l'espace-temps) et  $\|\vec{u}\| = 0$ .

En fixant l'origine du plan (ou mieux en regardant l'ensemble des vecteurs, ensemble qui forme un plan), les vecteurs de lumière forment deux droites sécantes qui divisent le plan en quatre parties, dont l'une est constituée des vecteurs isotropes ou du genre temps tels que  $t \geq 0$ , et du vecteur nul : ce sont les *vecteurs futurs*.

Plus généralement, cette notion de définit en dimension plus grande: dans un système de coordonnées, soit le vecteur  $\vec{u} = (x, y, z, t)$ . Son carré scalaire compté en unité de temps au carré vaut

$$\vec{u}^2 = t^2 - c^{-2}(x^2 + y^2 + z^2).$$

Le vecteur  $\vec{u}$  est appelé un vecteur futur si en ce sens  $\vec{u}^2 \geq 0$  et  $t \geq 0$ .

On peut maintenant définir la relation d'*ordre temporel* entre les points du plan pseudo-euclidien: étant donnés deux points  $M$  et  $N$ , on pose que  $M \leq N$  ( $M$  est antérieur à  $N$ ) si et seulement si  $\overrightarrow{MN}$  est un vecteur futur. On peut vérifier que ceci est bien une relation d'ordre.

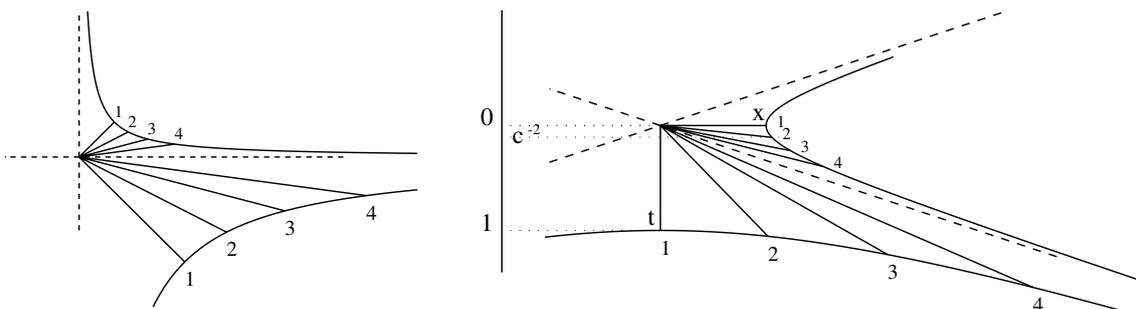
Si  $M \leq N$ , l'intervalle de temps (distance en mesure temporelle) de  $M$  à  $N$  est  $\sqrt{\overrightarrow{MN}^2}$ .

Une courbe est une ligne d'univers possible à condition que la relation d'ordre temporel ci-dessus définie entre les points de la courbe soit un ordre total; cet ordre coïncide alors avec l'ordre du temps vécu par le visiteur qui suit cette ligne. En fait, cette condition est équivalente à demander que la longueur de cette courbe soit une quantité bien définie.

Pour deux points  $M \leq N$ , parmi les lignes d'univers possibles de  $M$  à  $N$ , la ligne droite est la plus longue ligne d'univers possible et la seule dont la longueur en mesure temporelle soit égale à l'intervalle de temps de  $M$  à  $N$ . Les plus courtes lignes sont de longueur nulle et constituées de segments ayant des directions de lumière, mais ce ne sont pas des lignes d'univers possibles pour des observateurs ou tout objet de masse non nulle, qui ne peuvent pas atteindre la vitesse de la lumière: pour la ligne d'univers d'un photon les vecteurs tangents sont de carré scalaire nul tandis que pour celle d'une particule de masse non nulle ou d'un observateur, tous les vecteurs tangents sont strictement de genre temps (de carré scalaire non nul).

Les cercles de cette géométrie sont non plus des ellipses particulières mais des hyperboles particulières, de deux genres possibles (le genre d'un cercle étant celui de ses rayons). Celles d'un même genre sont là encore celles obtenues à partir de l'une d'elles (qu'on peut choisir arbitrairement) par homothéties et translations. Tous les cercles (des deux genres) sont les hyperboles ayant la même paire de directions des asymptotes, qui sont les directions de lumière. Chaque cercle est en deux branches de longueur infinie chacune, puisque la longueur d'un arc de cercle est proportionnelle à l'angle qui se voit aussi comme angle de rotation, et que les rotations peuvent toujours se composer (additionnant les angles) sans jamais boucler un tour.

Deux droites sont orthogonales si et seulement si on peut passer de l'une à l'autre par symétrie par rapport à une droite de lumière et parallèlement à l'autre direction de lumière. Les vecteurs de lumière sont les vecteurs orthogonaux à eux-mêmes.



### Les repères de lumière

Le plan pseudo-euclidien admet un nouvel outil d'étude par rapport à la géométrie euclidienne : l'utilisation d'un repère dont les deux axes sont suivant les directions de lumière. En notant  $(a, b)$  les coordonnées dans un tel repère, la rotation d'un angle  $\alpha$  autour de l'origine s'exprime par

$$\begin{cases} a' = a \exp(\frac{\alpha}{c}) \\ b' = b \exp(-\frac{\alpha}{c}) \end{cases}$$

On peut remarquer le rapport avec l'utilisation des nombres complexes en géométrie euclidienne, où la rotation d'un angle  $\alpha$  s'exprime par la multiplication par  $\exp(i\alpha)$ . En effet, en cohérence avec la définition de la nouvelle unité d'angle que nous avons introduite, abordons cette fois le problème en unité de temps, et en représentant le plan euclidien  $(x, y)$  par des coordonnées d'espace-temps  $(t, x')$  définies par  $(t = x, x' = k^{-1}y)$ . Cette déformation des coordonnées se répercute sur l'écriture des nombres complexes: on remplace l'emploi de  $i$  par celui d'un certain  $i'$  pour pouvoir toujours écrire le "nombre complexe" sous la forme  $z = x + iy = t + i'x'$ , d'où  $i'^2 = -k^2 = c^{-2}$ .

Le carré de  $i'$  étant alors positif, Cette formule admet donc deux solutions réelles  $i' = \pm c^{-1}$  qui correspondent aux deux coordonnées dans un repère de lumière. C'est ainsi le fait de pouvoir regarder les solutions de cette équation (au lieu de la conserver telle quelle lors des calculs) qui est nouveau par rapport à la géométrie euclidienne et permet, en imitant l'écriture en nombres complexes de la géométrie euclidienne, de rendre certains calculs en quelque sorte plus simples ou explicites.

Il est facile de voir que le coefficient  $\exp(\frac{\alpha}{c})$  est la mesure de l'effet Doppler (le nombre par lequel se multiplie la fréquence d'une onde électromagnétique lors de cette rotation).

L'ordre temporel  $M \leq N$  s'exprime dans un repère de lumière par le même ordre sur chacune des coordonnées :  $(a_M \leq a_N \text{ et } b_M \leq b_N)$ .

### Exemple de calcul concret

#### Enoncé

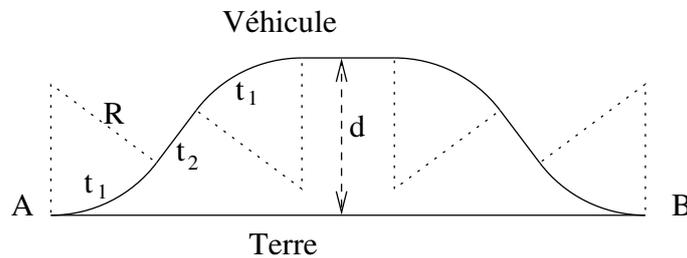
Imaginons un voyageur qui part de la Terre en subissant de son propre point de vue une accélération  $g$  pendant une durée  $t_1$ , puis continue sur sa lancée sans accélération pendant une

durée  $t_2$ , et enfin décélère suivant la même accélération qu'au départ en sens contraire. A quelle distance  $d$  de la Terre se trouve-t-il alors ? Et s'il revient ensuite sur Terre de la même manière, de combien sera la différence  $\Delta t$  entre la durée de son voyage pour lui et pour quelqu'un resté immobile sur la Terre ? (Précisons que pour que ce problème soit effectivement traité par la relativité restreinte sans faire appel à la relativité générale, il faut négliger d'une certaine manière la pesanteur terrestre, plus précisément la vitesse d'échappement du champ de gravitation terrestre).

*Solution*

Traduisons d'abord l'énoncé de ce problème dans le langage de la géométrie euclidienne.

Partant d'une route droite appelée "Terre", un véhicule (allant à une grande vitesse constante  $v = k^{-1}$ ) se met à dévier, le volant étant braqué dans une certaine position qui lui fait sentir une poussée latérale  $g$  pendant une durée  $t_1$ . Puis brusquement le volant est redressé, et reste ainsi pendant une durée  $t_2$ . Enfin, en braquant dans l'autre sens de la même manière on parvient après encore une durée  $t_1$  à redresser le véhicule sur une route droite parallèle à la Terre. A quelle distance de la Terre se trouve-t-on alors ? Si on revient de la même manière, quelle est la différence de durée mise pour y atteindre le point d'arrivée  $B$  à partir du point de départ  $A$ , par rapport à un véhicule qui les aurait reliés en restant sur la route droite "Terre" ?



Pendant la durée de la poussée, le véhicule parcourt un arc de cercle de longueur  $t_1$  en unité de temps ou  $vt_1$  en unité d'espace, d'angle  $\alpha = t_1g$  (ce qui donne en radians  $\alpha k = t_1gv^{-1}$ ) et de rayon (en unité d'espace)  $R = v^2g^{-1}$ .

On en déduit par une observation géométrique élémentaire

$$d = 2R(1 - \cos(k\alpha)) + vt_2 \sin(k\alpha) = 2v^2g^{-1}(1 - \cos(kt_1g)) + vt_2 \sin\left(\frac{t_1g}{v}\right)$$

donc en utilisant la "formule magique"  $v^2 = k^{-2} = -c^2$  on obtient le résultat au problème de relativité

$$d = 2c^2g^{-1}\left(\text{ch}\left(\frac{t_1g}{c}\right) - 1\right) + ct_2 \text{sh}\left(\frac{t_1g}{c}\right).$$

De même on calcule la différence de longueur en mesure temporelle entre la courbe et la ligne droite :

$$\Delta t = 4\left(t_1 - kR \sin(k\alpha)\right) + 2t_2(1 - \cos(k\alpha)) = 4\left(t_1 - cg^{-1} \text{sh}\left(\frac{t_1g}{c}\right)\right) + 2t_2\left(1 - \text{ch}\left(\frac{t_1g}{c}\right)\right)$$

qui est finalement une quantité négative, la ligne droite étant plus longue que la ligne courbe comme on l'avait annoncé.

**Note.** A partir d'ici il est possible de passer directement au chapitre suivant (mécanique classique).

### 3.5. Visions de la relativité de la simultanéité à une dimension

Dans les sections 3.5 et 3.6, nous allons étudier les propriétés de ce qu'on voit expérimentalement, par la lumière qui se propage dans le vide dans l'espace-temps de la relativité restreinte.

L'image du monde qu'un observateur situé en un point reçoit alors en un instant, est donnée par l'ensemble des photons qui lui parviennent. La ligne d'univers d'un photon est une droite de l'espace-temps de direction isotrope (dans tout référentiel galiléen, son vecteur vitesse est constant de norme  $c$ ).

Autrement dit, de son évènement  $O$  d'observation, il voit l'ensemble des points  $M$  tels que  $\overrightarrow{MO}$  est un vecteur futur de carré scalaire nul. Mais, sauf effet de perspective où la largeur de l'observateur lui permet de reconstituer une information sur la profondeur, il ne perçoit pas le point  $M$  lui-même mais seulement la direction de droite ( $OM$ ). On posera ici par convention  $c = 1$ .

Dans cette section 3.5, on se concentrera sur les visions expérimentales possibles de la géométrie du plan pseudo-euclidien, soit une droite au cours du temps.

Un premier cas est celui où l'observateur se trouve sur la droite même où ont lieu les phénomènes observés. On passe des coordonnées  $(t, x)$  aux coordonnées de lumière  $(t + x, t - x)$  dont l'une donne la mesure de la date d'observation par un observateur immobile ( $x_0 = \text{constante}$ ), suivant l'abscisse  $x$  de l'évènement observé (à une constante additive près  $\pm x_0$ ; de toute façon la mesure du temps n'a de sens physique qu'à une constante additive près) : la première  $(t + x - x_0)$  si  $x > x_0$ , la deuxième  $(t - x + x_0)$  si  $x < x_0$ .

Nous avons vu que ce système de coordonnées  $(t + x, t - x)$  présente l'immense avantage de réduire les "rotations" du plan pseudo-euclidien, à la multiplication de chaque coordonnée par une valeur de l'effet Doppler, l'une inverse de l'autre : ces deux axes étant de directions invariantes, une "rotation" s'obtient en étirant l'un et en contractant l'autre.

Que deviennent les vecteurs  $\vec{t}$  et  $\vec{x}$  d'une base orthogonale de ce plan lors d'une telle rotation ? Remarquons qu'il suffit de décrire au premier ordre d'approximation l'effet d'une rotation de petit angle  $\alpha$ , car une rotation d'angle quelconque s'en déduit en répétant une rotation de petit angle un grand nombre de fois. Dans l'étude déformée de la géométrie euclidienne, cette rotation qui fait avancer le vecteur  $\vec{t}$  d'environ  $\alpha\vec{x}$  fait avancer le vecteur  $\vec{x}$  de  $-k^2\alpha\vec{t}$ . Donc en posant  $c = 1$ , cela fait avancer le vecteur spatial  $\vec{x}$  (directeur des droites d'instantanéité), de  $\alpha\vec{t}$ . C'est l'effet nommé "relativité de la simultanéité". Nous savons que cette appellation ne doit pas nous abuser, car cette notion de simultanéité définie par une équation de la forme  $t = \text{constante}$  dans un référentiel, n'est au fond rien qu'un énoncé sur cette coordonnée  $t$  dont nous ne venons de voir qu'un rapport ténu à l'expérience, celui d'être défini par la formule  $t = \frac{1}{2}((t + x) + (t - x))$  à partir des coordonnées de lumière  $t + x$  et  $t - x$  qui correspondent à des mesures de temps de réceptions de photons par des horloges considérées comme fixes. Cette coordonnée de "temps" est bien différente des notions physiques de temps, les temps que mesurent les horloges d'une part, l'ordre de causalité temporel défini par le cône de lumière d'autre part.

C'est de cette construction abstraite, artificielle, de la représentation de la coordonnée  $t$  sous forme du "temps" intuitif, que résulte ce phénomène a priori paradoxal de "relativité de la simultanéité". On peut alors se demander si cette représentation est définitivement inutilisable comme aide à la compréhension des phénomènes. La réponse est non, à condition de fonder ses motivations expérimentales non sur des constructions artificielles mais en termes de ce que l'expérience elle-même nous propose. Alors, la pertinence de l'intuition temporelle (que nous avons jusqu'ici considérée comme nulle pour ne pas en abuser) sera, sans plus, celle qui résidera dans cette expérience.

Nous avons donc besoin d'une situation expérimentale donnant à un observateur de percevoir *naturellement* des évènements se produisant sur un segment d'une droite spatiale  $D$  au cours du temps (en fait on peut étendre automatiquement cette expérience à la perception d'un morceau de plan au cours du temps), suivant des sensations visio-temporelles conformes aux coordonnées orthogonales d'espace-temps  $(t, x)$  que nous avons utilisées dans l'étude géométrique.

Voici cette expérience. Les évènements à observer se produisant en apparence sur un segment de droite  $S \subset D$  et dans certaines limites de temps (donc théoriquement dans une certaine région d'un plan  $P$  pseudo-euclidien), soit un observateur  $O$  éloigné de  $D$  dans une direction autre que celle de  $D$  (par exemple perpendiculaire à  $D$  mais pas nécessairement) à une distance infiniment grande par rapport à la longueur de  $S$ . Autrement dit, posons  $S$  immobile par rapport à  $O$  et infiniment petit par rapport à sa distance à  $O$  (on peut dessiner le tout comme vu point de vue d'un autre observateur immobile par rapport à  $O$ , situé encore infiniment plus loin perpendiculairement au plan contenant  $O$  et  $D$ ).

On définit alors le temps des évènements sur  $S$  sous forme du temps qu'indique l'horloge de  $O$  lors de la perception en  $O$  des photons qu'ils émettent, et leur position par la direction visuelle de ces photons (dont la variation est infiniment petite, autour d'une direction donnée). Ceci forme

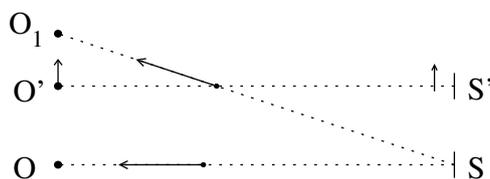
un système de coordonnées affine de  $P$ .

La question est alors : à quelle condition ce système de coordonnées d'espace-temps est-il orthogonal pour la géométrie pseudo-euclidienne, avec des mesures de temps et d'espace conformes à ce que nous avons décrit ? Réponse : il est orthogonal si et seulement si l'observateur  $O$  voit les points de  $S$  comme étant à égale distance de lui, autrement dit, s'il voit le plan  $P$  comme étant à chaque instant une droite  $D$  perpendiculaire à la direction de son regard (la direction allant de  $O$  vers  $S$ ; pour le dire théoriquement, le plan contenant la droite d'univers  $O$  et un rayon lumineux vu est perpendiculaire à  $P$ ). Sa mesure d'espace est alors correcte. Ensuite, dans ce cas, l'unité de mesure de temps sera correcte si et seulement si  $D$  est immobile par rapport à  $O$  (c'est-à-dire que sa distance à  $D$  est fixe au cours du temps, sans regarder les mouvements de coulissement de la droite  $D$  sur elle-même; dit théoriquement,  $O$  est parallèle à  $P$ ).

Remarquons que la position visuelle d'un point de  $D$  peut également se définir par une graduation de  $D$  immobile par rapport à  $O$  (chaque marque est théoriquement une droite de  $P$  parallèle à  $O$ ).

Nous allons maintenant rendre compte de la relativité de la simultanéité quand on passe à un deuxième observateur  $O'$  en mouvement par rapport à  $O$ , dans la direction de  $D$ .

Pour cela, partons de la notion de simultanéité dans  $S$  relative à l'observateur  $O$  par rapport auquel  $S$  est immobile et perpendiculaire à la direction qui le relie à  $O$ . La subtilité se trouve dans la manière suivant laquelle cette condition expérimentale se retrouve avec  $O'$ .



Soit  $S'$  un segment immobile par rapport à  $O'$ , coulissant le long de  $D$ , et rencontrant  $S$  lorsque les évènements s'y produisent. C'est ce segment  $S'$  qui doit être perpendiculaire à sa direction à  $O'$ . Or,  $S'$  est parallèle au mouvement, donc la direction de  $O'$  à  $S'$  est à chaque instant parallèle à celle de  $O$  à  $S$ . (En effet, si on imagine le système  $(O', S')$  comme étant une boîte rigide très allongée, elle est lancée dans la direction orthogonale à sa longueur, donc il n'y a pas de raison qu'une de ses extrémités devance l'autre, et sa direction reste fixe).

Pour que les conditions soient respectées,  $O'$  doit observer des évènements qui se produisent sur  $S'$ . Mais  $S'$  doit recouvrir  $S$  quand ils se produisent; pendant ce temps,  $O'$  rencontre  $O$ . Mais la lumière ne lui en parvient que plus tard, alors que  $O'$  est déjà éloigné de  $O$ . Les rayons lumineux arrivant en  $O'$  de ces évènements, bien que de direction spatiale perpendiculaires à  $D$  par rapport à  $O'$ , ne sont pas perpendiculaires à  $D$  par rapport à  $O$ . Cette perception pour  $O'$  correspond à la perception de tout observateur que  $O'$  rencontre alors, en particulier de  $O_1$  immobile par rapport à  $O$ . La condition d'orthogonalité est fautive pour  $O_1$  : les points en avant de  $S$  sont plus proches de  $O_1$  que ceux qui sont en arrière. Donc, pour des évènements simultanés par rapport à  $O$  qui se produisent sur  $S$ , ceux qui sont en avant de  $S$  sont perçus par  $O'$  avant ceux qui sont en arrière.

Ce raisonnement redonne la relativité de la simultanéité, et la forme de la rotation pseudo-euclidienne en première approximation, que nous avons mentionnées.

### 3.6. Transformation relativiste d'une photographie en relief

*Cas d'un espace-temps à 2+1 dimensions (qui se généraliserait aisément à 3+1 dimensions).*

Nous n'allons pas étudier ici les rapports de toutes les observations possibles mais considérer seulement la question suivante.

Soit l'image en relief (à deux dimensions, une dimension angulaire et une dimension de profondeur) qu'un observateur perçoit du monde en UN instant. C'est donc le cône de lumière passée représenté sous forme projetée sur le plan  $t = 0$  parallèlement à l'axe du temps. Soit un autre observateur en mouvement par rapport au premier mais le rencontrant à l'instant de l'observation. En cet instant, il observe également le monde en relief en recevant les mêmes rayons lumineux que

le premier (son cône de lumière passée est le même). Quelle est alors la loi de transformation qui fait passer de ce que voit l'un à ce que voit l'autre ?

Il s'agit d'une transformation du plan. L'origine  $O$  du plan y joue un rôle privilégié, puisque c'est l'image du sommet du cône de lumière. Voici quelques résultats:

**Théorème 1.** *Une telle transformation envoie toute ellipse de foyer  $O$  en une autre ellipse de foyer  $O$  ayant même mesure du petit axe; sa restriction à toute ellipse de foyer  $O$  coïncide avec une transformation affine.*

**Théorème 2.** *Sa restriction à toute demi-droite issue de  $O$  est linéaire.*

**Théorème 3.** *Quelles que soient deux ellipses de foyer  $O$  et de même mesure du petit axe, toute restriction à l'une d'une transformation affine qui l'envoie sur l'autre se prolonge de façon unique en une transformation possible produite par ce type d'expérience, pour un deuxième observateur de mouvement et de disposition convenables.*

Tout cela se déduit aisément du fait qu'une ellipse de foyer  $O$  est l'image de l'intersection du cône de lumière avec un plan, et que la longueur du petit axe est liée au volume délimité entre le cône et le plan, ou encore au carré scalaire du vecteur  $u = (a, b, c)$  qui définit ce plan comme étant celui d'équation "(produit scalaire par  $u$ ) = 1", soit  $at - bx - cy = 1$ .

Le théorème 2 s'obtient soit directement, soit comme limite du théorème 1, et suffit à construire le prolongement unique évoqué au théorème 3.

On peut ajouter pour la même raison :

**Théorème 4.** *La restriction d'une telle transformation à toute parabole de foyer  $O$ , de même qu'à toute branche d'hyperbole de foyer  $O$ , coïncide avec une transformation affine et  $O$  est toujours foyer de l'image. (La condition sur le petit axe a un équivalent ici mais qui est moins facile à exprimer).*

Attention : les deux branches d'une même hyperbole de foyer  $O$  se transforment en deux branches d'hyperboles différentes (on ne peut retrouver une correspondance qu'à travers une symétrie centrale...). Les droites doivent être rangées parmi les branches d'hyperboles de foyer  $O$  car elles sont échangées avec elles par ces transformations.

*Autre explication :* la preuve théorique étant déjà faite ci-dessus, ajoutons une interprétation en termes plus expérimentaux de ces faits. Comme ces transformations forment un groupe, il suffit de les justifier au premier ordre d'approximation pour des vitesses faibles. Pour cela, voyons le premier ordre d'approximation du résultat précédent à redémontrer, concernant l'effet sur un cercle  $\mathcal{C}$  de centre  $O$  : son image sera une ellipse de foyer  $O$  et de faible excentricité, donc approximativement un cercle de même rayon mais de centre décalé par rapport à  $O$ , et on passe de l'un à l'autre par translation. Pour justifier cela, considérons (en imaginant encore évidemment les choses du point de vue d'un troisième observateur immobile par rapport au premier et situé dans la troisième dimension, loin perpendiculairement au plan de la scène) une figure  $\mathcal{C}'$  définie par les deux propriétés suivantes: d'abord,  $\mathcal{C}'$  coïncidera avec  $\mathcal{C}$  lorsqu'il émettra sa lumière vers  $O$ , donc c'est un cercle par rapport au premier observateur  $O$  (donc à un instant précis pour  $O$ , qui précède l'évènement d'observation d'une durée égale au rayon du cercle); ensuite, il est immobile par rapport au deuxième observateur  $O'$  qui rejoindra le centre  $O$  de  $\mathcal{C}$  lors de l'observation. Donc il n'en est pas le centre, car il n'était pas au centre quand la lumière a été émise. Fin de redémonstration (d'où résulte la forme d'ellipse de  $\mathcal{C}'$  par propriété de groupe, c'est-à-dire déduite de la première approximation en répétant celle-ci, d'amplitude infinitésimale, un grand nombre de fois).

En reprenant les mêmes notations, on remarque ensuite la contraction relativiste des longueurs : cette ellipse  $\mathcal{C}'$  que  $O'$  observe se contracte en un cercle (de diamètre donné par le petit axe) en passant au premier observateur. On parle de contraction des longueurs puisque cet ensemble d'évènements est synchronisé par rapport à  $O$  mais non par rapport à  $O'$  : c'est l'ensemble des évènements d'émissions des photons d'un objet fixe par rapport à  $O'$  donc en mouvement par rapport à  $O$ , qui parviendront au centre à un instant donné, venus à la vitesse de la lumière de distances différentes, interprétés comme une observation abstraite instantanée par rapport à  $O$ .

### *La sphère de vision dans l'espace-temps à 3+1 dimension*

Commençons par motiver les choses géométriquement. Cette fois, en considérant l'image qu'on reçoit par la lumière, nous allons oublier la profondeur et ne retenir que l'image qu'on a des choses sur la sphère de vision. De toute façon, la transformation de la profondeur se déduit de celle de l'angle de vision d'un petit objet, puisque la largeur-image de cet objet (produit de la profondeur par l'angle de vision) se conserve (en effet, il n'y a pas de contraction des longueurs perpendiculairement au mouvement, ou on peut encore le voir en prenant une ellipse presque aplatie, de petit axe la largeur-image de l'objet). On remarque que ce rapport des profondeurs, inverse du rapport des angles de vision, coïncide également avec le rapport des fréquences des ondes reçues suivant l'effet Doppler (lié naturellement au rapport de perception du temps).

La question est alors : quelles sont donc les transformations de cette sphère par changement de référentiel ? La réponse sort en fait immédiatement de la remarque ci-dessus : la variation de la profondeur est liée à celle de la largeur angulaire (infinitésimale). Donc la variation de largeur angulaire ne dépend pas de la direction de cette largeur, mais seulement du lieu où elle se trouve sur la sphère de vision. Ceci montre que c'est une transformation conforme de la sphère ("conforme" signifie: qui préserve les angles: en effet, elle transforme dans la sphère de vision un petit disque en petit disque donc aussi les angles comme secteurs d'un tel disque).

Pour une étude plus approfondie de ces transformations, voir le texte no3.ps de ce site. Ce sont les transformations de Moebius de la sphère (celles qui transforment les cercles en cercles, et qui par la projection stéréographique correspondent aux transformations homographiques du plan complexe).

### **3.7. Introduction à la théorie des spineurs**

(Note: ce paragraphe est beaucoup moins corrigé que les autres; à long terme, il pourra être changé ou déplacé. Il est pour le moment inutile à la suite: on peut passer au chapitre 4).

Comme nous l'avons signalé, les tenseurs sont d'une importance capitale en physique théorique. Pour aborder sérieusement la physique mathématique, il faudrait étudier d'abord en détails la théorie du calcul tensoriel. Mais comme le présent document ne cherche pas à approfondir les notions mais se contente de les survoler superficiellement, un aperçu rapide et superficiel des tenseurs suffira.

A l'avenir peut-être ce chapitre sera complété, mais en fait ce n'est pas nécessaire car toutes les idées de cette approche superficielle ici en vue, suffisantes pour la suite, se trouvent dans le document présent sur ce site : <http://spoirier.lautre.net/tenseurs.dvi> (extraits de messages des newsgroups sur les tenseurs).

La théorie des spineurs est une théorie indispensable en physique des particules dans la définition des spins demi-entiers (équation de Dirac de l'électron par exemple). Nous supposons ici connues les notions usuelles d'algèbre linéaire.

Dans sa forme la plus générale, l'objet de la théorie des spineurs est le suivant: étant donné un espace quadratique  $E$  (espace vectoriel réel ou complexe muni d'une forme bilinéaire  $B$  symétrique non dégénérée), on trouve un espace vectoriel  $S$ , dit espace spinoriel associé à  $E$ , avec une application linéaire  $L$  de  $E$  dans l'espace des endomorphismes de  $S$ , tels qu'on ait l'identité  $L(x)L(y) + L(y)L(x) = -2B(x, y) \text{Id}$  (où  $\text{Id}$  est l'application identique sur  $S$ ).

Nous n'allons pas ici tenter de résoudre ce problème. Mentionnons simplement le fait qu'une solution théorique générale (existence avec une certaine forme d'unicité) est connue, pouvant même se construire relativement facilement, sauf que cela laisse un mystère : à quoi peut donc bien ressembler géométriquement la solution en fin de compte, "une fois effacés les traits de construction" ?

Il se trouve que pour un  $E$  de petite dimension, on peut raisonnablement la comprendre. En particulier, dans le cas qui intéresse la physique, où  $E$  est l'espace-temps de la Relativité restreinte, les structures géométriques de ce problème réalisent en quelque sorte une reconstitution de  $E$  à partir de son cône de lumière, autrement dit de "ce qu'on voit" expérimentalement en un point et à un instant donnés de l'espace-temps. Nous en verrons un aperçu. Malheureusement, nous ne pourrions pas présenter ainsi tous les aspects de la théorie, car la bonne présentation de certains aspects nécessiterait l'usage du calcul tensoriel (à moins bien sûr de simuler ce dernier par d'obscuracrobates).

Ce qui suit s'appuie sur les résultats des sections précédentes sur les expériences visuelles.

### *Spineurs d'un espace-temps à 2+1 dimensions*

Les transformations décrites précédemment de ce qu'on voit admettent une formulation élégante en termes de spineurs. Nous verrons cet outil appliqué d'abord au cas d'un espace à 2 dimensions spatiales plus le temps.

La propriété fondamentale du spin 1/2 est que quand l'espace physique fait un tour complet sur lui-même, ce spin ne fait qu'un demi-tour. Le modèle de base de ce phénomène est celui d'une variable complexe dont le carré s'identifierait à un vecteur d'un plan de l'espace physique. Prenons donc comme espace spinoriel  $S$  le plan "racine carrée complexe" du plan d'observation ci-dessus, ou image réciproque par l'opération de carré complexe. Quelle est donc l'image réciproque par le carré complexe d'une ellipse de foyer  $O$  ? Réponse : c'est une ellipse de centre  $O$  ! Et le type de transformations décrit précédemment est l'image par le carré complexe des transformations spéciales linéaires (linéaires de déterminant 1) de  $S$  (vu comme  $\mathbb{R}$ -espace vectoriel de dimension 2).

La mesure du petit axe de l'ellipse de foyer  $O$  est proportionnelle à l'aire de l'ellipse de centre  $O$  correspondante, dont on sait qu'elle est préservée par les transformations spéciales linéaires.

Ainsi, un vecteur du genre temps, vecteur intérieur au cône de lumière future, et correspondant à un référentiel, donne une ellipse de foyer  $O$  dans le plan de vision : c'est celle qui apparaît sous forme d'un cercle dans ce référentiel. L'ellipse de centre  $O$  correspondante dans le plan spinoriel donne une forme bilinéaire symétrique définie positive (qui est la structure euclidienne qui donnait à ce plan la structure de plan complexe utilisée plus haut).

Plus généralement, un vecteur spatio-temporel  $u = (t, x, y)$  s'identifie au tenseur symétrique dans le plan spinoriel (élément symétrique de  $S \otimes S$ ) de matrice

$$\begin{pmatrix} t+x & y \\ y & t-x \end{pmatrix}$$

De manière équivalente, on peut dire que la forme linéaire associée à  $u$ , produit scalaire par  $u$  d'un vecteur  $v = (t', x', y')$ ,  $u.v = tt' - xx' - yy'$  s'identifie à la forme bilinéaire symétrique dans le plan spinoriel, de matrice

$$A = \begin{pmatrix} t-x & -y \\ -y & t+x \end{pmatrix}$$

(Elle équivaut à la donnée de la solution  $L$  du problème des spineurs évoqué au début, en étant l'application qui à  $(v, w)$  associe  $\det(v, L(u)(w))$ ) (On retrouve ainsi le fait que  $t+x$  et  $t-x$  sont plus fondamentaux que  $t$  et  $x$  !) Une transformation spéciale linéaire de matrice  $P$  envoie cette forme bilinéaire sur celle de matrice  ${}^tPAP$ .

C'est l'expression de la rotation spatio-temporelle (transformation de Lorentz combinée avec des rotations spatiales) la plus générale avec deux dimensions spatiales !

On vérifie :  $\det A = t^2 - x^2 - y^2$  (invariant relativiste, carré scalaire du vecteur  $(t, x, y)$ ).

Donc la forme bilinéaire est dégénérée si elle appartient au cône de lumière; elle est dans le cône de lumière future si elle est positive. L'application du plan spinoriel dans le cône de lumière, qu'on a évoqué précédemment en disant que c'est le carré complexe, est finalement le carré tensoriel. De manière équivalente, à un spineur  $s$  correspond la forme bilinéaire symétrique correspondant au carré tensoriel de  $s$ , définie par  $s^2(v, w) = \det(s, v) \det(s, w) = {}^t_v B w$  où  ${}^t_s = (a, b)$  et

$$B = \begin{pmatrix} b^2 & -ab \\ -ab & a^2 \end{pmatrix}$$

(ou son opposé si on veut un vecteur du cône de lumière passée). On vérifie que  ${}^t_s A s$  est le 2 fois le produit scalaire des vecteurs définis par  $A$  et  $B$  (où  $(t, x, y).(t', x', y') = tt' - xx' - yy'$ )

### *Spineurs de l'espace-temps à 3+1 dimensions*

D'après ce qu'on a décrit, les transformations de la sphère de vision sont les transformations d'un espace projectif de dimension 1 complexe. L'espace vectoriel de dimension 2 complexe auquel il est associé sera notre espace spinoriel  $S$ .

Un vecteur  $v = (t, x, y, z)$  équivalent à la donnée de la forme linéaire (produit scalaire par  $v$ ), de composantes  $(t, -x, -y, -z)$ , se définit alors comme étant l'espace des éléments hermitiens de  $\bar{S} \otimes S$  (produit tensoriel sur le corps  $\mathbb{C}$  de l'espace conjugué de  $S$  par  $S$ ), ou si on parle en termes duaux, celui des formes sesquilinéaires hermitiennes sur  $S$ , de matrice

$$A = \begin{pmatrix} t - x & -y - iz \\ -y + iz & t + x \end{pmatrix}$$

(Contrairement au cas précédent, l'espace  $S$  mentionné ici n'est pas la solution du problème spinoriel dans lequel on définit  $L$ , qui serait plus compliquée) On vérifie :  $\det A = t^2 - x^2 - y^2 - z^2$ . La rotation spatio-temporelle la plus générale est donc de la forme  ${}^t\bar{P}AP$  où  $P$  est dans  $SL(2, \mathbb{C})$ . A quelques acrobaties près, cela redonne dans le cas purement spatial ( $t = 0$ ) la présentation quaternionique des rotations. L'application de l'espace spinoriel dans le cône de lumière s'exprime par :  $\bar{s} \otimes s(\bar{v}, w) = \det(s, v) \det(s, w) = {}^t v B w$  où  ${}^t s = (a, b)$  et

$$B = \begin{pmatrix} \bar{b}b & -\bar{b}a \\ -\bar{a}b & \bar{a}a \end{pmatrix}$$

On remarque que dans ce passage de  $s$  à  $B$  ce n'est pas seulement le signe qui est perdu mais tout l'argument : multiplier  $s$  par un nombre complexe de module 1 ne modifie pas le résultat. On voudrait donc donner un sens à cet argument. Ce n'est pas possible dans l'absolu, puisqu'une homothétie de  $S$  de rapport un nombre complexe de module 1 ne change rien. Mais il est possible de comparer les arguments de plusieurs vecteurs. Pour mieux voir ce qui se passe, fabriquons de toutes pièces une référence : c'est l'argument du déterminant de deux vecteurs dans  $S$  (qui n'intervenait pas ci-dessus puisqu'on avait le produit d'un déterminant et du conjugué d'un autre). Soit donc un spineur  $s$ . On cherche une image visuelle de son argument. Pour cela, prenons un vecteur  $s'$  tel que  $\det(s, s') = 1$  : ceci le définit modulo l'addition avec un multiple complexe de  $s$ . Pour le moment, cela semble ne rien apporter puisque  $s'$  peut avoir n'importe quelle direction dans  $S$ , mais l'information intéressante est la variation de  $s + \lambda s'$  pour un nombre réel  $\lambda$  positif au voisinage de  $\lambda = 0$ . Prenons par exemple  $s = (1, 0)$  et  $s' = (a, 1)$ . On a alors  $B(s + \lambda s') = B(1 + \lambda a, \lambda)$  et en négligeant les termes en  $\lambda^2$  :

$$\begin{pmatrix} 0 & -\lambda \\ -\lambda & 1 + \lambda(a + \bar{a}) \end{pmatrix}$$

soit  $2t = 2x = 1 + \lambda(a + \bar{a})$ ,  $y = \lambda$ ,  $z = 0$ . Comme on veut retenir l'information indépendante de  $a$ , c'est que le vecteur de lumière qui est parti de  $t = x = \frac{1}{2}$  peut s'allonger ou se raccourcir comme il veut mais il avance de manière déterminée dans la direction de  $y$ . Finalement, pour  $s$  fixé, l'application qui à  $s'$  associe  $\det(s, s')$  signifie une application du plan tangent à la sphère que forme le cône de lumière future, en l'image de  $s$ , dans  $\mathbb{C}$ , conservant la structure euclidienne (elle renverse l'orientation si on oriente cette sphère "de l'extérieur", par le vecteur sortant, ou encore si on prend la sphère de vision définie par le cône de lumière passée). Une mesure du module du résultat signifie qu'une flèche (direction et sens) est marquée sur la sphère de vision, en l'image de  $s$ . Lorsque  $s$  tourne d'un angle  $h$ , donc est multiplié par un complexe  $u = e^{ih}$ , cette flèche tourne de  $2h$ . en effet,  $s'$  doit tourner de  $-h$  pour que  $\det(s, s')$  reste égal à un. Puis, le nouveau  $s + \lambda s'$  vaut  $us + \lambda us' = u(s + \lambda u^2 s')$  dont l'image physique a bien tourné de  $2h$ . Cela manifeste encore une fois le phénomène de spin demi-entier : quand on le tourne d'un angle, son image dans l'espace physique tourne d'un angle double.

Pour finir, donnons la représentation spinorielle du champ électromagnétique en un point. D'abord, quelle est donc la forme d'un tel champ en un point ? C'est une forme bilinéaire anti-symétrique de l'espace vectoriel associé (tangent) à l'espace-temps, de matrice (où les coordonnées sont écrites dans l'ordre  $(t, x, y, z)$  :

$$\begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix}$$

Comme n'importe quelle forme bilinéaire antisymétrique (comme par exemple le moment cinétique), elle se représente dans l'espace spinoriel par une forme bilinéaire symétrique : soit  $M = E + iB$  (donc de composantes  $M_x = E_x + iB_x$ , etc), vecteur de l'espace quadratique  $\mathbb{C}^3$  de base orthonormée, qui s'incarne par la forme bilinéaire symétrique de matrice

$$\begin{pmatrix} -My + iMz & Mx \\ Mx & My + iMz \end{pmatrix}$$

La transformation du champ électromagnétique par changement de référentiel (ou rotation spatio-temporelle) en résulte, par l'action naturelle des transformations spécial linéaires de l'espace spinoriel sur l'espace de ses formes bilinéaires ( $A \mapsto {}^tPAP$ ). On retrouve l'invariant relativiste du champ par le carré scalaire  $M^2 = E^2 - B^2 + 2i(E.B)$  aussi égal à l'opposé du déterminant de la matrice.

## 4. Mécanique classique

### 4.1. Introduction à la géométrie différentielle

Pour aborder la mécanique nous aurons besoin de notions de géométrie différentielle. Nous avons introduit la géométrie affine qui est une géométrie plus souple que la géométrie euclidienne (avec plus de transformations et moins de structures). De même la géométrie différentielle est plus souple que la géométrie affine. Il n'y a plus de notion de droite, mais il reste la notion de courbe.

Nous allons d'abord introduire les principales notions de géométrie différentielle dans le cadre de la géométrie affine, puis nous nous affranchirons de ce cadre pour retenir ces notions en elles-mêmes. Nous ne chercherons pas à donner de cela une définition rigoureuse axiomatiquement mais l'accent sera mis sur les idées et notions de base qui sont utilisées en pratique.

On connaît la notion de dérivée d'une application de l'ensemble des nombres réels dans lui-même, ou plus généralement, la dérivée d'une application d'une droite affine dans une autre; la valeur d'une telle dérivée en un point est homogène au rapport d'une quantité associée à la deuxième droite (appartenant à la droite vectorielle qui lui correspond, ensemble des différences de deux éléments) sur une quantité associée à la première. Pour fixer les idées, disons qu'on travaillera avec les applications dérivables et dont la dérivée est continue, mais d'autres choix seraient possibles (de préférence une condition de régularité plus forte). Pour simplifier, les applications appartenant à la classe de régularité considérée (celle-là ou une autre) seront qualifiées de *lisses*.

Généralisons cela aux deux cas suivants.

– Le cas d'une courbe paramétrée, application lisse d'une droite affine  $D$  dans un espace affine  $\mathcal{E}$ . Sa différentielle est une application de  $D$  dans l'espace des vecteurs de  $\mathcal{E}$  divisés par les quantités associées à  $D$ . Si on se représente  $D$  comme étant la droite du temps, cette courbe paramétrée se voit comme étant le mouvement d'un point, et sa différentielle est la donnée du vecteur vitesse de ce point en chaque instant.

– Le cas d'un champ scalaire, application lisse d'un espace  $\mathcal{E}$  dans une droite  $D'$ . (Le nom de "scalaire" s'emploie aussi comme synonyme de "quantité", désignant un élément d'une droite vectorielle ou affine, traité comme un nombre, par opposition au cas d'éléments d'espaces vectoriels de dimension plus grande.) Sa différentielle en un point est une forme linéaire sur l'espace des vecteurs de  $\mathcal{E}$  multipliée par une quantité de  $D'$ . Et sa différentielle (tout court) est l'application qui à tout point de  $\mathcal{E}$  associe la différentielle en ce point.

En rassemblant ces deux notions, on peut dire que quand un point se déplace dans un espace muni d'un champ scalaire, la dérivée par rapport au temps de la valeur du champ rencontré à chaque instant est le résultat de la différentielle du champ en la position courante du point, appliquée à son vecteur vitesse. Par abus de convention on sous-entendra parfois dans la notion d'appartenance d'un vecteur à un espace vectoriel, le fait que ce peut être en réalité le produit d'un élément de cet espace vectoriel par une quantité.

Introduisons maintenant les espaces de la géométrie différentielle, qu'on appelle des *variétés*. Intuitivement, une variété de dimension  $n$  est un espace qui est approximativement un espace affine de dimension  $n$  au voisinage infinitésimal de chacun de ses points. Par exemple, on a les sous-variétés d'un espace affine (c'est-à-dire des parties de cet espace qui sont des variétés): les surfaces dans l'espace usuel sont les sous-variétés de dimension 2. Dans un espace affine  $\mathcal{E}$  de dimension  $n$  muni d'un champ scalaire  $f$ , l'ensemble  $S$  d'équation  $f = 0$  est une sous-variété de dimension  $n - 1$  à condition que la différentielle de  $f$  ne soit en aucun point de  $S$  égale à la forme linéaire nulle sur l'espace des vecteurs de  $\mathcal{E}$ . Un théorème dit que toute variété de dimension  $n$  est représentable comme sous-variété d'un espace affine de dimension  $2n + 1$  (ou toute dimension plus grande mais pas toujours moindre).

A chaque point  $x$  d'une variété  $S$  de dimension  $n$  est associé de manière continue un espace vectoriel  $T_x S$  de dimension  $n$  appelé l'*espace tangent* de la variété en ce point. Dans le cas d'une sous-variété d'un espace affine, on imagine facilement l'espace tangent en chaque point comme étant l'espace tangent en ce point qu'on connaît comme sous-espace affine de l'espace de départ, avec le point de tangence pris comme origine pour en faire un espace vectoriel. C'est d'ailleurs ainsi qu'on peut définir la notion de sous-variété de dimension  $n$  d'un espace affine: une sous-variété est une

partie qui en chaque point admet un espace tangent, c'est-à-dire un sous-espace affine approximant la partie près du point.

Mais dans le cas général d'une variété, les espaces tangents se définissent abstraitement, sans être vus comme inclus dans un même espace affine. C'est ainsi une notion indépendante de la manière dont on représente la variété comme sous-variété d'un espace affine.

Les notions de différentielle qu'on a vues pour un espace affine se généralisent au cas d'une variété de la manière suivante, en s'aidant de représentations comme sous-variétés d'un espace affine. Finalement, dans le cas contraire d'une variété abstraite, ces opérations tiennent lieu de définition des espaces tangents.

Une courbe paramétrée d'une sous-variété  $S$  de  $\mathcal{E}$  est une courbe paramétrée de  $\mathcal{E}$  qui est incluse dans  $S$ , autrement dit qui est une application  $c$  de  $D$  dans  $S$ . Sa différentielle en tout point  $t \in D$  (sa vitesse à tout instant) est un vecteur appartenant à  $T_{c(t)}S$  (divisé par une quantité de temps). Un champ scalaire sur  $S$  est la restriction à  $S$  d'un champ scalaire de  $\mathcal{E}$ . Sa différentielle en chaque  $x \in S$  est une forme linéaire sur  $T_x S$  (restriction à l'espace tangent de la différentielle du champ de  $\mathcal{E}$  auquel il correspond).

Par exemple, en choisissant un repère de  $\mathcal{E}$ , chaque coordonnée étant un champ scalaire sur  $\mathcal{E}$  donne un champ scalaire sur  $S$ . La donnée de ces champs scalaires sur  $S$  équivaut à la donnée de cette représentation de  $S$  comme sous-variété de  $\mathcal{E}$ . Une manière facile de construire d'autres représentations de  $S$  comme sous-variété d'espaces affines consiste à prendre une représentation donnée, donc une liste de champs scalaires modulo un choix de repère, et à ajouter à cette liste un ou plusieurs autres champs. Ou autrement dit, considérer le graphe de cet autre champ scalaire sur une sous-variété comme une autre représentation de cette variété. Ceci donne une représentation dans un espace affine ayant une dimension de plus que celle de départ.

## 4.2. Notion d'équilibre

Introduisons une grandeur physique (autrement dit une droite vectorielle) appelée *énergie* (ses éléments sont les quantités d'énergie).

### *Equilibre d'un point*

Considérons un point matériel libre de se déplacer dans une variété  $\mathcal{E}$  et soumis à des forces dues à des causes extérieures fixes. Par exemple,  $\mathcal{E}$  peut être une surface sur laquelle une bille est libre de rouler, soumise à la pesanteur et au champ magnétique d'un aimant fixe caché sous la surface. Ici, on ne s'intéressera pas au mouvement de rotation de la bille sur elle-même, qui pourrait aussi bien glisser, mais seulement au déplacement de son centre, réduisant l'idée de la bille à un point. Et plus précisément, on étudiera la condition pour qu'elle soit à l'équilibre.

Chaque force s'exprime sous la forme d'un champ scalaire  $V$ , appelé le *potentiel* et à valeurs dans la droite des énergies (ou une droite affine correspondant, ceci pour exprimer que "le potentiel n'est défini qu'à une constante additive près").

L'effet global d'un système de forces données par des potentiels est identique à celui d'une seule force dont le potentiel est égal à la somme des potentiels des différentes forces.

Il y a deux notions d'équilibre qu'on peut considérer.

**Equilibre stable.** *Dans une variété  $\mathcal{E}$  munie d'une force de potentiel  $V$ , un point d'équilibre stable est un minimum local de  $V$ .*

En effet, une particule en mouvement ne peut pas passer par des lieux où l'énergie potentielle est supérieure à l'énergie qu'elle possédait au départ. Donc en partant du voisinage d'un minimum de potentiel sans autre énergie elle ne peut pas en sortir.

On remarque qu'à moins de se trouver sur un bord de la variété, en un minimum local la différentielle s'annule. Plus généralement:

**Equilibre.** *Un point d'équilibre de  $\mathcal{E}$  muni d'une force de potentiel  $V$  est un point où la différentielle de  $V$  est nulle. Un point d'équilibre instable est un point d'équilibre qui n'est pas stable.*

La justification nécessiterait ici que le potentiel soit différentiable deux fois, car pour que la variation soit trop faible pour fournir à la particule qui commencerait infinimentesimement à s'écarter de cette position l'énergie nécessaire pour achever de s'éloigner nettement en un temps limité, il

est quasiment nécessaire (et suffisant) que cette variation ne dépasse pas en ordre de grandeur le carré de la distance.

Par exemple, pour un point sur un relief soumis à la pesanteur, le fond d'un creux est une position d'équilibre stable, tandis qu'un sommet et un col sont des positions d'équilibre instable.

Dans le cas où plusieurs forces s'exercent, la condition d'équilibre en un point est l'annulation de la différentielle de la somme des potentiels, qui est égale à la somme des différentielles des potentiels au même point.

**Force.** *Dans une variété  $\mathcal{E}$  où réside une force de potentiel  $V$ , la force qui s'exerce sur une particule située en un point donné de  $\mathcal{E}$  se définit comme étant moins la différentielle de  $V$  en ce point. La condition d'équilibre d'une particule en un point soumise à plusieurs forces est que la somme de ces forces soit nulle.*

### *Equilibre d'un système*

On imagine un système comme étant un assemblage de matériaux déformables pouvant être soumis à des forces extérieures, qu'on peut imaginer comme étant des champs ou l'effet de contacts avec d'autres matériaux déformables liés à des supports fixes. (En fait, il se trouve qu'en physique fondamentale, les champs se comprennent comme des déformations de sortes de milieux élastiques dans l'espace-temps qui interagissent par contact avec la matière).

**Espace de configuration.** *On appelle espace de configuration d'un système l'ensemble des manières dont il pourrait être disposé dans l'espace (appelées états), sans tenir compte de la question de son équilibre.*

C'est souvent une variété de dimension infinie. Chaque état serait en principe réalisable, soit en tant que cliché à un instant donné d'un système en mouvement, soit en ajoutant au système d'autres forces extérieures ajustées en sorte d'annuler exactement les forces présentes.

Cette notion permet de ramener la notion d'équilibre d'un système à celle de l'équilibre d'un point que nous avons introduite, à savoir un point qui se déplace dans l'espace de configuration. Les forces s'expriment par des potentiels définis sur l'espace de configuration, à savoir la mesure de l'énergie présente pour chaque état du système, et se répartissent en deux espèces: les forces internes d'une part, les forces extérieures d'autre part (dont l'énergie réside à l'extérieur du système).

### **4.3. Bilans des forces extérieures**

Nous allons voir maintenant comment on peut faire un bilan des forces extérieures qui exprime une condition nécessaire d'équilibre en oubliant les forces internes. L'espace tangent en un point de l'espace de configuration représente l'ensemble des petites modifications ou vitesses d'évolution, de l'état du système à partir d'un état donné. Les forces sont des formes linéaires sur cet espace, et la condition d'équilibre est que leur somme soit nulle. Une condition nécessaire d'équilibre est que la somme de leurs restrictions à un sous-espace donné soit nulle. Les forces intérieures peuvent alors être oubliées si on sait que leurs restrictions à ce sous-espace sont toujours nulles.

Or, on connaît des types d'évolutions d'un système suivant lesquels les forces internes ne peuvent pas agir, c'est-à-dire que leur potentiel y demeure constant: ce sont les mouvements qui ne déforment pas le système (et n'affectent pas son état interne) parce qu'ils sont définis par un déplacement géométrique de l'espace physique (translation ou rotation), comme celui d'un solide indéformable. Formalisons cela.

Soit  $G$  le groupe des déplacements de l'espace physique, ou encore une de ses sous-variétés contenant l'élément neutre  $\text{Id}$  (l'identité, le déplacement qui ne déplace rien). C'est une variété. Fixons un état  $x$  de l'espace de configuration  $\mathcal{E}$ . Soit alors l'application différentiable de  $G$  dans  $\mathcal{E}$  qui à tout déplacement associe l'état obtenu en appliquant ce déplacement à  $x$ . En la composant avec un champ de potentiel  $V$  sur  $\mathcal{E}$ , on obtient un champ de potentiel  $V_G$  défini sur  $G$ . La différentielle en  $\text{Id}$  de cette application de  $G$  dans  $\mathcal{E}$  est une application linéaire de  $L = T_{\text{Id}}G$  vers  $T_x\mathcal{E}$ . On dit que  $L$  est l'algèbre de Lie de  $G$  si  $G$  est un groupe. La composée de cette application linéaire par la force (comme forme linéaire en  $x$ ) coïncide avec la force de potentiel  $V_G$  en  $\text{Id}$ .

Ainsi, on pose alors comme condition nécessaire d'équilibre, l'équilibre de  $\text{Id}$  dans  $G$  soumis aux forces de potentiels  $V_G$ .

L'argument plus haut se traduit par le fait que le potentiel sur  $G$  correspondant à toute force interne est une constante, et donc sa différentielle en  $\text{Id}$  est nulle.

**Torseur d'une force.** *Le torseur d'une force extérieure appliquée à un système dans un état donné, est la forme linéaire sur l'algèbre de Lie du groupe  $G$  des déplacements géométriques de l'espace physique, construite comme force de potentiel  $V_G$  en  $\text{Id}$  (opposée de la différentielle de  $V_G$ ) défini ci-dessus.*

Ainsi, la condition nécessaire d'équilibre du bilan des forces extérieures est que la somme des torseurs de ces forces soit nulle.

On peut généraliser cette notion au cas du torseur d'un lien rigide, qui consiste à interdire tout mouvement à une partie du système (un bord, un point, un axe...) qu'on appellera le support de ce lien. Le problème est que les déplacements qui permettraient de calculer le torseur sont ici interdits, et que cette force ne s'exprime pas par un potentiel. Alors, on peut quand même la calculer d'après les forces internes, à deux conditions. D'abord, que le système soit à l'équilibre. Ensuite, qu'au voisinage du support il n'y ait pas d'autres liens rigides qui contraignent le système, et que les autres forces extérieures qui s'y exercent soient négligeables. En effet, l'équilibre implique que la force du lien est équilibrée par les forces internes voisines, à condition qu'elle ne soit pas directement compensée par d'autres forces extérieures qui ne passeraient pas par l'intérieur du système.

Voyons donc comment reconstituer le torseur de la force du lien rigide sur le système, en supposant connue l'énergie interne du système en toute circonstance, et appliquons cette connaissance au cas où le support du lien subirait un déplacement, contrairement à ce qui était supposé au départ. Remarquons que cette connaissance est en fait comprise dans celle des effets de mouvements conformes à l'hypothèse, puisqu'il s'agit de l'énergie interne qui est invariante par déplacements, et donc l'effet d'un déplacement du support seul avec son voisinage équivaut au déplacement inverse du reste du système par rapport à ce support.

Regardons donc la variation d'énergie interne du système lorsque seul le voisinage immédiat du support suit ce mouvement, réalisant une transition lisse avec les parties plus éloignées (celles soumises aux autres forces extérieures) qui ne suivent pas. Le torseur force du lien rigide sur le système, donc, se déduit des variations d'énergie interne du voisinage résultant d'un déplacement du support, par le fait que le support doit se trouver dans sa position d'équilibre entre la force extérieure donnée par le lien et la force intérieure du voisinage du lien dans le système: la somme de ces deux torseurs de forces étant nulle (appliquées au support), elles doivent être opposées.

Remarquons qu'apparaît ainsi le principe d'action-réaction entre deux objets distants interagissant par l'intermédiaire d'un système (pouvant être un milieu élastique ou un champ), ces objets étant vus comme deux parties du système: cette interaction s'exprime par la même énergie interne qui dépend des mouvements de ces deux objets l'un par rapport à l'autre. Le torseur force que  $A$  exerce sur  $B$  se lit à la variation d'énergie quand  $B$  se déplace par rapport à  $A$ , et celui de la force que  $B$  exerce sur  $A$  se lit à celle quand  $A$  se déplace par rapport à  $B$ . La somme des deux est nulle, l'énergie interne ne variant pas quand le même déplacement s'exerce sur le tout.

On a aussi:

**Théorème.** *Dans un système à l'équilibre, Le torseur des forces traversant une hypersurface  $S$  (surface de dimension  $n - 1$  dans un espace de dimension  $n$ ) ne dépend que du bord de  $S$ .*

En effet, soient deux surfaces disjointes ayant le même bord. Vérifions que le torseur des forces les traversant sont égaux (le cas de deux surfaces non disjointes s'en déduit en prenant une succession de surfaces de même bord où deux surfaces successives sont disjointes). Ces deux surfaces enferment un volume, et on peut alors faire le bilan des forces entrant dans ce volume. Le torseur des forces entrant dans ce volume vaut celui des forces entrant par la première surface, moins celui des forces sortant par la deuxième. Le résultat vaut zéro puisqu'il y a équilibre, donc les forces traversant les deux surfaces sont égales. En toute rigueur, pour que ces notions aient un sens il faudrait que le voisinage de ce bord soit vide afin de pouvoir considérer la déformation de la partie intérieure quand on déplace les deux surfaces l'une par rapport à l'autre, mais on admettra que les résultats sont bien définis à la limite quand on fait tendre vers zéro l'épaisseur de ce vide.

**Remarque.** La mesure de cette force traversant une hypersurface, qui se déduit des variations d'énergie interne de la partie du système contenu dans un volume dont cette hypersurface est une partie du bord, quand on la déplace par rapport au reste du bord, ne se retrouve pas inscrite dans le potentiel sur l'espace de configuration du plus grand système non bordé par elle. En effet, pour l'y retrouver il faudrait déplacer les deux surfaces l'une par rapport à l'autre du point de vue de l'espace du volume intérieur (pour déformer ce volume) sans affecter leurs dispositions du point de vue de l'espace extérieur, ce qui n'est possible qu'en théorie de la relativité générale. Ce qui explique d'ailleurs pourquoi c'est la relativité générale qui exprime un champ (le champ de gravité) comme naturellement induit par la disposition de la somme des forces matérielles elles-mêmes.

Abordons maintenant les détails de la forme du torseur force. Comme le groupe des rotations est compliqué, on décompose en pratique le torseur d'une force en les deux notions suivantes, restrictions de cette forme linéaire à deux sous-espaces vectoriels de  $L$  qui sont supplémentaires:

**Résultante d'une force.** *La résultante d'une force extérieure appliquée à un système est la forme linéaire sur l'espace des vecteurs de l'espace physique construite de la même manière par l'identification naturelle de cet espace vectoriel à l'algèbre de Lie du groupe des translations (espace des vitesses de translation).*

Pour le dire intuitivement, ceci assemble par translation dans le même espace vectoriel des vecteurs de l'espace physique, les forces exercées en tous points du système pour en faire le bilan.

**Moment d'une force par rapport à un point.** *Même construction appliquée au groupe des rotations autour de ce point (les déplacements qui laissent fixe ce point).*

En effet, une vitesse de déplacement solide quelconque se décompose de manière unique comme somme de la vitesse de translation égale à la vitesse d'un point par ce déplacement, et de la vitesse de rotation autour de ce point définie par la partie linéaire de ce mouvement.

Ces constructions semblent faire uniquement apparaître les forces comme des formes linéaires, à l'encontre de l'usage courant de représenter les forces par des vecteurs. Bien sûr, les notions de vecteur et de forme linéaire se correspondent rigoureusement par le produit scalaire. Ce qui les distingue c'est intuitivement la direction qu'ils désignent, orthogonale l'une à l'autre: dans l'espace euclidien de dimension 3, un vecteur désigne une direction de droite tandis que la forme linéaire correspondante désigne celle de son plan orthogonal. En mécanique classique, le vecteur force apparaît sous la forme expérimentale de l'accélération d'une particule soumise à cette force.

La question qui se pose alors est: la direction d'un vecteur force joue-t-elle géométriquement un rôle dans la notion d'équilibre ? La réponse est oui: il se trouve que dans la forme géométrique des torseurs par rapport à l'espace physique, la direction importante est celle du vecteur de la résultante et non de la forme linéaire. Cela ne doit pas nous étonner, puisque la notion de torseur est fondée sur le groupe des déplacements, qui représente la géométrie euclidienne de l'espace.

Nous allons maintenant étudier en particulier la géométrie des torseurs et leur équilibre dans le cas du plan, qui est la plus petite dimension où cette notion apparaît. Nous ne l'étudierons pas en dimension 3 comme voudrait l'usage, puisque ces efforts supplémentaires dans cet autre cas particulier plus compliqué cacheraient le fait que c'est finalement une notion générale qui serait traitable plus simplement en tant que telle à condition d'introduire au préalable le calcul tensoriel, ce que nous ne ferons pas dans l'immédiat.

#### 4.4. Géométrie des forces dans le plan

Décrivons d'abord la forme des torseurs dans le plan.

Un torseur peut se représenter sous forme du champ des moments, application qui à tout point du plan associe le moment du torseur par rapport à ce point. C'est-à-dire qu'on restreint le torseur en tant que forme linéaire sur l'algèbre de Lie du groupe des déplacements, à l'algèbre de Lie du groupe des rotations autour de chaque point. Ce faisant, on ne perd aucune information, et ceci est valable en général au-delà du plan en toute dimension, car les vitesses de déplacements quelconques s'obtiennent en ajoutant des vitesses de rotations: d'abord comme on a dit, une rotation avec une translation, puis une translation s'obtient en ajoutant deux vitesses de rotations de parties linéaires opposées mais de centres différents.

Dans le plan, le moment en un point est simplement un scalaire, parce que le groupe des rotations autour d'un point est de dimension 1. Ces rotations se mesurent par leur vitesse angulaire, à condition de choisir une orientation du plan (un sens de rotation). Donc, le champ des moments d'un torseur est l'application qui à tout point associe la puissance fournie par la force lors de la rotation de centre ce point à la vitesse angulaire d'un radian par unité de temps (le résultat est homogène à une énergie).

Quelle est alors la forme du champ des moments d'un torseur ?

**Torseurs du plan.** *Une orientation du plan étant choisie, l'algèbre de Lie des déplacements du plan s'identifie à l'espace des points pondérés du plan; le torseur d'une force dans le plan se représente par une forme affine (application affine du plan dans une droite vectorielle, ici l'ensemble des quantités d'énergie), dont la partie linéaire est la résultante tournée d'un angle droit dans le sens direct.*

Ces deux énoncés sont équivalents, puisque l'espace des formes affines et celui des points pondérés sont le dual l'un de l'autre.

On pourrait démontrer cela abstraitement, mais pour donner un aperçu plus concret de la situation nous allons la regarder sur l'exemple fondamental d'une force exercée en un point. Soit donc une force donnée par une forme linéaire  $f$  de l'espace physique, n'étant sensible qu'à la vitesse d'un point matériel  $M$  donné dans le système. Sa résultante est égale à  $f$ . Calculons le moment de cette force par rapport à tout point  $O$ . Soit  $r$  la rotation vectorielle d'un angle droit dans le sens direct suivant l'orientation choisie du plan. La rotation d'un radian par unité de temps autour de  $O$  dans le sens direct entraîne le point  $M$  suivant la vitesse  $r(\overrightarrow{OM})$ , donc la puissance fournie par  $f$  au système, qui définit son moment par rapport à  $O$ , vaut  $f(r(\overrightarrow{OM}))$ , soit en appliquant encore  $r$  sur le tout,  $r(f)(\overrightarrow{MO})$ . On trouve ainsi le résultat: le champ des moments est l'application qui à tout  $O$  associe  $r(f)(\overrightarrow{MO})$ . C'est donc la forme affine s'annulant en  $M$  et de partie linéaire  $r(f)$ , qui est la résultante  $f$  tournée d'un angle droit.

D'autre part il est naturel que l'ensemble des vitesses de rotations s'identifie à l'espace des points pondérés du plan, où la position d'un point pondéré correspond au centre de la rotation et sa masse correspond à la vitesse angulaire.

Nous allons compléter cette description par celle de la transmission des forces à travers les systèmes, sachant que les forces se transmettent uniquement de proche en proche, et qu'en tout point le torseur de la force qui est transmise par là s'annule en ce point. Voici:

**Théorème.** *Toute disposition des forces en équilibre dans le plan peut se décrire au moyen d'un champ scalaire  $f$  (défini à l'addition près par une forme affine) d'après la règle suivante: quels que soient deux points  $M$  et  $N$  du plan, le torseur de la force traversant de droite à gauche toute courbe reliant  $M$  à  $N$  se représente par la différence  $T(N) - T(M)$  des formes affines dont les graphes sont les plans tangents au graphe de  $f$  aux points  $M$  et  $N$ .*

Ce théorème n'a pas d'équivalent pour des forces en dimension plus grande.

On a déjà vu à la section précédente qu'à l'équilibre, le torseur d'une force traversant une courbe (cas particulier d'hypersurface) ne dépend que des extrémités de la courbe.

Fixons une origine  $O$  dans le plan. Pour tout point  $M$ , soit  $T(M)$  le torseur de la force traversant de droite à gauche toute courbe reliant  $O$  à  $M$ . Alors il est immédiat que pour tous points  $M$  et  $N$ , le torseur de la force traversant toute courbe reliant  $M$  à  $N$  vaut  $T(N) - T(M)$ . Il suffit en effet de prolonger une courbe de  $O$  à  $M$  par une courbe de  $M$  à  $N$  pour obtenir une courbe de  $O$  à  $N$ .

Posons maintenant pour tout  $M$ ,  $f(M) = T(M)(M)$  (le moment en  $M$  de  $T(M)$ ). Il ne reste qu'à vérifier que le graphe de  $T(M)$  est le plan tangent en  $M$  au graphe de  $f$ .

Calculons la valeur de  $f$  en un point  $N$  voisin de  $M$ :

$$f(N) = (T(N) - T(M))(N) + T(M)(N).$$

Or  $T(N) - T(M)$  est le torseur des forces traversant une courbe reliant  $M$  à  $N$ . Non seulement ce torseur est petit correspondant à la proximité des deux points, mais la droite où il s'annule passe

par le voisinage commun de ces points, d'où le fait que sa valeur en  $N$  est un infiniment petit du second ordre. En effet, son moment par rapport à  $N$  correspond à l'effet des rotations de centre  $N$  qui se traduisent par une faible vitesse sur la petite courbe reliant ces points directement dans leur voisinage.

Ainsi,  $T(M)(N)$  approxime  $f(N)$  au premier ordre au voisinage de  $M$ , ce qu'il fallait démontrer.

#### 4.5. Mécanique relativiste, introduction

Abordons maintenant l'étude de la mécanique relativiste. En un certain sens nous en avons déjà tout dit: c'est la théorie de l'équilibre que nous venons d'étudier, en employant dans le rôle d'espace physique l'espace à 4 dimensions de la relativité restreinte. On remplace "système" par "scénario", "énergie" par "action", et comme tout est à l'équilibre, la définition de l'équilibre (annulation de la différentielle de l'énergie) est érigée en principe, nommé le *principe de moindre action*.

Mais en fait, il reste plusieurs problèmes.

Un problème est de décrire, à l'intérieur de ce cadre général, quelles sont en particulier les propriétés mécaniques des objets physiques qui se rencontrent le plus couramment en pratique, à savoir les particules élémentaires ou non. Par mesure de simplicité on ne décrira pas d'autres choses comme les champs, en sorte qu'on ne verra que la notion élémentaire de particule sans rien préciser de ce qui fait en réalité la diversité des espèces de particules existantes.

Or, comme il s'agit là de décrire au mieux les propriétés des objets physiques en eux-mêmes sans s'occuper de leurs apparences expérimentales, le langage le mieux adapté pour l'intuition est le langage que nous avons appelé le langage théorique, à savoir celui de l'équilibre avec son énergie potentielle, et non celui de la mécanique relativiste. Il n'en reste pas moins que si ce langage est adapté, cela ne signifie pas que les choses soient réellement telles qu'elles semblent quand on les décrit dans ce langage. Cela signifie qu'il y a une correspondance entre les concepts de la réalité à décrire et ceux qui nous sont familiers par notre expérience des situations d'équilibre dans la vie de tous les jours, et que nous choisissons d'utiliser cette familiarité pour exprimer aisément les choses à décrire à travers cette correspondance. Or, tout irait bien dans cette affaire s'il n'y avait qu'une seule bonne manière de faire cette traduction, par laquelle les descriptions intuitives puissent s'élaborer sans hésitation. Hélas, ce n'est pas le cas car il y a deux possibilités. Comment est-ce possible ? Entre deux traductions, la différence consiste bien sûr en une retraduction de la représentation dans elle-même. Ici, dans la théorie de l'équilibre c'est le signe de l'énergie qui s'inverse.

En effet, la notion générale d'équilibre est conservée quand tous les signes des énergies sont également inversés. Seule la notion d'équilibre stable est perdue et remplacée par celle de l'équilibre le plus instable, celui de l'équilibre sur les sommets de potentiels, ce qui, il faut le dire, constitue malgré son extrême simplicité mathématique une gymnastique très inhabituelle pour notre intuition courante de l'équilibre. D'où la question ridicule en théorie mais grave en pratique pour nous pauvres humains devant savoir gérer au mieux les ressources de notre imagination: parmi ces deux traductions possibles, y en a-t-il une qui soit meilleure que l'autre, et suivant quels critères ?

Dans la section suivante, nous développerons une description théorique qualitative des propriétés mécaniques des particules en termes d'une traduction choisie suivant le fait qu'elle induit la plus grande fréquence de situations d'équilibre stable, en sorte de préserver le confort de notre intuition. Ce confort sera d'autant mieux préservé que la description utilisée se fera encore en termes de la géométrie euclidienne de l'espace, en dépit du fait que l'espace réel est pseudo-euclidien et que pour cette raison il n'existe aucune correspondance complètement cohérente et rigoureuse entre cette description et la réalité. On peut même aller plus loin en disant que ce problème du changement de géométrie relativise la définition d'un choix du signe de l'énergie dans les descriptions, car elle consiste justement à faire planer dans l'air un changement de signe qui remet en question les relations habituelles entre les signes des quantités en jeu dans les états d'équilibre, et leur rapport avec la notion de stabilité. D'où la difficulté pour l'intuition à appréhender la cohérence d'une description donnée lorsque la géométrie pseudo-euclidienne est prise en compte.

Puis, nous ferons quelques remarques sur la notion de stabilité, montrant qu'à travers la

correspondance entre théorie et expérience elle se comporte étrangement, en sorte qu'elle ne devrait même pas être prise comme référence et que le "principe de moindre action" porte mal son nom.

Enfin, nous établirons rigoureusement les lois générales de la mécanique relativiste exprimées en termes de leur formulation expérimentale usuelle, comme traductions des lois générales de l'équilibre précédemment établies appliquées à l'espace pseudo-euclidien suivant une certaine représentation théorique (dont la cohérence vient du fait qu'elle ne réutilise sans examen que ce qu'on a pu définir sans utiliser les spécificités de la géométrie euclidienne).

#### 4.6. Description des particules

Nous allons ici introduire une description théorique des propriétés mécaniques des particules comme dans un espace euclidien, déduite par traduction de leur comportement expérimental usuel sans tenir compte des questions d'accords de signe ou autres coefficients de proportionnalité liés à la vitesse de la lumière.

Qu'est-ce qu'une particule ? Expérimentalement, c'est un point animé d'une vitesse rectiligne uniforme quand il n'est soumis à aucune force; sa vitesse varie uniquement quand il est soumis à des forces de somme non nulle (ni avant, ni après).

Traduit théoriquement, cela donne une ligne qui est droite si aucune force extérieure ne s'exerce dessus, et qui est incurvée uniquement là où s'exercent des forces extérieures de somme non nulle. On trouve facilement dans la vie courante deux sortes d'objets se comportant ainsi: les ficelles et les élastiques. Ces deux objets présentent une résistance à l'élongation. Celle de l'élastique est variable tandis que celle de la ficelle est totale, pouvant se voir comme la limite d'un élastique dont la raideur tend vers l'infini. Par contre ils ne présentent aucune résistance à la courbure contrairement au fil de fer. Comme une telle résistance permettrait à la courbure de se prolonger en dehors du strict lieu d'application des forces extérieures contrairement à ce qu'indique l'expérience des particules, elle ne sera pas admise et le modèle des fils de fer sera donc exclu.

La seule question qui reste à préciser dans cette description est donc celle de la raideur de l'élastique, qui est sa résistance à l'élongation. Cette question se formule en termes du champ de potentiel sur l'espace de configuration: l'énergie de chaque partie d'un élastique s'exprime en fonction de sa longueur. La force de tension de l'élastique en chaque lieu est alors la dérivée de l'énergie par rapport à la longueur. Dans le cas d'une résistance infinie (un fil), la relation s'exprime en disant que la longueur est indépendante de la tension. Alors, qu'en est-il pour une particule ? Que dire de sa tension, à quoi correspond-elle expérimentalement ? Pour un élastique de tension donnée faiblement incurvé en un lieu par une force, la courbure (ou la différence de direction de part et d'autre du lieu d'application de la force) est proportionnelle à cette force, ce qui correspond bien expérimentalement à l'effet d'accélération proportionnel à la force exercée. Ainsi s'accordent les notions théorique et expérimentale de force comme étant proportionnelles. Plus précisément, la force expérimentale correspond à la densité de force théorique par rapport au temps.

Ensuite, pour un élastique de courbure donnée, la force nécessaire pour équilibrer cette courbure est proportionnelle à sa tension, de même que la force nécessaire pour donner à une particule une accélération donnée est proportionnelle à sa masse. Ainsi voit-on que la tension théorique d'une particule représente (est proportionnelle à) sa masse expérimentale.

Or, que nous dit l'expérience sur les masses des particules et leurs variation ? Elle nous dit que la masse de chaque espèce de particule élémentaire est une constante universelle. On a donc affaire à un élastique dont la tension est indépendante de sa longueur ! Ainsi son énergie est proportionnelle à sa longueur.

Connaissons-nous un tel phénomène dans la vie courante ? S'il n'est pas facile de trouver une substance visqueuse qui se comporte ainsi assez précisément quand on l'étire dans une dimension, on en connaît un équivalent en dimension 2: la surface de l'eau ou d'une bulle de savon véhicule aussi une tension superficielle constante, qui est cette fois une tension par unité de longueur du bord, et correspond à une densité surfacique d'énergie constante. En fait, comme on le verra plus loin, cette densité surfacique constante d'énergie dans sa description expérimentale correspond théoriquement à la composante temporelle du même phénomène de surface tendue en dimension 3, dans lequel cette tension ne privilégie aucune des directions de son extension dans l'espace-temps pseudo-euclidien. En effet, cette uniformité s'exprime par le fait qu'il n'y a pas de notion de vitesse

d'une partie ou d'un bord couissant dans la surface, qui ait des effets mécaniques mesurables dans la surface. Et donc, que les contractions et dilatations n'ont pas d'effet sur la tension.

Le même phénomène existe aussi en dimension 4, sauf qu'en l'occurrence celui qu'on observe semble s'étendre universellement sans avoir aucune frontière sur laquelle sa tension puisse s'exercer: c'est la *constante cosmologique*.

On peut expliquer ces propriétés des particules de la manière suivante: Prenons un tel élastique (une particule) attaché à ses extrémités (sa création et sa fin), saisissons-en un de ses points (événements) et tirons-le dans la direction de l'élastique. Le résultat précédent s'exprime ici par le fait que cet acte de tirer ainsi un point de l'élastique dans le sens de la longueur n'implique aucune transmission d'énergie par celui qui tire. Seule la partie avant de l'élastique transmet son énergie à la partie arrière.

Finalement, ce résultat apparaît naturel: car en saisissant un point et en le tirant, qu'avons-nous fait finalement ? En quoi le fait de saisir tel événement et de l'amener vers tel autre de la même ligne d'univers différerait-il du fait de saisir cet autre point dès le départ ? Car il s'agit là de la même particule !

Un élastique habituel est fait d'un nombre fini d'atomes, avec des nombres précis d'entre eux de part et d'autre du point qu'on saisira, en sorte que le choix du lieu où le saisit a une implication physique sur l'élastique, qui intervient lors de l'allongement que peuvent ensuite subir les parties de l'élastique de chaque côté de ce point. Mais il ne se passe rien de tel dans une particule: la forme géométrique de chacune de ses parties suffit à la décrire physiquement en toute circonstance.

On peut également réinterpréter ce phénomène en termes de rugosité: pour que saisir un point de l'élastique et le déplacer puisse réellement déplacer quelque chose dedans pour modifier l'état de la particule et donc son potentiel, il faudrait qu'il y ait quelque chose d'irrégulier dans sa longueur, que ce déplacement puisse affecter. Or, de cela il n'y a rien, puisque la particule est parfaitement lisse.

Bien sûr, dans une "grosse particule" formée d'un grand nombre d'atomes, comme par exemple un grain de sable, l'agitation thermique des atomes constitue dans la dimension temporelle une rugosité sur l'appui de laquelle on peut tirer. La force théorique qui prend alors appui sur cette rugosité pour tirer dans la dimension temporelle, et entraîner une variation de tension (masse) entre les deux parties (passé et futur), s'appelle expérimentalement la transmission d'une quantité de chaleur vers ce grain.

### *Le spin*

Citons pour finir la dernière propriété mécanique des particules, qui est leur spin. Mais en fait, il n'est pas exact que ce soit encore vraiment là une propriété mécanique dont on puisse rendre compte suivant l'expression de la mécanique classique (ou de l'équilibre) que nous avons employée jusqu'à présent. En effet, le spin est par nature une propriété quantique des particules, qui met en défaut certaines règles de la mécanique classique. Une attitude sage serait de passer simplement sous silence ce phénomène étrange en le négligeant, ce que nous aurions le droit de faire, puisqu'à l'échelle de la mécanique classique qui est celle des phénomènes que nous avons décrit et où les approximations générales nécessaires à la pertinence de cette description sont valables, le spin est également négligeable. Malgré cela, nous allons présenter ici le spin dans les termes de la mécanique classique, et ce faisant nous allons préciser en quoi cette mécanique classique par laquelle nous exprimons ce phénomène est mise en défaut.

Voici: le spin représente la torsion de l'élastique, c'est-à-dire, dans le cas où il suit une droite (la particule est isolée), le moment de la force véhiculée par rapport à cette droite elle-même. L'expérience quotidienne de ce phénomène est facile: on pend un objet à un fil que l'on tient par le haut, on roule le fil dans la main et l'objet se met à tourner. L'explication de cette expérience courante est facile: le nombre de tours que l'on a fait subir au fil sur lui-même se trouve enregistré dans les détails de la disposition de la matière que constitue le fil, et à cet état de torsion du fil ainsi affecté est associé une énergie, de laquelle dérive le moment de la force qui fera tourner l'objet suspendu.

Seulement, dans le cas du spin d'une particule, cette explication tombe à l'eau: son état en tant qu'élastique à l'équilibre dans l'espace pseudo-euclidien n'est pas affecté par le mouvement de

torsion qu'on inculque à une de ses extrémités; pourtant, une puissance est réellement transmise dans ce mouvement. Au fait, peut-on réellement tenir et rouler ainsi la particule sur elle-même ? En un sens non, puisqu'il n'y a pas de rugosité dans la largeur de la particule, sur laquelle ce mouvement pourrait s'exercer. Mais d'autre part, le moment de la force véhiculée est bien réel. En quoi peut-il alors consister ?

Ce modèle a cette part d'incohérence, qu'on peut "fournir une puissance au système" par un mouvement de son bord qui ne modifie pas l'état du système. Cela peut se comprendre par le fait que cet expérience doit être interprétée autrement, en disant que qu'il faut inclure dans l'état du système celui de son bord, à savoir ici, dans quelle direction latérale est orientée la main qui tient l'élastique en chaque extrémité. Cette main ne peut alors changer sa direction qu'en tournant et fournissant donc une énergie, dans le cas où cette prise définit la terminaison physique de l'extension de l'élastique. Si on fait faire un tour complet à cette direction autour de l'axe de l'élastique, alors on lui a donné ou pris une énergie précise correspondant au spin, en revenant à l'état initial. Et c'est là que l'insuffisance de ce mode de représentation apparaît.

Mais si au contraire l'élastique se prolonge plus loin, on ne peut définir la manière dont l'énergie se répartit de chaque côté d'une surface de coupe qu'en choisissant une direction latérale à cette coupe. Si on coupe suivant cette direction, cela donne aux deux extrémités issues de cette coupe cette même direction de prise.

### *Stabilité et instabilité, exemples*

Nous allons aborder la question de la stabilité de l'équilibre au sens théorique en mécanique relativiste, et son rapport avec la stabilité de l'équilibre expérimental pour le cas d'un système n'évoluant pas au cours du temps. La question est: y a-t-il des scénarios qui constituent des minima ou maxima d'action, et lesquels ? On verra en fait qu'en sortant de l'exemple restrictif des particules il n'y en a pas, d'autant plus que cette action qui est minimisée dans le cas des particules correspond à l'opposée de l'énergie potentielle. (Ceci a bien sûr l'avantage que la "force" issue d'une action se définit par la différentielle de l'action et non son opposé.) Comme les équilibres expérimentaux minimisent cette énergie, ils maximisent l'action par rapport aux scénarios voisins de fixité par rapport au temps. Ceci ruine tout espoir de voir l'action minimisée ou maximisée en général (et le changement de géométrie n'y change rien).

[Quelques exemples sont donnés ici; d'autres exemples complémentaires de ceux-ci pourront être ajoutés à l'avenir]

On pourrait imaginer un élastique (une particule) tendu au fond d'une vallée de potentiel (on suppose qu'il est lourd et ne décolle pas). Ce serait un état stable d'un point de vue théorique, mais que donne-t-il expérimentalement ? Pour le voir, il faut attacher ses extrémités ailleurs que précisément au fond de la vallée. Alors, en parcourant la vallée, que voit-on ? Qu'il reste quasiment au fond sauf à la fin (et de même au début symétriquement) où il s'écarte de manière accélérée. C'est donc, expérimentalement, un équilibre instable.

Veut-on faire le contraire et demander un équilibre stable d'un point de vue expérimental ? Le scénario de la particule restant fixe et les scénarios de la particule oscillant suivant n'importe quelle amplitude autour du point d'équilibre stable étant tous permis par la mécanique relativiste, on peut transiter continuellement d'un scénario à l'autre en fixant les extrémités (la particule partant et finissant en sa position d'équilibre). Alors l'action ne saurait varier puisque sa variation est nulle au premier ordre à toute étape de ce changement. C'est donc que la variation de l'action due à l'écart par rapport à la position d'équilibre d'après la force extérieure compense celle due à l'oscillation (action interne). Le rapport précis de compensation entre les deux est ce qui détermine la période de l'oscillation: une variation plus serrée du potentiel nécessite pour être compensée une oscillation plus serrée (rapide).

Donc, si on veut que le potentiel extérieur ressemble en théorie à son apparence expérimentale afin que les notions de stabilité se correspondent, il faut changer le signe par rapport à ce que nous avons décrit jusqu'à présent et considérer l'élastique tendu (la particule) comme un système instable, comme si on comprimait l'élastique et qu'il résiste du fait qu'il n'a pas décidé de quel côté il allait se plier. En fait, le point important ici n'est pas le signe de la tension mais l'instabilité fondamentale de la particule comme élastique théorique en prenant comme référence de stabilité

celle du point de vue expérimental.

C'est pourquoi, le "principe de moindre action" représentant la minimisation de l'action des scénarios d'une particule isolée, il a été convenu de donner à l'action le signe opposé à l'énergie au sens expérimental (le sens expérimental étant le seul à déterminer une correspondance entre les deux tandis que la correspondance théorique est au libre choix des goûts et des couleurs).

#### 4.7. Mécanique relativiste, formulation

Etablissons enfin les lois fondamentales de la mécanique relativiste comme traductions expérimentales des lois de l'équilibre.

Le point de vue expérimental part du choix d'un référentiel. Un référentiel s'exprime de manière équivalente par sa direction temporelle, ou par son vecteur temps noté  $\vec{t}$  (objet traité comme un vecteur mais qu'il faut en fait multiplier par une quantité de temps pour obtenir un vrai vecteur géométrique), ou son espace  $S$  orthogonal à  $\vec{t}$  (des vecteurs de l'espace usuel), ou la coordonnée de temps  $T$  (de même direction que  $\vec{t}$  au sens du produit scalaire), qui est une forme linéaire multipliée par une quantité de temps, autrement dit définie par les relations:  $(T(\vec{t}) = 1$  et une équation de  $S$  est  $T(M) = 0$ ).

Voici les conventions. On a un champ d'action sur l'espace de configuration théorique. Choisissons de représenter théoriquement le principe de moindre action en termes d'équilibre, en définissant l'énergie théorique comme étant l'opposée de l'action (ce qui dans notre approche constitue en fait la définition de l'action à partir de l'expérience), et donc la force théorique est égale à la différentielle de l'action. Nous définirons la force expérimentale comme étant la restriction à  $S$  de la densité de force théorique par rapport au temps. En effet, nous allons démontrer que cela redonne toutes les propriétés expérimentales usuelles de la force, et ceci nous sert de définition des conventions de signes théoriques en fonction des données expérimentales. (Ce choix de faire correspondre ainsi les forces théorique et expérimentale sans changement de signe semble en effet le choix de convention le plus naturel).

##### *Lois de conservation*

Nous allons obtenir les lois de conservation de la mécanique relativiste comme traduction expérimentale des bilans des forces extérieures que nous avons présentés en mécanique de l'équilibre, avec un choix de référentiel. Considérons un système en évolution (scénario). Chaque instant (élément de la droite d'arrivée de  $T$ ) définit un sous-espace de dimension 3 qui divise le scénario en une partie passée et une partie future. Alors, l'objet qui "se conserve" se définit comme étant le torseur (résultante et moment) des forces théoriques que la partie passée exerce sur la partie future.

En effet, considérons le système théorique délimité par les deux faces parallèles d'équations  $T(M) = t_1$  et  $T(M) = t_2$  (avec  $t_1 < t_2$  fixés) et des parois latérales. L'annulation de la somme des forces appliquées à ce système se réécrit en disant que ce qui sort par  $t_2$  moins ce qui entre par  $t_1$  est égal à la somme de ce qui entre par les parois latérales ou qui s'applique dans l'intérieur du système. C'est donc la variation expérimentale de l'objet associé au système entre les instant  $t_1$  et  $t_2$  qui est égale au torseur théorique des forces appliquées au système entre ces instants. Donc, la dérivée par rapport au temps de cette quantité est égale à la densité temporelle de l'exercice de ce torseur. En particulier, cette quantité se conserve au cours du temps dans le cas d'un système isolé (non soumis à des forces autres que les forces internes apparaissant dans les coupes du système en  $t_1$  et  $t_2$ ).

Enumérons à présent ces quantités qui se conservent dans le cas d'un système isolé:

- La résultante théorique  $f$  se décompose expérimentalement en sa partie spatiale (restriction de  $f$  à  $S$ ) appelée *quantité de mouvement* ou *impulsion* et notée  $p$ , et sa partie temporelle appelée *énergie*, définie par  $E = -f(\vec{t})$ . La conversion de ce  $f = p - ET$  en vecteur par le produit scalaire en unité d'espace définit le *quadrivecteur énergie-impulsion*  $\vec{p} + \frac{E}{c^2}\vec{t}$ . Sa norme en mesure temporelle s'appelle sa *masse*

- Le moment (par rapport à une origine arbitraire de l'espace-temps), qui se décompose en deux parties: d'une part, sa restriction aux rotations spatiales appelée le *moment cinétique* qui correspond aux rotations spatiales (celles de  $S$ ), et d'autre part ce qui représente le centre d'inertie du système (voir plus bas).

Il ne reste plus qu'à vérifier que tout cela redonne l'expression expérimentale usuelle des lois de la mécanique (classique ou relativiste).

Déjà, on tire immédiatement des définitions et résultats ci-dessus la

**Relation fondamentale de la dynamique.** *La dérivée par rapport au temps de l'impulsion d'un système est égale à la résultante des forces extérieures appliquées au système en cet instant.*

La difficulté dans la vérification de ce résultat est de montrer l'équivalence des définitions expérimentales des termes employés (force et impulsion) à leurs définitions usuelles. Comme d'autre part on admet la relation fondamentale de la dynamique comme principe physique, le résultat ci-dessus est en fait la confirmation de notre définition de la force expérimentale comme densité temporelle de force théorique, restreinte à l'espace.

Relions d'abord la force expérimentale à l'énergie expérimentale:

**Théorème de la puissance.** *La puissance mécanique expérimentale exercée par une force  $f$  sur un système évoluant à la vitesse  $v$  est égale à  $f(v)$ .*

Ici, la puissance désigne la densité temporelle (ou débit) d'énergie transmise par cette force; "mécanique" signifie que toute transmission d'une quantité de chaleur est retranchée de ce bilan;  $v$  est un vecteur tangent à l'espace de configuration des états expérimentaux; ces états sont définis théoriquement comme étant les coupes possibles de scénarios par un espace fixe  $S_1$  d'équation  $T(M) = \text{constante}$ .

Le vecteur tangent  $v$  représente la vitesse de variation des coupes successives suivant le temps  $T$ , ramenées abstraitement par translation temporelle à l'espace  $S_1$ . Mais nous allons l'utiliser comme vecteur tangent à l'espace de tous les scénarios par un prolongement arbitraire, dont on demande seulement que dans ce mouvement théorique, la coupe par  $S_1$  demeure contenue dans  $S_1$  et y coïncide avec le mouvement expérimental de vitesse  $v$  invoqué.

La force  $f$  représente la force expérimentale; notons  $f'$  la force théorique correspondante exercée entre deux instants infiniment proches l'un de l'autre, divisée par l'intervalle entre ces instants. (Ceci complète donc  $f$  par les composantes temporelles de cette force théorique, qui l'accompagnent naturellement.)

Quelle est donc la puissance transmise par cette force au système ? D'après la définition de l'énergie expérimentale donnée plus haut, cette puissance vaut  $-f'(\vec{t})$  où  $\vec{t}$  représente le mouvement théorique de translation dans la direction du temps, du passé vers le futur (et à la vitesse du temps qui passe). Pour montrer l'égalité voulue, remarquons que la différence entre ces termes vaut  $f'(\vec{t} + v)$ . Or  $\vec{t} + v$  représente un mouvement de coulissement du scénario sur lui-même, auquel la force  $f'$  doit être insensible si elle ne comporte pas d'échange thermique avec le système. Bien sûr, dans le cas contraire cette différence correspond à la quantité de chaleur transmise.

Ainsi est retrouvée la relation entre force et énergie constituant la théorie de l'équilibre dans le cas expérimental, comme conséquence du principe de moindre action.

*Exemple de calcul concret*

Soit une particule de masse  $m$  immobile au départ éclatant en deux particules de masses  $m_1$  et  $m_2$ . Calculer leurs énergies  $E_1, E_2$ .

Solution : soient  $p, p_1$  et  $p_2$  les quadrivecteurs énergie-impulsion des particules respectives. Leur conservation s'exprime par  $p = p_1 + p_2$ . La particule de départ étant déclarée immobile,  $p$  est dans la direction du vecteur temps; la composante temporelle de  $p_1$  étant  $\frac{E_1}{c^2}$ , on a le calcul suivant avec le produit scalaire en mesure temporelle :

$$\begin{aligned} m_2^2 &= p_2^2 = (p - p_1)^2 = m^2 + m_1^2 - 2m \frac{E_1}{c^2} \\ \implies E_1 &= \frac{c^2}{2m} (m^2 + m_1^2 - m_2^2) \end{aligned}$$

*Centre d'inertie*

Nous allons définir le centre d'inertie d'un système en termes de la mesure du moment du système sur les rotations théoriques d'axes de directions spatiales (axes de dimension 2 parallèles

à  $S$ ). Notons que ces rotations ne forment pas un groupe (la composée de deux telles rotations n'est pas une telle rotation). Mais rappelons que ce sont seulement les vitesses de rotation qui nous importent.

D'abord, donnons une description géométrique de ces rotations, en commençant par les rotations vectorielles (ou autrement dit, dont l'axe plan passe par un point origine fixé).

Voici: une rotation qui déplace le vecteur temps  $\vec{t}$  suivant une vitesse  $\vec{w} \in S$ , déplace tout  $\vec{v} \in S$  suivant la vitesse  $\frac{\vec{w} \cdot \vec{v}}{c^2} \vec{t}$ .

Ensuite, considérons l'ensemble des rotations d'axe contenu dans un même espace de simultanéité  $S(t)$  (espace euclidien de direction  $S$  formé des évènements d'une même date  $t$  dans le référentiel choisi). La propriété commune de toutes ces rotations est le fait que la vitesse de tout point de  $S(t)$  est colinéaire à  $\vec{t}$ . Or il suffit de décrire le mouvement de tout point de  $S(t)$  car le reste de la rotation s'en déduit (comme le mouvement d'un plan détermine celui des solides posés dessus). Ainsi, on représentera une telle rotation par l'application affine  $w$  définie sur  $S(t)$  qui définit la vitesse de tout point  $N \in S(t)$  par l'expression  $-\frac{w(N)\vec{t}}{c^2}$ .

Nous venons ainsi de décrire géométriquement un certain sous-espace vectoriel de l'algèbre de Lie du groupe des déplacements théoriques. Décrivons maintenant la forme de la restriction d'un tenseur à ce sous-espace, autrement dit une forme linéaire sur ce sous-espace.

A quoi ressemble donc une forme linéaire sur l'espace des formes affines sur l'espace euclidien  $S(t)$ ? C'est un point pondéré de  $S(t)$ , autrement dit, ou bien un vecteur auquel s'applique la partie linéaire de  $w$ , ou bien un point  $G$  pondéré par une masse  $m$ , désignant l'application linéaire qui à tout  $w$  associe  $mw(G)$ . Attention: ce que nous appelons ici et plus bas la masse ne doit pas être confondu avec la notion de masse que nous avons invoquée précédemment en parlant de la masse d'une particule, et qui est la définition communément admise en mécanique relativiste (elle désigne ordinairement la norme du vecteur de la résultante conservée, tandis que ce que nous appelons provisoirement ici la masse est sa composante temporelle).

Calculons la masse  $m$  de ce point pondéré: elle s'obtient en appliquant cette expression au cas de la forme affine constante  $w$  égale à 1. Ceci correspond au mouvement qui déplace tout point de  $S(t)$  suivant le même vecteur  $\frac{-1}{c^2} \vec{t}$ . C'est donc le mouvement de translation de vitesse  $\frac{-1}{c^2} \vec{t}$ .

Or, nous avons défini l'énergie  $E$  du système comme étant moins la composante temporelle de la résultante, autrement dit la puissance théorique transmise par le passé sur le futur lors d'une translation théorique de vitesse  $-\vec{t}$ .

D'où finalement

$$E = mc^2.$$

Mais bien sûr pour l'obtenir, nous avons triché en ajustant le coefficient dans la définition de  $w$  qui se répercute sur la définition de  $m$ . Il ne nous reste plus qu'à justifier ce coefficient par le fait qu'il en résulte la propriété suivante:

**Mouvement du centre d'inertie.** *Dans un référentiel donné, la vitesse de variation au cours du temps du centre pondéré d'inertie  $mG$  d'un système isolé est égale au vecteur impulsion  $\vec{p}$  du système (correspondant à la forme linéaire impulsion, par le produit scalaire spatial usuel).*

Nous allons donner deux démonstrations de ce résultat, l'une simple et intuitive, l'autre complète et rigoureuse.

*Démonstration 1.*

Reprenons la remarque démontrée dans le cas du plan que le moment par rapport à un point s'exprime par une forme affine dont la direction est celle du vecteur force. L'évolution du centre d'inertie constitue une droite de l'espace-temps dirigée par le quadrivecteur énergie-impulsion  $\vec{p} + \frac{E}{c^2} \vec{t}$ . Une droite dirigée par ce vecteur apparaît expérimentalement comme un point animé de la vitesse  $\frac{\vec{p}}{m}$  si ce  $m$  est défini par  $E = mc^2$ .

*Démonstration 2.*

On reprend une démonstration complète partant de la construction des tenseurs en termes de rotations.

Pour vérifier la relation énoncée entre centres pondérés d'inertie, fixons une forme affine  $w$  sur l'espace euclidien  $S$  du référentiel, qui est identifié à chacun des espaces  $S(t)$ , où la correspondance entre tous les  $S(t)$  se définit par la translation temporelle dans la direction de  $\vec{t}$ . Quelle est alors la relation entre  $mw(G(t))$  et  $mw(G(t'))$  ? C'est le même torseur  $\mu$  qui s'applique aux deux vitesses de rotations  $r(t)$  et  $r(t')$  définies par le même  $w$  sur les différents espaces  $S(t)$  et  $S(t')$ . On a donc:

$$mw(G(t')) - mw(G(t)) = \mu(r(t')) - \mu(r(t)) = \mu(r(t') - r(t)).$$

Ensuite, le mouvement  $r(t') - r(t)$  étant la différence de deux mouvements de rotation de même partie linéaire, est un mouvement de translation. Calculons le vecteur vitesse de ce mouvement. Une façon de le faire est de remarquer qu'on passe de l'objet géométrique  $r(t)$  à l'objet  $r(t')$  par la translation  $\tau$  de vecteur  $(t' - t)\vec{t}$ . Donc, pour tout point  $M$ ,

$$(r(t') - r(t))(M) = \tau(r(t))(M) - r(t)(M) = r(t)(\tau^{-1}(M) - M) = (t - t')r(t)(\vec{t}).$$

Enfin, la valeur de  $r(t)(\vec{t})$  se déduit de la définition du rapport entre  $w$  et  $r(t)$  et de la description donnée plus haut de ces rotations:  $r(t)(\vec{t}) = -\vec{w}$  où  $\vec{w}$  est le vecteur représentant la partie linéaire de  $w$ .

Donc,

$$mw(G(t')) - mw(G(t)) = \mu((t - t')r(t)(\vec{t})) = (t' - t)\mu(\vec{w}) = (t' - t)p(\vec{w}) = (t' - t)w(\vec{p}).$$

Ceci étant valable pour tout  $w$ , on conclut:

$$mG(t') - mG(t) = (t' - t)\vec{p}.$$

#### 4.8. L'espace des phases

La mécanique classique que nous avons exposé se traduit en une loi d'évolution, par laquelle l'état d'un système à un instant donné détermine son état à venir. Voyons les notions de base qui interviennent dans cette question.

L'état d'une particule classique consiste en sa position et sa vitesse, ou encore sa position et son impulsion. Plus généralement pour un système quelconque, son état en un instant se définit par la coupe du scénario par cet instant dans l'espace-temps, et l'ensemble de toutes les forces théoriques (les impulsions) qui sont véhiculées à travers cette coupe. En effet, pour un système donné à l'équilibre, je peux le couper sans perturber l'équilibre de la moitié qui m'intéresse à condition de maintenir en place toute la disposition de ce qui se trouve sur cette coupe et d'y exercer pour cela toutes les forces nécessaires. Si je m'intéresse à l'évolution que cette moitié aurait si le système évoluait sans avoir été coupé, je peux reproduire la même évolution en exerçant sur la coupe les mêmes mouvements et forces qui seraient véhiculées à travers ce plan si le système n'était pas coupé. En effet, il n'y a pas à prendre en compte de mouvements de cette coupe qui sortent de son plan, puisque tout est fait de particules qui, comme nous avons décrit, peuvent être prolongées par la simple traction de leur bord suivant leur long, sans en être affectées. Ainsi tout mouvement qui sortirait de ce plan est représentable de manière unique par un mouvement demeurant dedans. Pour la même raison, il n'est pas besoin de préciser la composante temporelle des forces théoriques (les énergies) mais seulement celles qui agissent suivant la direction du plan de coupe (impulsions), car cette composante temporelle est déterminée de manière unique par ces conditions.

On appelle espace des phases d'un système la variété constituée de tous les états qu'il pourrait avoir à un instant donné, où la notion de l'état d'une coupe est constituée de sa forme et celle des forces qui la traversent (impulsions).

Nous l'avons défini à l'aide d'une coupe, mais l'usage que nous en ferons pourrait s'appliquer à des manières beaucoup plus générales de maintenir le bord d'un système.

L'important, donc, est qu'on a un système en équilibre (un scénario) pouvant être soumis à des potentiels quelconques, et qu'il y a deux acteurs qui, chacun de son côté, maintient une partie du système, et peut agir dessus en la déplaçant et en modifiant la force qu'il exerce dessus.

Ces deux parties sont disjointes l'une de l'autre, en sorte que les acteurs interagissent uniquement par l'intermédiaire du système: les actions de l'un sur le système sont ressenties par l'autre et inversement. Si l'un choisit par exemple de maintenir fixe sa partie, l'autre pourra bouger la sienne et affectera la force à laquelle le premier doit résister pour maintenir cette fixité. Ou s'il décide de laisser sa partie suivre le mouvement suivant une certaine élasticité, les actions de l'autre la feront bouger.

Définissons l'espace des phases comme étant l'ensemble des états d'équilibre possibles pour le système dont on oublie les voisinages immédiats des deux acteurs, et dont la réalisation dépend du bon vouloir conjoint de ceux-ci. (On ne s'intéresse pas à la manière dont chacun se débrouille de son côté pour faire évoluer sa prise en sorte de s'accorder à chaque instant avec la position du système dans l'espace des phases). Et subjectivement, chaque acteur se représente l'espace des phases en décrivant tout état d'équilibre du système en dehors de lui sous la forme de ce qu'il peut ressentir de son côté lorsque le système se trouve dans cet état.

**Théorème.** *Les structures symplectiques attribuées par chaque acteur à l'espace des phases sont opposées (donc égales si l'un a le rôle d'émetteur et l'autre de récepteur).*

Une structure symplectique est (à peu de chose près) une application qui à toute courbe fermée munie d'un sens de parcours, associe une quantité d'action (que nous traduisons par le terme d'énergie dans notre langage théorique), d'une manière linéaire (condition que nous ne détaillerons pas; disons seulement que la même courbe parcourue dans deux sens opposés doit donner des résultats opposés). A toute courbe fermée de l'espace des phases, donc, chaque acteur associe le travail total qu'il doit fournir au système lorsque l'état du système effectue le parcours de cette courbe (infiniment lentement: l'état à chaque instant est supposé être un état d'équilibre), en supposant que cet acteur reprend à l'arrivée la même disposition qu'il avait au départ (ce qui est possible puisque le système parcourant une boucle se retrouve à la fin dans le même état qu'au départ et donc il ne tient qu'à l'acteur de se remettre dans le même état).

Ces deux structures symplectiques sont opposées car le système est supposé être uniquement soumis aux actions libres des deux acteurs d'une part, à des forces issues de potentiels d'autre part. Quand le système revient à sa configuration initiale après avoir fait un tour, donc, le total des énergies fournies au système par toutes les forces est nul à l'exception du travail fourni par les deux acteurs. Au bilan, le travail total donné par l'un est donc celui reçu par l'autre, ce qu'il fallait démontrer.

De cela découle le

**Théorème de Liouville.** *Le mouvement de l'espace des phases dans lui-même qui exprime la loi d'évolution d'un système au cours du temps, préserve le volume.*

Comme nous venons de voir qu'il préserve la structure symplectique, il suffit de donner une définition de la mesure du volume dans l'espace des phases qui dépend uniquement de cette structure symplectique. Note: la notion de volume n'a un sens mathématiquement cohérent que dans des variétés de dimension finie. Nous supposons donc que l'espace des phases est de dimension finie, ce qui est le cas en pratique pour un système de particules (il y a 6 dimensions par particule), mais non pour le cas d'un champ.

Dans le cas d'un espace des phases de dimension 2, l'aire d'une surface se définit par la quantité d'action associée à la courbe frontière de cette surface, une orientation étant donnée. Dans le cas général (en dimension finie), la définition est plus compliquée mais elle existe; elle consiste en une simple expression tensorielle dont nous admettons ici l'existence.

## 5. Mécanique statistique

### 5.1. Fondements de la mécanique statistique.

Il y a deux manières possibles de fonder la mécanique statistique: sur la mécanique classique d'une part, sur la mécanique quantique d'autre part. Comme la réalité est quantique et que la mécanique classique en est une approximation ayant un large domaine de validité pratique, ainsi la mécanique statistique n'est pas exactement la même suivant qu'on la fonde de l'une ou l'autre manière, et c'est celle fondée sur la mécanique quantique qui décrit plus précisément la réalité. Cela n'empêche pas de chercher à la décrire et l'expliquer dans les termes de la mécanique classique, ce qui est une approche valable dans nombre de situations pratiques.

Contrairement à ce à quoi on pourrait s'attendre a priori, son fondement sur la physique classique comporte des difficultés mathématiques, notamment des problèmes de divergence, qui n'existent pas avec la physique quantique. Ainsi peut-on dire que sous certains aspects, le fondement sur la physique quantique est plus aisé pour arriver quasiment aux mêmes résultats. Comme il correspond mieux à la réalité, il est plus satisfaisant.

De quoi s'agit-il ? Nous avons montré que la loi d'évolution au cours du temps en mécanique classique s'exprime par une transformation de l'espace des phases  $\mathcal{E}$  qui préserve son volume (théorème de Liouville).

Le passage à la physique statistique consiste à dire qu'on ne connaît plus l'état  $x \in \mathcal{E}$  du système dans l'espace des phases, mais que cet état est aléatoire suivant une loi de probabilité bien déterminée appelée l'état statistique du système, qui s'exprime par une densité de probabilité: on peut imaginer la densité de probabilité comme la densité d'une substance présente en quantité totale égale à 1, comme une cuillère de café dans la marmite de l'espace des phases. La densité de probabilité en chaque point est le rapport de la probabilité de présence dans un petit voisinage de ce point sur le volume de ce voisinage.

De là on peut définir l'entropie comme une mesure de la dilution de cette probabilité:

$$S = - \int_{\mathcal{E}} p \ln p$$

moyenne par rapport aux possibilités élémentaires de se retrouver en les différents lieux de  $\mathcal{E}$  (donc intégrale suivant  $p$  fois l'élément de volume) de la mesure  $-\ln p$  de cette dilution.

Un des problèmes que cela pose est le problème de la divergence, notamment du fait que dans le cas d'un champ classique l'espace des phases est de dimension infinie et donc cela pose des problèmes à la notion d'élément de volume. Concrètement, si on se détourne de la simple manie de pinaillage d'analyse des mathématiciens pour adopter les grandes démarches d'approximations chères aux physiciens, il n'en reste pas moins que les divergences en question devraient se retrouver concrètement dans les conséquences expérimentales de ces hypothèses, ce qu'on a appelé la "catastrophe ultraviolette": en considérant l'espace des phases de dimension infinie du champ électromagnétique, le rayonnement thermique du corps noir d'une température donnée devrait avoir une intensité tendant vers l'infini au fur et à mesure qu'on inclut dans le calcul ces dimensions correspondant aux fréquences de plus en plus élevées, qui sont autant de voies de rayonnement supplémentaires.

Un autre inconvénient de cette présentation est qu'elle semblerait indiquer une conservation et non une augmentation de l'entropie, le mouvement déplaçant les concentrations de probabilité et ne les diluant pas. Ceci peut être surmonté sous forme de l'approximation d'un mélange (résultat de mouvement où la concentration demeure mais est dispersée un peu partout) par une dilution, du fait qu'il est pratiquement impossible de réeffectuer à l'inverse ce mouvement assez précisément pour effectivement défaire le mélange de façon constatable en pratique.

Un autre défaut encore de l'approche par la physique classique est qu'elle n'admet pas la possibilité d'une variation du nombre de particules présentes. En effet, la dimension de l'espace vaut six fois le nombre de particules présentes, donc une création ou une annihilation de particules devrait s'exprimer par une transformation préservant le volume d'un espace de phases vers un autre de dimension différente, ce qui apparaît quelque peu absurde. En effet, si à la limite on

pourrait considérer une création de particules comme plongeant le premier espace en une sous-variété du suivant (définie par le point exact où la nouvelle particule a été créée), la densité de probabilité qui avait une dilution volumique dans le premier n'aurait plus une dilution volumique dans le second mais une concentration sur une sous-variété, ce qui donnerait alors à l'entropie finale la valeur de moins l'infini (!). Par exemple, si on interprète classiquement la lumière comme système de particules et non comme onde, toute émission de lumière apparaît comme une création de particules.

Par rapport à cela, que nous dit la physique quantique ? Qu'il y a une limite à la division du volume de l'espace des phases (et donc une limite à la concentration de la probabilité). En effet, dans toute direction de surface de dimension 2 dans l'espace des phases, on a une mesure des aires donnée par la restriction de la structure symplectique à cette surface. Une aire  $y$  étant ainsi homogène à une action, on a comme unité de mesure naturelle de cette aire la constante de Planck  $\hbar$ . Alors, la possibilité de division de cette surface est limitée comme ne pouvant pas être plus fine qu'un découpage en cellules d'aire  $h = 2\pi\hbar$  chacune (cela se comporte comme tel découpage bien que les frontières entre les cellules ne soient pas localisables en termes de la géométrie usuelle).

Ainsi s'exprime le principe d'incertitude de Heisenberg: l'état d'un système dans l'espace des phases n'est pas plus précisément défini suivant chaque direction de surface, que par la donnée de la cellule qui le contient dans un tel découpage flou. Par conséquent, le volume minimal qu'on peut considérer dans un espace des phases de dimension  $2k$  vaut  $h^k$ . En fait, la réalité qui se cache derrière l'apparence classique d'une portion de l'espace des phases de volume  $nh^k$ , est un espace de Hilbert de dimension  $n$ , et ce qui se cache derrière une apparente évolution incompressible de ce volume, est une rotation l'espace de Hilbert.

Alors, nous allons ici développer la mécanique statistique à partir d'un modèle de départ mathématiquement simple, intuitivement proche de la mécanique classique tout en reflétant assez bien les propriétés essentielles de la réalité quantique.

Il consiste à partir d'un ensemble  $\mathcal{E}$  discret (fini ou dénombrable) nommé "ensemble des états classiques". Il s'interprète classiquement comme l'ensemble des cellules de volume  $h^k$  en lesquelles on aurait découpé l'espace des phases, ou quantiquement comme celui des vecteurs d'une base choisie dans l'espace de Hilbert. Nous oublierons totalement la forme géométrique de l'espace des phases, son rapport avec l'espace physique et ses points d'équilibre au sens expérimental classique.

On considère qu'on ne connaît pas la valeur  $x \in \mathcal{E}$  de l'état classique du système mais que cet état est aléatoire suivant une loi de probabilité  $p$  bien déterminée: à chaque état classique  $x$  est associé sa probabilité  $p(x)$  d'être réalisé. Cette loi de probabilité  $p$  est appelée l'*état statistique* (ou thermodynamique) du système. Un état statistique sur un espace  $\mathcal{E}$  est donc une application  $p$  de  $\mathcal{E}$  dans  $[0, 1]$  dont la somme des valeurs vaut 1:

$$\sum_{x \in \mathcal{E}} p(x) = 1.$$

On passe ainsi de la notion d'état classique à la notion d'état statistique en remplaçant l'espace  $\mathcal{E}$  par l'ensemble des lois de probabilité sur  $\mathcal{E}$ . On passe alors de la notion d'équilibre classique à celle d'équilibre statistique (ou thermodynamique). Un état d'équilibre statistique est un état statistique qui ne varie pas au cours du temps. L'équilibre statistique n'implique pas l'équilibre classique car il énonce seulement que la diminution de chaque  $p(x)$  par l'évolution possible de l'état  $x$  vers d'autres états sera exactement compensé par son augmentation due aux évolutions des autres états vers  $x$ , en sorte que  $p(x)$  demeure constant.

Pour pouvoir établir ses propriétés, il faut partir d'un postulat physique que nous allons énoncer maintenant. Introduisons d'abord des notations.

Soit un système  $\mathcal{S}$  isolé, d'état statistique  $p_1$  sur un espace  $\mathcal{E}$  à l'instant  $t_1$ , en évolution vers un espace  $\mathcal{E}'$  à l'instant  $t_2 > t_1$ . Pour tous  $x \in \mathcal{E}$  et  $y \in \mathcal{E}'$ , notons  $p(x, y)$  la probabilité pour que, si son état est  $x$  à l'instant  $t_1$ , il passe à  $y$  à l'instant  $t_2$ . On a évidemment  $\forall x \in \mathcal{E}, \sum_{y \in \mathcal{E}'} p(x, y) = 1$ , et cet objet détermine la transformation de l'état statistique du système de  $p_1$  à l'instant  $t_1$  à  $p_2$  à l'instant  $t_2$  par la formule

$$\forall y \in \mathcal{E}', p_2(y) = \sum_x p(x, y)p_1(x).$$

L'axiome de base dont nous allons partir, s'énonce alors par la formule :

$$\forall y \in \mathcal{E}', \sum_{x \in \mathcal{E}} p(x, y) \leq 1$$

ou bien cette même formule mais comme une égalité, en fonction du contexte pour la raison suivante: fondamentalement, la loi est celle d'une égalité; mais en pratique, on peut se restreindre à l'expression des parties qui interviennent effectivement:  $\mathcal{E} \subset \mathcal{E}_0$ ,  $\mathcal{E}' \subset \mathcal{E}_0$  avec :  $\forall x \in \mathcal{E}_0, p_1(x) \neq 0 \Rightarrow x \in \mathcal{E}$  d'une part,  $p_2(x) \neq 0 \Rightarrow x \in \mathcal{E}'$  d'autre part. Le cas d'une inégalité stricte provient donc du fait que certains états initiaux dans  $\mathcal{E}_0$  sont considérés comme a priori impossibles et exclus de  $\mathcal{E}$  à cause des conditions initiales de l'expérience qui ne leur permet pas de se réaliser.

Cet axiome se comprend comme reflet du théorème de Liouville (et il reflète également un phénomène correspondant de physique quantique) de la manière suivante: à  $t_1$ , l'espace des phases est découpé en cellules et l'état classique du système se résume à la cellule dans laquelle il se trouve. Il n'y a pas d'information plus précise, la densité de probabilité dans chaque cellule étant donc "uniforme". Puis on efface le découpage et on laisse le système évoluer suivant un mouvement incompressible de l'espace des phases, jusqu'à l'instant  $t_2$  où on fait un redécoupage en de nouvelles cellules. Ainsi, chaque découpage en cellules a pour effet d'opérer un parfait mixage à l'intérieur de chaque cellule, augmentant son entropie comme nous allons voir.

## 5.2. Loi de Boltzmann

L'état d'équilibre statistique vers lequel tend un système en contact thermique avec un milieu ambiant de température  $T$ , s'exprime par la formule

$$p(x) = e^{-\frac{E(x) - E_0}{kT}}$$

où  $k$  est la constante de Boltzmann servant de correspondance entre les énergies et les températures,  $E(x)$  est l'énergie de l'état  $x$  et  $E_0$  est la quantité qui sert à la normalisation :

$$\sum_x p(x) = 1 \Leftrightarrow e^{-\frac{E_0}{kT}} = \sum_x e^{-\frac{E(x)}{kT}} \Leftrightarrow E_0 = -kT \ln \left( \sum_x e^{-\frac{E(x)}{kT}} \right).$$

En fait, nous verrons que  $E_0$  est l'énergie libre du système dans cet état d'équilibre thermodynamique.

*Dans la suite on fixera la constante de Boltzmann à  $k = 1$  par définition de l'unité de température.*

En voici la démonstration, à partir de notre axiome sous sa forme d'égalité et de l'hypothèse suivant laquelle l'énergie est la seule quantité physique qui se conserve dans le cas d'un système isolé. (Remarque: il existe des situations physiques comportant d'autres paramètres qui se conservent, ce qui aboutit à des lois d'équilibre plus compliquées).

D'abord, montrons que la condition d'équilibre thermique d'un système isolé est que  $p(x)$  doit ne dépendre que de l'énergie  $E(x)$ . En effet, en se restreignant à un niveau d'énergie donné, la plus forte probabilité finale ne peut excéder la plus forte probabilité initiale, et de même la plus faible ne peut pas diminuer. Si elle demeurerait constante, le nombre d'états ayant cette même probabilité diminuerait: il ne peut augmenter et s'il ne diminuait pas, en répétant le même raisonnement sur les états restants menant à un état final d'équilibre  $p$ , alors cette fonction  $p$  serait une quantité qui se conserve, contrairement à l'hypothèse que l'énergie (et donc aussi toute fonction de l'énergie) est la seule quantité qui se conserve.

En tout cas, si les probabilités sur un niveau d'énergie sont égales à l'instant  $t$ , elles le sont nécessairement aussi à l'instant  $t'$ .

Ensuite, la relation entre les probabilités pour des énergies différentes vient concrètement de la possibilité d'échanger cette énergie avec des parties du milieu ambiant. Soit une particule du milieu ambiant pouvant se trouver dans un état  $e$  d'énergie  $E_1$  avec la probabilité  $p_1$  ou dans un état  $e'$  d'énergie  $E_2$  avec la probabilité  $p_2$ , et venant interagir avec notre système  $\mathcal{S}$ . On

suppose que ces probabilités  $p_1$  et  $p_2$  ne dépendent pas de l'état de  $\mathcal{S}$  (on dit que la particule n'est pas corrélée à  $\mathcal{S}$ ; l'hypothèse sous-jacente est que le milieu ambiant est infiniment plus vaste et qu'il peut donc échanger avec  $\mathcal{S}$  une quantité de chaleur quelconque sans modifier ses propriétés générales). Comme on l'a supposé en équilibre thermique avec le milieu ambiant, le système formé de  $\mathcal{S}$  et de la particule doit être en équilibre thermique, et en particulier deux états de même énergie doivent avoir la même probabilité. Donc, quels que soient deux états  $x$  et  $x'$  de  $\mathcal{S}$  tels que  $E(x') - E(x) = E(e') - E(e)$ , les combinaisons d'états  $(x, e')$  et  $(x', e)$  ayant la même énergie  $E(x) + E(e') = E(x') + E(e)$  doivent avoir la même probabilité  $p(x)p_2 = p(x')p_1$ , autrement dit

$$\frac{p(x')}{p(x)} = \frac{p_2}{p_1}.$$

Donc le rapport des probabilités de deux états ne dépend que de la différence d'énergie entre ces deux états et est identique au rapport des probabilités de deux états d'une particule du milieu ambiant ayant la même différence d'énergie. (Au lieu d'une particule on peut prendre un grand système, les conséquences sont les mêmes). La démonstration de la loi de Boltzmann est maintenant terminée.

On peut remarquer la propriété suivante, qui se déduit de notre axiome et le généralise. Soit un système  $\mathcal{S}$  d'état statistique  $p$  sur un espace  $\mathcal{E}$  à l'instant  $t$ . On le met en contact thermique avec un système extérieur  $\mathcal{S}'$  initialement non corrélé à  $\mathcal{S}$  et à l'équilibre de température  $T$  défini par la loi de Boltzmann. On suppose que  $\mathcal{S}$  évolue vers un espace  $\mathcal{E}'$  d'états possibles à l'instant  $t' > t$ , tandis que l'ensemble des états possibles de  $\mathcal{S}'$  reste le même avec la même loi d'énergie, et que le tout est isolé de l'extérieur. Alors on a la formule suivante sur la loi d'évolution  $p(x, y)$  de  $\mathcal{S}$  tandis qu'il interagit avec  $\mathcal{S}'$  :

$$\forall y \in \mathcal{E}', \sum_{x \in \mathcal{E}} p(x, y) e^{\frac{E(y) - E(x)}{T}} \leq 1$$

ou bien la même formule avec égalité, pour les mêmes raisons.

### 5.3. Energie, énergie libre

Nous allons calculer l'énergie libre  $F$  du système dans l'état thermodynamique  $p$ , c'est-à-dire l'énergie qu'on peut en extraire par un travail mécanique (réversible), sans tenir compte des échanges de chaleur avec le milieu ambiant à température  $T$ .

D'abord, définissons l'énergie  $U$  d'un état statistique  $p$  : c'est la moyenne des énergies des états particuliers, suivant leur probabilité :

$$U(p) = \sum_x p(x)E(x).$$

Ensuite, pour calculer l'énergie libre il faut choisir une référence pour laquelle on convient que cette énergie est nulle. Admettons pour le moment qu'elle est nulle dans le cas où le système est en équilibre suivant la loi de Boltzmann avec  $E_0 = 0$  (et nous verrons pourquoi cette supposition est bonne; en attendant, prenons cela comme une définition) :  $p(x) = e^{-\frac{E(x)}{T}}$ . Tout autre état  $p$  peut (très théoriquement !) se ramener à cette situation par le travail "mécanique" suivant.

On commence par isoler  $\mathcal{S}$  thermiquement du milieu ambiant et interdire toute évolution thermique de  $\mathcal{S}$ . Autrement dit, le découpage en cellules de l'espace des phases est transporté par l'évolution en sorte qu'elle s'y exprime par  $p(x, y) = 1$  si  $x = y$  et 0 sinon. Etant donné un état statistique  $p$  fixé (supposé connu, tandis qu'on ne connaît pas l'état particulier  $x$  du système), modifions mécaniquement la fonction d'énergie  $E(x)$  pour la remplacer par la fonction  $E'$  définie par

$$E'(x) = -T \ln p(x).$$

A l'arrivée, l'état  $p$  étant resté le même, son énergie libre s'est annulée, d'après notre hypothèse. Comme cette transformation est adiabatique (i.e. sans échange de chaleur avec l'extérieur) et

réversible, l'énergie libre  $F(p)$  du système au départ est égale à l'énergie retirée sous forme de travail lors de cette transformation :

$$F(p) = U(p) - U'(p) = \sum_x p(x)(E(x) + T \ln p(x)).$$

On pourrait considérer que ce raisonnement n'est pas valable du fait qu'il faudrait travailler avec l'énergie réellement échangée et non son espérance, mais nous allons confirmer plus loin ce résultat par d'autres moyens. Déjà, le fait que dans le cas de l'équilibre de Boltzmann le  $E_0$  coïncide avec l'énergie libre ainsi définie est immédiat (on a tout fait pour).

Ensuite, et c'est là le plus intéressant, voyons comment cette énergie libre se modifie lorsque le système subit une transformation infinitésimale, dans laquelle les fonctions  $E(x)$  et  $p(x)$  se modifient infinitésimalement de  $\delta E(x)$  et  $\delta p(x)$  respectivement.

Là encore par définition, la part d'effet direct pour  $p$  fixé de  $\delta E(x)$  sur l'énergie libre  $F(p)$ , de même que sur l'énergie totale  $U(p)$ , coïncide avec le travail fourni de l'extérieur sur le système,

$$W = \sum_x p(x)\delta E(x).$$

D'autre part, l'effet de  $\delta p$ , qui apporte à l'énergie  $U$  une quantité de chaleur

$$Q = \sum_x E(x)\delta p(x)$$

modifie  $F$  d'une ampleur de

$$\sum_x \delta p(x)(E(x) + T \ln p(x))$$

puisque

$$\sum_x p(x)\delta(\ln p(x)) = \sum_x \delta p(x) = \delta\left(\sum_x p(x)\right) = 0.$$

On remarque alors que la condition pour que les variations de  $F$  au premier ordre ci-dessus soient nulles quel que soit  $\delta p$  (qui vérifie toujours  $\sum_x \delta p(x) = 0$ ) pour une loi d'énergie  $E$  fixée, est que  $E(x) + T \ln p(x)$  soit indépendant de  $x$ , donc que  $p$  soit défini par la loi de Boltzmann de température  $T$ .

Ainsi, l'équilibre statistique coïncide avec l'équilibre de l'état  $p$  dans l'espace des états statistiques, en prenant en guise d'énergie potentielle l'énergie libre  $F$ .

On remarque que cet unique état d'équilibre est précisément celui où l'énergie libre prend sa valeur minimale.

En effet (en cherchant des  $p$  tels que  $F$  soit la plus petite possible), on remarque qu'elle ne prend pas sa valeur minimale en un état statistique qui soit au bord (une valeur limite) de l'ensemble des états statistiques possibles, c'est-à-dire en l'occurrence un état qui s'annule pour un certain  $x$  ( $\exists x, p(x) = 0$ ). En effet, en s'approchant d'un tel bord,  $\ln p(x)$  approche  $-\infty$  (moins l'infini), donc  $\delta p(x)$  donne une contribution d'énergie libre importante de signe contraire  $\delta p(x)(E(x) + T \ln p(x))$  ce qui prouve que  $F$  augmente lorsqu'on s'approche du bord du domaine. Donc  $F$  doit prendre sa valeur minimale pour une position de  $p$  intérieure à son domaine (i.e. un  $p$  qui ne s'annule pas) donc une valeur autour de laquelle ses variations sont nulles au premier ordre, ce qu'on vient d'étudier.

Vérifions maintenant que l'énergie libre est effectivement la mesure du travail qu'on peut retirer réversiblement d'un système à l'équilibre statistique en l'amenant à un état pur  $x_0$  (défini comme état statistique par  $p(x_0) = 1$  et  $p(x) = 0$  pour tout autre  $x$ ) grâce à des échanges thermiques avec le milieu extérieur de température  $T$  : le procédé consiste à amener le système d'un état statistique à l'autre en modifiant la loi d'énergie  $E$  (et donc lui fournissant un travail) de façon progressive pour le laisser à chaque instant rejoindre (par contact avec le milieu extérieur) l'équilibre thermique des lois d'énergie successives. Alors la part thermique de la variation de l'énergie libre devient

d'autant plus négligeable, et l'énergie mécanique fournie se rapproche d'autant plus du montant de cette variation que cette évolution est progressive (la somme des travaux se rapprochant de son espérance du fait que, plus l'évolution est progressive, plus le travail se divise en contributions indépendantes, la contribution de chaque instant étant aléatoire d'après la probabilité  $p$  d'équilibre de l'instant considéré sans corrélation avec les contributions des instants précédents).

Il reste la question : au cours d'une transformation quelconque d'un système, l'énergie libre  $F$  diminue-t-elle ? Comme l'énergie  $U$  se conserve pour un système isolé, la question se ramène à celle de la différence entre les deux.

#### 5.4. Entropie.

La formule de l'entropie  $S$  d'un état statistique  $p$  se déduit de celle de l'énergie libre que nous avons établie, suivant la formule  $F = U - TS$  :

$$S(p) = - \sum_x p(x) \ln p(x)$$

(à multiplier par la constante de Boltzmann que nous avons ici fixée à 1). L'entropie d'un système formé de deux sous-systèmes non corrélés ( $p(x, y) = p_1(x)p_2(y)$ ) est la somme des entropies  $S_1 = S(p_1)$  et  $S_2 = S(p_2)$  de ces sous-systèmes :

$$\begin{aligned} -S(p) &= \sum_{x,y} p_1(x)p_2(y) \ln(p_1(x)p_2(y)) = \sum_{x,y} p_1(x)p_2(y) \ln(p_1(x)) + \sum_{x,y} p_1(x)p_2(y) \ln(p_2(y)) \\ &= -(S_1 + S_2). \end{aligned}$$

Dans le cas de systèmes corrélés, l'entropie du tout  $S = S(p)$  est inférieure à la somme des entropies  $S_1$  et  $S_2$  des états  $p_1(x) = \sum_y p(x, y)$  et  $p_2(y) = \sum_x p(x, y)$ . En effet, en attribuant à ces deux sous-systèmes les énergies respectives  $E_1(x) = -T \ln p_1(x)$  et  $E_2(y) = -T \ln p_2(y)$ , on constate que  $U(p) = U(p')$  où  $p'$  est l'état non corrélé correspondant  $p'(x, y) = p_1(x)p_2(y)$ ; mais  $p'$  étant l'état d'équilibre statistique de cette énergie, il possède la plus grande entropie:  $F(p') < F(p) \Rightarrow S(p) < S(p') = S_1 + S_2$ .

Vérifions maintenant que l'entropie d'un système isolé ne peut qu'augmenter au cours du temps, à partir de notre axiome.

On a pour une évolution de  $p_1$  à  $p_2$  d'un système isolé

$$\begin{aligned} -S_1 &= \sum_x p_1(x) \ln p_1(x) = \sum_{x,y} p(x, y)p_1(x) \ln p_1(x) \\ -S_2 &= \sum_y p_2(y) \ln p_2(y) \end{aligned}$$

Or, pour tout  $y$  on a

$$\sum_x p(x, y)p_1(x) \ln p_1(x) \geq p_2(y) \ln p_2(y)$$

parce que l'application qui à  $u$  associe  $u \ln u$  étant convexe, la moyenne pondérée par  $p(x, y)$  des images des  $p_1(x)$  est supérieure à l'image de leur moyenne  $p_2(y)$ .

Maintenant que la croissance de l'entropie est établie, on en déduit la décroissance de l'énergie libre d'un système  $\mathcal{S}$  lors d'un échange thermique avec un milieu  $\mathcal{S}'$  de température  $T$  ainsi :

L'énergie totale  $U(\mathcal{S}+\mathcal{S}')$  se conserve, et l'entropie du tout  $S(\mathcal{S}+\mathcal{S}')$  augmente, donc  $F(\mathcal{S}+\mathcal{S}')$  diminue. Les systèmes  $\mathcal{S}$  et  $\mathcal{S}'$  étant initialement décorrélés,  $F(\mathcal{S}+\mathcal{S}')$  est égal à  $F(\mathcal{S}) + F(\mathcal{S}')$  au départ mais lui est supérieur à la fin. Comme  $\mathcal{S}'$  est initialement à l'équilibre de température  $T$ ,  $F(\mathcal{S}')$  ne peut qu'augmenter. Donc  $F(\mathcal{S})$  diminue.

## 6. Introduction à la physique quantique

### 6.1. Introduction: une théorie en deux parties

La physique quantique est riche en théories mathématiques intéressantes, en paradoxes aussi. Ses mystères ne sont pas “explicables” par les concepts de mécanique classique ordinaires, ni même par les principes de réalité ordinaires. Certains de ses aspects lui sont spécifiques, et d’autres sont reliés à la physique classique qui leur donne son langage et ses concepts.

On peut considérer la physique quantique comme une conjonction de deux théories aux développements plus ou moins séparés, suivant la divergence de ses deux axiomes fondateurs communément notés  $U$  et  $R$ . Elles sont toutes deux étranges et paradoxales et en manque de fondements clairs, mais pas du tout pour les mêmes raisons. Il faut préciser que, du moins suivant les connaissances actuelles, ces deux théories sont très éloignées l’une de l’autre au niveau fondamental, tandis qu’elles interagissent en permanence au niveau de la physique appliquée.

Les mystères de  $U$  sont le travail exclusif des physiciens mathématiciens, et sont d’une bien trop grande difficulté technique pour que des non-spécialistes puissent seulement en avoir une idée; c’est uniquement sur cette théorie que se portent les recherches en physique des particules. La théorie  $R$  par contre, une fois énoncées ses formules de base, ne laisse que des mystères de nature philosophique (ou métaphysique) sans possibilité d’épreuve expérimentale, et qui n’intéressent généralement plus les physiciens depuis longtemps.

Rare en son genre, Roger Penrose rêve d’unifier  $U$  et  $R$  comme étant deux approximations d’une théorie plus vaste.

#### *La théorie $U$*

C’est celle du monde quantique en lui-même, de ses ondes et de ses particules qui sont une même chose, et dont on n’a actuellement pas fini de sonder les profondeurs mathématiques (sauf approximations), lesquelles sont très lointaines car cette exploration mathématique, tout en étant fructueuse, repousse ses fondements extrêmement loin. On espère en trouver un jour enfin le fond, notamment avec la théorie des supercordes.

Cette théorie est entièrement déterministe, de même que dans une théorie probabiliste les valeurs calculées des probabilités sont considérées comme parfaitement certaines. Le monde de la théorie  $U$  est un monde platonicien, et la notion de réalité factuelle lui est étrangère. Son temps est réversible, en dehors des propriétés de quelques particules exotiques (exception sans rapport intrinsèque avec le sens de variation de l’entropie).

Elle peut se “construire” de façon plus ou moins magique à partir de la physique classique suivant un procédé nommé “quantification”.

Précisons que ce n’est pas là un lien exact et rigoureux, et ne peut pas en être un, bien que l’idée soit là, puisque le monde quantique n’est pas réductible au monde classique. Non que le monde quantique soit différent de ce qui sera décrit ici [partie en projet], mais ce qui sera décrit ici n’est pas vraiment décrit, car cette description sera vague, éludant ses nombreux problèmes ainsi que d’autres notions fondamentales de la physique quantique, complémentaires de ceux-ci (dont les théories de jauge qui décrivent “les propriétés mécaniques particulières” des diverses sortes de particules, cette expression étant à prendre avec beaucoup de guillemets; les pires problèmes de divergence étant dépendants de ces propriétés particulières, les omettre peut simplifier beaucoup la présentation). En un mot, elle sera non-mathématique (normal, ce n’est ici qu’une introduction !). Non qu’elle soit dépourvue de mathématiques, ou que les mathématiques ne puissent y pénétrer, ou qu’elles n’y jouent pas un rôle fondamental (bien au contraire !) mais que sa pleine réduction à une base mathématique serait extraordinairement ardue et généralement incomplète. La formalisation détaillée de la théorie pose, en effet, de graves problèmes. On pourrait exprimer cette situation en disant que la théorie ne repose pas sur un fondement bien solide contenu dans quelques axiomes clairs et qui expliquerait ou impliquerait tout en détail. Mais faute de fondement sur lequel se reposer, on peut se repérer à l’aide du plafond (la mécanique classique qui en est une approximation et dont les objets sont, eux, clairs).

Ici encore, les idées qui seront énoncées ne sont pas vraiment nouvelles par rapport à ce qui est connu, car elles sont sous-jacentes à tout le formalisme de la théorie, même si en général on ne les trouve pas facilement explicitées ainsi dans les cours sur le sujet.

La restriction de l'étude de la physique quantique à des phénomènes non-relativistes (aux vitesses faibles devant celle de la lumière) est le bon moyen couramment employé pour limiter dans une première étape la profondeur de la théorie, ce modèle approché étant entièrement réductible à un système axiomatique sans trop de difficulté. Cette première approche axiomatique claire et maîtrisable de la physique quantique, appelée "mécanique quantique de base" ou "première quantification", est suffisante dans beaucoup de situations.

Mais nous n'allons pas procéder ainsi car notre but ne sera pas de résoudre des problèmes standardisés mais de donner un panorama plus large et qualitatif de la situation, en montrant au maximum le sens et l'élégance des choses, quitte à amener plus de questions que de réponses.

Voici donc le procédé de quantification qui passe de la mécanique classique à la mécanique quantique: il est semblable à la construction de l'équilibre statistique comme fondé sur la mécanique classique, mais appliqué au niveau de la représentation théorique à l'espace de configuration des scénarios non soumis au principe de moindre action dans l'espace pseudo-euclidien. (On prend l'espace de configuration au lieu de l'espace des phases: la différence ne consiste qu'en le fait que l'espace de phase y ajoute la donnée des impulsions, liée au rapport à un temps qui n'existe plus et qui n'a en fait de toute manière pas d'influence sur le reste de l'équilibre statistique). Le rôle de la température est joué par la constante de Planck et l'action correspond finalement à une quantité imaginaire d'énergie: la loi de Boltzmann  $e^{-\frac{E}{kT}}$  devient  $e^{i\frac{A}{\hbar}}$  où  $A$  est l'action (normalement l'action est notée  $S$  mais nous avons déjà trop usé la lettre  $S$ ).

Dans cette analogie, à la fonction probabilité correspond la fonction *amplitude* qui n'est plus une probabilité, puisqu'elle peut prendre des valeurs négatives voire complexes, que sa somme n'est plus contrainte à valoir 1, et qu'elle ne sera pas la probabilité qui interviendra concrètement.

### La théorie $R$

C'est celle qui traite du problème du lien à l'expérience avec sa réalité factuelle, ses mesures, son hasard, son rôle de l'observateur. Ses fondements mathématiques sont parfaitement définis et sans mystère, permettant de calculer les probabilités expérimentales en fonction précise des résultats de la théorie  $U$ . En fait, c'est cette théorie  $R$  elle-même qui est la mécanique statistique quantique que nous avons évoqué. Son écriture en fonction de  $U$  n'est affectée que de façon superficielle par la grande reformulation que traverse  $U$  quand il passe de la mécanique quantique de base à son expression relativiste.

Le problème est qu'elle arrive sur la théorie  $U$  sans y être invitée; elle constitue une rupture par rapport à l'évolution gouvernée par  $U$ , que  $U$  est par nature incapable de produire lui-même (quelle que soit la complexité de l'environnement qu'on pourrait prendre en compte). De plus, aucune règle ne semble pouvoir indiquer quand ni comment elle intervient. Mais elle intervient d'une telle manière que la question des paramètres particuliers de son intervention, n'ayant (jusqu'à présent du moins) aucune influence mesurable sur les phénomènes, ne se pose finalement pas : elle intervient comme si elle n'intervenait pas. Elle est issue de nulle part et elle est introuvable (sans qu'il y ait pour autant la moindre ambiguïté pratique sur les résultats de ses prédictions). De la manière la plus remarquablement subtile (et facilement démontrable), elle est réinterprétable comme fondue à l'intérieur de la théorie  $U$  jusqu'au dernier moment : tant que ce sont les appareils de mesure qui observent, on peut interpréter ces appareils de mesure comme n'ayant pas déclenché  $R$  mais comme étant de simples objets quantiques déterminés par  $U$ , et comme si seuls les humains (ou plus généralement les organismes vivants) avaient le pouvoir de la faire intervenir, ce qui d'ici là résoud le problème de son intervention comme n'ayant pas eu lieu.

D'un autre côté, dès qu'un système un tant soit peu complexe et thermiquement agité comme un organisme vivant ou même un simple appareil de mesure entre dans le coup, son agitation thermique engendre (conformément à  $U$ ) un phénomène statistique de perte de contrôle de certains paramètres du système nommé *décohérence*, tel qu'il ne reste plus aucune chance pour que la différence entre l'hypothèse de la non-intervention de  $R$  ou de son intervention d'une manière correspondant précisément au type d'observation effectuée, puisse encore finalement jouer quelque rôle observable que ce soit. Cette sorte d'intervention virtuelle donne ainsi en pratique le droit d'interpréter  $R$  comme ayant eu lieu.

Mais, alors qu'au niveau microscopique (fondamental) l'intervention de  $R$  passe pour inexis-

tente ou virtuelle, son existence réelle globale est quasiment incontestable : c'est elle qui, une fois pour toutes, fixe à partir de rien l'orientation du temps thermodynamique (celui de la croissance de l'entropie, équivalent au choix de l'ordre de causalité temporel) et qui fait se réaliser un seul des futurs possibles suivant les probabilités données par les calculs lorsque ce que la théorie  $U$  détermine est en fait une combinaison d'avenirs ne se ressemblant pas du point de vue de ce qui est observé subjectivement à l'observateur (par exemple dans la fameuse expérience du chat de Schrödinger, une combinaison (chat mort)+(chat vivant)). (A la rigueur on n'en a pas de preuve, on pourrait supposer qu'ils se réalisent tous conformément à  $U$  dans des univers parallèles qui à cause de la décohérence perdent toute relation entre eux).

Et en fait, la situation est encore pire que cela. Cette théorie  $R$  ne fait pas que laisser parfois un intervalle de temps et de circonstances possibles pour son intervention, entre lesquelles on ignorerait quelle hypothèse d'intervention serait la vraie. Elle va jusqu'à démontrer, si on admet quelques principes physiques difficilement contestables, qu'aucune des hypothèses "raisonnables" de loi générale pouvant déterminer les circonstances particulières de son intervention dans l'intervalle des possibilités, ne peut finalement tenir la route, cet intervalle de possibilités étant finalement réduit à l'ensemble vide d'après ces principes sans qu'on puisse pour autant la prendre au piège de ses contradictions. Voilà ce que nous montrera le paradoxe EPR : non seulement on ne sait pas comment s'effectue l'intervention de  $R$ , mais en plus, alors que les prédictions expérimentales de la théorie demeurent parfaitement cohérentes et sans ambiguïté, le raisonnement montre qu'aucun scénario particulier d'intervention de  $R$  ne peut être considéré comme "le vrai" sans rompre arbitrairement une symétrie du problème qu'aucune expérience ne vient mettre en doute par ailleurs. Aucun, sauf l'impensable interprétation en univers parallèles (suivant laquelle  $R$  n'a jamais lieu et le chat de Schrödinger demeure dans la superposition mort+vivant), toujours en cohérence avec toutes les interprétations géocentriformes particulières.

## 6.2. Le paradoxe EPR

C'est l'un des paradoxes les plus étranges de la physique quantique (théorie  $R$ ), qui, au-delà de la simple dualité onde/particule ou du principe d'incertitude, exprime assez nettement le problème que cette physique pose au réalisme classique. Et il semble que sa signification précise n'ait pas été toujours vulgarisée très parfaitement : on trouve souvent l'affirmation d'une influence instantanée d'une particule sur l'autre, comme si cette influence était réelle. D'où prétexte pour dire que la théorie de la relativité restreinte, qui entraîne l'impossibilité d'une relation de cause à effet propagée plus vite que  $c$ , serait fautive. Or, dans ce phénomène quantique il n'y a rien, ni théoriquement ni expérimentalement, qui permette de situer effectivement dans l'espace-temps quelque influence que ce soit plus rapide que  $c$ . Seul le réalisme classique peut nous y conduire, en déposant là son échec qu'il n'a pas voulu avouer ailleurs.

Et pour mieux cerner sa signification, nous nous concentrerons ici sur l'énoncé exact des prédictions expérimentales de la théorie quantique dans les purs termes des conséquences observables de  $R$  oubliant tous ses liens avec  $U$  par lesquels ces résultats se démontrent habituellement, en sorte de ne pas confondre les mystères de la réalité avec les intermédiaires de calcul. Ces prédictions auraient été vérifiées expérimentalement d'après les nouvelles de ces dernières années. (Mais je ne m'intéresse pas précisément aux détails expérimentaux du problème, et le but de cette page est uniquement d'en présenter ses aspects théoriques.)

### *Quelles particules utiliser ?*

L'expérience a été faite avec des photons corrélés. D'où l'erreur classique de certaines personnes n'ayant pas de connaissances solides en physique, de supposer que ce phénomène de corrélation s'expliquerait par le fait que la durée du voyage d'un photon est pour lui-même toujours nulle, ce qui préserverait un lien de causalité réciproque de ces photons sur leur origine commune. Ceci est faux, la vitesse de la lumière n'ayant rien à faire avec tout cela : le phénomène est en principe indépendant du type de particule utilisé et de leurs vitesses.

A ceux qui font cette erreur, je conseille de s'efforcer de mieux comprendre ce que dit la relativité restreinte à ce sujet : la mesure des durées est une chose, mais l'ordre de causalité temporel en est une tout autre. Entre deux événements distincts reliés par une droite de direction de lumière, bien que l'écart en mesure de temps soit nul, l'ordre temporel n'a lieu que dans un seul

sens. Ce qui peut influencer un évènement se trouve dans son cône de lumière passée (intérieur et frontière), tandis que les points de son cône de lumière future ou de l'hyperplan  $t = 0$  en sont bien loin.

La même expérience peut se concevoir avec des électrons, sauf que dans ce cas elle devient beaucoup plus difficile à réaliser techniquement. Voici pourquoi : la réussite de l'expérience suppose que la particule (ou plus précisément le paramètre qui nous intéresse dans la particule, ici le spin de l'électron) n'a subi aucune interaction thermique avec son environnement, qui ferait perdre en la dispersant dans la nature cette information sur le spin qu'on veut étudier.

Or, l'électron présente un risque d'interaction avec l'environnement beaucoup plus fort que le photon, parce qu'il interagit avec les photons (le champ électromagnétique) tandis que le photon n'interagit qu'avec des charges électriques, or il est beaucoup plus difficile de vider l'espace de ses photons que de le vider de ses charges : le rayonnement thermique, qu'on pourrait d'ailleurs appeler "l'agitation thermique du champ électromagnétique", est en général relativement important, et il a d'autant plus le temps de perturber un des électrons avant qu'ils ne parviennent à la distance voulue l'un de l'autre, que leur vitesse est faible devant celle de la lumière. Donc probablement le seul moyen de faire le vide de rayonnement serait de réfrigérer les parois près du zéro absolu.

Cependant nous décrirons cette expérience dans le cas d'électrons, pour des motifs théoriques :

1) Dans le fond, "à équivalence mathématique près", cela ne change rien : toute expérience (et les paradoxes qui en découlent) avec le spin d'un électron est exactement traduisible en une expérience avec la polarisation d'un photon et inversement. Cette équivalence vient du fait que le nombre d'états quantiques possibles du spin de chaque particule (ce qui désigne la dimension de son espace de Hilbert) est le même, à savoir 2 (il en faut au moins 2 pour qu'il se passe quelque chose, à partir de 3 ce serait plus compliqué).

2) Entre ces deux cas, c'est le cas de l'électron, et uniquement lui, qui manifeste dans la géométrie de l'espace physique la forme des symétries internes du problème quantique considéré, ou si on préfère celle de ses manifestations expérimentales. Ces propriétés géométriques sont remarquables (son rapport à  $U$  s'exprime par la théorie des spineurs), plus que dans la description de la polarisation du photon.

3) Par la même occasion, le récit de cette expérience permettra de se familiariser avec les propriétés du spin de l'électron et la notion de doublet d'électrons qu'on rencontre en chimie.

Nous allons donc maintenant entrer dans quelques manifestations de la théorie  $R$ . N'oublions pas qu'il n'existe aucune séparation qu'on puisse faire entre cette théorie et la mécanique statistique. Les vecteurs que nous emploierons seront uniquement des vecteurs de l'espace euclidien à 3 dimensions, un référentiel galiléen étant fixé une fois pour toutes.

### *Mesure du spin d'un électron*

Prenons d'abord le cas d'un seul électron. C'est une particule de spin  $\frac{1}{2}$ . Que représente son spin ? Dans le langage de la physique classique, c'est son moment cinétique de rotation sur lui-même. C'est bien là l'objet quantique qui se conserve dans  $U$  et qui, dans la correspondance entre physique classique et physique quantique, traduit la notion de moment cinétique qui se conserve.

Or, du point de vue de la physique classique, en choisissant conventionnellement une orientation de l'espace, on peut représenter ce moment cinétique par un vecteur  $\vec{\sigma}$ . Comme un moment cinétique est homogène à une action, c'est donc  $\hbar$  lui servira d'unité de mesure de la norme de  $\vec{\sigma}$  en physique quantique, et la propriété de spin  $\frac{1}{2}$  signifie que "classiquement",  $\|\vec{\sigma}\| = \frac{1}{2}$ .

Mais, du point de vue quantique, il n'y a que deux états possibles du spin d'un électron, pouvant donner un seul bit d'information : ce spin vaut un tel vecteur de norme  $\frac{1}{2}$ , ou bien son opposé. Et pas un autre vecteur. Il est orienté dans un sens, mais il n'a pas de direction. S'il avait une direction, cette direction serait une propriété du spin de l'électron plus riche que cette seule alternative à deux possibilités : ce serait un paramètre porté par l'électron pouvant varier continuellement, en prenant une infinité de valeurs différentes au passage. Mais une telle diversité d'états n'existe pas. (On peut voir avec la téléportation quantique qu'en un certain sens, l'information sur l'état de spin de l'électron peut être classiquement contenue dans deux bits d'information, soit un état parmi quatre possibles.)

Alors, comment se manifeste le sens du spin de l'électron, sans qu'il ait une direction ? C'est l'expérimentateur qui, pour observer ce sens, est libre d'en décider la direction comme il veut. Il choisit une droite, ou ce qui revient au même une paire de vecteurs opposés  $\{\vec{\sigma}, -\vec{\sigma}\}$  de norme  $\frac{1}{2}$ , de manière symétrique : il ne propose pas un vecteur puis son opposé, mais les deux en même temps pour n'en privilégier aucun a priori. Son expérience consiste à demander à l'électron : "Autour de cette droite, dans quel sens tournes-tu ?" (Ou pour le dire plus rigoureusement, il lui demande non le vecteur moment cinétique mais sa composante suivant la direction donnée, sans s'occuper des autres composantes). Alors, nécessairement, sans l'ombre d'une hésitation ni quelque autre humeur que ce soit, celui-ci donnera sa réponse, qui sera l'un des deux vecteurs  $\vec{\sigma}$  et  $-\vec{\sigma}$  (en fait, l'une des deux valeurs  $\pm\frac{1}{2}$  de cette composante). Mais alors, comment cette réponse dépend-elle de la direction choisie ?

Voici ce que dit la théorie dans le cas d'un électron dans un état spatial précis, c'est-à-dire une disposition de la fonction d'onde dans l'espace physique, qui englobe l'information en positions et impulsions de cet électron (ceci afin d'éviter toute question de corrélation entre l'état spatial et l'état de spin):

1) Pour tout électron d'état spatial donné issu d'un processus dont les conditions initiales seraient entièrement connues en un certain sens (à savoir qu'il n'existe aucune influence extérieure qui aurait pu être analysée et qui ne l'a pas été; à ce titre, un état statistique est considéré comme entièrement connu), il existe un vecteur  $\vec{\sigma}$  de norme  $\|\vec{\sigma}\| \leq \frac{1}{2}$  qui résume l'influence mesurable de ces conditions sur le spin de l'électron considéré séparément du contexte de l'expérience, au sens de la règle suivante: quel que soit le vecteur unitaire  $\vec{u}$ , si l'expérimentateur décide alors de mesurer le spin suivant la direction de  $\vec{u}$ , la probabilité d'obtenir  $\frac{\vec{u}}{2}$  sera égale à  $\frac{1}{2} + \vec{\sigma} \cdot \vec{u}$ . Ce vecteur  $\vec{\sigma}$  est alors appelé le *spin de l'électron produit par ce processus*.

On remarque que cet énoncé est cohérent avec la possibilité de décider au hasard du processus qui produira le spin: le processus qui consiste à décider suivant une probabilité  $p$  de faire appel à un processus produisant le spin  $\vec{\sigma}$  et suivant la probabilité  $1-p$  de faire appel à un autre processus qui produit le spin  $\vec{\sigma}'$ , n'est finalement rien d'autre qu'un processus produisant le spin  $p\vec{\sigma} + (1-p)\vec{\sigma}'$ .

2) Quelle qu'ait pu être la situation auparavant, si on mesure le spin d'un électron suivant la paire  $\{\vec{\sigma}, -\vec{\sigma}\}$ , après la mesure le spin de l'électron au sens du 1) sera égal au résultat trouvé par la mesure (cette mesure détruit donc entièrement toute autre information qui a pu exister auparavant dans ce spin).

Ainsi, une fois obtenu le résultat d'une mesure, par exemple  $\frac{1}{2}\vec{u}$ , si on lui pose à nouveau exactement la même question (suivant la même direction  $\vec{u}$ ), il donnera nécessairement la même réponse  $\frac{1}{2}\vec{u}$ . Par contre, si on lui pose ensuite la question de sa direction par rapport à un autre vecteur unitaire  $\vec{v}$ , il répondra par une nouvelle valeur de  $\vec{\sigma} = \pm\frac{1}{2}\vec{v}$  suivant la probabilité égale à  $(\frac{1}{2} + \vec{\sigma} \cdot \vec{u})$ . Si après cela on lui redemande sa direction suivant  $\vec{u}$ , il aura oublié sa première réponse, mais sa nouvelle réponse dépendra uniquement de la dernière qu'il avait donnée, suivant la même loi de probabilité.

### *Préparation de la paire corrélée*

On part d'un doublet d'électrons. C'est un ensemble de deux électrons qui sont dans un même état spatial, mais "dans des états de spin opposés" au sens de ce que nous allons décrire. Pour l'obtenir, il suffit de les contraindre à se mettre dans le même état spatial, car le principe d'exclusion de Pauli les empêche d'être à la fois dans le même état spatial et le même état de spin (ou pour le voir autrement, il implique que deux électrons de même état spatial sont de spin opposés). (En fait, le principe d'exclusion de Pauli et le principe d'incertitude de heisenberg ne se pas des principes mais des conséquences d'autres lois).

Enfin, cela n'est pas parfaitement exact : ils sont bien dans des états de spin opposés mais leurs états spatiaux ne sont pas identiques mais seulement assez proches (c'est-à-dire que les probabilités de résultats de toute mesure de cet état spatial sur les deux électrons sont des probabilités voisines l'une de l'autre et de faible corrélation. . . en les distinguant l'un de l'autre suivant leur spin bien entendu), à cause de la répulsion électrostatique qui empêche cette identité d'être parfaite. Naïvement, on pourrait penser que cette répulsion entre les électrons les empêche de se rencontrer, et rend ainsi superflu le principe d'exclusion de Pauli. Ce n'est pas vrai, à cause d'un phénomène

ondulatoire quantique qui se manifeste par ce qu'on appelle l'*effet tunnel* : alors qu'il faudrait classiquement une énergie infinie pour que des charges ponctuelles de même signe se rencontrent, en mécanique quantique la densité de probabilité du système dans l'espace de configuration d'un doublet ne tend pas vers 0 lorsque la distance  $r$  entre eux tend vers 0.

Pour que le principe d'exclusion de Pauli puisse se mettre en œuvre, il faut le rendre dominant sur la répulsion électrostatique. Dans le cas d'une paire d'électrons seuls mobiles dans un environnement fixe, le seul moyen est de réduire leur étalement spatial  $\Delta x$ . En effet, l'énergie de répulsion électrostatique est proportionnelle à  $\frac{1}{r}$  tandis que l'énergie de répulsion due au principe d'exclusion se lit sur le principe d'incertitude de Heisenberg avec  $r \approx \Delta x$  : l'impulsion  $p$  est de l'ordre de  $\frac{\hbar}{\Delta x}$  et l'énergie cinétique d'une particule d'impulsion  $p$  et de masse  $m$  vaut  $\frac{p^2}{2m}$ , soit une énergie de l'ordre de  $\frac{\hbar^2}{m r^2}$ . Pour des distances assez faibles, il domine donc sur la répulsion électrostatique.

En pratique on n'a pas besoin que le principe d'exclusion devienne infiniment plus grand que la répulsion électrostatique : si les électrons sont confinés dans un espace réduit, la décomposition en états distincts suivant les niveaux d'énergie donne d'une part des doublets (où les spins sont exactement opposés même si les états spatiaux peuvent être très différents), et d'autre part des états où les spins des deux électrons s'ajoutent en un certain sens. Le plus bas niveau d'énergie des doublets est plus bas que celui des paires d'électrons où les spins s'ajoutent; il suffit donc de laisser le système perdre son énergie pour obtenir un doublet. En fait, ce n'est pas exactement vrai à une distance encore plus petite à cause du moment magnétique de l'électron qui vient perturber ce phénomène par l'interaction entre les spins des deux électrons, mais c'est une bonne approximation.

En particulier dans le cas de l'atome d'hélium, le puits de potentiel dû au noyau a cet effet de confiner les électrons dans un même état spatial, créant un doublet.

Maintenant qu'on a le doublet, il faut séparer les électrons. Bon, ici ce n'est pas pratique, mais imaginons qu'on arrive à virer le noyau sans influencer sur le spin (en le bombardant par un neutron de haute énergie par exemple, ou encore en appliquant un fort champ électrique qui envoie les électrons dans un sens et le noyau dans l'autre sens; bien sûr on parle ici de théorie et non de la méthode la plus pratique pour tester EPR).

Voilà donc nos deux électrons du doublet subitement libres de s'éloigner l'un de l'autre. Maintenant, on veut qu'ils se retrouvent dans deux boîtes distinctes qu'on éloignera ensuite l'une de l'autre, un électron dans chaque boîte (les boîtes sont elles aussi très théoriques, pour supposer qu'elles puissent contenir et transporter chacune un électron sans perturber son spin). Comment peut-on faire ? Il ne faut surtout pas trier les électrons d'après leur spin, il n'y aurait plus d'effet EPR. Il faut que jusqu'au dernier moment, aucune mesure ou autre influence extérieure ne vienne interagir avec leur spin.

Il faut les trier d'après leur position spatiale. Pour cela, on peut par exemple disposer une boîte de chaque côté, attendre que les électrons puissent venir dedans, fermer les boîtes et compter s'il y a bien, oui ou non, exactement un électron dans chaque boîte. Sinon, on recommence (avec un autre doublet, ou bien en rouvrant les boîtes pour laisser une seconde chance à notre doublet), jusqu'à ce que cette condition voulue se réalise.

Maintenant, ça y est, notre paire EPR est prête.

### *Mesure des spins d'une paire EPR d'électrons*

Nous avons maintenant deux expérimentateurs avec chacun son électron dans une boîte, les deux électrons étant issus d'un même doublet. Et chacun peut, quand il veut, mesurer le spin de son électron dans la direction qu'il veut.

Quels sont alors les probabilités de résultats ? La théorie énonce à ce sujet les prédictions suivantes :

**Proposition 1.** Le premier qui mesure le spin, quelle que soit la direction choisie, obtient indifféremment l'un ou l'autre résultat suivant la probabilité  $\frac{1}{2}$  chacun, autrement dit son spin est nul avant la mesure au sens que nous avons énoncé. Notons  $\vec{\sigma}$  le résultat de la mesure. Alors ce résultat fait connaître l'autre électron comme étant de spin opposé: si l'autre observateur mesure le spin de son électron par rapport à un autre vecteur  $\vec{v}$ , alors la connaissance du premier résultat permet d'estimer la probabilité d'obtenir  $\frac{1}{2} \vec{v}$  pour le deuxième à la hauteur de  $(\frac{1}{2} - \vec{\sigma} \cdot \vec{v})$ .

**Proposition 2.** Si les deux expérimentateurs mesurent chacun leur spin de manière indépendante, suivant des directions librement choisies au dernier moment par eux (peu importe comment), alors, le choix de leurs directions de mesure étant donné, la probabilité que le couple des réponses soit un certain  $(\vec{\sigma}, \vec{\sigma}')$  est égale à  $(\frac{1}{4} - \vec{\sigma} \cdot \vec{\sigma}')$ .

**Proposition 3.** Comme on le voit, les propositions 1 et 2 ne sont que des formulations équivalentes d'un même résultat, interprété suivant des points de vue différents.

*Consistance du phénomène d'un point de vue quantique.*

D'une manière générale (quels que soient les systèmes considérés avec un plus grand nombre d'états), le résultat de mesures successives ne peut différer du cas de mesures indépendantes que dans la situation où une mesure viendrait affecter l'autre partie du système par une interaction *matérielle* (quelle qu'en soit la forme, champ électromagnétique ou autre). Or, conformément à la relativité restreinte, une telle influence ne se produit d'une mesure sur l'autre que suivant l'*ordre temporel* entre ces deux événements, autrement dit pas plus vite que la lumière. Si l'intervalle entre les mesures est de type espace, ces mesures sont indépendantes.

Chaque observateur peut interpréter le phénomène en disant que les mesures ont lieu dans l'ordre où il les observe, c'est-à-dire l'ordre dans lequel il en reçoit l'information par des signaux lumineux. Donc dans le cas EPR, si les deux mesures sont indépendantes (séparées par un intervalle de type espace), chacun des expérimentateurs peut considérer que c'est lui qui effectue la première mesure qui a une chance sur deux de donner l'un ou l'autre résultat. Ce résultat obtenu, il peut alors considérer (les chances ultérieures de résultats sur) l'autre électron éloigné exactement comme s'il l'avait déjà mesuré suivant la même direction et qu'il en avait obtenu le résultat opposé. En un certain sens, on pourrait dire que c'est le même électron dont la ligne d'univers va d'une mesure à l'autre en passant par le doublet initial. (Ceci ne remet nullement en question l'orientation du temps : les lignes d'univers des particules sont décrites par la théorie  $U$  qui ignore l'ordre temporel en tant qu'ordre de causalité.)

Le lien qui existe dans la loi de probabilité entre les résultats de mesures, n'est pas une interaction mais une *corrélacion quantique*, qui étend la notion classique de corrélation entre deux parties d'un système (que nous avons définie en mécanique statistique: une loi de probabilité  $p(x, y)$  qui n'est pas de la forme  $p_1(x)p_2(y)$ ). En effet, quelle que soit la situation, toute identité algébrique entre probabilités valable dans le cas des corrélacions classiques est également valable dans le cas de corrélacions quantiques, et les seules propriétés qui peuvent être mises en défaut sont des inégalités sur ces probabilités. La corrélation quantique ne fait que prolonger l'espace des corrélacions classiques, utilisant les mêmes paramètres et leur permettant simplement d'aller "un peu plus loin" sur la même lancée. On sait bien qu'une corrélation classique n'a absolument aucune chance de transmettre quelque information que ce soit, ne faisant qu'exprimer un hasard qui s'était déjà produit avant que les sous-systèmes ne soient séparés. En allant un peu plus loin dans la corrélation, on ne fait que multiplier par quelque chose une interaction instantanée à distance qui était nulle, ce qui donne toujours une interaction nulle.

Un système statistique classiquement corrélé devrait avoir une entropie d'au moins  $\ln 2$  pour pouvoir présenter un spin nul à la première mesure, jusqu'à la valeur maximale de  $2 \ln 2$  correspondant au cas d'un système non corrélé. Par contre, l'entropie d'une paire EPR d'électrons reste nulle tant qu'aucune mesure de spins n'est effectuée.

Précisément, d'après la théorie quantique, les paires d'électrons suivantes sont dans le même état concernant le spin pour des états spatiaux donnés "dans deux boîtes" (c'est-à-dire que quel que soit le type d'expérience qu'on puisse concevoir à partir d'une paire d'électrons, les lois de probabilités sur les résultats observables seront toujours identiques d'une paire à l'autre):

- D'une part, une paire d'électrons fournis par un processus ayant secrètement tiré au sort un vecteur  $\vec{u}$  uniformément sur la sphère, puis fourni une paire d'électrons mesurés comme étant de spins respectifs  $\frac{1}{2}\vec{u}$  et  $-\frac{1}{2}\vec{u}$  (l'information sur ce vecteur étant ensuite oubliée).

- D'autre part, une paire d'électrons fournis par un processus ayant secrètement tiré au sort le choix entre donner une paire EPR (issue d'un doublet) suivant la probabilité  $\frac{1}{3}$ , ou donner une paire d'électrons de spins complètement aléatoires (spins nuls et indépendants) suivant la probabilité  $\frac{2}{3}$ .

Plus généralement, des théorèmes similaires peuvent se démontrer pour tout système quantique

corrélé, quels que soient le nombre de parties du système et la forme (nombre d'états quantiques possibles supposé fini) de chaque partie.

### *Échec de l'interprétation classique*

Classiquement, on serait tentés de faire le raisonnement suivant pour deux mesures indépendantes.

Chaque expérimentateur, une fois sa mesure effectuée, peut prédire le résultat de l'autre d'une manière probabiliste, et même certaine si les directions des mesures sont identiques. Et comme aucune des deux mesures ne peut réellement affecter l'autre, il peut considérer que le résultat de l'autre est déterminé d'avance quoi qu'il fasse. Du moins, c'est ce qui vient à l'esprit quand on considère le cas où les mesures se font suivant la même direction. Or, chacun peut choisir sa direction au dernier moment sans que cela ait la moindre conséquence sur (les probabilités de mesure de) l'autre particule. Donc, puisque c'est vrai quand les directions sont les mêmes et qu'un choix de direction n'a aucune influence à distance sur l'autre électron, cela doit être vrai de manière générale.

Ainsi, les chances des résultats qu'on peut obtenir suivant les directions choisies sont déterminées avant sa mesure. On peut alors comprendre les propositions sur les probabilités de mesure en disant que sur l'ensemble des chances initiales du système, une moitié  $M$  sont des chances d'obtenir un vecteur  $\vec{\sigma}$  si on mesure dans cette direction, et d'autre part une moitié  $M'$  également sont des chances d'obtenir pour l'autre électron un vecteur  $\vec{\sigma}'$  si on le mesure dans cette autre direction.

Quand on compare  $M$  à  $M'$ , une proportion de  $(\frac{1}{2} - 2 \vec{\sigma} \cdot \vec{\sigma}')$  de chances issues de  $M$  se retrouvent dans  $M'$  (ce sont les chances dans  $M \cap M'$ ), et les autres qui auraient donné le résultat opposé (celles de  $M - M'$ ) sont remplacées dans  $M'$  par un autre ensemble  $(M' - M)$  de chances issues du complémentaire de  $M$ .

D'autre part, si au lieu de cela on effectue la deuxième mesure suivant une autre direction, le même raisonnement s'applique avec un ensemble  $M''$  de chances d'obtenir un vecteur  $\vec{\sigma}''$ . Et on est alors amenés à penser qu'entre une direction ou une autre pour la deuxième mesure, si on compare l'ensemble  $M'$  des chances qu'on avait d'obtenir  $\vec{\sigma}'$  avec celui  $M''$  de celles d'obtenir  $\vec{\sigma}''$ , la quantité de remplacements de  $M'$  à  $M''$  ne devrait pas dépasser la somme des quantités de remplacements de  $M'$  à  $M$  et de  $M$  à  $M''$ . C'est ce qu'on appelle *l'inégalité de Bell*.

Or, cette conclusion est réfutée par les résultats énoncés plus haut de la physique quantique. On peut le voir par un simple exemple, ou qualitativement de la manière suivante. Entre des mesures des deux électrons suivant deux directions proches formant un angle  $\theta$ , on est presque sûr d'obtenir les résultats opposés, la proportion d'exceptions étant égale à  $\frac{1-\cos\theta}{2}$ , soit une quantité nulle au premier ordre au voisinage de  $\theta = 0$ . Les chances ne varient donc pas au premier ordre d'approximation lorsqu'on bouge  $\vec{v}$  au voisinage de  $\vec{u}$ . Mais comme la position de  $\vec{u}$  ne devrait avoir aucune influence sur la proportion de chances qui sont remplacées pour le résultat de la deuxième mesure, cette annulation au premier ordre de la quantité de remplacements lors d'un déplacement de  $\vec{v}$  devrait être valable pour toute position de  $\vec{v}$ , donc finalement de proche en proche en additionnant les quantités de remplacements successifs, pour toute variation de  $\vec{v}$ , ce qui est absurde puisque le résultat doit s'inverser lorsque  $\vec{v}$  va jusqu'à se retourner vers la position opposée.

Ainsi, le raisonnement classique sur les "causes" des résultats observés mène à des conclusions contradictoires, ce qui prouve que les objets invoqués dans le raisonnement (les ensembles de chances préexistantes qui déterminent les résultats) ne correspondent pas à des objets réels. Ceci vient du fait que parler du résultat qu'on obtiendrait si on effectuait une certaine mesure n'a pas de sens tant que cette mesure n'est pas réellement effectuée. Et pour mesurer l'état des spins de la paire, on n'a que deux essais à disposition : mesurer une fois le spin de l'un, et une fois le spin de l'autre. Il est impossible de produire un triplet d'électrons corrélés de la même manière (cela n'est pas concevable dans le formalisme de la mécanique quantique).

En conclusion, à cause des inégalités qui ne sont plus valables, le hasard quantique de l'observation d'un spin ne peut pas se comprendre suivant l'intuition classique comme étant déterminé par une cause (un "jet de dés") qui existait déjà dans l'environnement *local* de chaque observateur avant qu'il n'effectue sa mesure.