PENSER
AVEC LES
SCIENCES

L'espace physique entre mathématiques et philosophie

Coordonné par Marc LACHIÈZE-REY

EDP
SCIENCES

« Penser avec les sciences »

Collection dirigée par
Michel Paty et Jean-Jacques Szczeciniarz

Ouvrage paru :

Sur la science cosmologique, Jacques Merleau-Ponty
Philosophie, langage, science, Gilles-Gaston Granger

ISBN 2-86883-821-9

TABLE DES MATIÈRES

Introduction

Marc Lachièze-Rey

Cet ouvrage présente les comptes-rendus du colloque qui s'est tenu au Centre d'Études scientifiques de Cargèse du 29 janvier au 9 février 2001. Il est consacré aux réflexions sur l'espace physique, tel qu'il apparaît au sein des diverses théories physiques passées ou présentes : en particulier la physique quantique et la relativité, mais aussi les approches plus modernes et plus spéculatives.

Une longue liste de philosophes, épistémologues et physiciens (Leibniz, Kant, Mach, Poincaré, Einstein...) n'ont cessé de discuter la notion d'espace, et d'en critiquer le statut, sinon la pertinence : réalité ou illusion, objet physique ou entité métaphysique ? Le statut philosophique de l'espace physique reste indéterminé.

Le but de cet ouvrage est de reprendre certaines de ces réflexions, en les actualisant à la lumière des théories physiques modernes, y compris celles encore en gestation. L'intention est ici de poursuivre aussi loin que possible des réflexions de fond sur la notion d'espace, et celles qui lui sont liées. C'est pourquoi les contributions des différents auteurs oscillent constamment entre physique et mathématiques, épistémologie et philosophie.

Le texte de Pascal Nouvel peut ici être vu comme une introduction à l'épistémologie. S'il se présente comme fondateur d'un nouveau type d'épistémologie, on y trouvera une approche de cette discipline elle-même, et une classification des diverses tendances qui sont exprimées lors de ses applications à la physique. Partant de la question : « l'affect est-il le continent oublié de l'épistémologie », Nouvel suggère d'envisager la science du point de vue de celui qui la fait. Il reste à appliquer ce point de vue fructueux à l'épistémologie de l'espace. Les contributions suivantes apportent de substantielles contributions dans cette voie.

Les notions d'*invariance* et de *symétrie* sont intimement liées à celle de géométrie. Les invariances d'un espace géométrique, et/ou d'une théorie sont représentées par un groupe, auquel est associée une algèbre, et ces structures caractérisent de manière essentielle la théorie. Physiquement, la notion d'invariance est liée à celle de reproductibilité des phénomènes (Claude Comte). Et certains aspects se représentent comme des échanges entre référentiels. Bien entendu, la notion d'invariance est également liée à celle de relativité, galiléenne ou relativiste.

La plupart des contributions évoquent les différents outils qui permettent d'aborder la géométrie. En particulier, la théorie des groupes joue un très grand rôle, notamment les groupes de Lorentz, Poincaré et de Sitter. Ils sont longuement évoqués par Jean-Pierre Gazeau, Jacques Renaud et Luciano Boi.

Ces outils sont indispensables à l'expression de la physique actuelle, mais aussi à l'élaboration de nouvelles théories. Parmi elles, les espaces, ou variétés, munis d'un grand nombre de dimensions, jouent des rôles essentiels, malgré les difficultés de représentation et d'approche intuitive. C'est à ce propos que l'article de Jean-Marc Lévy-Leblond se révèle précieux : en présentant des généralisations multi-dimensionnelles de résultats simples (théorème de Pythagore, calculs de surfaces et de volumes...) il permet une appréhension intuitive de quelques propriétés caractéristiques de ces variétés à grand nombre de dimensions. Cette généralisation de résultats classiques d'une géométrie « à la physicienne » fait ressortir le côté très particulier de nos géométries à « petit » nombre de dimensions.

L'espace

La notion d'espace sous-tend toute la physique, et lui est apparemment indispensable. Elle fut introduite formellement par Newton, dans ses *Principia*, après les travaux de nombreux prédécesseurs, dont notamment Descartes. Newton énonce l'existence de l'espace physique, assimilé au seul espace mathématique connu à l'époque, *l'espace euclidien*, ainsi baptisé parce que sa géométrie correspond aux postulats énoncés par le géomètre grec.

Mais la notion fut bouleversée au XIX^{e}, avec la découverte des « espaces non euclidiens », ou *variétés* (non euclidiennes). Dès lors, la question se posait de savoir laquelle de toutes les variétés possibles convenait le mieux pour décrire le Monde physique. Au début du XX^{e} siècle, la géométrie non euclidienne fut incorporée dans la physique par les théories de la relativité : la relativité restreinte énonce qu'il faut remplacer l'espace et le temps séparés par une variété lorentzienne à quatre dimensions, l'espace-temps de Minkowski. Ensuite, la relativité générale énonce que, pour tenir compte de la gravitation, il faut considérer une variété plus générale, toujours à quatre dimensions et lorentzienne, mais avec une *courbure* qui, précisément, représente la gravitation.

Très peu de temps après, apparaît la physique quantique, qui implique de nouveaux types d'espaces, les *espaces de Hilbert*. Évoluant afin de s'accorder avec la relativité restreinte, elle devient la *théorie quantique des champs*. Cette dernière (sous la forme des *théories de jauge*) implique de nouveaux types d'espaces géométriques, les *espaces fibrés*. Une nouvelle géométrie est en jeu, qui concerne ce que l'on qualifiera d'*espace interne* (fibre en terme mathématique) pour le distinguer de l'espace-temps (base du fibré).

Le présent ouvrage a pour ambition de répondre à plusieurs questions :

– Quel sont les espaces géométriques (dans un sens très large, variétés ou autres) les mieux adaptés pour décrire les différentes branches de la physique, y compris dans les nouvelles théories physiques explorées actuellement ?
– Quels sont les accords et les antagonismes entre eux ?
– Est-il possible de les harmoniser, ou de les raccorder, en particulier en ce qui concerne l'antagonisme quantique/relativiste ?
– Quels sont les statuts – ontologique, métaphysique, épistémologique – de ces différentes notions d'espace ?

Les questions de l'ontologie et de la réalité de l'espace, du temps et/ou de l'espace-temps se reposent dans le cadre de chaque nouvelle théorie physique. Déjà, dès le XVII^{e} siècle, les positions de Leibniz s'opposaient à celles de Newton. Un siècle plus tard, l'approche philosophique de Kant s'attaque en profondeur au statut de l'espace. Sa conception constitue le sujet de l'article de Jean-Michel Besnier. Il montre en détail l'argumentation de Kant qui conduit à considérer l'espace et le temps comme des *catégories de la pensée*. Il expose didactiquement les liens entre ces dernières et la sensibilité, les positions de Kant vis-à-vis du réalisme et de l'idéalisme, par rapport à celles de Berkeley, Descartes, Hume, Leibniz, Newton. Restant toujours dans une approche très pédagogique, il termine en évoquant la réception de l'analyse kantienne par Cassirer et Heidegger.

La question de l'actualité du kantisme, notamment vis-à-vis de la nouvelle physique relativiste (sans parler des propositions les plus récentes) reste ouverte. C’est le sujet traité par Jean-Jacques Szczeciniarz : peut-on encore accorder une pertinence aux conceptions kantiennes de l’intuition de l’espace ?

Sous un titre étrange et provocateur, Mario Novello présente une vision originale de l’espace et de la géométrie en relativité générale. Il montre comment, du point de vue de la propagation de la lumière, les propriétés géométriques de l’espace (sa courbure) peuvent être décrites par un *indice de réfraction*, dans une géométrie sans caractéristiques particulières (sans courbure). Cette vision permet de considérer (au moins localement) la relativité générale comme une théorie non particulièrement géométrique. Réciproquement, la propagation de la lumière dans un espace-temps plat, mais dans un environnement particulier, peut être décrite comme suivant les géodésiques d’une certaine métrique, caractérisant un espace-temps courbe (fictif). Novello s’intéresse au statut de ce dernier. Outre un éclairage épistéomologique original sur la notion de géométrie, cela suggère certains procédés expérimentaux qui pourraient permettre de tester des aspects de la relativité générale.

Christiane Vilain s’intéresse à la manière dont est construite la notion d’espace : comment pouvons-nous l’acquérir ? Quel est le rôle de nos *sensations* ? Un des enjeux sera alors de savoir si ces notions correspondent nécessairement à la notion euclidienne, ou si elles s’accordent avec les visions relativiste et quantique (voir aussi la question du rôle des perceptions pour nos conceptions de l’espace, chez Michel Paty). Après avoir rappelé ce qui s’est passé au moment de l’introduction de l’espace en physique (aux XVI^{e} et XVII^{e} siècles), Christiane Vilain s’intéresse aux conceptions de Poincaré : celui-ci montre que notre notion de l’espace se construit à partir de la notion de *déplacement* plutôt que de celle de sensation. Du point de vue psychologique, on pourrait sans doute insister sur le caractère « volontaire » de cette acquisition. Mais c’est surtout le point de vue mathématique qui trouve ainsi sa justification, et en particulier la notion de groupe (de Lie) fondée sur celle de déplacement.

Michel Paty s’intéresse, en épistémologue, au statut de l’espace et des notions géométriques qui l’accompagnent, dans la physique quantique. Il situe d’emblée le problème au sein du « monde quantique ». Cela suppose que ce monde quantique possède bien une existence *réelle*. Cette vision, qui devient la vision moderne de la physique quantique, contredit l’ancienne *intepétation de Copenhague*. Paty nous rappelle que rien ne nous assure que le concept d’espace physique conserve sa validité à l’échelle microscopique. Il y a là un prélude aux idées aujourd’hui développées en gravité quantique.

Paty prend comme point de départ la critique de l'espace physique qui résulte de la physique quantique. Elle est liée au problème de la localisation, également abordée, d'un point de vue plus technique, par Jean-Pierre Gazeau, André Heslot et Serge Reynaud. Paty rappelle et souligne certaines caractéristiques fondamentales de l'approche quantique : impossibilité de localisation, non-séparabilité, et non-localité. Il suggère que, bien que la physique quantique soit effectivement décrite dans le cadre géométrique de l'espace euclidien (ou de l'espace-temps de Minkowski pour la théorie quantique des champs), une autre conception de l'espace la sous-tend de fait. Un paradoxe se cache en effet dans l'approche quantique : bien qu'elle soit formulée dans le cadre géométrique de la physique newtonienne, ou minkowskienne, l'impossibilité de localisation interdit d'exploiter la structure sous-jacente de variété continue, supposée présente. Ce paradoxe fait dire à Paty que les propriétés quantiques empêchent de considérer véritablement l'espace (ou l'espace-temps) comme une *variété* avec ses propriétés habituelles.

La physique quantique rend particulièrement crucial le lien entre espace et *vide*. La difficulté des deux notions remonte déjà au très ancien problème de l'éther, déjà évoqué à l'époque de Newton. Serge Reynaud souligne l'impossibilité d'identifier le vide quantique à un pur néant. Autrement dit, l'espace ne peut être dissocié de l'état des champs quantiques qui, nécessairement, l'occupent. C'est toute la question du *vide quantique*, à laquelle Reynaud consacre sa contribution, en particulier à la question controversée de l'énergie du vide. Il montre que, dès que l'on atteint la précision des mesures quantiques, on ne peut traiter le mouvement et la gravitation d'un point de vue purement géométrique. On doit tenir compte du vide quantique et de tous ses effets. Autrement dit, cela implique que tout ce que l'on attribue à l'espace (ou à l'espace-temps) doit en fait être attribué à la combinaison espace + vide(s) quantique(s). La question reste ouverte de bien connaître tous les effets en question ; et aussi de savoir ce qu'il convient, dans ces conditions, d'appeler exactement « espace » et « vide ».

Bien entendu, la plupart des problèmes intéressants (et difficiles) proviennent des différences de traitement de l'espace, du temps et de l'espace-temps en relativité et en physique quantique. Ceci est lié à la difficile question de la définition des *opérateurs de localisation* en physique quantique, qui constitue le thème principal des contributions de André Heslot et Jean-Pierre Gazeau. Les approches des deux auteurs sont similaires : tenter de définir une possibilité de localisation en physique quantique, qui reste compatible avec la notion de causalité. Cette dernière est naturellement implantée en relativité restreinte, mais qu'en est-il en mécanique quantique ? Les deux auteurs explorent en détails les manières de résoudre les paradoxes liés à cette difficulté.

André Heslot nous présente la construction, formelle, d'une théorie de mécanique quantique, où l'espace-temps n'existe pas au départ : ce dernier apparaît comme le résultat d'une *construction* de la théorie, résultant précisément de l'application d'opérateurs de localisation. Ceux-ci sont élaborés de manière à assurer la covariance (ici, l'invariance de Lorentz). Heslot peut montrer comment les contraintes que la causalité impose à cette construction nécessitent l'existence d'états d'énergie négative (voir aussi à ce propos la contribution de Jacques Renaud). Ceci apporte *a posteriori* une justification à la théorie quantique relativiste de l'électron (de Dirac).

Jean-Pierre Gazeau pose des questions similaires, mais répond dans un autre formalisme, qui rapproche l'approche quantique, exprimée en termes d'*états cohérents*, de la théorie du signal. Il consacre son étude à l'examen de la possibilité de définir des opérateurs de localisation présentant de « bonnes propriétés », notamment, ici encore, la causalité dans l'espace-temps. Au passage, cela fournit une interprétation nouvelle et originale de l'espace, ou de l'espace-temps géométrique, comme « plongé » à l'intérieur d'un espace (de Hilbert) formel d'opérateurs. Gazeau montre les difficultés issues des tentatives pour rendre les opérateurs de localisation compatibles avec la causalité. Il en déduit que la bonne localisation se déroule plutôt dans l'espace des phases.

Jacques Renaud étudie la compatibilité entre quantification et covariance. Si celle-ci ne pose guère de problème pour l'espace-temps de Minkowski, il n'en est pas de même dans un espace-temps courbe. Renaud s'intéresse ici à celui de de Sitter. Il montre d'abord comment la quantification canonique usuelle est incompatible avec la covariance. Il introduit alors une nouvelle méthode de quantification (basée sur celle de Gupta et Bleuler) qui résoud ce problème. Il peut paraître surprenant aux spécialistes de la physique quantique que cette nouvelle procédure ne soit pas fondée sur un espace de Hilbert, mais sur une généralisation de ce dernier. Plus surprenant encore, elle implique des états d'énergie négative. Prenant en compte son avantage essentiel, à savoir de résoudre les problèmes d'énergie infinie rencontrés en théorie des champs usuelle, Jacques Renaud montre que l'on peut tout à fait s'accommoder de ces caractères paradoxaux.

Malgré leur cohérence et leur « élégance », les théories de la relativité ne sont pas indemnes de problèmes quant aux conceptions de l'espace. Luciano

Boi s'intéresse aux conceptions de l'espace et de l'espace-temps, telles qu'elles ressortent des deux théories de la relativité, restreinte et générale, mises en perspective avec celles de la physique newtonienne. Dans la foulée, il souligne également les aspects nouveaux qui semblent jouer un rôle important dans les nouvelles théories de la physique en cours d'élaboration. La question du lien entre espace et matière est examinée avec une attention particulière, dans les cadres de la physique newtonienne, de la relativité restreinte et de la relativité générale.

Boi porte un intérêt particulier à l'antagonisme entre les conceptions de l'espace en relativité et en physique quantique. En historien des sciences, il rappelle que des incompatibilités du même genre se sont déjà présentées dans l'histoire de la physique : leurs résolutions ont mené à de nouvelles théories physiques avec de nouvelles géométries, impliquant de nouvelles formes d'espace. L'intention première de Boi est de souligner l'importance du rôle de la géométrie (de l'espace-temps) dans la physique, et en particulier sa relation avec les lois physiques. Il discute de la position limite qui consiste à déclarer que *tout* dans la physique – interactions et particules – se réduit à de la géométrie (comme, déjà, la gravitation en relativité générale).

Plusieurs articles sont consacrés à la notion de *référentiel* (ou de repère, ou de système de coordonnées), très liée à la conception de l'espace. Les quantités géométriques véritablement intrinsèques sont indépendantes de tout référentiel. C'est ce qu'exprime la notion de *covariance* en relativité. Ainsi, la physique doit rester invariante par rapport au choix d'un référentiel. Pourtant, que ce soit dans les variétés (espace ou espace-temps), ou dans les espaces de Hilbert en physique quantique, et dans toutes les branches de la physique, il est presque toujours indispensable de définir et d'utiliser un référentiel. Outre la commodité des calculs ainsi obtenue, cela permet d'accéder à certaines quantités observables liées à la présence d'un observateur.

La question des référentiels en physique quantique constitue le fondement de l'article de Jean-Pierre Gazeau. Celle de covariance, c'est-à-dire, l'indépendance vis-à-vis du choix d'un référentiel, constitue le point de départ de la nouvelle méthode de quantification proposée par Jacques Renaud. La notion de référentiel est aussi le fondement de l'approche épistémologique de Claude Comte.

Du point de vue relativiste, la question est déjà abordée dans l'article de Luciano Boi. Elle l'est de manière plus détaillée et plus technique dans celui de Marc Lachièze-Rey. Les théories de la relativité ont énoncé que le cadre convenable pour la physique était l'*espace-temps*, plutôt qu'un espace et un temps séparés. Cette exigence s'accorde avec le principe de covariance, fondement de la théorie. Pourtant, on a souvent besoin d'évoquer la notion d'*espace* : soit pour retrouver un langage familier afin d'interpréter tel ou tel résultat relativiste ou cosmologique ; soit lorsque l'on veut incorporer des notions quantiques non locales, ce qui exige que l'on dispose de notions d'espace et de temps séparés. Ceci équivaut au choix d'un référentiel, indispensable donc aux *interprétations* de la physique quantique et de la cosmologie.

Du point de vue mathématique et géométrique, la notion est parfaitement définie. Il existe cependant, dans une variété donnée (espace-temps), des référentiels en nombre illimité. Lachièze-Rey montre comment le choix au sein de cette disparité, équivalent à la définition de l'espace (et du temps) au sein de l'espace-temps, doit nécessairement dépendre de l'observateur : un découpage de l'espace-temps en espace + temps ne saurait être fait de manière covariante et absolue. Mais Lachièze-Rey propose une manière canonique (unique) d'effectuer un tel découpage, *du point de vue d'un observateur*. Ce découpage est accompli grâce à la notion de *synchronisation* ici étendue, de manière parfaitement opérationnelle (et en accord avec les notions déjà énoncées par Einstein), à la relativité générale. Il en résulte une manière unique de définir espace et temps dans l'espace-temps, pour un observateur donné.

Ce découpage peut être utilisé pour définir des variables canoniques à la base d'une quantification de la relativité générale. Sans aller jusqu'à une quantification complète de la gravitation, la *cosmologie quantique* se donne un progamme plus restreint, qui consiste à quantifier une évolution de l'Univers, en demeurant dans le cadre d'une classe de modèles cosmologiques classiques (non quantiques) donnés. Ainsi, il est inutile de définir des variables canoniques pour le champ de gravitation (tenseur de courbure) complet, mais seulement pour les paramètres, en nombre réduit, qui sont pertinents pour la cosmologie. C'est à ce propos qu'Edgard Elbaz propose une méthode originale, concurrente de l'équation de Wheeler–de Witt usuellement invoquée. Son choix différent de variables canoniques conduit à une quantification qui permet de suivre la fonction d'onde du fluide cosmique dans ses différents états (vide quantique, puis radiation, puis matière), avec un raccordement harmonieux.

Qu'il me soit permis de remercier le Centre d'Études Scientifiques de Cargèse, le Commissariat à l'Énergie Atomique, et la Fondation Louis de Broglie, grâce à qui ce colloque a pu se tenir.

À propos de André Heslot

André Heslot, Maître de Conférences à l'Université Paris 7-Denis Diderot (*Laboratoire de Physique Théorique de la Matière Condensée*, Université Paris 7-Denis Diderot), nous a quitté en juin 2000. Esprit libre, il avait engagé son intelligence, incisive, rigoureuse, sur les fondements des théories quantiques. Ce sont là des questions que peu de théoriciens contemporains considèrent comme prioritaires, peut-être parce qu'elles sont simplement hors de leur temps. Ce texte est repris de notes qu'André avait préparées pour un séminaire. Elles datent de 1989 et n'avaient jamais été publiées.

Liste des participants

- **Jean-Michel Besnier**
  Université de Paris IV-Sorbonne, France
- **Luciano Boi**
  École des Hautes Études en Sciences Sociales,
  Centre de Mathématiques, 54, boulevard Raspail, 75006 Paris, France
- **Claude Comte**
  Équipe REHSEIS (UMR 7596), CNRS et Université Paris 7-Denis Diderot,
  Centre Javelot, 75251 Paris Cedex 05, France
- **Edgar Elbaz**
  Université Claude Bernard, Lyon-1, France
- **Sylvain Fautrat**
  Université de Marne-la-Vallée, France
- **Antoine Folacci**
  Université de Corse, Corte, France
- **Tarik Garidi**
  Laboratoire de Physique Théorique de la Matière Condensée
  Boite 7020, Université Paris 7-Denis Diderot,
  75251 Paris Cedex 05, France
- **Jean-Pierre Gazeau**
  Laboratoire de Physique Théorique de la Matière Condensée
  Boite 7020, Université Paris 7-Denis Diderot,
  75251 Paris Cedex 05, France

- **Éric Huguet**
  Laboratoire APC, Université Paris 7-Denis Diderot, France
- **Étienne Klein**
  DSM/DIR, CE Saclay, 91191 Gif-sur-Yvette Cedex, France
- **Joseph Kouneiher**
  CNRS UMR 8102,
  Observatoire de Paris-Meudon, France
- **Marc Lachièze-Rey**
  Service d'Astrophysique, CE Saclay,
  91191 Gif-sur-Yvette Cedex, France
  et laboratoire APC
- **Jean-Marc Lévy-Leblond**
  Université de Nice, France
- **Jean-Paul Longavesne**
  Professeur à l'ENSAD
- **Pascal Nouvel**
  Université Paris 7, France
- **Mario Novello**
  Centro Brasileiro de Pesquisas Físicas,
  Rua Dr. Xavier Sigaud, 150, Urca
  22290-180, Rio de Janeiro, RJ, Brazil
- **Michel Paty**
  Équipe REHSEIS (UMR 7596), CNRS et Université Paris 7-Denis Diderot,
  Centre Javelot, 75251 Paris Cedex 05, France
- **Jacques Renaud**
  LPTMC, Université Paris 7, 75251 Paris Cedex 05, France
- **Serge Reynaud**
  Laboratoire Kastler Brossel
  UPMC case 74, Jussieu, 75252 Paris Cedex 05,
  Laboratoire de Physique Théorique de l'ENS,
  Laboratoire du CNRS de l'École Normale Supérieure et de l'Université Paris-Sud, 24 rue Lhomond, 75231 Paris Cedex 05, France

- **Christophe Salini**
  Équipe REHSEIS (UMR 7596), CNRS et Université Paris 7-Denis Diderot, Centre Javelot, 75251 Paris Cedex 05, France
- **Jean-Jacques Szczeciniarz**
  Professeur de l'Université Paris 7, France
- **Roland Triay**
  Centre de Physique Théorique, Marseille-Luminy, France
- **Christiane Vilain**
  DARC, Observatoire de Paris-Meudon, France

01 Orientations de l'épistémologie contemporaine : vers une épistémologie des affects

Pascal Nouvel

À côté de la très vaste littérature scientifique qui s'est accumulée au cours du XXe siècle, s'en est constituée une autre qui prend la science dans son ensemble ou dans l'une de ses parties comme objet d'études et de réflexions. Moins abondante que la première, elle n'en est pas moins diverse et contrastée. On peut y distinguer deux grands groupes, deux grands styles d'études.

1. Examen général de la littérature sur la science

Le premier groupe est critique. Comme exemple de ce type de discours, on peut citer, bien sûr, les textes de Martin Heidegger sur la science[1] ou encore *L'homme unidimensionnel*[2] d'Herbert Marcuse. La tonalité critique y est perceptible d'emblée et il y est souvent allégué que, pour juger la science sans parti pris, il faut se situer en dehors de la science[3]. Ces contributions, souvent riches de perspectives profondes et d'aperçus métaphysiques ne font pas l'objet de la présente analyse.

Le second type apparaît, par contraste, comme bienveillant à l'égard de la science (quoiqu'il lui arrive de développer des critiques non moins aiguës).

1 – Par exemple : [6] « Ce que nous voudrions savoir, c'est ce que ceux-là [les scientifiques] non seulement ne veulent pas savoir, mais peut-être même sont à jamais incapables de savoir, en dépit de toute leur science et de toute leur habileté artisanale. » (p. 21.)

2 – « Ce que j'essaie de montrer c'est que la science, en vertu de sa propre méthode et de ses propres concepts, a projeté un univers au sein duquel la domination sur la nature est restée liée à la domination sur l'homme et qu'elle l'a aidé à se développer – et ce lien menace d'être fatal à l'univers dans son ensemble. » [15].

3 – Voir par exemple M. Heidegger pour qui science et technique sont généralement confondues : « L'essence de la technique n'est absolument rien de technique. Aussi ne percevrons-nous jamais notre rapport à l'essence de la technique, aussi longtemps que nous nous bornerons à nous représenter la technique et à la pratiquer, à nous en accommoder ou à la fuir. » ([7], p. 9.)

Il se développe à partir d'un intérêt de compréhension et d'élucidation : compréhension des objectifs, élucidation des méthodes de la science. À ce second type appartiennent, par exemple, les textes de Pierre Duhem sur la théorie physique[4], ceux de Bachelard sur le rationalisme[5], et beaucoup d'autres. À s'intéresser de plus près à ces textes, on s'aperçoit qu'ils ne forment pourtant pas un ensemble homogène. On s'aperçoit même que l'on peut de nouveau les diviser en deux grandes catégories.

Dans la première catégorie, les progrès des sciences sont expliqués par des motifs internes à ces sciences, c'est-à-dire essentiellement par des motifs conceptuels, intellectuels, théoriques. *La logique de la découverte scientifique* [19] de Karl Popper, livre publié en 1934 en allemand puis, en traduction anglaise en 1959, représente un bon exemple de ce type d'études. Popper cherche à préciser la nature de la connaissance scientifique en l'opposant à la connaissance non scientifique ; il cherche à établir un critère de « démarcation » (selon son expression) entre ces deux formes de connaissances. C'est à travers ce type d'études que le mot « épistémologie » a acquis les traits distinctifs qui sont aujourd'hui les siens. Ces traits peuvent être résumés de la manière suivante : l'épistémologie est l'examen conceptuel des résultats de la science, elle vise à préciser la nature de la connaissance scientifique. Ce genre d'approche débouche sur une histoire des sciences elle-même théorique et conceptuelle. Si, par exemple, on veut faire l'histoire de la découverte de la structure de l'ADN, on cherchera à préciser les diverses conceptions du support matériel de l'hérédité qui ont pu précéder cette découverte. On fera une histoire des sciences centrée sur les concepts et leurs transformations progressives.

Dans la seconde catégorie, les progrès des sciences sont expliqués par des motifs externes à la science, c'est-à-dire essentiellement par des motifs sociaux, politiques, économiques, culturels. Dans ce contexte, l'histoire de la découverte de l'ADN deviendra une histoire sociale et politique : on cherchera à savoir d'où venaient les crédits qui ont permis aux scientifiques de faire leurs travaux, comment était constitué le laboratoire où ils opéraient, sur la base de quels critères se sont faits leurs recrutements et en vue de quels objectifs, etc. Tout ceci formera une explication « externe » des progrès scientifiques, en ce sens qu'elle fera essentiellement intervenir des facteurs non spécifiques à la science. L'explication peut aller jusqu'à s'affranchir entièrement du contenu

4 – On peut citer, en particulier le grand livre de Pierre Duhem sur l'épistémologie de la physique [2].

5 – Entre 1949 et 1955, Bachelard fait paraître une série d'études sur le rationalisme [1] qui manifeste un intérêt renouvelé pour les question épistémologiques dans les travaux du philosophe.

même des connaissances scientifiques. À ce second type appartiennent les discours de la sociologie des sciences. Un livre comme celui de Steven Shapin et Simon Shaeffer, *Le Léviathan et la pompe à air, Hobbes et Boyle entre science et politique* [20] constitue un exemple de ce type de littérature : les auteurs ne s'intéressent pas au contenu de connaissance mais à la manière dont ces connaissances sont produites, échangées, reconnues, etc. On examine la signification sociale du savoir, non sa signification intrinsèque.

Ce type d'analyse a été inauguré par Robert Merton qui a consacré sa thèse, soutenue en 1938, à l'étude des rapports entre science, technologie et société dans l'Angleterre du XVIII^e siècle. Les études de Merton ont rapidement été qualifiées d'externalistes par opposition aux études conceptuelles qui s'intéressent aux contenus même de la connaissance scientifique. Merton développe la thèse selon laquelle la connaissance scientifique est le produit de facteurs sociaux (externes) et de facteurs cognitifs (internes). Sur ce premier point, rares sont ceux qui le contredisent. Mais il ajoute (et sur ce second point, en revanche, certaines critiques se feront bientôt entendre, comme on va le voir), il ajoute que les facteurs sociaux sont une condition nécessaire à la mise en place des facteurs cognitifs mais qu'ils ne les influencent pas. Selon lui, deux domaines distincts et complémentaires doivent ainsi se constituer : externalisme et internalisme doivent constituer deux types d'approches qui ne se contredisent pas mais, au contraire, se complètent.

C'est avec la publication, en 1962, du livre de Thomas Kuhn intitulé *La structure des révolutions scientifiques* [9] que ce second point sera discuté et critiqué. Ce livre développe l'idée selon laquelle le progrès des sciences peut s'analyser comme une succession de paradigmes entre lesquels aucune mesure de pertinence ne peut être rigoureusement effectuée (les paradigmes sont incommensurables entre eux). Il laisse entendre que les facteurs sociaux, politiques, culturels (bref, les facteurs identifiés comme externes par Merton) jouent un rôle dans la constitution même du savoir scientifique, et non pas seulement un rôle d'accélérateur ou de ralentisseur de son développement. Il fait ainsi apparaître tout le potentiel polémique que contiennent les conceptions externalistes de la science. À partir du moment de sa publication, approches internalistes et externalistes de la science ne cesseront plus de se combattre. En témoigne le livre édité par Imre Lakatos et Alan Musgrave sous le titre *Critisism and the growth of khnowledge* publié en 1970 ; livre qui reprend les actes d'un colloque qui, dès 1965, avait opposé les partisans de la première approche à ceux de la seconde ([10], p. 57).

2. Le concept de vérité dans les versions internalistes et externalistes de la science

L’affrontement s’est focalisé sur quelques grands concepts de la philosophie traditionnelle et tout spécialement sur le concept de *vérité*. Du côté des internalistes, la vérité, conformément à une tradition qui remonte à Platon, est regardée comme *dévoilement*. La science serait ce travail de la pensée par lequel les voiles qui nous masquent la vue du réel sont progressivement levés : « le but [de la science] est de trouver des théories qui, grâce à la discussion critique, nous approchent de la vérité »[6]. Selon les interprètes, les voiles sont plus ou moins opaques, plus ou moins nombreux, pour certains même, ils sont infinis en nombre (Popper est de ceux-là). Mais au-delà de ces différences, ce qui importe est que derrière les voiles se trouve la présence consistante du *réel*. C’est pourquoi ces approches se reconnaissent généralement dans une conception *réaliste* de la vérité.

Du côté des externalistes, la vérité est regardée comme *construction* : elle est socialement construite. Elle est une croyance partagée et n’a aucun privilège épistémologique spécial. En particulier, l’idée qu’elle pourrait se distinguer de l’erreur par un lien plus adéquat au réel est regardée comme une naïveté. La vérité est *relative* au groupe social qui lui accorde foi, voire aux individus singuliers qui leur accordent leur crédit. Ces conceptions se sont donc reconnues comme *relativistes*[7].

Il existe des formes intermédiaires, des réalistes externalistes et des relativistes internalistes. Mais la ligne d’affrontement principale a opposé et oppose encore des réalistes partisans d’une approche internaliste qui soulignent la rigueur des savoirs scientifiques (les opposant souvent à d’autres genres de savoirs) à des relativistes partisans d’une approche externaliste, qui font ressortir le caractère de construction sociale attachée à toute description du monde.

6 – Popper, K., *Normal science and its danger, in* [10], p. 51-58.

7 – Kuhn s’est toujours défendu d’être relativiste, précisant par exemple, dans sa postface à *La structure des révolutions scientifiques* : « Les théories scientifiques de date récente sont meilleures que celles qui les ont précédées [...]. Ce n’est pas là une position relativiste, et elle précise en quel sens je crois fermement au progrès scientifique. » [9], p. 279. Cependant, dès sa publication, son livre sera regardé comme une contribution majeure à l’épistémologie relativiste. Kuhn, d’ailleurs emploie parfois des formules qui se laissent facilement interpréter comme l’expression d’une forme de relativisme : « Celui qui adopte un nouveau paradigme à un stade précoce doit souvent le faire au mépris des preuves fournies par les résolutions de problème. Autant dire qu’il lui faut faire confiance au nouveau paradigme pour résoudre les nombreux et importants problèmes qui sont posés [...]. Une décision de ce genre ne relève que de la foi. » [9], p. 216.

La considération des arguments échangés par les deux tendances qui viennent d'être évoquées n'ont guère varié depuis la publication du livre de Lakatos et Musgrave. Ce fait, à lui seul, doit nous faire conclure que le débat est dans une impasse théorique. Je définis une impasse théorique par le *tourner en rond* des arguments. Lorsque des débats, qui ont pu apparaître, un temps, importants, peut-être même décisifs, tournent en rond sous nos yeux en produisant constamment les mêmes arguments sans que jamais ces arguments ne fassent évoluer la position de ceux qui sont engagés dans le débat, nous tenons là quelques-uns des signes qui permettent de juger que le débat est dans une impasse théorique.

3. Que faire quand un débat est dans une impasse ?

Pour sortir d'une impasse, il faut reculer ; prendre du recul, quitte à perdre de vue provisoirement les problèmes qui ont conduit à adopter ainsi la démarche de l'écrevisse. Il faut aborder la question sous un angle neuf. Où trouver un angle neuf sur la question ? Nous savons déjà que ce ne sera ni du côté internaliste ni du côté externaliste ni dans une composition de ces deux approches. Où alors ? Où trouver un point de vue qui s'affranchisse de cette distinction ? À cette question on peut proposer une réponse pour ainsi dire géométrique, ou mieux, topologique : pour s'affranchir de la distinction entre externalisme et internalisme, l'analyse doit se placer sur la frontière délimitant ces deux domaines. Cette frontière, en quoi consiste-t-elle ? Il n'est pas difficile de le concevoir : la frontière, c'est le savant lui-même.

S'installer sur ce poste d'observation présente des difficultés spécifiques. Il est intéressant de constater que ces difficultés, qui ont parfois été entrevues, ont été utilisées comme un argument pour éviter de s'engager dans cette voie. Lakatos par exemple, dans son livre *Histoire et méthodologie des sciences* [11] explique qu'il serait certainement utile de concevoir ce qu'est la science du point de vue du savant lui-même, mais, croît-il pouvoir affirmer, c'est là une chose impossible à réaliser[8]. Rien ne prouve pourtant que ces difficultés soient insurmontables. Il convient, à tout le moins, d'examiner la possibilité de les surmonter.

8 – [11], p. 148 : « Pour moi, cette démarcation kantienne entre la logique de l'évaluation et la psychologie de la découverte va de soi. »

Résumons l'objectif que nous venons d'identifier : nous disons qu'il faut se placer sur la frontière entre externalisme et internalisme et cette frontière, nous disons que c'est le savant lui-même. En d'autres termes, il faut *voir la science du point de vue de celui qui la fait*. Voilà établie l'orientation générale de notre programme.

4. Mise en œuvre du programme

Maintenant, comment mettre en œuvre un pareil programme ? Comment faire pour voir la science du point de vue de celui qui la fait ? Cette démarche est explicitée en détail dans *L'art d'aimer la science, psychologie de l'esprit scientifique* [16]. On peut la résumer de la manière suivante : voir la science du point de vue de celui qui la fait va inévitablement impliquer un travail d'imagination. Il va falloir « se mettre à la place » du chercheur. Il faut donc trouver les moyens d'orienter correctement cette imagination. Quels seront ces moyens ?

En premier lieu, il importe de réaliser un travail scientifique. Non pas simplement aller dans un laboratoire comme l'ont fait certains sociologues des sciences, mais entreprendre un travail scientifique effectif. On allègue en général les difficultés de la formation. Mais la vraie difficulté se situe bien plus dans le choix de la question que l'on va poser. Or, éprouver cette difficulté est déjà une manière de comprendre la nature de l'activité scientifique. En effet, tout chercheur est confronté à la question du choix de ses propres questions ([16], p. 65). Il doit décider lui-même quelle question est intéressante et quelle autre secondaire : c'est là une composante essentielle de l'activité scientifique, car elle détermine, pour des périodes généralement assez longues, l'orientation des travaux d'un chercheur donné. Le chercheur décide de ses questions et il ne dispose d'aucune méthode pour en évaluer la valeur. La méthode sera requise au moment d'entrer dans la réalisation du travail qui sera effectué en vue de l'obtention d'une réponse à une question donnée, mais non pas dans le choix de la question elle-même. Par conséquent, c'est dans la confiance qu'il a en lui-même que le chercheur puise l'essentiel de l'énergie qui lui permet de déterminer et de préciser la question qu'il fera sienne. Réciproquement, ce sont les sentiments qui le lient à ses propres questions qui rendent la science intéressante pour lui. Confiance dans sa manière d'aborder les questions : c'est là une notion simple mais capitale (qui a cependant été jusqu'ici très peu soulignée par l'épistémologie). Donc, en premier lieu, se confronter à la situation du chercheur et en tirer quelques notions simples sur la nature des problèmes qui se posent à lui.

En second lieu, il convient de disposer d'un matériau empirique pour mettre à l'épreuve les idées qui auront été tirées de cette pratique. On pourra, par exemple, s'appuyer sur certains témoignages ou récits et montrer, à partir d'eux, comment est concrètement réalisé le travail scientifique. Le livre *L'art d'aimer la science* [16] effectue une telle démarche en s'appuyant sur l'ouvrage de James Watson intitulé *La double hélice* [22]. Ce livre raconte la découverte de la structure de la molécule d'ADN, une des plus importantes découvertes du vingtième siècle en biologie. Mais — et c'est là ce qui est intéressant —, il la raconte du point de vue singulier de son principal protagoniste. C'est un compte-rendu *personnel*. À partir de tels témoignages et informé par la sensibilité qu'on acquiert en ayant pratiqué soi-même une science, il devient possible de répondre à l'objectif indiqué : « voir la science du point de vue de celui qui la fait ».

Peut-on, au moyen d'une telle démarche, qualifier et préciser la nature du travail scientifique ? Peut-on, en particulier, la caractériser sous le rapport du rationnel et de l'intuitif ? Sur ce point, la réponse que fournit le livre de Watson ne présente aucune ambiguïté : l'homme de science inventif *est* un homme intuitif. Et on pourrait ajouter d'autres témoignages analogues. Voici par exemple Max Planck : « Lorsque le pionnier de la science lance les antennes tâtonnantes de sa pensée, il doit posséder une vive imagination intuitive, car les idées neuves ne sont pas engendrées par déduction, mais bien par une imagination artistiquement créatrice. » [17]. Voici encore François Jacob : « De toute évidence, le monde de la science, aussi bien que celui de l'art ou de la religion, était un monde créé par l'imagination humaine, mais à l'intérieur de contraintes très strictes imposées à la fois par la nature et par notre cerveau. » [8]. Chacun de ces témoignages mériterait une analyse spécifique. Contentons-nous ici de remarquer que pour celui qui a fait l'expérience d'une découverte scientifique, le mot de « méthode » n'est pas celui qui vient le plus spontanément à l'esprit lorsqu'il se remémore les conditions de sa découverte.

Comme nous l'avons déjà remarqué, le choix de ses questions est laissé aux soins de l'intuition du chercheur, et la persévérance qu'il manifeste dans son travail ne peut être produite que par sa confiance en lui-même, donc par le sentiment qu'il a d'être sur la bonne voie. Plus encore, le fait d'être en mesure d'évaluer les approches concurrentes possibles sur son sujet, de saisir leur pertinence et d'opter néanmoins pour la voie qu'il s'est choisie, tout ceci n'est pas susceptible d'une justification intégralement rationnelle, mais relève bien plutôt d'une évaluation intuitive[9]. Ainsi ce qu'on appelle « méthode scientifique » provient de certains affects qui font que l'homme de science fait ce qu'il fait, qu'il persévère dans son action. Des affects et non une méthode : voilà ce que nous trouvons à la racine de l'activité scientifique. L'activité scientifique produit des concepts, mais elle ne peut entretenir cette production que parce qu'elle s'alimente de certains affects. La formule « L'art d'aimer la science » doit être prise comme un concept qui désigne l'ensemble des affects qui rendent la science possible. Donc, premier bénéfice de ce questionnement, premier résultat : la science repose bien davantage sur des affects que sur une méthode.

5. Provenance de ces affects

Si c'est bien chez ceux qui font la science que l'on peut montrer l'action des affects qui rendent la science possible, inutile, en revanche de s'adresser à eux pour savoir d'où proviennent ces affects. Pour eux, en effet, ces affects apparaissent comme spontanés. C'est pourquoi, lorsqu'ils en parlent, ils évoquent le plus souvent un « besoin fondamental de connaissance », une « passion pour la vérité », ou encore, un « désir insatiable de connaître » : autant d'expressions qui signalent que l'affect est repéré mais aussi que sa source n'est pas clairement identifiée.

Pour identifier une pareille source, il faut examiner la situation d'une civilisation qui n'a pas connu ces affects. Et parmi celle-ci les Grecs sont particulièrement intéressants. Les Grecs, de qui nous avons hérité une part importante de notre culture, connaissaient-ils le genre d'affects qui sont à l'œuvre dans la science ? Ils connaissaient le calcul. Leur géométrie étonne encore de nombreux savants. Leur astronomie était fausse, sans doute, mais elle n'en était pas moins sophistiquée et précise dans ses observations. Geoffrey Lloyd, le grand historien de la science grecque admet que l'affirmation courante selon laquelle la science est née en Grèce n'est pas dénuée de fondement. Car, explique-t-il, la science se caractérise par deux choses essentiellement : en premier lieu le retrait des dieux du domaine de l'explication, en second lieu le recours à la discussion et à l'argumentation pour parvenir à l'explication

9 – On a souvent fait valoir l'importance des controverses en sciences (voir par exemple B. Latour [12]) alors que celles-ci sont relativement rares. Cette orientation de l'attention a conduit à négliger, à contrario, l'importance de la compétion entre groupes de recherche et aussi entre individus d'un même groupe, compétition qui est pourtant la source d'une part considérable des affects qui animent les chercheurs dans leur travail.

correcte [14]. De ces deux principes essentiels de l'explication scientifique on trouve des traces non ambiguës chez les philosophes de Milet, et en particulier chez Thalès alors qu'elles font défaut auparavant et ailleurs à la même époque. On voit par la suite cette attitude se répandre et s'amplifier. Ainsi, par exemple, les textes de Platon manifestent un sens aigu de l'argumentation. Par eux, et par d'autres textes de la même époque, on peut conclure que les Grecs savaient comparer, observer, contrôler un résultat. Tout l'éventail méthodique et conceptuel qui est, de nos jours encore, mis en œuvre dans le développement de la science paraît, dès cette époque, pleinement acquis. Nous devons pourtant constater que les Grecs ne développent pas une science qui, comme la nôtre, transforme le monde dans lequel ils vivent. Cela provient-il de ce qu'il leur manque des concepts, des méthodes, des modèles pour développer une pareille science, ou bien faut-il penser que ce qui leur a manqué est principalement le temps ? Ce point est fortement controversé par les historiens. Certains estiment en effet que la science expérimentale est déjà en marche dans la Grèce antique[10]. L'Europe du XVII^e^ siècle, Galilée, Kepler, Newton n'auraient fait alors que développer une tradition déjà ancienne. D'autres au contraire estiment que quelque chose de radicalement nouveau apparaît avec la naissance des sciences expérimentales et mécaniques en Europe vers l'an 1600, quelque chose que les Grecs n'ont pas connu[11]. D'autres enfin tentent de concilier ces points de vue en élaborant des systèmes explicatifs complexes. Mirko Grmek, par exemple, pense pouvoir identifier cinq niveaux épistémologiques dans le développement de la méthode expérimentale. Les Grecs auraient eu connaissance des deux premiers et auraient pressenti le troisième, mais ils ne seraient pas parvenus à concevoir les deux derniers [5].

Ces contradictions se résolvent sans difficulté dès lors qu'on conçoit que l'objet de la discussion, l'activité scientifique, peut se comprendre comme une composition de deux éléments essentiels : un élément de méthode et un élément de goût. La méthode est connue dès l'Antiquité, mais le goût ne se développe que vers le XVII^e^ siècle en Europe. Si, en effet, au cours d'un voyage ethnologique, nous rencontrons des peuples qui se nourrissent surtout de produits crus et d'un petit nombre seulement de produits cuits et que, par la suite, nous disons d'eux qu'ils ignorent les procédés de la cuisson, nous faisons là un témoignage inexact. Nous déduisons plus justement lorsque nous disons qu'ils ont peu de goût pour les aliments cuits. De même, de ceux chez

10 – Par exemple : Edelstein L. [3], p. 573-604 ; Lejeune A. [13].

11 – Par exemple : [21].

qui la science ne se développe que dans des proportions modestes, comme paraissent l'avoir été les Grecs, nous devons dire qu'ils n'ont eu que peu de goût pour l'élucidation des causes matérielles des phénomènes[12].

Pour proposer ici une opposition grossière mais, dans l'ensemble, utile : la science expérimentale opère une nette distinction entre les choses et les personnes (distinction qu'elle exprime le plus souvent à travers celle, classique, du fait et de l'interprétation : le fait est dans la chose ; l'interprétation est dans la personne), distinction qui n'était sans doute pas aussi radicale pour les Grecs. Les analyses de Michel Foucault ont bien mis en relief cette différence. Ces analyses suggèrent que l'idée d'une distinction entre choses et personnes ne va nullement de soi, et en particulier qu'elle n'était pas courante dans la pensée grecque[13], mais qu'elle est au contraire apparue à l'époque où la science moderne commençait à se développer, vers le début du XVII^e^ siècle[14].

6. Dépassement du dilemme réalisme-relativisme

Les conditions épistémologiques préalables au développement d'une connaissance scientifique reposent dans l'affirmation d'une distinction entre les choses et les personnes, affirmation que nous trouvons exprimée, de manière explicite ou (le plus souvent) implicite, dans tous les domaines de la science. Mais ces conditions ne peuvent s'établir durablement que si des sentiments congruents en stabilisent la présence dans la pensée. La connaissance d'une chose est toujours à la fois affective et cognitive et jamais seulement cognitive. Nous retrouvons ici, affirmée cette fois sur un plan épistémologique, la composition évoquée précédemment de la science comme méthode et affect qui avait permis de résoudre le paradoxe d'une science expérimentale apparaissant dans l'Antiquité mais ne se développant que bien plus tardivement : une part de méthode et une part de goût ; une part de concept et une part d'affect. La connaissance sans affect n'est rien, elle équivaut à l'oubli.

12 – Cette question a été abordée sous une forme un peu différente par Mirko Grmeck.

13 – Foucault, M. [4] : « Pendant toute l'Antiquité [. . .] jamais le thème de la philosophie (comment avoir accès à la vérité ?) et la question de la spiritualité (quelles sont les transformations dans l'être même du sujet qui sont nécessaires pour avoir accès à la vérité ?), jamais ces deux questions n'ont été séparées. » (p. 18.)

14 – [4], p. 19 : « L'âge moderne de l'histoire de la vérité commence à partir du moment où ce qui permet d'accéder au vrai, c'est la connaissance elle-même et elle seule. C'est-à-dire à partir du moment où, sans qu'on lui demande rien d'autre, sans que son être de sujet ait à être altéré pour autant, le philosophe (ou le savant, ou simplement celui qui cherche la vérité) est capable de reconnaître, en lui-même et par ses seuls actes de connaissance, la vérité et peut avoir accès à elle. »

Dans un film de François Truffeau on voit deux personnages évoquer une scène qu'ils ont vécu ensemble quelques mois auparavant. La jeune femme comprend soudain que l'homme ne se rappelle plus de certains détails dont elle a, elle, gardé fidèlement le souvenir. Cette observation, pourtant, ne la conduit pas à douter de la qualité de la mémoire de son compagnon : elle plonge dans une sombre et mélancolique rêverie et nous, spectateurs, nous comprenons très bien le mouvement de sa pensée : du souvenir manquant elle conclut à l'absence de sentiment. Nous savons donc très bien conclure de la connaissance à l'affect dans ce genre de situation. Pourquoi devrions-nous oublier de le faire lorsque nous parlons d'épistémologie ? L'affect est le continent oublié de l'épistémologie.

Nous pouvons maintenant envisager la manière dont le paradoxe entre réalisme et relativisme qui a été précédemment évoqué pourrait trouver sa solution : ce qui est relatif dans la connaissance, ce n'est pas l'objet ni sa description mais l'intérêt pour cet objet. En d'autres termes, une connaissance, en tant qu'elle vise à décrire le réel est susceptible de plus ou moins de vérité (au sens classique de l'adéquation de la pensée à la chose), mais en tant quelle désigne aussi un sentiment, elle est néanmoins relative. Lorsque nous disons par exemple : « voici le matériel génétique des êtres vivants » et nous montrons une double hélice, nous disons une vérité, certes, mais seulement parce que nous supposons acquis le goût pour cette connaissance. Au besoin, on fera remarquer que la chose reste vraie qu'on le croit ou non. Et par cet argument, on compte redoubler la vivacité de l'intérêt pour cette vérité. Cet intérêt, pourtant, demeure un sentiment.

7. Comment constituer une épistémologie des affects ?

L'analyse initiale de la situation de la réflexion sur la science à la fin du XXe siècle nous a conduit à mettre en place le projet d'un « voir la science du point de vue de celui qui la fait ». Ce projet à son tour débouche sur la notion d'une épistémologie des affects qui est à construire. Par épistémologie des affects, il faut entendre le *logos*, le discours, sur la connaissance envisagé dans ses rapports avec les affects qui rendent possibles cette connaissance.

Il faut remarquer que du point de vue de l'épistémologie traditionnelle cette formule – épistémologie des affects – recèle une contradiction dans les termes. L'épistémologie, en effet, s'est constituée précisément en écartant les affects du cercle de ses préoccupations. Non que le mot *épistémologie* implique une notion qui entraînerait la mise à l'écart de l'affect, mais parce que les travaux les plus significatifs qui ont marqué le développement de cette discipline ont établi comme préalable une distinction entre aspect logique et aspect psychologique de la science au gré de laquelle ce second aspect était considéré comme n'appartenant pas au champ propre de l'épistémologie[15]. Une tradition désormais solide a ainsi orienté le développement de l'épistémologie dans un sens qui l'éloigne des affects pour des raisons qu'il serait certainement intéressant d'élucider. Ce n'est pas cependant ici le lieu de développer de telles considérations. Concluons simplement qu'une révision critique de ce sur quoi l'épistémologie traditionnelle a posé ses fondements doit constituer le point de départ d'une orientation de l'épistémologie dans laquelle les affects, les sentiments, les émotions, les passions occupent une juste place.

Bibliographie

[1] Bachelard, G., *L'activité rationaliste de la physique contemporaine*, PUF, Paris, 1951. *Le matérialisme rationnel*, PUF, Paris, 1953. *Le rationalisme appliqué*, PUF, Paris, 1949.

[2] Duhem, P., *La théorie physique. Son objet et sa structure*, Chevalier et Rivière, Paris, 1906.

[3] Edelstein, L., Recent trends in the interpretation of ancient science, *in J. Hist. Ideas*, 13, 1952.

[4] Foucault, M., *L'herméneutique du sujet, cours au collège de France, 1981-1982*, Gallimard/Le Seuil, Paris, 2001.

[5] Grmek, M.D., *Le chaudron de Médée, l'expérimentation sur le vivant dans l'antiquité*, Institut Synthélabo pour le progrès de la connaissance, Paris, 1997.

[6] Heidegger, M., *Qu'est-ce qu'une chose ?* (1962), Trad. fr. J. Reboul et J. Taminiaux, Gallimard, Paris, 1971.

15 – Cette distinction est marquée de la manière la plus nette par Popper [18] « La connaissance au sens objectif est connaissance sans connaisseur ; elle est connaissance sans sujet connaissant. » (p. 184). Popper en fait d'ailleurs un motif de disqualification de l'ensemble de la réflexion sur la connaissance de Locke à Russel, réflexion qui, selon lui, n'aurait pas convenablement marqué la distinction entre connaissance objective et connaissance subjective : « l'épistémologie traditionnelle, celle de Locke, Berkeley, Hume et même de Russel, est hors sujet en un sens assez strict. » Que la locution « épistémologie traditionnelle » soit ici utilisée dans un sens quelque peu anachronique n'empêche pas de situer la pensée de Popper : il s'agit bien de mettre à l'écart la psychologie et, à travers elle, toute réflexion sur les affects qui sont liés à la connaissance.

[7] Heidegger, M., La question de la technique in *Essais et conférences* (1954), Trad. fr. A. Préau, Gallimard, Paris, 1958.

[8] Jacob, F., *La statue intérieure*, Paris, Éditions Odile Jacob, 1987.

[9] Kuhn, T., *La structure des révolutions scientifiques* (1962), Trad. fr. L. Meyer, Flammarion, Paris, 1970.

[10] Lakatos, I. and Musgrave, A. (Eds.), *Criticism and the growth of knowledege, Proceedings of the international colloquium in the philosophie of science*, Cambridge university Press, Cambridge, 1970, p. 57.

[11] Lakatos, I., *Histoire et méthodologie des sciences, programme de recherche et reconstruction rationnelle* (1978), Trad. fr. C. Malamoud et F. Spitz, PUF, Paris, 1994.

[12] Latour, B., *La science en action, introduction à la sociologie des sciences* (1989), Trad. fr. M. Biezunski, Gallimard, Paris, 1995.

[13] Lejeune, A., La science grecque a-t-elle atteint le stade expérimental ? *Rev. Quest. Scient.* (1957) 321–343.

[14] Lloyd, G.E.R., *Une histoire de la science grecque*, Trad. fr. J. Brunschwig, La Découverte, Paris, 1974–1990.

[15] Marcuse, H., *L'homme unidimensionnel* (1964), Trad. fr. M. Wittig, Éditions de minuit, Paris, 1968.

[16] Nouvel, P., *L'art d'aimer la science, psychologie de l'esprit scientifique*, PUF, Paris, 2000.

[17] Planck, M., Signification et limites de la science, in *Autobiographie scientifique* (1941), Trad. fr. A. George, Paris, Flammarion, 1960.

[18] Popper, K.R., *La connaissance objective* (1972-1979), Trad. fr. J.-J. Rosat, Flammarion, Paris, 1991.

[19] Popper, K., *La logique de la découverte scientifique* (1959, version anglaise remaniée de *Logik der Forshung*, paru en 1935), Trad. fr. N. Thyssen-Rutten et Ph. Devaux, Payot, Paris, 1973.

[20] Shapin, S., Shaffer, S., *Le Léviathan et la pompe à air, Hobbes et Boyle entre science et politique* (1985), Trad. fr. T. Piélat, La Découverte, Paris, 1993.

[21] Vernant, J.-P., Remarques sur les formes et les limites de la pensée technique chez les Grecs, *Rev. Hist. Sci.* 10 (1957) 205–225.

[22] Watson, J.D., *The double helix, a personal account of the discovery of the structure of DNA* (1968), A Norton critical edition, W.W. Norton & Co. New-York, London (1980) Trad. fr. H. Joël, Hachette, Paris, 1984.

02 Le statut de l'espace dans la *Critique de la raison pure* de Kant

Jean-Michel Besnier

Parmi les objets de réflexion qui ont contribué à la maturation de la *Critique de la raison pure* (1781), la question de l'espace a été centrale pour Kant. Un petit article, paru en 1768, intitulé *Du premier fondement de la différence des régions de l'espace*[1], en témoigne. Kant y affronte les deux conceptions philosophiques alors en présence : celle de Leibniz qui considère l'espace comme un pur système de relations idéales et celle de Newton qui en fait une réalité distincte des objets qui s'y trouvent contenus. À dire vrai, son choix en faveur des thèses de Newton semble être bien arrêté et il est convaincu que l'espace possède une réalité qui résiste à une approche strictement géométrique. Le paradoxe des figures symétriques lui est l'occasion d'argumenter l'idée selon laquelle l'espace serait doté de « régions » différentes, exigeant par conséquent qu'on le conçoive de manière non homogène et anisotrope.

La position de Newton se résume pour lui à ceci : « Imaginons que la première chose créée ait été une main humaine. Il faut alors que cette main soit nécessairement ou bien une main droite ou bien une main gauche, et pour produire l'une des deux, il fallait un autre acte de la cause créante que celui par lequel a pu être fait son contraire ». Assurément, la création par Dieu de ces objets symétriques non congruents – c'est-à-dire qui peuvent être conceptuellement identiques sans pouvoir être enfermés dans les mêmes limites spatiales – suppose celle d'un espace différencié préalable, jouant le rôle d'espace de référence. « Pour Leibniz, poursuit Kant, cela n'aurait fait aucune différence si Dieu avait créé en premier une main droite plutôt qu'une main gauche. Il faut aller un pas plus loin dans la création du monde pour que se présente une différence. Si Dieu, au lieu de créer d'abord une main droite et

1 – On trouve cet article dans [4], p. 607 sq.

ensuite une main gauche, avait commencé par une main droite, puis créé une *seconde* main droite, ce ne serait pas par son premier acte qu'il aurait changé le plan de l'univers mais bien par le second, en créant une main orientée de façon identique, plutôt que de façon opposée à celle du premier spécimen créé. » C'est dire que l'espace n'a pas de réalité propre, pour Leibniz, et qu'il n'existe qu'avec la relation de symétrie établie entre deux objets congruents (en l'occurrence : deux mains droites, en tant que telles, superposables). Sans ces objets établissant des relations réciproques, l'espace n'existerait pas. Pour Newton, au contraire, c'est parce que l'espace existe – à la manière d'un absolu (« véritable organe de Dieu ») indépendant de la matière et condition de toute composition corporelle – que les objets peuvent établir des relations réciproques. Kant pense d'abord que le paradoxe des figures symétriques suffit à démontrer la supériorité de la thèse de Newton. Par la suite, il considérera que la différence des objets symétriques ne s'explique pas par la seule disposition de leurs parties dans un espace absolu et il en viendra à conclure à leur caractère intuitif. S'imposera alors à lui de rapporter l'espace au sujet, ce qui aura en outre l'avantage de lever les difficultés (les « antinomies ») que fait émerger la conception newtonienne d'un espace objectif [2]. L'Esthétique transcendantale développera en effet la thèse de l'idéalité de l'espace, à la fois contre Leibniz et Newton. Auparavant, Kant aura publié *La Dissertation de 1770* (ainsi qu'on la désigne) pour soutenir la distinction d'un monde sensible et d'un monde intelligible, et contrer Leibniz qui considère le sensible comme de l'intelligible confus. Il aura adressé en 1772 à Marcus Herz une fameuse lettre posant le problème critique dans ces termes : comment le concept peut-il donc synthétiser des représentations sensibles ? Bref, il aura mûri une théorie de la connaissance, dont la *Critique de la raison pure* offrira l'exposé complet et dont il convient de retracer les grandes lignes, si l'on veut comprendre le statut qu'il accorde à l'espace.

Au moment où il rédige son *opus magnum*, Kant a forgé une solution pour résoudre le problème de la représentation formulé dans la lettre à Marcus Herz, solution qu'il a lui-même décrit en la comparant à la révolution copernicienne : on ne saurait se représenter l'objet en général comme une chose-en-soi qu'il s'agirait de faire devenir pour soi ; il faut en revanche partir de la structure cognitive et expliquer comment l'objet s'y trouve déterminé comme « règle de synthèse des représentations ». De sorte que l'objectivité désigne un rapport à la connaissance et non la restitution de l'objet tel qu'il est, un quelque chose qui se laisse quantifié, qualifié et agencé dans des catégories relationnelles, un quelque chose qui se règle donc sur l'esprit.

La *Critique de la raison pure* s'est promise de mettre de l'ordre, de soumettre au « Tribunal de la Raison » les systèmes métaphysiques qui se font la guerre depuis toujours et de les contraindre à plier leurs prétentions à la question *Quid juris ?* La « critique » signifie avant tout l'acte de séparer les instances dont le brouillage est source de conflits, de délimiter les espaces de compétence des facultés émettrices de prétentions spéculatives, bref : d'établir une « topologie » des éléments qui interviennent dans toute connaissance. La construction de la *Critique* offre d'une part une « théorie transcendantale des éléments » – avec l'« esthétique transcendantale », la « logique transcendantale » (analytique des concepts et analytique des principes) et la « dialectique transcendantale ») et d'autre part, une assez brève « théorie transcendantale de la méthode ». Cette construction obéit à une démarche nouvelle dans la tradition philosophique : elle commence par une théorie de la finitude (et non par une définition de l'absolu) pour terminer par une réfutation des systèmes dogmatiques ; elle interroge d'abord les conditions de la vérité scientifique avant de s'engager à déconstruire les erreurs et de remettre à sa juste place la métaphysique.

En quoi consistent donc ces conditions de la vérité ? À découvrir comment un donné quelconque peut s'inscrire sous un concept, à déterminer les règles universelles et nécessaires qui définissent l'objectivité – à savoir : les catégories qui font que ce donné doit avoir une quantité extensive – un nombre –, être qualifiable – posséder un degré de réalité – et se présenter avec une certaine permanence, en interaction, ou comme la cause possible d'un autre donné. Cette approche des conditions de la vérité conduisait Jacques Rivelaygue à commenter en ces termes ce que Kant nommait « l'objet transcendantal = X » : « Il n'est pas une chose, il n'est rien d'autre que les méthodes par lesquelles la science constitue un objet. »

La question nodale de la *Critique de la raison pure* est la suivante : « Comment des jugements synthétiques *a priori* sont-ils possibles ? » C'est qu'en effet, dans tout processus de connaissance, on est amené à énoncer des jugements de trois types : (1) des jugements analytiques – par exemple : « tous les corps sont étendus » – dans lesquels le prédicat ne nous fait pas sortir du concept du sujet ; on est là proche de la tautologie ; (2) des jugements synthétiques – par exemple : « tous les corps sont pesants » – dans lesquels le prédicat en dit davantage que le sujet ; on est là dans une position qui peut exiger une expérience pour accroître la connaissance ; (3) des jugements synthétiques *a priori* dans lesquels le prédicat ajoute quelque chose au concept du sujet, mais sans rien devoir à l'expérience pour cela : ces jugements nous conduisent par exemple à affirmer *a priori* qu'un objet doit avoir une cause et à admettre la nécessité pour les catégories (les concepts purs de l'entendement) de s'appliquer au monde sensible (c'est-à-dire à l'expérience). La question portant sur ces derniers jugements est centrale si l'on veut comprendre comment la physique est possible comme science universelle et nécessaire (c'est-à-dire : non soumise principiellement à l'expérience qui ne saurait jamais fonder une vérité *universelle et nécessaire*). Élucider cette question permettra de mettre en évidence le critère qui différencie la science et la métaphysique : toutes deux reposent en effet sur des jugements synthétiques *a priori*, mais seule la première (la science) satisfait en outre aux exigences d'une « expérience possible ».

C'est pour expliquer la notion d'« expérience possible » que Kant développe une théorie des facultés intervenant dans toute connaissance. Cette théorie distingue (1) la sensibilité (réceptivité des données provenant de la sensation) ; (2) l'entendement (activité des concepts qui classent et informent les données sensorielles) ; et (3) la raison (pouvoir des Idées qui organisent systématiquement les concepts de l'entendement). Cette théorie des facultés cognitives souligne en tout premier lieu la nécessité de l'espace et du temps comme intuitions pures, précédant tout travail de l'entendement et ne pouvant se laisser déduire par aucune expérience. Le temps et l'espace ne sont donc ni des concepts de l'entendement ni des intuitions d'objets offerts à la sensibilité, mais ce qui conditionne l'application ou le remplissement des uns par les autres. Ils ne sont pas représentables mais ils sont au fondement de toute représentation. Ils sont *a priori* et transcendantaux, puisqu'ils rendent la connaissance et la science possibles. C'est ce que démontre l'Esthétique transcendantale.

Pour définir le critère de l'expérience possible (auquel satisfait la science mais dont manque la métaphysique), il faut cependant aller au-delà et établir les conditions *a priori* de l'exercice de l'entendement. C'est-à-dire : examiner les 12 catégories que Kant établit à partir de la table des jugements logiques dont la fonction est précisément de lier concepts et intuitions (= la « déduction métaphysique des catégories ») ; les examiner aussi pour comprendre comment elles peuvent s'appliquer à l'expérience pour permettre de la penser ou, comme dit Kant, d'en « épeler les phénomènes » (= la « déduction transcendantale des catégories »).

Cet examen fait surgir l'exigence d'un troisième terme entre le concept pur (la catégorie) et l'intuition (l'objet sensoriel) – troisième terme destiné à éviter qu'on puisse affirmer avec les métaphysiciens ou bien que le concept crée l'objet (idéalisme subjectif) ou bien que l'objet crée le concept (empirisme dogmatique). L'« imagination transcendantale » joue le rôle de ce troisième terme en garantissant que le concept pur ne crée pas l'objet, mais qu'il détermine seulement la forme de l'objectivité et donc la possibilité de l'expérience (idéalisme transcendantal). L'imagination transcendantale se situe donc comme l'intermédiaire entre l'entendement et la sensibilité. Elle permet aux catégories (formes *a priori* de l'entendement) de s'appliquer au temps (forme *a priori* de la sensibilité), et au temps de se laisser déterminer par les catégories. Exemple : le concept de causalité (catégorie de la relation) s'applique au temps et explique la succession nécessaire d'un événement A et d'un événement B ; la « temporalisation » de la cause justifie l'irréversibilité rendue manifeste par l'expérience. Autre exemple du rôle joué par l'imagination transcendantale : le concept de quantité appliqué au temps explique le nombre et ses propriétés...

Un dernier élément s'ajoute à ceux qu'on vient de passer en revue : le sujet transcendantal destiné à assurer l'unité originaire de l'« aperception ». Ce sujet n'est pas psychologique (comme l'aurait voulu Hume) mais épistémique : il est le garant de l'universalité du fonctionnement cognitif à l'œuvre en tout être humain ; c'est par lui qu'on peut imputer des représentations à une identité stable (quelque chose comme la conscience). C'est cette instance qui, aidée par l'imagination transcendantale, « appréhende » le divers livré par les intuitions, pose quelque chose comme un objet, retient le passé de la sensation éprouvée par le sujet, le reproduit et l'associe au présent. Kant décrit ainsi comme « appréhension », « reproduction » et « recognition » les trois fonctions dévolues au sujet transcendantal dans le processus de connaissance. C'est donc une conception fonctionnaliste qu'il propose : la conscience n'est pas définie en termes de contenu mais comme l'acte de réunir selon des règles.

L'imagination transcendantale et le sujet transcendantal permettent enfin de caractériser cette opération présentée par Kant comme « un art caché dans les profondeurs de l'âme humaine »[2] et qu'il a nommée « le schématisme transcendantal ». Qu'on se borne à dire ici qu'il s'agit de l'acte par lequel les catégories se trouvent temporalisées et autorisent ainsi la construction de l'objectivité. Un exemple fera toucher du doigt la portée théorique du schématisme transcendantal : considéré de son point de vue, un triangle ne s'expliquera pas comme une idée générale qui viendrait, de manière toute

2 – In [5], *Analytique des principes*, p. 226 : Kant ajoute : « dont nous arracherons toujours difficilement les vrais mécanismes à la nature pour les mettre à découvert devant nos yeux ».

platonicienne, s'appliquer dans la réalité sensible mais il sera conçu comme la série des actes temporels nécessaires à sa construction dans l'espace. Par là même, les catégories apparaissent comme des méthodes de construction – comme des « schèmes ». Jacques Rivelaygue l'exprimait ainsi : « la catégorie est la règle de liaison, la façon dont les choses doivent être reliées pour qu'il y ait objet, et le schème est la représentation de l'application de cette règle, de la méthode de liaison, autrement dit la synthèse figurée. La catégorie, comme règle de liaison, est de droit, elle désigne la façon de procéder pour qu'il y ait objet, ce qui ne signifie pas qu'elle soit consciente (pas plus que ce n'est nécessairement le cas de la règle d'un jeu). Au contraire, le schème est la prise de conscience du schéma temporel selon lequel il faut appliquer cette règle pour la rendre effective[3]. »

Le sujet transcendantal mérite donc d'être décrit comme le système des catégories de l'entendement qui s'appliquent grâce au sens interne (le temps), dans le sens externe (l'espace). Il est donc clairement subordonné à la réceptivité (la sensibilité) et, à ce titre, il doit être considéré comme « fini » et passif (il ne produit pas les catégories). Il est conditionné par l'espace et le temps. Il est certes capable de « penser », c'est-à-dire de produire des concepts qui ne s'appliquent à aucune intuition, qui sont inschématisables ; il peut par exemple envisager un système qui regrouperait tous les concepts de l'entendement et négligerait l'exigence de trouver une application dans l'expérience sensible. Mais il est surtout préposé à « connaître », en respectant le critère de « l'expérience possible », en s'en tenant aux phénomènes et en résistant à la tentation de prospecter la chose-en-soi qui est irreprésentable pour la sensibilité à laquelle il reste soumis.

C'est sans doute par là qu'on peut le plus efficacement comprendre que l'idéalisme transcendantal est un réalisme empirique : Kant se situe en effet à la fois contre Berkeley et contre Descartes[4]. Pour le premier, l'espace et le temps étant les propriétés réelles des choses, il en résulte qu'aucune substance matérielle n'est concevable, que seuls les mots par lesquels nous désignons des objets offrent l'illusion d'interrompre le continuum spatio-temporel et qu'une déréalisation des objets est inévitable, dés que l'on y regarde d'un peu près. Kant a évité l'immatérialisme de Berkeley en montrant que l'espace est seulement la forme dans laquelle paraît la matière de l'intuition, que les

3 – J. Rivelaygue [6], cité dans J-M. Besnier, [1], p. 279.

4 – *Cf.* [5], p. 381, *Dialectique transcendantale*, 4^{e} *Paralogisme de la psychologie transcendantale* où Kant réfute à la fois l'idéaliste dogmatique – « Celui qui nie l'existence de la matière » – et l'idéaliste sceptique – « Celui qui la met en doute, parce qu'il la tient pour indémontrable ».

phénomènes sont fondés dans la seule structure universelle et nécessaire du sujet épistémique. Contre Descartes, qui croyait le moi plus certain que le monde, Kant soutient que les phénomènes du sens interne (le temps) sont indissociables de ceux du sens externe (l'espace). Ce qui ne signifie rien d'autre que ceci : le moi a besoin du monde extérieur pour réaliser le point fixe, la permanence et les repères qui le caractérisent. Rivelaygue commentait ce dernier trait en soulignant qu'une psychologie scientifique requiert la base d'une physique – ce qui ne saurait contrarier les efforts pour « naturaliser » l'esprit que déploient aujourd'hui les sciences cognitives.

L'essentiel de l'économie de la théorie kantienne de la connaissance étant ainsi résumé, on en peut à présent dégager les caractéristiques de l'espace et du temps. Kant livre une « exposition métaphysique » en quatre points pour l'espace et cinq points pour le temps, pour montrer qu'ils ne proviennent pas de l'expérience et qu'ils ne sont donc pas des intuitions empiriques ; pour montrer aussi qu'ils ne sont pas non plus des concepts[5].

J'énumère seulement les arguments les plus saillants en les inscrivant sous cinq points :

1. Toute perception suppose que l'on ait le sentiment d'une extériorité ainsi que celui d'une simultanéité : elle exige donc l'espace. De même, elle suppose que l'on ait le sentiment d'une succession d'impressions : elle exige donc le temps. On voit ainsi que la moindre expérience sensorielle implique le temps et l'espace comme conditions préalables.
2. L'espace et le temps sont une représentation nécessaire et *a priori*. La preuve : on peut sans doute *penser* un espace et un temps vides, mais nullement se forger une représentation qui ne soit ni temporelle ni spatiale. C'est pourquoi le temps et l'espace sont bel et bien au fondement de toute représentation et méritent d'être dits *a priori*.
3. L'espace n'est pas un concept mais une intuition pure. Il n'est pas discursif. Il n'y a qu'un seul espace à l'intérieur duquel on peut déterminer des parties – ce qui est le propre d'une intuition, car une représentation qui ne peut être donnée que par un seul objet est une intuition. Le tout y précède les parties, à la différence d'un concept dont l'analyse invite à parcourir un à un les individus particuliers qui le composent.

   De son côté, le temps est assurément *a priori* sans quoi les axiomes qui sont au fondement de la cinématique seraient inconcevables. Il n'y aurait que des sciences expérimentales et aucune ne pourrait jamais

5 – Les considérations qui suivent gagneraient à être étayées par la lecture du passage correspondant de l'Esthétique transcendantale. *Cf.* [5], p. 120 sq.

rien affirmer d'universel et de nécessaire. La fonction transcendantale du temps tient au fait qu'il rend possibles des jugements synthétiques *a priori* (par exemple celui-ci : « des temps différents ne sont pas simultanés mais successifs »).

4. L'espace est une intuition *a priori*. Il contient une infinité de représentations possibles, alors qu'un concept se présente comme le caractère commun construit à partir d'une multitude de représentations données. S'il n'était pas cette intuition pure *a priori*, les jugements synthétiques *a priori* de la géométrie ne seraient pas possibles.

   De son côté, le temps est une totalité non conceptuelle : c'est un tout qui préexiste à ses parties, et non pas, comme le concept, une synthèse de propriétés qui lui préexistent. On ne peut d'ailleurs démontrer conceptuellement les propriétés du temps – celle, par ex., qui veut que deux instants différents ne soient pas simultanés. C'est donc bien une intuition pure *a priori*. S'il ne l'était pas, aucune science ne serait possible. La preuve en peut être donnée de manière apagogique : supposons que le temps soit un concept, alors il doit obéir aux règles de la logique formelle dont le principe de base est la non-contradiction. Or, le temps n'est-il pas justement ce qui permet de penser deux attributs contradictoires dans le même sujet ? Le changement ne se définit-il pas comme la possibilité pour un même corps de demeurer le même en se voyant attribuer telle propriété à l'instant t et telle autre à l'instant t' ? Le temps (comme l'espace) est un *a priori* d'ordre non-conceptuel, c'est-à-dire d'ordre sensible. C'est par là qu'il est exigé (comme l'espace) pour toute représentation.

5. L'espace et le temps sont des grandeurs infinies données. Leurs parties s'obtiennent par limitations, alors que dans un concept, l'ordre impose toujours d'aller des parties au tout.

En conséquence, il convient de retenir que toute représentation est forcément spatio-temporelle et que, sans elle, ni le temps ni l'espace n'auraient de signification. Kant peut donc, dès lors, soutenir la thèse dite de « l'idéalité du temps et de l'espace » :

– idéalité, car nous n'avons pas à faire, avec eux, à une chose ou à un être ;
– transcendantale, car ils sont des *a priori* qui rendent possible la connaissance objective.

Je signale, sans développer, que Kant a été conduit, dans *L'Analytique des principes*[6] à ranger l'espace et le temps sous les figures du néant. Il en parle alors comme des « êtres d'imagination » (*ens imaginarium*). On a l'impression qu'il renoue là avec la tradition métaphysique qui s'emploie à dissiper le temps et l'espace en les caractérisant comme des déficits ontologiques.

En quoi l'espace et le temps apparaissent-ils donc comme des néants ? Réponse : ils ne sont pas des objets, comme on vient de le rappeler, mais ils ne sont que la forme des objets, leur condition formelle. Ils ne sont pas un « quelque chose », ce qui permet à Kant de conclure, dans *L'Analytique des principes*, qu'ils sont « rien », qu'ils voisinent avec le *nihil privativum*, ce néant privatif qui est l'objet vide d'un concept, c'est-à-dire le concept d'une absence d'objet. Qu'on songe à la notion d'obscurité : elle peut bien être dite un concept, à condition d'ajouter que ce concept ne désigne positivement aucun objet, mais seulement une absence de lumière.

Dans *L'Analytique des principes*, donc, à la différence de *L'Esthétique transcendantale* qui les définit comme des « positions » en tant que telles condition de l'« existence » d'objets, Kant range l'espace et le temps du côté de la négativité. Si l'on conjoint les deux approches, on dira que la négativité + la position = la position d'une absence d'objet. En d'autres termes, l'espace et le temps sont une position qui ne pose aucun objet. « C'est la position de l'existence sans l'existence d'un quelque chose qui existerait », écrit Rivelaygue qui évoquait, à l'appui de cette définition, celle proposée par Heidegger de « l'être de l'étant » qui n'est pas lui-même un étant mais bien plutôt un « rien ».

Un célèbre débat entre Heidegger et Cassirer[7] l'a souligné : le kantisme est lourd d'une équivoque, selon qu'on interprète le statut de l'espace et du temps à partir de *L'Esthétique transcendantale* ou bien à partir de *L'Analytique*

6 – Voir notamment *Les anticipations de la perception*, [5], p. 242 sq., *Les analogies de l'expérience*, p. 253 sq. et, finalement, la dernière page de *L'Analytique transcendantale*, p. 328, où Kant propose la « table de la division du concept de *rien* » , après avoir défini l'espace et le temps purs comme *ens imaginarium* et suggéré leur parenté avec le *nihil privativum*.

7 – L'enjeu de ce débat de Davos (mars 1929) entre Heidegger et Cassirer, restitué dans [2], était le suivant : la *Critique de la raison pure* est-elle vraiment une théorie de la connaissance ? N'est-elle pas plutôt la reprise du projet de constituer une ontologie qui résisterait à la métaphysique de la subjectivité d'inspiration cartésienne, c'est-à-dire qui ne céderait pas à l'oubli de la question de l'être au profit de celle des étants ? On sait que le nerf de l'argumentation de Heidegger était de faire valoir que la seconde édition (1787) amenderait la première (1781) dans le sens d'un repli frileux : Kant y élaborerait avec l'imagination transcendantale la thèse d'une « réceptivité originaire » qui interdirait de miser sur l'aptitude des hommes à fonder tout projet rationnel et donc un humanisme. L'imagination transcendantale, cet art mystérieux déjà évoqué en ces termes, limiterait les prétentions de l'entendement en l'obligeant à compter avec une certaine passivité pour recevoir les images particulières auxquels ses concepts peuvent s'appliquer. La finitude serait donc radicale, ce qui engagerait à anticiper une « destruction de la raison », dont Heidegger poursuivrait le programme.

transcendantale : dans un cas, ils se trouvent interprétés dans la perspective d’un idéalisme subjectif : on dira en ce sens qu’ils n’appartiennent pas à l’« étant », qu’ils sont subjectifs – autrement dit : de simples apparences, des êtres d’imagination. Dans l’autre cas, ils sont définis au contraire comme l’être même de l’étant sensible : ce sont des « riens » qui sont à la source de la position de tout étant. Le temps et l’espace semblent voués, chez Kant, à osciller entre l’être (la position) et le néant (« l’être d’imagination »). Heidegger se reconnaîtrait plutôt dans la première lecture, Hegel dans la seconde[8].

Bibliographie

[1] Besnier, J.-M., *Histoire de la philosophie moderne et contemporaine*, tome 2, Le livre de poche, 1998.

[2] Fritsch, V., *La gauche et la droite. Vérités et illusions du miroir*, Flammarion, 1967.

[3] Heidegger, M., Cassirer, E., *Débats sur le kantisme et la philosophie* (1929), éditions Beauchesne, 1972.

[4] Kant, E., *Œuvres*, vol. 1, coll. La Pléiade, Gallimard, 1985.

[5] Kant, E., *Critique de la raison pure* (1781), Trad. française A. Renaut, Aubier, 1997.

[6] Rivelaygue, J., *Leçons de métaphysique allemande*, tome 2, Grasset, 1992.

8 – De son côté, Husserl exigera de refaire *L’Esthétique transcendantale* qui échoue selon lui, chez Kant, à rendre pensable l’unité de la chose perçue. Si le temps et l’espace sont continus, où la chose perçue peut-elle donc trouver son unité ? Comment ne va-t-elle pas éclater dans tous les sens ? Kant a eu besoin de renvoyer à l’entendement pour assurer cette unité, c’est là sa faiblesse qui impose que l’on prolonge l’effort pour constituer une véritable ontologie du sensible.

03

L'espace physique vu du monde quantique : une approche épistémologique

Michel Paty

1. Introduction

Il me faut, en premier lieu, expliciter un tant soit peu ou définir sommairement ce que j'entends par chacune des deux expressions figurant dans le titre de cet exposé : *L'espace physique vu du monde quantique*, quitte à y revenir ensuite avec plus de détails.

Prenons, d'abord, la première expression, qui est la plus familière, l'*espace physique*. J'entends, par espace physique, l'espace de notre expérience du comportement des objets macroscopiques (y compris notre propre corps), qui est aussi l'espace de la physique macroscopique, et qui peut être décrit soit comme espace des sensations et de la perception, soit comme espace géométrique (topologique et métrique), reconstruit intellectuellement à partir du précédent, et indépendant du sujet de la connaissance. C'est sur cet espace de la géométrie que la physique classique puis relativiste a fondé ses élaborations, le prenant pour cadre ou siège des phénomènes physiques, et on le considère généralement, en conséquence, comme étant l'espace physique : nous reviendrons sur ses divers caractères, acquis au cours de ces élaborations.

Considérons maintenant la seconde expression. Parler d'un *monde quantique* (ou d'un *domaine* quantique : celui des phénomènes ou des objets quantiques), cela veut dire accorder aux éléments d'un tel monde qu'ils possèdent une consistance propre, ce qui peut aussi se dire autrement : qu'ils possèdent des propriétés physiques, ou encore qu'ils sont réels. C'est cela qui est nié généralement, notamment par la position qui correspond à l'interprétation philosophique de la mécanique quantique connue comme celle de l'École de Copenhague : selon cette vue, les propriétés des systèmes quantiques sont inéluctablement tributaires des moyens de leur approche par les appareils classiques de mesure, eux-mêmes fondés sur les concepts et les théories de la physique classique[1].

Admettre que l'on puisse légitimement parler d'un monde (de propriétés) quantiques, cela suppose qu'il soit possible de se placer d'un point de vue proprement quantique, par abstraction du sujet perceptif et des moyens et dispositifs qui lui fournissent les éléments de perception sensible à partir desquels il établit une connaissance intellectuelle de ce domaine ou de ce monde. Cela suppose que l'on puisse considérer une description théorique de ce domaine qui soit à la fois physique et autoconsistante ; c'est-à-dire que, d'une part, elle se rapporte (exprimée, bien entendu, au moyen de représentations symboliques) aux phénomènes concernés du monde physique (le monde réel, tel qu'on l'appelle communément) ; et que, d'autre part, l'on n'ait pas besoin de recourir pour elle à un référent externe tel que celui de l'observation ou la perception, qui sont anthropocentriques. Il faut donc supposer que l'on dispose d'une théorie propre ou intrinsèque du domaine quantique, et sortir, évidemment, du « cercle » de l'interprétation observationaliste ou « complémentariste », en adoptant par principe une perspective réaliste sur le monde physique et sur la théorie physique comme connaissance effective de ce domaine de la nature. Nous reviendrons également sur cette question préliminaire et fondamentale.

Le problème que je voudrais tenter d'éclairer du point de vue épistémologique est celui du rapport entre ces deux notions, en les supposant fondées : *espace physique, monde quantique*. En particulier, nous savons que les phénomènes et les systèmes quantiques impliquent une critique de l'espace tel que la physique classique et relativiste le considère, par exemple, l'espace lié au temps de la théorie de la relativité (restreinte ou générale). Dans ces théories, les systèmes physiques sont toujours considérés comme sous-tendus par la notion de points matériels individuels localisés dans cet espace-temps, liés entre eux par une relation de causalité spatio-temporelle : en relativité restreinte, deux points physiquement, c'est-à-dire causalement, reliés ou susceptibles de l'être, appartiennent à une même région de « type temps » du cône de lumière, telle que $c^2t^2 - r^2 \leq 0$. Les systèmes « classiques » (dans ce sens, c'est-à-dire non quantiques) *étendus* sont tels que leurs structure ponctuelle élémentaire

1 – Bohr [8].

est toute entière située dans cette région de type temps (par exemple le front d'une onde sonore, le front d'une onde lumineuse[2], ou un bâton rigide).

La critique de l'espace physique par la physique quantique est portée par les relations d'indétermination (distribution spectrale des variables conjuguées d'espace et d'impulsion, $\Delta x . \Delta p > \hbar$), par la dualité onde – corpuscule (diffraction, interférences, qui impliquent une extension dans l'espace des « objets physiques » considérés), par la non-localité des systèmes quantiques, par la non-séparabilité locale de systèmes ayant interagi dans le passé, par l'indiscernabilité des « particules » identiques (fermions ou bosons) composant un système quantique. Qui dit ici *critique de l'espace*, dit en réalité critique du *concept d'objet-dans-l'espace*[3]. Nous nous interrogerons, en particulier, sur la réponse traditionnelle de l'interprétation orthodoxe de la mécanique quantique en termes d'approche statistique, et sur la signification à cet égard de la connaissance, relativement récente, de systèmes quantiques individuels.

Que devient donc le concept d'espace de la géométrie et de la physique des corps matériels et des champs classiques avec la physique quantique ? En demeure-t-il quelque chose après cette critique qui revient à limiter ses conditions de validité, et peut-être même sa signification ? On doit se rappeler que notre concept d'espace physique nous vient de notre connaissance des propriétés des objets macroscopiques (non quantiques) et que rien ne nous assure qu'il soit encore pertinent dans le domaine atomique et infra-atomique[4] ; mais que, d'un autre côté, l'élaboration et la formulation de la physique quantique continue de faire appel à ce concept, non seulement comme concept générateur (voir la façon dont les opérateurs quantiques pour représenter les variables dynamiques sont construits dans la théorie, en particulier celui de position)[5], mais aussi dans son sens le plus classique.

À cet égard, on doit accorder une attention privilégiée à la réflexion sur la notion de propagation d'un système quantique dans l'espace, en rapport au principe de superposition, à la question de la dispersion ou de l'étalement

2 – Les points d'un front d'onde lumineuse sont reliés causalement à leur commune origine, et donc sont situés sur le cône de lumière de l'espace-temps.

3 – Einstein lui-même reconnaissait que la physique quantique « ne se propose aucunement de donner une représentation mathématique en termes d'espace et de temps » [22]. Voir Paty [57], chap. 5, p. 188-193 ; [76] chap. 6.

4 – Un collaborateur de l'*Encyclopédie* de Diderot et d'Alembert, M. Guénault, le rappelait au XVIII[e] siècle [36] ; *cf.* Paty [66]). Une telle idée a été également émise à propos des phénomènes quantiques, entre autres par Paul Langevin [43] ; Einstein lui-même en a évoqué la possibilité (voir notamment : [18, 21, 22]).

5 – Dirac [17]. *Cf.* Paty [55].

(comme pour une onde), ainsi qu'à celle de la cohérence d'un système quantique isolé dans l'espace. On se demandera aussi comment se caractérise ou s'engendre l'espace occupé par un système purement quantique de relativement grandes dimensions tel qu'un condensat de Bose-Einstein. Nous ne prétendons pas apporter une solution à de telles questions, mais seulement les approcher sous l'éclairage épistémologique, qui est celui de leur signification possible.

2. Penser l'espace et penser dans l'espace

2.1. La physique tributaire des notions d'espace et de géométrie

La pensée physique – et, jusqu'à un certain point, la pensée mathématique, du moins géométrique – s'est toujours appuyée sur la notion d'espace. Celle-ci a toujours été en même temps au centre de la réflexion philosophique, de la théorie aristotélicienne des lieux naturels à l'espace homogène et isotrope de la géométrie et de la physique de Galilée, Descartes, Huygens, Newton, Leibniz, et jusqu'à la physique contemporaine du continuum spatio-temporel. Songeons, pour nous en tenir à la période ouverte par la révolution scientifique du XVII^e siècle, à la conception des idées innées de Descartes qui accompagna sa *Géométrie*, à sa critique par Locke, pour qui l'origine des idées se trouve dans les sensations, à l'esthétique transcendantale de Kant (l'espace et le temps comme formes intuitives *a priori* de la sensibilité), aux analyses épistémologiques faites au XIX^e siècle par Hermann von Helmholtz, Ernst Mach, William James, Henri Poincaré, fondées en partie sur les études psycho-physiologiques, et enfin à celles, au XX^e, que l'on doit à Albert Einstein, Hermann Weyl, Ernst Cassirer, Rudolf Carnap, Hans Reichenbach et d'autres, sans oublier les travaux des psychologues et des neurophysiologues, de Jean Piaget à Jean-Pierre Changeux, Alain Berthoz...

Nous en retiendrons ici que la notion d'espace s'est constituée dans un premier temps comme « espace perceptif » et « représentatif » à partir des données de la perception, qui tiennent à notre constitution physiologique et neuropsychologique mettant en jeu les différents sens (du toucher, de la vision, de l'ouie, ...) ; cet « espace représentatif », obtenu par la coordination des différents espaces perceptifs, n'est ni homogène et isotrope ni même à trois dimensions, si ces dimensions sont à rapporter aux différents paramètres qui font la perception[6]. On peut cependant considérer, avec Mach, que ces dimensions sont d'origine à la fois physique et physiologique, résultant de l'organisation spatiale des organismes vivants supérieurs eux-mêmes, suivant les orientations naturelles dues à la pesanteur, à la locomotion (haut-bas, avant-arrière), et à une asymétrie droite-gauche [51, 52]. Pour Poincaré, ces trois dimensions sont plutôt caractérisées au niveau de la coordination des espaces sensoriels, à travers une analyse de la constitution de la notion de continu, qui fonde l'« Analysis situ », étude de l'espace (géométrique) en tant qu'il est continu et à trois dimensions[7].

Quant à l'*espace géométrique*, qui s'est constitué par les opérations de l'entendement à partir de l'espace représentatif, il peut être considéré soit comme *espace métrique*, soit comme *espace topologique*[8]. C'est à partir de lui que sont définies les grandeurs et les figures de la géométrie (avec, en premier lieu pour les grandeurs, les distances spatiales), et c'est à lui qu'est identifié, de Galilée, Descartes, Newton, à Riemann, Maxwell, Helmholtz, Poincaré (pour ne citer qu'eux), l'« espace physique ». Après la reconstruction axiomatique de la *Géométrie* de Hilbert [39], une dissociation s'est opérée entre la géométrie purement mathématique, et la géométrie de l'espace physique : Einstein est un bon témoin de cette nouvelle perspective, en définissant, dans ses deux théories de la relativité, l'espace physique par des grandeurs géométriques (coordonnées et distances) prises sur des corps physiques, moyennant des relations de correspondance ou de coordination, qui reviennent à la construction d'une « géométrie physique » élémentaire par celle d'un « espace physique de référence ». Cette notion lui a été indispensable pour penser la cinématique de la relativité restreinte indépendamment de la dynamique, et pour s'affranchir, avec la relativité générale, des limitations de l'espace euclidien[9].

Il faudrait compléter ces remarques en évoquant les enseignements de l'histoire des sciences qui montrent comment le concept d'espace de la physique et de la géométrie s'est progressivement élaboré, en rupture avec les conceptions communes et par construction intellectuelle, mettant en œuvre les ressources de la pensée rationnelle, mathématique et physique : des lieux aristotéliciens qualitatifs à l'espace continu homogène, isotrope et infini de

6 – Tels que les divers mouvements musculaires dirigeant le toucher ou l'orientation du globe oculaire pour la vision : voir Poincaré [79], chap. 3, 4 et 5.

7 – Ce terme, « Analysis situ », inventé par Leibniz, fut repris par Riemann pour désigner la topologie, qu'il fonda. Voir Poincaré [80], chap. 4 ; [82], chap. 3.

8 – Voir Paty [60], chap. 6 et 7 ; [66].

9 – Voir les travaux désormais classiques d'Einstein sur la relativité (Einstein [26], vols. 2 à 8). Voir aussi Weyl [90], Schrödinger [88] ; *cf.* Paty [60], chap. 3, 4, 5

Bruno, Galilée, Descartes, Gassendi, et à la pensée infinitésimale et différentielle du continu spatial, de Newton, Leibniz, Euler, d'Alembert, Lagrange, à Riemann, Ricci, Levi-Civita. L'espace, en physique comme en géométrie, est, depuis lors, un concept construit mathématiquement à l'aide de grandeurs appropriées à sa structure et à ses propriétés, qui concernent les entités physiques définies dans l'espace (*cf.* Paty [66]).

L'idée de *grandeur mathématique* prend son point de départ dans la définition de la ligne droite, avec ses points, mais aussi, après Newton et Leibniz, avec ses éléments différentiels. Dans les *Règles pour la direction de l'esprit* (env. 1628), en particulier à la règle 14, Descartes analyse la notion de grandeur, soumise à l'ordre et à la mesure (au sens de rapports, de proportions), c'est-à-dire géométrisable ou arithmétisable, en se fondant sur la dimension spatiale prise pour archétype[10]. Kant dans la *Critique de la raison pure* (2e éd., 1787), fait de cette même dimension spatiale le point de départ de sa description des *grandeurs extensives* qui, avec les *intensives*, appartiennent à ce qu'il appelle, dans son *Analytique transcendantale*, les « axiomes de l'intuition » et les « anticipations de la perception », qui figurent parmi les *principes synthétiques de l'entendement pur*, rendant possible la connaissance et, en particulier, l'application des mathématiques aux phénomènes [41].

Dans sa dissertation inaugurale de 1854, *Sur les hypothèses qui servent de fondements à la géométrie*[11], Riemann accorde une place privilégiée aux trois dimensions de l'espace ou, plutôt, à l'espace à trois dimensions, dans son étude systématique des grandeurs ou variétés à *n* dimensions, en l'entendant comme l'*espace physique*, dont les propriétés topologiques et métriques ne sont pas données par les propriétés générales des grandeurs, mais par les « concepts empiriques sur lesquels sont fondées les déterminations métriques de l'étendue », à savoir « le concept de corps solide et celui de rayon lumineux » [86]. Il précisait que ceux-ci « cessent de subsister dans l'infiniment petit », et ajoutait : « Il est donc très légitime de supposer que les rapports métriques de l'espace dans l'infiniment petit ne sont pas conformes aux hypothèses de la géométrie, et c'est ce qu'il faudrait effectivement admettre, du moment où l'on obtiendrait par là une explication plus simple des phénomènes. La question de la validité des hypothèses de la géométrie dans l'infiniment petit est liée à la question du principe intime des rapports métriques dans l'espace. » [86].

10 – Descartes [15]. *Cf.* Paty [74].

11 – Riemann [86]. *Cf.* Paty [60], chap. 6.

Cette remarque concerne directement notre propos : elle peut se traduire en une question sur la structure de l'espace dans les dimensions qui correspondent au domaine quantique. On sait que, concernant le domaine opposé, celui des très grandes dimensions, qui correspond en fait à l'astronomie et à la cosmologie (cette dernière n'existait pas encore comme science à l'époque), Riemann avait préparé les conditions de la pensée de l'espace par la théorie de la relativité générale. En caractérisant la différence entre la *topologie* et la *métrique*, Riemann était amené à distinguer l'*illimité* (ce qui n'a pas de bornes), qui appartient aux rapports d'étendue (topologiques), et l'*infini*, qui appartient aux rapports métriques, ce qui faisait voir d'emblée la possibilité de géométries non euclidiennes (en particulier, par la possibilité de concevoir des espaces finis sans limites, comme un plan sphérique).

Ces évocations nous rappellent que la pensée physique, et la mathématisation de la physique à travers la notion de *grandeur d'expression mathématique*, ont toujours été, jusqu'à la théorie de la relativité générale inclue, tributaires d'une pensée de l'espace. Elles nous rappellent que les phénomènes physiques et les systèmes (ou corps) physiques qui en sont le siège ou la source sont, en général, pensés dans l'espace, même s'ils ont contribué à modifier la pensée de l'espace et à réformer notre « intuition spatiale ».

De réceptacle indifférent de la physique newtonienne, le concept d'espace est ensuite revenu à l'indissociabilité d'avec les corps qu'il renferme, avec les phénomènes dont il est le siège : l'espace (et, plus exactement, l'espace-temps) est structuré par la matière – énergie qu'il contient (leçon de la relativité générale chaque jour confirmée, notamment par la connaissance la plus précise des objets astronomiques).

Jusqu'à la théorie de la relativité générale, celle-ci comprise, les conceptions des systèmes physiques et des phénomènes dont ils sont le siège ou l'origine ont été tributaires d'une pensée de l'espace, et ces phénomènes et ces systèmes sont, en général, pensés dans l'espace, même s'ils ont contribué à modifier la pensée de l'espace et s'ils ont obligé à reformuler notre « intuition spatiale ».

2.2. Est-il possible de penser sans l'espace ?

Ainsi la physique pensait-elle *dans l'espace* et pensait-elle l'*espace*, comme elle pense dans le temps (première définition, selon Newton, après l'invention du temps physique par Galilée, de la mécanique, qui fut la première des sciences classiques à être mathématisée, *cf.* [61, 62]), jusqu'à l'avènement de la physique quantique. La physique quantique a remis en question cette certitude tranquille en mettant la pensée de l’espace et dans l’espace devant le dilemme suivant : *ou bien* garder le concept d’espace physique dans le sens précédent qui peut sembler naturel (mais dont on a vu qu’il ne l’est pas plus que les autres, puisqu’il est construit) et renoncer à l’idée que les « systèmes quantiques » pourraient être représentés directement dans cet espace comme des systèmes dotés de propriétés qui leur soient propres ; *ou bien* admettre que cela a un sens de parler de systèmes quantiques ayant des propriétés et *enlever à l’espace son rôle premier* dans la représentation de ces systèmes.

En prenant ce second point de vue, on est amené à *penser* la physique autrement que « *dans l’espace* » au sens habituel ; et cependant, on a toujours, à un moment ou à un autre, à envisager les systèmes physiques quantiques en référence à cet espace, qui est celui des corps macroscopiques et du cosmos dans son état actuel, notamment lorsqu’il faut raccorder les systèmes quantiques, considérés intrinsèquement, avec les systèmes d’observation par lesquels nous avons connaissance, même si c’est très indirectement, de leurs propriétés quantiques elles-mêmes. Se demander *comment une telle pensée est possible* amène donc, en même temps, à s’interroger sur le rapport entre le monde quantique et l’espace physique de notre expérience[12]. Cette interrogation passe par un inventaire des modifications du concept d’espace requises par la description physique des systèmes quantiques.

On peut aussi se demander, comme les corps macroscopiques sont constitués de systèmes quantiques (particules élémentaires, atomes, etc.), et comme leurs propriétés macroscopiques résultent de l’effet sous-jacent des propriétés quantiques[13], et sont comme « émergentes » à leur niveau propre à partir de ces dernières, si le concept d’espace physique lui-même ne serait pas « émergent » à ce niveau[14], que l’on supposerait constitué à partir de l’effet de propriétés quantiques qu’on pourrait appeler *pré-spatiales*. On pourra envisager cette « émergence » de l’espace physique, à partir des propriétés des systèmes quantiques, de deux points de vue : un point de vue *épistémologique*, et un point de vue plus directement physique, en fait *cosmogénétique*.

12 – Sur le rapport, d’une manière générale, entre les pensées respectives du domaine quantique et du domaine classique, voir Paty [72].

13 – C’est un postulat que nous admettons, comme il est d’ailleurs généralement admis par les physiciens. Nous ne pouvons discuter ici les problèmes épistémologiques et philosophiques qu’il soulève : disons simplement qu’il s’appuie sur une conception moniste du monde naturel, qui ne se confond pas avec un réductionnisme de principe, que l’introduction des notions de *niveaux de structuration* des corps physiques et d’émergence permet d’éviter.

14 – Sur cette notion d’« émergence », voir Paty [57], chap. 8, p. 283-284, et une discussion par Zahar [93].

Du point de vue *épistémologique*, on constate la caractérisation d'un certain niveau de constitution ou d'organisation par des concepts propres à partir d'un niveau sous-jacent, lui-même caractérisé par des concepts signifiants dans son domaine ; les deux séries de concepts théoriques pour chacun des niveaux sont supposés se raccorder en principe l'une à l'autre par une correspondance, même si celle-ci n'est pas toujours acquise, soit par un défaut de connaissance, soit par une trop grande complexité des effets. Du point de vue *cosmogénétique*, on se préoccupe de caractériser et de formuler théoriquement la formation et l'expansion de l'espace physique des objets macroscopiques et du cosmos.

La difficulté ou l'impossibilité de penser sans l'espace est bien illustrée par le passage suivant de la *Critique de la raison pure* de Kant, dans l'*Esthétique transcendantale* : « L'espace est une représentation (*Vorstellung*) nécessaire *a priori* qui sert de fondement à toutes les intuitions (*Anschauungen*) externes. On ne peut jamais se représenter qu'il n'y ait point d'espace, quoiqu'on puisse bien penser (*denken*) qu'il ne s'y trouve aucun objet. L'espace est donc considéré comme la condition de possibilité des phénomènes (*Erscheinungen* : ce qui apparaît dans la perception), et non pas comme une détermination qui en dépende, et il est une représentation *a priori* servant nécessairement de fondement aux phénomènes externes[15]. » Remarquons en passant que l'indication selon laquelle on peut penser qu'il ne se trouve aucun objet dans l'espace indique combien la pensée kantienne de l'intuition de l'espace est tributaire de l'espace absolu newtonien (contenant de tous les corps, qui n'est pas affecté par eux). Il en fait un forme *a priori* de la sensibilité, et non plus un concept mathématique ou physique. On sait que le contenu du concept newtonien d'espace (absolu) diffère de celui du concept cartésien correspondant : l'espace cartésien (relatif aux corps), constitue une propriété essentielle des corps. C'est le détachement par Newton de l'espace par rapport aux corps, et sa prééminence sur eux, traduite en antériorité, qui permet à Kant d'en faire la pierre angulaire de son édifice philosophique sur la connaissance.

Mais ce qu'il m'importe ici de souligner c'est l'*absolue nécessité*, pour Kant, de l'*espace pour concevoir les phénomènes*. Si l'on veut, comme nous venons de le suggérer, remplir un programme pour la connaissance qui ne mette pas l'espace au premier plan, il nous faudra passer outre à l'« interdit kantien », ou alors concevoir autrement, plus indirectement, l'effet sur la connaissance de quelque intuition spatiale.

15 – Kant [41] : Esthét. transc., p. 785. Voir le commentaire que donne de ce passage Gilles-Gaston Granger dans son livre sur *La pensée de l'espace* [32], p. 10.)

La définition kantienne du phénomène, *Erscheinungen* (« ce qui apparaît dans la perception »), limite ce dernier au domaine de la perception seule. Si l’on veut étendre la notion de *phénomène* en dépassant la perception directe, pour atteindre ce qui serait conçu et donné plus directement par l’entendement, et rapporté seulement de manière très indirecte (et complexe) dans la perception, on devra nécessairement sortir du schéma kantien. Remarquons, cependant, que l’on peut continuer de trouver, dans la philosophie kantienne, la trame élémentaire, relative pour ainsi dire, la « condition élémentaire de possibilité » d’une pensée rationnelle de la science. En transportant le contenu du concept newtonien d’espace dans le contenu de l’intuition sensible correspondante, Kant liait étroitement, il est vrai, sa conception (son « esthétique transcendantale ») à la conception newtonienne, et devait encourir pour cela la critique empiriste, par exemple celle de von Helmholtz, plus tard celles de Reichenbach et de Carnap. Mais, d’un autre coté, ce qu’il constituait n’étant pas une physique (mais une « esthétique transcendantale »), sa conception échappait en un sens à cette même critique empiriste. C’est ce qui permit à Cassirer de proposer, dans ses ouvrages *Substance et fonction* et *La théorie de la relativité d’Einstein*, de modifier la signification de l’*a priori* kantien, en « fonction de spatialité », compatible avec les constructions des mathématiques et de la physique modernes. L’inspiration fondamentale de l’esthétique transcendantale peut alors être considérée comme compatible avec une pensée critique, même radicale, de la notion d’espace.

Faisons à cet endroit une autre remarque. Que l’espace (et le temps) soi(en)t la condition de possibilité des phénomènes, même sous une forme atténuée, c’est ce qui est en cause avec la physique contemporaine. L’espace possède-t-il à cet égard un privilège sur d’autres notions liées fondamentalement à la perception ? (par exemple, des notions liées à des grandeurs intensives et non plus extensives, comme le son, la lumière, les couleurs, la chaleur, etc.). De Galilée et Descartes à Leibniz, Newton et Kant, l’espace et le temps ont un privilège sur les autres grandeurs, mais surtout en ceci qu’elles sont à l’origine de la pensée de ces dernières (cela est explicite dans la Règle 14 de Descartes mentionnée plus haut). Si l’on considère les données actuelles de la physiologie cérébrale et neuronale et de la psychologie cognitive, on peut admettre de maintenir ce privilège, au moins comme un choix de représentation fondamentale de référence. L’histoire des connaissances et des sciences fait également apparaître la représentation spatiale comme étant effectivement fondamentale et première, depuis les objets et des images de l’homme préhistorique jusqu’à la géométrie. Quant à la représentation temporelle, elle sembla également une donnée première, encore que l'on puisse se demander si c'est la notion de temps ou celle de mouvement qui s'est imposée en premier ; mais cela revient en fin de compte au même (le temps peut être considéré comme une combinaison de mouvement et d'espace, comme le mouvement est une combinaison d'espace et de temps).

Concluons (provisoirement) sur ces questions en indiquant que Kant, malgré les restrictions qu'on a vues, laisse peut-être, par le rationalisme affirmé de sa philosophie, la porte ouverte sur une possibilité de dépasser, dans la connaissance, les limitations du sensible. Gilles Granger a donné à ce sujet une précision éclairante, à propos du rapport entre l'*intuition sensible* (forme *a priori* de la sensibilité), qui conditionne la possibilité de connaissance des objets, et la connaissance elle-même de ces objets, en faisant remarquer que Kant l'attribue au « caractère constructif de la mathématique en général, qui s'oppose alors à la philosophie, et de la géométrie en particulier » ([32], p. 10-11). Cette géométrie est pour Kant, « la science qui détermine synthétiquement et pourtant *a priori* les propriétés de l'espace » [41]. D'autres, allant au-delà de Kant, parleront d'éléments rationnels de connaissance, inventés par la pensée créatrice, choisis librement par la raison pour éclairer notre expérience. . .[16]

2.3. Une physique sans l'espace est-elle possible ?

Les discussions sur la critique des conceptions kantiennes suscitées par les géométries non euclidiennes et la relativité générale avaient amené un philosophe néo-kantien comme Ernst Cassirer à nuancer l'affirmation de Kant, en proposant de considérer comme inhérente à la pensée non pas la notion d'espace, mais une « fonction de spatialité », comme on vient de le rappeler, permettant l'élaboration de concepts moins rigides. Pourtant les physiciens (ni les autres, d'ailleurs. . .) n'envisagent pas volontiers, en général, la perspective d'abandonner les concepts spatio-temporels dans les éléments de base des descriptions de la physique, pour peu du moins qu'ils se préoccupent du raccord de leurs théories aux phénomènes physiques directement accessibles.

On pourrait rassembler tout un florilège de déclarations d'auteurs, physiciens ou philosophes, faisant état de l'impossibilité de penser une physique sans espace (voir par exemple Louis de Broglie[17]), ou du moins de la difficulté à le faire.

16 – Entre autres, Poincaré et Einstein. Voir Paty [68].

17 – Par exemple, de Broglie (1966), p. 62-63, cité dans Paty [57].

Einstein, qui pouvait le concevoir, selon ce qu'il déclara à plusieurs reprises à propos des systèmes de la physique quantique, avouait cependant ne pas savoir comment penser ainsi[18]. Évoquant les problèmes de la théorie quantique des champs, il écrivait, dans *Physique et réalité* : « On a certes fait observer que l'introduction d'un continuum d'espace-temps pouvait, eu égard à la structure moléculaire de tout événement dans l'infiniment petit, être déjà considérée comme allant contre la nature. Le succès de la méthode de Heisenberg ne nous renverrait-il pas à un mode de description de la nature purement algébrique ? Ne nous invite-t-il pas à éliminer de la physique les fonctions continues ? Mais, dans ce cas, il faudrait également renoncer, sur le plan des principes, à faire usage du continuum d'espace-temps. Il n'est pas impensable que l'esprit humain, à force de perspicacité, finisse par découvrir des méthodes permettant de franchir ce pas. Mais, pour le moment, une telle ambition équivaut à vouloir respirer dans un local sans air. » ([27], vol. 5, p. 148.)

Et encore, dans une lettre à un correspondant, à propos de la possibilité d'envisager une théorie du discontinu, il exprimait la difficulté à penser ainsi dans les termes suivants : « Dans une telle théorie [du discontinu, sans rapports différentiels], il n'y a pas place pour l'espace et le temps, mais uniquement pour des nombres, des constructions numériques et des règles pour les former sur la base de règles algébriques excluant le processus limite (...). Dans une théorie algébrique de la physique (...), il sera particulièrement difficile, par exemple, de déduire un semblant d'ordre spatio-temporel (...). Je ne vois pas à quoi pourrait ressembler la structure axiomatique d'une physique comme celle-là, et je n'aime pas qu'on en parle par allusions obscures. Mais je crois tout à fait possible que l'évolution nous y conduise (...)[19]. »

Du côté des physiciens qui considèrent que l'on doit en venir à un dépassement des grandeurs d'espace, on notera la conception du « second » David Bohm, après son abandon du programme des variables cachées, envisageant de substituer à l'ordre selon l'espace et le temps un ordre différent, « impliqué » ou « emplié » ([5], chap. 5, 6, 7 et [6]).

18 – Einstein [18, 22], cité dans Paty [57].

19 – Albert Einstein, Lettre à Hans S. Joachim (14.8.1954), *in* Einstein [27], vol. 1, p. 256-257.

3. Caractérisation d’un « monde quantique »

3.1. La pensée physique des phénomènes et des systèmes quantiques

Lorsque les physiciens étudient les propriétés des systèmes qui sont à l’origine des phénomènes quantiques, avant de les ramener aux caractéristiques des dispositifs expérimentaux par lesquels ces phénomènes sont mis en évidence, ils travaillent dans le système conceptuel de la théorie quantique elle-même, avec ses grandeurs exprimées mathématiquement et leurs relations propres. Cela est particulièrement évident lorsqu’il s’agit de travail théorique à proprement parler, mais cela se marque aussi au niveau de l’étude « phénoménologique » de cette physique même, tant théorique qu’expérimentale. Il s’agit de comprendre tel ou tel processus de transformation ou d’interaction de particules et de champs (quantiques). Du point de vue de ce travail et de cette pensée théorique, le raccord avec le monde phénoménal est assuré par le fait que, pour un processus donné, les calculs théoriques sur les quantités comme les opérateurs d’interaction et les vecteurs d’état donneront, par exemple, les amplitudes de transition entre les états initial et final qui fourniront, en définitive, combinées au volume d’espace de phases disponible[20], des quantités telles que des sections efficaces d’interaction ou des durées de vie de désintégration (c’est-à-dire des fréquences statistiques) directement mesurables expérimentalement. Mais, si ce sont ces fréquences que l’on détermine effectivement par la mesure, l’objet de la pensée physique sur lequel elles nous informent reste bien l’*état quantique* désigné par la représentation théorique, communément appelée « le formalisme », bien qu’il s’agisse de théorie physique : c’est-à-dire telle « particule quantique », avec sa charge, son spin orienté (polarisé), ses autres « nombres quantiques »...

Tous ces raisonnements portent donc fondamentalement sur les grandeurs du « formalisme » de la théorie quantique, considérés de fait comme significatifs du point de vue physique, et fonctionnant comme les véritables éléments de la pensée physique, de la théorie à l’expérience. On représente, par exemple, un électron et un neutrino par leurs *spineurs* respectifs, qui sont les fonctions d’état de ces particules quantiques à spin demi-entier, régies par le formalisme de l’équation de Dirac et celui de la théorie quantique des

20 – Plus précisément, le module de l’amplitude de transition entre un état initial et un état final, élevé au carré et intégré sur le volume d’espace de phases disponible des variables, fournit la probabilité du processus considéré. Sur la théorie quantique des champs, voir, p. ex., Itzykson et Zuber [40].

champs[21]. En les associant adéquatement aux fonctions d'état d'un neutron et d'un proton, libres ou liés dans un noyau atomique, la théorie fournit ainsi les propriétés de la désintégration β du neutron ou, inversement, de l'interaction (faible) d'un neutrino sur un neutron, donnant un proton et un électron dans l'état final. En calculant un tel processus, les physiciens raisonnent comme s'ils pensaient directement ces processus, et ils le font avec d'autant plus de bonne conscience que les expériences leur révèlent bien ces particules, identifiées dans les détecteurs, et qu'elles leur confirment (ou non) leurs prédictions théoriques sur le comportement de ces réactions (leur structure).

Certes, les expériences ne leur donnent accès qu'à des traces indirectes dans les détecteurs (grains d'argent révélés d'émulsion photographique, gouttelettes de liquide ou bulles gazeuses, étincelles, signaux électriques, ...) qui, en dernier ressort, sont des effets macroscopiques relevant de la physique classique. Mais les règles de la physique quantique leur permettent de passer de l'un à l'autre, d'effectuer, pour ainsi dire, le « transit » entre les deux mondes : celui de leur représentation des systèmes quantiques considérés en eux-mêmes, et celui du monde macroscopique, de la physique classique des dispositifs expérimentaux. En absence d'interactions avec un environnement matériel diffus, idéalement dans le vide, des particules quantiques gardent au long de leur parcours leur caractère quantique (exprimé, pour l'essentiel, par le principe de superposition linéaire des fonctions d'état, c'est-à-dire la cohérence des phases de ces dernières), et elles le maintiennent encore lorsqu'elles entrent en collision avec une autre : toute la physique des particules et des champs quantiques est fondée sur cette idée.

Ce sont ces phénomènes se produisant entre des systèmes quantiques, quand ils sont *livrés à eux-mêmes et non pas projetés ou réduits* sur des caractères ou des concepts classiques (onde, corpuscule ; position, impulsion ; etc.), exprimant leur « dualité » et passibles des « relations d'incertitude » ou inégalités de Heisenberg, qui constituent ce que l'on est bien en droit, finalement, d'appeler le « *domaine quantique* » ou le « *monde quantique* ». Pour les physiciens qui sont habitués à travailler sur lui, ce monde est aussi réel que le domaine ou le monde classique (c'est-à-dire des objets décrits par les moyens conceptuels et théoriques de la physique classique). Il peut être tout aussi bien *pensé* que l'autre, mais seulement au moyen des concepts et de la théorie quantique, et de la manière la plus précise. Simplement, sa relation aux impressions

21 – Les matrices γ_μ de l'équation de Dirac jouent un rôle fondamental : sur elles sont construits les opérateurs d'interaction des champs des interactions électromagnétique, faible et forte.

sensibles, sources initiales de nos connaissances, est plus indirecte, et réclame davantage de médiations.

Telle est du moins la manière dont les physiciens d'aujourd'hui auront tendance à répondre spontanément, si on les interroge sur ce que sont les « objets » sur lesquels ils travaillent. Ils rapporteront, par exemple, les résultats obtenus en détectant les neutrinos en provenance du Soleil, en termes d'oscillations des neutrinos, durant leur propagation, de son état initial aux deux autres[22]. La connaissance de ce « monde quantique » présente des particularités qui ont longtemps fait problème et suscité des interprétations diverses. Nous ne nous appesantirons pas ici sur ces questions. Nous indiquerons simplement que l'acceptation « pratique » de ce monde, fondée à la fois sur la pensée théorique et sur les résultats d'expérience, doit s'accompagner d'une élucidation conceptuelle, épistémologique, de son mode de représentation. Cette élucidation n'est pas toujours acquise, et des ambiguïtés demeurent souvent quant au statut de réalité physique de ce « monde » et au caractère pleinement *physique* de sa *représentation*.

Pour prendre un exemple, les cosmologues quantiques se trouvent, me semble-t-il, dans la nécessité de résoudre la contradiction, dont ils ne se rendent pas tous compte, entre les deux attitudes ou positions suivantes. D'une part, ils professent une conception observationaliste et subjectiviste de la physique quantique, mettant en avant l'observation comme référent supposé de la théorie et le prétendu « principe » de réduction de la mesure. D'autre part, ils paraissent souscrire à une conception objective et réaliste fondamentale de cette même physique quantique dès qu'il est question de la faire servir à la considération de l'Univers dans son ensemble.

Selon moi, l'élucidation dont je parle est possible et demande une réflexion et une analyse épistémologique sur la signification physique des grandeurs théoriques utilisées en physique quantique et sur leur signification, du point de vue physique. En l'absence d'une telle élucidation, l'attitude courante reste extrêmement ambiguë quant à la signification du rapport de la théorie à l'expérience ; l'on évite en général toute discussion en rapportant les grandeurs théoriques au « formalisme » et à ses règles d'interprétation. Et, pour ces dernières, on invoque encore bien souvent la référence de l'observation, héritage de l'« interprétation orthodoxe » ou « philosophie de la complémentarité ».

22 – Les neutrinos du type ν_e, produits dans les réactions nucléaires du Soleil, se transforment en partie (semble-t-il) en ν_μ ou ν_τ, ce qui est possible par le principe de superposition (voir, p. ex., Paty [63]), et les résultats rapportés récemment (en 2000 et en 2001) dans les publications spécialisées. Voir aussi, sur les développements récents de la physique subatomique, Bimbot et Paty [4].

Les grandeurs du « formalisme quantique » sont-elles des grandeurs mathématiques, ou des grandeurs physiques, mathématisées, comme on le considère pour les autres théories physiques ? Admettre la deuxième réponse (que *ce sont des grandeurs physiques*) permet de constater un déplacement du problème de l'interprétation. Au lieu d'être un problème sur la nature de la connaissance (observationalisme contre réalisme) il s'agit désormais, de manière plus banale, de garantir le caractère physique approprié d'une grandeur formulée mathématiquement.

L'intelligibilité du « monde quantique » se trouve grandement simplifiée si l'on adopte le point de vue selon lequel il est légitime de parler d'un « domaine quantique » propre. Elle demande, pour cela, un examen des divers aspects de la représentation théorique, concordants avec la connaissance expérimentale des phénomènes, et des problèmes d'interprétation qui leur sont reliés. Mentionnons-les ici pour mémoire (nous avons étudié certains d'entre eux par ailleurs[23]) : le concept d'état et de fonction, ou vecteur, d'état et les grandeurs dynamiques appelées « observables » ; la non-séparabilité locale ; le rapport entre état quantique et probabilités et la double signification de ces dernières (théorique et statistique) ; le rapport entre les domaines quantique et classique ; la pensée théorique de la physique emmenée par les formes mathématiques, pour aboutir à une représentation légitimement physique ; le problème de l'espace physique, vu du monde quantique (objet du présent exposé) ; la décohérence et la mesure ; puis le problème de la gravitation quantique et de la matière de la cosmologie.

Nous considérons, dans ces analyses, que l'on doit admettre comme physique, parlant de concepts ou de représentation, ce qui est théoriquement significatif en considération des phénomènes décrits. En ce sens, des grandeurs physiques exprimées mathématiquement sont légitimées par l'intelligibilité qu'elles procurent dans un domaine physique donné par l'expérience des phénomènes. Il convient donc de concevoir une extension de sens de la notion de grandeur physique, au-delà des limitations généralement admises pour cette notion aux seules grandeurs « à valeurs numériques » et « directement mesurable ». On admettra dès lors que la fonction d'état (« ou fonction d'onde »), définie comme vecteur dans un espace de Hilbert, soumis au principe de superposition, et les variables dynamiques correspondantes (les « observables ») décrites sous forme d'opérateurs agissant sur la première sont des grandeurs physiques décrivant le système quantique. Ces grandeurs sont les grandeurs

23 – Pour une discussion plus détaillée, dans cette perspective, de ces questions d'« interprétation physique », voir Paty [57, chap. 6, 68, 71, 78 ; 56, 58, 70, 72, 73, 75].

appropriées à une telle description, au contraire des grandeurs classiques correspondantes (quand il y en a), qui n'en décrivent que des projections sur des grandeurs classiques.

En fait, pour le dire d'un mot, les termes du débat quantique traditionnel apparaissent archaïques aujourd'hui, si l'on veut bien prêter attention à la connaissance effective des systèmes quantiques dont on dispose, qui va très au-delà de ce qu'autorisait la conception orthodoxe « de la complémentarité ». Les connaissances récentes acquises sur des systèmes quantiques individuels, en particulier, montre comment la théorie quantique implique et décrit directement de tels systèmes. De même, dans l'autre sens, les systèmes collectifs à grand nombre de composants identiques indiscernables, tels que la condensation de Bose-Einstein, récemment observée, qui paraissaient naguère une vue de l'esprit, qui correspondrait plutôt à un défaut de connaissance, sont démontrés exister effectivement, objectivement, dans la nature, remplissant exactement les traits que la théorie leur prévoyait.

3.2. La signification philosophique d'une extension de sens de la notion de « grandeur physique »

L'extension de sens proposée pourrait simplifier considérablement notre compréhension des domaines de connaissance correspondants. La simplification serait même radicale en ce qui concerne les problèmes d'interprétation. Elle rétablirait, comme conséquence immédiate, la signification et l'usage de la notion d'*objets* d'une théorie conçue comme description et représentation. Le genre de réalisme que nous considérons ici est, certes, un *réalisme critique*, celui de *constructions symboliques* pour *représenter la « réalité »*, et conçu comme un programme pour l'élaboration scientifique[24].

Prenons les concepts d'*état* et de *grandeur quantiques*. Étant donné que la théorie *quantique* permet d'expliquer tant de groupes de phénomènes et d'en donner des modèles puissants, il est tentant de la concevoir comme une théorie fondamentale d'un *monde d'objets*. Et telle est bien, de fait, comme nous l'avons vu, la manière spontanée dont les physiciens la pratiquent, bien qu'ils se heurtent à des difficultés quand (mais seulement alors) ils en viennent à s'interroger sur la *transition* de ce *domaine quantique* au *domaine classique* des appareils de mesure. L'interprétation « courante » (« de Copenhague ») professe qu'il n'existe pas, scientifiquement, de *signifié conceptuel* désigné dans la théorie, pouvant être considéré comme un objet (c'est-à-dire une entité

24 – Paty [57], en particulier les chap. 1 et 10 ; [60], chap. 9.

ayant des *propriétés*) et que ce *signifié conceptuel* (ce qu'on appelle « objet » de la théorie, soit l'état du système) n'existe qu'en relation à des conditions données (et optionnelles) de préparation pour la mesure.

Mais, en réalité, les scientifiques qui travaillent avec des objets physiques quantiques récusent par leur pratique (même s'ils ne l'osent pas de façon explicite concernant des questions de « philosophie ») cette interprétation courante, en élaborant une *nouvelle objectivité* conçue sur le mode de l'ancienne, mais faisant appel à des concepts et à des grandeurs dotées d'un *sens plus large* que les grandeurs classiques. Il est vrai qu'un *état quantique* n'est accessible à l'expérience qu'indirectement, mais ceci n'affecte pas la possibilité d'*en avoir connaissance*. La modification épistémologique essentielle aura consisté, en vérité, en une *extension de sens* (restée implicite) du concept de *grandeur* ou de *quantité physiques* (en particulier, celui d'*état physique*), à des entités qui ne sont pas simplement à valeur numérique. Cette extension est légitimée par les *phénomènes*, dans une acception de ce terme qui ne les réduit pas à leur appréhension par la *perception*, mais qui les conçoit selon l'*entendement*, c'est-à-dire selon leur capacité à être portés à notre *connaissance*, et elle est réalisée par l'essentiel du formalisme même de la théorie quantique.

Cette extension a été, de fait, préparée par les travaux des physiciens théoriciens de la physique quantique sensibles aux propriétés formelles, mathématiques, de la théorie, comme Max Born, Werner Heisenberg, Paul Adrian Dirac, John von Neumann, Hermann Weyl et d'autres, où les grandeurs physiques classiques étaient remplacées par des « grandeurs quantiques » différentes d'elles en premier lieu par leurs propriétés formelles (superposition pour les fonctions d'état, non-commutation pour les opérateurs, ...). Par exemple, les *nombres-q*, non commutatifs, proposés par Dirac pour remplacer les *nombres-c* ordinaires, suggèreraient immédiatement une extension de sens comme celle que nous venons d'indiquer[25]. Mais ces pionniers n'avaient cependant pas cru devoir proposer d'emblée ces constructions *formelles* comme de simples extensions de sens des grandeurs *physiques* parce que les questions d'interprétation alors soulevées ne paraissaient pas les y autoriser. De telles grandeurs restaient seulement mathématiques, leur rapport aux phénomènes physiques étant réglé par l'« interprétation ». La pierre d'achoppement était

25 – Voir les travaux de Dirac de 1926. *Cf.* Mehra et Rechenberg [53], vol. 4, p. 162 et suiv., Darrigol [14].

alors essentiellement le passage du classique au quantique, avec la question de la « mesure » au sens quantique[26].

Considérer comme grandeur physique au sens plein du terme des *vecteurs d'état quantiques* dans la forme de superpositions linéaires et des « opérateurs observables » ne commutant pas nécessairement, pourvus de valeurs propres affectées de probabilités, cela implique de renoncer à la connexion étroite, voire à l'identification, de *propriétés* avec *ce qui est ou pourrait être mesuré* au sens classique et numérique, et de concevoir différemment ce qu'est une *propriété*, en la rapportant au système ou à l'état *tel qu'il est construit intellectuellement* par un processus d'abstraction et d'élaboration théorique intégrant des données factuelles. Des propriétés conçues de la sorte *ne sont plus contextuelles* et peuvent être dites *intrinsèques* : telles sont, dans cette perspective, les propriétés des « particules » quantiques élémentaires (photon, électron, quark, etc.), munies de leur charge, de leur spin, etc., et des champs quantiques.

Ces *propriétés* ne dépendent aucunement des circonstances de leur observation, mais elles sont *reconstituées* à partir d'observations expérimentales fournissant des valeurs de grandeurs correspondant à des propriétés contextuelles, affectées de probabilités mesurées par des fréquences d'événements. À cet égard, les probabilités, loin de constituer une limitation de la connaissance, *permettent la détermination* des grandeurs intrinsèques, qui sont celles dont la théorie se soucie principalement, à partir de la distribution spectrale de leurs composantes.

4. La physique quantique et le concept « classique » d'espace

4.1. Les limitations du concept d'espace pour la physique quantique

Nous ne passerons pas ici en revue de manière détaillée les différentes limitations que la théorie quantique oblige à mettre au concept d'espace : probabilité de présence et non pas localisation certaine ; absence de trajectoire définie ; inégalités de Heisenberg entre la position et l'impulsion ; caractère étendu, c'est-à-dire non local, d'un système quantique capable d'interférer

26 – Paty [67] ; [77]. Sur la formalisation de la mécanique quantique et l'introduction des grandeurs mathématiques abstraites, voir, en particulier Dirac [17], von Neumann [54], Weyl [91], etc. Sur les diverses manière d'aborder le problème de la mesure en mécanique quantique, d'abord posé par von Neumann [54], voir Wheeler et Zurek [92].

avec lui-même dans un dispositif de diffraction ; indiscernabilité des identiques qui gouverne des comportements collectifs comme s'il s'agissait d'« influences à distance » ; non localité et corrélations quantiques à distance (ou non séparabilité locale des systèmes quantiques)[27]... Ces caractéristiques n'étaient dues, selon l'interprétation orthodoxe, qu'à l'impossibilité de s'affranchir, pour les phénomènes quantiques, de la référence à l'observation et de considérer des systèmes quantiques en eux-mêmes[28]. Pourtant, si l'on admet la signification, du point de vue de la pensée physique (et de la réalité physique correspondante décrite par la première), d'*un monde autoconsistant de propriétés et de systèmes quantiques*, ces limitations doivent être conçues comme des caractères objectifs. Elles signifient, d'une part, que le concept d'espace de la géométrie et de la physique classique-relativiste ne peut être utilisé tel quel comme concept de la physique quantique et, d'autre part, que des systèmes quantiques venant à occuper une région d'espace donnée, précisément défini ou macroscopique, confèrent (tant qu'ils conservent ces propriétés, en cohérence de phase) à l'espace occupé des propriétés « quantiques », différentes des propriétés habituellement attachées à un tel espace, et notamment sa divisibilité en points élémentaires.

Les grandeurs classiques qui servaient d'éléments de base à la théorie physique classique et ont permis d'élaborer sa formulation lagrangienne-hamiltonienne ont perdu, avec la physique quantique, leur signification physique directe. La théorie quantique a gardé, dans sa construction, la part « formelle », c'est-à-dire la formulation hamiltonienne, à la condition d'abandonner la signification, et même la forme, des grandeurs initiales. À part le temps, qui n'a pas d'équivalent comme opérateur, les autres variables ont été remplacées, dans l'« analogie formelle » qui a gouverné la formulation de la mécanique quantique, par des grandeurs dynamiques sous forme d'opérateurs (les « observables »). Les équivalents, ou plutôt les correspondants classiques (éventuels), des variables de la théorie physique classique fournis par la théorie quantique sont les valeurs propres de ces opérateurs correspondant à un état propre solution de l'équation d'état (et soumis au principe de superposition, de telle façon que toute superposition linéaire d'états propres est aussi un état possible du système). Ainsi un système quantique n'a-t-il une position dans l'espace que sous les conditions qu'octroie le statut de valeur propre. La position spatiale d'un système quantique n'est attribuable que de manière

27 – Sur la non séparabilité, voir Bell [3], Aspect [1], ...

28 – Sur l'interprétation « orthodoxe » ou « de Copenhague », voir notamment : Bohr [8], Rosenfeld [85].

contextuelle, et soumise ainsi aux limitations et conditions rappelées plus haut. Elle ne fait pas partie des concepts « propres » des systèmes quantiques, et l'on doit, pour la considérer, préciser à quelles circonstances relatives elle est conditionnée. L'absence de trajectoire déterminée, la non-localité, l'extension spatiale, des systèmes quantiques résultent de ces considérations.

4.2. Engendrement d'espace, cosmogenèse, et espace vide de la propagation des systèmes quantiques

Lorsqu'on parle de systèmes quantiques tels que les décrit la théorie quantique, on est amené à considérer que leurs « propriétés », au sens plein du terme (c'est-à-dire ce qui les caractérise en propre, indépendamment de tout « contexte » au sens expérimental du terme), sont *autodéfinies* par cette théorie. Mais, en même temps, on les considère en fait le plus souvent dans le cadre d'un espace déjà donné, qui est celui admis antérieurement de notre expérience, de l'ensemble de nos expériences, du monde environnant et du cosmos. C'est l'espace (-temps) classique, de la physique relativiste (restreinte et générale) et de la cosmologie : celui où ces systèmes quantiques se propagent ; ou encore celui où ils s'agrègent ensemble pour former des atomes et des molécules et des systèmes de telles entités, qui, au niveau macroscopique, apparaissent très bien définis spatialement.

Dans le premier cas, celui de la propagation, l'espace classique (celui de la distance parcourue, par exemple par des particules d'accélérateurs, ou des neutrinos en provenance du Soleil) est pris, avec le temps, comme l'une des variables suivant lesquelles on détermine l'évolution du système, bien qu'elle ne soit pas inhérente à ce dernier : du moins correspond-elle à une grandeur physique issue de l'expérience, mais qui reste extérieure à la définition du système quantique.

Dans le second cas, l'espace n'est pas donné à l'avance, et le volume d'occupation se détermine en fonction de l'agrégation des constituants : il est d'autant moins défini que le nombre de constituants quantiques est petit. À cet égard la pensée des systèmes quantiques et de leur agrégation en « grands » systèmes qui occupent un espace défini paraît assez semblable à la constitution de l'espace dans une cosmogenèse, sans espace préexistant, l'espace se créant au fur et à mesure du déploiement de l'Univers. À ceci près, cependant, que la cosmologie quantique suppose toute la matière concentrée, sans espace ou quasiment, avant son déploiement dans l'espace, la notion qui est ici en cause étant celle d'espace comme contenant. L'espace de l'Univers s'engendre avec le déploiement (et la dilution d'énergie) de la matière. La cosmogenèse quantique (et ce trait appartient aussi bien à la relativité générale) correspond à l'engendrement de l'espace (et de l'espace-temps) par la matière. Selon la cosmologie quantique (qui n'est encore qu'une série de modèles théoriques, rassemblés dans le « modèle standard »), cette matière initiale est purement quantique, et cet engendrement de l'espace et du temps se fait par différenciation des états physiques [70] et découplage progressif des interactions, sauf l'interaction gravitationnelle.

Au début du processus se situe l'état de « gravitation quantique », sur lequel nous n'avons aucune certitude[29]. En considérant l'état actuel de l'Univers, l'espace physique (et le temps), issus de cette genèse, nous sont donnés, et les systèmes quantiques élémentaires s'y propagent, et parfois s'y agrègent en se localisant. On doit constater cet état de choses, sans bien savoir comment raccorder cette coexistence de deux types d'« objets physiques » aussi différents que *des systèmes quantiques*, pour lesquels le concept d'espace reste étranger à leur définition, et de l'*espace vide environnant*, dans lequel ils « existent ». On doit, en tout cas, considérer ici deux réalités physiques : les systèmes quantiques, et l'espace physique, créé dans la genèse cosmique, dans lequel ils se trouvent. À quoi s'ajoute que cet espace vide n'est pas vraiment vide, conceptuellement et physiquement, puisqu'il comprend le vide quantique, qui est matériel, représenté en termes de champs (et de particules virtuelles, corollaires de ces champs), en sorte que les deux entités s'appellent l'une l'autre, la primauté conceptuelle revenant, du point de vue adopté, au concept de champ quantique.

Il faudrait analyser davantage le rapport de ces deux « régimes », ou de ces deux représentations du monde physique, les systèmes quantiques et l'espace physique. Que peut-on dire d'un système quantique qui se propage ? Il ne se localise que lorsqu'il interagit avec le monde de l'espace-matière classique, par exemple : l'impact sur l'écran dans une expérience d'interférence ; la localisation par détection d'une particule d'un faisceau ; le confinement par champ magnétique, électrique, gravitationnel. Laissons provisoirement la question en suspens.

4.3. L'espace et les systèmes quantiques individuels

La réponse traditionnelle de l'interprétation « complémentariste » de la mécanique quantique pour maintenir l'espace classique et la spécificité

29 – Sur les problèmes d'une théorie quantique de la gravitation, voir, par exemple, Kouneiher [42].

quantique faisait appel à une dualité, ou une complémentarité, entre une représentation *causale* et une représentation spatio-temporelle, qui se justifiait en invoquant le caractère statistique des prédictions de la théorie quantique. Ce dernier caractère résultait du rôle fondamental, de principe et référentiel, de l'observation. Cependant, la connaissance non seulement *par*, mais *selon* l'observation, apparaît de plus en plus comme une conception *ad hoc*, qui se satisfait d'une demi-connaissance, et dont l'insuffisance se révèle avec celle d'une conception ambiguë de la signification de résultats probabilistes, dont elle est solidaire, avec l'affirmation de la « solution statistique ». N'existe-t-il pas de possibilité de penser « directement » les systèmes quantiques indépendamment de leur observation ? Cette revendication de la pensée théorique apparaît conaturelle au développement d'une véritable pensée de ces systèmes[30]. L'exemple des discussions évoquées plus haut sur la distinction entre l'espace perceptif et l'espace géométrique (construit par l'entendement) peut être ici utile pour nous suggérer une marche à suivre, vers la conception d'un système relationnel de grandeurs différent, ayant pris ses distances par rapport aux données immédiates de la perception et de l'observation, et conçu en connaissance de cause par le seul entendement (grâce au rôle exemplaire de la mathématique, maîtresse en rationalité). On peut aussi considérer, de manière plus spécifique, que la spatialité concernée par la physique quantique peut être reconstruite et pensée, rendue intelligible, par l'entendement, au-delà de la perception.

Pour en revenir à la réponse traditionnelle de l'interprétation « orthodoxe » de la mécanique quantique par l'approche statistique et la dualité entre le concept de système physique individuel et l'espace continu[31], son caractère insuffisant est également avéré aujourd'hui par les faits eux-mêmes. Il est clair, en effet désormais, que la théorie quantique (par la fonction d'état, les grandeurs dynamiques et les équations d'état) peut caractériser des systèmes individuels qui correspondent bien à des entités physiques (particule unique, atome unique, photon unique).

30 – John Bell considérait qu'une théorie quantique satisfaisante devrait faire appel à des « be-ables » plutôt qu'à des « observ-ables » [3]. Mario Bunge, de son coté, parlait de quantons, systèmes proprement quantiques, sans réduction classique à une dualité de l'onde ou du corpuscule (1973). Jean-Marc Lévy-Leblond s'est efforcé aussi de penser une théorie « proprement quantique », en commençant par « nettoyer » ses pseudo-concepts en faisant pleinement droit à ces quantons, et en modifiant en conséquence la manière d'enseigner la « Quantique » (Lévy-Leblond [46–48], Balibar et Lévy-Leblond [2]). On pourrait allonger la liste avec, notamment, les essais significatifs de formuler une théorie quantique *sans réduction par la mesure*, donc en se libérant du référent contraignant de l'observation.

31 – Voir Bohr, Born, Pauli...

En fait, un certain nombre de physiciens quantiques de la mouvance « orthodoxe » avaient bien conscience de la possibilité de rapporter les propriétés des systèmes quantiques à des systèmes individuels. Mais ils s'interdisaient de concevoir de tels systèmes indépendamment de leurs conditions d'observation ; il faudrait, pensaient-ils, pour s'assurer du caractère individuel d'un système quantique, le compter et, pour cela, le détecter (par interaction) suivant son aspect corpusculaire, donc détruire son aspect ondulatoire dual, ou tout simplement son caractère quantique. De toute façon, la conception « orthodoxe » au sens strict n'envisageait l'éventualité d'un système quantique individuel que si celui-ci était dûment « chaussé », pour ainsi dire, de l'instrument pour le détecter, et celui-ci le liait indissolublement à l'espace physique au sens classique-relativiste. Le problème d'une éventuelle difficulté à penser l'espace pour les systèmes quantiques ne se posait donc pas, puisque ces systèmes ne pouvaient être conçus qu'immergés dans cet espace, cadre de l'observation et donc des phénomènes rapportés à des « systèmes-en-observation », seuls concevables. L'espace connaissait seulement, en raison des nécessités de l'observation particulière aux phénomènes quantiques (gouvernée par l'indivisibilité du quantum d'action), des limites d'utilisation, dont nous avons parlé (fondamentalement, celle des relations d'inégalité de Heisenberg).

Cette vue, soit dit en passant, nous fait d'ailleurs comprendre pourquoi la réponse de Bohr à l'argument EPR (d'Einstein, Podolsky et Rosen) aurait été à elle seule strictement incapable de formuler clairement l'inséparabilité quantique comme concept et propriété [7, 28]. Celle-ci fut, en fait, énoncée par Einstein, en négatif, comme difficulté d'une *séparabilité locale pour des systèmes individuels*, précisément parce qu'il avait pensé le problème de la représentation de tels systèmes (ce qu'il appelait des « états physiques réels »)[32]. En bref, avec la conception « observationaliste » orthodoxe, il était inutile, hors de question et, somme toute, impensable, de pousser plus avant la pensée critique de l'espace à propos des systèmes quantiques.

En ce qui concerne les systèmes quantiques individuels, la réalisation, par construction, de faisceaux où les particules quantiques sont « égrenées » une à une, avec une résolution temporelle significativement contrôlable, permet, de nos jours, de faire des expériences avec des particules quantiques individuelles, et implique donc, même pour les physiciens (ou les épistémologues) observationalistes les plus invétérés, la nécessité d'une pensée physique de ces systèmes. À la caractérisation théorique d'individualité pour

32 – Voir, notamment : Einstein [23]. Sur les développements et la signification de cette question, voir Paty [56, 57, 76].

des systèmes quantiques, taxée naguère de purement formelle et sans portée physique, correspond désormais le donné factuel : la possibilité d'étudier effectivement des « particules quantiques » individuelles, qui était niée dans une perspective strictement opérationaliste, est désormais acquise comme un fait d'expérience[33]. C'est donc bien pour de telles particules quantiques individuelles, auxquels on doit désormais rapporter la description théorique quantique, que le concept habituel d'espace apparaît comme limité, et donc comme insuffisant.

Mais qu'est-ce que l'espace physique pour les systèmes quantiques ? Les difficultés épistémologiques suscitées par le concept « classique » d'espace physique nous invitent à ne pas le mettre directement en avant. Il n'y a pas d'« espace physique » qui nous soit donné indépendamment des phénomènes. Autrement dit, pour les phénomènes quantiques, nous admettrons au mieux qu'il y ait un « espace physique quantique » différent de l'« espace physique classique » (qui inclut celui au sens de la théorie de la relativité).

Il nous faut donc considérer, pour ce qui est des phénomènes quantiques, le concept et la grandeur espace, sous deux aspects. Concernant le premier : que nous dit la physique quantique, considérée comme physique d'un « monde quantique » (ou « monde des phénomènes quantiques »), sur l'espace physique ? autrement dit : la physique quantique définit-elle (ou caractérise-t-elle) un « espace quantique » ayant des caractéristiques particulières ? Concernant le second aspect : comment s'effectue le raccord entre l'espace selon le monde quantique et l'espace du monde macroscopique, ou espace physique classique ?

À la première question, la théorie quantique offre une réponse, par la manière même dont elle s'est constituée. À partir de la notion première de groupe de transformation et d'invariance sont formés les opérateurs, à l'aide de générateurs infinitésimaux : c'est la manière minimale dont intervient la notion de spatialité dans la formation des grandeurs quantiques. À la seconde interrogation, la réponse traditionnelle (mais insuffisante au regard de notre attente sur les systèmes quantiques), est celle de la mesure ; nous avons déjà évoqué un autre aspect, celui de la propagation dans l'espace physique au sens classique d'un système quantique, et mentionné la dualité, au sens de la coexistence, entre les systèmes quantiques et l'espace physique, classique dans lequel ils se propagent. Reste un troisième aspect, celui d'un espace

33 – Voir, entre autres, Grangier [33]. Sur la notion d'état quantique, sa description théorique et sa contrepartie physique, telle qu'on peut la caractériser aujourd'hui, voir Paty [77].

émergent à partir de certains types d'agrégats de systèmes quantiques. Nous l'avons déjà brièvement mentionné, et nous y revenons pour terminer.

5. Pré-espace physique quantique et espace physique classique

Les changements de concepts et de représentations, tels que ceux qui ont abouti à la formulation de la théorie quantique, s'ils ont été rendus nécessaires par l'expérience et les moyens « classiques » de celle-ci, ont été effectués au niveau de l'entendement, par le jeu des objets de pensée et des opérations sur eux, permis par la puissance des propriétés relationnelles de grandeurs mathématiques. Leur adéquation aux problèmes physiques considérés leur donne valeur de représentation physique effective. Ces transformations dans les structures mêmes de la connaissance (et non pas seulement dans ses contenus) correspondent à des élargissements successifs des éléments de rationalité dans la pensée, qui permettent d'assimiler notre expérience des données, tant des formes (objets formels) que du monde de l'empirie. Ainsi la physique a-t-elle pu s'abstraire de l'espace physique classique pour penser un « monde quantique » qui n'est plus organisé autour de cette notion, selon un processus exemplaire où, tout en entrant à un certain titre dans la construction des grandeurs quantiques l'espace y subit une mutation de sens. Cette mutation n'est pas de l'indifférence, car les systèmes quantiques ont très souvent à voir avec l'espace physique, selon ce que nous avons entrevu. Nous allons considérer, pour terminer, un cas particulier de ces rapports, lorsque des systèmes quantiques en parfaite cohérence de phase s'agglomèrent pour constituer une certaine région de l'espace physique, tout en étant totalement quantique.

5.1. Une expérience astronomique par la pensée : existe-t-il des étoiles « condensats de Bose-Einstein » ?

Comment pouvons-nous nous représenter l'espace de systèmes quantiques se trouvant en cohérence de phase ? Considérons le cas d'un grand nombre de particules ou systèmes quantiques identiques se trouvant agglutinés dans le même état physique : elles engendreraient à coup sûr de l'espace, occupant, comme on le dit pour la superfluidité ou pour la condensation de Bose-Einstein, « tout l'espace à leur disposition ». Parmi les condensats obtenus jusqu'ici, à partir de milliards d'atomes identiques « tombés » sur le même état quantique, l'un a été observé sous la forme d'un ellipsoïde. Quelle est la nature, et la structure, de ce volume fini d'espace occupé par un tel atome quantique multiple ? Il n'est pas constitué de points, puisque le même état occupe tout l'espace, d'une extrémité à l'autre, sans division possible.

On pourrait pousser l'image plus loin encore. Imaginons une « expérience par la pensée » (ou, plutôt une « observation par la pensée », puisqu'elle porte sur l'astronomie), d'un effet de condensation de Bose-Einstein : que serait un espace de dimensions astronomiques empli de matière quantique indiscernable et cohérente, telle un condensat de Bose-Einstein, sur la surface (ou dans le volume) d'une « étoile de Bose-Einstein », à supposer que de tels objets puissent exister ?

Mais, tout d'abord, l'existence de tels objets est-elle possible ? L'idée m'en était venue en réfléchissant à la propriété fondamentale des systèmes quantiques qu'est l'indiscernabilité des systèmes identiques, et qui se présente de deux façons : celle des fermions et celle des bosons, avec les propriétés de symétrie statistique correspondante (antisymétrique pour les fermions, avec pour conséquence le principe d'exclusion de Pauli ; symétrique pour les bosons, avec comme effet la condensation de Bose-Einstein de particules identiques dans un même état quantique). Le principe d'exclusion est responsable de l'organisation de la matière atomique (la classification périodique des éléments, par la répartition des électrons en états différents, dont chacun ne peut être occupé que par un seul d'entre eux), mais il détermine également la structuration d'étoiles qui sont constituées de gaz dégénérés de fermions, comme les naines blanches (électrons) et les étoiles à neutrons...

Ne pourrait-on concevoir, de manière symétrique, des objets célestes dont la structuration serait directement expliquée par la statistique de Bose-Einstein, par exemple sous la forme d'un condensat, avec ou sans effondrement de l'étoile ? Je demandai à un astrophysicien de mes amis ce qu'il pensait d'une telle éventualité. Il me répondit qu'il lui semblait, à première vue, qu'il ne pourrait « y avoir aucune région de l'Univers aussi froide que cela ». Il s'en expliquait ainsi : « Puisque le rayonnement à 3 K est partout présent, aucun système ne peut être inférieur à cette température (à moins qu'il n'interagisse ni avec les photons ni avec les neutrinos)[34]. » Toutefois, cette fin de recevoir ne me satisfaisait pas. En effet, il y a désormais au moins un endroit de l'Univers où un tel froid à été réalisé, à savoir la Terre, certes artificiellement,

34 – Marc Lachièze-Rey, communication privée par courrier électronique (8 nov. 1999).

en laboratoire, par les êtres humains, et durant un laps de temps fini (et très petit). La nature livrée à elle-même devrait pouvoir faire, pour le moins, aussi bien : on pourrait imaginer des fluctuations de température dans une région donnée de l'Univers. On pourra, certes, objecter que les interactions avec le reste du fond cosmique fossile réchaufferaient le système local, empêchant une agrégation un peu stable de grandes quantités de matière. Mais on peut aussi imaginer une production, par quelque processus naturel encore, d'une sorte de puits de basse température qui tendrait à maintenir cet agrégat dans le froid quasi absolu tandis que le réchauffement compensatoire se ferait à l'extérieur : une sorte de machine thermique naturelle, un réfrigérateur stellaire, en quelque sorte.

On pourrait encore imaginer une sorte de matière bosonique, indifférenciée, ou sous forme de paires de quarks et de gluons, dont l'état le plus bas (*zero point energy*) pourrait être le centre d'une condensation de Bose-Einstein. Des particules supersymétriques, à supposer qu'elles existent (ce que rien ne permet encore de dire), ne pourraient-elles s'agréger en étoiles, symétriquement à ce que font les particules ordinaires, et éventuellement, pour les bosons supersymétriques (contrepartie des fermions ordinaires, les nucléons et les électrons), se condenser sur un état d'énergie minimale par effet Bose-Einstein ? (et l'on pourrait envisager un condensat pris dans tout le volume de l'objet céleste). Interrogé, Pierre Fayet, dont les travaux sur les supersymétries font autorité, m'a répondu qu'une telle situation était effectivement pensable, d'autant plus que les particules supersymétriques contreparties des quarks (qui sont des fermions), à savoir les *squarks*, sont des bosons, susceptibles de se condenser de la sorte : il en avait d'ailleurs eu l'idée, brièvement évoquée dans l'un de ses premiers articles sur le sujet[35]. L'hypothèse qu'il puisse exister dans le cosmos de nombreuses étoiles de Bose-Einstein de cette sorte ne serait pas si folle...

Considérons donc une région d'espace physique, par exemple, la surface d'une étoile compacte ultrafroide, aux environs de 0 Kelvin ou au plus à la « température de Fermi », dont un point de la surface constituerait un centre localisé de condensation, et des atomes bosoniques identiques (quels qu'ils soient) refroidis et portés à l'état du « point d'énergie zéro » sur lequel ils tomberaient tous ensemble. Si cela est pensable, c'est donc possible et cela devrait exister, comme dirait Leibniz, voire cela existe déjà quelque part dans l'Univers, et il suffirait d'aller l'y observer. L'état du condensat occuperait toute

35 – Pierre Fayet, communication privée, décembre 2000. Voir Fayet [30].

l'étendue à sa disposition, peut-être toute la superficie (macroscopique) de l'étoile, l'ensemble des milliards de milliards d'atomes mobilisés étant agrégés en un seul état quantique se trouvant très largement délocalisé (aux dimensions du volume occupé).

L'espace physique occupé par le condensat serait un espace de points en phase... et l'on serait curieux de savoir de quoi un tel état aurait l'air, en comparaison à un espace physique ordinaire de superficie équivalente. On peut le concevoir comme extension à un très grand domaine spatial de ce qui a été déjà observé pour la superfluidité (un fluide parfait sans viscosité remontant les parois du vase), ou pour les premiers condensats purs de Bose-Einstein produits[36]. Pour l'observer, les physiciens pourraient s'équiper d'un télescope « analyseur quantique » braqué sur l'espace étrange de cet objet lointain, ou encore armer un vaisseau spatial en vue d'une expédition vers ces contrées du *wilderness* quantique inviolé pour en examiner de près les propriétés peu communes. On pourrait même imaginer en faire un film, en cette année 2001, « *A quantum star odyssey* », « Une odyssée vers l'étoile quantique »[37].

5.2. Et d'autres rapports d'espace...

Il est beaucoup plus difficile de se représenter l'espace des systèmes « corrélés à distance », qui est en fait, étendu à volonté, puisque les distances de corrélations peuvent y être arbitraires. Il est clair que ces corrélations ne peuvent pas se concevoir en termes de communication ou de jonction dans l'espace : comme s'il y avait une tige rigide entre les deux sous-systèmes pour les corréler, où une communication instantanée se produirait entre les extrémités. Mais ce serait revenir à l'action à distance newtonienne, et nier la relativité. En vérité, si l'on examine comme la théorie décrit de tels systèmes, elle ne fait tout simplement pas intervenir l'espace dans leur corrélation, et la relativité restreinte n'est pas concernée. Il faut admettre que cette corrélation, cette non séparabilité locale est un principe premier, ou une propriété première de ces systèmes, et qu'elle pré-existe à la considération de l'espace. Mais nous avons affaire ici à un très petit nombre de particules ou systèmes (deux, trois, pas beaucoup plus), dont la définition n'appelle pas la notion d'espace (ce qui est physiquement significatif, ce sont les variables propres de leur champ).

36 – Sur la condensation de Bose-Einstein, voir [13].

37 – Une version modernisée et futuriste des antiques peplums...

Prenons encore l'espace physique du vide quantique [16], qui n'est agité constamment de soubresauts virtuels que par manière de parler, parfaitement inexacte : les images complaisantes faisant appel à des intuitions macroscopiques sont ici totalement inopérantes. C'est la théorie quantique des champs et elle seule qui peut nous aider à nous représenter ce vide prompt à réagir aux excitations de la matière et des champs.

Considérons, enfin, l'espace de la propagation des systèmes quantiques en cohérence de phase. Le phénomène limite de la jonction de l'espace quantique avec l'espace classique est évidemment celui de la *décohérence*, récemment produit et observé en laboratoire [37], qui se produit naturellement (les caractères quantiques étant progressivement, quoique très rapidement, dissous dans le milieu ambiant des interactions), plutôt que celui de la mesure, qui est imposée d'une manière contraignante par les caractéristiques de l'appareillage classique (imposant ainsi une « réduction » des caractères quantiques à des caractères classiques choisis à l'avance) [67, 72].

On retiendra cependant ce trait, qui n'était pas acquis voici peu de temps encore, qu'un état de superposition de type quantique pour un système « mésoscopique » d'atome et de champ imbriqués a été observé se propageant sur un certain parcours fini. On pourra faire remarquer que cet état de chose est admis sans autre forme de procès en physique des particules élémentaires, avec les particules électriquement neutres comme les mésons K^0 et les neutrinos, voire les neutrons, pour de grandes distances où ces systèmes quantiques restent isolés de toute matière environnante jusqu'à leur interaction, et qu'il constitue l'un des principes d'explication de la physique des champs quantiques de ce domaine [77]. La différence tient ici à la quasi visualisation du phénomène, avec des atomes géants et des champs à la frontière du classique, phénomène qui est presque à la limite de ce qui est observable directement : un chat de Schrödinger mésoscopique, cela ne passe pas inaperçu !

Bibliographie

[1] Aspect, A., *Trois tests expérimentaux des inégalités de Bell par mesure de polarisation de photons*, Thèse de doctorat ès-sciences physiques, Université Paris-Sud, Orsay, 1983.

[2] Balibar, F., Lévy-Leblond, J.-M., *Quantique*. Tome 1 : *Rudiments*, Paris, 1998.

[3] Bell, J.S., *Speakable and unspeakable in quantum mechanics*, Cambridge University Press, Cambridge, 1987.

[4] Bimbot, R., Paty, M., Vingt cinq années d'évolution de la physique nucléaire et des particules, *in* Yoccoz, J. (Éd.), *Physique subatomique : 25 ans de recherche à l'IN2P3, la science, les structures, les hommes*, Editions Frontières, Gif-sur-Yvette, p. 12–99, 1996.

[5] Bohm, D., *Wholeness and the implicate order*, Routledge and Kegan Paul, London, 1980.

[6] Bohm, D., Hiley, B.J., *The Undivided Universe: An Ontological Interpretation Of Quantum Theory*, London, Routledge, 1993.

[7] Bohr, N., Can quantum-mechanical description of physical reality be considered complete ? *Physical Review* 48, 696–702, 1935.

[8] Bohr, N., *Atomic physics and human knowledge*, New York, Wiley, 1958 ; trad. fr. par Edmond B., Roland O., *Physique atomique et connaissance humaine*, Paris, Gauthier-Villars, nouv. éd. établie par Catherine Chevalley, Paris, Gallimard, 1991.

[9] Cassirer, E., *Substanzbegriff und Funktionsbegriff*, B. Cassirer, Berlin, 1910. Trad. angl., *Substance and function, in* Cassirer 1923 (éd. 1953), p. 1–346.

[10] Cassirer, E., *Zur Einstein'schen Relativitätstheory*, Bruno Cassirer, Berlin, 1921. Trad. angl., *Einstein's theory of relativity considered from the epistemological standpoint, in* Cassirer 1923 (éd. 1953), p. 347–460.

[11] Cassirer, E., *Substance and function and Einstein's theory of relativity*. Trad. angl. par William Curtis Swabey and Mary Collins Swabey, Open Court, Chicago, 1923 ; Dover, New York, 1953 [édition utilisée].

[12] Comte, C., Les probabilités quantiques sont-elles géométriques ?, *in* Flament, D. (éd.), *Histoire de géométries*, Textes du Séminaire de 1997, Document de travail, Maison des Sciences de l'Homme, Paris, p. 1–38, 1998.

[13] Cornell, E., Wiemann, C., The Bose-Einstein condensate, *Scientific American*, 3, 26-31, 1998. Trad. fr., La condensation de Bose-Einstein, *Pour la Science*, 247, 92–97, 1998.

[14] Darrigol, O., *From* c-*Numbers to* q-*Numbers. The classical Analogy in the History of Quantum Theory*, University of California Press, Berkeley, 1992.

[15] Descartes, R., *Regulœ ad directionem ingenii*, in Descartes, *Opuscula Posthuma*, Amsterdam, 1701 ; in *Œuvres de Descartes*, publiées par C. Adam et P. Tannery, 11 volumes (1ère éd., 1896-1913) ; nouvelle édition révisée, 1964-1974 ; ré-éd., 1996 : vol. 10, p. 349–486 ; Trad. fr. J. Sirven, *Règles pour la direction de l'esprit*, Vrin, Paris, 1970.

[16] Diner, S. and Gunzig, E., *Univers du tout et du rien*, Editions de l'Université de Bruxelles, Bruxelles, 1998.

[17] Dirac, P.A.M., *The Principles of Quantum Mechanics*, Clarendon Press, Oxford, 1930. 4th ed., 1958. Trad. fr. par Alexandre Proca et Jean Ullmo, *Les principes de la mécanique quantique*, Presses Universitaires de France, Paris, 1931.

[18] Einstein, A., *On the method of theoretical physics*, The Herbert Spencer lecture, Oxford, june l0, 1933 ; repris *in* Einstein [25], p. 263–270. Original en allemand : Zur Methodik der theorischen Physik, *in* Einstein [19], ed. 1960, p. 113–119. Trad. fr., Au sujet de la méthode de la physique théorique, *in* Einstein [20], p. 163–173 ; autre trad. fr., Sur la méthodologie de la physique théorique, in *Œuvres*, vol. 5 : Science, *Ethique, Philosophie*, Seuil, Paris, p. 102–107, 1991.

[19] Einstein, A., *Mein Weltbild*, Querido, Amsterdam, 1934. Ré-édition (modif. et augm. de textes plus récents), Herausgeben von Carl Seelig, Europa-Verlages, Zurich, 1953 ; Ullstein Bücher, West-Berlin, 1960.

[20] Einstein, A., *Comment je vois le monde*, Flammarion, Paris, 1934. (Trad. fr., par le Colonel Cros, de Einstein [19]). Nouvelle trad. par Maurice Solovine, Flammarion, Paris, 1958. Nouv. trad. par Régis Henrion, Flammarion, Paris, 1969.

[21] Einstein, A., Physik und Realität, *Franklin Institute Journal*, CCXXI, 1936, 313–347. Trad. angl., Physics and reality, *Franklin Institute Journal*, CCXXI, 1936, 349–382 ; également *in* Einstein [25], p. 283–315. Trad. fr., Physique et réalité, *in* Einstein [24], p. 66-108. Autre trad. fr., Physique et réalité, *in* Einstein [27], vol. 5 : *Science, Ethique, Philosophie*, p. 125–151.

[22] Einstein, A., Considerations concerning the fundaments of theoretical physics, *Science*, n.s., XCI, 1940 (may 24), 487–492 ; repris sous le titre The fundaments of theoretical physics, *in* Einstein [25], p. 315–326. Trad. fr., Les fondements de la physique théorique, *in* Einstein [24], p. 109–123.

[23] Einstein, A., Quantenmechanik und Wirklichleit, *Dialectica* 2, 1948, 35-39. Trad. fr., Mécanique quantique et réalité, *in* Einstein [27], p. 244–249.

[24] Einstein, A., *Conceptions scientifiques, morales et sociales*, trad. fr. par Maurice Solovine, Flammarion, Paris, l952. Reéd abrégée, trad. fr. revue par Daniel Fargue, Flammarion, Paris, 1990.

[25] Einstein, A., *Ideas and Opinions*, transl. by Sonja Bergmann, Crown, New-York, 1954. Re-éd. Laurel, New-York, 1981.

[26] Einstein, A., *The Collected Papers of Albert Einstein*, ed. by John Stachel *et al.*, puis par M. Klein *et al.*, Princeton University Press, Princeton, New Jersey, 8 vols. parus à partir de 1987.

[27] Einstein, A., *Œuvres choisies*. Trad. fr. sous la dir. de Françoise Balibar, Seuil, Paris, 6 vols, 1989-1993.

[28] Einstein, A., Podolsky, B, and Rosen, N., Can quantum-mechanical description of physical reality be considered complete ? *Physical Review*, ser. 2, XLVII, 1935, 777–780. Trad. fr., Peut-on considérer que la mécanique quantique donne de la réalité physique une description complète ? *in* Einstein [27], vol. 1 : *Quanta : Mécanique statistique et physique quantique*, p. 224–230, 1989-1993.

[29] Einstein, A. and Born, M., *Briefwechsel 1916-1955*, Nymphenburger Verlagshandlung, München, 1969. Trad. fr. par Pierre Leccia, *Correspondance 1916-1955*, commentée par Max Born, Seuil, Paris, 1972.

[30] Fayet, P., La supersymétrie et l'unification des interactions fondamentales, *La Recherche* 19, 197 (1988) 334–345.

[31] Freire Jr., O., *A emergencia da totalidade. David Bohm e a controversia dos quanta*, Tese de doutorado, Dpto de História, Universidade de Sõ Paulo, 1995.

[32] Granger, G.-G., *La pensée de l'espace*, Odile Jacob, Paris, 1999.

[33] Grangier, P., *Étude expérimentale de propriétés non-classiques de la lumière ; interférences à un seul photon*, Thèse de doctorat ès sciences physiques, Université Paris-Sud, Orsay, 1986.

[34] Griffin, A., Snoke, D.W., Stringari, S. (Eds.), Bose-Einstein condensation, Cambridge University Press, 1995.

[35] Grunbaum, A., *Philosophical problems of space and time*, Knopf, New York, 1963. Second, enlarged ed., Reidel, Dordrecht, 1973.

[36] Guénault, M., Étendue, *in* d'Alembert, Jean & Diderot, Denis (Éds.), *Encyclopédie ou Dictionnaire raisonné des sciences, des arts et des métiers*, vol. 6 (1756), p. 171–174.

[37] Haroche, S., Brune, M., Raimond, J.-M., Experiments with single atoms in a cavity: entanglement, Schrödinger's cats and decoherence, *Philosophical Transactions of the Royal Society of London* A, 355, 2367–2380, 1997.

[38] Helmholtz, H. von, *Schriften zur Erkenntnistheorie*, herausgegeben und erlaütert von Paul Hertz und Moritz Schlick, Springer, Berlin, 1921. Trad.

angl. *Epistemological writings*. The Paul Hertz/Moritz Schlick centenary edition of 1921 with notes and commentary by the editors ; newly transl. by Malcolm F. Lowe. (Ed.) with an Intoduction and bibliography by R.S. Cohen, Y. Elkana, Reidel, Dordrecht and Boston, 1978.

[39] Hilbert, D., *Grundlagen der Geometrie* (1899) ; 7e éd. avec appendices, Teubner, Berlin, 1930 ; trad. fr. (partielle) par P. Rossier, *Les fondements de la géométrie*, Dunod, Paris, 1971.

[40] Itzykson, C., Zuber, J.B., *Quantum field theory*, Mc GrawHill, New York, 1980.

[41] Kant, I., *Critik der reinen Vernunft*, J.F. Hartknoch, Riga, 1781 ; 2e ed., 1787. Trad. fr. par Alexandre J.L. Delamarre, F. Marty, *Critique de la raison pure, in* Kant, Emmanuel, *Œuvres philosophiques*, vol. 1, Gallimard, Paris, p. 705–1470, 1980.

[42] Kouneiher, J., *Épistémologie et histoire récente de la gravitation quantique*, Thèse de doctorat en épistémologie et histoire des sciences, Université Paris 7-Denis Diderot, 1998.

[43] Langevin, P., *La notion de corpuscules et d'atomes*, Hermann, Paris, 1934.

[44] Leite Lopes, J., Escoubès, B., *Sources et évolution de la physique quantique. Textes fondateurs*, Masson, Paris, 1995.

[45] Leite Lopes, J., Paty, M. (Éds.), *Quantum Mechanics Half a Century Later*, Reidel, Dordrecht, 1977.

[46] Lévy-Leblond, J.-M., Towards a Proper Quantum Theory, *in* Leite Lopes, José et Paty, Michel (Eds.), *Quantum Mechanics Half a Century Later*, Reidel, Dordrecht, 1977.

[47] Lévy-Leblond, J.-M., The Picture of the Quantum World: from Duality to Unity, *International Journal of Quantum Chemistry* 12, Suppl. 1, 415, 1977.

[48] Lévy-Leblond, J.-M., Neither Waves, nor Particles, but Quantons, *Nature* 334, 6177, 1988.

[49] Lévy-Leblond, J.-M., Quantum Theory at Large, *in* Beltrametti, E., Lévy-Leblond, J.-M. (Éds), *Adv. in Quantum Phenomena*, Plenum, New York 1996.

[50] Mach, E., *Die Mechanik in ihrer Entwicklung historisch-kritisch Dargellstellt*, Leipzig, 1883 ; autres éditions augm. de 1888 à 1933. Trad. angl. (de la 2e éd. allemande), par Thomas J. McCormack, *The science of mechanics. A critical and historical exposition of its principles*, Open Court, Chicago, 1893 ; autres éd. augm., de 1902 à 1960. Trad. fr. (de la 4e éd.

allemande) par E. Bertrand, *La mécanique. Exposé historique et critique de son développement*, Hermann, Paris, 1904 ; re-éd., 1923.

[51] Mach, E., *Erkenntnis und Irrtum. Skizzen zur Psychologie der Forschung*, J.A. Barth, Leipzig, 1905 ; 2e éd. augm., 1906 ; autres éd. : 1917, 1920, 1926. Trad. fr. (abrégée, de la 2e éd. allemande) par Marcel Dufour, *La connaissance et l'erreur*, Flammarion, Paris, 1908 ; re-éd., 1922. Trad. angl. (de la 5e éd. allemande) par Paul Foulkes et Thomas J. McCormack, *Knowledge and error*, Reidel, Dordrecht, 1976.

[52] Mach, E., *Space and geometry, in the light of physiological, psychological and physical inquiry*, transl. from the German by Thomas J. McCormack, Open Court, La Salle (Ill.), 1906.

[53] Mehra, J., Rechenberg, H., *The Historical Development of Quantum Theory*, Vol. 4, Parts 1 and 2, Springer-Verlag, New-York, Berlin, 1992.

[54] Neumann, J. von, *Mathematische Grundlagen der Quantenmechanik*, Springer, Berlin, 1932. Trad. fr. par Alexandre Proca, *Les fondements mathématiques de la mécanique quantique*, Librairie Alcan et Presses Universitaires de France, Paris, 1947.

[55] Paty, M., Symétrie et groupes de transformation dans les théories contemporaines de la matière : jalons épistémologiques, *Colloque Abel-Galois, Lille, 21-25 février 1983, Première partie*, Institut de Recherches de Mathématiques Avancées (IRMA), Lille, fasc. 5, 1985.

[56] Paty, M., La non-séparabilité locale et l'objet de la théorie physique, *Fundamenta Scientiae* 7, 47–87, 1986.

[57] Paty, M., *La matière dérobée. L'appropriation critique de l'objet de la physique contemporaine*, Archives contemporaines, Paris, xx + 442 p., 1988.

[58] Paty, M., L'inséparabilité et la mesure des systèmes quantiques, *in* Jacob, A. (Éd.) *Encyclopédie philosophique universelle*, volume 1 : A. Jacob (Éd.), *L'Univers philosophique*, Presses Universitaires de France, Paris, 1989, p. 1172–1177. Repris dans [59], Chap. 12, p. 164–175.

[59] Paty, M., *L'analyse critique des sciences, ou le tétraèdre épistémologique (sciences, philosophie, épistémologie, histoire des sciences)*, L'Harmattan, Paris, 1990.

[60] Paty, M., *Einstein philosophe. La physique comme pratique philosophique*, Presses Universitaires de France, Paris, 1993.

[61] Paty, M., Le caractère historique de l'adéquation des mathématiques à la physique, *in* Garma, S., Flament, D., Navarro, V. (Eds.), *Contra los titanes de la rutina. Contre les titans de la routine*, Comunidad de Madrid/C.S.I.C., Madrid, p. 401–428, 1994.

[62] Paty, M., Sur l'histoire du problème du temps : le temps physique et les phénomènes, *in* Klein, Etienne et Spiro, Michel (Éds.), *Le temps et sa flèche*, Editions Frontières, Gif-sur-Yvette, 1994, p. 21–58 ; 2e éd., 1995 ; Coll. Champs, Flammarion, Paris, p. 21–58, 1996.

[63] Paty, M., The nature of Einstein's objections to the Copenhagen interpretation of quantum mechanics, *Foundations of physics* 25, 183–204, 1995.

[64] Paty, M., « Mathesis universalis » e inteligibilidad en Descartes, Trad. en español por Martha Cecilia Bustamente, *in* Albis, V.R., Charum, J., Sanchez, C.H., Serrano, G. (Eds.), *Memorias del Seminario en conmemoración de los 400 años del nacimiento de René Descartes*, Academia Colombiana de Ciencias Exactas, Físicas y Naturales, Coleccion Memorias, n° 9, Bogotá, 1997, p. 135–170. Version en portugais : « Mathesis universalis » e inteligibilidade em Descartes, Trad. em português por Maria Aparecida Corrêa-Paty, *Cadernos de História e Filosofia da Ciência* (Campinas), Série 3, vol. 8, 1998, 9–57. Également, in *Seminário sobre O Cartesianismo*, Centro de Estudos de História e Filosofia da Ciência, Centros de Investigação da Universidade de Évora, Évora, Portugal, 2000, p. 145–200. [Original en français : « Mathesis universalis » et intelligibilité chez Descartes, *in* Karine Chemla, Siegmund Probst, Agnès Erdély et Antonio Moretto (Eds.), *Ceci n'est pas un festschrift pour Imre Toth*, à paraître.

[65] Paty, M., Le vide matériel, ou : La matière crée l'espace, *in* Diner, S., Gunzig, E. (Éds.), *Univers du tout et du rien*, Editions de l'Université de Bruxelles, Bruxelles, p. 22–44, 1998.

[66] Paty, M., Les trois dimensions de l'espace et les quatre dimensions de l'espace-temps *in* Flament, D. (Éd.), *Dimension, dimensions I*, Série Documents de travail, Fondation Maison des Sciences de l'Homme, Paris, p. 87–112, 1998.

[67] Paty, M., Are quantum systems physical objects with physical properties ? *European Journal of Physics* 20, 373–388, 1999 (Special issue on « Unsolved problems of physics ».)

[68] Paty, M., La création scientifique selon Poincaré et Einstein, *in* Serfati, M. (Éd.), *La recherche de la vérité*, Coll. L'Écriture des Mathématiques, ACL-Éditions du Kangourou, Paris, p. 241–280, 1999.

[69] Paty, M., Einstein et la pensée de la matière, *in* Monnoyeur, F. (Éd.), *La matière des physiciens et des chimistes*, Le Livre de poche, Hachette, Paris, p. 213–252, 2000.

[70] Paty, M., Cosmologie et matière quantique, *in* Seidengart, Jean et Szczeciniarz, Jean-Jacques (Éds.), *Cosmologie et philosophie. Hommage à Jacques Merleau-Ponty*, numéro spécial de *Épistémologiques, philosophie, sciences, histoire. Philosophy, science, history* (Paris, São Paulo) 1, 219–249, 2000.

[71] Paty, M., Interprétations et significations en physique quantique, *Revue Internationale de Philosophie* (Bruxelles) 212, 199–242, 2000.

[72] Paty, M., The quantum and the classical domains as provisional parallel coexistents, *Synthese*, 125, 179–200, 2000 (*in* French, S., Krause, D., Doria, F.A. (Eds.), *Festschrift in honor of Newton da Costa on the occasion of his seventieth birthday*).

[73] Paty, M., Physical quantum states and the meaning of probability, *in* Costantini, Domenico, Galavotti, Maria Carla and Suppes, Patrick (Eds.), *Stochastic Causality*, CSLI Publications, Cambridge University Press, 2001.

[74] Paty, M., La notion de grandeur et la légitimité de la mathématisation en physique, *in* Espinoza, M. (Éd.), *De la science à la philosophie : Hommage à Jean Largeault*, L'Harmattan, Paris, 2001, sous presse.

[75] Paty, M., La fisica cuantica, o el arrastre del pensamiento fisico por las formas matematicas, *Congreso Internacional « 100 ans de mecanica cuantica »*, Madrid, 22-25 novembre 2000 [La physique quantique ou l'entraînement de la forme mathématique sur la pensée physique].

[76] Paty, M., *Einstein, les quanta et le réel. Critique et construction théorique*, à paraître.

[77] Paty, M., The concept of quantum state: new views on old phenomena, *in* Cohen, Robert S., Howard, D., Renn, J., Sarkar, S., Shimony, A. (Eds.), *John Stachel Festschrift*, Boston Studies in the Philosophy and History of science, Kluwer, Dordrecht, in press. (Vers. fr. non publiée : Le concept d'état quantique : un nouveau regard sur d'anciens phénomènes.)

[78] Paty, M., Construction d'objet et objectivité en physique quantique, Intervention à la Journée sur la question des « conditions d'application » de l'objectivisme et du constructivisme à la description de certains phénomènes physiques ou sociaux, Université Paris 7-Denis Diderot (Centre Javelot), Paris, mardi 22 mai 2001, à paraître.

[79] Poincaré, H., *La science et l'hypothèse*, Flammarion, Paris, 1902, 1968.

[80] Poincaré, H., *La valeur de la science*, Flammarion, Paris, 1905, 1970.

[81] Poincaré, H., *Science et méthode*, Flammarion, Paris, 1908.

[82] Poincaré, H., *Dernières pensées*, Flammarion, Paris, 1913 ; reéd. 1963.

[83] Reichenbach, H., *Philosophie der Raum Zeit Lehre*, de Gruyter, Berlin,

1928. Trad. angl. par Maria Reichenbach et John Freund, *The philosophy of space and time*, Dover, New York, 1957.

[84] Reichenbach, H., *Selected writings*, Cohen R.S., Reichenbach M. (Eds.), 2 vols., Reidel, Dordrecht, 1978.

[85] Rosenfeld, L., *Selected papers*, Cohen R.S., Stachel J. (Eds.), Reidel, Dordrecht, 1979.

[86] Riemann, B., Über die Hypothesen, welche der Geometrie zugrunde liegen [Mémoire présenté le 10 juin 1854 à la Faculté philosophique de Göttingen], *Abhandlungender königlischen Gessellschaft der Wissenschaften zu Göttingen*, vol. 13, 1867. Trad. fr. par J. Houel, « Sur les hypothèses qui servent de fondement à la géométrie », *in* B.R., Œuvres mathématiques, trad. fr. par L. Laugel, avec une préface de M. Hermitte et un essai de M. Félix Klein, Paris, 1898 ; nouveau tirage, Paris, 1968., p. 280–299. Trad. angl. *in* B.R., *Collected works*, Weber H., Dedekind R. (Eds.), with supplement by M. Noether and W. Wirtinger, New York, 1953.

[87] Schilpp, P.-A., *Albert Einstein: philosopher-scientist*, The library of living philosophers, Open Court, Lassalle (Ill.), 1949, Re-éd. 1970. Trad. en allemand, *Albert Einstein als Philosoph und Naturforscher*, Kohlhammer Verlag, Stuttgart, 1955 (textes originaux d'Einstein et de Pauli).

[88] Schrödinger, E., *Space-time structure*, Cambridge University Press, Cambridge, 1950 ; repr. with corrections, 1960 ; 1988.

[89] Seidengart, Jean et Szczeciniarz, Jean-Jacques (Éds.), *Cosmologie. En hommage à Jacques Merleau-Ponty*, numéro spécial de *Épistémologiques. Philosophie, sciences, histoire. Philosophy, Science, History* (Paris, São Paulo) 1, (n° 1–2), 2000.

[90] Weyl, H., *Raum, Zeit, Materie*. [1e édition, cours professé à l'École Polytechnique fédérale de Zurich en 1917.] 3e édition remaniée, 1919. 4e édition, augmentée, 1921. Trad. fr. sur la quatrième édition allemande, par Gustave Juvet et Robert Leroy, *Temps, espace, matière. Leçons sur la théorie de la relativité générale*, Blanchard, Paris, 1922. Nouveau tirage augmenté de commentaires par Georges Bouligand, Blanchard, Paris, 1958, 1979.

[91] Weyl, H., *Gruppentheorie und Quantenmechanik*, Hirzel, Leipzig, 1928 (2e éd. 1931). Trad. angl. par H. P. Robertson, *The theory of groups and quantum mechanics*, Methuen, London, 1931 ; Dover, New York, 1950.

[92] Wheeler, J.A., Zurek, W.H. (Eds.). *Quantum theory of measurement*, Princeton University Press, Princeton, 1983.

[93] Zahar, E., On the alleged. . . Paty. . ., *Archives d'Histoire des sciences*.

À consulter : Studies in History and Philosophy of Modern Physics

Auyang, Sunny Y., Spacetime as a fundamental and inalienable structure of fields, SHPMP (2001) 205–216.

Cao, Tian Yu, Prerequisites for a consistent framework of quantum gravity, SHPMP (2001) 181-204.

De Regt, Henk W., Spacetime visualisation and the intelligibility of physical theories, SHPMP (2001) 243–266.

Dieks, Dennis, Introduction, SHPMP (2001) 151–156.

Diers, Dennis, Space and time in particle and field physics, SHPMP (2001) 217–242.

Hartmann, Stephen, Effective field theories, reductionism and scientific explanation, SHPMP (2001) 267.

’T Hooft, Gerard, Obstacles on the way towards the quantisation of space, time and matter and possible resolutions, SHPMP (2001) 157-180.

SHPMP, *Studies in History and Philosophy of Modern Physics* 32B(2), June 2001, Special Issue: *Spacetime, Fields, and Understanding Perspectives On Quantum Field Theory*.

04

Espaces et référentiels

Claude Comte

Le problème du référentiel est d'une importance capitale pour la construction de la théorie physique, car il est intimement lié à la question des conditions de validité des principes et des lois. En effet, la formulation de lois à partir de l'expérience n'est possible que dans la mesure où il existe une classe suffisamment large de situations du laboratoire, telles que les expériences soient reproductibles ou bien, ce qui revient au même, un ensemble suffisamment étendu de points de vue d'observation équivalents par rapport aux phénomènes. Les transformations spatio-temporelles de passage entre ces situations équivalentes constituent le groupe d'invariance des lois. Certaines théories physiques peuvent être reformulées dans l'esprit du programme d'Erlangen de Félix Klein : toute géométrie est caractérisée par la donnée d'un groupe d'invariance, dont les éléments sont les transformations assurant la reproductibilité des propriétés des figures. Les propriétés métriques, c'est-à-dire, les « lois quantitatives » de la géométrie, découlent de la donnée de ce groupe. L'objectif principal de cet article est :

(i) de montrer que la mutation opérée par Felix Klein peut être étendue à différents domaines de la théorie physique en amplifiant le rôle joué par les principes d'invariance ;

(ii) de faire apparaître la *structure de groupe* comme le *fondement rationnel* de l'existence de certaines grandeurs aussi bien en physique classique qu'en physique quantique ;

(iii) de décrire les propriétés des espaces propres aux objets de ces théories, définis comme l'ensemble des situations spatio-temporelles au sens large qui sont accessibles à ces objets.

La notion d’invariance est inhérente à toute théorie physique, puisque l’énoncé de lois présuppose la *reproductibilité* des phénomènes. La construction de la théorie physique exige donc au préalable de connaître l’ensemble des situations du laboratoire – ou *référentiels* – telles que cette condition soit remplie avec une précision suffisante.

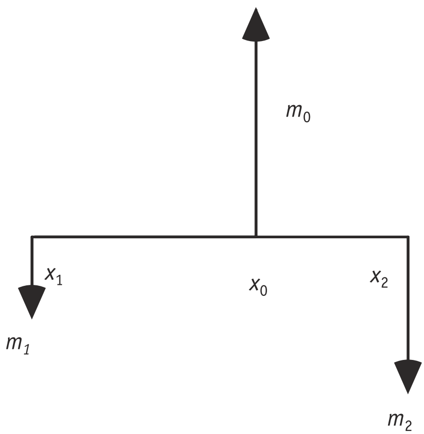
Galilée a énoncé l’impossibilité de détecter le mouvement d’un navire voguant à vitesse uniforme sur une mer calme par des expériences internes, en observant que *toutes les expériences qui pouvaient être effectuées à l’intérieur de la cabine du navire étaient reproductibles et se déroulaient de la même manière que sur la terre ferme* : le navire de Galilée réalisait ainsi, approximativement, les conditions du système physique isolé. Il en a déduit que les lois de la chute libre ne dépendaient pas du mouvement de translation uniforme, et donc que la pierre lâchée verticalement du haut du mât du navire tombait toujours au bas du mât : son raisonnement peut être considéré comme l’introduction du *principe de relativité* en physique, comme une hypothèse de nature expérimentale et à partir de laquelle de multiples conséquences pouvaient être dégagées.

Un vaisseau spatial en mouvement libre dans le champ de gravitation, et dont on a arrêté la rotation par rapport à l’environnement d’étoiles, le critère étant l’annulation de la force centrifuge à l’intérieur, constitue le mode de réalisation moderne du *référentiel galiléen*. Dans l’espace interne à ce vaisseau, nous disposons d’un laboratoire idéal, tel que les expériences – imaginons qu’elles se déroulent sur un banc expérimental – soient reproductibles par *translation spatiale ou temporelle, par rotation spatiale*, et également lorsque l’on communique au banc expérimental un *mouvement de translation uniforme* par rapport aux parois du vaisseau. Le système solaire dans son ensemble réalise avec une bonne approximation les conditions du vaisseau spatial galiléen. Nous montrons dans la première partie, en reformulant la théorie du levier, par des considérations immédiatement transposables à la chrono-géométrie et à la théorie classique des collisions de particules ponctuelles, que le groupe des invariances galiléennes entraîne une déviation par rapport aux lois d’Archimède, qui devient sensible aux grandes distances, et une forme des lois de la géométrie correspondant à un espace à courbure constante.

Ce premier modèle appelle une généralisation effectuée dans la seconde partie, où l’on expose la théorie d’un *système de grandeurs définies par un schéma barycentrique invariant sur un groupe de transformations*. On est conduit à envisager des objets tels que leur description complète requiert *plusieurs référentiels* « *complémentaires* ».

Dans la troisième partie, le modèle général est appliqué à la physique quantique. L'objectif est de démontrer qu'il est possible de construire cette théorie à partir de faits d'observation de nature qualitative concernant les probabilités, ce qui permet d'inclure l'interprétation de cette théorie dans les principes physiques énoncés, et d'éviter ainsi le recours à des considérations philosophiques extérieures à la théorie pour résoudre les problèmes d'interprétation. Il s'agit ici d'une première approche, qui sera complétée et approfondie ultérieurement. Nous montrons que les dispositifs expérimentaux permettant la manifestation des objets quantiques (des « quantons ») jouent le rôle de « *référentiels* » pour ceux-ci, et sont tout comme les référentiels galiléens doués de symétries internes.

1. Théorie de l'équilibre



Considérons l'équilibre d'un levier auquel on applique des forces perpendiculaires m_1, m_2, $-m_0$ aux points x_1, x_2, x_0 respectivement. On suppose qu'il existe deux conditions d'équilibre, qui s'expriment comme la conservation d'une fonction paire $\varepsilon(x)$, la « résultante », et d'une fonction impaire $p(x)$, le « moment des forces appliquées » :

$$m_0\varepsilon(x_0) = m_1\varepsilon(x_1) + m_2\varepsilon(x_2) \tag{4.1}$$

$$m_0 p(x_0) = m_1 p(x_1) + m_2 p(x_2). \tag{4.1'}$$

Les fonctions inconnues $\varepsilon(x)$ et $p(x)$ peuvent être déterminées directement à partir du **principe de relativité** : nous postulons soit l'invariance des lois (4.1, 4.1′) par une translation $-y$ quelconque de l'origine des coordonnées le long de la droite support du levier, soit la reproductibilité de l'équilibre par la translation $+y$; dans les deux cas, les coordonnées x sont remplacées par $x+y$, et les conditions suivantes doivent être satisfaites également

$$m_0\varepsilon(x_0+y) = m_1\varepsilon(x+y_1) + m_2\varepsilon(x_2+y) \tag{4.2}$$

$$m_0 p(x_0+y) = m_1 p(x_1+y) + m_2 p(x_2+y). \tag{4.2′}$$

Ces conditions sont extrêmement contraignantes pour la forme des fonctions $\varepsilon(x)$ et $p(x)$. En effet, des équations fonctionnelles peuvent être établies comme conditions nécessaires découlant de l'étude d'un cas particulier : des bras de levier égaux $x = x_0 - x_1 = x_2 - x_0$ et des forces égales appliquées aux extrémités : $m = m_1 = m_2$. Dans le « référentiel » où $y = 0$, le moment des forces est nul et la condition (4.1) s'écrit

$$2m\,\varepsilon(x) = m_0\varepsilon(0). \tag{4.3}$$

On a $\varepsilon(0) \neq 0$, car $\varepsilon(0) = 0 \Rightarrow \varepsilon(x) = 0\ \forall x$, et l'on pourra donc poser $\varepsilon(0) = 1$ dans la suite pour la commodité. Dans le référentiel déplacé de $-y$ les conditions d'équilibre s'écrivent

$$m[\varepsilon(y+x) + \varepsilon(y-x)] = m_0\,\varepsilon(y) \tag{4.4}$$

$$m[p(y+x) + p(y-x)] = m_0\,p(y). \tag{4.5}$$

Les quotients de (4.4) par (4.3) et de (4.5) par (4.3) donnent

$$\varepsilon(y+x) + \varepsilon(y-x)] = 2\varepsilon(x)\varepsilon(y) \tag{4.6}$$

$$p(y+x) + p(y-x)] = 2\varepsilon(x)p(y). \tag{4.7}$$

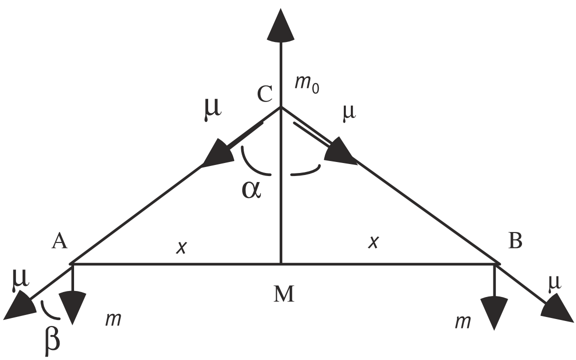
On peut démontrer que les solutions intégrables de (4.6) et (4.7) sont nécessairement continues et dérivables à tout ordre [1]. Par conséquent, ces équations fonctionnelles peuvent être ramenées à des équations différentielles. Avec un peu d'algèbre on s'assure que la solution générale est déterminée à une constante près, et l'on obtient deux types de solutions :

$$\varepsilon(x) = \cosh(x/R), \quad p(x) = R\sinh(x/R)$$
$$\text{(solutions hyperboliques)} \tag{4.8}$$

$$\varepsilon(x) = \cos(nx), \quad p(x) = R\sin(nx)$$
$$\text{(solutions elliptiques).} \tag{4.9}$$

Le recours à l'expérience est indispensable pour connaître la valeur de la constante R ou n, qui doit être considérée comme une constante structurelle de l'espace physique associé au référentiel galiléen. La loi d'Archimède n'est retrouvée que dans la limite des petites distances $x/R \to 0$ ou $n\,x \to 0$.

L'objectif est à présent de dégager les conséquences des lois de la mécanique contenant la constante structurelle R (ou n) sur la géométrie de l'espace (la matière est l'architecte de sa demeure). Pour cela, nous considérons l'expérience bidimensionnelle dans laquelle le triangle solide isocèle ACB est soumis à deux forces égales m appliquées en A et B perpendiculairement à AB. L'équilibre est assuré par l'action de la force m_0 appliquée au sommet C du triangle solide ABC, nécessairement dirigée par raison de symétrie dans la direction de la hauteur MC.



L'effet des deux forces m en A et B est identique à celui de deux forces $\mu = m/\cos\beta$ dirigées selon CA et CB. L'équilibre des lignes solides AC et BC implique que les forces appliquées en leurs extrémités sont opposées. La force m_0 en C doit donc se décomposer en deux forces de valeurs égales μ dans les directions CB et CA, d'où la condition d'équilibre

$$m_0 = 2m\cos\alpha/\cos\beta. \tag{4.10}$$

Soit M le milieu de AB. L'équilibre de la ligne solide MC, soumise à la force m_0 en C requiert l'application d'une force $-m_0$ en M, et finalement la condition d'équilibre de AB s'écrit comme précédemment :

$$m_0 = 2m\cosh x/R \quad \text{ou bien} \quad m_0 = 2m\cos nx. \tag{4.11}$$

La compatibilité des deux conditions d'équilibre ainsi obtenues entraîne la relation purement géométrique

$$\cos\alpha/\cos\beta = \cosh(x/R) \quad \text{ou bien} \quad \cos\alpha/\cos\beta = \cos nx. \tag{4.12}$$

La somme des angles du triangle ACM est égale à $\pi + \alpha - \beta$. Comme nous avons $\alpha < \beta$ dans le cas hyperbolique, $\alpha > \beta$ dans le cas elliptique, et $\alpha = \beta$ dans la limite $1/R \to 0$ (ou $n \to 0$), nous obtenons la classification suivante des géométries compatibles avec l'existence d'équilibres mécaniques reproductibles par déplacement :

(i) R réel : géométrie de Lobachevski,
(ii) n réel : géométrie sphérique de Riemann,
(iii) $1/R \to 0$ ou $n \to 0$: géométrie d'Euclide.

La constante R ou $1/n$ est la courbure de l'espace.

2. Relations barycentriques invariantes

2.1. Définitions et hypothèses

(H1) Soit une *classe C de* ***référentiels équivalents*** désignés par r, qui sont des systèmes physiques dont les situations spatio-temporelles (au sens large) sont reliées entre elles par un ***groupe de transformations*** *G continu à d* paramètres. On choisit arbitrairement dans C un « *référentiel origine* » ρ auquel on associe la transformation identique ; la situation du référentiel r est alors complètement définie par les paramètres de la transformation de passage, que l'on convient de désigner par r également, de ρ à r. On suppose que la situation et la configuration géométrique des objets physiques peuvent être décrites au moyen de vecteurs appartenant à l'espace E associé à une représentation réelle (irréductible) du groupe G. De plus, nous faisons l'hypothèse que *les référentiels et les systèmes physiques étudiés ici sont des objets de même nature et doués de symétries internes*, de telle sorte que toutes les grandeurs physiques qui leur sont attachées ne dépendent que de la donnée d'un seul vecteur pour chacun d'eux. Le vecteur décrivant la situation spatio-temporelle du référentiel r sera noté $\underline{r}$; celle-ci pourra également être repérée par l'opérateur mathématique R correspondant à la transformation physique $r \in G$ dans la représentation considérée et tel que $\underline{r} = R\underline{\rho}$.

Nous nous intéresserons particulièrement aux *fonctions de corrélation entre deux référentiels et à valeurs positives*, pour lesquelles nous convenons d'adopter la même notation quelles que soient les variables utilisées : $f(r, r') \equiv f(\underline{r}, \underline{r}') \equiv f(R, R')$: c'est le cas, par exemple, de l'énergie d'une particule classique, dont le référentiel galiléen de repos est r' (et dont la vitesse est $\underline{r}'$ par rapport à ρ) mesurée avec un appareil fixe par rapport au référentiel galiléen r (par exemple un calorimètre de vitesse $\underline{r}$ par rapport à ρ, auquel l'énergie de la particule est transférée) ; ou bien de la probabilité de transition quantique d'un état quantique de moment cinétique $\hbar m$, $m = -j$, $-j+1, \ldots, +j$, manifesté par l'appareil de Stern et Gerlach d'orientation $\underline{r}$ vers l'état $\hbar m'$ manifesté par l'appareil analogue d'orientation $\underline{r}'$. On remarque que la définition de R par la condition $\underline{r} = R\underline{\rho}$ n'est pas univoque, car on peut substituer SR à R, où $S \in g_r$, g_r étant le sous-groupe tel que $\underline{r} = S\underline{r}$: ce sous-groupe est appelé le *sous-groupe des symétries internes du référentiel* r. On s'assure aisément que $g_r = R\,g_\rho\,R^{-1}$.

De plus, R (tel que $\underline{r} = R\underline{\rho}$) dépend du choix du référentiel origine : lorsqu'on effectue le changement de référentiel origine $\rho \to \rho'$ avec $\underline{\rho} = T\underline{\rho}'$, on a $R \to RT$ et $R' \to R'T$.

Nous nous intéresserons particulièrement au cas où f ne dépend que de la *situation relative* des deux référentiels. Il découle du **principe de relativité** (étendu à l'espace de la représentation E) que :

(i) **(H2)** *la grandeur f est invariante par transport du système physique isolé constitué de l'objet observé* ($R' \to TR'$) et *de l'appareil de mesure* ($R \to TR$) : on a donc $f(TR, TR') = f(R, R')$, car les expériences de physique sont reproductibles par G dans tout système isolé ;

(ii) **(H2')** *la forme mathématique de la loi exprimant la grandeur f est indépendante du choix du référentiel origine*, car toutes les situations spatio-temporelles de référentiels de la classe C sont équivalentes. Soit donc le changement $\rho \to \rho'$ avec $\underline{\rho} = T\underline{\rho}'$; on aura donc également $f(RT, R'T) = f(R, R')$.

En posant $\varphi(S) = f(I, S)$, où I désigne la transformation identique, ces deux conditions impliquent $f(R, R') = \varphi(R^{-1}R') = \varphi(R'R^{-1})$, ou de manière équivalente $\varphi(R'RR'^{-1}) = \varphi(R)$: $\varphi(R)$ est donc une *fonction centrale sur le groupe* ; on écrira pour la simplicité $f(R)$ au lieu de $\varphi(R)$.

(À côté des fonctions centrales, il existe également d'autres grandeurs s'exprimant comme des combinaisons linéaires de fonctions centrales, et formant des sous-espaces fonctionnels invariants : l'impulsion, définie comme la dérivée de l'énergie par rapport à la vitesse en mécanique newtonienne, en est un exemple.)

(H3) Le schéma barycentrique : soit un processus très général d'interaction entre des objets physiques désignés par l'indice i ($i = 1, \ldots, i$max, où $i$max est arbitraire) ; l'état matériel et dynamique d'un objet individuel est décrit par L paramètres réels et indépendants condensés dans la notation p_i, les paramètres de « *polarisation* », qui sont mesurés dans le référentiel origine ρ. Ces objets se présentent dans l'état initial avec les « *fréquences* » N_i, nombres *entiers ou rationnels positifs* indépendants du référentiel ; ils donnent ensuite par combinaison des objets d'une seule espèce, de fréquence N_0 et de polarisation p_0 dans l'état final. Un exemple est fourni par la coalescence d'un faisceau de particules classiques de types et de vitesses différents en un faisceau homocinétique de particules d'un seul type : la fréquence est le nombre de particules d'une certaine espèce, et la polarisation comprend à la fois les composantes de la vitesse et les paramètres matériels (la masse) associés à une particule d'un type donné.

On se restreint ici au cas où la fréquence résultante est $N_0 = \sum_i N_i$ et on suppose que la polarisation p_0 est déterminée de manière univoque en fonction des données initiales p_i, N_i par les relations barycentriques

$$N_0\, w(m, r|p_0) = \sum_{i=1}^{i\max} N_i\, w(m, r|p_i), \quad m = 1, \ldots, n, \tag{4.13}$$

qui doivent être vérifiées quel que soit le référentiel d'observation r, et constituent donc un système infini d'équations dont les conditions de compatibilité, examinées plus loin, permettront de déterminer la forme mathématique nécessaire du « ***champ d'intensités*** » $w(m, r|p)$ *à valeurs positives* associé à un objet de polarisation p. Ce champ s'étend sur toutes les situations spatio-temporelles r ; les composantes du champ sont désignées par l'indice m et sont supposées vérifier la condition d'invariance **(H3')**

$$w(m, T^{-1}r|p) = w(m, r|Tp).$$

On suppose que la somme dans le membre de droite de (4.13) existe pour toute distribution physique des grandeurs p_i, N_i, et donc que la fonction $w(m, r|p)$ est *intégrable* par rapport à p, et également à r, sur toute réunion d'intervalles finis. Dans le cas d'un spectre continu de valeurs de m, nous supposons que l'hypothèse d'intégrabilité s'étend à m aussi.

Remarque : La justification des relations barycentriques à partir de faits d'observation en physique quantique est donnée dans la référence [2] : p décrit les conditions de préparation d'un objet quantique individuel – un *quanton* – et N le nombre de répétitions de l'expérience consistant à faire interagir le quanton « p » avec un référentiel r fixe, et à compter le nombre de fois où le quanton se manifeste dans l'un des *états matériels qualitativement distincts* $m = 1, \ldots, n$ correspondant au référentiel r : à l'état physique m associé au référentiel r correspond la probabilité $w(m, r|p)$ qui dépend exclusivement des conditions de la préparation initiale p. On a ainsi un ensemble statistique d'expériences identiques, dont le nombre d'éléments est N. Pour un ensemble statistique constitué du mélange de sous-ensembles statistiques de quantons dont les états initiaux sont décrits par N_i et p_i, $i = 1, \ldots, i\text{max}$, la probabilité de l'état m manifesté par le référentiel r est donnée par la *règle des probabilités conditionnelles*, ce qu'expriment bien les relations (4.13), mais on fait en plus l'hypothèse cruciale que *la forme mathématique de la probabilité reste invariable* – seules changent les valeurs des L paramètres (désignés par p_0) dont elle dépend – *et s'exprime donc toujours par la même fonction* $w(m, r|p_0)$, *quels que soient le nombre et la nature des sous-ensembles statistiques mélangés*. *Cette propriété est caractéristique des objets quantiques*, et peut être établie expérimentalement d'une manière directe ; en effet, un critère simple est donné plus loin : la matrice des $w(m, r|p)$, que l'on peut construire à partir des données expérimentales, doit être de rang L ; autrement dit, *les probabilités quantiques sont à variabilité limitée*. Le lecteur accoutumé à raisonner dans le cadre du formalisme usuel de la quantique pourra se convaincre que cette *propriété qualitative* est implicitement contenue dans la description formelle la plus générale d'un système quantique, qui est donnée non par la fonction d'onde, mais par la matrice densité : cette matrice hermitique de trace unité contient au plus $n^2 - 1$ paramètres indépendants, quel que soit le mode de préparation.

La polarisation résultante p_0 peut être interprétée comme l'ensemble des paramètres qui déterminent la *propension variable du quanton individuel à se manifester dans les différents états matériels correspondant à un référentiel r quelconque* : elle est la moyenne pondérée des *polarisations – ou propensions –* p_i avec les poids N_i. Les propensions p_i et p_0 jouent ici le rôle des points d'application des poids dans le problème analogue du levier d'Archimède.

Pour un quanton individuel dont l'état initial est défini par la propension p, il est utile de remarquer que sa manifestation ultérieure dans l'état matériel m du référentiel r correspondra à l'apparition d'un nouvel état, auquel sera associée la nouvelle propension définie par $p' = (r, m)$ qui se substituera à p. Le seul procédé permettant la mesure expérimentale du *champ de probabilités* $w(m, r|p)$ d'un quanton individuel de propension p est donc le suivant : on sépare un ensemble statistique de quantons de préparation identique p en sous-ensembles statistiques disjoints (correspondant tous au même p) sur lesquels on fait agir des référentiels de paramètres différents r_1, r_2, r_3, etc.

(H4) Il existe au moins un système d'équations *indépendantes*. pour la détermination de p_0 (L inconnues). Il est fourni par un ensemble de Γ valeurs de r différentes, auxquelles sont associées les valeurs de l'indice $m = 1, \ldots, n$.

On doit donc avoir $L = n\Gamma$. Les référentiels $r_1, r_2, \ldots, r_\Gamma$ sont dits « *complémentaires* ».

Remarque : il ressort des relations linéaires (4.1) que les grandeurs $w(m, r|p)$ sont les composantes d'un tenseur que l'on peut convenir de désigner par la notation $|p)$; l'action de la forme linéaire tensorielle $(m, r|$ donne la composante $(m, r|p) = w(m, r|p)$. Si le tenseur $|p)$ est de rang k, alors m apparaît comme une notation condensée pour l'ensemble des k indices associés à une composante particulière, et r comme un symbole désignant le repère associé au référentiel r. On note qu'il est impossible de trouver un tenseur dont toutes les composantes restent positives dans tous les systèmes de coordonnées, mais qu'*on peut construire des tenseurs tels qu'un certain sous-ensemble de composantes restent positives dans tous les systèmes de coordonnées ; nous conviendrons de les appeler « tenseurs définis positifs »*. C'est le cas de la quatrième composante du quadrivecteur énergie-impusion en relativité restreinte, par exemple : ce quadrivecteur peut être défini soit par la donnée de ses quatre composantes dans un seul système de coordonnées, soit par la donnée de sa quatrième composante dans quatre systèmes de coordonnées différents. L'existence d'une composante vectorielle de signe déterminé est une caractéristique du groupe de Lorentz orthochrone. Plus généralement, dans le cas d'un groupe G quelconque, on pourra construire un tenseur défini positif à partir d'un tenseur quelconque de rang l quelconque sur un espace de dimension s, comportant s^l composantes $A_{i,j,\ldots}$: c'est le tenseur de rang $k = 2l$ à s^{2l} composantes $B_{i,i',i''\ldots,j,j',j''\ldots} = A_{i,i',i''\ldots}A_{j,j',j''\ldots}$, dont les $n = s^l$ composantes diagonales $B_{i,i',i'',\ldots,i,i',i''\ldots} = A_{i,i',i''\ldots}A_{i,i',i''\ldots}$ sont positives dans tous les systèmes de coordonnées ; on pose $m = (i, i', i'', \ldots)$. Le tenseur $B_{i,i',i''\ldots,j,j',j''\ldots} = A_{i,i',i''\ldots}A_{j,j',j''\ldots}$ est donc complètement déterminé par la donnée de ses $n = s^l$ composantes positives dans $\Gamma = s^l = n$ systèmes de coordonnées. Le procédé de construction ci-dessus n'est cependant pas exhaustif. Nous montrerons plus loin comment le problème de la construction des tenseurs définis positifs et à valeurs bornées peut être résolu en passant par la décomposition en séries de Fourier.

(H5) L'objet générique de polarisation p est toujours une combinaison d'objets « simples », en nombre indéterminé pour le moment, et définis comme suit (ce sont les « états purs » de la « mécanique quantique usuelle ») : l'objet simple est tel que sa polarisation se réduise à la donnée d'un référentiel r' unique, appelé le « *référentiel de repos* », et d'*un état matériel* unique désigné par l'indice m' ($m' = 1, \ldots, n$). Le champ d'intensités d'un objet simple est un champ de *fonctions centrales* $w(m, r|m', r')$, *invariantes par les* *transformations de symétrie interne des référentiels r* ***et*** *r′*, et vérifiant en outre les *conditions de stabilité*

$$w(m, r|m', r) = \delta_{m,m'}. \tag{4.14}$$

Ainsi, il existe toujours au moins une décomposition

$$N_0\, w(m, r|p) = \sum_{ki} N_k\, w(m, r|m'_k, r'_k). \tag{4.15}$$

Les intensités forment un ensemble convexe dont les objets simples sont la bordure (ce point sera développé dans une prochaine publication).

2.2. Forme générale des lois de conservation

Soit la matrice $w(m, r|p)$ dont les vecteurs colonnes désignés par $|p)$ correspondent à un ensemble dénombrable de polarisations différentes arbitrairement choisies. Les vecteurs lignes $(m, r|$ sont tels que les L premiers soient associés à L équations indépendantes pour la détermination des L paramètres de polarisation inconnus p_0 par les équations (4.1) ; d'après **(H4)**, il est possible de réaliser cette condition avec au moins un choix de Γ valeurs de r différentes, l'indice m parcourant les valeurs $m = 1, \ldots, n$. Les lignes suivantes se subdivisent en sous-ensembles successifs de L lignes, qui se déduisent des L premières par la substitution de $T_i r_j$ à r_j ($j = 1, \ldots, \Gamma$), où T_i ($i = 1, 2, \ldots$) est une suite d'éléments de G. Chacun de ces sous-ensembles correspond également à L équations indépendantes : en effet, l'ensemble des référentiels Tr_j ($j = 1, \ldots, \Gamma$) est globalement équivalent à l'ensemble des référentiels r_j ($j = 1, \ldots, \Gamma$) dont il se déduit par le déplacement T. *Et il est impossible que l'on puisse observer, de leurs points de vues complémentaires, des propriétés différentes pour un même système physique, d'après le* ***principe de relativité***. L'indépendance linéaire du système d'équations, qui équivaut à l'indépendance de certaines composantes du champ d'intensités, doit être considérée comme une propriété physique.

L'ordre maximum d'un déterminant non nul que l'on peut extraire de la matrice $w(m, r|p)$ est donc L (ce qui fournit un critère simple permettant de distinguer les probabilités quantiques : la matrice est de rang L). Il s'ensuit que :

(i) L'espace Λ des vecteurs colonnes est de dimension L. Il découle alors de **(H5)** qu'il existe une base de Λ constituée de vecteurs associés à des polarisations d'objets simples, sur laquelle tout vecteur $|p)$ pourra être décomposé avec des coefficients c_k (de signes quelconques) :

$$|p) = \sum_{k=1}^{L} c_k \,|r_k, m_k). \tag{4.16}$$

On obtient ainsi un ensemble de paramètres de polarisation, sous la forme des coefficients c_k, $k = 1, \ldots, L$.

(ii) L'espace Λ^* des vecteurs lignes est également de dimension L. Toute ligne étant combinaison linéaire des L premières, on a les relations

$$\forall T \in G, \quad \big(m_j, Tr_j\big| = \sum_{k=1}^{L} (m_k, r_k|\, D_{k,j}(T), \quad j = 1, \ldots, L \tag{4.17}$$

et, pour les composantes,

$$\forall p, \quad w(m_j, Tr_j|p) = \sum_{k=1}^{L} w(m_k, r_k|p)\, D_{k,j}(T), \quad j = 1, \ldots, L, \tag{4.18}$$

où la matrice des coefficients $D_{k,j}(T)$ est inversible et indépendante de p.

La polarisation p étant fixée, nous effectuons le déplacement T' de tous les référentiels intervenant dans la relation (4.18), de telle sorte que leurs *situations relatives* restent inchangées (posant $r_2 = R\, r_1$, on a $T' r_2 = (T' R T'^{-1}) T' r_1$: les couples de vecteurs reliés par les opérateurs d'une même classe de conjugaison du groupe G sont donc dans la même situation relative). *Le principe de relativité* peut être étendu au cas où plusieurs référentiels interviennent dans l'énoncé d'une loi : *les lois de la physique sont invariantes par les changements du système de référentiels tels que les situations relatives de ces référentiels pris deux à deux soient préservées.* On a donc des relations identiques à (4.17) ou (4.18), avec les mêmes valeurs des coefficients :

$$\begin{aligned}(m_j, T'(Tr_j)| &= \sum_{k'=1}^{L} \big(m_{k'}, T' r_{k'}\big| D_{k',j}(T) \\ &= \sum_{k,k'} (m_k, r_k| D_{k,k'}(T') D_{k',j}(T).\end{aligned}$$

Par ailleurs, on a $\big(m_j, (T'T) r_j\big| = \sum_{k=1}^{L} (m_k, r_k| D_{k,j}(T'T)$. Par comparaison et compte tenu de l'indépendance des vecteurs $(m_k, r_k|$, $k = 1, \ldots, L$, on obtient les équations fonctionnelles

$$D(T_2\, T_1) = D(T_2)\, D(T_1). \tag{4.19}$$

D'où le **théorème 4.1.** *Les composantes du champ d'intensités sont les fonctions de base d'une représentation réelle du groupe d'invariance.*

On aboutit à une proposition analogue en considérant les colonnes de la matrice $w(m, r|p)$. Puisque toute colonne qui suit les L premières est une combinaison linéaire de celles-ci, et qu'il existe, comme on l'a établi, une base constituée de polarisations simples, on a les relations

$$|Tr_i, m_i) = \sum_{j=1}^{L} \left|r_j, m_j\right) D_{j',i}(T) \tag{4.20}$$

$$w(m, r|Tr_i, m_i) = \sum_{j=1}^{L} w(m, r|r_j, m_j) D_{j',i}(T). \tag{4.21}$$

Ainsi, les grandeurs physiques correspondant au système décrit par $|Tr_i, m_i)$ sont déterminées par la donnée des grandeurs attachées aux systèmes $|r_j, m_j)$. La relation (4.21) est une loi de la physique établie dans le référentiel r, qui, d'après le *principe de relativité*, est invariante par le changement de référentiel $r \to T'^{-1}r$. Cela permet d'écrire, compte tenu de **(H3')**,

$$w(m, r|T'Tr_i, m_i) = \sum_{j=1}^{L} w\left(m, r|T'r_j, m_j\right) D'_{j,i}(T). \tag{4.22}$$

On aurait abouti à la même conclusion en remarquant que l'expérience consistant à vérifier la relation (4.20) est reproductible par tous les déplacements qui préservent les situations relatives des systèmes physiques pris deux à deux. Ainsi, par une démarche analogue à la déduction de (4.19), on arrive aux équations

$$D'(T_2\, T_1) = D'(T_2)\, D'(T_1). \tag{4.23}$$

Les représentations $D(T)$ et $D'(T)$ du groupe G sont équivalentes. En effet, si l'on désigne par M la matrice carrée $L \times L$ d'éléments $w(m_i, r_i|r_j, m_j)$, posant $\overline{D}_{i,j}(T) = D_{j,i}(T^{-1})$, et compte tenu de (H3'), on obtient aisément

$$\overline{D}(T) = M\, D'(T) M^{-1}. \tag{4.24}$$

Les relations linéaires (4.18) et (4.22), avec (4.19), (4.23) et (4.24) constituent la généralisation de la transformation de Lorentz.

À partir des matrices de la représentation $D(T)$, on obtient les fonctions de base (de la variable R, avec $\underline{r} = R\underline{\rho}$) des espaces Λ^* et Λ. Ce sont des combinaisons linéaires des lignes de la matrice $D_{i,j}(R^{-1})$, avec des coefficients α_i quelconques, ou bien des colonnes de la matrice $D_{i,j}(R)$, avec des coefficients β_j quelconques : on forme ainsi des *fonctions de base contravariantes et covariantes* respectivement. Leur contraction fournit des fonctions satisfaisant la condition **(H3')**. Cette méthode est appliquée dans la référence [2]. Nous résumons ici les résultats obtenus.

Plaçons-nous par exemple dans le cas du groupe des rotations O_3. Les intensités $w(m, \underline{r}|\underline{r}', m')$ associées aux objets simples sont, comme on l'a établi, des fonctions centrales sur O_3. Celles-ci peuvent être décomposées sur un ensemble complet de fonctions centrales, désignées par $P_l(\underline{r} \cdot \underline{r}')$, qui appartiennent aux différentes représentations irréductibles désignées par l'indice l (il en existe une seule par représentation irréductible : c'est le polynôme de Legendre de première espèce d'ordre l). On obtient donc

$$w\,(m, \underline{r}|\underline{r}', m') = \sum_l \alpha_{m,l} P_l(\underline{r} \cdot \underline{r}')\beta_{l,m'} \quad \text{où} \quad m = 1, \ldots, n, \tag{4.25}$$

et $m' = 1, \ldots, n$. Ces relations doivent être inversibles. Posant $(s^{-1})_{l,m} = (\alpha^{-1})_{l,m}$, on a

$$\sum_{m=1}^{m=n} (s^{-1})_{l,m}\, w(m, \underline{r}|\underline{r}, m') = P_l(\underline{r} \cdot \underline{r}')\beta_{l,m'}. \tag{4.26}$$

On reconnaît la *valeur moyenne, mesurée dans le référentiel* $\underline{r}$*, de la grandeur physique* (d'indice l) *dont le spectre de valeurs est* $(s^{-1})_{l,m}$, et dont le *caractère vectoriel ou tensoriel.* apparaît dans le membre de droite. La condition **(H5)** de *stabilité des quantons simples* entraîne

$$\sum_{m=1}^{m=n} (s^{-1})_{l,m}\, w(m, \underline{r}|\underline{r}, m') = \sum_{m=1}^{m=n} (s^{-1})_{l,m}\delta_{m,m'} = (s^{-1})_{l,m'} = P_l(\underline{r} \cdot \underline{r})\beta_{l,m'} = \beta_{l,m'}$$

D'où l'expression générale des intensités :

$$w(m, \underline{r}|\underline{r}', m') = \sum_l s_{m,l} P_l(\underline{r} \cdot \underline{r}')(s^{-1})_{l,m'}. \tag{4.27}$$

D'où le **théorème 4.2.** *Le spectre des valeurs propres de la matrice des intensités* $w(m, r|m'r')$ *est un ensemble de* n *fonctions spéciales* $P_l(\underline{r} \cdot \underline{r}')$ *et ses vecteurs propres sont indépendants des référentiels* r *et* r'.

Cette proposition sera retrouvée au paragraphe suivant par une méthode plus puissante. Dans le cas des intensités positives [2], le choix des éléments de matrice $s_{m,l}$ est unique et correspond à la *solution quantique* ; et de plus, les spectres des grandeurs physiques $s_{m,l}$ (pour l fixe) sont des *constantes à valeurs entières* proportionnelles aux n charges électriques qu'il faut disposer à distances égales sur une ligne droite pour réaliser un monopole ($l = 0$), un dipole ($l = 1$), un quadripole ($l = 2$), etc.

3. Le théorème de Féjer-Riecz et le formalisme quantique

Féjer a publié en 1915 une démonstration due à Riecz du théorème suivant :

Tout polynôme de Fourier réel et non négatif est le carré du module d'un polynôme de Fourier du même ordre, dont les coefficients sont des nombres complexes qui ne sont soumis à aucune condition supplémentaire [5, 6].

Soit

$$w(\tau) = \sum_{m=-n}^{m=+n} c_m \exp(im\tau) \geq 0, \quad c_m = c^*_{-m} \ ,$$

et le polynôme $\Phi(z) = c_n^* + \ldots + c_1^* z^{n-1} + c_0 z^n + c_1 z^{n+1} + \ldots + c_n z^{2n}$, tel que

$$\begin{aligned} &w(\tau) = \exp(-in\tau)\Phi(\exp(i\tau)) \ ; \\ &w(\tau) > 0 \Longrightarrow w(\tau) = |\Phi(\exp i\tau)|. \end{aligned}$$

Les zéros de $\Phi(z)$ sont appariés :

$$\Phi(z_k) = 0 \Longrightarrow \Phi(z_k^{*-1}) = 0.$$

On a donc

$$\Phi(z) = c_n \prod_{k=1}^{n} (z - z_k)(z - z_k^{*-1}).$$

Pour $z = \exp i\tau$, compte tenu des égalités

$$\begin{aligned} \left|e^{i\tau} - z_k\right| \left|e^{i\tau} - z_k^{*-1}\right| &= \left|e^{i\tau} - z_k\right| \left|e^{i\tau}\right| \left|z_k^* - e^{-i\tau}\right| \left|z_k^*\right|^{-1} \\ &= \left|e^{i\tau} - z_k\right| \left|e^{i\tau}\right| \left|e^{i\tau} - z_k\right| |z_k|^{-1} = \left|e^{i\tau} - z_k\right|^2 |z_k|^{-1} , \end{aligned}$$

on obtient finalement $w(\tau) = |\Psi(\tau)|^2$, où

$$\Psi(\tau) = \left|z_1 \cdot z_2 \ldots z_n \cdot c_n^{-1}\right|^{-1/2} \prod_{k=1}^{k=n} (\exp i\tau - z_k) \qquad (\text{mod. } i\varphi).$$

Ici, $\Psi(\tau)$ est un *polynôme de Fourier complexe de même ordre que $w(\tau)$, défini à un facteur de phase $e^{i\varphi}$ près*, ce qu'exprime l'écriture mod. $i\varphi$.

On peut se convaincre que le théorème de Féjer-Riecz est applicable à un champ d'intensités $w(m, r|r', m')$ positives, pourvu que ces fonctions soient bornées, ce qui est toujours le cas pour des probabilités. Nous donnons ici les idées-clés de la démonstration, qui sera approfondie dans une publication ultérieure.

Posons $|m) = |\rho, m)$. Les invariances des fonctions centrales permettent d'introduire la notation simplifiée $(m|Tm') = w(m, r|r', m')$ où $r' = Tr$.

Soit τ le paramètre du sous-groupe continu g, monoparamétrique et simplement connexe, contenant la transformation T. Pour $T(\tau) \in g$, on pose $w_{m,m'}(\tau) = (m|T(\tau)m')$. Un théorème bien connu affirme que g est susceptible d'une paramétrisation additive, de telle sorte que les représentations irréductibles sont unidimensionnelles et vérifient l'équation fonctionnelle $f(\tau+\tau') = f(\tau)f(\tau')$; les solutions bornées sont de la forme $f_\gamma(\tau) = e^{i\gamma\tau}$, avec γ réel. Par conséquent, les $w_{m,m'}(\tau)$ sont des combinaisons linéaires des $f_\gamma(\tau)$. Ce sont donc des séries ou des intégrales de Fourier, qui peuvent toujours être approchées par un polynôme trigonométrique de la forme appropriée pour l'application du théorème de Féjer-Riecz :

(i) Toute série de Fourier peut être approchée par un polynôme trigonométrique par rapport à la variable angulaire τ, quelle que soit la précision désirée, pourvu que l'ordre du polynôme soit choisi suffisamment grand.

(ii) Si l'on impose des conditions aux limites périodiques sur un certain domaine, les fonctions $w_{m,m'}(\tau)$ sont développables en séries de Fourier. En élargissant les dimensions de ce domaine, on passe progressivement des séries aux intégrales de Fourier qui décrivent des fonctions arbitraires non périodiques.

Il existe donc une fonction désignée par $\langle m, r|r', m'\rangle$, définie à une phase près, appelée « *amplitude* », et telle que

$$w(m, r|r', m') = \left|\left\langle m, r|r', m'\right\rangle\right|^2 . \tag{4.28}$$

(À ce stade, les $\langle m, r|r', m'\rangle$ ne sont pas encore des « bra-kets de Dirac ».)

Les $\langle m, r|r', m'\rangle$ sont des polynômes trigonométriques qui peuvent être développés sur un ensemble complet de fonctions trigonométriques, dont les variables sont les paramètres de certains sous-groupes de G (dans le cas du groupe O_3, ces variables sont les angles d'Euler). *Comme ces fonctions de base se transforment linéairement dans les transformations du groupe G, il en sera de même pour les $\langle m, r|r', m'\rangle$, qui appartiendront donc à une représentation du groupe G, dont le type reste à préciser*.

Il est possible de répéter, à propos de la matrice $\langle m, r|r', m'\rangle$, dont les lignes sont les classes d'équivalence (appelées « rayons ») constituées de vecteurs $e^{i\gamma}\langle m, r|$, et les colonnes les « rayons » $e^{-i\gamma'}|r', m'\rangle$, les raisonnements effectués au paragraphe 2 sur la matrice $w(m, r|r', m')$. Comme la substitution de Tr à r opère linéairement sur les amplitudes, il en découle (nous l'admettrons ici) qu'elles sont les fonctions de base d'une *représentation complexe définie à une phase près* :

$$U(T')U(T) = e^{i\varphi} U(T'T). \tag{4.29}$$

C'est le principe mathématique posé par H. Weyl [7] dans sa formulation de la mécanique quantique à partir de la théorie des groupes. Notre démarche permet de le justifier par des raisonnements physiques. La dimension de U est $L' = n\Gamma'$, avec $\Gamma' = 1$, parce qu'il n'est pas nécessaire (comme c'était le cas pour les probabilités) de surmonter l'inaccessibilité physique de certaines composantes tensorielles (due à la condition de non-négativité) par le recours à plusieurs référentiels complémentaires. Par suite, la représentation D associée aux intensités est le produit direct $D = U \times U^*$. Il en découle (nous l'admettrons) que $L \leq n^2 - 1$. *Il n'existe aucune contrainte supplémentaire sur U*, de telle sorte que la construction des représentations complexes vérifiant (4.29) permet d'obtenir tous les champs de probabilité quantiques possibles.

Pour les probabilités, nous avons en plus la relation

$$\forall T \in G, \quad \sum_m w(m, r|r', m') = \sum_m w(m, r|Tr', m') = 1. \tag{4.30}$$

Montrons qu'elle entraîne l'existence d'un *produit scalaire invariant* : on a en effet $\sum_i |x_i|^2 = \sum_i |x_i'|^2$ lorsqu'on pose $x_i = \langle m, r|r', i\rangle = |x_i|e^{i\alpha_i}$ et $x_i' = \langle m, r|Tr', i\rangle = |x_i'|e^{i\alpha_i'}$. Soit la transformation T appartenant à une famille de transformations connectées continûment avec l'identité. Dans la limite $T \to I$, on a $|x_i'| \to |x_i|$ et $\alpha_i' \to \alpha_i$. L'action *linéaire* de $U(T)$ sur les amplitudes s'écrit dans cette limite

$$\left|x_j'\right| e^{i\alpha_j'} = \sum_k \left(\delta_{j,k} + i\,\eta\,\tilde{U}_{j,k} + \ldots\right) |x_k|\, e^{i\alpha_k}, \text{ où } \eta \ll 1. \tag{4.31}$$

Et l'on établit aisément que la matrice $\tilde{U}_{j,k}$ est hermitique :

$$\tilde{U}_{j,k} = \tilde{U}^*_{k,j}.$$

L'opérateur $U(T)$ associé à $T(\tau)$, τ réel, peut être construit par itération de la transformation infinitésimale précédente N fois ($N \to \infty$), en posant $\eta = \tau/N$. On obtient par ce procédé l'opérateur *unitaire*

$$U(T) = \lim_{N\to\infty} \left(I + i\frac{\tau}{N}\tilde{U}\right)^N \equiv e^{i\tau\tilde{U}} \qquad (\text{mod. } i\varphi). \tag{4.32}$$

Ainsi, l'amplitude $\langle m, r|r', m'\rangle$ devient le produit scalaire du vecteur colonne (ou « ket ») $|r', m'\rangle$ et du vecteur ligne (ou « bra ») $\langle m, r|$; la condition $w(m, r|r, m') = \delta_{m,m'}$ entraîne alors les *relations d'orthogonalité*

$$(m, r|r, m') = \delta_{m,m'} \qquad (\text{mod. } i\varphi). \tag{4.33}$$

Dans une base de « kets », les éléments de la matrice unitaire U s'écrivent

$$U_{m,m'}\left(R^{-1}R'\right) = \left\langle m, r|r', m'\right\rangle \qquad (\text{mod. } i\varphi), \tag{4.34}$$

de telle sorte que l'on retrouve la proposition connue sous le nom de « principe de superposition » :

$$|r, m\rangle = \sum_{m'} \left|r', m'\right\rangle \left\langle m', r'|r, m\right\rangle \qquad (\text{mod. } i\varphi). \tag{4.35}$$

Le symbole r correspond donc à ce que l'on appelle couramment une « représentation » dans le formalisme usuel de la mécanique quantique. Choisissons la « représentation » particulière qui correspond à notre référentiel origine ρ, et posons $|\rho, m\rangle = |m\rangle$, $\langle m, \rho| = \langle m|$. Désignons par $\hat{R}$ l'opérateur *unitaire* associé à la transformation physique R telle que $r = R\rho$. D'après (4.34), on a alors pour les éléments matriciels (à un facteur de phase près) :

$$\left\langle m, r|r', m'\right\rangle = \left\langle m \left| \left(\hat{R}^{-1}\hat{R}' \right| m'\right\rangle\right) \equiv \left\langle m \left| \hat{R}^{-1}\hat{R}' \right| m'\right\rangle \equiv \left\langle m \left| \hat{R}^{+}\hat{R}' \right| m'\right\rangle. \tag{4.36}$$

Considérons maintenant le sous-groupe g_ρ des *symétries internes du référentiel* ρ, et soit $T \in g_\rho$. La condition d'invariance des intensités, $w(m', r'|\rho, m) = w(m', r'|T\rho, m)$, est équivalente à l'invariance des quantités $|\langle m'|\hat{R}'^{-1}\hat{T}|m\rangle| = |\langle m'|\hat{R}'^{-1}|m\rangle|$, quels que soient m' et R'. Il s'ensuit que les vecteurs $|m\rangle$ et $\hat{T}|m\rangle = e^{i\tilde{T}}|m\rangle$ sont identiques à un facteur de phase e^{it} près, t réel. Par conséquent $\tilde{T}|m\rangle = t|m\rangle$, et il est alors facile de démontrer que les opérateurs $\tilde{T} \in g_\rho$ sont commutatifs.

D'où le **théorème 4.3.** *Les kets associés aux états matériels qualitativement distincts ($m = 1, \ldots, n$), correspondant au référentiel (r), sont les vecteurs propres communs à l'ensemble des opérateurs hermitiques commutatifs associés aux symétries internes de ce référentiel. Afin d'assurer l'existence d'une grandeur et d'en autoriser la mesure, le référentiel doit nécessairement présenter la symétrie correspondant à cette grandeur.*

Il reste à établir la relation avec les considérations de la seconde partie de cet article. Pour cela, on effectue la décomposition de $D = U \times U^*$ en somme directe de représentations irréductibles :

$$w\,(m, r|m', r') = U^{(j)}_{m,m'}\left(R^{-1}R'\right) U^{(j)*}_{m,m'}\left(R^{-1}R'\right) = \sum_{l=0}^{l=2j} s_{m,l}\, P_l\,(\underline{r} \cdot \underline{r}')\, s^{-1}_{l,m'}$$

Cette expression est de la forme annoncée par le théorème 4.2. On trouve alors que *les vecteurs propres de la matrice des probabilités de transition sont les coefficients de Clebsch-Gordan* $s_{m,l} = s^{(-1)}_{l,m'} = \langle j, j, m, -m|l, 0\rangle$.

4. Conclusion

La possibilité de déduire le formalisme de la mécanique quantique à partir d'hypothèses physiques a été démontrée grâce au théorème de Fejér-Riecz, qui jette un pont entre deux rives : la probabilité d'une part, qui se trouve du côté de l'expérience, et l'amplitude d'autre part, qui est l'outil mathématique le plus commode. Nous avons seulement donné ici les premiers éléments d'un programme de recherche, dont l'objectif sera de reformuler la théorie quantique comme une physique statistique d'une nature particulière, traitant d'objets – doués de propriétés reproductibles et indépendantes des observateurs – qui ne sont autres que des *champs de probabilités géométriques* intégralement conditionnés par la symétrie des systèmes physiques, les « référentiels » qui les génèrent. Il sera intéressant d'explorer tout ce qu'implique l'existence du schéma barycentrique : quel statut particulier ce mode de définition confère-t-il aux probabilités quantiques, qui les distingue des probabilités classiques définies comme des mesures sur des ensembles ? Est-il possible de déduire directement l'impossibilité des variables cachées et des grandeurs « contrafactuelles » ? Une fois ce programme accompli, il restera encore à percer un mystère encore plus grand : par quel processus – brisure de symétrie à l'échelle macroscopique ? – les objets-champs de probabilités quantiques peuvent-ils s'organiser collectivement pour former la substance matérielle des « référentiels » présentant les symétries macroscopiques requises pour assurer l'existence de ces champs ?

Bibliographie

[1] Comte, C., *Eur. J. Phys.* 7, 225 (1986).
[2] Comte, C., *Riv. Nuovo Cimento*, 111 B, N. 8, 937 (1996).

[3] Vilenkin, N. Ja., *Fonctions Spéciales et Théorie de la Représentation des Groupes*, Dunod, Paris, 1969.

[4] Hamermesh, M., *Group Theory and its Application to Physical Problems*, Addison Wesley, 1964.

[5] Fejér, L., J. Reine und Angewandte Mathematik, 146, 53 (1915).

[6] Riecz, F., Sz-Nagy, B., *Vorlesungen über Funktionalanalysis*, VEB Verlag der Wissenschaften, Berlin, 1973, p. 108-109.

[7] Weyl, H., *Gruppentheorie und Quantenmechanik*, Hirzel, Leipzig, 1928.

05

Nouvelles dimensions mathématiques et épistémologiques du concept d'espace en physique, de Riemann à Weyl et à Witten

Luciano Boi

« La question de la validité des hypothèses de la Géométrie dans l'infiniment petit est liée avec la question du principe intime des rapports métriques dans l'espace. »
(B. Riemann, *Sur les hypothèses qui servent de fondement à la géométrie*, 1854)

« ... A geometry comes into being, which explain not only the phenomenon of gravity, but also of the electromagnetic field. According to the established theory, both phenomena arise from the same source. In these theory, all physical magnitudes have a geometrical meaning. »
(H. Weyl, *Gravitation and Elektrizität*, 1918)

« It seems reasonable to think that general relativity and quantum theory are intrinsically incompatible and that, rather than merely developing technique, what is required is some fundamental breakthrough in our understanding of the relationship between space-time structure and quantum process. »
(E. Witten, *Physics and Geometry*, 1987)

1. Introduction : rappels historiques et mathématiques

Nous voudrions proposer ici une analyse, d'abord du concept mathématique d'espace tel qu'il a été exploré, depuis la découverte des géométries non euclidiennes, par Gauss, Lobatchevski et Riemann, ensuite, des nouvelles conceptions géométriques élaborées notamment par H. Poincaré, E. Cartan et H. Weyl concernant la structure de l'espace-temps et la nature de l'espace physique. Il s'agira en même temps de montrer la nécessité à la fois historique et épistémologique d'une étude renouvelée des grands concepts géométriques et physiques développés aux XIXe et XXe siècles, comme ceux de variété, groupe, courbure, connexion et formes spatiales.

C'est à partir de ce travail herméneutique sur la pensée mathématique de l'espace, que l'on cherchera à dégager la véritable dimension et signification philosophiques de ces concepts. Aussi, on sera amené à considérer le rapport en quelque sorte fondateur que la géométrie entretient avec notamment la physique. La question de la géométrisation des phénomènes naturels, et notamment la question des rapports entre géométrie et physique, retiendra tout particulièrement notre attention. Dans ce contexte, on discutera les idées de Riemann et Clifford, puis d'Einstein, Weyl et Cartan. Commençons par mettre en évidence trois points importants ayant trait au développement mathématique du concept d'espace.

1.1. La découverte de différents modèles d'une même géométrie

Le premier point est que la découverte de différents modèles de géométrie non euclidienne, qu'ils soient isomorphes[1], ou pas, ont permis de montrer deux propriétés mathématiques importantes.

(i) Premièrement, il existe une pluralité de modèles d'une même théorie, ou une catégorie d'objets mathématiques équivalents ; ou encore, ce qui revient au même, une pluralité d'interprétations sémantiques compatibles d'un même domaine d'objets mathématiques.

(ii) Deuxièmement, ces mêmes objets mathématiques ne peuvent pas être pensés comme des entités absolues existant dans un univers platonique dont l'objectivité serait donnée avant toute intuition, idéalisation et formalisation. Ces opérations se révèlent en réalité nécessaires afin d'appréhender les modes d'existence et les diverses significations de ces objets mathématiques.

La construction de ces modèles nous apprend en effet qu'un même objet mathématique, non seulement peut être susceptible de plusieurs interprétations différentes, toutes éventuellement consistantes, mais il peut également connaître différents niveaux d'actualisation. On en veut par preuve le fait, que Leibniz avait déjà bien compris, que ces différentes interprétations, loin d'avoir une simple signification linguistique ou symbolique, traduisent le fait

1 – On dit que deux théories ou deux modèles sont *isomorphes* quand on peut traduire l'un dans l'autre à l'aide d'un dictionnaire sans que les relations essentielles entre les objets de la théorie et leur signification changent.

qu'il existe une pluralité de mondes réels possibles, ou différents états du même monde, par exemple de notre univers (voir plus loin pour quelques exemples).

Observons en outre qu'en dépit de cette pluralité et diversité de modèles, les mathématiques possèdent une unité profonde, et ce, essentiellement pour deux raisons. La première de ces raisons tient à ce que tout objet ou concept mathématique n'est pas une entité qui existe séparément à d'autres entités, à savoir comme un objet en soi, ou quelque chose d'absolument singulier, mais bien plutôt en relation avec d'autres objets et concepts mathématiques ; cette relation est à la fois réelle et structurelle (ou formelle). Pour éclairer ce qui précède, citons deux exemples.

1. On sait qu'il est possible de caractériser n'importe quelle figure géométrique, soit par ses propriétés métriques, soit par ses propriétés projectives, ou encore, par ses propriétés topologiques. Si on se place, par exemple, dans le cas projectif, il est facile de montrer que l'on peut passer des premières de ces propriétés aux deuxièmes en introduisant de nouveaux objets géométriques appelés *coniques* et *quadriques*, et également, certaines opérations qui permettent de relier les propriétés métriques à cet espace nouveau donné, en l'occurrence, par une conique ou par une quadrique. On obtient ainsi de nouvelles propriétés, appelées *métrico-projectives*, pour les objets géométriques que l'on a considérés au départ, qui sont différentes des propriétés euclidiennes usuelles.
2. Selon qu'un objet géométrique est plongé (je reviendrai dans un instant sur les concepts de *plongement* et d'*immersion* qui jouent un rôle essentiel notamment en topologie) dans l'espace euclidien R^3 ou dans l'espace a quatre dimensions R^4, ou encore dans un espace de dimension supérieure à 4, ses propriétés se conservent (soit entièrement, soit en partie) invariantes, et d'autres peuvent apparaître.

Le second point appelle des considérations plus approfondies. À ce propos, prenons l'exemple de l'objet topologique nœud. En termes mathématiques, un *nœud* est l'image d'un plongement d'une courbe C^1 ou d'une sphère S^2 dans l'espace R^3. En des termes plus abstraits, il peut être défini comme un sous-ensemble K d'un espace X, lorsque K est homéomorphe à une sphère S^p, où p est ≥ 2. K est une chaîne si K est homéomorphe à un espace qui résulte de l'union disjointe de $S^{p_1} \cup \ldots \cup S^{p_r}$ d'une ou plusieurs sphères. Deux nœuds ou chaînes K, K' sont (topologiquement) *équivalents* s'il existe un homéomorphisme $h : X \to X$ tel que $h(K) = K'$; autrement dit, $(X, K) = (X', K')$.

Il convient maintenant de donner quelques définitions élémentaires, utiles pour mieux comprendre ce qui va suivre.

Définition 5.1. On dénote l'espace à n dimensions $R^n = (x_1, x_2, \ldots, x_n)$, où chaque x_i est un nombre réel ; on désigne alors R^4 par (x, y, z, w). La sphère à n dimensions S^n est définie par la relation $S^n = (x_1, x_2, \ldots, x_n) \in R^{n+1}/(x_1^2 + \ldots + x_1^{n+1} = 1)$. La boule à n dimensions B^n est le sous-ensemble $B^n = (x_1, x_2, \ldots, x_n) \in R^n/(x_1^2 + \ldots + x_1^n \langle 1)$. S^n est alors le bord de B^{n+1}, et la sphère à 1 dimension, S^1, est le cercle unité dans le plan, c'est-à-dire donc le bord de B^2.

Définition 5.2. Puisque nous adoptons ici un point de vue topologique, par « un cercle dans R^3 » il faut entendre n'importe quel ensemble (ou espace) homéomorphe à un cercle et non pas nécessairement à une courbe circulaire tracée dans le plan. « Un cercle dans R^3 » peut dès lors être défini comme *un chemin fermé qui ne présente pas d'intersections*.

Définition 5.3. Une surface sera ici toujours considérée comme une sous-variété compacte de dimension 2. Ce concept remonte au mathématicien allemand Bernhard Riemann (1826-1866). Par exemple, la *bande de Möbius* est une surface dont le bord est un cercle.

Définition 5.4. Soient X et Y deux espaces topologiques (tous les espaces dont nous venons de parler sont des espaces topologiques ou des espaces d'Hausdorff) qui possèdent les propriétés suivantes : (i) Un *plongement* de X dans Y est une application continue bijective qui envoie X sur Y – c'est-à-dire un homéomorphisme de X dans $f(X)$. On appelle l'espace Y *l'espace ambiant du plongement*. Si X et Y sont deux variétés suffisamment lisses, une immersion de X dans Y est la donnée d'une application lisse, de X dans Y, qui est localement (mais pas forcément globalement) bijective. On peut également définir une immersion comme une application lisse dont la différentielle, df, est partout non singulière. On appelle alors *variété immergée* l'image d'une immersion. Par exemple, le limaçon, c'est-à-dire le graphe en coordonnées polaires de $r = 1/2 + \cos\theta$, est le résultat d'une immersion d'un cercle dans le plan.

Définition 5.5. Voici une dernière définition, essentielle pour comprendre tout ce qui précède. D'un point de vue topologique, une *relation d'équivalence* fondamentale entre des sous-ensembles (ou entre des sous-espaces appartenant à un espace de dimension plus élevée) est celle d'*isotopie*.

Deux sous-ensembles A_1 et A_2 de R^n seront dit isotopes s'il est possible de se donner une application (ou plus intuitivement, un mouvement élastique) dans R^n qui amène A_1 à coïncider avec A_2. D'une manière plus générale, soit A un objet quelconque, on peut alors considérer une famille continue de plongements, $f_t : A \rightarrow R^n$ avec $f_1(A) = A_1$ et $f_2(A) = A_1$, où $t \in I$, I étant l'intervalle unité. Une telle application correspond à une isotopie de A. On dira donc que deux sous-ensembles ou sous-espaces de R^n sont équivalents s'il existe une isotopie entre eux.

Afin d'illustrer intuitivement le concept d'isotopie considérons un pneu de forme annulaire qui, en évitant tout éclatement et/ou toute déchirure, peut prendre la forme topologique d'un tore ou d'une sphère de Riemann à un trou. C'est là un exemple « concret » d'isotopie entre deux sous-variétés de R^n : l'anneau et le tore. L'idée de la démonstration consiste à montrer les deux surfaces conservent leur propre degré de connexité une fois qu'elles sont soumises à une application isotope. Mais tandis que l'anneau et le tore sont multiconnexes, la sphère, elle, est simplement connexe : la preuve est que la sphère et le tore ne sont pas isotopes.

Terminons ces brèves remarques en donnant quelques résultats relatifs à la théorie des nœuds.

(i) Il y a des exemples de cercles noués dans R^3. De plus, on peut nouer un cercle dans R^3 de plusieurs manières différentes. Par contre, il paraît impossible de nouer des cercles dans le plan. Pour des raisons différentes, il est impossible d'obtenir des cercles noués dans R^4. En conclusion, seulement l'espace à trois dimensions R^3 permet de nouer des cercles.

(ii) Par ailleurs, il est impossible de nouer une sphère dans R^3, tandis que dans R^4 on peut nouer des sphères non triviales. En revanche, il est impossible de nouer une sphère dans R^5. On voit donc que l'espace R^4 est la seule dimension qui permet de nouer des sphères.

1.2. Géométrie et relativité

Nous voudrions maintenant parler de la naissance de la géométrie différentielle moderne grâce à Gauss et surtout à Riemann, naissance qui coïncide avec l'introduction du concept mathématique général de variété à n dimensions. Ce concept a permis une refonte, non seulement du concept d'espace et donc de la géométrie, mais des mathématiques dans leur ensemble. Avec celui de groupe, introduit par F. Klein dans son célèbre *Programme d'Erlangen* de

1870, le concept de variété a marqué le développement des mathématiques au XIX^e siècle. Ces deux concepts sont profondément liés car, on peut toujours définir un groupe en termes de variété différentiable et, inversement, faire agir sur cette dernière tel et tel groupe de transformations.

Mais il faut aussi souligner que leur importance n'est pas que mathématique. En fait, il est apparu de plus en plus clairement qu'ils sont en même temps des principes d'intelligibilité des phénomènes naturels, tant à l'échelle microscopique que macroscopique. La relativité générale d'Einstein, par exemple, est fondée sur le concept de variété différentiable de Riemann qui, rappelons-le, peut être douée d'une métrique non euclidienne (hyperbolique, elliptique ou autre), et posséder des objets géométriques un peu plus compliqués que l'on appelle des tenseur de courbure. Ce sont des objets mathématiques qui ont aussi une signification physique, puisqu'ils correspondent aux potentiels gravitationnels de la relativité générale. Cela veut dire, en d'autres termes, que les propriétés des phénomènes qui se déroulent à l'échelle de notre Univers reposent sur la structure géométrique et topologique d'une variété pseudo-riemannienne, dont l'espace-temps d'Einstein constitue le modèle physique par excellence.

On peut donc affirmer que la géométrie et la topologie différentielles apparaissent encore aujourd'hui comme l'un des langages les plus universels pour comprendre la réalité physique. Ceci à toute échelle : la notion de connexion, par exemple, est essentielle aussi bien pour construire des modèles d'Univers que pour décrire l'intérieur du proton et, plus en général, les particules élémentaires et leurs interactions. Nous reviendrons sur ce thème lorsqu'on parlera de la géométrisation de la physique, et notamment de la relativité générale et de la théorie quantique des champs de jauge.

1.3. Groupes invariants et symétries physiques

Faisons quelques remarques sur le concept de groupe, développé autour des années 1870 par S. Lie et F. Klein dans le cadre des nouvelles géométries non euclidiennes, après qu'E. Galois et C. Jordan l'eurent déjà introduit en algèbre et en analyse respectivement. La théorie abstraite des groupes se développa rapidement, d'abord au sein des mathématiques pures, et ce grâce notamment à Killing, E. Cartan et H. Weyl, ensuite en physique grâce en particulier à Poincaré, qui a compris le premier que l'on pouvait faire reposer entièrement la relativité restreinte d'Einstein sur un groupe de transformations à dix paramètres (dit « groupe de Poincaré ») qui agit sur l'espace de

Minkowski, en l'absence totale de gravitation. Ce groupe définit donc une géométrie de l'espace-temps, et c'est cette géométrie qui a permis de décrire les phénomènes physiques étudiés par la relativité restreinte.

Le rôle fondamental de la théorie des groupes en physique a été mis en évidence, entre autres, par E. Nother, E. Wigner et H. Weyl. Ces deux derniers ont tout particulièrement étudié les groupes qui interviennent en mécanique quantique. On distinguera généralement deux grandes catégories de groupes. À la première appartiennent les groupes globaux de Lie, qui s'appliquent à l'ensemble des transformations (ou des symétries) de l'espace-temps ; ils sont donc de nature géométrique. Les groupes de la seconde catégorie sont non compacts et de dimension infinie, et servent à décrire un certain nombre de variables ou de propriétés physiques, comme la phase et le moment cinétique, qui caractérisent les particules de matière et leurs interactions. À la différence des premiers, ces groupes, généralement appelés « groupes quantiques », sont de nature dynamique et ils sont formés de « symétries internes » au lieu qu'uniquement spatiales.

Le théorème de Nother, du nom de la mathématicienne qui l'a énoncé en 1918, a joué un rôle fondamental dans la compréhension du rapport profond qui existe entre géométrie et physique ; loin d'être contingent, ce rapport doit être expliqué par la nature des objets géométriques d'une part, et des grandeurs physiques de l'autre. En gros, le théorème de Nother dit que, étant donné un système physique mécanique, relativiste ou même quantique, on peut associer au groupe dynamique qui représente l'ensemble des symétries directement liées aux grandeurs physiques du système, une propriété de conservation. Autrement dit, à chaque groupe continu de symétries correspond une loi de conservation d'une quantité physique ; ainsi, par exemple, la conservation de l'énergie correspond à l'invariance de la théorie par rapport aux translations temporelles, et la conservation du moment cinétique à son invariance par rapport aux rotations. Les lois de conservation sont donc essentielles pour l'étude des systèmes physiques. Qui plus est, elles jouent un rôle fondamental dans la constitution même des phénomènes naturels.

Il convient peut-être de donner une idée au moins intuitive de ce qu'est une symétrie. On dit qu'on objet est symétrique si on peut lui faire subir une ou plusieurs transformations qui commutent entre elles, sans que cela modifie ses propriétés essentielles, géométriques et/ou physiques. Par exemple, on peut faire subir une rotation de 60 degrés à un flocon de neige sans modifier son apparence. On peut d'ailleurs le faire tourner d'un angle multiple de 60 degrés, ou lui faire subir successivement plusieurs rotations, et le résultat sera le même. Déjà ce simple exemple, qui peut toutefois facilement se généraliser à des situations physiques beaucoup plus complexes où interviennent un grand nombre de paramètres, montre qu'un même objet ou phénomène peut demeurer invariant vis-à-vis de plusieurs transformations différentes. Dans un tel cas, on dira que l'ensemble de ces transformations possède une structure mathématique de groupe et qu'il forme le groupe de symétries de l'objet. Philosophiquement, les principes de conservation expriment bien le fait qu'il doit y avoir de l'ordre caché dans le désordre apparent, du permanent et du stable dans ce qui apparaît changeant et instable.

2. Les modèles non euclidiens d'espace et d'espace physique

Il est maintenant bien connu que certaines propriétés ou « vérités » qu'Euclide avait attribuées à l'espace étaient loin d'être vraies et évidentes. Une de ces vérités occupait une place tout à fait spéciale dans son système, il s'agit du 5e postulat de ses *Éléments* : « *Si on a une droite qui rencontre deux autres droites situées dans le même plan, et qui fait d'un même côté des angles intérieurs dont la somme résulte moindre que deux angles droits, alors les deux droites prolongées indéfiniment se rencontrent du côté dont la somme est inférieure à deux angles droits*. » (On peut donner au moins deux définitions équivalentes de ce même postulat : « *par un point extérieur à une droite donnée dans le plan, on ne peut tracer qu'une et une seule passant par ce point et parallèle à cette droite* » ; « *la somme des angles intérieurs d'un triangle est égale à deux angles droits* ».)

Or on sait qu'en niant la vérité de ce postulat, on arrive à une nouvelle géométrie, non euclidienne ou de Lobatchevski (du nom du mathématicien russe qui l'a inventée au début du XIXe siècle), qui non seulement est parfaitement cohérente sur le plan mathématique, mais également fondée d'un point de vue physique. En effet, c'est sur une telle géométrie que repose la relativité restreinte d'Einstein.

Dans la géométrie non euclidienne de Gauss et Lobatchevski, la somme des angles d'un triangle est toujours inférieure à deux angles droits, contrairement à la géométrie d'Euclide où cette somme est toujours égale à 180 degrés. De plus, dans la géométrie non euclidienne la différence par rapport à deux angles droits est proportionnelle à l'aire de la surface. Cela montre bien que la somme angulaire n'est pas une grandeur mathématique absolue, mais qu'elle dépend de la nature et de la dimension de l'espace dans lequel on se place, et également, de la forme de l'univers physique dans lequel on effectue les mesures. Les « habitants » d'un univers de Lobatchevski peuvent tracer au moins deux lignes parallèles passant par n'importe quel point extérieur à une ligne droite (le diamètre) d'un cercle non euclidien. Dans ce cercle, les lignes droites sont représentées par des arcs de cercles qui rencontrent la limite ou la circonférence du cercle à angle droit. Ils peuvent en outre y tracer des triangles courbes formés d'arcs de grands cercles, et montrer que leur somme angulaire est une quantité qui varie en fonction de l'aire de la surface. On voit par là qu'une propriété de mesure concernant en l'occurrence la somme des angles dépend d'une propriété pour ainsi dire qualitative qui est la forme de la surface.

Ce n'est là que le début d'un long développement de la pensée mathématique, qui a permis des résultats fondamentaux notamment en physique mathématique. Un de ces résultat affirme (nous en donnons pour le moment un énoncé intuitif) que « *des théories physiques — par exemple certaines théories physiques des champs quantiques — sont indépendantes de la métrique de l'espace, c'est-à-dire que, quelle que soit la forme métrique adoptée, les propriétés essentielles de ces théories physiques (et notamment leur symétries) demeurent invariantes, alors que ces mêmes théories physiques dépendent de la structure topologique qui caractérise l'espace-temps dans lequel elles se déroulent* ».

Une des raisons qui fait l'intérêt de la géométrie non euclidienne de Lobatchevski (ou hyperbolique selon la nomenclature qui remonte à F. Klein), est le fait que son groupe de symétries (le groupe de Poincaré) correspond exactement au groupe de Lorentz des rotations ou des transformations linéaires de l'espace-temps laissant invariant un cône de lumière, qui est le groupe de la relativité restreinte, ou le groupe de symétries des cônes de lumière de la relativité.

Le postulat d'Euclide est également contredit par la géométrie elliptique de Riemann (à ne pas confondre avec sa géométrie différentielle). En fait, si on prend un espace elliptique, ou une sphère de Riemann finie et à bords comme modèle de notre Univers, on constate d'abord que deux grands cercles de la sphère, qui remplacent les droites euclidiennes, peuvent renfermer un espace, ensuite, que par un même point pris sur la circonférence de la sphère peuvent passer plus qu'un grand cercle (en fait, une multitude de grands cercles), c'est-à-dire de droites euclidiennes.

On voit ainsi que le modèle d'espace conçu par Riemann est « fermé » et « ouvert » à la fois, plus précisément, topologiquement fini mais métriquement illimité. Sa topologie, c’est-à-dire essentiellement son ordre de connexité, détermine en grande partie sa structure interne, alors que sa métrique, bien qu’elle puisse être rigoureusement définie, n’admet pas pour autant de détermination mathématique unique, car elle peut prendre des valeurs qui dépendent, du moins en partie, des champs physiques qui évoluent dans l’espace. Si on définit une variété riemannienne comme une variété sur laquelle on fixe un champ de tenseurs g_{jk}, vérifiant les conditions $g_{kj} = g_{jk}$, $\det(g_{jk}) \neq 0$, alors la métrique riemannienne de l’espace-temps est donnée par l’expression quadratique $ds^2 = g_{jk}\, dx_j\, dx_k$, qui est symétrique et définie positive.

Einstein a montré que la structure riemannienne d’un espace-temps à 4 dimensions permettait de se représenter d’une manière très précise le phénomène physique de la gravitation, pourvu que les fonctions g_{jk} correspondent au champ gravitationnel. La relativité générale repose sur un principe épistémologique dont la portée est grande. Selon ce principe, les phénomènes physiques sont représentés par des objets géométriques associés au groupe de difféomorphismes de la variété riemannienne X. Le groupe de difféomorphismes de la variété X comprend toutes ces transformations qui peuvent s’appliquer à la variété et qui laissent sa forme métrique invariante ; tout élément du groupe est différentiable, c’est-à-dire que les fonctions g_{jk} possèdent des dérivées partielles de tous les ordres. Dans le cas, par exemple, d’une surface générale S, ces difféomorphismes ont la propriété de transformer des courbes tracées sur S qui sont tangentes en courbes tangentes, des courbes osculatrices en courbes osculatrices, et bien d’autres propriétés comme celles qu’on appelle des propriétés de contact.

Le concept de variété et la géométrie différentielle qui la caractérise permettent une classification fine des phénomènes physiques, c’est-à-dire qu’ils permettent de ramener un très grand nombre de variables physiques à un petit nombre d’objets géométriques. Par exemple, température et vitesse sont associées pour définir un vecteur tangent à l’espace-temps ; chaleur et production d’entropie définissent une densité vectorielle, etc. De ce fait la relativité générale, loin de se réduire à une simple théorie de la gravitation, apparaît comme un instrument d’analyse universel pour la physique, du moins à l’échelle globale de notre Univers. Cela ne doit pas cependant nous faire oublier le point essentiel : la gravité est une théorie qui mesure la courbure de l’espace-temps et qui explique les phénomènes physiques à l’échelle macroscopique en termes de cette courbure. De ce point de vue, la courbure est l’objet géométrique fondamental de la structure d’un espace-temps à 4 dimensions. Sa signification physique réside en ce que la trajectoire des « lignes d’Univers » des particules, correspondant aux géodésiques tracées sur la variété riemannienne, et la façon dont a lieu une distorsion de leurs chemins, sont en réalité un produit de la courbure de l'espace-temps.

3. Remarques épistémologiques sur l'espace

Les remarques précédentes montrent clairement qu'on peut concevoir l'espace physique comme une réalité qui émerge de la structure mathématique du monde. En d'autres termes, on peut dire que le comportement du monde dans lequel nous vivons est en quelque sorte engendré par des structures mathématiques riches et profondes. Certains mathématiciens et physiciens, comme Roger Penrose et Alain Connes, pensent à ce propos que plus nous avançons dans la compréhension du monde physique et pénétrons davantage dans les lois de la nature, plus on est forcés d'admettre que le monde physique en tant que tel tend, pour ainsi dire, à s'évaporer en ne laissant de stable et de permanent que les structures mathématiques sous-jacentes. Cependant, on pourrait penser que les rapports entre le monde des êtres mathématiques et l'univers des phénomènes physiques sont plus complexes qu'une telle conclusion ne le laisse croire, et que, de toute façon, l'univers physique ne se réduit pas au monde mathématique.

Il n'en demeure pas moins que plus on cherche à comprendre les lois de la physique, plus on rencontre un monde constitué de concepts et de structures mathématiques. Considérons brièvement l'exemple des rapports entre topologie et théorie quantique de champs. Il est clair qu'il existe une connexion profonde entre les concepts mathématiques et les objets physiques appartenant à cette théorie, un rapport qui n'est ni de l'ordre de l'application d'un langage mathématique à une réalité physique, ni de la démonstration rigoureuse, conforme aux méthodes « standard » des mathématiciens, de résultats heuristiques trouvés par les physiciens. Il s'agit bien plutôt de comprendre que certains « objets » mathématiques, comme ceux de connexion, espace fibré, invariants de nœuds, surfaces de Riemann, variétés de Calabi-Yau, groupes quantiques, représentent des modèles intelligibles des phénomènes physiques. Qui plus est, les symétries et les invariants topologiques des espaces sont directement impliqués dans la façon même dont les théories physiques constituent leurs propres objets.

Dans l'*Esthétique transcendantale*, Kant a raisonné comme si la nature tridimensionnelle de l'espace était une condition *a priori* et nécessaire de son existence pour nous. Son fondement résiderait dans ce que l'intuition subjective par laquelle nous percevons et pensons l'espace doit forcément s'accompagner d'une représentation tridimensionnelle des relations spatiales entre les objets. Pour Kant, en effet, le fait qu'on ne peut penser ces derniers que dans un tel cadre, vient de ce que les limites de nos représentations des objets dans l'espace coïncident avec le caractère tridimensionnel de ces mêmes représentations, au-delà desquelles notre intuition de l'espace cesse d'être fondée dans la réalité sensible pour acquérir le statut d'une intuition formelle.

Mais on pourrait se demander, premièrement, pourquoi Kant ramène les différentes façons qu'en en principe on a de penser l'espace aux limites qui caractérisent notre intuition (ou perception) sensible de l'espace ambiant ? Deuxièmement, pourquoi admet-il que les propriétés spatiales incluses dans les axiomes et postulats d'une géométrie dérivent directement et nécessairement des caractères qu'on attribue à notre représentation sensible de l'espace ? Troisièmement, qu'est-ce qui lui fait affirmer que les axiomes géométriques expriment un contenu évident, *a priori* et nécessaire de l'espace ? Ce sont là, nous semble-t-il, autant de difficultés majeures de la conception kantienne de l'espace et de la géométrie qui devraient conduire à proposer une révision profonde de son système et de sa théorie de la connaissance concernant la structure de l'espace mathématique et de l'espace physique.

Contrairement à Kant, Gauss avait rejeté l'idée que l'espace puisse simplement être considéré comme une forme pure de notre intuition sensible. Pour lui, l'espace a une réalité différente et indépendante des représentations que nous en avons, car il possède des lois et des propriétés objectives inhérentes à sa nature même. Donnons deux exemples qui peuvent aider à mieux comprendre ce qui sépare leurs conceptions. Kant avait étudié le phénomène de la symétrie gauche/droite et il en avait conclu qu'elle prouverait la nature *a priori* de l'espace, tandis que pour Gauss la symétrie en question prouverait exactement le contraire, à savoir que l'espace est régi par certains principes de symétrie ayant une signification objective, liés à telle et telle propriété géométrique intrinsèque ou à tel et tel contenu physique de l'espace. Ce qui, dans ce cas, renforce davantage le point de vue de Gauss, c'est le fait qu'une telle symétrie n'est pas que relative, c'est-à-dire qu'elle ne dépend pas du point de vue de l'observateur ou du caractère conventionnel de ses mesures : en effet, on peut facilement montrer que la gauche et la droite ne sont pas, dans beaucoup de cas, interchangeables, elles sont chirales. Il s'agit donc d'une contrainte objective qui caractérise l'organisation ou la structure spatiale de nombreux objets et phénomènes à la fois dans le monde physique et dans le règne du vivant.

L'autre exemple est celui de la réflexion que l'on peut faire subire à une figure géométrique ou à une surface. Pensez à deux images d'un même objet : d'abord à son image « réelle », puis à l'image que l'on voit reflétée dans le miroir. Or les deux images ne sont pas symétriques ce qui, en termes plus mathématiques, revient à dire que l'objet en question n'est pas invariant vis-à-vis d'une réflexion ou d'une symétrie miroir. Imaginons maintenant l'exemple d'une surface appelée *bande* (ou *ruban*) de *Möbius*. On sait que sa principale caractéristique est le fait d'être non orientable (il s'agit d'une propriété topologique de la plus haute importante). Cela signifie que quelqu'un qui voyagerait le long d'une courbe fermée se retrouverait, à la fin du voyage, avec une orientation locale qui n'est plus la même que celle qu'il avait au départ, c'est-à-dire que la droite est maintenant à gauche, et la gauche à droite. Essayons en outre d'imaginer deux images différentes de cette même surface : soit l'une d'entre elles, disons M_1, une bande circulaire à laquelle on a imprimé une demi-torsion à gauche ; soit l'autre, que l'on désignera par M_2, la même bande circulaire, mais à laquelle on a imprimée une demi-torsion à droite. Les deux situations sont possibles, et il y en a pas une que notre représentation sensible de l'espace privilégie par rapport à l'autre. On peut en fait montrer que M_1 et M_3 sont des bandes de Möbius non isotopes dans R^3. À la question de savoir s'il y a une préférence pour l'une ou pour l'autre de ces deux bandes dans R^3, la réponse est non ! Mais ceci n'exclut pas qu'en dimension 4, ou 5, ou plus grande encore, une telle préférence ne puisse effectivement se produire. Cela prouverait entre autres, à l'encontre du point de vue kantien, que la dimension est l'une des propriétés essentielles sur laquelle repose la constitution objective de l'espace.

La grande nouveauté des conceptions de Gauss, Riemann, Clifford, Klein, Poincaré, et d'autres, est d'avoir montré que l'espace possède, d'une part, des structures mathématiques intrinsèques indépendantes de l'espace euclidien usuel, de l'autre, une réalité physique qui renvoie à des propriétés dynamiques qui ne sont pas directement accessibles à notre perception. On en est arrivé ainsi à postuler, non seulement qu'il peut exister plusieurs géométries mathématiquement possibles, mais également que la géométrie euclidienne n'est pas la seule géométrie valable pour l'espace physique. Même quand on se limite à ne considérer que la physique (classique) newtonienne, il est toujours possible, comme E. Cartan l'a bien montré, de généraliser la théorie de la gravitation de Newton dans le cadre du formalisme géométrique de la relativité générale, en sorte que la théorie de Newton doit être en réalité considérée comme un cas-limite de la théorie einsteinnienne de la gravitation. Or une telle généralisation ne saurait être possible dans le cadre de la physique classique. On appelle ici « physique classique » cette théorie dans laquelle les lois mécaniques, cinématiques et dynamiques des corps, formulées par des équations mathématiques à l'aide du calcul différentiel et intégral sont considérées comme existant séparément de la structure géométrique de l'espace-temps ; l'espace, lui, est essentiellement considéré comme une entité absolue et extérieure aux lois de mouvement des corps et aux propriétés physiques des phénomènes. Ces mathématiciens et physiciens avaient bien compris que la question de savoir quelle géométrie s'accorde avec l'espace physique n'est pas décidable *a priori*, mais par des expériences fondamentales.

4. Relativité générale et cosmologie

La relativité générale, qui n'était au départ qu'une théorie mathématique extrêmement précise et symétrique, s'est avérée être profondément liée à la physique, dans le sens que sa structure mathématique correspond à la structure du monde réel, et notamment de notre Univers. En d'autres termes, la relativité générale nous dit que le comportement du monde physique à l'échelle cosmologique est régi par une structure géométrique réalisée sur une variété espace-temps à 4 dimensions de type riemannien avec métrique et courbure. Ses prédictions ont été d'ailleurs confirmées par la plupart des expériences physiques effectuées jusqu'à maintenant. La relativité générale considère l'Univers comme un tout. Ses développements mathématiques ont permis d'élaborer trois modèles (standard) d'Univers. Pour simplifier, on dira que ces trois modèles sont définis par un paramètre, k, qui est la courbure spatiale de l'Univers. Un autre paramètre qui apparaît dans les équations de la relativité générale a été introduit par Einstein sous le nom de *constante cosmologique*.

Si la constante cosmologique a valeur nulle, on obtient trois modèles d'Univers décrits par le paramètre k : (i) le modèle d'espace-temps d'un Univers plat en expansion constante avec des sections spatiales euclidiennes, et avec courbure nulle ($k = 0$) ; (ii) le modèle d'espace-temps d'un Univers courbe d'abord en expansion, puis en contraction, mais avec des sections spatiales sphériques et de courbure positive ($k = +1$) ; (iii) le modèle d'espace-temps d'un Univers courbe en expansion avec des sections spatiales non euclidiennes (lobatchevskiennes) et avec courbure négative ($k = -1$). Dans tous les trois modèles, on est en présence d'un état singulier initial, le *Big Bang*, qui marque le commencement de l'Univers. Mais tandis que dans le cas où k est nulle, l'Univers connaîtra très probablement une expansion très grande, maximale, pour ensuite retomber en un *Big Crunch*, dans le cas où k est négative, l'Univers restera vraisemblablement indéfiniment en expansion. Plusieurs physiciens et cosmologistes semblent privilégier le modèle d'Univers à courbure constante négative, et ce essentiellement pour deux raisons, l'une de nature physique, l'autre plutôt mathématique. La raison physique invoquée est que, puisque selon la relativité générale la courbure de l'espace semble être déterminée par la densité de matière présente dans l'Univers, il se pourrait que la quantité de cette matière ne soit pas suffisante pour que l'Univers ait une géométrie de type sphérique, autrement dit, pour qu'il soit spatialement fermé. L'autre raison concerne une propriété mathématique de la géométrie non euclidienne hyperbolique. Elle peut être décrite intuitivement en disant que la géométrie qui caractérise les modèles d'espace non euclidien hyperbolique présente les mêmes propriétés partout, tout aussi bien au centre qu'à la périphérie de l'espace. En d'autres termes, la géométrie valable localement se reproduit à l'infini et reste valable à l'échelle globale : l'espace se comporte comme un tout, de façon auto-similaire mais avec des variations de taille.

5. La physique vue comme déformation de l'espace géométrique

Dans cette partie, nous aimerions d'abord proposer deux énoncés épistémologiques, ensuite, donner les exemples physiques correspondants pour essayer d'en mettre en évidence leur signification profonde. Commençons par les énoncés épistémologiques.

1. Il nous semble que *le physique peut être essentiellement compris comme du géométrique en acte*. Ce qui ne veut pas dire pour autant que les propriétés physiques des phénomènes se réduisent aux principes géométriques des théories, ni évidemment que la signification de ces mêmes principes s'épuise dans les propriétés physiques des phénomènes qu'ils permettent de décrire et d'expliquer.
2. Nous pensons que *c'est le réel qui est géométrique, et non pas la géométrie qui est réelle ou empirique*. Cela veut dire qu'il est beaucoup moins

important de corroborer ou d'infirmer *a posteriori* une théorie physique par l'expérience, que de montrer que les structures, les concepts et les invariants mathématiques de cette même théorie sont précisément ceux auxquels obéissent les phénomènes du monde réel.

Passons maintenant à une brève présentation historique de l'exemple physique. En 1918, le mathématicien et physicien Hermann Weyl a proposé une nouvelle théorie mathématique (il s'agit d'une généralisation de la géométrie riemannienne où notamment la condition d'intégrabilité locale n'est plus satisfaite et où à la place intervient un facteur scalaire non intégrable) dans le but d'unifier la gravitation avec l'électromagnétisme, c'est-à-dire la relativité générale d'Einstein avec la théorie de Maxwell. Cette théorie fut cependant rejetée, à commencer par Einstein lui-même, sur la base de quelques observations expérimentales qui semblaient en contredire ses principes essentiels. Cependant, une dizaine d'années plus tard, la théorie de Weyl fut redécouverte dans le cadre des premières formulations des théories quantiques des champs par Pauli et Dirac, qui ont montré qu'elle permettait d'unifier d'une façon mathématiquement et physiquement cohérente les principes de la mécanique quantique avec les équations du champ électromagnétique.

Mais c'est au début des années 1960 qui se produit quelque chose de révolutionnaire en physique, qui aboutira peu de temps après dans la construction du « modèle standard » par Salam, Ward, Glashow et Weinberg. Ce modèle a permis d'unifier la force « électrofaible », qui englobe les forces nucléaires faibles et l'électromagnétisme, avec la force nucléaire forte qui est aussi une théorie de jauge. Le modèle standard correspond à la symétrie $SU(3) \times SU(2) \times U(1)$. Une étape fondamentale dans la construction du modèle standard a été la découverte importante faite en 1954 par les physiciens théoriciens Yang et Mills, montrant que la force nucléaire faible et la force électromagnétique sont toutes les deux des forces de jauge. Rappelons que selon la théorie quantique des champs, toutes les forces – de jauge ou pas – sont produites par des échanges de particules que l'on appelle communément « les messagers ». Ces particules messagers doivent toujours avoir un spin entier (0, $1h$, $2h$, etc.), alors que les particules de matière (les « sources ») comme les électrons, les neutrinos les protons et les neutrons, décrites par l'équation de Dirac, ont un spin $h/2$.

La propriété fondamentale des forces de jauge, qui les distingue des autres forces, réside en ce qu'elles sont produites par des messagers de spin 1. (Le principe de jauge est l'expression d'un comportement universel qui relie l'intensité des forces de jauge avec la notion de charge. Ainsi par exemple, l'intensité du couplage de jauge entre le photon et un noyau d'hélium est deux fois celui du photon avec un noyau d'hydrogène parce que la charge électrique du noyau d'hélium est deux fois celle du noyau d'hydrogène.) Le prototype de toutes les forces de jauge est l'électromagnétisme. Dans ce cas, le messager de spin 1 est le photon γ, le quantum de lumière. Pour se représenter la force entre des objets « source » tels qu'un proton (p_+) et un électron (e_-), transportant chacun une charge unité (mais des signes opposés), on peut s'imaginer un proton incident émettant un photon après une décélération (ou un changement de direction) conformément aux idées de Maxwell (figure 5.1, à gauche). Dans un deuxième temps, le photon est absorbé par l'électron qui est accéléré (figure 5.1, à droite), toujours selon les idées de Maxwell. Ce qui en résulte est un échange d'impulsion entre le proton et l'électron. C'est ainsi qu'opère la force électromagnétique.

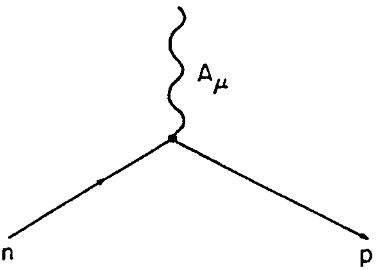


FIG. 5.1 – La force électromagnétique entre un proton (p) et un électron (e).

Yang et Mills ont eu le mérite, non seulement de redécouvrir l'idée originaire de Weyl d'invariance de jauge locale, mais ils ont en plus montré qu'elle constitue un principe général valable dans (ou presque) tous les champs de la physique. Premièrement, ils ont réalisé que l'interaction nucléaire forte peut être obtenue d'une façon similaire à une théorie des champs comme l'électromagnétisme, qui est, on l'a vu, invariante par changement de jauge. Deuxièmement, quelque temps après avoir formulé leur théorie, ils ont compris quelque chose de tout à fait fondamental, à savoir qu'un champ de jauge se comporte comme un objet mathématique, précisément comme une connexion sur un fibré. Ils ont alors postulé que le groupe de jauge local était le groupe isospin SU(2), appelé aussi le sous-groupe unitaire à 2 paramètres. L'idée était révolutionnaire car elle comportait un changement important dans la façon de concevoir l'« identité » d'une particule élémentaire. L'idée selon laquelle la connexion isospin, et par conséquent le potentiel de jauge, agit de la même façon que le groupe de symétrie SU(2), est le résultat le plus important de la théorie de Yang-Mills. Ce concept est en effet au cœur de toute théorie locale de jauge. Il montre clairement comment le groupe de symétrie de jauge est construit au sein même de la dynamique de l'interaction entre particules et champs. Qui plus est, bon nombre de propriétés physiques importantes du champ peuvent être directement déduites de la connexion, qui peut être conçue comme une combinaison linéaire des générateurs du groupe SU(2). De plus, on peut associer cette opération formelle à des processus physiques réels (figure 5.2). La théorie de Yang-Mills permet en effet de décrire les propriétés internes, les nombres quantiques comme le spin, des particules élémentaires. Cela veut dire que l'isospin est impliqué dans la détermination de la forme fondamentale de l'interaction.

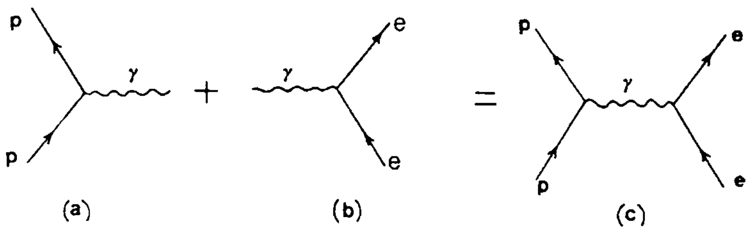


FIG. 5.2 – Dans la théorie de Yang-Mills, un neutron n est transformé en un proton p par absorption du champ de jauge de Yang-Mills A_μ. Ce dernier agit comme un opérateur isospin.

6. Théories de jauge et groupes de symétries

Terminons ces remarques sur la théorie de jauge en en rappelant quelques aspects mathématiques. La première chose à souligner est que toutes les théories physiques fondamentales connues jusqu'à présent sont en réalité des théories de jauge. Ce qui veut dire que chacune d'entre elles est d'abord et avant tout la donnée d'un groupe de symétries, ou, comme dans certains cas, d'un semi-groupe de symétries, ou encore, d'un groupe de représentations. Afin de préciser cette dernière affirmation, considérons le groupe orthogonal SO(3). Il décrit les rotations de l'espace euclidien à 3 dimensions. En d'autres termes, SO(3) agit sur $\mathbf{R}^3$, ce qui signifie que chaque élément de SO(3) définit une transformation linéaire de $\mathbf{R}^3$. Ceci peut se généraliser en admettant qu'un groupe G agit sur un espace de vecteurs V s'il existe une application ρ de G vers les transformations linéaires de V tel que

$$\rho(gh)u = \rho(g)\,\rho(h)u$$

pour tout $u \in V$. On dit que ρ est une représentation du groupe G sur l'espace de vecteurs V. En fait, une représentation n'est qu'un type spécial d'homéomorphisme. Soient G et H deux groupes, on a alors que l'application $\rho : G \to H$ est un homéomorphisme si

$$\rho(gh) = \rho(g)\,\rho(h).$$

Si l'on définit le groupe linéaire général GL(V) de sorte à ce qu'il soit le groupe de toutes les transformations linéaires inversibles de V, une représentation de G sur V n'est rien d'autre que l'homéomorphisme

$$\rho : G \to \mathrm{GL}(V).$$

Par conséquent, quand G est un groupe de Lie, on ne considérera que des représentations où V est de dimension finie et où ρ est une application lisse entre des variétés différentiables.

L'un des faits les plus remarquables de la physique contemporaine est d'avoir compris que les différents groupes de Lie permettent de définir des systèmes d'équations, appelés équations de Yang-Mills, qui décrivent très bien les diverses forces dans le modèle standard. Ce groupe est appelé le groupe de symétries ou le groupe de jauge des forces en question. Ainsi, l'électromagnétisme a comme groupe de jauge le groupe unitaire à une dimension U(1). Ce groupe est commutatif, ou abélien, ce qui revient au même. Un groupe G est abélien si

$$gh = hg$$

pour tout g, $h \in G$. Il convient de préciser que les équations de Yang-Mills sont linéaires, précisément lorsque le groupe de jauge est abélien. Par contre, dans le modèle standard, la force nucléaire forte a comme groupe de jauge le groupe unitaire spécial SU(3), qui est non abélien. Cela veut dire que cette force se comporte de façon non linéaire, et que sa structure est donc plus fine que celle qui caractérise la force électromagnétique. À partir des groupes G et H, on peut former un seul groupe en appliquant l'opération $G \times H$, qu'on appelle produit direct de G et H, et, en effet, dans le modèle standard les forces électromagnétique et faible sont traitées d'une façon unitaire, et elles ont comme groupe de jauge le groupe $\mathrm{SU}(2)\times\mathrm{U}(1)$, qui est aussi non abélien.

Le modèle standard est fondé sur le groupe de jauge complet $SU(3) \times SU(2) \times U(1)$. Pour être précis, partiellement complet, puisque la gravité n'y est pas comprise.

Dans le modèle standard, chaque particule possède une charge, non seulement dans le sens de la charge électrique usuelle, mais aussi dans le sens que la charge détermine la façon dont la particule interagit avec les forces nucléaires faible et forte. La connexion profonde entre la théorie des groupes et la charge consiste en ce que la charge d'une particule est d'abord et avant tout la donnée du choix de la représentation pour le groupe de jauge en question.

7. L'espace physique, de la relativité générale à la mécanique quantique

Formulons quelques idées générales à partir desquelles il pourrait être possible d'élaborer une théorie des champs quantiques qui soit invariante sous une certaine classe topologique de difféomorphismes. Une telle théorie repose sur l'idée que l'espace de configurations et leurs déformations sont constitutifs de plusieurs classes de phénomènes physiques. Nous pensons à ce propos qu'il faut abandonner la théorie locale des champs quantiques, dans laquelle les opérateurs physiques sont marqués par des régions correspondantes d'une certaine variété métrique, au profit d'une théorie globale des champs quantiques, dans laquelle les opérateurs physiques sont, en revanche, des objets géométriquement invariants sous les difféomorphismes. Mentionnons d'abord deux exemples importants qui montrent bien comment la structure géométrique et topologique de l'espace influe sur les phénomènes physiques.

(i) Dans le cas de l'électromagnétisme, on peut empiriquement montrer que le champ peut être identiquement nul, alors que des effets physiques peuvent encore être détectés. Ces effets sont dus au fait que le transport parallèle doit être non trivial au cas où la région de l'espace ne serait pas simplement connexe, et en effet on connaît à présent plusieurs modèles d'espaces non euclidiens qui se caractérisent par le fait d'être multiconnexes et dont le contenu physique est très riche. Géométriquement, le fait que la courbure (le champ de jauge) s'annule, donne l'information nécessaire sur le transport parallèle le long des chemins (géodésiques) fermés. Dans le langage de la physique, le transport parallèle est généralement décrit en termes d'un facteur de phase non intégrable. Or la propriété de non intégrabilité se réfère localement à l'existence d'un champ non nul, tandis qu'à grande échelle cette même propriété est de nature topologique et peut survenir même dans le cas de champs nuls. Il importe ainsi de souligner que la connexion est un objet géométrique, alors que le potentiel (des équations du champ) doit être considéré comme étant de nature physique. C'est le choix de la jauge grâce à laquelle on décrit le potentiel qui n'a pas de signification physique, ce qui correspond au fait que l'espace fibré (géométrique) qui « porte » la connexion ne possède pas des sections horizontales.

(ii) Un autre exemple qui montre que des objets géométriques, une fois qu'on les fait agir sur l'espace qu'ils « habitent », peuvent produire des effets physiques même en l'absence de champ physique, est le suivant. On doit sérieusement considérer la possibilité que sur des espaces-temps topologiquement non triviaux, le modèle de fibré principal de la théorie des champs soit hautement non trivial. En effet, en 1930, Paul Dirac introduira la notion de *monopole magnétique*, c'est-à-dire d'un champ électromagnétique avec une singularité isolée dans l'espace. Depuis il a été possible d'observer que l'intégrale du champ sur une sphère bidimensionnelle bornant la singularité, pouvait prendre des valeurs réelles non nulles. Ces entiers proviennent en fait de la première classe de Chern du fibré unitaire U(1) sous-jacent, et le fait qu'ils ne soient pas nuls, prouve le caractère topologiquement et physiquement non trivial de ce fibré.

Une question centrale sur laquelle nous voudrions brièvement revenir concerne le statut et le rôle des symétries dans les théories physiques récentes. D'abord, il est important de remarquer que le concept de symétrie implique à la fois les idées de continuité et de discontinuité dans la nature. D'une part, les formes ou configurations symétriques sont des objets étendus qui requièrent l'existence d'un espace ambiant continu. De l'autre, elles manifestent également un ordre intrinsèque, réalisé mathématiquement par un ensemble d'opérations (discrètes ou continues) qui conserve cet objet. Continuité et discontinuité sont toutes deux présentes dans la théorie des groupes de Lie, où l'on a des groupes continus comme les groupes des rotations, les groupes euclidien, etc. Mais dans le monde physique « réel », on suppose que la réalité « ultime » (microscopique) est constituée d'éléments discrets, particules ou atomes. Soulignons par ailleurs que dans la physique classique les symétries n'avaient qu'un rôle auxiliaire permettant de simplifier raisonnements ou calculs, tandis que le développement contemporain de la physique tend à leur faire jouer un rôle de plus en plus fondamental, car elles apparaissent comme un principe organisateur et constitutif de la réalité physique elle-même. Certes, le champ d'investigation des symétries s'est dans les derniers temps considérablement élargi, et la plupart des nouvelles symétries que les théories de jauge non abéliennes et la théorie des supercordes ont contribué à mettre en évidence, ne sont plus accessibles à notre expérience immédiate. Il suffit d'évoquer le rôle dynamique des symétries dans les domaines des particules subatomiques, et également le fait que les symétries « brisées » sont essentielles dans la description et la compréhension de nombreux phénomènes de natures d'ailleurs très diverses.

L'étude de certaines propriétés de la « supersymétrie » montre bien comment elles ouvrent la voie à une unification plus profonde des forces existant dans la nature. D'un point de vue mathématique, la supersymétrie est un nouveau type d'objet géométrique. Au lieu de se placer sur une variété différentiable comme dans le cas de la géométrie différentielle classique, on se place d'emblée sur un espace fibré analogue à l'espace fibré des spineurs où interviennent des variables non commutatives. Un grand changement se produit dans la dynamique de la théorie physique et également dans sa structure mathématique lorsqu'on remplace les composantes (les particules) censées former la matière par des objets étendus comme les cordes. D'abord, ces objets sont pourvus d'une structure interne. Mais plus important encore, est le fait que les propriétés physiques des particules et de leurs interactions apparaissent ainsi comme autant des caractéristiques des formes spatiales engendrées par les modes d'oscillation de ces cordes. La charge électrique, par exemple, peut être vue comme un effet produit par le mouvement de la corde dans son espace plutôt que comme quelque chose qui s'ajoute à la particule comme objet « fondamental ».

Une des questions centrales des développements récents de la physique théorique et en particulier de la théorie des champs quantiques concerne la structure et la nature de l'espace-temps. Ces développements exigent que l'on élargisse de façon importante notre conception de l'espace, du temps et de l'espace-temps. Au regard de la réalité aussi bien microscopique que cosmologique, cette conception ne peut que être très éloignée de l'image de l'espace euclidien comme d'un espace plat (sans courbure), isotrope et homogène, mais aussi du modèle d'espace riemannien de la relativité générale, c'est-à-dire d'un espace continûment différentiable et dont la structure est déterminée par une métrique définie toujours positive. En effet, les principales théories physiques développées ces dernières décennies, comme la théorie conforme des champs, la gravité quantique, la supergravité, la théorie des supercordes et la géométrie non commutative, malgré leurs différences significatives, semblent néanmoins avoir en commun une tendance à l'unification de l'espace-temps et de la dynamique, de la structure géométrique et des interactions physiques.

On sait que la relativité générale et la mécanique quantique ont conduit à un changement profond de nos conceptions de l'espace et de l'espace-temps, de la structure de la matière et des rapports entre les observateurs et leurs instruments de mesure et les phénomènes. Cependant, elles aboutissent à des conclusions opposées quant à la nature de l'espace-temps. La relativité générale est fondée sur un modèle mathématique qui est un continuum à 4 dimensions doué d'une métrique pseudo-riemannienne. La mécanique quantique, en revanche, admet que l'espace physique à l'échelle subatomique, modélisé mathématiquement par un espace de Hilbert sur lequel on définit une certaine structure d'algèbre d'opérateurs, se comporte de manière discontinue et que sa métrique fait intervenir des termes infinis. Les tentatives récentes pour parvenir notamment à une théorie de la gravité quantique reposent sur l'idée que l'espace-temps à l'échelle quantique doit avoir une structure essentiellement discrète. Il est possible en effet de montrer, par des raisonnements à la fois mathématiques et physiques, que pour les phénomènes existant à l'échelle de Planck 10^{-33} cm, les effets des fluctuations quantiques sur la courbure de l'espace-temps sont assez grands pour produire des modifications importantes dans sa topologie. Cela suggère l'idée d'un certain type de superposition non linéaire des différentes topologies, qui ressembleraient très peu à ce que d'ordinaire on appelle une variété lisse. Plus précisément, on peut faire l'hypothèse que la structure de l'espace-temps à l'échelle quantique n'est pas celle d'une variété différentiable C^{∞}, mais vraisemblablement l'équivalent d'un espace topologique non classique construit à partir d'une surface de Riemann compacte complexe à bords. Une des conséquences de cette hypothèse, en particulier dans la théorie des champs quantiques, c'est que les propriétés topologiques globales de la variété M (lorentzienne ou pseudo-riemannienne) jouent un rôle fondamental et, partant, plusieurs effets physiques quantiques proviennent de la structure géométrique globale (et de la topologie) de M.

Terminons ces réflexions en mettant en évidence quelques idées nouvelles que les développements récents de la physique théorique ont permis de dégager. Elles conduisent à repenser de manière profonde la conception qu'on s'est faite jusqu'à maintenant de la nature de l'espace et de l'espace-temps. Il n'est pas exagéré d'affirmer à ce propos qu'on est en présence d'une transformation fondamentale de la pensée scientifique et philosophique, tout aussi importante sinon plus importante que la révolution einsteinnienne du début du siècle dernier.

(i) La structure géométrique de l'espace-temps est à l'origine du comportement dynamique, et non seulement cinématique, des phénomènes physiques qui se déroulent en son sein. Cela ne concerne plus désormais uniquement le champ gravitationnel, mais également l'électromagnétisme et les autres champs de matière.

(ii) Il apparaît clairement qu'aussi bien les symétries spatio-temporelles que physiques (internes), dictent les différentes interactions entre forces et entre particules, et que leurs lois sont à l'origine des propriétés essentielles des phénomènes physiques.

(iii) L'invariance locale de jauge semble être un principe universel qui gouverne les interactions fondamentales entre les particules et les champs de matière. En fait, toutes les théories physiques connues de nos jours peuvent être formulées en accord avec ce principe, bien qu'il reste toujours à démontrer qu'elles appartiennent au même groupe de symétries de jauge.

(iv) Les théories de jauge et des supercordes illustrent un des caractères (et mystères) les plus profonds de la réalité physique, qui est que plus on cherche à comprendre les principes fondamentaux qui régissent le comportement des phénomènes à l'échelle microscopique, plus on découvre que ces principes renvoient à des structures mathématiques sous-jacentes toujours plus riches en invariants.

8. Théorie quantique des champs, modèle standard et théorie des cordes

Commençons par quelques rappels sur la théorie quantique des champs, qui est fondée sur la notion d'interaction. Cette théorie unit la mécanique quantique et la relativité restreinte. Dans ce cadre, toutes les forces qui existent dans la nature sont décrites par des échanges de particules dites *médiatrices*. L'électrodynamique quantique décrit les phénomènes électromagnétiques ; les forces qui résultent de l'interaction électromagnétique correspondent à des échanges de *photons*. Les forces nucléaires qui lient les protons aux neutrons du noyau, correspondent à l'échange de particules médiatrices entre quarks : les *gluons*. Les protons et les neutrons sont soumis à l'interaction forte, force qui diffère de l'interaction électromagnétique qui s'exerce entre protons et électrons. Elle est beaucoup plus intense que la force électromagnétique ; en revanche, elle n'a de portée qu'à très courte distance : 10^{-15} m seulement.

Le modèle standard représente la théorie la plus complète en physique des particules. Il a fourni un cadre théorique cohérent grâce auquel on parvient à décrire toutes les interactions des particules élémentaires. Dans la nature existent des particules n'ayant pas de structure interne et qui ne sont pas soumises aux interactions fortes, ce sont les *leptons* qui comprennent les électrons, les nuons, le tau et trois types de neutrinos. L'*interaction faible* regroupe les leptons et les quarks ; c'est l'interaction responsable des désintégrations radioactives (bêta) des noyaux atomiques, caractérisées par l'émission d'un électron ou de son antiparticule, le positon. Comme les autres forces, l'interaction faible est due aussi à l'échange de particules médiatrices, les particules *W* et *Z*. À la différence des photons et des gluons de l'interaction forte, les particules médiatrices de l'interaction faible ont une masse élevée : près de 100 000 giga-électronwolts (un giga-électronvolt est égal à un milliard d'électronvolts). Les interactions électromagnétiques faibles peuvent être réunies dans le cadre d'une théorie unifiée, appelée théorie des interactions électrofaibles. Le modèle standard actuel des particules élémentaires est donc la synthèse des théories des interactions électrofaibles et des interactions fortes.

Mais le modèle standard présente une limite fondamentale, car il n'inclut pas la gravitation qui, dans ce cadre, apparaît incompatible avec les trois autres interactions : électromagnétiques, faible et forte. En fait, deux conceptions s'opposent : d'un côté, le modèle standard avec ses théories quantiques et ses interactions locales entre particules et, de l'autre, la relativité générale, où la gravitation résulte d'une déformation de l'espace-temps. La question ainsi se pose de savoir s'il faut considérer la relativité générale telle que nous la connaissons comme la théorie classique d'une théorie quantique plus générale, la gravitation quantique ; mais il faut alors accepter qu'une théorie quantique de la gravitation puisse être en même temps une théorie quantique de l'espace et du temps. Or, le modèle standard et la plupart des théories physiques sont bâties sur l'espace et le temps, et toute modification des fondations de cet espace et de ce temps pourrait conduire à remettre (du moins partiellement) en cause le modèle standard de la physique. Il est en tout cas possible qu'aux petites distances inférieures à 10^{-35} m, les incertitudes dues à la théorie quantique perturbent vraisemblablement la structure de l'espace-temps. Si un électron est réellement une singularité de l'espace-temps avec une masse (une particule infiniment petite), cette singularité devrait disparaître à cette échelle. À quoi ressemble alors un électron ?

Tout récemment, la théorie des supercordes semble donner une réponse plausible à cette question. Selon cette théorie, les leptons et les quarks seraient des manifestations de minuscules objets unidimensionnels, les « supercordes ». Sortes d'infimes structures en formes de lacet, les supercordes seraient moins singulières qu'un point (image d'une particule dans la théorie des champs). De ce point de vue, la formulation de la gravitation quantique pourrait mieux se faire dans le cadre d'une théorie des supercordes. Pour cela, il faut supposer que l'espace-temps décrit à l'aide de 4 dimensions (trois d'espace et une de temps) acquiert six dimensions supplémentaires. Cela semble *a priori* contredire nos observations, mais seules quatre des dix dimensions auraient des effets macroscopiques ; les autres ne seraient perceptibles qu'à petite distance, et y rendraient la gravitation plus intense.

Prenons une simple image : si nous enroulons une feuille de papier sous forme de cylindre, elle ressemble, vue de loin, à une droite à une dimension. Cela s'explique d'abord physiquement, de trop grandes ou trop petites distances, et physiologiquement, du fait des limites qui caractérisent nos systèmes sensoriels. La deuxième dimension, celle qui permet de tourner autour du cylindre, n'est perçue qu'à courte distance. Ainsi en va-t-il des dimensions cachées de l'Univers : elles sont peut-être là, mais si petites que nous ne les voyons pas. Si notre espace possède effectivement d'autres dimensions à petite échelle, elles pourraient être mises en relation avec certains phénomènes observés par les physiciens des particules. En d'autres termes, peut-être qu'il existe une explication géométrique des structures fondamentales du modèle standard, et peut-être aussi que les leptons et les quarks (tenus ensemble par l'interaction faible) ne sont pas ponctuels.

Que la géométrie soit intimement liée à la physique est prouvé notamment par le fait que le nombre de dimensions de l'espace et de l'espace-temps joue un rôle fondamental dans la physique et notamment dans les théories unifiées, depuis en particulier les théories de Minkowski et Einstein, la théorie de Kaluza-Klein, la supergravité et la théorie des cordes. L'idée sous-jacente à ces théories est que l'espace tridimensionnel existe sous plusieurs modes, dans le sens qu'il pourrait être également « plongé » dans un espace plus vaste, à quatre dimensions. C'est cette idée qui a conduit à certaines révisions fondamentales des théories physiques fondées sur les concepts d'espace et d'espace-temps. En effet, ce que la physique théorique propose aujourd'hui est l'extension de cette idée aux dimensions supérieures. Il s'agit en particulier de penser notre espace tridimensionnel comme un objet géométrique plongé dans un univers qui comporte une ou plusieurs dimensions supplémentaires.

Ces dimensions « cachées » pourraient avoir une influence sur la gravité à petite distance, et en même temps pourraient permettre de mieux comprendre certains phénomènes dus à la nature quantique de la gravitation à très courte échelle.

On a vu que le modèle standard est parvenu à unifier l'interaction électromagnétique et l'interaction faible. Cependant, le modèle standard n'explique toujours pas les effets de certaines fluctuations quantiques prévues par la théorie ; ces effets sont de l'ordre de l'énergie de Planck. En particulier, si le modèle standard décrit bien tout ce qui a lieu au niveau de l'échelle électrofaible (de l'ordre de 10^{-19} m), il semble par ailleurs inapte à expliquer l'unification des interactions électrofaibles avec la gravitation à l'échelle de Planck. Le principal propos de la physique microscopique est donc celui de la construction d'une nouvelle théorie insensible aux (ou invariante vis-à-vis des) fluctuations quantiques. L'une des approches possibles de ce problème consiste à considérer une symétrie beaucoup plus générale de celle du modèle standard – nommée *supersymétrie*. Cette symétrie semble avoir l'avantage de décrire par un même formalisme trois forces de la Nature : les interactions électromagnétiques, faibles et fortes. (Je rappelle que l'interaction forte est responsable de la cohésion des noyaux d'atome ; protons et neutrons sont donc soumis à l'interaction forte ; elle est en plus beaucoup plus intense que la force électromagnétique, mais elle ne porte qu'à très courte distance : 10^{-15} m seulement.)

Il existe cependant une approche radicalement différente de celle qui a abouti au modèle standard, et qui consiste à modifier la structure géométrique de l'espace-temps, la nature de la gravité et l'échelle de Planck elle-même. On suppose en particulier que l'échelle de Planck doit être conçue comme l'énergie à laquelle les forces de la Nature s'équilibrent. Dans ce modèle, on introduit des dimensions supplémentaires pour donner aux forces de gravitation une intensité supérieure aux courtes distances, de sorte que l'échelle de Planck serait bien supérieure à 10^{-35} m. On appelle ces dimensions ajoutées *cycliques*, et leur présence, à petite échelle, fait apparaître des propriétés physiques nouvelles. Une telle idée a été proposée dans les années 1920 par le mathématicien polonais Theodor Kaluza et le physicien suédois Oskar Klein ; ils développèrent en effet une remarquable théorie qui unifiait la gravitation et l'électromagnétisme à l'aide d'une dimension spatiale supplémentaire. L'idée à été reprise depuis dans la théorie de la « Supergravité » et, plus récemment, dans la théorie des cordes, avec cependant la différence importante suivante : alors que dans les théories citées on attribuait aux dimensions cycliques supplémentaires un rayon proche de l'échelle de Planck habituelle, dans la théorie proposée récemment, au contraire, les dimensions cycliques sont à l'intérieur de cercles dont le rayon, selon les modèles, est compris entre 10^{-14} mm.

La question évidemment se pose de savoir où pourraient se trouver les dimensions cachées d'aussi petite taille ? Et pourquoi nous ne les voyons pas ? À ce propos, supposons que la matière et les trois forces faible, forte et électromagnétique soient confinées dans une membrane, qui serait l'espace tridimensionnel classique. Nous vivrions alors comme les figures géométriques de *Flatland* imaginées en 1884 par le révérende anglais Edwin Abbott ; les particules décrites par le modèle standard ne seraient présentes que dans la membrane, et seules les lignes de force gravitationnelles s'étendraient dans l'espace complet : autrement dit, seules les particules responsables de l'interaction, c'est-à-dire les gravitons, se déplaceraient librement dans les dimensions cycliques. Dans un tel Univers, la présence de dimensions supplémentaires ne serait perceptible que par les effets de la gravité. L'idée des particules restreintes à une membrane provient de la théorie des supercordes, où on identifie des membranes particulières nommées *D-branes* (la lettre *D* est l'initiale du mathématicien allemand Peter Gustav Lejeune Dirichlet (1805-1859), mathématicien qui explora les propriétés mathématiques d'espaces tels que les *D-branes*). Selon la théorie des supercordes, les électrons et les photons sont comme des minuscules bouts de corde dont les extrémités restent attachées à un tel espace-membrane ; les gravitons, en revanche, sont des minuscules boucles de cordes qui se déplacent dans toutes les dimensions.

D'un point de vue physique, on suppose dans cette théorie que, même si la gravité est renforcée à petite distance, son intensité n'égale celle des autres forces qu'aux distances inférieures à 10^{-19} m ; aux grandes distances, elle demeure très faible. D'après la théorie des dimensions cachées, les gravitons interagiraient de façon sensible avec les autres particules (cela revient à dire que l'interaction gravitationnelle est plus intense à petite distance), de sorte que les gravitons devraient être produits en nombre supérieur lors des collisions des particules. En outre, la très grande majorité d'entre eux s'échapperaient dans les dimensions cachées. Remarquons enfin que les gravitons pourraient être couplés avec des particules massives du type de celles prédites par Kaluza-Klein. Elles joueraient un rôle identique à celui du photon, le médiateur des interactions électromagnétiques.

Si la théorie des dimensions cachées sera un jour vérifiée expérimentalement, alors notre conception de l'espace et notre notion de l'Univers changeront profondément. Ce dernier ne serait donc qu'une mince membrane d'un espace aux multiples connexions ; et dans des feuillets passant par les dimensions supplémentaires notre Univers n'apparaîtrait que comme un point unique et infinitésimal sur chaque tranche, entouré de vide. Ce qui reviendrait à découvrir que nos trois dimensions ne sont pas le centre (ou l'essentiel) de l'Univers, car des membranes d'autres Univers tridimensionnels pourraient être parallèles à la nôtre et placées à moins d'un millimètre de la nôtre. En outre, bien que toutes les particules du modèle standard soient nécessairement scellées à la membrane qui constitue notre Univers, on peut supposer que des particules autres que le graviton ne fassent pas partie du modèle standard et se propagent dans les dimensions supplémentaires. Loin d'être vides, les dimensions cachées contiendraient une multitude de structures intéressantes, voire d'autres Univers.

Il n'est pas à exclure que la théorie des espaces multidimensionnels et multiconnexes puisse contribuer à élucider deux mystères de la physique des particules et de la cosmologie. Le premier concerne la détermination de la masse du neutrino. La théorie des dimensions cachées prévoit que les neutrinos interagissent avec une particule qui habiterait les dimensions cachées. Cette interaction apparaîtrait faible dans nos trois dimensions d'espace, car elle aurait surtout lieu dans les autres dimensions. C'est pourquoi le neutrino n'aurait qu'une masse faible. L'autre mystère concerne l'explication de la masse cachée de l'Univers, c'est-à-dire du caractère invisible de la « matière noire » ; ceci s'expliquerait par le fait que, les photons étant confinés à notre membrane, la matière contenue dans les dimensions parallèles n'enverrait aucune lumière vers nos yeux.

Les développements contemporains en physique théorique posent avec force la question fondamentale suivante : est-ce que l'espace-temps doit être considéré comme un cadre fixé au départ, un « objet » qui surdétermine la physique ? Ou est-ce qu'il ne pourrait plutôt être repensé comme un concept lui-même dérivé, c'est-à-dire comme une approximation résultant de la physique ? Les développements les plus récents de la physique suggèrent que la réponse à cette question réside très vraisemblablement dans une réunification de la dynamique de la théorie non perturbative des champs quantiques avec la géométrie de l'espace-temps aux petites échelles.

À ce propos, Edward Witten a notamment montré que, si l'on remplace les diagrammes de Feynman ordinaires (c'est-à-dire affectés par des singularités) par les diagrammes de cordes, on n'a plus besoin à ce moment-là d'un espace-temps qui soit donné comme un cadre abstrait, mais d'une théorie du champ à deux dimensions qui décrit la façon dont les cordes se propagent. Plus précisément, on peut obtenir l'espace-temps à partir de (ou en même temps que) cette théorie du champ. Cette idée fait apparaître une différence essentielle par rapport aux conceptions antérieures : tandis que dans la physique classique on parle d'espace-temps et des champs qu'il peut contenir ; dans la théorie des cordes, on a une théorie du champ auxiliaire à 2 dimensions qui encode l'essentiel de l'information sur l'espace-temps. On peut dire ainsi que le programme des supercordes constitue une étape majeure dans les tentatives faites depuis un siècle et demi pour comprendre la nature des rapports entre la structure géométrique de l'espace et les lois des interactions qui régissent le comportement des phénomènes à toutes les échelles de la Nature, tentatives qui, comme on l'a vu, avaient été au centre des préoccupations de quelques grands esprits comme Riemann, Clifford, Poincaré et Einstein.

Remerciements : Je suis reconnaissant à Marc Lachièze-Rey pour l'invitation à présenter un exposée au colloque de Cargèse de 2001 sur *L'espace physique, entre mathématiques et philosophie*, d'où cette contribution est issue. Mes remerciements vont aussi à Jean-Pierre Luminet, Mario Novello, Joseph Kouneiher, Jean-Jacques Szczeciniarz, Dominique Lambert et Jean-Michel Alimi pour les discussions stimulantes durant le colloque et dans d'autres occasions.

Bibliographie

[1] Berger, M., Gostiaux, B., *Géométrie différentielle : variétés, courbes et surfaces*, Presses Universitaires de France, Paris, 1987.

[2] Bennequin, D., Questions de physique galoisienne, *in Passions des formes*, Hommage à R. Thom, sous la dir. de M. Porte, ENS Editions Fontenay-St Cloud, 1994, 311–410.

[3] Boi, L., L'espace : concept abstrait et/ou physique ; la géométrie entre formalisation mathématique et étude de la nature, *in 1830-1930: A Century of Geometry. Epistemology, History and Mathematics*, Boi L., Flament D., Salanskis J.-M. (Eds.), Springer-Verlag, Heidelberg et Berlin, 1992, 65–90.

[4] Boi, L., Mannigfaltigkeit und Gruppenbegriff. Zu den Veränderungen der Geometrie im 19. Jahrhundert, *Mathematische Semesterberichte* 41(138) (1994) 1–16.

[5] Boi, L., Die Beziehungen zwischen Raum, Kontinuum und Materie im Denken Riemanns ; die Äthervorstellung und die Einheit der Physik. Das Entstehen einer neuen Naturphilosophie, *Philosophia Naturalis* 30(2) (1994) 171–216.

[6] Boi, L., *Le problème mathématique de l'espace. Une quête de l'intelligible*, Springer-Verlag, Berlin & Heidelberg, 1995.

[7] Boi, L., Géométrie de l'espace-temps et nature de la physique : quelques réflexions historiques et épistémologiques, *Manuscrito* 23(1) (2000) 31–97.

[8] Boi, L., Theories of Space-Time in Modern Physics, *Synthese* 136 (2003) 36–88.

[9] Boi, L. Geometrical and Topological Foundations of Theoretical Physics: From Gauge Theories to String program, *International Journal of Mathematics and Mathematical Sciences* (2003), à paraître.

[10] Boi, L., Space-Time and Physics, *in The Cambridge History of Philosophy 1870-1945*, T. Baldwin (Ed.), Cambridge University Press, Cambridge, 2003, 207–219.

[11] Boniolo, G. (Ed.), *Filosofia della fisica*, Bruno Mondadori, Milan, 1997.

[12] Cao, T.Y., *Conceptual Developments of 20th Century Field Theories*, Cambridge University Press, Cambridge, 1997.

[13] Cartan, E., *Leçons sur la géométrie des espaces de Riemann*, Paris, Gauthier-Villars, 1928.

[14] Chatelet, G., *Les enjeux du mobile*. Mathématique, Physique, Philosophie, Seuil, Paris, 1993.

[15] Clifford, W.K., *Lectures and Essays I*, Macmillan and Co., London, 1879.

[16] Connes, A., *Géométrie non commutative*, InterEditions, Paris, 1990.

[17] Damour, Th., General Relativity and Experiment, *in* Iagolnitzer D. (Ed.), *Proc. of the Xth Intern. Congress of Mathematical Physics*, International Press, Boston, 37–46.

[18] De Broglie, L., *La physique nouvelle et les quanta*, Flammarion, Paris, 1937.

[19] DeWitt, B.S., Graham, N., *The Many-Worlds Interpretation of Quantum Mechanics*, Princeton University Press, Princeton, 1973.

[20] Dubrovine B., Novikov S., Fomenko A., *Géométrie contemporaine. Méthodes et applications*, 2 vol., trad. française, Editions Mir, 1982.

[21] Eddington, A.S., *The Nature of the Physical World*, Cambridge University Press, Cambridge, 1929.

[22] Ehlers, J., Einführung der Raum-Zeit Struktur mittels Lichtstrahlen und Teilchen, *in Philosophie und Physik der Raum-Zeit*, J. Audretsch et K. Mainzer (Eds.), Wissenschaftsverlag, Mannheim, 1994.

[23] Feynman, R., *Leçons sur la physique*, Odile Jacob, Paris, 2000.

[24] Friedmann, A., Lemaître, G., *Essais de Cosmologie*, précédés de *L'Invention du Big Bang*, avec une introduction de J.-P. Luminet, Seuil, Paris, 1997.

[25] Green, B., *L'univers élégant*, Robert Laffont, Paris, 2000.

[26] Green, M.B., Schwarz, J.H., Witten E., *Superstring Theory*, 2 vol., Cambridge University Press, Cambridge, 1987.

[27] Gunzig, E., Du vide à l'univers, *in Le vide – Univers du tout et du rien*, édité par E. Gunzig et S. Diner, Revue de l'Université de Bruxelles, Editions Complexe, 1997.

[28] Hawking, S., Penrose, R., *The nature of space and time*, Princeton University Press, Princeton, 1996.

[29] Heisenberg, W., *Les principes physiques de la théorie des quanta*, Éditions J. Gabay, Paris, 1990.

[30] Hooft, G.'t, *In search of the ultimate building blocks*, Cambridge University Press, Cambridge, 1997.

[31] Isham, C., La gravitation quantique, *in La nouvelle physique*, sous la dir. de P. Davies, Flammarion, Paris, 1993.

[32] Itzykson, C., Zuber, J.-B., *Quantum Field Theory*, McGraw-Hill, Singapore, 1985.

[33] Kouneiher, J., Helein, F., On the Soliton-Particle Dualities, *in New Interactions of Mathematics with Natural Sciences and the Humanities*, Boi L. (Ed.), Springer-Verlag, Heidelberg et Berlin, à paraître.

[34] Lachièze-Rey, M., *Au-delà de l'espace et du temps*, Le Pommier, Paris, 2003.

[35] Landau, L., Lifchitz, E., *Physique théorique*, tome 2, *Théorie des champs*, 4^{e} éd. revue et complétée, Editions Mir, Moscou, 1989.

[36] Lochack, G., *La géométrisation de la physique*, Flammarion, Paris, 1994.

[37] Luminet, J.-P., *L'univers chiffonné*, Fayard, Paris, 1996.

[38] Manin, Yu I., *Mathematics and Physics*, Birkhäuser, Boston, 1981.

[39] Novello, M., *Le cercle du temps*, Atlantica, Paris, 2001.

[40] O'Raifeartaigh, L., *The Dawning of Gauge Theory*, Princeton Univ. Press, Princeton, 1997.

[41] Pais, A., *Inward bound. Of matter and forces in the physical world*, Oxford University Press, Oxford, 1986.

[42] Paty, M., *Einstein philosophe*, Presses Universitaires de France, Paris, 1993.

[43] Pauli, W., *Physique moderne et philosophie*, Albin Michel, Paris, 1999.

[44] Penrose, R., Structure of Space-Time, *in Battelle Rencontres*, DeWitt C.M., Wheeler J.A. (Eds.), Benjamin, New York, 1968, 121–235.

[45] Redhead, M., *From Physics to Metaphysics*, Cambridge Univ. Press, Cambridge, 1995.

[46] Riemann, B., Über die Hypothesen, welche der Geometrie zu Grunde liegen (mémoire d'Habilitation, 1954), *Abh. K.Gesell. Wiss. Gött.* 13 (1867) 133–152 ; reimprimé dans B. Riemann, *Gesammelte Mathematische Werke, Wissenschaftlicher Nachlass und Nachträge*, nouv. éd., Narasimhan R., Teubner B.G. (Eds.), Springer-Verlag, Leipzig et Berlin, 1990.

[47] Salam, A., *La grande unification. Vers une théorie des forces fondamentales ?*, Seuil, Paris, 1991.

[48] Smolin, L., *Three roads to quantum gravity*, Basic Books, New York, 2001.

[49] Souriau, J.-M., Physique et géométrie, *in La Pensée Physique Contemporaine*, Diner S., Fargue D., Lochak G. (Eds.), Ed. Augustin Fresnel, Paris, 1982, 343–364.

[50] Torretti, R., *Philosophy of Geometry from Riemann to Poincaré*, Reidel, Dordrecht, 1984.

[51] Weyl, H., *Gesammelte Abhandlungen*, vol. 2, Chandrasekharan K. (Ed.), Springer-Verlag, Heidelberg & Berlin, 1968.

[52] Weyl, H., *Raum, Zeit, Materie*, 7e éd. revue et complétée par J. Ehlers, Springer-Verlag, Berlin et Heidelberg, 1988 (1re éd., 1918).

[53] Witten, E., Topological Quantum Field Theory, *Commun. Math. Phys.* 117 (1988) 353–386.

06

Variations *N*-dimensionnelles sur des thèmes de Pythagore, Euclide et Archimède

Jean-Marc Lévy-Leblond

1. Introduction

La culture moderne se caractérise par un vif et actif intérêt pour son passé. Ce ressourcement permanent est rendu possible par la nouveauté radicale d'une situation où l'humanité, pour la première fois dans son histoire, peut accéder en permanence à la plupart de ses œuvres. Les technologies actuelles permettent aux productions littéraires, musicales, artistiques de tous les temps d'être reproduites et distribuées à bas coût, ouvrant à tout un chacun (au moins dans les parties prospères du monde) l'ensemble du patrimoine culturel humain. Mais il y a certainement une autre raison à cette archéophilie, à savoir l'affaiblissement des valeurs culturelles communes sous l'impact des profonds changements historiques que nous vivons à l'échelle de la planète entière. Dans une situation de grande incertitude quant à l'avenir, il est tout naturel de se tourner vers le passé pour y trouver inspiration et sens — et tout simplement réconfort. La Renaissance est l'exemple type d'un tel mouvement de retour riche de progrès. Sans passé, pas d'avenir. D'où l'importance d'une fréquentation toujours renouvelée des chefs d'œuvre : Euripide et Shakespeare, Cervantès et Hugo, Monteverdi et Schubert, Giotto et Delacroix nous aident toujours à vivre, aimer — et mourir. À la condition cependant que ces grandes œuvres soient effectivement re-créées (re-présentées), non dans l'illusoire intention de retrouver leur sens originel, mais, à l'opposé, d'en dégager de nouvelles significations. Il s'agit bien d'écouter, lire, regarder ce qui date d'hier avec nos oreilles, yeux et esprits d'aujourd'hui ; aussi bien, Bach n'est plus le même après Stravinsky, ni Titien après Picasso.

Si ces idées relèvent de l'évidence en ce qui concerne l'art, elles semblent moins banales quant à la science, tout au moins quand on la considère dans son rapport (ou absence de...) avec la culture. De fait, la science, au moins depuis le début du XXe siècle, s'est souvent targuée de sa modernité absolue, et a revendiqué une contemporanéité radicale, voire même une amnésie constitutive, reléguant tout intérêt pour son passé au rang de suppléments d'âme facultatifs. L'absence de culture historique chez la plupart des scientifiques d'aujourd'hui n'a d'équivalent dans aucune autre profession intellectuelle. Certes, les esprits les plus créateurs de la science entretiennent souvent une relation active avec leurs prédécesseurs, et nombre des grandes avancés du dernier siècle témoignent d'un dialogue explicite avec le passé. Einstein était parfaitement conscient de se confronter directement avec Galilée et Newton, et, pour prendre un exemple plus spécifique, Robinson, en développant l'analyse non-standard, se référait explicitement à Leibniz. Mais au niveau plus humble et plus commun de l'enseignement, de la vulgarisation, et même de la recherche, un tel lien avec le passé est pour le moins exceptionnel. Lors même que l'histoire de la science est enseignée ou convoquée, c'est en général sans connexion avec la pratique scientifique effective. Cela est fort regrettable, car des avancées modernes peuvent donner à des résultats anciens de significations inédites et des extensions nouvelles qui éclairent ces développements récents, tout comme une production moderne d'*Antigone* ou du *Roi Lear* peut révéler des sens nouveaux et avoir un impact actuel.

Après cette bien trop pompeuse entrée en matière, je voudrais offrir quelques exemples tirés de ce qui est, après tout, l'un des plus vieux métiers du monde, à savoir la géométrie, conçue à la physicienne, c'est-à-dire comme une « mesure de l'espace ». On montera comment certains résultats classiques (et même antiques) de géométrie dans l'espace usuel trouvent d'intéressantes généralisations dans l'espace à N-dimensions, ce qui, on l'espère, à la fois montera la longue portée de ces théorèmes traditionnels et, peut-être, aidera à forger une meilleure intuition des hautes dimensionalités spatiales. En fonction de l'amnésie collective incriminée plus haut, il est très difficile de savoir ce qui, dans les développements présentés ici peut prétendre à une relative originalité, sinon sur le fond, du moins quant à la forme. Que la plupart des collègues mathématiciens et physiciens consultés n'aient pu donner de références relève sans doute d'abord de notre absence de mémoire collective. De fait, seul un heureux hasard m'a permis de retrouver l'un au moins des résultats de ce papier (section 2) dans un article déjà ancien (1935) du grand géomètre H. S. M. Coxeter. C'est avec intérêt et gratitude que j'accueillerais toute information historique sur les questions abordées ici.

2. Pythagore et l'orthosimplexe à *N* dimensions

Commençons par revisiter le très ancien théorème de Pythagore, que nous allons généraliser en dimension quelconque. Bien entendu, il ne s'agit pas ici de l'interprétation métrique de ce théorème, désormais transformé en axiome définitoire des espaces euclidiens, mais d'un résultat plaisant sinon profond[1]. Présentons d'abord l'énoncé classique sous une forme légèrement modifiée qui permettra son extension.

Soit donc dans le plan euclidien deux droites orthogonales se coupant en O. Pour tout segment AB appuyé sur ces deux droites (figure 6.1), le carré de sa longueur est somme des carrés des longueurs des deux segments OA et OB qu'il intercepte sur les droites — ses projections orthogonales :

$$AB^2 = OA^2 + OB^2. \tag{6.1}$$

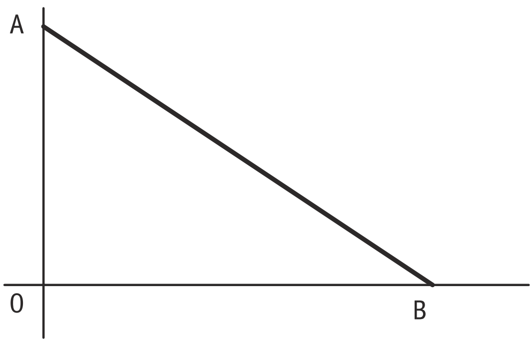


FIG. 6.1 – Pythagore dans le plan.

Considérons maintenant dans l'espace euclidien à trois dimensions, trois droites orthogonales concourant en O. Soit ABC un triangle ayant un sommet

1 – J'avais d'abord écrit « et apparemment inédit », avant de trouver par chance une référence à l'article [3]. Les auteurs y démontrent exactement le même résultat, mais avec une méthode peut-être moins générale.

sur chacune de ces droites, et se projetant respectivement en OAB, OBC et OCA sur les plans définis par deux de ces trois droites (figure 6.2). Alors :

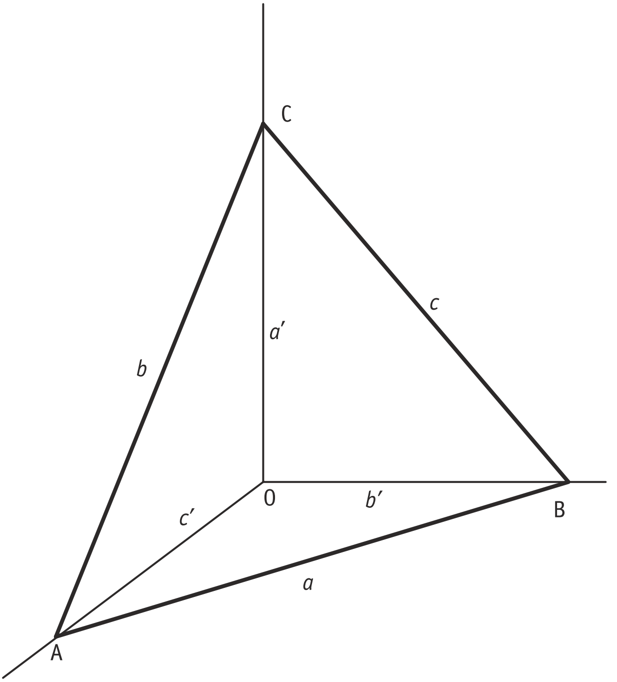


Fig. 6.2 – Pythagore dans l’espace.

Théorème de Pythagore 3-D. *Le carré de l’aire du triangle ABC est égal à la somme des carrés des aires de ses trois projections :*

$$(\text{Aire ABC})^2 = (\text{Aire OAB})^2 + (\text{Aire OBC})^2 + (\text{Aire OCA})^2 \,. \qquad (6.2)$$

On peut le prouver sans trop de mal à partir de la formule de Héron donnant l’aire d’un triangle quelconque en termes des longueurs (a, b, c) de ses côtés :

$$A^2 = \frac{1}{4}(a+b+c)(b+c-a)(c+a-b)(a+b-c) = \frac{1}{4}(2b^2c^2-a^4+\text{perm. circ.}). \quad (6.3)$$

Le théorème de Pythagore (standard) fournit les côtés du triangle ABC en fonction de ceux de ses projections :

$$a^2 = b'^2c'^2, \quad b^2 = c'^2 + a'^2, \quad c^2 = a'^2 + b'^2 \qquad (6.4)$$

avec les notations de la figure 6.2. Soit donc, en reportant dans (6.2)

$$A^2 = \frac{1}{4}\left(a'^2 b'^2 + b'^2 c'^2 + c'^2 a'^2\right)$$

ce qui est bien le résultat annoncé (6.1).

En vérité, comme souvent, une démonstration plus pertinente permet de comprendre la généralité de ce théorème. L'aire orientée du triangle ABC s'exprime en effet comme un vecteur orthogonal à son plan *via* le produit vectoriel de deux (quelconques) de ses côtés :

$$\mathbf{A} = \frac{1}{2}\mathbf{a} \wedge \mathbf{b}. \tag{6.5}$$

Il suffit de noter que

$$\mathbf{a} = \mathbf{c}' - \mathbf{b}', \; \mathbf{b} = \mathbf{a}' - \mathbf{c}', \; \mathbf{c} = \mathbf{b}' - \mathbf{a}' \tag{6.6}$$

avec des notations vectorielles évidentes, pour voir que

$$\mathbf{A} = \frac{1}{2}(\mathbf{c}' - \mathbf{b}') \wedge (\mathbf{a}' - \mathbf{c}') = \frac{1}{2}\mathbf{c}' \wedge \mathbf{b}' + \frac{1}{2}\mathbf{a}' \wedge \mathbf{c}' + \frac{1}{2}\mathbf{b}' \wedge \mathbf{a}'. \tag{6.7}$$

Les trois termes sont les aires vectorielles respectives des trois triangles OAB, OBC et OCA. Puisqu'elles sont mutuellement orthogonales, le théorème de Pythagore sous sa forme usuelle fournit immédiatement le résultat (6.2).

Mais cette démonstration elle-même relève d'une interprétation plus générale. Appelons « aire vectorielle » d'une surface Σ, le vecteur

$$\mathbf{A}_\Sigma := \int_\Sigma \mathbf{n}\, d^2 s \tag{6.8}$$

où $\mathbf{n}$ est la normale à la surface. Alors, pour toute surface fermée, l'aire vectorielle totale est nulle. L'interprétation physique en est simple[2] : pour un vecteur quelconque $\mathbf{u}$, la quantité $\mathbf{u}.\mathbf{A}_\Sigma$ est le flux du champ constant $\mathbf{u}$ à travers la surface fermée Σ, quantité nulle (on peut, si nécessaire pour s'en convaincre transformer ce flux en l'intégrale de volume de sa divergence — nulle). Appliqué à un tétraèdre quelconque, ce résultat veut donc dire que la somme des aires vectorielles des quatre faces est nulle. Si trois de ces faces sont deux à deux orthogonales, ce qui est le cas du tétraèdre OABC considéré ici, il en résulte que l'aire vectorielle de la quatrième face est égale à la somme vectorielle des aires des trois autres faces, mutuellement orthogonales, d'où,

2 – Merci à Jean-Paul Marmorat pour cette remarque.

d'après le théorème de Pythagore sous sa forme tridimensionnelle usuelle, le résultat (6.2).

Soit maintenant N droites orthogonales concourantes en O dans un espace euclidien à N dimensions, et un N-simplexe formé par N points $(\mathrm{P}_1, \mathrm{P}_2, \ldots \mathrm{P}_N)$ pris respectivement sur chaque droite. En y adjoignant le point O, on obtient un $(N+1)$-simplexe, dont tous les angles en O sont droits, et que nous appellerons donc « orthosimplexe ». Sa face $(N-1)$-dimensionnelle $\mathrm{H} := (\mathrm{P}_1, \mathrm{P}_2, \ldots \mathrm{P}_N)$ sera appelée face « hypoténusale », et ses N faces $\mathrm{F}_k := (\mathrm{O}, \mathrm{P}_1, \mathrm{P}_2, \ldots \cancel{\mathrm{P}_k}, \ldots \mathrm{P}_N)$[3], avec $k = 1, 2, \ldots N$, seront appelées « faces droites » (ce sont les projections de la face hypoténusale parallèlement aux axes — ce sont également des orthosimplexes, en dimension $(N-1)$).

Alors : **théorème de Pythagore N-D** *Le carré du volume $(N-1)$-dimensionnel de la face hypoténusale d'un orthosimplexe est égal à la somme des carrés des volumes de ses N faces droites* :

$$(\mathrm{Vol}_{N\text{-}1}\ \mathrm{H})^2 = \sum_1^N (\mathrm{Vol}_{N\text{-}1}\ \mathrm{F}_k)^2. \tag{6.9}$$

La démonstration est quasiment triviale dès lors que l'on recourt au calcul extérieur. Notant $\mathbf{r}_k = \mathrm{OP}_k$ le vecteur définissant le point P_k ($k = 1, 2, \ldots N$), on peut exprimer le volume $(N-1)$-dimensionnel du N-simplexe $\mathrm{P}_1\mathrm{P}_2\ldots\mathrm{P}_N$ sous forme d'un vecteur orthogonal au sous-espace de dimension $(N-1)$ qui le contient, en prenant le produit extérieur de $(N-1)$ de ses arêtes, par exemple celles qui partent du point P_1 :

$$\mathbf{A} = (\mathbf{r}_2 - \mathbf{r_1}) \wedge (\mathbf{r}_3 - \mathbf{r_1}) \wedge \ldots \wedge (\mathbf{r}_N - \mathbf{r_1}). \tag{6.10}$$

Il est aisé de voir en utilisant une méthode de récurrence, que l'antisymétrie du produit extérieur permet d'écrire le membre de droite comme une somme des N termes donnés par le produit extérieur, au signe près, de $(N-1)$ des N vecteurs $\mathbf{r}_k$ ($k = 1, 2, \ldots N$) :

$$\mathbf{A} = (\mathbf{r}_2 \wedge \mathbf{r}_3 \wedge \ldots \wedge \mathbf{r}_N) - (\mathbf{r}_1 \wedge \mathbf{r}_3 \wedge \ldots \wedge \mathbf{r}_N) \ldots \pm (\mathbf{r}_1 \wedge \mathbf{r}_2 \wedge \ldots \wedge \mathbf{r}_{N-1}) \tag{6.11}$$

Le premier de ces termes exprime le volume de la face F_1, le N-simplexe $\mathrm{OP}_2\mathrm{P}_3\ldots\mathrm{P}_N$, sous forme d'un vecteur orthogonal $\mathbf{A}_1$ à cette face. Il en va de même pour chacun des autres termes correspondant respectivement à chaque face droite. On a donc (avec un choix adéquat des orientations) :

$$\mathbf{A} = \mathbf{A}_1 + \mathbf{A}_2 + \ldots + \mathbf{A}_N. \tag{6.12}$$

3 – La notation indique clairement que la face F_k est définie par les points $(\mathrm{O}, \mathrm{P}_1, \mathrm{P}_2, \ldots \mathrm{P}_N)$ à l'exclusion du point P_k.

Exactement comme dans le cas tridimensionnel, ce résultat vaut pour tout simplexe.

Mais le cas de l'orthosimplexe considéré ici a ceci de particulier que les vecteurs correspondants aux volumes des projections (les faces droites) étant orthogonaux ($\mathbf{A}_k$ est parallèle à l'axe k), le théorème de Pythagore usuel implique la formule (6.9).

Notons cependant que la généralité de ce résultat est limitée par le fait que les N-simplexes définis par N points sur N droites orthogonales sont, dès que $N > 3$, très particuliers. De fait, ils sont définis par la donnée des N paramètres $(r_1, r_2 \ldots, r_N)$, alors que le N-èdre le plus général dépend de $N(N-1)/2$ paramètres (la longueur de ses arêtes par exemple).

3. Euclide et la pyramide à N dimensions

La maîtrise conceptuelle de l'espace passe par la géométrie, conçue, au sens premier, comme une science de la mesure des formes, c'est-à-dire au fond une physique de l'espace. L'estimation des longueurs, aires et volumes des objets les plus simples de l'espace tridimensionnel en est la base [5]. On ne saurait à cet égard surestimer le rôle des formules permettant de calculer l'aire d'un triangle et le volume d'une pyramide ; ce sont, après tout, les premiers exemples d'intégrations à 2 et 3 dimensions, comme on le comprendra bien plus tard. Or le raisonnement géométrique qui permet d'établir ces deux résultats se généralise aisément à N dimensions, fournissant à la fois une nouvelle démonstration de résultats élémentaires, et, sans doute, une approche qui permet de développer quelque intuition N-dimensionnelle.

Le cas du triangle est vite réglé, puisque tout triangle a, de façon évidente, une aire égale à la moitié de celle du rectangle de même base et de même hauteur (figure 6.3). Mais, pour une généralisation aisée, on préfère partir ici du cas particulier d'un triangle rectangle isocèle, moitié du carré de même côté. La formule qui donne son aire :

$$\text{Aire Triangle} = \frac{1}{2}\ \text{Hauteur} \times \text{Longueur Base} \tag{6.13}$$

s'étend alors à un triangle quelconque par des transformations affines adéquates, qui biaisent le triangle en dilatant ses dimensions, tout en maintenant la formule (6.13) et en en respectant le facteur 1/2 (figure 6.4).

L'extension à 3 dimensions peut se faire aisément en considérant un cube, et en utilisant sa symétrie d'ordre 3 autour d'un axe joignant deux sommets

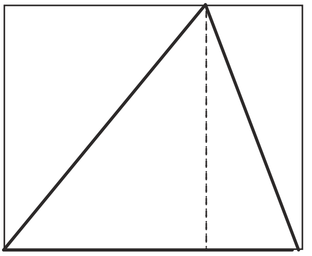


Fig. 6.3 – L'aire du triangle.

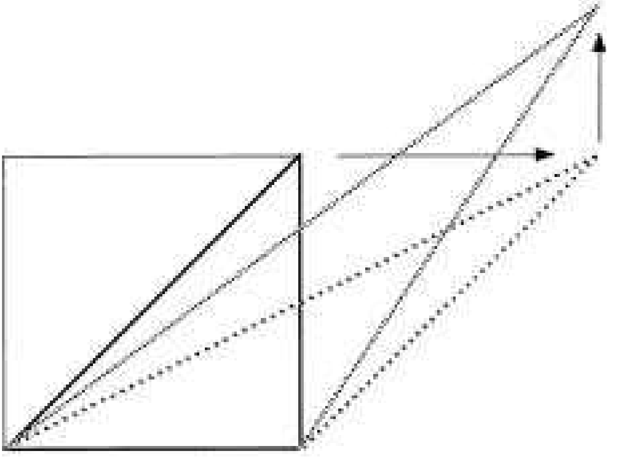


Fig. 6.4 – Encore l'aire du triangle.

opposés. Un astucieux découpage en trois pyramides égales jointes le long d'une telle grande diagonale montre que le volume d'une pyramide droite de base carrée et de hauteur égale au côté du carré est bien le tiers du volume du cube (figure 6.5) :

$$\text{Volume Pyramide} = \frac{1}{3}\ \text{Hauteur} \times \text{Aire Base}. \tag{6.14}$$

À nouveau, des transformations affines adéquates permettent de transformer cette pyramide en n'importe quelle autre de base parallélipipédique. Enfin, une comparaison à la Cavalieri d'une telle pyramide avec une autre de même hauteur et de base quelconque, mais ayant même aire que la première

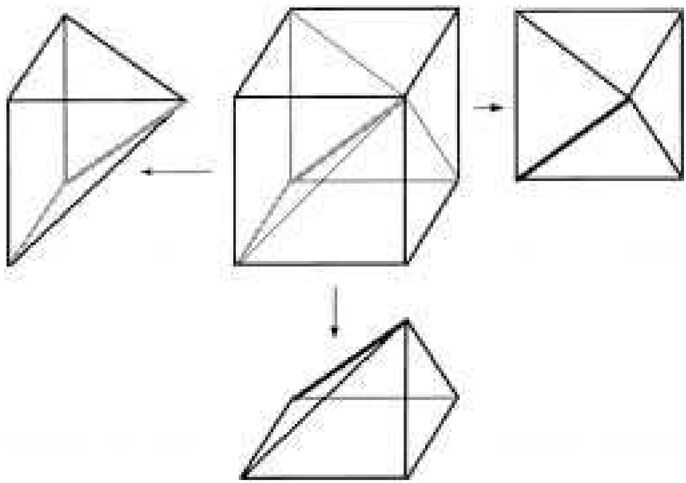


FIG. 6.5 – La dissection du cube et le volume de la pyramide.

(figure 6.6), achève de montrer que la formule (6.14) vaut pour n'importe quelle pyramide[4].

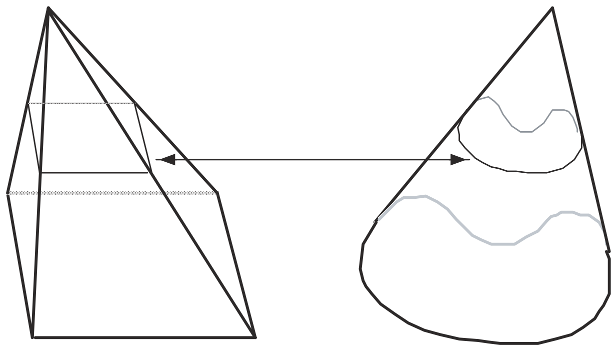


FIG. 6.6 – Le volume de la pyramide quelconque.

Au vu du facteur 1/2 qui s'introduit en dimension 2 et du facteur 1/3 qui intervient en dimension 3, la généralisation est tentante. En dimension 4, la chose est encore visualisable. L'hypercube de dimension 4 (souvent appelé « tessaract » dans les ouvrages ésotériques) se laisse aisément représenter

4 – L'utilisation sans vergogne, « à la physicienne », d'arguments infinitésimaux implicites, permet ici de court-circuiter brutalement la solution négative du troisième problème de Hilbert : l'impossibilité (contrairement au cas à deux dimensions) d'un découpage fini de tétraèdres de même volume en éléments congruents.

projeté en dimension 2 (figure 6.7). On *voit* ainsi qu'il est possible de le découper, effectivement, en 4 hyperpyramides égales, à base cubique (tridimensionelle), jointes le long d'une grande diagonale de l'hypercube. Le volume de chacune est donc égal au quart du volume de l'hypercube, et ce résultat se généralise à n'importe quelle hyperpyramide, comme ci-dessus.

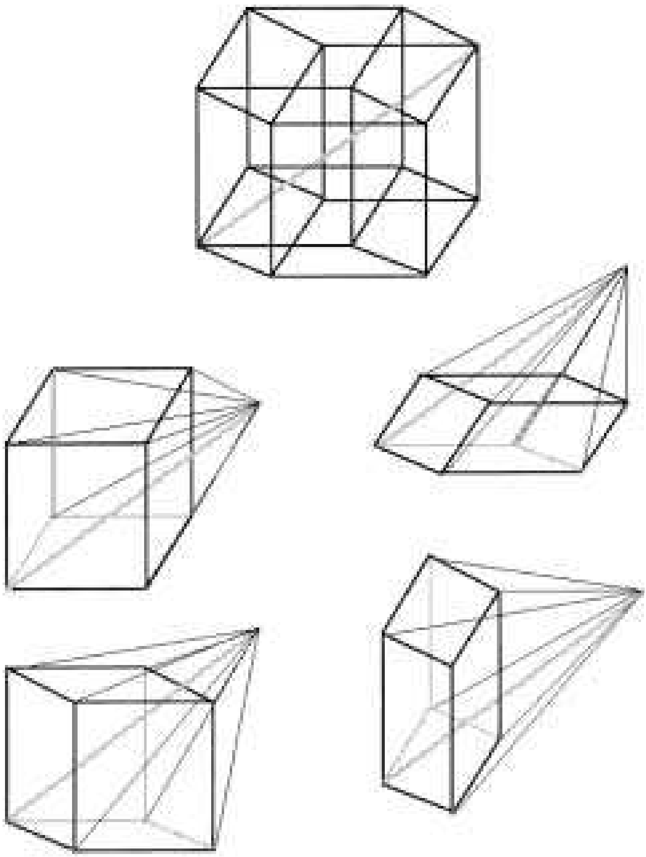


FIG. 6.7 – La dissection de l'hypercube et le volume de l'hyperpyramide.

Une assez simple démonstration formelle en dimension N quelconque se présente alors. Soit dans l'espace euclidien $\mathcal{E}^N$ muni d'une base orthonormée de sommet 0 l'hypercube $\Gamma^{(N)}$ de côté unité dont les 2^N sommets sont définis par leurs coordonnées $(a_1, a_2, \ldots a_N)$, avec $a_k = 0$ *ou* 1 $(k = 1, 2, \ldots N)$. Notons $\Gamma_k^{(N-1)}$ la face hypercubique $(N-1)$-dimensionnelle de $\Gamma^{(N)}$ définie par $a_k = 1$; ces N hypercubes sont les N faces de $\Gamma^{(N)}$ qui se rejoignent au sommet $0' = (1, 1, \ldots 1)$ opposé à l'origine $0 = (0, 0, \ldots 0)$. Et considérons enfin les N hyperpyramides $\Pi_k^{(N)}$ ($k = 1, 2, \ldots N$) ayant pour sommet 0 et pour bases respectives les N hypercube $\Gamma_k^{(N-1)}$. Ces hyperpyramides égales, qui ont $00'$ pour arête commune, fournissent une partition en N volumes égaux de l'hypercube $\Gamma^{(N)}$. En effet, tout point $(x_1, x_2, \ldots x_N) \in \Gamma^{(N)}$ (c'est-à-dire tel que $0 \leq x_k \leq 1 \ \forall k = 1, 2, \ldots N$), appartient à une et une seule hyperpyramide, à savoir $\Pi_l^{(N)}$ telle que $x_l = \sup\,(x_1, x_2, \ldots x_N)$. Il en résulte évidemment que le volume de chacune de ces hyperpyramides est égal à $1/N$. L'extension de ce résultat à une hyperpyramide quelconque se fait comme ci-dessus, démontrant donc bien que

$$\text{Volume Hyperpyramide}^{(N)} = \frac{1}{N}\,\text{Hauteur} \times \text{Volume Hyperbase}^{(N-1)}. \qquad (6.15)$$

En termes algébriques, ce résultat est évidemment équivalent à l'intégration d'un monôme de puissance $(N-1)$. En effet, considérons la section de l'hyperpyramide $\Pi_N^{(N)}$ (par exemple) par le plan $x_N = t$ $(0 < t < 1)$: c'est l'hypercube de dimension $(N-1)$, ayant pour sommets les points $(ta_1, ta_2, \ldots ta_{N-1}, t)$ avec $a_k = 0$ *ou* 1 $(k = 1, 2, \ldots N-1)$; son volume vaut t^{N-1}. Le volume total de la pyramide s'obtient donc en intégrant sur la cote t du plan de section. En définitive, nous avons montré par des moyens purement géométriques que

$$\int_0^1 dt\, t^{N-1} = \frac{1}{N}. \qquad (6.16)$$

Le calcul du volume de la pyramide et l'obtention du facteur 1/3 par trisection pyramidale du cube n'est pas exactement la méthode présentée par Euclide. Suivant une démonstration plus ancienne due à Eudoxe, mais dont l'idée est généralement attribuée à Démocrite[5], c'est la trisection d'un prisme en pyramides de même volume (quoique inégales) qui fournit le résultat. On laissera au lecteur le plaisir de s'assurer que cette méthode se généralise aussi à N dimensions.

4. Archimède et la sphère à N dimensions

4.1. Le double saut de π

Parmi les nombreux mystères qu'offre le nombre π [4], le moindre n'est pas le suivant : non content de régir la circonférence et la surface du cercle[6],

5 – Voir la longue note de Thomas L. Heath, [6] vol. III, pp. 365-368.

6 – . . .comme on disait quand j'étais petit. Mais je sais, il faut dire maintenant « l'aire du disque » — c'est l'air du temps.

le même nombre intervient dans les formules donnant l'aire de la sphère et son volume[7]. Mais il y a mieux : les expressions de l'aire et du volume de la sphère à N dimensions n'exigent pas d'autre nombre transcendant, ce qui serait après tout, ou plutôt avant tout calcul, concevable. Ces formules se contentent de faire appel à π, élevé à une puissance qui saute de deux en deux avec la dimensionalité, généralisant ainsi le comportement observé en dimensions 2 et 3. Le tableau 6.1 exhibe cet étrange phénomène[8] :

TAB. 6.1 – Aire et volume de la sphère à N dimensions.

N	0	1	2	3	4	5	6	7
A_N	0	2	$2\pi R$	$4\pi R^2$	$2\pi^2 R^3$	$\frac{8}{3}\pi^2 R^4$	$\pi^3 R^5$	$\frac{16}{15}\pi^3 R^6$
V_N	1	$2R$	πR^2	$\frac{4}{3}\pi R^3$	$\frac{1}{2}\pi^2 R^4$	$\frac{8}{15}\pi^2 R^5$	$\frac{1}{6}\pi^3 R^6$	$\frac{16}{105}\pi^3 R^7$

Le calcul explicite le plus courant n'éclaire guère ces résultats. Le « truc » classique (un jeu de Gauss...) pour calculer l'aire de la sphère en dimension N consiste à intégrer dans l'espace euclidien une fonction gaussienne qui présente la double propriété spécifique de se factoriser en fonctions (gaussiennes) d'une seule variable et d'être à symétrie sphérique.

La première de ces propriétés permet d'écrire :

$$\int_{\mathfrak{R}_N} d^N r \, \exp(-r^2) = \int_{\mathfrak{R}} \ldots \int_{\mathfrak{R}} dx_1\, dx_2 \ldots dx_N \exp(-x_1^2 \ldots - x_N^2)$$
$$= \left[\int_{\mathfrak{R}} dx \, \exp(-x^2) \right]^N = [\Gamma(1/2)]^N = \pi^{N/2}. \qquad (6.17)$$

La seconde fournit l'expression :

$$\int_{\mathfrak{R}_N} d^N r \, \exp(-r^2) = \int_0^\infty a_N\, r^{N-1} dr \exp(-r^2) = \tfrac{1}{2} a_N \Gamma(N/2). \qquad (6.18)$$

D'où le résultat final (pour le volume, on utilise (6.5)) :

$$a_N = 2\frac{[\Gamma(1/2)]^N}{\Gamma(N/2)} \quad \text{et} \quad v_N = 2\frac{[\Gamma(1/2)]^N}{N\,\Gamma(N/2)} = \frac{[\Gamma(1/2)]^N}{\Gamma(1+N/2)}. \qquad (6.19)$$

7 – Pardon : le volume de la boule.

8 – « ...seeing that the circumference of a circle is $2\pi R$ while the surface of a sphere is $4\pi R^2$, we might be tempted to expect that the hypersurface of a hypersphere [in 4 dimensions] to be $6\pi R^3$ or $8\pi R^3$. It is unlikely that the use of analogy, unaided by computation, would ever lead us to the correct expression, $2\pi^2 R^3$. » H. S. M. Coxeter [2], p. 119. Lors de son cours de Méthodes Mathématiques de la Physique, dans les années 1960, Laurent Schwartz se taillait toujours un joli succès en faisant part de son regret de ne pas vivre dans un espace à 6 dimensions où la formule de la surface est « si belle ».

On peut expliciter, en distinguant suivant le caractère pair ou impair de la dimension. Dans le premier cas, $N = 2p$, le dénominateur est une factorielle à valeur entière, et il y a au numérateur N facteurs $\sqrt{\pi}$, soit p facteurs π. Dans le second cas, il s'introduit au dénominateur un facteur $\sqrt{\pi}$ qui compense l'un des $(2p+1)$ facteurs $\sqrt{\pi}$ du numérateur, ne laissant finalement encore que p facteurs π. C'est de cette compensation, qui semble ici parfaitement contingente, que surgit le « double saut » de π, sans que sa nécessité géométrique soir compréhensible.

On voudrait ici présenter deux autres méthodes de calcul permettant d'éclairer le mystère du double saut des puissances de π, dont l'une est fondée sur une très simple et élégante formule de récurrence qui a la vertu de généraliser des résultats vieux de plus de deux millénaires.

4.2. À 3 dimensions (Archimède)

Commençons par rappeler le calcul classique de l'aire de la sphère ordinaire (celle de *notre* espace). C'est Archimède qui montra le premier [1] que l'aire d'une sphère est exactement égale à celle de l'aire latérale du cylindre circonscrit de même rayon R et de hauteur égale au diamètre de la sphère (figure 6.8).

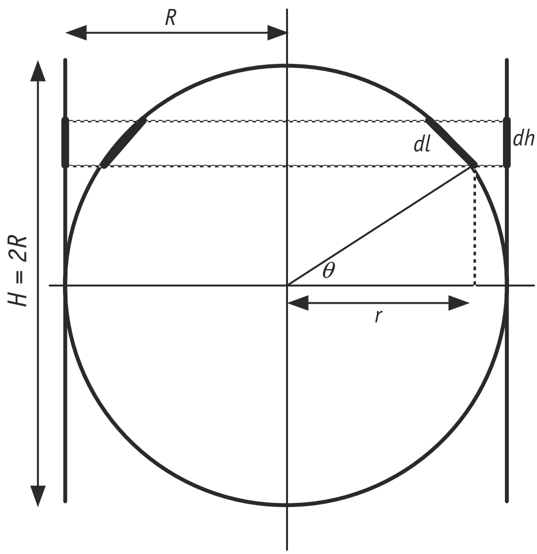


FIG. 6.8 – L'aire de la sphère.

La démonstration la plus simple consiste à considérer sur la sphère de rayon R, le parallèle de latitude définie par l'angle polaire θ. La zone infinitésimale comprise entre ce parallèle et un autre distant (sur la sphère) de dl, a pour aire $dA = 2\pi R \sin\theta\, dl$, puisque le rayon du cercle parallèle à cette latitude est $r = R \sin\theta$. Mais les plans parallèles contenant les deux cercles découpent sur le cylindre circonscrit à l'équateur un ruban de hauteur $dh = \sin\theta\, dl$. L'aire de ce ruban est donc $dA' = 2\pi R\, dh = dA$, égale à celle de la zone sphérique infinitésimale. La sphère et le cylindre ont ainsi même aire totale, d'où :

$$A = 2\pi RH = 4\pi R^2. \tag{6.20}$$

Quant au volume intérieur de la sphère, il suffit de le considérer comme découpé en pyramides infinitésimales ayant leurs sommets au centre et leurs bases sur la surface ; pour chacune de ces pyramides, on a évidemment $dV = 1/3 R\, dA$, et donc pour la sphère entière (après tout une sorte de pyramide généralisée), $V = 1/3 R\, A$, soit

$$V = \frac{4\pi}{3} R^3. \tag{6.21}$$

Archimède formulait ce résultat en énonçant l'égalité des volumes de la sphère et du cylindre évidé, privé des cônes ayant le centre pour sommet appuyés sur les cercles de base du cylindre[9]. Archimède était, dit-on, si fier (à juste titre) de ces résultats qu'il avait demandé que fût gravée sur sa tombe la figure-clé de la sphère avec son cylindre circonscrit et le cône. Trois siècles après son assassinat lors du siège de Syracuse, Cicéron avait pu reconnaître à ce signe sa sépulture, hélas disparue depuis.

La correspondance établie entre la sphère et le cylindre circonscrit n'est autre que la projection cartographique « cylindrique parallèle » (ou « normale »), dont nous avons montré au passage qu'elle respecte les aires localement (et pas seulement globalement) ; mais, comme on sait, elle distord les angles, d'autant plus que l'on se rapproche des pôles où elle devient singulière[10].

9 – On peut démontrer directement cette égalité en remarquant l'égalité des aires des sections de ces deux volumes par tout plan perpendiculaire à l'axe du cylindre (respectivement un disque et un anneau). La théorie des indivisibles à la Cavalieri (mais connue dès avant le XVII^e^ siècle), forme heuristique d'une théorie rigoureuse de l'intégration, suffit alors pour affirmer que les deux solides (la sphère et le cylindre évidé), formés de feuilletages de même aire, ont même volume.

10 – Contrairement à une idée reçue, cette projection n'est pas celle de Mercator. Ce grand cartographe inventa en 1569 une projection cylindrique conforme (respectant les angles), beaucoup moins triviale. Très curieusement, et malgré l'antique résultat archimédien, la projection cylindrique normale ne semble pas avoir été utilisée en cartographie avant les travaux de J. H. Lambert. Voir John P. Snyder [7].

4.3. À N dimensions

Considérons, dans l'espace euclidien $\mathcal{E}_N$ de dimension N, la sphère $\mathcal{S}_N$ de rayon R, dont nous notons $A_N(R)$ l'aire et $\mathcal{B}_N$ sa boule intérieure de volume $V_N(R)$. Notant encore a_N et v_N l'aire et le volume de la sphère de rayon unité, on a pour des raisons d'homogénéité évidentes :

$$\begin{cases} A_N(R) = a_N R^{N-1} \\ V_N(R) = v_N R^N \end{cases} . \tag{6.22}$$

Enfin, une procédure d'intégration élémentaire, généralisant immédiatement le cas de l'espace à 3 dimensions, permet d'écrire :

$$V_N(R) = \int_0^R A_N(r)\, dr = \int_0^R a_N\, r^{N-1}\, dr = a_N \frac{1}{N} R^N, \tag{6.23}$$

d'où

$$v_N = \frac{1}{N} a_N, \text{ ou encore } V_N(R) = \frac{1}{N} R\, A_N(R). \tag{6.24}$$

Après ces préliminaires, venons-en au cœur de notre argumentation, qui consiste à établir une projection cylindrique archimédienne généralisée. À cette fin, considérons le point courant $r = (x_1, x_2, \ldots, x_N) \in \mathcal{E}_N$, et faisons un paramétrage polaire de ses deux premières coordonnées en posant :

$$\begin{cases} x_1 = \rho \;\cos\varphi \\ x_2 = \rho \;\sin\varphi \\ (x_3, x_4, \ldots, x_N) = r' \in \;\mathcal{E}_{N-2} \end{cases} . \tag{6.25}$$

Avec ces nouvelles coordonnées, l'élément de volume s'écrit

$$d^N r = dx_1\, dx_2 \ldots dx_N = \rho\; d\rho\; d\varphi\; d^{N-2} r'. \tag{6.26}$$

On établit alors une bijection entre la sphère $\mathcal{S}_N$ (privée de ses pôles pour éviter la singularité du paramétrage en $\rho = 0$) et le produit cartésien de la boule $\mathcal{B}_{N-2}$ (paramétrée par r', avec $|r'| < R$) et de la sphère $\mathcal{S}_2$ (c'est-à-dire le cercle paramétré par l'angle φ), autrement dit, un tore :

$$\left\{ r = (\rho, \varphi, r') \;;\; |r| = R \right\} \in \mathcal{S}_N \leftrightarrow \begin{cases} \varphi \in \mathcal{S}_2 \\ r' \in \mathcal{B}_{N-2} \;;\; |r'| < R \\ \rho = \sqrt{R^2 - |r'|^2} \end{cases} . \tag{6.27}$$

Il s'agit bien d'une généralisation de la projection cylindrique habituelle : dans le cas de la dimension 3, la sphère ($\mathcal{S}_3$) est projetée sur le cylindre, produit cartésien du cercle ($\mathcal{S}_2$) et du segment ($\mathcal{B}_1$), que l'on peut aussi bien considérer comme un tore. Dans le cas général aussi, la projection conserve la mesure, puisque l'intégrale sur la sphère $\mathcal{S}_N$ (munie de sa mesure uniforme) d'une fonction F quelconque, s'écrit, à l'aide de la distribution de Dirac δ :

$$\begin{aligned}\int_{\mathcal{S}_N} d^{N-1}\sigma\, F(r) &= \int_{\Re_N} d^N r\, \delta(|r| - R)\, F(r)\\ &= \int \rho\, d\rho\, d\varphi\, d^{N-2}r'\, \delta\left(\sqrt{|r'|^2 + \rho^2} - R\right)\, F(\varphi, r')\\ &= R \int d\varphi \int_{|r'|<R} d^{N-2}r'\, F(\varphi, r'),\end{aligned} \tag{6.28}$$

où l'on a utilisé le résultat classique $\delta[u(x) - u(a)] = [u'(a)]^{-1}\, \delta(x - a)$.

En intégrant la fonction unité ($F = 1$), on obtient l'aire de la sphère $\mathcal{S}_N$ comme produit de l'aire de la sphère $\mathcal{S}_2$ par le volume de la boule $\mathcal{B}_{N-2}$, soit la charmante formule :

$$A_N(R) = 2\pi\, R\, V_{N-2}(R), \quad \text{ou encore } a_N = 2\pi\, v_{N-2}. \tag{6.29}$$

Cette formule constitue notre résultat essentiel ; malgré sa simplicité, je ne l'ai trouvée dans aucun livre[11].

Compte-tenu de (6.24), on peut alors écrire les formules de récurrence suivantes :

$$a_N = \frac{2\pi}{N-2} a_{N-2} \quad \text{et} \quad v_N = \frac{2\pi}{N} v_{N-2} \tag{6.30}$$

qui conduisent très aisément au tableau 6.1[12], et permettent également de retrouver les formules (6.19).

Ces résultats permettent de comprendre immédiatement le saut de 2 en 2 de la puissance de π. En effet, les dimensions 0 et 1, triviales, ne font pas intervenir π ; ensuite, on passe de 0 à toutes les puissances paires successives avec intervention d'un facteur π supplémentaire à chaque étape (on peut d'ailleurs noter que la formule (6.29) vaut déjà pour $n = 2$), et de même, à partir de 1, pour les puissances impaires.

11 – Coxeter ([2] p. 126) donne une formule équivalente (la première de (6.30) en fait), mais en la déduisant des expressions générales établies par la méthode « gaussienne » (*cf.* Appendice), et sans en commenter le sens géométrique.

12 – On note en particulier le cas de la sphère à 4 dimensions, dont l'aire est bien égale au volume du tore à 3 dimensions ; la cartographie correspondante peut d'ailleurs être un outil géométrique intéressant.

4.4. Un π par plan !

On peut jeter un autre éclairage sur cette curieuse différence de nature entre dimensionalité paire et dimensionalité impaire de l'espace qui sous-tend notre problème. Le volume de la sphère $\mathcal{S}_N$ peut s'écrire :

$$V_N = \int_{\mathcal{E}_N} d^N r\, \theta\left(R^2 - r^2\right) = \int_{\mathcal{E}_N} dx_1\, dx_2 \ldots dx_N\, \theta\left(R^2 - x_1^2 - x_2^2 \ldots - x_N^2\right) \quad (6.31)$$

où θ est la fonction de Heaviside.

Supposons la dimension *paire*, $n = 2p$. On fait alors un changement de variable consistant à passer en coordonnées polaires dans chacun des p plans définis par deux axes de coordonnées :

$$x_{2k-1} = \rho_k \cos\varphi_k, \;\; x_{2k} = \rho_k \sin\varphi_k \qquad (k = 1, 2, \ldots p). \qquad (6.32)$$

On note encore

$$t_k = \rho_k^2 \qquad (k = 1, 2, \ldots p). \qquad (6.33)$$

Il vient alors immédiatement, d'après (6.31) :

$$V_{2p} = \int_{\mathcal{E}_N} \prod_{k=1}^{p} \rho_k d\rho_k d\varphi_k\, \theta\left(R^2 - \sum_1^p \rho_k^2\right) = \left(\frac{2\pi}{2}\right)^p \int \prod_{k=1}^{p} dt_k\, \theta\left(R^2 - \sum_1^p t_k\right), \qquad (6.34)$$

où le domaine d'intégration dans l'espace à p dimensions est l'intérieur de la pyramide définie par les demi-axes positifs ($0 < t_k$, $k = 1, 2, \ldots p$) et par l'hyperplan les coupant en $t_k = R^2$ ($k = 1, 2, \ldots p$). Le volume correspondant est $(p!)^{-1} R^{2p}$, et l'on obtient la formule finale :

$$V_{2p} = \frac{\pi^p}{p!}\, R^{2p}. \qquad (6.35)$$

Si la dimension est *impaire*, $n = 2p + 1$, on fait le même changement de variables que précédemment dans les p plans correspondant à p couples de coordonnées, mais il reste une coordonnée célibataire. Il vient alors, au lieu de (6.34) :

$$\begin{aligned} V_{2p+1} &= \left(\frac{2\pi}{2}\right)^p \int \left(\prod_{k=1}^{p} dt_k\right) dx_{2p+1}\, \theta\left(R^2 - \sum_1^p t_k - x_{2p+1}^2\right) \\ &= \frac{\pi^p}{p!} \int_{-R}^{R} \left(R^2 - x_{2p+1}^2\right)^p dx_{2p+1} = 2\frac{\pi^p}{p!} R^{2p+1} \int_0^1 \left(1 - u^2\right)^p du, \end{aligned} \qquad (6.36)$$

soit enfin :

$$V_{2p+1} = \frac{2^{2p+1} p!\ \pi^p}{(2p+1)!} R^{2p+1}, \tag{6.37}$$

conformément aux expressions générales.

On peut considérer les changements de variable précédents comme projetant, avec conservation de la mesure euclidienne :

- En dimension paire, la boule $\mathcal{B}_{2p}$ sur le produit cartésien de p sphères $\mathcal{S}_2$ (cercles) et de l'intérieur d'un polyèdre à p dimensions.
- En dimension impaire, la boule $\mathcal{B}_{2p+1}$ sur le produit cartésien de p sphères $\mathcal{S}_2$ (cercles) toujours et de l'intérieur d'un segment de paraboloïde à $p+1$ dimensions (défini par $\sum_{k=1}^{p} t_k + x_{2p+1}^2 0 \leq R^2,\ 0 < t_k$ $(k = 1, 2, \ldots p)$).

Dans les deux cas, les volumes du polyèdre ou du paraboloïde sont rationnels, et il y a donc autant de facteurs π que de cercles.

Au fond, la situation est toute simple : il y a autant de facteurs π dans l'expression du volume (et de l'aire) de la sphère à N dimensions que de « circularités indépendantes » dans l'espace, si l'on entend par là le nombre de façons indépendantes de tourner en rond, c'est-à-dire tout simplement le nombre de plans indépendants. C'est évidemment la partie entière de la moitié de la dimension[13].

4.5. L'effet de surface

Il peut être intéressant de montrer comment l'expression du volume de la sphère à N dimensions éclaire la nature physique des espaces de haute dimensionalité. C'est ce qu'on pourrait appeler l'« effet de surface » : plus la dimensionalité est élevée, plus les points d'un corps proches de sa surface externe sont nombreux — en un certain sens que les exemples suivants vont éclairer. Cette assertion vaut, bien entendu, pour la plupart des corps N-dimensionnels suffisamment réguliers, et le cas de la sphère n'a que la vertu de permettre des calculs explicites aboutissant à des résultats simples.

Posons-nous d'abord la question suivante : quelle est le rayon R' de la boule concentrique interne qui comprend la moitié du volume d'une boule de rayon R ? On doit donc avoir $V(R') = 1/2V(R)$, soit $v_N R'^N = 1/2 v_N R^N$, ou

13 – Il est loisible de se demander si la différence de comportement ici mise en évidence entre les espaces de dimensions paires et impaires est liée à la caractéristique d'Euler-Poincaré qui les distingue également.

encore :

$$R' = 2^{-1/N} R. \tag{6.38}$$

On voit que lorsque N augmente indéfiniment, le rayon R' se rapproche du rayon R. Il est plus éloquent encore de considérer la coquille externe, comprise entre la boule interne et la boule externe, et de volume aussi égal à la moitié du volume de la cette dernière. Plus la dimensionalité est élevée, plus l'épaisseur $\Delta R = R - R'$ de cette coquille est faible par rapport au rayon R, comme le montre le tableau 6.2.

Tab. 6.2 – L'effet de surface.

N	1	2	3	4	5	...	$N \gg 1$
$\Delta R/R$	0,5	0,29	0,21	0,16	0,13	...	$\ln 2/N \ll 1$

On voit donc que la moitié externe du volume de la boule est répartie dans une couche de plus en plus mince. La raison de cet état de choses est évidemment la disponibilité d'un nombre élevé — à savoir $(N-1)$ — de directions orthogonales à la direction radiale, qui permettent un « étalement » transversal, autorisant un grand volume sous une faible épaisseur.

Dans le même ordre d'idée, il est intéressant de comparer la surface de la sphère $\mathcal{S}_N$ au volume de sa « section principale », soit la boule $\mathcal{B}_{N-1}$, en calculant le rapport $A_N/V_{N-1} = a_N/v_{N-1}$ (qui ne dépend pas du rayon pour les raisons d'homogénéité mêmes qui rendent cette quantité intéressante). D'après les formules (6.19), il vient :

$$a_N/v_{N-1} = \frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{N-1}{2}\right)}{\Gamma\left(\frac{N}{2}\right)} = B\left(\frac{1}{2}, \frac{N-1}{2}\right) \tag{6.39}$$

soit une fonction eulérienne, dont le tableau 6.3 donne les premières valeurs.

Tab. 6.3 – L'effet de surface (bis).

N	1	2	3	4	5	...	$N \gg 1$
a_N/v_{N-1}	2	π (3,14...)	4	$3\pi/2$ (4,71...)	16/3 (5,33...)	...	$2\sqrt{\pi N}$ ($\gg 1$)

Il s'agit d'une fonction monotonément croissante, qui montre bien que l'aire d'une sphère est proportionnellement d'autant plus grande (par rapport au volume de sa section) que sa dimensionalité est élevée.

Une autre façon de voir les choses est la suivante. Soit une densité de probabilité uniforme dans la boule unité de dimension N. Demandons-nous quelle est la distance moyenne $\bar{r}_N$ entre un point pris au hasard dans la boule et son centre. Sa valeur est donnée, tout naturellement, par :

$$\bar{r}_N = \frac{\int_0^1 dr\, r\, A_N(r)}{\int_0^1 dr\, A_N(r)}, \tag{6.40}$$

soit encore, d'après (6.22) :

$$\bar{r}_N = \frac{a_N \int_0^1 dr\, r^N}{a_N \int_0^1 dr\, r^{N-1}} = 1 - \frac{1}{N}. \tag{6.41}$$

En d'autres termes, la distance au centre moyenne, pour les grandes valeurs de la dimension N, est très proche du rayon de la boule, ce qui veut bien dire que la plupart des points sont près du bord...

Cela signifie que le rayon de la boule donne une assez mauvaise idée de sa taille effective, plus faible qu'on ne pourrait le penser, le nombre des dimensions disponibles, compensant, en quelque sorte, la concentration des points à proximité de la surface.

Une illustration quelque peu surprenante de cette interprétation est la suivante. Considérons qu'une bonne idée de la « taille effective » d'une boule de dimension N, est donnée par la longueur a_N du côté d'un hypercube de volume égal à celui de la boule, que nous prenons de rayon unité. D'après (6.19), cette longueur vaut :

$$a_N := (v_N)^{1/N} = \frac{\Gamma(1/2)}{[\Gamma(1+N/2)]^{1/N}}, \tag{6.42}$$

qui se trouve être une fonction *décroissante* de N (figure 6.9). On note les valeurs particulières évidentes $a_1 = 2$ (en dimension 1, la boule de rayon unité et le cube de côté 2 coïncident) et $a_2 = \sqrt{\pi}$. Pour les hautes dimensionalités, cette longueur tend asymptotiquement vers zéro, avec le comportement

$$a_N \approx \sqrt{2\pi e}\, N^{-1/2} \qquad (N \to \infty). \tag{6.43}$$

Si un hypercube de côté tendant vers zéro peut avoir le même volume que l'hypersphère de rayon unité, c'est qu'il déborde de la sphère par tous ses « coins », puisque les sommets de l'hypercube sont à la distance $d_N = N^{1/2} a_N$ du centre, toujours supérieure à l'unité, et qui tend d'ailleurs vers une constante :

$$d_N \to d_\infty = \sqrt{2\pi e} \qquad (N \to \infty) \tag{6.44}$$

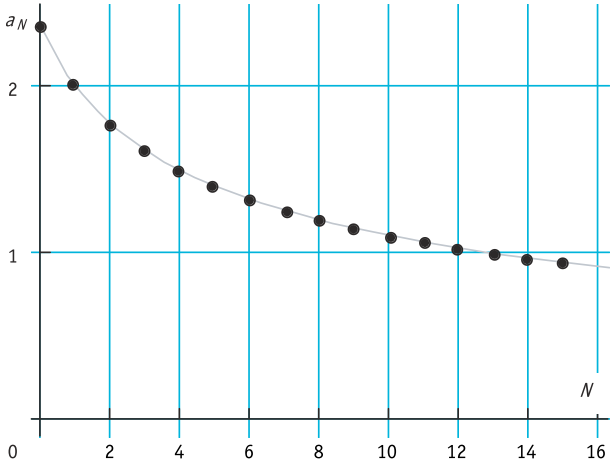


FIG. 6.9 – La taille effective de la boule de rayon unité en dimension N.

(autrement dit, à la limite $N \to \infty$, l'hypercube a un côté nul, mais est inscrit dans une sphère de rayon fini...).

L'expression (6.42) est évidemment indéfinie en dimension nulle. Mais elle admet une limite bien définie pour $N \to 0$, à savoir :

$$a_0 = \sqrt{\pi \; e^{\gamma}} = 2{,}3656\ldots, \qquad (6.45)$$

où γ est la constante d'Euler. Force est donc bien de considérer la curieuse constante (6.45) comme la taille effective de la boule de dimension nulle...

5. Dimensionalité et orthogonalité

Terminons par quelques considérations géométriques élémentaires sur l'espace euclidien $\mathcal{E}_N$ à N dimensions, qui mettent en évidence... l'espace considérable dont on y jouit dès que la dimension devient grande. Pour le dire simplement et naïvement, les vecteurs d'un espace de haute dimensionalité tendent à être « de plus en plus indépendants » (ce qui est évident), et même « de plus en plus orthogonaux » (ce qui l'est moins). De fait, les cosinus directeurs de toute direction (par rapport à un système d'axes orthogonaux) devant avoir l'unité pour somme, si leur nombre N est grand, la situation générique sera celle où chacun de ces cosinus est petit, et l'angle correspondant voisin de l'angle droit. Vérifions-le sur quelques cas particuliers intéressants.

5.1. Équiaxialité

Considérons un système de N axes orthogonaux dans $\mathcal{E}_N$, et appelons « équiaxes » de ce référentiel les 2^{N-1} droites faisant des angles égaux avec chacun des axes ; à 2 dimensions, les équiaxes sont les deux bissectrices de l'angle entre les axes. Notons α_N l'angle (aigu) entre un équiaxe et un axe de référence ; on l'appellera « angle équiaxial » (figure 6.10). Le vecteur unitaire porté par un équiaxe a donc $\pm \cos \alpha_N$ pour projection sur chacun des axes de référence, de sorte que l'angle α_N est défini par la relation $N \cos^2 \alpha_N = 1$, soit

$$\cos \alpha_N = \frac{1}{\sqrt{N}}. \tag{6.46}$$

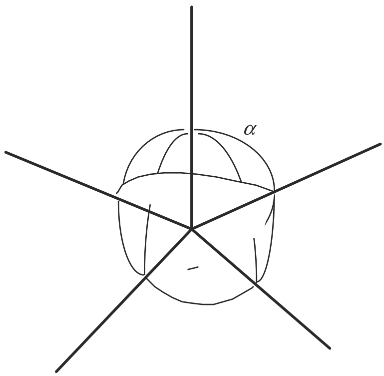


Fig. 6.10 – L'angle équiaxial.

Il est intéressant de calculer les valeurs numériques de l'angle équiaxial pour les basses dimensionalités (tableau 6.4).

Tab. 6.4 – L'angle équiaxial.

N	1	2	3	4	5	...	$N \gg 1$
α_N	0	45°	54,7°	60°	63,4°	...	$\approx \frac{\pi}{2} - \frac{1}{\sqrt{N}}$

On voit que les équiaxes sont de plus en plus proches de l'orthogonalité avec les axes.

5.2. Isogonalité

D'un intérêt comparable est la figure, que nous appellerons « isogonale », formée dans $\mathcal{E}_N$ par $N+1$ droites faisant deux à deux des angles égaux. À deux dimensions, on obtient le trigone à la Mercedes, à trois, les directions des sommets d'un tétraèdre régulier vus depuis son centre ; dans le cas général, ce sont les directions des sommets du simplexe régulier vus depuis son centre. Notons β_N l'« angle isogonal » que forment deux quelconques de ces droites (figure 6.11). Par symétrie, les $N+1$ vecteurs unitaires $\mathbf{u}_k$ ($k = 1, 2, \ldots N, N+1$) portés par les droites isogonales satisfont la relation de dépendance linéaire

$$\sum_{1}^{N+1} \mathbf{u}_k = 0. \tag{6.47}$$

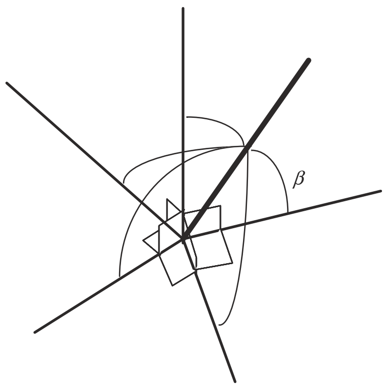


Fig. 6.11 – L'angle isogonal.

En projetant cette relation sur l'un quelconque des vecteurs, on obtient la relation définissant l'angle isogonal, soit

$$\cos \beta_N = -\frac{1}{N}\cdot \tag{6.48}$$

D'où le tableau de valeurs (tableau 6.5).

Tab. 6.5 – L'angle isogonal[14].

N	1	2	3	4	5	...	$N \gg 1$
β_N	180°	120°	109,4°	104,5°	101,5°	...	$\approx \frac{\pi}{2} + \frac{1}{N}$

14 – On note que la relation bien connue $\beta_3 = 2\alpha_3$, qui régit plusieurs aspects des symétries spatiales de notre monde (par exemple en cristallographie), est tout à fait particulière au cas tridimensionnel. L'équation $\mathrm{Arc}\cos(N^{-1}) = 2\,\mathrm{Arc}\cos(N^{-1/2})$, dont la (seule) solution est $N = 3$, s'ajoute ainsi aux autres particularités de *notre* espace, par exemple le fait que $N = N(N-1)/2$(ce qui fait du produit extérieur de deux vecteurs un « produit vectoriel »), ou $N + 1 = 2^{N-1}$ (ce qui permet aux droites équiaxiales d'être aussi isogonales).

Ici encore, on observe la tendance à l'orthogonalité quand croît la dimensionalité.

5.3. Uniformité

Soient la sphère à N dimensions $\mathcal{S}_N$. Quelle est la distance angulaire moyenne entre deux points pris au hasard sur cette sphère, munie de sa mesure uniforme ? Un simple argument de symétrie suffit à affirmer qu'il s'agit d'un angle droit. Mais la répartition probabiliste de cet angle θ, soit $\rho_N(\theta)$, mérite attention. Paramétrons la sphère en coordonnées sphériques généralisées, soit :

$$\begin{aligned} &x_1 = \cos\theta_1, \quad x_2 = \sin\theta_1\cos\theta_2, \quad x_3 = \sin\theta_1\sin\theta_2\cos\theta_3, \ \dots \\ &\dots \ x_{N-1} = \sin\theta_1\sin\theta_2\dots\cos\theta_{N-1}, \quad x_N = \sin\theta_1\sin\theta_2\dots\sin\theta_{N-1}. \end{aligned} \tag{6.49}$$

Avec ces coordonnées, la mesure uniforme sur la sphère $\mathcal{S}_N$ s'écrit :

$$d^{N-1}\sigma = d\theta_1\, d\theta_2 \dots d\theta_{N-1} \ \sin^{N-2}\theta_1 \sin^{N-3}\theta_2 \dots \ \sin\theta_{N-2}. \tag{6.50}$$

Une intégration sur tous les angles donnerait évidemment l'aire de la sphère $\mathcal{S}_N$ calculée ci-dessus. Nous nous intéressons ici à l'angle que fait la direction courante, prise au hasard, avec une direction fixe. Choisissons l'axe zénithal $0x_1$ pour direction de référence ; c'est donc la variable aléatoire θ_1 qui nous intéresse. En intégrant sur toutes les autres variables angulaires, nous voyons ainsi que la densité de probabilité cherchée, qui détermine la répartition aléatoire de l'écart angulaire entre deux directions quelconques, est donnée par

$$\rho_N(\theta) = K\sin^{N-2}\theta. \tag{6.51}$$

C'est une distribution de valeur moyenne $\overline{\theta} = \pi/2$, de plus en plus « piquée » au fur et à mesure que la dimensionalité N croît (figure 6.12) ; pour les grandes valeurs de N, sa largeur est $\delta\theta \approx O(N^{-1/2})$, et elle tend à la limite $N \to \infty$ vers la distribution de Dirac $\delta(\theta - \pi/2)$.

Autrement dit, plus un espace euclidien est de dimension élevée, plus deux directions *quelconques* sont proches de l'orthogonalité.

Cette idée n'est sans doute pas sans effets sur l'intuition que nous pouvons avoir (ou pas, justement !) d'un espace de dimension infinie, tel l'espace hilbertien de la théorie quantique.

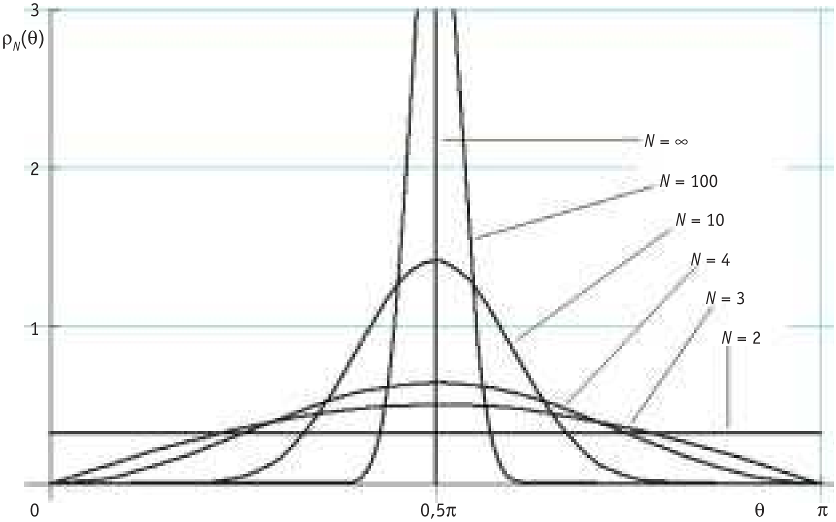


FIG. 6.12 – La densité de probabilité en distance angulaire pour deux points au hasard sur une sphère de dimension *N*.

Bibliographie

[1] Archimède, *Œuvres complètes*, vol. 1, La sphère et le cylindre, Mugler, C. (Ed.), Les Belles Lettres, Paris, 1970.

[2] Coxeter, H.S.M., *Regular polytopes*, MacMillan, 1963.

[3] Coxeter, H.S.M., Donchian, P.S., A N-dimensionnal extension of Pythagoras theorem, *Math. gazette* 19, 206, 1935.

[4] Delahaye, J.-P., *Le fascinant nombre Pi*, Belin, 1999.

[5] Euclide, *Elements*, Livre XII.

[6] Heath, T.L., *The thirteen books of Euclide'elements*, Dover, New York.

[7] Snyder, J.-P., *Flattening the Earth (two thousands years of map projections)*, Chicago University Press, 1993.

07 Espaces physiques : pluralité, filiation, statut

Sylvain Fautrat

> « Fixons notre attention hors de nous-mêmes, autant que possible ; hasardons notre imagination dans les cieux, ou jusqu'aux limites ultimes de l'univers : en réalité, nous n'avançons pas d'un degré au-delà de nous-mêmes et ne pouvons concevoir aucune sorte d'existence hormis les perceptions qui sont apparues dans ces étroites limites. C'est l'univers de l'imagination et nous n'avons d'autres idées que celles qui y sont produites.
>
> *Le plus que nous puissions faire dans la recherche d'une conception des objets extérieurs supposés différer* spécifiquement *de nos perceptions, c'est d'en former une idée relative, sans prétendre y enfermer les objets qui y sont reliés.* »
>
> (David Hume, *Traité de la nature humaine*- Livre I, 1739, [11], p. 124)

1. Introduction

La notion d'espace physique donne une occasion de visiter la problématique de la genèse et de la constitution des diverses formes de connaissance. La présente étude défend plusieurs thèses portant sur le mode de formation des constituants de connaissances qu'on peut appeler, selon les disciplines, concepts scientifiques ou schémas cognitifs, ainsi que sur le mode de fonctionnement de ces connaissances au regard de l'expérience. En résumé, cela correspond à la question de savoir comment une recherche peut se constituer en connaissance.

Les relations entre espace perçu par un individu et espace(s) des théories physiques peuvent être reconsidérées en faisant entrer dans l'étude certains apports des sciences cognitives. Cette question rejoint naturellement la question plus générale du rôle des représentations humaines pré-scientifiques dans le processus de formation des concepts scientifiques. Puisqu'il s'agit ici d'espace, la géométrie et les mathématiques, ainsi que leur statut, ne peuvent pas non plus rester en dehors du problème.

La notion d'espace renvoie également, peut-être de manière plus perturbante que d'autres notions, à la disjonction entre deux manières de penser le monde : pour l'une, le monde est prédéfini avec son espace, y vivent et s'ébattent des sujets percevants qui s'en font la représentation qu'ils peuvent ; pour l'autre approche, il faut prendre en compte que toutes les définitions un tant soit peu précises du monde ne peuvent être que subordonnées, que dérivées à partir des vies subjectives des sujets connaissants, si bien que le monde n'a pas à être conçu comme déterminé d'avance, et cela s'applique au même titre à la notion d'espace.

La recherche d'ontologie a généré diverses positions concernant l'espace, depuis l'espace-réceptacle de corps ou d'événements, jusqu'à l'espace défini secondairement comme système de relations entre corps, entre sensations, ou entre d'autres éléments[1]. Par ailleurs, dans l'usage scientifique, l'expression « espace physique » recouvre un ensemble de concepts insérés dans des théories ; cet ensemble suit l'évolution des dites théories, qui naissent, se multiplient, se perpétuent, sans atteindre une synthèse. Cette dernière constellation va pouvoir servir de point de départ pour une approche plus épistémologique qu'ontologique. Ainsi, il ne s'agit pas de viser l'essence d'un Espace qui serait pour nous un donné, mais de se tourner vers les savoirs en usage, de les prendre comme des données à interroger, et à partir desquels commencer le travail.

La première étape consiste à identifier quelques occurrences courantes du concept qui nous intéresse, sans idée de classer sur une quelconque échelle de valeur les différentes versions de l'espace que proposent telles ou telles théories physiques ou mathématiques. Il va falloir prendre acte de la coexistence de plusieurs conceptions, chercher à comprendre la possibilité de cette coexistence, proposer un type d'articulation entre ces conceptions. Pour ce faire, il faut disposer d'un cadre d'intellection approprié ; c'est pour construire un tel cadre que, dans un deuxième moment, certaines idées générales seront introduites et présentées sous forme d'un *modèle* (ici, un modèle dans le champ de l'épistémologie). Selon ce modèle, chaque situation de connaissance est à décomposer en un *schéma formel* d'une part, et ses *cas d'application* d'autre part. Point essentiel, selon cet axe de pensée, un cas d'application ne peut

1 – Voir par exemple [26], chap. 36 et 37.

être connu qu'au travers d'un schéma formel, *via* un certain processus d'application. De leur côté, les schémas formels sont générés en présence de certains cas d'application déjà rencontrés ou en cours d'introduction. Les deux entités entretiennent ainsi des rapports constitutifs dynamiques ; parallèlement, comme on va le voir sur des exemples, deux couples de ce genre peuvent s'influencer mutuellement. C'est sous cet auspice que se trouvent rapprochées deux questions apparemment distinctes : celle du rapport entre théorie scientifique et expérience et celle du rapport entre schéma cognitif et sensation. Dans les deux cas, on n'a jamais affaire à une découverte à partir de rien, mais à des entreprises de connaissance se constituant selon le même mode.

C'est dans cette perspective que les instances d'espace seront commentées et que la question de la coexistence d'une multiplicité de notions d'espace sera reprise en même temps que celle de leur articulation.

2. Quelques questions

2.1. Quelques instances d'espace

Si on se questionne à partir des mots « espace » et « physique » et qu'on se demande à quels référents ils renvoient, on ne trouve pas des réponses homogènes et encore moins une réponse unique.

Citons quelques exemples (liste non exhaustive) :

(a) espace visuel, espace auditif, etc. ;
(b) espace de la géométrie euclidienne ;
(c) espace de la Mécanique classique ;
(d) espaces-temps des Relativités restreinte et générale ;
(e) espace des états de la Mécanique quantique ;
(f) espaces de dimensions supérieures (supercordes).

Faisons quelques commentaires sur chacun.

Les sensations auditives, visuelles, tactile, musculaire, etc., prennent place dans des espaces perceptifs (a). Ces sensations ne sont pas toutes indépendantes – par exemple, je vois mon doigt toucher un verre et je sens le contact avec le verre – si bien que l'on peut envisager, soit des espaces perceptifs différents mais coordonnés de diverses manières, soit un seul espace perceptif dans lequel toutes les sensations sont censées avoir leur place. Quoi qu'il en soit, on a bien affaire à (au moins) un *espace* ; le qualificatif *physique* se justifie en ce sens que les sensations sont couramment attribuées à des causes appelées objets physiques.

L'espace euclidien (b) est souvent pris comme entité *mathématique*, présente la sphère intellectuelle, du raisonnement, de l'idéalité. Pourtant, les contenus rationnels de l'espace euclidien (concepts, théorèmes) ont leurs pendants pratiques. Ces résultats sont utilisés dans de nombreux corps de métier non scientifiques (géomètres-arpenteurs, maçons, menuisiers, ...) et dans la vie courante.

La Mécanique classique (c), science rationnelle devenue également partie intégrante de l'ingénierie, contient la notion d'espace euclidien comme partie. C'est peut-être le premier exemple d'espace physique auquel de nombreux physiciens penseraient.

Avec la Relativité restreinte (d), l'espace se voit couplé au temps pour former l'espace-temps, lui-même couplé à la matière par la Relativité générale. À l'échelle locale, et même à l'échelle globale en relativité restreinte, *via* un système de traduction (les transformations de Lorentz), pour chaque référentiel l'espace euclidien est conservé.

En mécanique quantique (e), l'évolution des systèmes physiques est conçue dans un espace défini mathématiquement (espace d'Hilbert), dont les propriétés, en particulier le recours aux nombres complexes et le nombre de dimensions, ne peuvent se réduire à celle de l'espace euclidien. L'espace « classique » n'est pourtant pas absent du panorama quantique : il est là pour fournir les valeurs possibles pour les variables de position (résultats de mesure de position).

Avec les théories exploratoires actuelles de type supercordes (f), les entités physiques sont décrites dans un espace ayant plus de trois dimensions spatiales. Les propriétés des dimensions supplémentaires sont aujourd'hui l'objet d'études, mais d'ores et déjà on fixe comme règle que leur extension doit être suffisamment limitée pour ne pas remettre en question la fameuse tridimensionnalité spatiale. C'est un legs que chaque théorie semble recevoir de la théorie précédente sans remise en cause.

2.2. Relation(s) entre ces différents espaces

Finalement, à une certaine échelle spatiale, la structure euclidienne à trois dimensions n'est pas soluble dans la nouveauté ; toutes les contorsions semblent permises dès lors qu'elles se réalisent en dehors de ce noyau dur. De quelle échelle spatiale s'agit-il ? Manifestement, celle à laquelle l'homme ordinaire mène sa vie de tous les jours. Pourquoi cette échelle ? Quelle relation y a-t-il entre cette échelle et les théories physiques qui, pour certaines, ne parlent pas de la vie de tous les jours de l'homme ordinaire ?

Une autre question concerne les mathématiques. Quand on parcourt la liste précédente, les théories semblent de moins en moins s'intéresser au perceptif en même temps qu'elles intègrent une quantité croissante de mathématiques. Ce déplacement progressif est-il le signe d'un changement de nature de la représentation de l'espace ? Est-il lié à une opposition entre percepts et idéalités mathématiques ? Doit-on penser avec H. Poincaré que nos conceptions scientifiques de l'espace sont déterminées par des structures universelles et mathématiques déjà inscrites dans notre esprit ?

2.3. Quel statut d'objectivité pour les différents espaces ?

On connaît le Réalisme physique selon lequelle la réalité s'identifie à ce que décrit la physique, et pour qui tous les écarts ressentis dans la subjectivité par rapport à cette description sont à interpréter comme autant de mirages psychologiques, d'illusions ou d'erreurs du sujet. Selon cette doctrine, le problème de la connaissance de l'espace se ramène à reconnaître la bonne théorie physique, celle qui permet de faire le partage entre le vrai et le faux. Après la diffusion de la mécanique de Newton, on a cru nécessaire d'en expliquer les succès par une objectivité absolue de ses concepts de base, l'espace manié par cette mécanique se devait d'être exactement l'espace du monde lui-même, ou encore être, toujours exactement, l'espace défini dans l'esprit de ses usagers. À l'avènement de la Relativité einsteinienne, la doctrine a dû trouver une stratégie pour se perpétuer : on a demandé d'admettre qu'il y avait eu erreur, que le véritable espace n'était pas celui qu'on croyait, qu'il se présente en réalité sous la forme de l'espace-temps (quatre dimensions, métrique pseudo-riemanienne). Ainsi, à chaque fois, l'avènement d'une nouvelle théorie où l'espace est impliqué s'accompagne d'un discours ontologique et normatif : « Dans le passé, nos prédécesseurs se sont trompés, ils croyaient connaître l'espace ; or, ce n'est qu'aujourd'hui, avec la nouvelle théorie qu'on le connaît enfin vraiment. » La période actuelle qui nous annonce une nouvelle mutation ne fait pas exception. À cet égard, la notion d'espace n'est pas un cas particulier, et l'histoire des sciences est bien sûr faite de revirements répétés affectant des notions et des catégories, fondamentales ou secondaires.

Pour le scientifique au travail, la culture du réalisme possède une vertu certaine : elle est très stimulante. Il y a aussi un prix à payer, car il ne suffit pas à une idée nouvelle de montrer sa valeur empirique et sa cohérence, il lui faut échapper au soupçon d'être une fiction.

Sur le plan épistémologique, il faut tirer un bilan de ces changements de théories et des discours qui les accompagnent, de ces changements qui font l'histoire des sciences. Alors, comment ne pas être un réaliste de l'espace ? Plutôt qu'argumenter en faveur de telle ou telle théorie qui décrirait, ou s'approcherait, d'un espace qui existerait, en tant qu'espace indépendant des représentations d'êtres connaissants, nous pouvons essayer de prendre le problème sous l'angle de l'origine ou de la genèse de la notion, et suivre ses interventions lors de la construction des différentes théories. Même sans pousser très loin ce travail, on doit pouvoir justifier une remise en perspective, un déplacement depuis la question de l'objectivité vers les questions de construction(s) de théories, de leur articulation, de la constitution de la connaissance.

3. Un modèle épistémologique[2]

Dans cette partie sont présentées des idées générales, à partir de l'analyse de cas particuliers choisis dans les champs de la physique et de la psychologie cognitive.

3.1. La question de la relation entre résultats expérimentaux et théorie

Lors de certaines mesures de laboratoire, la physique permet d'obtenir des résultats expérimentaux qui ne s'écartent que de très peu des résultats théoriques. Cet excellent accord est souvent interprété comme le signe que la théorie physique porte en elle une description objective de la réalité, ou pour dire plus précisément, que le monde possède une structure, indépendante de notre connaissance, et que cette structure se retrouve dans la théorie efficiente.

Pourquoi ne pas se contenter de cette conception ? D'une part, parce que l'histoire ne nous a pas livré une seule théorie physique mais un ensemble de théories (mécanique du point, hydrodynamique, optique, électromagnétisme, thermodynamique, mécanique quantique, mécaniques relativistes restreinte et générale, chromodynamique quantique, etc.) qui ne s'articulent entre elles

2 – Les idées de cette section seront développées dans une publication à venir.

que partiellement : les concepts peuvent différer foncièrement d'une théorie à l'autre, et celles-ci ne sont pas efficaces dans les mêmes situations expérimentales. On n'a donc pas affaire à un ensemble théorique homogène porteur d'une structure unique et cohérente.

D'autre part, il ne faut pas rejeter l'idée que les résultats expérimentaux peuvent ne pas être indépendants de la théorie. Comment mettre à jour et penser une telle dépendance ? Pour analyser le type de processus qui permet à une théorie d'entrer en contact avec les situations empiriques, des études de terrain, de nature technique, utilisant les ressources des savoirs établis, de la logique, de l'Histoire des sciences, sont nécessaires ; elles doivent s'accompagner d'un travail de formalisation.

Selon un schéma de pensée assez répandu, le monde empirique fournit à l'expérimentateur des valeurs numériques, la théorie produit de son côté ses propres résultats, et par comparaison des deux types de résultats, on peut évaluer la pertinence de la théorie en regard de la réalité. La thèse de l'objectivité stricte constituée par la description théorique suppose que les résultats expérimentaux sont obtenus sans intervention de la théorie. C'est ce point qu'on va évaluer.

3.2. L'exemple des référentiels inertiels en mécanique classique

3.2.1. Position du problème

L'expression « référentiel inertiel » figure dans le monde théorique, dans les livres de physique. Mais un correspondant dans le monde extra-théorique, un référentiel déterminé par rapport aux corps matériels connus, n'est pas donné d'emblée. Quelles sont les données théoriques qui peuvent servir de support pour la recherche d'un référentiel inertiel empirique ?

En mécanique classique, un référentiel peut, par exemple, être représenté par un trièdre cartésien (trois axes de coordonnées spatiales, rectilignes et infinis, perpendiculaires deux à deux) associé à une chronométrie (axe de coordonnée pour la variable temporelle). De nombreuses grandeurs mécaniques (position, vitesses, accélérations, ...) sont définies par rapport à des référentiels et varient selon le référentiel choisi. Les référentiels inertiels, ou galiléens, sont ceux pour lesquels les grandeurs force, masse, et accélération, relatives à un mobile donné, satisfont la « relation fondamentale de la dynamique »[3] : $\mathbf{f} = m\,\mathbf{a}$. Or, cette relation ne dérive d'aucune autre relation où les termes qu'elle contient trouveraient une définition indépendante, un ancrage propre ; c'est dire que la relation en question apparaît comme une définition implicite à la fois pour la force, la masse et les référentiels inertiels. À ce stade, on ne peut pas envisager un passage à l'application empirique en procédant linéairement, c'est-à-dire, pour un mobile donné, déterminer une des trois entités, puis une deuxième, puis la troisième, puisque dès la première étape on a besoin des résultats des étapes suivantes. On est alors amené à développer la théorie pour obtenir des énoncés moins généraux susceptibles d'être mis en correspondance avec les objets empiriques qu'on connaît par ailleurs. Ainsi, on sait que notre environnement est peuplé de nombreux corps que la mécanique doit concerner, et qu'on n'a pas affaire par exemple au cas simpliste d'un mobile seul dans l'univers, ne subissant aucune force, et pour lequel la théorie a pourtant déjà quelque chose à dire.

3.2.2. Mise en contact avec le monde empirique

À partir de la relation fondamentale de la dynamique, un certain nombre de propositions peuvent être obtenues.

(a) Les différents référentiels inertiels (r.i.) sont uniquement en translations uniformes les uns par rapport aux autres.

(b) Pour un ensemble E de corps n'éprouvant que des forces d'interaction satisfaisant l'égalité de l'action et de la réaction, le barycentre de masse est en mouvement rectiligne uniforme par rapport aux différents r.i. En vertu de la proposition (a), il est alors fixe par rapport à un des r.i.
Ce résultat amorce une piste pour construire un ancrage empirique de la notion de r.i., même si l'absence d'action d'origine extérieure semble pour le moment nous restreindre à un cas plutôt artificiel.

(c) Si on considère maintenant qu'il existe des corps extérieurs à l'ensemble E précédent, et entretenant entre eux, et avec les éléments de E, le même type de force, le mouvement du barycentre de E est modifié ; il est en général accéléré par rapport aux r.i. sous l'action des forces d'origine extérieure. Néanmoins, si on élargit la définition de l'ensemble d'étude pour y inclure les corps supplémentaires, on retrouve le cas (b), si bien que le nouveau barycentre est encore lié à un certain r.i. Dans une situation intermédiaire où l'on inclut une partie seulement des corps extérieurs, on réduit d'autant l'accélération du barycentre, c'est-à-dire

3 – Les lettres en caractères gras désignent des vecteurs.

qu'on s'approche du cas (b), et ce d'autant plus qu'on aura inclus un plus grand nombre de corps dans l'ensemble.

On en tire la leçon suivante en vue de l'application : un r.i. sera d'autant mieux approché par le barycentre[4] d'un ensemble de corps connus, qu'on aura choisi cet ensemble plus vaste, plus englobant par rapport aux corps de l'univers. Considérons les ensembles suivants :

E_1 = la Terre,

E_2 = le Soleil, la Terre et les planètes,

E_3 = le Soleil, la Terre et les planètes, toutes les étoiles de la Galaxie.

Selon le principe précédent, cette suite d'ensembles permet de définir des référentiels[5] dont le caractère inertiel va croissant (une liste de ce genre n'a pas de fin si le nombre de corps interagissant dans l'univers n'est pas fini). La démarche permet donc de faire un classement, selon le critère « d'inertialité », parmi l'infinité des référentiels définissables empiriquement. La notion abstraite de référentiel inertiel a maintenant une prise dans le concret.

La description de la procédure s'est centrée sur la notion de r.i. Il reste que la loi de force, les masses et positions (sur lesquelles est calculé le barycentre), les accélérations (et l'échelle de temps dont elles dépendent), etc., doivent être déterminées conjointement dans la même démarche. Le programme est rempli par étapes, et à chaque étape il faut déterminer, par récurrence, les valeurs empiriques des différents paramètres pointés par la théorie.

Une même grandeur peut souvent être exprimée par deux formules différentes en fonction de grandeurs différentes, et susceptibles de mesures séparées. L'écart entre les deux résultats peut être jugé trop important. Dans ce cas, il est possible d'espérer un meilleur ajustement en reprenant la démarche : pousser plus loin l'élargissement de l'ensemble E, poser une loi de force modifiée, . . . (voire en changeant des éléments essentiels au sein de la théorie, c'est-à-dire passer à une autre théorie).

4 – Il convient de préciser que le barycentre, étant un point, ne suffit pas à définir complètement un référentiel. Un point suffit pour fixer, par exemple, une origine du système de coordonnées spatiales, mais il reste à préciser la direction de ces axes. Le mouvement de rotation correspondant se réduit *selon la même démarche* que celle qu'on est en train de suivre pour le mouvement de translation, c'est-à-dire qu'il faut se donner des énoncés théoriques suffisamment avancés pour être mis en correspondance avec des situations empiriques effectives : pendule de Foucault, gyroscope, etc. Il s'agit ici de dégager un processus actif dans la connaissance, non de propager l'analyse sur toute la physique ou toute la science.

5 – Il s'agit respectivement des référentiels dits : pour E_1, « terrestre » ou « géocentrique », selon la manière de fixer l'orientation des axes (voir note 4) ; « héliocentrique » ou « de Copernic » pour E_2 ; « galactocentrique » pour E_3.

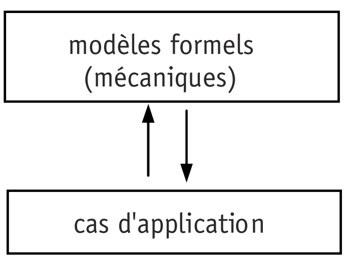
3.2.3. Rôle et statut de la théorie

Au cours de la procédure, la théorie (définie par les formules qu'on trouve dans les livres, ainsi que par la compréhension qu'on en a) joue elle-même le rôle de guide de recherche pour celui qui veut en faire usage. Il n'y a pas la théorie d'un côté et un guide d'un autre côté : il y a la théorie, et c'est un guide. Pour l'observateur, l'expérimentateur, l'ingénieur, la théorie se présente, certes, comme un système de concepts et de formules mathématiques, mais aussi, d'une certaine manière, comme un plan de travail, un mode d'emploi des formules dans les contextes extra-théoriques. La connaissance (la théorie et sa compréhension) organise l'expérience dans un sens qui lui est favorable, elle pilote la construction de situations où elle se peut se révéler efficace. L'intervention de la théorie dans la production des résultats expérimentaux peut donc bien être affirmée.

C'est sur la seule théorie que repose la définition d'un référentiel inertiel, point de passage obligé pour la détermination des valeurs des grandeurs mécaniques (vitesse, quantité de mouvement, etc.) ; ces valeurs permettent ensuite d'exhiber une belle cohérence entre théorie et applications. Mais rien dans le processus décrit ne pousse à imaginer que des forces, des masses, des référentiels inertiels, des quantités de mouvement, préexistent et nous attendent. La pertinence de la théorie résulte de l'application d'une procédure qui ne ressemble pas à l'exploitation d'un isomorphisme préexistant entre une structure théorique et une structure extra-théorique.

3.2.4. Schéma simple

On commence à apercevoir un mode d'organisation particulier existant entre les éléments d'une théorie et les situations qu'ils permettent d'appréhender. Les relations entre théorie et cas d'application peuvent être visualisées ainsi :

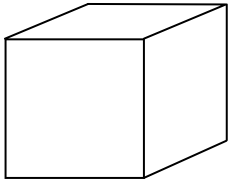


Dans ce schéma, le mot « modèle » peut être compris comme « ensemble des notions sollicitées pour aborder une situation ». Pour l'exemple de la mécanique, on peut penser à l'ensemble des concepts et formules de cette science. Dans d'autres cas, les modèles formels pourront se présenter sous d'autres figures : une liste d'instructions, un processus de traitement, etc. Dans la suite on utilisera aussi bien « modèle » que « processus ».

L'expression « cas d'application » ne renvoit pas à une entité bien définie. On veut parler des situations auxquelles les modèles vont pouvoir s'appliquer, bien que ces situations ne soient pas connues ni définies d'avance. C'est l'effectivité du processus d'application à partir des modèles, qui fait émerger le domaine de pertinence : une théorie est efficace dans les situations du type... de celles où elle est efficace ; et c'est pour ces cas qu'on peut parler après coup de cas d'application.

Les modèles doivent être pensés ici comme étant de nature formelle, au sens où une même forme (un même modèle, un même processus) peut s'appliquer à des cas multiples et variés, et qui peuvent être vus comme assurant un correspondant matériel au modèle.

3.3. L'exemple de la vision d'un cube



Considéré d'une certaine manière, ce dessin ne porte que neuf segments concourants.

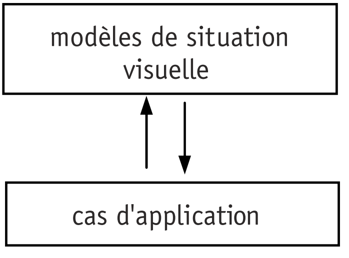
Pour y voir un cube, objet géométrique dans un espace à trois dimensions, il faut quelque chose en plus. Que faut-il ? Répondre précisément à cette question engage certainement assez loin. Au minimum, il faut que s'applique, d'une manière ou d'une autre, un certain modèle formel de cube.

L'exemple du cube n'a ici rien de particulier (le fait, par exemple, qu'il soit susceptible d'une définition mathématique n'intervient pas) ; il est pris comme exemplification du fait que ce qui est effectivement vu ne peut pas se déduire de la seule connaissance des images rétiniennes actuelles, ou, si l'on veut, de la distribution d'énergie lumineuse sur le fond de l'œil[6].

Que dire sur ce modèle de cube ? Il ne s'agit pas ici d'un modèle décrit dans un livre, comme pouvaient l'être les modèles de nature scientifique de la Mécanique. Il ne peut s'agir non plus d'un modèle de cube déposé en mémoire, vu sous un angle particulier avec des arêtes ayant des inclinaisons et des proportions particulières. Il ne peut s'agir non plus d'une équation mathématique. Alors de quoi peut-il s'agir ? Ce n'est pas le lieu d'engager des hypothèses sur la nature de ce modèle formel ; chaque postulat sur cette nature (par exemple, existence d'un prototype utilisé pour des comparaisons, procédure programmée de traitement de l'information, émergence à partir de processus neuronaux *a priori* non finalisés pour produire ce modèle précis, ...) peut produire une voie de recherche et c'est le terrain des sciences cognitives. Ce qui importe dans notre perspective, c'est que, d'une certaine procédure (inconsciente), résulte une présentation non plane de cube[7] ; c'est dans ce sens non engagé qu'il faut entendre ici le mot de modèle.

En définitive, c'est le modèle convoqué qui définit ce qu'il y a précisément à voir.

On retrouve l'organisation du schéma précédent.



6 – En principe, toute perception peut être l'occasion d'une réflexion sur cette différence qui existe entre ce qu'on peut appeler (trop) rapidement données sensorielles et ce qui est finalement perçu, mais c'est avec des exemples simples préparés à cette fin qu'on se place plus facilement en position de réflexion. Pour d'autres exemples visuels, on peut se reporter par exemple à [16].

7 – Tel ou tel modèle est-il « spontané » ? Une personne sans éducation géométrique, dont l'attention n'a jamais été dirigée vers cette figure répertoriée, aura certainement peu de chance de voir un cube plutôt que, par exemple, neuf traits positionnés de manière particulière. Il faut admettre que, pour l'individu, le modèle de cube, ou la procédure de traitement de l'information qu'il représente, résulte d'une histoire. Il faut envisager que dans cette histoire, qui va d'un départ où il n'a jamais été question de cube jusqu'au moment où une vision claire et nette d'un cube (et apparemment spontanée) est réalisée, s'inscrivent en fait des phases de création et de modification, d'ajustements.

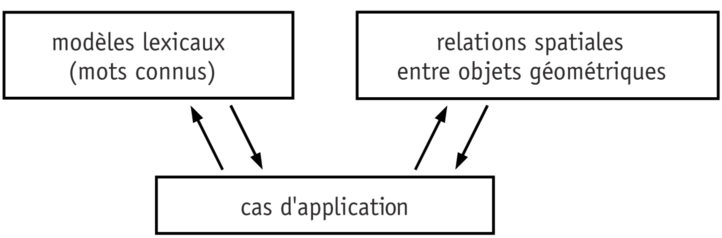
3.4. Influence entre deux types de modèles formels

Plus haut, il a été fait allusion à « des corps qu'on connaît par ailleurs ». On peut se demander ce qu'est cette connaissance et si elle entre dans le schéma précédent. On va réfléchir sur un exemple simple.

Le plus souvent, en regardant la figure ci-dessus ([25], p. 8), on reconnaît deux mots de la langue anglaise : the cat. Pourtant, si on regarde plus longuement le 2e et l'avant-dernier caractères, on constate qu'ils sont identiques *sur le plan géométrique*. De ce point de vue, ils ne peuvent être à la fois H et A. Pourtant, dans l'optique de lire des mots, ces ressemblances géométriques n'entrent pas en ligne de compte.

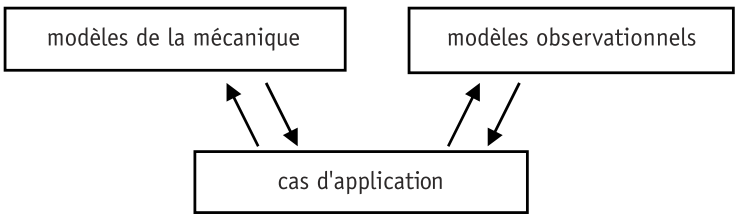
Ainsi, un regard vers une même feuille de papier peut participer à deux processus de connaissance différents. (1) Quand on cherche à faire la lecture d'un texte, nos références sont les mots figurant dans notre lexique personnel, ceux qui peuvent être reconnus dans une situation de lecture. C'est ainsi que cette opération se constitue en situation linguistique effective de déchiffrage d'entités lexicales (lettres, mots, ...). (2) Quand on ausculte, avec une visée de type géométrique, les zones noires sur le papier, on utilise d'autres références : comparaison de distances entre extrémités de bâtons verticaux, horizontaux, obliques, etc.

L'un des deux processus peut influencer l'autre ; par exemple, on peut être conduit à réviser la lecture initiale de certains mots à la suite d'une investigation de nature géométrique. Deux entreprises de connaissance peuvent ainsi être couplées lorsqu'elles partagent quelque chose sur le plan de l'application (dire purement et simplement qu'elles partagent le même cas d'application serait contradictoire). En résumé :



Revenons aux objets étudiés par la mécanique, et plus généralement par les sciences expérimentales ou observationnelles. Dans la section 3.2., les objets (Terre, Soleil, Planètes, etc.) auxquels les modèles formels de la mécanique sont à appliquer, étaient désignés par l’expression « les objets empiriques qu’on connaît par ailleurs ». On peut maintenant revenir sur ce « par ailleurs ». Sans recours immédiat à la mécanique, ces objets sont connus par usage des yeux, des télescopes, etc. Est-ce à dire, sans recours à des modèles ? non, il s’agit simplement d’autres modèles, ceux intervenant lors de la perception visuelle, lors de l’usage des instruments (modèles scientifiques permettant d’en interpréter les données, de corriger ces dernières pour tenir compte d’effets perturbateurs prévus dans la théorie de ces instruments, etc.). Ainsi, la prise de connaissance en question entre bien dans le cadre général Modèles formels/applications.

Finalement, mécanique et savoirs observationnels sont couplés selon le schéma :



4. Retour sur les questions précédentes

4.1. Perspectives de travail ouvertes par le modèle précédent

S’inspirer de ce qui précède consiste à identifier des contenus formels en position de négociation dynamique avec leur cas d’application, deux types différents de contenus pouvant également s’influencer par le biais de leurs applications.

4.2. Perceptions spatialisées

4.2.1. Travaux de Mach et Poincaré

On sait que, à la suite d'autres savants du XIXe siècle, E. Mach et H. Poincaré ont traité la question des relations entre physiologie, psychologie, mathématiques et espace euclidien de la Physique.

La physiologie et les sciences cognitives se sont depuis développées, les théories physiques traitant d'espace se sont multipliées, mais des problèmes, comme celui des relations entre ces différentes sphères, n'ont pas été pour autant effacés par la science.

Rappelons les idées travaillées par ces chercheurs.

Mach[8] insiste sur les différences patentes entre l'espace euclidien qu'enseigne la géométrie d'une part, et l'espace, ou les espaces, issus de nos sens.

L'espace géométrique est abstrait, illimité, isotrope, homogène, métrique, les corps y ont des tailles fixes.

Par contraste, l'espace « physiologique » procure un sentiment de substantialité de l'espace. Il est limité en ouverture (au sens où, par exemple en bougeant les yeux, on découvre ce qui était hors des limites du champ visuel initial). Il n'est pas homogène et est sensible aux trois directions, verticale, latérale, et antéro-postérieure définies par notre corps (par exemple, les sensations tactiles liées à notre main gauche sont qualitativement différentes de celles liées à notre main droite ; placé au centre d'un hémisphère, nous le voyons aplati selon la direction verticale, etc.). Il n'est pas métrique (les objets n'ont pas de taille déterminée, ils grossissent ou rétrécissent lorsqu'ils s'approchent ou s'éloignent de nos yeux ; au-delà de quelques dizaines de mètres, nous ne pouvons plus déterminer avec précision la distance qui nous sépare d'eux ; à courte distance, les évaluations précises de longueurs à l'œil ou au toucher ne viennent qu'après la pratique des mesurages de type physique, c'est-à-dire après éducation au contact d'instruments).

Les 3 dimensions de l'espace viennent, non de la vue seule, mais des sensations musculaires, en particulier celles liées au déplacement des membres. Les sensations forment « un système, que nous apprenons peu à peu à faire correspondre à l'espace géométrique » ([12], p. 330). Les sensations d'espace « se fondent par association » pour des raisons d'efficacité qui tiennent à la conservation de l'espèce biologique ([13], p. 167).

8 – Mach [13],[12], chap. XX et XXII ; voir également les articles de M. Paty [17, 18].

Reprenant ce tableau à deux types d'espace, dont l'existence conjointe est présentée sous forme problématique, Poincaré[9] prolonge l'approche de Mach en introduisant des mathématiques, qu'il voit impliquées dans la mise en correspondance elle-même, celle dont parle Mach. Partant d'un empirisme à la Hume, dont il adopte le vocabulaire (impressions, associations des idées, ...), il recourt ensuite au postulat selon lequel les mathématiques sont implantées dans l'esprit.

Parmi les sensations (visuelles, musculaires, du toucher), certaines régularités, associations, lois, se constatent. Par exemple, selon Poincaré, lors de l'observation d'un objet, il y a toujours cohérence constatée entre sensation d'accommodation visuelle et sensation de convergence des deux yeux ; ainsi l'espace visuel n'a-t-il pas seulement les deux dimensions de la rétine, mais trois, la troisième venant des variations conjointes des sensations musculaires d'accommodation et de convergence. De ce point de vue, et selon un critère basé sur les aspects physiologiques, l'espace visuel lui-même n'est pas homogène puisqu'il mêle sensations optiques et musculaires (Mach ([13], p. 166 et 167), lui, nous demande de les reconnaître homogènes, car participant d'un tout unique : le comportement biologique global). Tous les autres muscles apportent par ailleurs leurs lots de sensations. Finalement, l'espace « représentatif » (à quelque chose près « l'espace physiologique » de Mach) peut fournir, sous forme de relations entre séries de sensations, des lois empiriques. Ainsi, certaines séries de sensations, reconnues comme internes à l'individu, car « volontaires », peuvent compenser, c'est-à-dire ramener à leur état initial, certaines autres sensations : un objet en mouvement donne une sensation visuelle se modifiant dans le temps, mais un déplacement des yeux ou du corps (sensations internes) permet de récupérer la vision initiale (en suivant l'objet des yeux, par exemple).

Un espace digne de la géométrie peut-il sortir de cette sphère de la sensibilité ? Pour Poincaré, c'est dans l'entendement, instance distincte, que préexistent les mathématiques, les groupes, la géométrie[10]. Un espace géométrique (euclidien ou non), ou, aussi bien, le groupe des transformations associé, est alors susceptible d'être sollicité pour organiser « commodément » les

9 – Voir [21], chap. II à V ; [22], chap. IV et V. Voir également [19], livre II, chap. I ; [20].

10 – « Ce qui est l'objet de la géométrie, c'est l'étude d'un « groupe » particulier ; mais le concept général de groupe préexiste dans notre esprit au moins en puissance . Il s'impose à nous, non comme forme de notre sensibilité, mais comme forme de notre entendement. » [21], p. 93. « Dans notre esprit préexistait l'idée latente d'un certain nombre de groupes ; ce sont ceux dont Lie a fait la théorie. », p. 107. « Je crois donc que si par espace on entend un continu mathématique à trois dimensions, fût-il d'ailleurs amorphe, c'est l'esprit qui le construit, mais il ne le construit pas avec rien, il lui faut des matériaux et des modèles. Ces matériaux et ces modèles préexistent en lui. » [22], p. 98.

sensations. Ce qui n'était pas possible (espace infini, homogène, isotrope, ...) dans la sphère de la représentation, l'est dans l'entendement. L'hypothèse redoutée selon laquelle la géométrie, et la physique utilisant cette géométrie, dépendrait de l'individu considéré, est écartée car l'entendement est supposé être le même pour tous les humains (c'est « l'Entendement », sans précision, non de celui de Pierre ou de Paul).

4.2.2. Commentaires

Où s'arrête la perception ? Dans une conception simpliste de la perception, où des données sensorielles objectives sont transmises plus ou moins directement à une conscience, il n'y a pas de place pour l'intervention de quelconques modèles. En fait, depuis le début, les travaux de recherche en psychophysique exhibent des différences notables entre ce qui est perçu et ce que pourrait laisser supposer l'observation des seules données physiques présentées au sujet percevant sur lequel on expérimente ; on est nécessairement conduit à une conception bien moins simple de la perception. Au minimum, il faut reconnaître que la perception n'est pas directe, qu'elle possède ses modalités propres.

Prenons quelques exemples liés au contexte de cet article.

Exemple 7.1. Sur le dessin de la section précédente, il est bien possible de voir un cube, en tant que cube, et pas seulement neuf segments noirs. Pour considérer cette vision, on ne peut se contenter d'un ensemble de points de différentes couleurs ou luminosités ; dans cette vision, le cube est présent. Plus précisément encore, pendant la perception, le caractère carré des faces ne fait pas problème alors même que certaines de ces faces se présentent sur le dessin, et sur l'image rétinienne, comme des losanges.

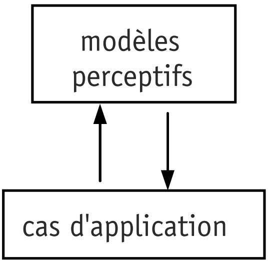
Exemple 7.2. Lorsque la nuit, nous regardons les étoiles, nous pouvons, soit les voir comme disposées sur une sphère (la sphère des fixes des anciens), soit les voir sans distances déterminées par rapport à l'observateur. Cela dépend fortement des discours que nous aurons entendus et qui auront structuré nos conceptions astronomiques.

Exemple 7.3. Lorsqu'on demande à un sujet d'ajuster la taille d'un disque situé à distance fixe jusqu'à ce qu'il le voie de la même taille que tel autre disque-cible situé à une certaine distance, l'expérience révèle que l'égalité des tailles, estimée dans la perception, ne coïncide pas avec une égalité des tailles des images rétiniennes. Les écarts sont importants ([28], p. 162 et 163) et, de plus, dépendent de la consigne, c'est-à-dire de la manière de présenter verbalement au sujet ce qu'on lui demande de réaliser ([24], p. 131 et 132). Les expériences de ce genre poussent à reconnaître que nous percevons visuellement les objets avec une idée de leur taille propre, ou pour dire autrement, qu'une certaine taille « intrinsèque » est présente au sein de la perception. Pour un objet donné, des estimations différentes de sa taille peuvent être obtenues selon les conditions ; par exemple, si, dans la perception, une estimation de l'éloignement de l'objet cible est possible (voir à ce sujet un peu plus loin dans cette section), l'objet peut être vu avec une taille pratiquement indépendante de cet éloignement (il y a « constance de taille »), comme s'il y avait à chaque fois prise en compte, et donc suppression, de l'effet attendu de diminution de taille apparente avec l'augmentation de distance. Selon une approche physicienne, la diminution de taille découle de constructions d'optique géométrique et devrait être patente ; la perception suit d'autres logiques.

Exemple 7.4. N.R. Hanson ([9], chap. 1) considère J. Kepler et Tycho Brahé, pris en tant que figures paradigmatiques de l'histoire de l'astronomie. Ils sont convaincus, l'un que le Soleil est fixe et que la Terre se meut autour, l'autre que c'est la Terre qui est fixe. Hanson envisage que ces deux observateurs ne devaient pas voir la même chose lorsqu'ils tournaient chacun leur regard vers notre Soleil, leur perception incluant sous une certaine forme leur conviction.

On arrive à la question : jusqu'où faut-il faire entrer des éléments théoriques dans la perception ? Cachée dans cette question se loge l'idée que la perception possède dans l'absolu une définition bien circonscrite, qui va de soi, et qui ne doit pas être transgressée. Or, l'observation ou l'expérimentation dans ce domaine comportent toujours deux aspects différents qu'on doit examiner. On distingue : (a) un objet ou un phénomène (un dessin, un son, une scène, etc.) donnés à percevoir à un sujet, et définis physiquement (longueur, distance, fréquence, intensité, etc.) par référence à des savoirs de type scientifique et objectif ; (b) une autre entité, disons ce qui est perçu par le sujet, que l'on cherche à décrire, à classer, à relier à l'entité (a) par des lois. Or, si, dans un premier temps, (a) peut être considéré comme bien défini par construction, il n'en va pas de même de (b) puisqu'il est l'objet de la recherche. Si un sujet entend le carillon de Big Ben à chaque fois qu'il est placé en face d'une surface jaune à pois verts, aussi inattendu cela soit-il, il faudra bien en prendre acte et avoir une définition suffisamment ouverte de la perception pour y intégrer ce résultat ; dans le même ordre d'idée, il faut avoir une conception de la perception suffisamment ouverte pour laisser entrer au sein des percepts certains aspects qu'on serait classiquement tenté de catégoriser autrement, par exemple en les classant comme intellectuels, théoriques, etc. Par ailleurs, l'entité (a) elle-même est à considérer de plus près ; on doit se rappeler qu'elle résulte d'un certain mode de production. Par exemple, la notion de longueur (d'un trait sur un dessin) résulte à la fois de l'histoire de la géométrie et de la physique, faite de créations conceptuelles (discutée plus loin dans ce texte pour ce qui concerne la géométrie), et de processus d'application, régulés par le haut (*i.e.* par les modèles formels de ces sciences). Le sentiment d'objectivité vient de l'efficacité, indiscutable, de ce dispositif, non de ce que la longueur serait un donné pur. C'est ainsi se joue, au sein de notre culture, un combat entre valeurs : valeurs portées par la science (« la longueur mesurée du trait est objective ») contre valeurs portées par le sujet (« ce que je vois est ce que je vois »).

On va maintenant s'intéresser plus particulièrement aux situations de perception visuelle et garder présent à l'esprit le schéma suivant.



On peut récapituler quelques indices à partir desquels peut être construite la sensation d'espace à plus de deux dimensions, c'est-à-dire d'espace avec profondeur et relief[11]. Dans la naissance de la sensation de profondeur, contrairement à ce qu'avançait Poincaré, mais en accord avec Mach ([13], p. 201), les recherches actuelles ne conduisent pas à attribuer un rôle direct aux sensations mêmes d'accommodation et de convergence[12]. D'autres

11 – Berkeley notait bien, en 1709, dans l'*Essai d'une théorie nouvelle de la vision*, que la distance entre l'observateur et l'objet observé n'est pas donnée par la vision ; elle doit être dérivée.

processus de traitement des entrées perceptives peuvent exploiter, d'ailleurs indépendamment les uns des autres, d'une part la différence entre les deux images rétiniennes à un instant donné (vision stéréoscopique), d'autre part différents indices déjà présents dans l'image rétinienne d'un seul œil (vision monoculaire), ou encore la différence entre les images se succédant dans le temps. L'exploitation des légers décalages entre les images en vision stéréoscopique ([2], p. 34) est clairement établie par la fabrication de stéréogrammes plans ([16], chap. 8), permettant pourtant de voir des objets situés à différentes profondeurs. Par ailleurs, lorsqu'il y a mouvement relatif entre les objets observés et les yeux (soit que la scène contient des objets mobiles, soit que l'observateur bouge la tête), la différence de vitesse de déplacement des deux images rétiniennes des différents objets permet, *via* l'application d'une procédure de traitement, de créer la sensation de plus ou moins grand éloignement (chaque objet étant d'autant moins éloigné que la différence des deux vitesses est plus grande). Pour ce qui concerne les images rétiniennes issues d'un seul œil, il y a déjà plusieurs possibilités. La différence de texture ou de luminosité des différentes parties de la surface d'un corps, de ses différentes facettes, peut être reliée à la différence d'exposition à la lumière, c'est-à-dire aussi à la différence d'orientation des facettes dans l'espace, d'où la possibilité de construction, à partir de l'état de surface, d'une perception en relief du corps avec ses creux et bosses. (Ce lien entre luminosité et relief était déjà étudié méthodiquement par Mach ([13], p. 182 et 185).) Comme indice utilisable on peut aussi citer l'agrandissement au cours du temps de l'image rétinienne d'un objet, ou son rétrécissement, qui peuvent être traités dans la perception comme un rapprochement ou un éloignement de l'objet par rapport à l'observateur. Pour les scènes statiques, peuvent également être exploités des effets de perspective ; par exemple, des lignes qui sont parallèles sur l'objet ne le sont pas sur l'image rétinienne (les rails de la voie ferrée se rencontrent au loin) et cette différence peut être exprimée en termes de profondeur ; de même, un objet A dont l'image est partiellement cachée par celle d'un objet B, est à placer derrière B. Il va de soi que le processus, pour ces comparaisons entre objet connu et image rétinienne, n'est possible que si le sujet percevant dispose d'une certaine référence interne concernant l'objet en question, que s'il dispose d'un modèle de l'objet.

12 – Voir [2], p. 71. F. Varela chap. VII montre aussi comment l'accommodation peut entrer dans l'explication d'un aspect de la perception visuelle (la constance de taille). L'explication repose sur la considération du fonctionnement global du système nerveux ; elle est sans rapport avec l'hypothèse qu'on trouve chez Poincaré selon laquelle une corrélation entre accommodation et convergence serait constatée et utilisée.

Concernant l'appréciation du mouvement au sein de la perception, on peut faire une remarque. Lorsque nous regardons un paysage statique (un mur), tout en bougeant la tête, ou encore lorsque nous regardons le même spectacle filmé par une caméra en mouvement panoramique, nos images rétiniennes sont foncièrement changeantes dans le temps ; pourtant, le plus souvent le paysage est perçu comme fixe. Ainsi, ce qui, dans la perception, est fixe ou mobile n'est pas déterminé de manière univoque par la fixité ou la mobilité dans l'image rétinienne. Lorsque nous sommes dans un train qui démarre sans accélération perceptible, nous pouvons d'ailleurs basculer d'une perception à l'autre : dans la première nous voyons le paysage bouger dans l'encadrement fixe de la fenêtre, dans l'autre nous voyons l'intérieur du train et son contenu se déplacer dans un paysage perçu comme fixe. Pour revenir à l'exemple de Hanson, on peut ainsi comprendre que Tycho Brahé et Kepler puissent ne pas voir la même chose en regardant le Soleil : le matin, par exemple, l'un peut voir le Soleil s'élever sur l'horizon terrestre, par référence à son modèle l'autre peut voir l'horizon s'abaisser sous le Soleil ([9], p. 182 note 6).

À chaque fois, les processus qui viennent d'être évoqués font appel à la reconnaissance et au suivi d'objet. On comprend que, si un objet peut être identifié, puissent être également traités son suivi dans le temps, son occultation partielle par un autre objet, sa variation de taille, etc. Pour reconnaître un objet unique en rotation derrière les modifications qu'il subit dans la vision, plutôt qu'en rester à la succession des vues, certes connectées par des relations de proximités de forme mais ne définissant pas un objet pour cela, il faut disposer d'un modèle assurant non seulement la reconnaissance mais aussi la continuité dans le temps. Il est donc justifié d'interroger la notion même d'objet visuel et les modalités de sa reconnaissance ; actuellement différentes voies de recherche sont explorées ([28], chap. VI ; [24], chap. 2C).

Toujours concernant les caractéristiques des modèles qu'on peut postuler pour la perception visuelle, il convient de noter que les détecteurs de lumière, les cellules cônes et bâtonnets dont les densités varient sur la surface de la rétine et qui sont pour partie d'entre elles, déjà connectées entre voisines, n'en sont pas moins des entités discrètes. Or, dans la perception, rien ne correspond aux séparations spatiales entre les détecteurs : le champ visuel ne présente pas la même structure discrète, les objets vus ne sont pas systématiquement lacunaires ou granuleux. Sur le même thème : les cellules rétiniennes sont absentes dans la région de la « tache aveugle », là où le nerf optique se rassemble en un faisceau pour partir vers le cerveau. Ici encore, sauf à se placer dans des conditions particulières ([14], p. 5–7), la perception fournit un champ visuel où cette lacune n’est pas perçue. En termes d’application de modèles, il faut donc admettre quelque chose comme un processus de lissage topologique.

L’intervention de processus perceptifs[13] est également évidente lorsqu’on se rappelle que, tout simplement, deux yeux fournissent deux images rétiniennes (légèrement) différentes, alors que la vision n’est pas, sauf cas particuliers, dédoublée.

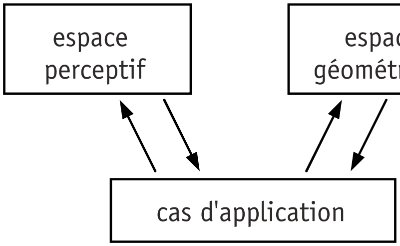
Poincaré en appelle aux régularités, aux lois empiriques, qui sont constatées dans la sensibilité, etc. Il en est de même de toutes les approches empiristes qui veulent se fonder sur de telles relations, censées être données directement lors de l’observation. Par contraste, l’analyse de la perception en termes de modèles, introduit, avant même que puissent être notées des régularités ou des lois concernant des objets, une couche de procédures préalables du type lissage, reconnaissance d’objet, suivi d’objet, etc. Dans ces conditions, l’image et le statut de lois qui traduiraient des réalités objectives sont fortement modifiés ; les lois sont subordonnées à des modèles, correspondant probablement à des possibilités parmi d’autres, de générer la perception. Ce caractère contingent est patent dans les cas où, pour une même situation, nous oscillons entre deux perceptions différentes (comme lors du démarrage du train évoqué plus haut, ou lorsqu’on regarde une figure « ambiguë », « réversible », dans laquelle on peut voir aussi bien un objet qu’un autre selon l’approche qu’on a du dessin).

L’effet du type constance de taille (exemple 3) est à rapprocher d’un autre l’effet, celui de constance de forme. Cette expression correspond au fait que lorsque nous regardons un objet de forme connue, nous sommes souvent peu sensibles au changement de forme effectif de son image sur notre rétine : l’assiette posée sur la table est vue circulaire bien que, ne nous trouvant pas au-dessus d’elle, la projection en donne une image rétinienne fortement ovalisée ([24], p. 128 et 130). Nous n’avons pas d’une part une vision consciente où figurerait explicitement une assiette ovale, et à côté une opération fabriquant la forme ronde ; pour une large gamme de conditions expérimentales, nous voyons une assiette qui, pour nous, est ronde. La perception travaille

13 – Sous l’angle de l’étude du système nerveux, F. Varela formule les choses ainsi [27], p. 166 : « La perception est un processus de compensation qu’effectue le système nerveux au cours d’une interaction. Un espace perceptif est une classe de processus compensatoires qu’un organisme peut subir. La perception et les espaces perceptifs ne reflètent pas les caractéristiques de l’environnement, mais *l’invariance de l’organisation anatomique et fonctionnelle du système nerveux au cours de ses interactions*. » Un peu plus loin : « L’organisme n’extrait pas la distance des caractéristiques de l’environnement ; au contraire, par un processus de compensation des perturbations, il *engendre* les distances perçues, comme un mode de comportement compatible avec l’environnement. »

avec, associée à l'objet, une espèce de forme absolue (il est rond ; il possède des faces perpendiculaires ; etc.). De même, un objet connu peut être associé dans la perception à une taille sans qu'une réflexion intellectuelle digne de ce nom n'intervienne. Ces propriétés perceptives ou cognitives préfigurent les caractéristiques des corps indéformables de la géométrie métrique. On est loin de l'opposition frontale que Mach et Poincaré concevaient entre espace géométrique à concepts métriques et espace perceptif non métrique[14].

Avec l'habitude, un charpentier par exemple peut *voir* les poutres en bois comme ayant chacune telle ou telle section (parmi les valeurs en usage dans le métier), et ceci indépendamment de la distance à l'objet, qu'elle soit de trois ou de dix mètres. Des compétences métriques de ce genre peuvent être développées dans de nombreux secteurs d'activité. Naturellement, comme le notait Mach, une pratique soutenue est nécessaire, et cette pratique doit faire appel de manière répétée à des références métriques précises, telles des échantillons de tailles connues ou des appareils de mesure. Du point de vue des modèles formels de connaissance, un tel apprentissage correspond à un enrichissement des modèles perceptifs par introduction de contenus issus de la géométrie et de la physique.



Ce n'est d'ailleurs pas seulement dans les corps de métiers que la géométrie diffuse ; mêlée à des savoirs issus de la physique, on la rencontre en de nombreuses occasions de la vie courante aussi banales que la lecture d'un plan de rues ou des panneaux de signalisation routière, etc. (Aujourd'hui, il est même difficile d'imaginer abstraire la géométrie de la vie moderne, tant

14 – Le cheminement de Poincaré va d'une situation problématique (espace représentatif et espace géométrique sont séparés et incompatibles) qu'il décrit en détails, puis rencontre des notions mathématiques (la théorie des groupes) qui constituent pour lui un mur derrière lequel il n'y a plus rien à voir, un point où l'on peut cesser la quête. Or, on peut souligner que, d'une part, Poincaré ne cherche pas à se donner l'espace « représentatif » le plus réaliste possible (contrairement à Mach) ; d'autre part la géométrie à laquelle il pense est déjà réécrite sous une forme qui la subordonne aux groupes de transformations ; on a donc affaire à un cadre préparé pour asseoir les arguments. (Précisons que relever ici un certain mathématisme chez Poincaré – attitude par ailleurs assez répandue – ne signifie pas déconsidérer le travail en question ; au contraire, ses efforts pour conceptualiser et formuler des arguments ont produit des textes parmi les plus stimulants.)

notre habitat, les objets qu'on rencontre, nos déplacements, sont structurés pour partie par la géométrie.)

Ainsi, si des modèles perceptifs ont pu avoir une influence sur la formation des concepts géométriques, en sens inverse, ces derniers peuvent modifier, et modifient effectivement les contenus perceptifs[15].

4.2.3. Conclusion

Certains traits, que l'on pourrait penser propres aux concepts géométriques et à la réflexion intellectuelle, sont déjà partiellement présents dans la perception. Dans la conception d'une origine de la géométrie, il faut sans doute ouvrir plus la porte à la perception et moins au seul exercice du raisonnement.

De plus, le complexe « géométrie-physique », lorsqu'il a déjà pris corps, peut à son tour intervenir dans la perception. Pour réemployer le mot de Mach, les sensations ne sont pas données d'emblée en tant que système, il faut des modèles pour les structurer. Pour un individu éduqué à la géométrie et à la physique, des modèles issus de ces sciences participent également à cette structuration.

4.3. L'espace vu à travers la science

Plaçons-nous au niveau de la géométrie élémentaire et considérons l'hypothèse selon laquelle dans la formation de l'idée de droite géométrique s'est trouvé impliqué un aller-retour entre une notion de droite et une expérience minimale en optique. On sait qu'à l'époque d'Euclide, la géométrie a pu être présentée comme fondée sur l'usage « de la règle et du compas »[16]. Pourquoi ne pas s'appuyer plutôt sur la notion de droite elle-même, la droite idéale, plutôt que sur un objet matériel comme la règle ? C'est que la droite idéale ne semble pas aller de soi. Dans l'histoire des sciences, on ne trouve pas trace de l'introduction, reconnue commme évidente, de la notion de droite idéale infinie qui sera enseignée sans discussion bien des siècles plus tard. On trouve plutôt une idée d'abord assez floue qui se constitue et se précise par étapes. Par exemple, les tentatives pour en donner une définition (c'est-à-dire ici un

15 – Les descriptions qu'on trouve dans la littérature sont en général linéaires et n'envisagent pas d'effets de rétroaction. C'est bien le cas pour Poincaré : le mouvement est en sens unique, de la sensibilité vers l'entendement.

16 – Remarquons que cette expression se rencontre à une époque où la géométrie est déjà assez avancée en tant que système déductif ; elle ne dit rien sur les stades de développement très antérieurs.

énoncé ouvrant sur une signification) sont multiples et aucune n'est vraiment concluante. Dans la période grecque, on trouve par exemple [5] :

(1) est droit ce dont le centre fait écran aux deux extrémités ;
(2) la droite seule couvre une distance égale à celle qu'il y a entre ses points ;
(3) de toutes les lignes de mêmes extrémités, la plus courte est la droite ;
(4) la droite est une courbe tendue à l'extrême.

Dans la première définition, l'optique est déjà sous-entendue derrière le mot écran.

La deuxième et la troisième font appel à la notion de distance, qui demande à être explicitée et qui porte ses propres difficultés. (Il s'agissait alors de ne pas être à nouveau confronté aux incommensurables ; il faut également éviter d'introduire dans la définition de la longueur la référence à une unité qui, peu ou prou, serait déjà supposée rectiligne. En toute rigueur, il reste aussi le problème de l'existence et de l'unicité du chemin minimal.)

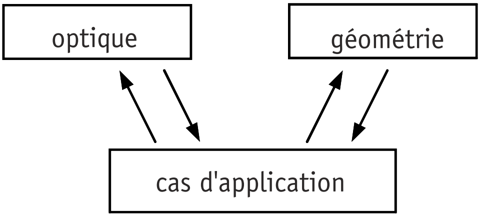
La quatrième définition renvoie directement à un vécu empirique de l'élasticité.

Le statut sommaire de la droite (par rapport aux conceptions postérieures) se lit aussi dans la non-affirmation de son caractère infini ; tout au plus c'est une des « demandes » d'Euclide de pouvoir prolonger toute droite actuellement tracée entre deux points.

Finalement, la définition qui apparaît chez Euclide (une ligne droite est celle qui est placée de manière égale par rapport aux points qui sont sur elle) n'est pas plus éclairante. Pourtant, les démonstrations ne peuvent pas être construites à l'aveuglette, et les significations, pour qu'elles puissent jouer à la fois leurs rôles de guide et de vigile, doivent être complétées. D'où le recours à des objets matériels, tels la règle et le compas, propres à porter un peu plus loin le sens des mots « droite » ou « cercle » ; par l'intermédiaire des constructions réglées qu'ils permettent, l'activité intellectuelle voit son champ d'action se resserrer, se préciser.

Rentrons dans le détail au sujet d'une telle règle, en bois si l'on veut. Puisque c'est sur elle que repose l'idée de droite, ou plutôt à ce stade, l'idée de rectitude, il est essentiel, pour l'esprit et pour la production des figures, qu'elle se distingue d'un bâton quelconque dont la rectitude ne serait pas assurée. Usuellement, on peut contrôler sa bonne rectitude en visant, à l'œil, le long de l'arête servant à tracer les traits. Ce faisant, on utilise les rayons lumineux ; d'une certaine manière, on fait entrer la lumière dans la fabrication de l'instrument. À cette occasion, la situation empirique qu'est la visée optique informe la règle, la met en forme (au sens géométrique, mais aussi au sens figuré en ce qui concerne la notion naissante de droite ; ici les deux sens fusionnent). Se souvenant que c'est typiquement cette règle matérielle qui sert à engendrer les figures géométriques offertes comme point de départ à l'activité intellectuelle, aux raisonnements, à la production de nouveaux contenus formels, on peut dire qu'en pensant aux théorèmes portant sur les triangles et polygones abstraits, on pense autant à des théorèmes portant sur des assemblages de rayons lumineux concourants.

Et, si les phénomènes optiques ont participé à la formation de la géométrie en déposant leur empreinte dans les concepts initiaux, en retour la géométrie, constituée en corps de doctrine organisé a fourni les matériaux conceptuels pour étudier et formuler les lois de l'optique géométrique.



Insistons : s'il y a collaboration entre ces deux disciplines, celle-ci ne doit pas être conçue comme un processus où des énoncés tout faits, issus d'une des deux disciplines, sont repris à titre de conventions par l'autre. Il faut se représenter une solidarité conceptuelle résultant d'une période pendant laquelle les modèles (modèles formels de situation) de l'une et de l'autre disciplines se forment en s'influençant.

L'histoire de la collaboration entre droite géométrique et optique ne s'arrête d'ailleurs pas à cette époque. Au XIX^e siècle, pour asseoir les conceptions également naissantes des géométries non-euclidiennes, puis la généralisation des diverses géométries, des géomètres insistent sur une référence au monde extra-mathématique. Deux citations [1] :

> C.F. Gauss (lettre à Olbers, 1817) : « Ainsi, la géométrie ne peut être mise à côté de qui est de nature *a priori*, mais plutôt à côté de la mécanique. »
>
> N.I. Lobatchevski (*Nouveaux principes de la géométrie*, 1838) : « L'inanité des efforts tentés depuis l'époque d'Euclide, dans le cours de deux mille ans, m'a conduit à soupçonner que la vérité à établir n'était pas impliquée dans les notions antérieures et que pour la démontrer, il

fallait, ainsi que cela a eu lieu pour d'autres lois de la nature, recourir à des expériences, par exemple aux observations astronomiques. »

Chez B. Riemann, en arrière plan de sa réflexion *Sur les hypothèses qui servent de fondement à la géométrie* (1854), figure l'appel répété à l'expérience, pensé comme repère solide pour le mathématicien créateur : « Il nous reste maintenant à examiner comment, à quel degré et avec quelle extension ces hypothèses sont confirmées par l'expérience. » Plus loin : « La réponse à ces questions ne peut s'obtenir qu'en partant de la conception des phénomènes, vérifiée jusqu'ici par l'expérience, et que Newton a pris pour base, et en apportant à cette conception les modifications successives, exigées par les faits qu'elle ne peut pas expliquer. »

Or, au début du XX^e^ siècle, la Relativité générale a trouvé dans la nouvelle droite (la géodésique de la géométrie généralisée, géométrie riemannienne) les contenus formels pour décrire la trajectoire de la lumière. C'est dans ces conditions que cette théorie physique, à laquelle la géométrie donne forme également, a pu prétendre traiter valablement d'optique à l'échelle astronomique et cosmologique. En première analyse, on pourrait dire que les vues de ces géomètres du XIX^e^ siècle, selon qui la valeur de la géométrie est liée aux réalités empiriques, étaient pénétrantes et justifiées. On peut aussi voir les choses autrement : ils ont eux-mêmes participé à la genèse des contenus formels qui structurèrent cette Relativité générale, et l'anticipation est pour partie auto-réalisatrice.

En considérant les anciennes définitions de la droite, on a trouvé qu'à côté de l'optique d'autres contenus empiriques sous-tendaient les idéalités de la géométrie élémentaire : fil tendu, corps matériels rigides, etc. Pour chacun, on peut établir des rapports de même nature, c'est-à-dire des rapports constitutifs et dynamiques entre les concepts de la géométrie et ceux des sciences empiriques correspondantes. Pour construire aussi bien la Cinématique des corps indéformables, l'Élasticité des fils et des corps solides, la Mécanique des milieux continus, il a fallu des modes d'expression formelle pour les déplacements et les déformations qu'on envisage pour les objets matériels. Et, pour ce faire, c'est précisément dans la géométrie qu'ont été puisées les idéalités (tels le trièdre indéformable, les écarts spatiaux par rapport à un tel trièdre, etc.) qui jouent le rôle de références pour exprimer avec précision la fixité, le déplacement, les déformations, les lois de comportement élastique des matériaux, etc.

4.4. Sur l'existence d'une bipolarité épistémologique

Souvent, dans les argumentations épistémologiques sont opposées deux sphères : l'une où l'on rapproche les idées de concret, de perception (conçue comme immédiate), de matériel, de monde réel, d'empirique, etc., et l'autre sphère avec l'abstrait, les idéalités, les mathématiques, le formel, le rationnel[17].

Or, selon l'étude précédente, une idée abstraite ne naît pas en tant qu'idée abstraite ; elle le devient progressivement en prenant corps au sein d'un modèle formel de connaissance. Par ailleurs, des données « empiriques » ne sont identifiées que *via* la mise en forme par des modèles formels (et en premier lieu, par les cadres formels de la perception). On n'a donc pas affaire à deux pôles qui gardent leur distance, mais plutôt à deux facteurs de l'activité cognitive.

17 – À diverses occasions, ces deux tendances sont énoncées explicitement. Pour la géométrie, certains auteurs militent du côté de conceptions empiristes (Clairaut, *Eléments de géométrie,* 1753 ; Camus, *Eléments de géométrie,* 1764 ; on trouvera dans, entre autres, [8] des commentaires sur les approches de plusieurs de ces géomètres : Clairaut, Bertrand, Legendre, Gauss, Bolyai, Lobatchevski). Chez d'autres auteurs, c'est au contraire l'autonomie du rationnel qui est prévaut.
Pendant de longues périodes chacun des deux pôles peuvent rester dans sa position retranchée, sans véritable affrontement. Dans certains cas, comme l'arrivée de la Relativité, les deux positions se rapprochent dangereusement. Effectivement, le cadre formel de cette théorie peut être revendiqué par le clan rationnel : espace et temps associés dans une variété (pseudo-)riemannienne de dimension 4, de métrique à signature inhomogène, courbure locale dépendant de paramètres de répartition de matière ; mais elle se présente aussi avec une validité éprouvée empiriquement (par observation d'une étoile lors d'une éclipse de soleil, du mouvement du périhélie de Mercure). Il y a nécessité d'un arbitrage : un discours va se faire jour, qui vante les vertus clarificatrices d'une distinction tranchée entre géométrie « pure » (abstraite, mathématique) d'une part, et géométrie « physique » d'autre part. (Voir : Einstein [4] ; [15], [23] par exemple paragraphe 3 ; [3], chap. 18). Selon ce schéma, il n'y a pas une géométrie mais deux. Il revient à la première géométrie de générer, selon un processus qui reste dans l'ombre, des structures formelles sans connnexion avec l'autre sphère. La deuxième géométrie est en fait l'affaire des physiciens dont la tâche revient à mettre en correspondance l'une ou l'autre des structures avec des faits empiriques. Cette mise en correspondance peut bien être creusée épistémologiquement, mais on ne prévoit pas l'existence de relations génétiques entre formation de la structure géométrique utile et mise en forme des faits géométrico-physiques. Ce discours s'articulait avec la recherche, à la même époque, d'une formalisation radicale des mathématiques. Du côté de la géométrie ainsi formalisée, droite, plan, etc., ne reçoivent plus de définitions visant la signification ; on cherche à les remplacer par des caractères écrits et des règles d'écriture. On passe ainsi d'une « axiomatique matérielle » d'Euclide à une « axiomatique formelle » (voir par exemple [6]). La présentation de la géométrie euclidienne qu'a publiée D. Hilbert en 1899 [10] est l'une des étapes significatives de cette transition. Peut-on espérer récupérer le contact avec les situations empiriques à la fin de la construction formelle ? Gonseth, à l'occasion de l'étude de la reconstruction ensembliste, met bien en avant les limites inhérentes à la construction purement formelle : poser seulement un formalisme ne pourra jamais générer de significations extérieures ni constituer de définition implicite de quoi que ce soit d'extérieur ou d'antérieur [7]. Il faut noter que l'abandon des significations extérieures par les mathématiques se comprend aussi comme un déplacement des frontières, une redistribution des tâches, entre secteurs scientifiques et cognitifs. Ainsi, les physiciens récupèrent les problèmes d'application et de signification ; à eux de dire, dans les cas d'application externes au formalisme (c'est-à-dire ceux où interviennent d'autres éléments que des symboles écrits), à quoi correspondent : une droite (qui, dans cette culture mathématique, n'est plus ce qui est tracé par une règle), une géodésique, les composantes d'un tenseur métrique, etc.

4.5. Le raccordement

4.5.1. Origine du problème de raccordement entre les deux types d'espace

D'un côté, l'espace décrit par la science peut être présenté comme construit *en partant* de l'espace perceptif ; l'espace perceptif est alors premier logiquement par rapport à l'espace de la science. Mais, on pense aussi à l'espace perceptif comme au faisceau d'une lampe torche qui explorerait un espace global *préexistant*, commun à tous les individus connaissant ; dans cette seconde description, l'espace de la science est premier chronologiquement.

Les antériorités logique et chronologique ne coïncidant pas, comment concilier les deux approches ?

Pour clarifier, il faut mettre à jour de nouvelles différences entre ces deux types d'entités.

C'est un fait que la manière de penser l'espace commun change selon qu'on se réfère à telle théorie plutôt qu'à telle autre ; il suit en quelque sorte l'histoire de la science. Ce n'est donc pas une entité déterminée ; or, lorsqu'on dit qu'il préexiste, on entend une entité non contingente. Finalement, on ne fait guère plus qu'*imaginer* un espace global, et d'une manière qui dépend de la théorie en vigueur au moment considéré.

Seul l'espace perçu par l'individu est un *véritable* espace, c'est-à-dire possédant les caractères substantiels d'espace, et pas seulement avec des définitions formelles énoncées dans des livres. Le sens spatial transite par le vécu physiologique de l'individu. L'espace des théories physiques, lui, n'est pas en premier lieu ressenti pas un individu ; il n'y a personne pour *voir* l'espace décrit par la science, pas d'entité percevante pour qui l'espace de la science est un espace de nature perceptive.

Ainsi, le mot espace ne recouvre pas la même idée dans les deux expressions.

4.5.2. Schéma de pensée proposé

L'espace perceptif dépend de contenus de modèles formels propres à un individu. L'espace décrit par la science correspond à des contenus d'autres modèles formels, propres à aucun sujet en particulier, sinon à des entités moins nettes comme la communauté humaine, la communauté scientifique, etc.

Les cas d'application ne sont pas les mêmes pour les deux classes de modèles formels. D'un côté, il s'agit de cas appartenant à la vie intime d'un individu ; les modèles liés à des sons, des couleurs, un éblouissement visuel, ..., des propositions linguistiques, etc. De l'autre côté, il s'agit de cas appartenant à la pratique scientifique, impliquant des objets communs, ou qu'on travaille à rendre communs : planètes, cyclotrons, trains, appareils de mesures, etc. (ces objets sont alors considérés dans le rôle dans lequel les place telle théorie ; par exemple, l'avion en tant que cas d'application de la Mécanique du vol, et non en tant que moyen de transport). Les cas d'application pour ces deux types d'espace ne sont pas donc les mêmes, du fait même qu'ils sont engendrés par des systèmes de modèles formels situés dans des plans différents.

Dans les expressions « espace géométrique » ou « espace décrit par la science », comment le mot espace se justifie-t-il ? Par un acte intellectuel, nous projettons sur la notion théorique, rationnalisée d'espace, le sens spatial que nous avons pu rencontrer dans le vécu physiologique. Il y a donc emprunt de la qualité spatiale dans le domaine perceptif au profit du domaine géométrique et physique. Au cours de l'histoire de la science, le transfert s'est fait de proche en proche, au fur et à mesure de la construction de la science : d'abord la géométrie, l'optique, ..., la physique classique, puis la relativité, la mécanique quantique, ..., jusqu'aux théories d'essai actuelles. Il y a en quelque sorte traçabilité du sens spatial. C'est d'ailleurs parce qu'ils partagent de cettte manière le même sens spatial que des modèles scientifiques et des modèles individuels peuvent s'influencer.

Nos différents mondes subjectifs étant distincts et séparés, on pourrait s'étonner que les modèles que les différents individus se constituent pour eux-mêmes lors de l'apprentissage d'une science présentent le plus souvent un bon niveau de compatibilité (les physiciens se comprennent relativement bien). Il faut se rappeler que les modèles formels scientifiques sont constamment régulés au cours de l'activité sociale. On comprend que, dans une large mesure, les mathématiques ou la physique puissent passer pour indépendantes de l'individu considéré, sans que cela fasse problème.

4.6. L'objectivité des divers espaces

Chacun des espaces de la liste précédente correspond à des modèles formels spécifiques. Chacun entretient des relations, du type couplage, avec ses cas d'application, relations dont qu'il tire son efficacité et sa légitimité (sa validité). Or, ce type de relation ne peut suffire à fonder rigoureusement la conception d'une entité extérieure douée d'une structure déterminée, et qui coïnciderait avec celle définie par les contenus de nos modèles formels. Pour cette raison, aucun des espaces ne peut être considéré comme reflet fidèle d'une partie déterminée du monde, c'est-à-dire qu'aucun ne peut être dit objectif en ce sens.

On pourrait peut-être encore envisager que la série des différentes théories portant sur l'espace, prises dans leur ordre historique de formation, converge progressivement vers une éventuelle réalité spatiale. Selon ce schéma, une théorie apparaissant à une certaine période devrait englober dans ses cas d'application l'ensemble des cas d'application des théories antérieures. Or, nos théories traitent plutôt de situations spécifiques, avec leurs concepts respectifs, et ne se recouvrent que partiellement. (On connaît même la thèse épistémologique de l'incommensurabilité selon laquelle les différentes théories ne peuvent que s'ignorer mutuellement.) L'établissement d'une correspondance est en général réalisé au prix d'une redéfinition des concepts qui doivent se correspondre, et d'un redécoupage du domaine d'application des théories antérieures. Par exemple, pour donner à la mécanique relativiste la possibilité d'être présentée comme englobant la mécanique classique comme cas limite aux basses vitesses, il faut soutenir que la mécanique classique ne concerne pas les vitesses arbitrairement élevées, bien que cette limitation soit en fait d'origine externe et ne se lise nulle part dans son formalisme ; il faut aussi oublier que les effets gravitationnels s'obtiennent dans les deux théories par l'usage de concepts foncièrement différents (force active dans un cas, courbure de l'espace-temps dans l'autre), que dans une même situation étudiée, les référentiels qui sont déclarés inertiels ou non-inertiels ne sont pas les mêmes dans les deux théories, etc. Finalement, on modifie l'objet-théorie afin qu'il présente les ressemblances voulues. En fait, chaque théorie présente un certain domaine de pertinence, que la théorie suivante, sans le remettre en cause, peut éventuellement aider à préciser.

5. Conclusion

5.1. On peut accepter une démultiplication dès lors qu'on dispose d'un modèle intégrateur garantissant l'intelligibilité

Dans la conception de la connaissance comme découverte et description d'un monde préstructuré, la coexistence de plusieurs théories traitant d'espace ne peut au mieux être ressenti que comme une situation transitoire dont on espère l'issue ; au pire, c'est un échec dans la recherche de connaissance vraie.

Si, au contraire, la connaissance est prise comme processus producteur de modèles se validant par couplage avec des cas d'application, la coexistence de théories, représentant autant d'angles d'attaque différents, ne pose pas problème.

La question n'est pas tant de savoir quelle théorie (ici de l'espace) est objective, que de se donner une bonne théorie de l'objectivité.

5.2. La cognition perceptive a donné une direction initiale

Les activités rationnelles qui ont produit les théories scientifiques se sont couplées à la perception avant de se coupler entre elles (tout en étant couplées, bien sûr, à leurs cas d'application). Par ce processus, en fournissant le sens spatial, la perception (la cognition) a orienté vers un type, ou certains types, de théories physiques susceptibles d'intégrer la notion d'espace.

Se faire une idée la plus informée possible sur l'espace réclame d'embrasser l'ensemble de ses instances, qui va de la perception jusqu'aux théories en gestation, en passant par les théories scientifiques bien établies ; ensemble elles constituent une histoire où des filiations sont signifiantes. Sous cet angle, la science peut être regardée comme l'étude des possibilités de prolongement, au delà du domaine perceptif, de la notion d'espace (à côté d'autres notions, bien sûr).

Qu'une mutation vienne nous priver de la qualité spatiale au sein de nos représentations, que de nouveaux modèles perceptifs nous proposent une vision du monde où l'idée d'étendue serait absente (en ne recourant qu'à des entités discrètes, par exemple) et c'est une bonne partie de la science qu'il faut réinterpréter voire recommmencer. Quelle physique de l'espace développerait, ou développe, une espèce vivante n'ayant pas nos yeux d'humains, nos bras, nos jambes ? Si connaître, c'est entrer dans un processus à couplage (au sens de la section 3), leurs « espaces », pour autant qu'il faille employer ce mot, peuvent être à la fois bien différents des nôtres, et d'égale valeur épistémologique.

Bibliographie

[1] Bkouche, R., *Euclide, Klein, Hilbert, et les autres*, *in* La rigueur et le calcul, Cedic, 1982.

[2] Bonnet, C., Ghiglione, R., Richard J.-F., *Traité de psychologie cognitive*, 1, Dunod-Bordas, 1982.

[3] Carnap, R., *The philosophical foundations of physics : An introduction to the philosophy of sciences*, 1966, édition Dover, 1955. Trad. française, Les fondements philosophiques de la physique, Armand Colin, 1973.

[4] Einstein, A., *La géométrie et l'expérience*, 1921, *in* Réflexions sur l'électrodynamique, l'éther, la géométrie et la relativité, Gauthiers-Villars, 1972.

[5] Euclide, *Les éléments*, vol. 1, Introduction de M. Caveing, livres I-V, avec commentaires de B. Vitrac, PUF, 1990.

[6] Eves, H., *Foundations and fundamental concepts of mathematics*, PWS-KENT publishing company, 1990.

[7] Gonseth, F., *Les mathématiques et la réalité – Essai sur la méthode axiomatique*, section 98, 1936, Librairie Sci. et Tech. Albert Blanchard, 1974.

[8] Gonseth, F., *La géométie et le problème de l'espace*, livre II, section 52, éditions du Griffon, 1946, livre IV, sections 189-193, éditions du Griffon, 1955.

[9] Hanson, N.R., *Patterns of discovery*, Cambridge University Press, 1958.

[10] Hilbert, D., *Les fondements de la géométrie*, Trad. française, Jacques Gabbay, 1997.

[11] Hume, D., *Traité de la nature humaine*, livre 1, 1739, Trad. française, Flammarion, 1995.

[12] Mach, E., *La connaissance de l'erreur*, 1905, Trad. française, Flammarion, 1908.

[13] Mach, E., *L'analyse des sensations*, Trad. française d'après la 6^{e} éd. 1911, éditions Jacqueline Chambon, 1996.

[14] Maturana, H., Varela, F., *L'arbre de la connaissance*, Addison-Wesley France, 1994.

[15] Nicod, J., *La géométrie dans le monde sensible*, 1923, introduction B. Russell, éd. française, PUF, 1962.

[16] Ninio, J., *L'empreinte des sens*, Odile Jacob, 1996.

[17] Paty, M., *Mach*, *in* Encyclopedia universalis.

[18] Paty, M., *Espace : la critique de Mach*, *in* Dictionnaire d'histoire et de philosophie des sciences, PUF, 2000.

[19] Poincaré, H., *Science et méthode*, 1908, Flammarion 1947.

[20] Poincaré, H., *Dernières pensées, pourquoi l'espace a trois dimensions*, Flammarion, 1963

[21] Poincaré, H., *La science et l'hypothèse*, 1902, Flammarion, 1968.

[22] Poincaré, H., *La valeur de la science*, 1905, Flammarion, 1970.

[23] Reichenbach, H., *The philosophy of space and time*, 1927, Dover, 1958.

[24] Roulin, J.-L. Coord., *Psychologie cognitive*, Bréal, 1998.

[25] Rumelhart, D.E., MCClelland, J.L., the PDP Research group, *Paralell distributed processing*, vol. 1, The MIT Press, 1986.

[26] Russell, B., *L'analyse de la matière*, 1927, Payot, 1965.

[27] Varela, F., *Autonomie et connaissance – Essai sur le vivant*, Seuil, 1989.

[28] Weil-Barais, A. (sous la Dir. de), *L'homme cognitif*, PUF, 1993.

08

Les théories spatiales de Poincaré à l'épreuve de l'Histoire classique

Christiane Vilain

Dans son premier livre de philosophie des sciences : *La Science et l'Hypothèse*, écrit en 1902, Henri Poincaré montre que l'espace géométrique, en tant qu'étendue vide, infinie, homogène et isotrope, est bien différent de l'espace perceptif dont nous tirons nos premières représentations. La raison en est, dit-il, que cet espace vide dans lequel nous sommes habitués à placer les objets qui nous entourent n'est pas engendré par nos sensations considérées comme passives. Il est en revanche engendré de façon active par nos déplacements, qui nous permettent de mettre en rapports nos sensations entre elles. Selon Poincaré, ce sont ces déplacements, et non pas nos sensations ou perceptions, qui sont homogènes et isotropes.

Or c'est à l'époque classique, celle de Galilée, Descartes, Huygens, Leibniz et Newton, entre autres, que l'espace géométrique de la tradition euclidienne devient espace physique propre à la représentation des phénomènes naturels. Ces phénomènes étaient auparavant représentés sur la scène diversifiée et limitée des corps de l'univers aristotélicien, univers totalement plein, hétérogène et centré sur la Terre. La scène vide sur laquelle on déplace sans les déformer les figures géométriques des pythagoriciens et des milésiens est replacée dans le domaine de la pure pure imagination, pour Aristote et ses successeurs. Il s'agit en tout cas d'entités qui ne se trouvent pas dans la nature. Il en est alors de même de l'espace vide et infini dans lequel tombent les atomes de Leucippe et Démocrite, si bien décrits par Lucrèce. Ce sont, pour Aristote, de pures fictions qui ne rendent pas compte de ce que nous observons.

Il est alors bien tentant de caractériser l'époque de la physique classique comme celle qui refonde la connaissance sur les déplacements de l'individu, sur une prise en compte de la façon dont les observations sont modifiées lorsque l'individu se déplace, et qui décentre ainsi l'ensemble de la construction aristotélicienne pour l'inscrire dans un espace vide indéfini. Une telle interprétation évoque la conception plus générale d'une science nouvelle dans laquelle l'expérience courante cède la place à l'expérimentation, c'est-à-dire à un processus d'observation actif, fabriqué par l'homme. L'accent mis sur les déplacements en est un aspect plus intéressant que celui de l'expérimentation proprement dite en ce qu'il s'inscrit plus aisément dans l'ensemble des activités humaines. Il n'est pas nécessaire de connaître à fond l'histoire de cette période et de la « Renaissance » qui l'a précédée, pour savoir à quel point les gens se déplacent, au cours de voyages terrestres et maritimes. On peut franchir l'océan ou simplement observer ce qui se passe sur un navire en mouvement régulier, sur une rivière par exemple. Le déplacement peut avoir lieu effectivement ou par la pensée, chaque expérience étant aisément extrapolée, jusqu'à dépasser les limites de l'univers visible. Il est même possible d'élargir cette notion de déplacement à toute activité permettant de faire varier les paramètres d'une observation, à l'aide d'un dispositif inspiré des anciennes machines simples, comme le levier ou le plan incliné. Cette intervention de l'individu, du savant devenu mobile et entreprenant – voire un peu agité ! – prend la place de la contemplation passive de « ce qui se passe ». Ce qui se passe est aussi ce qui passe, mais l'activité du savant lui permet de reproduire des expériences de façon quasiment identique. Il se déplace et peut revenir au même endroit. Son activité lui permet donc en quelque sorte d'abolir le temps : il a ainsi conquis un espace qui, comme celui du géomètre, implique une permanence et la reproduction illimitée des mêmes expériences.

L'intervention nouvelle de l'activité humaine dans l'étude de la nature ne peut être mise en doute, ni l'importance de la mobilité de l'observateur. Mais qu'il en résulte immédiatement un espace de type géométrique comme espace physique est plus difficile à prouver de façon générale à la lumière de l'histoire. L'ensemble des représentations des phénomènes montre en effet au XVII^{e} siècle une diversité décourageante : Descartes n'admet pas d'espace vide. Huygens et Leibniz en font, chacun à leur manière une donnée toute relative, tandis que Newton transforme cet espace vide en repère absolu possédant une puissance quasi divine, ou tout au moins une puissance dynamique. Les célèbres débats entre Descartes et Henri More, puis, plus tard entre Leibniz et Clarke, montrent à quel point le sujet est d'actualité[1].

1 – Nous en avions parlé plus longuement dans un article antérieur sur le même sujet [8], p. 149-167. Nous n'y reviendrons donc pas non plus ici en détail.

Mais nous voulons laisser de côté ici ces débats explicites trop connus par ailleurs pour tenter de regarder de près comment l'appréhension des phénomènes les inscrit, ou non, dans un espace. Or la notion même de « phénomène » intéressant pour l'étude de la nature n'a rien d'absolu ni de définitif. L'époque classique est « mécaniste » et privilégie les chocs, à la suite de Descartes. Huygens est le premier à établir les lois générales des chocs élastiques par une méthode « du navire en mouvement » qui le conduira à une notion d'espace physique vide et indifférencié, totalement relatif, identique à l'espace géométrique[2]. Cela suffirait à confirmer sans discussions les thèses de Poincaré. Mais ce travail de Huygens est trop spécifique, et trop mal compris par ses contemporains pour représenter la physique du XVII^e^ siècle. Revenons donc aux phénomènes en général. On n'observera plus le changement en général puisque le choc en constitue le prototype, le modèle, mais il y a tout de même un phénomène qui garde, à la suite des études galiléennes, une certaine autonomie par rapport au mécanisme : c'est la chute des corps. Mersenne et Huygens continuent à la traiter comme un phénomène continu, indépendant de la physique des chocs. C'est donc à travers l'étude de la chute des corps que nous allons maintenant mettre à l'épreuve nos interrogations sur l'espace.

1. La « délocalisation » de la chute des corps et sa « géométrisation »

Au Moyen Âge, on remarque que l'augmentation de vitesse du corps au cours de sa chute est sans doute plutôt fonction de la distance au point de départ que de la distance au but, qui est le centre de la Terre. À une image globale des corps en mouvement vers un centre ou vers la périphérie, selon des couches sphérique, on va substituer insensiblement une image locale de corps lâché à partir du repos, pour constater que le phénomène ne dépend pas visiblement de l'endroit d'où le corps tombe ni de l'instant de début de la chute. Ces deux représentations sont qualitatives et aucune mesure n'intervient, ni ne peut intervenir pour remplacer la première par la seconde. C'est bien plutôt un glissement d'intérêt dû au fait que l'on regarde tomber des

2 – Nous avions également traité en détail ce sujet dans l'article [8] et n'y reviendrons pas ici. Huygens utilise l'argument galiléen du navire en mouvement uniforme, dans lequel on effectue un choc dont on connaît le résultat, pour observer depuis la berge le résultat d'un choc correspondant à d'autres conditions initiales. Il en déduit l'ensemble des règles, et les lois de conservations que nous connaissons aujourd'hui. Il a donc fait de l'invariance galiléenne une méthode qui permet de trouver des lois, comme étant les seules possibles, et non plus seulement la possibilité du mouvement de la Terre.

corps que l’on lâche ou jette soi-même plutôt que de considérer l’ensemble des phénomènes naturels indépendamment de nous.

On observe également des jets de flèches, dont le mouvement persistant conduit Buridan à la physique de l’*impetus*. Tout corps lancé persiste dans son mouvement de lui-même et l’action du milieu ambiant tend plutôt à le freiner qu’à l’accompagner dans son mouvement. De nombreuses expériences viennent à l’appui de cette affirmation, dont certaines sont purement imaginaires déjà. Ce qui importe est qu’il ne s’agit plus seulement d’observer les phénomènes mais de les susciter pour tester une explication. La conclusion de Buridan lui permet ensuite de considérer la chute libre et de déduire une augmentation régulière de vitesse à partir du point de départ : si le corps garde naturellement la vitesse déjà acquise et que la gravité continue à agir, une nouvelle vitesse s’ajoute continuellement et celle-ci ne peut qu’augmenter.

Qu’y a-t-il de plus alors chez Galilée au sujet de l’accélération de la chute des corps ? Une mise en forme quantitative : si la vitesse croît comme le temps, alors les distances parcourues sont comme les carrés des temps (figure 8.1) ([3], p. 140-141).

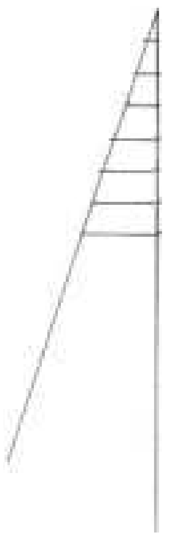


Fig. 8.1 – La ligne verticale représente le temps, et les lignes horizontales, les vitesses. Les distances sont donc proportionnelles aux triangles successifs.

Cela ne ressemble guère au raisonnement de Buridan puisqu’on pose l’hypothèse préalable de l’augmentation uniforme de la vitesse. C’est Christian Huygens qui reprend en fait vraiment le raisonnement de Jean Buridan pour lui donner une forme rigoureuse (figure 8.2)[3]. L'espace BE parcouru pendant le deuxième intervalle de temps et décomposé en un espace identique DE au premier et un espace BD inertiel, qui serait parcouru à vitesse uniforme si la gravité n'était plus là.

FIG. 8.2 – AB est un premier espace de la chute. BD, l'espace parcouru pendant le même temps, avec la vitesse acquise v. La gravité ajoute l'espace DE = AB, et une nouvelle vitesse v, etc.

Huygens a évité l'hypothèse préalable que Galilée avait dû avancer, pour la déduire de principes plus généraux, qui sont d'ailleurs des principes galiléens. Il s'agit de l'inertie et de la composition des mouvements par simple addition ou superposition. Ce sont ces principes mêmes qui permettent l'invariance longuement développée par Galilée dans la deuxième journée du *Dialogue* ([4], p. 133-281).

Or l'invariance galiléenne est évidemment liée aux possibilités de déplacement de l'individu, en bateau de préférence. Non seulement on peut laisser tomber des corps en tout temps et en tout lieu, mais aussi sur un navire en mouvement. Si ce mouvement est assez régulier, tout se passe comme au

3 – Huygens s'est en fait intéressé à ce problème dès son plus jeune âge, en 1647, pour conclure dans sa publication la plus importante, en 1673 ([5], p. 124-135).

repos. C'est de ce fait que l'on tire la conclusion suivante : non seulement la gravité agit de la même façon en tout temps et en tout lieu, mais aussi quel que soit le mouvement (uniforme) du corps sur lequel elle agit.

Tout cela, excepté la présentation proposée par Huygens, est assez connu. Si nous l'avons rappelé c'est pour montrer que de la simple possibilité du mouvement de l'individu, on déduit des lois de la nature : ce sont les seules lois invariantes possibles. Huygens en déduira également les règles des chocs élastiques, en effectuant par la pensée des chocs sur un navire en mouvement : le résultat de l'expérience effectuée sur le navire est supposé connu mais le spectateur qui se trouve sur la berge en voit une autre, dont il peut voir le résultat, ou plutôt le déduire par simple composition des vitesses du résultat qu'il connaissait préalablement.

L'observateur mobile établit donc des lois que n'aurait pu trouver un observateur immobile, qui voit pourtant tout aussi bien tomber, ou se choquer, des corps. On peut rapprocher ceci du fait que l'on démontre, un peu plus tard, des théorèmes de géométrie à deux dimensions à partir de figures à trois dimensions.

L'espace de la chute des corps et des chocs est maintenant celui d'un observateur qui se déplace : il est donc homogène et isotrope, bien que la gravité elle-même fasse apparaître une direction privilégiée. L'invariance est en effet la même le long de cette direction verticale, et les mouvements s'y composent de la même façon qu'horizontalement. La direction verticale est privilégiée pour l'action de la gravité, mais non pour les lois de la physique en général. C'est ce qu'ajoute Huygens au discours de Galilée.

Toutes les machines simples font intervenir la gravité, qui n'existe évidemment pas en géométrie. Mais un pendule, un levier ou une pompe à vide, fonctionnent de la même façon ici ou là, un peu plus haut ou un peu plus bas, à condition cependant de ne pas monter sur le Puy de Dôme ! L'homogénéité et l'isotropie sont locales. La physique du XVII^e siècle peut justement se caractériser par sa localité, propriété qui permet paradoxalement de la délocaliser puisque cela permet de dire que tout se passe de la même façon quelle que soit la situation par rapport à la Terre ou à l'Univers. Il faut bien avoir établi les lois locales pour en montrer les limites et l'on ne reviendra pas à une description globale centrée sur la Terre ou sur le Soleil.

L'espace de Huygens est celui de nos relations avec des corps en mouvement, il est relatif et indéfini ; on peut le dire « indifférencié ».

2. L'espace géométrique de Poincaré

Les analyses de Poincaré, à l'inverse de celles de Mach, ne font pas référence à l'histoire. Mach ne parle de la constitution d'un espace que dans son *Analyse des sensations* et non dans sa *Mécanique*. Il est alors question du mouvement volontaire des yeux et non de la chute des corps ni des machines. Poincaré analyse d'abord soigneusement les espaces de nos perceptions, visuelles et autres, pour conclure : « Ainsi l'espace représentatif, sous sa triple forme, visuelle, tactile et motrice, est essentiellement différent de l'espace géométrique. Il n'est ni homogène, ni isotrope ; on ne peut même pas dire qu'il ait trois dimensions. » ([7], chap. IV.)

Au lieu de s'intéresser au statut historique ou scientifique de l'espace géométrique, Poincaré tente, comme l'avait fait Mach, d'en décrire la genèse pour tout individu quel qu'il soit. Il s'agit alors, bien entendu, d'un espace géométrique général et universel dont on se demandera à quelle condition et pourquoi il peut être celui de la géométrie d'Euclide. Poincaré est tellement clair dans ses propos qu'il vaut mieux lui laisser largement la parole : « Aucune de nos sensations, isolée, n'aurait pu nous conduire à l'idée de l'espace, nous y sommes amenés seulement en étudiant les lois suivant lesquelles nos sensations se succèdent. Nous voyons d'abord que nos impressions sont sujettes au changement ; mais parmi les changements que nous constatons, nous sommes bientôt conduits à faire une distinction. Nous disons tantôt que les objets, causes de ces impressions, ont changé d'état, tantôt qu'ils ont changé de position, qu'ils se sont seulement déplacés. Qu'un objet change d'état ou seulement de position, cela se traduit toujours pour nous de la même manière : par une modification dans un ensemble d'impressions. Comment donc avons-nous été amenés à les distinguer ? » [7].

Nous pouvons les distinguer parce que nous pouvons compenser les seconds par un déplacement de notre propre corps, et ceci n'est possible que parce qu'il existe des corps que nous appelons « solides ». Ce sont ces déplacements, semblables aux déplacements des corps solides mais réciproques, qui constituent vraiment l'espace dans lequel nous nous représentons ensuite les phénomènes. « Ainsi se trouve définie, grâce à cette réciprocité, une classe particulière de phénomènes que nous appelons déplacements. Ce sont les lois de ces phénomènes qui font l'objet de la géométrie [...]. S'il n'y avait pas de corps solides dans la nature, il n'y aurait pas de géométrie. »

Notre espace est donc celui de la géométrie grâce à notre mobilité qui nous permet d'envisager celle des corps comme relative. Si nous considérons qu’un déplacement ne nous modifie pas fondamentalement, il ne modifie pas non plus les corps pour lesquels on peut retrouver le même point de vue en nous déplaçant nous-même. Nos déplacements compensatoires sont, de plus, indépendants de la direction par rapport à la Terre ou l’ensemble de l’Univers. C’est pourquoi nous nous les représentons dans un espace vide, homogène et isotrope. Cet espace est donc à la fois celui des figures géométrique et de notre mobilité personnelle locale, de nos rapports avec les corps proches qui nous entourent.

C’est en ce sens que la démarche de Poincaré éclaire celles de Galilée et Huygens, celles des mécaniciens de l’époque classique en général. Au lieu de contempler le Monde, on observe plus modestement ce qui se passe à côté de nous, plus modestement mais plus activement. On n’hésitera pas, en effet, à intervenir pour modifier le point de vue et établir les lois qui régissent les rapports entre ces points de vue. La quête de la connaissance du monde qui nous entoure est donc celle d’un ensemble de relations et non plus la quête de l’être qui, lui, n’est pas dans l’espace. L’espace est celui de certaines relations entre les corps, et de nos relations avec ceux-ci.

Il nous reste à nous demander si nous pouvons trouver à l’époque classique des discours explicites appuyant ce point de vue, chez des auteurs qui seraient plus bavards que Galilée et Huygens à ce sujet.

3. Choses, espace et relations au XVII^{e} siècle

Selon Francis Bacon : « . . . Nous ne nous bornons pas à constituer une histoire de la nature libre et déliée (telle qu’elle se manifeste dans son cours spontané et dans l’accomplissement de son œuvre propre). . . ; mais avant tout une histoire de la nature contrainte et tourmentée, telle qu’elle se manifeste quand l’art et l’assistance de l’homme l’arrachent à son état, la pressent et la façonnent. » ([2], p. 83.)

Cette phrase est étonnante par sa violence. L’intervention de l’homme y va bien au-delà de la simple mobilité. Bacon peut faire allusion aux machines simples qui, comme le pendule ou le plan incliné de Galilée, montrent les aspects de la chute libre en la modifiant. Il faut considérer maintenant que les expériences qui requièrent l’assistance de l’homme sont conformes au cours de la Nature, bien que les machines comme le levier ou la poulie œuvrent « contre-nature », puisqu’ils permettent de soulever de gros poids avec de petites forces. La redécouverte des textes d’Archimède est un facteur important de cette évolution. Mais il faut également envisager le déploiement plus général de pratiques nouvelles. Le regard se fait alors plus actif, plus unificateur, en liaison avec des pratiques diverses auxquelles il est lié, sans être conditionné par des finalités précises. Il s'agit toujours de comprendre la nature, mais cela peut se faire en la transformant, car cela ne modifie en rien les principes premiers.

Bacon continue ainsi : « C'est pourquoi toutes les expériences des arts mécaniques, toutes celles qui relèvent de la partie opérative des arts libéraux, toutes les expériences de ces nombreuses activités pratiques auxquelles manque encore le lien d'un art défini, s'y trouvent consignées. [...] car la nature des choses se livre davantage à travers les tourments de l'art que dans sa liberté propre. » ([2], p. 83.)

On ne peut être plus clair, et la critique constante que Bacon fait de la perception montre que l'activité, les pratiques, sont pour lui les vraies sources de connaissances même si leur but premier était l'amélioration de la vie humaine.

On a beaucoup critiqué la tendance à considérer Bacon comme le représentant, le porte-parole, de la nouvelle science. Galilée ne le connaît sans doute pas. Descartes l'a lu, mais ne s'y réfère pas. Quand à l'influence de ses textes sur la *Royal Society* de Londres après 1660, c'est une autre histoire. On peut cependant considérer que Bacon se fait l'écho d'une pensée générale qu'il est le seul à exprimer d'une façon aussi explicite. L'ensemble de son système lui est sans doute assez personnel, mais les phrases que nous avons données nous semblent emblématiques et pourraient avoir été proférées par Galilée.

Il ne sera jamais dit que cette activité, l'art des artisans, inscrit la représentation des phénomènes naturels dans un espace homogène et isotrope, un espace de type géométrique. Une telle affirmation n'aurait aucun sens à l'époque. L'affirmation même par Galilée de la nature géométrique des caractères du livre qu'est l'Univers ne nous dit rien de tel. Nous préférons en tout cas expliquer ici la géométrisation de la chute des corps par une pratique que par une décision qui serait de l'ordre du discours.

Le discours cartésien renvoie par ailleurs d'avantage à l'activité intellectuelle du sujet qu'à son activité pratique ou sa mobilité. Il en est de même de la *Logique de Port-Royal*, écrite en 1662, d'inspiration nettement cartésienne [1]. La « géométrisation du mouvement » par Descartes ne réside que dans l'assimilation de la matière à l'étendue, sans aucun appel à un espace vide comme scène sur laquelle se dérouleraient les mouvements.

La géométrisation effective du mouvement par Galilée, Huygens et Newton, son inscription dans un espace géométrique s'effectue cependant complètement, avant que le discours des équations différentielles et des champs ne vienne supplanter l'espace absolu de la dynamique newtonienne. Mais plus que la représentation dans une scène vide et homogène, la mobilité et l'activité du sujet ont d'abord déplacé l'intérêt sur les relations entre plusieurs phénomènes d'un ensemble, plutôt que sur un phénomène donné avec ses variantes accidentelles. Ce sont maintenant les variations, les divers aspects d'un phénomène qui seront considérés comme essentiels à sa compréhension. À l'appui de cette affirmation, nous laisserons le mot de la fin à Edme Mariotte, qui déclare dans son *Essai de Logique,* en 1678 ([6], p. 94) : « On demande quelquesfois, ce qu'une chose est en elle-mesme ; mais il est presque toujours impossible de satisfaire à cette demande : car puisque nous ne connoissons les choses naturelles que par les effets qu'elles font en nous, ou sur les autres choses ; ou par les effets que nous faisons en elles, ou qui sont faits en elles par d'autres choses ; et que les effets ne se font que par le rapport que les choses ont les unes aux autres, il est évident que nous ne pouvons savoir ce qu'elles sont en elles mesmes, et qu'il suffit de connoître ce qu'elles sont à nostre égard, et par rapport aux autres choses. »

Si les phénomènes ne sont plus observés que dans leurs relations, il en est de même des choses dont ils sont censés provenir. L'ontologie cède la place à une connaissance des relations, dont le cadre est conçu par l'intuition au lieu d'être simplement perçu.

Bibliographie

[1] Arnauld, A., Nicole, P., *La logique ou l'art de penser*, coll. Tel, Gallimard, 1992.

[2] Bacon, F., *Novum organum*, Trad. française et notes Malherbe M. , Pousseur J.-M., PUF, 1986.

[3] Galilée, *Discours concernant deux sciences nouvelles*, A. Colin, 1970.

[4] Galilée, *Dialogue sur les deux systèmes du monde*, Trad. française Fréreux R. avec le concours de F. De Gandt, Seuil, 1992.

[5] Huygens, C., *Horologium oscillatorium*, *in* Œuvres complètes, vol. XVIII, Société hollandaise des sciences, La Haye, 1889-1950.

[6] Mariotte, E., *Essay de Logique*, textes revus par Picolet G., avec la collab. d'A. Gabbey, Fayard, 1992.

[7] Poincaré, H., *La science et l'hypothèse*, 1902, Flammarion, 1968.

[8] Vilain, C., *L'espace classique*, Acta Cosmologica, fasciculus XXIV-1, 1998, 149-167.

09

Espaces mathématiques, espaces philosophiques

Jean-Jacques Szczeciniarz

Nous restons marqué par la question posée par un de nos éminents collaborateurs : « Jusqu'où la reconstitution de la vie de la Rome antique peut-elle aller dans l'exactitude ? Existe-t-il un théorème de limitation en sciences humaines fondées sur la structure physique du temps ? » Cette question qui reste ouverte mérite à elle seule une autre session qui deviendrait alors la session du peplum.

1. Introduction

Il existe certainement une question que pose l'espace. Notre perception suppose une appréhension dans des conditions contraignantes, il y faut toujours une forme de l'extériorité. L'espace est ce par quoi il peut se former de l'extériorité. Nous pouvons nous représenter mentalement cette extériorité, elle conditionne toute appréhension de nos objets, ce par quoi des formes même peuvent leur être attribuées. La question classique que nous allons poser est la suivante : l'apparition et le développement des structures mathématiques grâce auxquelles nous avons conçu les contraintes qui gouvernent notre façon de situer, de déplacer des objets dans cette forme ou par cette forme d'extériorité, nous ont-ils éclairé sur la nature de l'espace ? Ou au contraire n'ont-ils fait que nous permettre de prélever certaines propriétés de l'espace, par exemple dans certains cas son homogénéité ? Ces propriétés que nous désignons, au moins pour les construire, en les formalisant, font-elles connaître l'espace dans sa structure essentielle ? Ou bien assimilons-nous faussement ces particularités locales à tout l'espace, en globalisant indûment, ou en les considérant comme un aspect irréductible de l'espace ? N'avons-nous pas pris pour une définition ce qui n’est au fond qu’une construction abstraite permettant une certaine forme d’exploration qui reste à la surface de cette extériorité ?

Chaque fois qu’une réflexion sur l’espace démarre, elle pose toujours son concept comme une sorte d’irréductible, comme une chose en soi, que les concepts de la science ne sauraient réduire. Comme pour les concepts majeurs de la philosophie, la matière, le temps, le concept philosophique de l’espace ne peut être construit par les sciences et plus particulièrement les mathématiques.

Ce problème est thématisé par toute philosophie, et de façon particulièrement puissante par celle de Kant. Les affirmations philosophiques sont des affirmations de spécificité ou d’irréductibilité des concepts philosophiques aux théories scientifiques qu’ils impliquent pourtant. Pourquoi ? Parce qu’il est toujours question, dans les réflexions sur l’espace, de la façon dont une pensée se déploie dans ce qui n’est pas elle, dont elle s’efforce de maîtriser rationnellement sa situation, d’intervenir, ou au contraire dont elle s’arrête dans un repos contemplatif ; l’espace est toujours ce dans quoi je suis, un contenant dans lequel je suis immergé, qui dépasse et appelle en même temps une structuration conceptuelle. Il a souvent été conçu comme pré-mathématisé, ou bien comme de l’ante-mathématique ; il paraît incontestable que c’est la forme à travers laquelle je suis en rapport avec l’extérieur, mais c’est aussi à travers le concept d’espace que je conçois, comme dit Hegel, la quantité indifférente précisément dans ce qu’elle comporte d’extériorité. Il est nécessaire de commencer par aborder l’indifférence du quantitatif, l’espace suppose un rapport insupressible à la quantité. Mais dans cette optique encore, l’espace semble condamné à rester sur le seuil de la formalisation. Il peut éventuellement l’alimenter, il reste une matière inépuisable que n’entament pas les concepts reproduisant ses propriétés.

Cet article sera consacré à cette question examinée d’abord à travers le concept kantien d’espace, puis à travers le changement de point de vue, radical selon moi, produit par l’intervention de la géométrie et de l’algèbre de la fin du XIX^e siècle : nous devons concevoir l’espace d’abord et surtout à travers le concept de groupe de transformations. Mais je maintiendrai toujours cette question : pour concevoir et se représenter la notion même de transformation ne faut-il pas s’être donné déjà un certain concept d’espace pour que soit construite et réalisée la transformation ? Il y a l’espace que permet de concevoir la transformation elle-même, l’espace qui rend possible cette transformation, et l’espace auquel permet d’accéder le concept de groupe de transformations : les notions d'homogénéité et d'isotropie subsistent-elles, et de quelle façon ? Ou sont-elles épuisées par les propriétés des groupes opérant ? Des questions du même genre se posent à propos de l'orientation, de la direction, ou encore de la profondeur et de l'horizon, que je traiterai dans un prochain exposé.

2. L'espace de Kant

L'espace est une forme *a priori* de notre intuition, et il est possible de le décrire. Il faut entendre par là que l'espace n'est pas de nature intellectuelle ou conceptuelle. Il est un irréductible intuitif, par quoi nous avons rapport immédiatement et dans une certaine forme à l'extériorité. Le fait qu'il soit non réductible au concept se prouve pour Kant par ses spécificités. Il est infini ce que ne peut être le concept, il est un tout donné d'abord par quoi sont ensuite appréhendées les parties. Et surtout il est pour Kant tridimensionnel. Enfin il possède, dans son appréhension, la distinction irréductible de la droite et de la gauche. Je dois insister sur la remarque suivante : il reste selon moi toujours vrai aujourd'hui dans les mathématiques que la tridimensionnalité comporte une spécificité importante ; même $\mathbf{R}^3$ n'est pas exactement $\mathbf{R}^n$ pour $n = 3$. De même que la différence de la droite et de la gauche reste une propriété qui est explorée différemment mais qui se propage à travers toutes les dimensions. Ce sont ces formes disparates de l'intuition que je laisserai se détacher tout au cours de cet article.

Kant appelle exposition métaphysique la représentation claire, quoique, non détaillée de ce qui appartient à un concept, et c'est à cette exposition qu'il se livre dans la première partie de l'Esthétique transcendantale elle-même première grande subdivision de la *Critique de la Raison pure*.

L'espace n'est pas un concept empirique dérivé d'expériences extérieures. Pour que je puisse rapporter certaines sensations à quelque chose d'extérieur à moi ; et de même pour que je puisse me représenter les choses comme en dehors et à côté les unes des autres et par conséquent comme n'étant pas seulement différentes mais placées en des lieux différents, il faut que la représentation de l'espace soit déjà posée comme fondement. L'espace est une représentation externe *a priori* qui sert de fondement à toutes les intuitions externes. C'est à ce titre que l'espace n'est pas un concept discursif mais une forme de l'intuition pure. Toutes ces affirmations demandent des explications.

Kant sépare de la pensée l'*intuition*, qui est la faculté par laquelle nous nous rapportons immédiatement au monde extérieur. Le monde extérieur est ainsi une donnée, mais nous devons la reconquérir en exposant les conditions par lesquelles la perception de ce donné est possible. Il procède à une analyse dans ce dessein, mais il suppose que l'on peut, par son intermédiaire, isoler dans la pensée une faculté distincte des autres qui est l'intuition.

Distinguons à la suite de Kant matière et forme. « De toutes les intuitions, aucune n'est donnée *a priori*, si ce n'est la simple forme des phénomènes : espace et temps... Mais la matière des phénomènes, ce par quoi des choses nous sont données dans l'espace et dans le temps, ne peut être représentée que dans la perception, donc *a posteriori*. » (T.P. 572.)

« Cette faculté d'intuition *a priori* ne concerne pas la matière du phénomène, c'est-à-dire ce qui est sensation en lui, car c'est là l'élément empirique, mais seulement sa forme. » (*Prolégomènes* § 11, G. 46.)

« Cette forme pure de la sensibilité peut encore s'appeler intuition pure. Ainsi quand je détache de la représentation d'un corps ce qui en est pensé par l'entendement, comme la substance, la force, la divisibilité, etc. et aussi ce qui appartient à la sensation comme l'impénétrabilité, la dureté la couleur, etc., il me reste pourtant encore quelque chose de cette intuition empirique : l'étendue et la figure. Celles-ci appartiennent à l'intuition pure qui réside *a priori* dans l'esprit, même indépendamment d'un objet réel des sens ou de toute sensation, en qualité de simple forme de la sensibilité. » (T.P. 64.)

L'intuition est sensible, on ne le répétera jamais assez. C'est là une des grandes thèses kantiennes. Au sens où il n'est pas pour Kant possible que nous disposions d'une intuition intellectuelle par laquelle nous pourrions appréhender conceptuellement et immédiatement un objet intellectuel. En ce cas, selon Kant, nous serions en position non pas de récepteurs passifs, comme c'est le cas lorsque nous recevons les sensations du monde extérieur, mais en position de créateurs actifs, ce que nous ne sommes évidemment pas. La forme de l'intuition n'est pas un objet car elle n'est pas un phénomène. Mais c'est — en simplifiant — le cadre où les phénomènes prennent place. Elle n'est pas perçue mais appartient à la constitution de la sensibilité. « La simple forme de l'intuition, sans substance, n'est pas un objet en elle-même, mais la simple condition formelle de cet objet (comme phénomène), comme l'espace et le temps pur, qui, tout en étant quelque chose en qualité de formes de l'intuition, ne sont pas eux-mêmes objet d'intuition (*ens imaginarium)*. » (T.P. 288.)

Dans l'*Introduction* à la *Critique de la Raison pure*, Kant indique comment obtenir la forme spatiale en décomposant la représentation d'un corps. « Enlevez peu à peu du concept expérimental que vous avez d'un corps tout ce qu'il y a d'empirique : la couleur, la dureté, ou la mollesse, la pesanteur, l'impénétrabilité, il reste cependant l'espace qu'occupait le corps (maintenant disparu) et que vous ne pouvez pas faire disparaître. » (T.P. 42B.) Kant s'est livré à une exposition métaphysique du concept d'espace qui consiste à montrer essentiellement comment cette forme de l'intuition est *a priori*. Cette exposition doit être complétée par une exposition transcendantale.

Celle-ci montre comment un concept est un principe capable d'expliquer la possibilité d'autres connaissances synthétiques *a priori*. Le concept de synthèse est un autre concept-clé qui est au centre de la doctrine kantienne. Un jugement est synthétique s'il apporte une information supplémentaire de celle qui est enfermée dans les concepts mis en jeu dans ce jugement. Exemple très connu et posant problème à lui seul, le corps est pesant. Un tel jugement doit pouvoir être *a priori* c'est-à-dire indépendant de l'expérience, pour que la science soit possible. L'expérience doit pouvoir le réaliser et donc le confirmer. Que doit donc être la représentation de l'espace pour que la géométrie soit possible ? Il faut que ce soit une intuition, seule capable d'expliquer que nous puissions dépasser ce qui se trouve dans le seul concept. « Ainsi notre explication seule rend compréhensible (begriflich) la géométrie comme connaissance synthétique *a priori*. » (T.P. 68.)

« Entre deux points la ligne droite est le plus court chemin. » Dans le concept du droit il n'y a rien qui se rapporte à la quantité. À y regarder de plus près on pourrait objecter que le concept de chemin le plus court suppose un concept de *minimum* ou plutôt d'*extremum* qui est lié à des concepts de quantité et donc qui doivent se trouver dans l'entendement. Cette affirmation sur laquelle je reviens à la fin fait partie d'une axiomatique de l'espace.

Cette description de l'espace se fait par un mouvement qui se déploie et engendre une partie de l'espace. Le mouvement comme description d'un espace est un acte pur de la synthèse successive du divers (la synthèse demande donc de composer et d'unifier, elle est successive ce qui veut dire que cet acte est scandé par la succession temporelle) et cette action du sujet est absolument nécessaire pour penser et les figures géométriques et le temps lui-même. Comment se représenter une ligne ou un cercle, sans les tracer par la pensée ? Et comment se représenter le temps sinon en décrivant une ligne, ce qui nous assure qu'il est une multiplicité à une dimension ?

Cette description définit la nature de ce que Kant appelle les *grandeurs extensives*, *i.e.*, celles pour qui la représentation des parties rend possible la représentation du tout (par opposition avec les grandeurs intensives). Cette synthèse renvoie, dans la théorie kantienne, à une autre faculté qui est l'imagination, transcendantale car elle rend possible *a priori* la construction des concepts. Elle est productrice, car elle produit du nouveau. Elle est, comme il se doit pour l'imagination créatrice, fabricatrice de ce que le concept peut organiser de la matière imaginaire. Elle produit ce que Kant appelle des schèmes. C'est sur cette synthèse de l'imagination productrice que se fonde la géométrie avec ses axiomes.

Il faut insister sur ce point que Kant reconnaît et analyse lui-même : l'espace suppose, pour être conçu ou analysé, le concept de mouvement. Pour mesurer l'espace, « le progrès de l'intuition doit être reconnu sans limites », d'où le fait que l'espace soit infini donné. Or que signifie ce progrès sinon que je dois infiniment pouvoir porter l'unité arbitraire de longueur que j'ai choisie, l'aune ou le pied, le long d'un segment ? L'unité de mesure est par conséquent mobile. Naturellement les propriétés intrinsèques du mouvement de ce mobile comme la vitesse, l'accélération, etc. n'interviennent pas ici et sont réservées à la cinématique ; mais l'effet spatial des mouvements est retenu. Prenons le concept de mouvement au sens large, comme catégorie philosophique.

En effet « tous les autres concepts qui appartiennent à la sensibilité (à part l'espace et le temps), même celui de mouvement qui unit ces deux éléments, présupposent quelque chose d'empirique. Car le mouvement présuppose quelque chose de mobile. Mais dans l'espace considéré en lui-même il n'y a rien de mobile ; c'est pourquoi le mobile doit être quelque chose qui ne peut être trouvé dans l'espace que par l'expérience, par conséquent un *datum* empirique. Dès lors l'Esthétique transcendantale ne peut compter le concept de temps parmi ses *data* ; car le temps lui-même ne se transforme pas mais seulement ce qui est dans le temps... »

Le mouvement est synthèse de l'espace et du temps au sens objectif et subjectif. En effet il y a le mouvement que l'on perçoit dans l'espace, mouvement d'un mobile, et le mouvement comme description d'un espace qui suppose en tant que description perçue la mise en œuvre des facultés.

Les constructions de concepts en effet sont des mouvements. Kant différencie philosophie et géométrie en ce que la dernière procède par construction de concepts et la première par concept. La construction se fait dans l'espace *in concreto*. On peut donc manipuler les mouvements (abstraits), les additionner : « Le mouvement d'un objet dans l'espace n'appartient pas à une science pure, et par conséquent pas non plus à la géométrie, parce qu'on ne peut pas connaître *a priori*, mais seulement par expérience, que quelque chose est mobile. Mais le mouvement comme description d'un espace est un acte de la synthèse successive du divers dans l'intuition externe en général par l'imagination productrice... et appartient non seulement à la géométrie, mais encore à la philosophie transcendantale... » (T.P. 133 note.)

L'analyse kantienne concerne aussi bien le temps. Pour pouvoir penser le changement, même interne, nous devons figurer le temps comme forme du sens interne par le tracé d'une ligne... Mais le tracé lui-même suppose quelque chose qui bouge et quelque chose qui reste ; chaque morceau tracé subsiste indéfiniment une fois passé au suivant. Ceci implique que sous le changement il y ait du permanent, lequel est produit par une affection externe, donc un mouvement. « (...) Si je tire une ligne par la pensée ou que je veuille penser le temps d'un midi à un autre ou même seulement me représenter un certain nombre, il faut d'abord nécessairement que je saisisse une à une dans ma pensée ces diverses représentations. Si je laissais toujours échapper de ma pensée les représentations précédentes (les premières parties de la ligne, les parties antérieures du temps, ou les unités représentées successivement), et si je ne les reproduisais pas à mesure que j'arrive aux suivantes, aucune représentation entière... ne pourrait jamais se produire. » (T.P. 115.)

3. Une question sur la théorie des groupes

L'espace kantien n'implique-t-il pas une théorie des groupes de transformations ? Kant allègue trois exemples pour prouver la spécificité intuitive de l'espace. Le théorème prouvant que la somme des angles d'un triangle est égale à deux droits. Sa démonstration : prolongement d'un côté, tracé de la parallèle à un autre par un sommet repose sur le postulat que la translation dans le plan de ces angles n'altère pas leur valeur. Autrement dit, ni la translation ni la rotation n'altèrent la grandeur ou la forme des figures. Dans la construction de figures homothétiques, par laquelle on conserve invariable un rapport entre deux grandeurs bien que ces grandeurs varient. On prétend que l'homothétie conserve les formes des figures et non leurs grandeurs, et c'est cette indétermination de la grandeur par rapport à la forme qui fait, selon Kant, de la synthèse géométrique une synthèse générale. Kant utilise enfin le fameux exemple des objets symétriques, les mains qui ne peuvent coïncider que par retournement, la symétrie conservant certaines propriétés de la figure mais en l'inversant. Ceci prouve selon Kant le caractère irréductible de l'espace au concept de l'entendement et donc son caractère intuitif. On sait que laisse invariant la forme et la grandeur des figures, l'ensemble des opérations qui forment le groupe des déplacements. Il suffit de prendre le point d'affixe z, (dans le plan complexe) le déplacement est en fait une translation qui amène le point M, en le point M' d'affixe $z' = z + a$. La rotation amène ce point z' en le point M d'affixe $z =$ ï$+z\, \mathrm{e}^{i\alpha}$. De la même façon, l'homothétie correspond aux opérations du groupe linéaire de la variable complexe $z' = az + b$, avec a et b nombres complexes. Le premier cas est un cas particulier du second, le groupe des déplacements étant un sous-groupe du groupe des similitudes. Quant au dernier cas, c'est celui du groupe des inversions qui conservent les grandeurs tout en changeant leur orientation. Ce groupe ayant comme formule avec un pôle d'inversion quelconque : $Z - z_0 = \mu/(z - z_0)$.

L'espace est clairement soumis aux possibilités de transformations, de déplacement et rotations. Son caractère euclidien pose que l'espace doit comporter la possibilité d'ajouter une unité de mesure que l'on peut donc déplacer *ad infinitum*. On le lui a reproché, cet espace est euclidien. Mais là encore le caractère euclidien de cet espace se résume dans des propriétés d'espace vectoriel sous-jacent à l'espace affine à trois dimensions. Les coefficients de l'équation linéaire par laquelle nous déterminons un point de cet espace nous sont donnés par une base de cet espace vectoriel, par l'intermédiaire de laquelle nous engendrons n'importe quel vecteur de l'espace qui, par la liaison qu'il entretient aux trois vecteurs indépendants de la base choisie, est linéairement dépendant de cette base.

En conséquence, tous les traits qui visaient à spécifier l'espace, en lui donnant un caractère intuitif irréductible aux concepts de l'entendement, tous ces traits peuvent être conceptualisés en termes de la théorie des groupes. Ce qui fait de ce point de vue, la spécificité de l'espace euclidien et qui rend compte de ses aspects intuitifs, se retrouve dans les propriétés que fait saisir un certain groupe de transformations. L'espace euclidien est celui qui est invariant pour l'action du groupe des déplacements.

Il est tout à fait possible d'analyser un déplacement en termes tirés de la théorie kantienne. La notion d'invariant, pour telle ou telle action d'un groupe, peut être considérée comme une transposition de la nécessité de poser une forme de permanence. La différence tient toute entière dans le fait que cet invariant est produit comme une forme extrasubjective. Il peut en outre être nécessaire que, de façon plus primitive encore dans l'acte de production de l'invariant, il subsiste une forme de permanence mais ce n'est plus de cet invariant-là qu'il s'agit.

Quant aux autres concepts qui proviennent de celui de groupe, on peut les soumettre à une analyse du même genre. Il importe ici d'abord que l'on puisse composer les opérations : deux déplacements donnent encore un déplacement. On considère cette composition de deux points de vue. C'est une autre forme de stabilité qui énonce que nous pouvons composer des opérations (ou des éléments) de sorte que nous restions dans l'ensemble des opérations ou des éléments initiaux (loi de composition interne). Il faut considérer que nous avons affaire à une synthèse accomplie. La loi de composition objective la synthèse. Comme on le sait, une loi de composition est une abstraction de l'addition ou de la multiplication dont on ne retient que les conditions d'exercice sans que l'on ait pour autant à expliciter une unité de mesure qui homogénéise l'opération. Ici deux rotations, par exemple, se composent pour donner une rotation. Le second point de vue est le point de vue subjectif : la composition est une promotion de l'activité originairement synthétique de l'entendement au sens kantien. Si je décris ce qu'est la synthèse du point de vue de la synthèse transcendantale, je dois tomber sur une forme de composition interne. Les deux points de vue se retrouvent aussi bien dans le concept d'un inverse qui caractérise un élément du groupe, que dans celui d'élément neutre. Un déplacement composé avec son inverse donne une immobilité. Il semble bien, si l'on admet la possibilité de ces deux points de vue d'analyse, que le concept de groupe n'est pas incompatible avec les analyses kantiennes, en particulier celles qui portent sur les synthèses produisant les concepts géométriques.

4. L'intuition disparaît-elle ?

La question de l'espace se pose-t-elle encore si l'on tient compte des transformations de son concept mathématique ?

Si je reprends les profondes analyses de Vuillemin sur Riemann, il apparaît clairement que la forme intuitive de l'espace telle que la défendait Kant est recouverte par des déterminations de nature conceptuelle. On sait que Riemann introduit le concept de mutiplicités n-fois étendues[1]. Il faut faire le départ entre les axiomes concernant les grandeurs n-fois étendues et ceux qui concernent les relations métriques entre ces grandeurs. Lorsque Riemann parle de multiplicités, il spécifie que les relations métriques « ne peuvent être

1 – Riemann conçoit la quantité n-fois étendue comme injectable dans R^m (m plus grand que n). Mais il ajoute qu'une quantité n-fois étendue peut être conçue comme composée de pièces dont chacune peut être appliquée injectivement sur une partie de R^n. Riemann conçoit également ce que nous appelons le concept d'*atlas*. Il explore de façon plus intuitive comment une variété de dimension $n+1$ peut être construite à partir d'une variété de dimension n et une variété de dimension un. Et comment une variété de dimension n peut se décomposer en sous-variétés de dimension $n-1$ et un [8].

étudiées que dans les concepts de grandeurs abstraites, et leur dépendance ne peut se représenter que par des formules. Dans certaines hypothèses, cependant, elles sont décomposables en rapports qui, pris séparément, sont susceptibles d'une représentation géométrique, et par là il devient possible d'exprimer géométriquement les résultats du calcul » ([8], p. 285). On passe avec Riemann à l'analyse des structures et à l'axiomatique, de l'intuition donc au concept ([11], p. 408). Si cette analyse ainsi présentée semble incontestable, la rupture conceptuelle sous la forme du passage à une abstraction conceptuelle plus générale est-elle une rupture avec l'intuition ?

Riemann nous donne une réponse nette. Il procède par genre et différence spécifique. « Les axiomes de la géométrie, comme les marques d'une définition scolastique, apparaissent comme les déterminations successives de concepts génétiques se terminant avec l'espace euclidien. » Je reprends les commentaires de Vuillemin. Il relève deux couples d'oppositions conceptuelles : discontinu – continu ; et la division en parties indéterminées et déterminées. Si l'on associe déterminées et continues, on obtient le concept de multiplicités de *quanta* continus. Ce concept possède les propriétés suivantes : (1) la multiplicité est n-fois continue et ce nombre n de dimensions lui est essentiel ; (2) c'est là sans doute le plus important, il est possible de réduire la détermination de lieu à une détermination quantitative.

D'où une formulation en un axiome, que commente Vuillemin : dans une multiplicité à n dimensions, les déterminations peuvent être exprimées par n déterminations de grandeur. (L'espace est une multiplicité numérique à n dimensions.) Cette proposition est le résultat d'une extraordinaire abstraction : il a fallu réduire le lieu à une détermination numérique et produire un concept de dimension lié à la quantité et disjoint de la notion de mesure. L'intuition a reflué dans le quantitatif pour permettre la production d'un nouveau concept, celui de dimension *topologique*. Il reste une détermination en termes de quantité, mais en dehors de la constitution de condition de mesure. Un espace dépendant de pures conditions quantitatives est né ainsi. Mais en même temps il reste évidemment dépendant de conditions de situation : on peut dire qu'une telle grandeur est plus grande qu'une autre si elle est lui est intérieure. Corrélativement, Riemann peut mettre en évidence l'autre caractéristique de la grandeur étendue, ses propriétés métriques qui, elles, sont indépendantes du lieu. « Le système de couleurs est une multiplicité à trois dimensions, dans la mesure où, conformément aux recherches de Thomas Young et de Clark Maxwell, chaque couleur peut être représenté par un mélange de trois couleurs primitives... De même, nous pouvons considérer un système de sons simples comme une multiplicité à deux dimensions, si nous ne considérons que la hauteur et l'intensité laissant de côté les différences de timbre. Cette généralisation de l'idée est bien faite pour exprimer la distinction entre l'espace à trois dimensions et les autres multiplicités. Nous pourrions, comme nous l'apprend l'expérience quotidienne, comparer la distance verticale de deux points avec la distance horizontale de deux autres parce que nous pouvons appliquer une mesure à la première paire et ensuite à la seconde. Mais nous ne pouvons pas comparer la différence de deux sons de hauteur égale et d'intensité différente avec deux sons d'intensités égales et de hauteurs différentes. » [1].

D'où le second axiome posé par Riemann : les longueurs des lignes sont indépendantes de leur situation. D'après le premier axiome, les « déterminations » données, ou points, peuvent « être assignées comme des fonctions de n variables, $x_1, x_2, \ldots, x_n$. D'après le deuxième, ces points peuvent se déplacer, c'est-à-dire que ces n grandeurs sont données comme fonction d'une variable t. ([11], p. 409.)

Il pose donc les conditions qui permettent de mesurer les grandeurs x en unités. Dans ce dessein, Riemann suppose continus les rapports entre les changements de ces grandeurs dx. Un élément de ligne, dimension un, sera défini par la direction de la courbe ou dérivée, $\mathrm{d}x_1/\mathrm{d}x_2$ qui est constante à la limite [11]. Pour exprimer la longueur d'un élément d'arc ds en fonction des x et des dx, Riemann suppose que la longueur de cet élément, abstraction faite des grandeurs du second ordre, est invariante quand tous ses points subissent le même changement de lieu infiniment petit. L'élément d'arc entre deux points infiniment proches est une fonction homogène quelconque de degré 1, des $\mathrm{d}x_i$, invariante quand tous les $\mathrm{d}x_i$ changent de signe et où les coefficients sont des fonctions continues des grandeurs x_i. Dans les présentations systématiquement modernisées, on procède ainsi. Soit $c(t)$ un arc ; sa longueur est donnée par une intégrale $\int f.c(t)\mathrm{d}t$.

On prend pour fonction f la fonction qui, à chaque point $c(t)$, assigne la longueur du vecteur tangent $c(t)$. Ce vecteur a une longueur si l'espace vectoriel auquel il appartient est normé. Puisque les concepts d'espace tangent et de vecteur tangent sont intrinsèques, on a bien les réquisits riemanniens. En particulier, celui selon lequel la norme doit être une fonction positive homogène de degré un correspond à la définition de la norme. L'autre condition correspond à la définition de ce que nous appelons une variété riemannienne.

C'est l'expression sans doute la plus connue de la conception riemannienne de l'espace :

$$\mathrm{d}s^2 = \Sigma a_{ij}\mathrm{d}x_i\mathrm{d}x_j \quad a_{ij} = \phi_{ij}(x_1, x_2, \ldots, x_n).$$

Plusieurs remarques s'imposent sur ce processus de création riemannien qui a pour effet de repousser les fonctions de l'intuition au plus haut du côté des éléments de la subjectivité, ou plus loin du côté objectif. La constitution d'une mesure possible, ou d'une métrique, implique que l'on puisse normer la variation au-delà de ce qui pouvait sembler irréductible dans la construction de Kant. Il faut mesurer, donc disposer d'unité de mesure et donc décomposer les lignes en éléments. Il faut pour ce faire que les rapports entre les variations de grandeurs soient constants et que la longueur des éléments soit invariante quand tous ses points subissent le même changement de lieu infiniment petit. Riemann construit les conditions de possibilité de l'établissement d'une métrique en la structurant dans la construction de l'espace, dans l'infiniment petit. De ce point de vue l'invariant n'est plus ce qui reste dans la réalisation intuitive, mais le rapport des grandeurs infiniment petites évidemment construites sur l'espace de courbes, puisqu'il s'agit de directions de droites tangentes et d'invariant sous des « changements infinitésimaux ». De plus, ces variations sont vues comme des variations de variables coordonnées. De sorte que l'intuition est soumise non plus aux facultés subjectives transcendantales, mais, ici, aux concepts mathématiques de variations en coordonnées et de dérivées comme applications tangentes. Le fait de pouvoir construire de cette façon, héritée de Gauss, la longueur d'arc fait dépendre l'unité de longueur, comme unité de mesure, d'une construction intrinsèque ayant absorbé l'espace initial. Riemann a réussi à expulser hors de la subjectivité les conditions d'exercice de la mesure, pour les installer dans la construction spatiale elle-même. On peut comme précédemment essayer de cerner une nouvelle zone d'intuition. L'invariance des rapports des différentielles et celle du ds supposent une construction dans une forme d'extériorité beaucoup plus restreinte. L'homogénéité, par exemple, concerne d'abord les infiniment petits du premier ordre. C'est suivant une inspiration leibnizienne dans la différentielle elle-même que cherche à se réfugier l'intuition. La différentielle, selon un thème constant de recherche chez Riemann, devient un instrument d'exploration géométrique.

Considérons maintenant les dernières caractéristiques de cette multiplicité n-fois étendue. Riemann expose la notion de fonction de distance dans cette multiplicité. Soit donc la fonction $\Omega(x_1, \ldots, x_n; x_1^0 \ldots, x_n^0)$ qui a la même valeur pour tous les points à égale distance d'un point $P(x_1^0, \ldots, x_n^0)$. Riemann suppose cette fonction croissante au voisinage de P, dans toutes les directions et que ses dérivées existent. La différentielle première de Ω s'annule pour $x = x_0$, et la différentielle seconde est toujours positive quand on fait la substitution $x = x_0$. D'où l'axiome suivant qui pose qu'il existe une fonction $\Omega(x_1, \ldots, x_n; x_1^0, \ldots, x_n^0)$ telle que, pour $x_\nu = x_\nu^0$,

$$\mathrm{d}\Omega = \sum_{\mu=1}^{\mu=n} \frac{\partial \Omega}{\partial x_\mu} \mathrm{d}x_\mu = 0 \tag{9.1}$$

$$\mathrm{d}^2\Omega = \sum_{\mu,\nu=1}^{\mu,\nu=n} \frac{\partial^2 \Omega}{\partial x_\mu \partial x_\nu} \mathrm{d}x_\mu \mathrm{d}x_\nu > 0. \tag{9.2}$$

Quelle est la signification de cet axiome ?

Il existe une fonction de distance, c'est-à-dire que nous disposons d'un moyen de calculer une distance à l'intérieur de la multiplicité n-fois étendue, distance qui dépend de cette multiplicité elle-même (appelée, aujourd'hui une variété) ou qui en est une détermination. L'expression du second ordre est proportionnelle quadratiquement aux ds et dx. Elle est égale à une constante fois ds^2. Riemann explicite ici la manière de se donner une distance (et d'obtenir de la sorte la plus courte distance que l'on appellera géodésique). C'est une fonction des coordonnées conçue de telle sorte qu'on puisse y laisser libres des déterminations concernant la géométrie de la variété.

Surtout, l'intégrande qui donne la longueur d'arc est égal à la racine carrée d'une expression quadratique caractéristique. La valeur de cette expression est invariante par changement de coordonnées. Cette invariance la fait sortir encore davantage d'un conditionnement par l'intuition. La forme quadratique prise par l'élément de droite sur les variétés est utilisée par Riemann pour les caractériser. Il considère le cas le plus simple et se restreint dans sa dissertation aux R-variétés. Il observe que l'expression, dans une carte, d'un élément de droite dépend de $n(n+1)/2$ fonctions arbitraires g_{ij}, et que les transformations de coordonnées dépendent de n coordonnées. Donc $n(n-1)/2$ relations fonctionnelles ne dépendent pas du choix d'une carte, et suffisent à caractériser la variété. C'est alors qu'il peut fixer les relations métriques sur la variété. Et il montre que l'élément de droite prend une forme euclidienne pour un choix très particulier de carte. Le concept d'espace euclidien est donc très loin d'être coextensif à celui d'une R-variété de dimension trois.

Je ne retiens dans la suite de ce développement que le concept de courbure que Riemann est amené à introduire. Il explique que c'est seulement si l'espace est une variété de courbure constante que l'on peut maintenir que l'existence de corps, et non pas seulement de lignes sans épaisseur, ne dépend pas de la manière dont ils se trouvent sur la variété. Helmholtz considérait l'existence de corps rigides comme une condition nécessaire pour que l'on puisse mesurer une distance dans l'espace physique. Il existe une formulation mathématique du problème de Helmholtz. Il faut donner une formulation mathématique de l'idée que l'existence d'un corps est indépendante de sa situation dans l'espace. Tout corps géométrique placé dans une situation arbitraire peut être copié isométriquement en un point quelconque et une direction quelconque. De telles copies sont réalisables dans une variété de courbure constante « car les relations métriques y sont exactement les mêmes autour d'un point dans toutes les directions que dans les directions autour d'un autre et donc toutes les constructions peuvent être effectuées à partir de l'un ou l'autre » [8]. Riemann étend ainsi la notion de courbure gaussienne à une variété à n dimensions. La mesure de l'élément d'arc devient, dans la dissertation riemannienne,

$$\mathrm{d}s^2 = \frac{\sum_{\nu} \mathrm{d}x_{\nu}^2}{1 + \frac{K}{4}\sum_{\mu} x_{\mu}^2}.$$

On obtient un espace euclidien avec $K = 0$ partout ; si $K > 0$, on a un espace riemannien et $K < 0$ donne la géométrie de Lobatchevski.

La notion de courbure, l'une des plus déterminantes pour la géométrie, achève de définir l'espace hors de la subjectivité intuitive, dans le sens où elle produit des déterminations qui nous paraissent les plus extérieures et les plus expérimentables intuitivement à partir de propriétés intrinsèques à l'espace ou à la variété elle-même. Elle fait dépendre la forme de notre perception de la figure multidimensionnelle des propriétés de cette figure elle-même. La forme de notre intuition est encore absorbée par sa propre production.

5. L'intervention de la théorie des groupes

Helmholtz a construit la géométrie sur des axiomes élémentaires qui n'ont plus besoin de la notion d'élément d'arc ni celle d'intégration. Et surtout, ce sont les mérites que lui reconnaît Lie, il opère avec le concept de groupe de mouvements. « Riemann suppose au point de départ l'expression algébrique qui représente sous la forme la plus générale la distance de deux points infiniment proches et il en déduit les conditions de la mobilité des figures rigides. Pour ma part, partant du fait observé selon lequel le mouvement de figures rigides est possible dans notre espace avec le degré de liberté que nous connaissons, je déduis la nécessité de l'expression algébrique prise par Riemann comme un axiome. » ([11], p. 414.) Je cite les six axiomes de Helmholtz.

H$_1$ *Axiome de continuité et de dimension* : « L'espace à n dimensions est une multiplicité n-fois étendue, c'est-à-dire que l'élément particulier déterminé en lui, le point, est déterminable par la mesure de grandeurs variables quelconques, continues et indépendantes les unes des autres (coordonnées) au nombre de n. Chaque mouvement d'un point est donc accompagné d'un changement continu de coordonnées. Si les exceptions devaient se présenter, dans lesquelles, ou bien le changement fut discontinu, ou bien aucun changement dans l'ensemble des coordonnées n'eut lieu en dépit du mouvement, ces exceptions seraient cependant limitées à quelques lieux bornés par une ou plusieurs équations (donc à des points, des lignes, des surfaces, etc.), qu'on peut immédiatement exclure de la recherche. »

H$_2$ *Axiome de l'existence des corps rigides mobiles*. Avec la définition suivante de corps solide : « Entre les n coordonnées de chaque paire de points, qui appartient à un corps indéformable, il y a une fonction indépendante du mouvement de ce dernier, qui est la même pour toutes les paires de points congruents. Des paires de points sont congruentes quand elles peuvent coïncider simultanément ou successivement avec la même paire de points de l'espace. »

H$_3$ *Axiome de la libre mobilité des corps solides*, qui se décompose en :

> **H**$_{31}$: *Axiome des invariants essentiels*. Tout point peut être mu dans une autre position. Quand un point est fixe, un autre point quelconque peut prendre ∞^2 positions. Quand deux sont fixes, un autre peut prendre ∞ positions. Et, quand trois points sont fixes, aucun mouvement n'est possible.
>
> **H**$_{33}$ *Axiome de l'invariant unique entre deux points*. Lorsque les m points $P_1, \ldots, P_m$, sont fixes entre les n coordonnées d'un autre point quelconque P, il doit y avoir précisément m équations et m seulement. Cet axiome a pour conséquence que deux points ont un invariant et un seul.

H$_4$ *Axiome de la monodromie*. Si $n - 1$ points d'un corps quelconque demeurent fixes, en sorte que tout autre point est assujetti à décrire une certaine courbe, cette courbe est fermée.

Helmholtz ajoute dans un second article les deux nouveaux axiomes.

H$_5$ L'espace a trois dimensions.

H$_6$ L'espace est infini.

En suivant les analyses de Vuillemin on constate que la traduction par Lie des axiomes de Helmholtz donne le résultat suivant.

Dans l'espace à deux dimensions, lorsque la libre mobilité est supposée possible universellement, les groupes qui correspondent à la géométrie euclidienne et aux deux géométries non-euclidiennes de Lobatchewski-Bolyai, et Riemann vérifient ces axiomes. Mais si la libre mobilité n'est requise que dans une région déterminée de l'espace, il existe un autre groupe possible dans lequel la courbe décrite par un point quelconque dans une rotation n'est pas une courbe fermée mais une spirale logarithmique. D'où l'axiome de monodromie pour exclure ce cas.

Dans l'espace à trois dimensions, si l'on suppose la libre mobilité seulement à l'intérieur d'une région déterminée :

– Ou bien les axiomes de Helmholtz sont valables sans exception dans cette région et l'axiome de monodromie est alors superflu ; les trois premiers axiomes suffisent à caractériser les trois géométries de courbure constante.
– Ou bien la libre mobilité dans une certaine région n'est requise que pour les points de position générale, tandis que les points sur une certaine ligne se meuvent, quand un point est fixe, sur cette ligne et non sur la surface.

D'autres groupes que ceux qui caractérisent les trois géométries sont possibles. Lie et Poincaré ont complété, et même résolu de façon complète, la question de la caractérisation des trois géométries dans le langage de la théorie des groupes. De façon que les trois groupes (groupe des mouvements euclidiens et groupes des mouvements non-euclidiens) soient distingués de tous les autres groupes possibles de mouvements d'une multiplicité numérique ([3], vol. 3, p. 397).

Poincaré a résumé ainsi l'œuvre de Lie ([6], vol. 11, p. 92-113). « Il a cherché de quelle manière peuvent se combiner les divers mouvements possibles d'un système quelconque, ou plus généralement les diverses transformations possibles d'une figure. Si l'on envisage un certain nombre de transformations et que l'on les combine ensuite de toutes les manières possibles, l'ensemble de ces combinaisons formera ce qu'il appelle un groupe. À chaque groupe correspond une géométrie, et la nôtre, qui correspond au groupe des déplacements d'un corps solide, n'est qu'un cas très particulier. Mais tous les groupes que l'on peut imaginer possèderont certaines propriétés et ce sont précisément ces propriétés communes qui limitent le caprice des inventeurs de géométries ; ce sont elles d'ailleurs que Lie a étudiées toute sa vie. »

6. Un retour à Kant est-il possible ?

Ce parcours effectué, nous en tirons la conclusion suivante. La géométrie et l'espace comme concept ne se conçoivent plus qu'au moyen de la théorie des groupes. L'espace euclidien est par exemple l'invariant pour l'action du groupe des déplacements. Peut-on alors considérer la théorie kantienne de l'espace comme caduque ? Nous avons déjà, en particulier dans l'analyse de Riemann, essayé de déplacer l'analyse kantienne pour lui assigner un champ de validité restreint mais réel.

La proposition de Vuillemin est de considérer que la théorie kantienne contient *in nuce* une théorie des groupes de transformations. Il y a le mouvement, comme description d'un espace, qui est un acte pur de la synthèse du divers et c'est sur cette synthèse successive de l'imagination productrice quand elle engendre les figures que se fonde la Géométrie. Mais c'est là le point, il manque dans l'*Esthétique transcendantale* de Kant un postulat concernant la continuité de l'espace. L'intuition est incapable à elle seule de prouver que l'espace a trois dimensions car la notion de dimension implique celle de continuité. C'est là ce que Lie exprimait dans son premier axiome. L'autre problème concerne les espaces de courbure non constante.

L'exposition transcendantale de l'*Esthétique* (transcendantale veut dire ici qui concerne la possibilité d'autres sciences et donc l'application de la géométrie à la physique) conduit aux quatre axiomes :

> $\mathbf{T}_1$ La représentation de l'espace est intuitive.
>
> $\mathbf{T}_2$ La représentation de l'espace est *a priori*.
>
> $\mathbf{T}_3$ La représentation de l'espace permet de déterminer *a priori* le concept des objets.
>
> $\mathbf{T}_4$ La représentation de l'espace est une propriété formelle du sujet.

L'axiome T_3 est équivalent à deux axiomes. Le premier dit qu'il est impossible de faire abstraction de l'espace pour obtenir comme résidu les objets. Il vise la doctrine leibnizienne. Le second dit que l'espace est la forme nécessaire de l'extériorité. Au sens précis de la géométrie, cet axiome tombe sous le coup de la critique, avec T_3 car ils reviennent tous deux à affirmer que l'espace est la forme de l'extériorité et qu'on est en droit de distinguer la forme de l'intuition de sa matière. Et donc, je suis ici toujours l'analyse de Vuillemin, Kant préjuge du caractère propre à la métrique. La métrique doit être communiquée par la matière à l'espace d'une multiplicité continue. T_4 tombe sous le coup de la critique puisqu'il est posé pour éviter les antinomies qui pourraient naître de l'affirmation de l'espace comme grandeur infinie donnée alors que ces antinomies n'ont pas lieu.

Reste donc pour ce qui nous concerne :

> **T**$_1$ La géométrie procède synthétiquement c'est-à-dire par intuition sensible.
>
> **T**$_2$ La géométrie procède apodictiquement c'est-à-dire *a priori*.
>
> **M** L'espace est représenté comme une grandeur infinie.
>
> **M'** L'espace est représentée comme grandeur infinie donnée.

Le premier axiome tombe lui aussi sous les coups de la réflexion théorique géométrique. Dans la *Critique*, pour ne prendre que cet exemple, Kant oppose parmi les principes de la géométrie pure ceux qui sont réductibles à la logique (principe d'identité et de non-contradiction), et ceux qui requièrent une intuition tels que « la ligne droite est le plus court chemin entre deux points ». Il y a nécessairement de l'arbitraire dans la répartition de la logique et de l'intuition aux fondements de la géométrie. Vuillemin cite l'analyse de Hilbert qui démembre la notion de ligne droite en axiomes concernant la connexion des éléments (points, droites et plans) concernant le concept de segment, et de suite de points d'une droite, et l'axiome de continuité. De même rien n'est moins sûr que la thèse selon laquelle les géométries se réfèreraient en dernière instance à l'espace euclidien. T_2 ne tient pas non plus. Kant maintient comme seul exemple, dans la seconde édition de la *Critique*, le caractère tridimensionnel de l'espace. Ce problème est de nature topologique et il est vrai que Kant a pressenti le lien qui unit continuité et dimension. Mais on doit aller alors vers une conception bien plus primitive que celle d'un espace métrique, qui n'est plus sur le même plan que l'axiome des dimensions. L'affirmation selon laquelle (M) l'antériorité d'un tout indéfini par rapport à ses parties qu'il contient sans les désigner, qui fait de la représentation de l'espace une intuition, tombe sous les coups russelliens de la théorie des ensembles. La pensée d'un ensemble infini est légitime et requiert seulement la notion de correspondance biunivoque. Le dernier axiome (M') affirme le caractère *actuellement donné* de l'infinité de l'espace. Kant préfère l'infini à l'illimité pour des raisons qui ne sont pas géométriques, tirées des antinomies que l'on encourt dans le cas contraire. Et l'infinité n'est pas nécessairement liée à la géométrie euclidienne que promeut Kant. Il apparaît que l'intuition se trouve toujours obligée de recourir à des concepts et des arguments que Kant plaçait dans l'entendement.

7. Conclusion

Avec l'axiome des dimensions, du plus court chemin (géodésique) et celui de l'impossibilité de faire abstraction de l'espace, la théorie de Kant a anticipé, selon Vuillemin, les trois axiomes de Helmholtz : axiomes des dimensions, de la libre mobilité des figures et de l'invariance de la distance pour le groupe des déplacements. Mais la difficulté principale de la doctrine tient à son concept d'intuition. « Elle n'intervient pas nécessairement dans les démonstrations géométriques, elle n'appartient pas à la signification intrinsèque des axiomes et, si elle peut servir d'adjuvant pour la représentabilité des propositions, elle ne permet de distinguer Euclide qu'à la condition d'être complète et normale » (au sens général où les objets y intervenant, les fonctions sont continues ainsi que leurs dérivées premières) ([11], p. 414). Je voudrais terminer en nuançant la réponse de Vuillemin.

L'intuition sensible kantienne peut-elle se déplacer suffisamment pour conserver une effectivité ? Cette question reste interne à la philosophie. Car même si je fais des mathématiques, je suis amené à prendre des décisions métaphysiques. Elles ne font ici que se déplacer : l'intuition kantienne doit pouvoir se déplacer, et au prix d'un démembrement des facultés, continuer de jouer son rôle en remontant dans la théorie des groupes.

Si l'on considère quelques-uns des aspects de la Théorie des groupes que j'ai rappelés, on peut se demander comment y faire jouer des formes de l'intuition ? Nous avons affaire à une extraction d'opérations thématisées à un certain niveau d'abstraction. Ces opérations se composent, se neutralisent et supposent l'existence d'une opération neutralisée.

Dans tous ces cas, nous gardons à disposition des « tracés » représentés abstraits. Il y a, dans la déduction même, une trajectoire tracée. Il reste des formes de combinaisons qui respectent leur neutralisation possible. Je tiens qu'il y a là une forme nécessaire de réflexion, réflexivité. Mais le fait d'aller dans un sens, et de pouvoir revenir sur ses traces, fait appel à ce mouvement descriptif dont Kant a donné une présentation dans la Méthodologie. De même, trouver des opérations qui respectent ces contraintes (opérations de groupe) et qui laissent invariants les objets sur lesquels elles travaillent, permet de mettre en œuvre une autre forme intuitive correspondant à l'espace.

Il me semble que cette recherche possède une signification qui correspond à l'établissement de formes de stabilité dans l'espace, mais qui constituent ou reconstituent l'espace. C'est de cette manière qu'elles nous livrent une substantialité de l'espace mais, si je puis dire, une substantialité intrinsèque qui y puise une matière intuitive. Ce que Kant nous apprend même encore dans le travail algébrique et géométrique, c’est que l’intuition doit remonter dans le concept lui-même, qu’elle devient ainsi un support constituant du formalisme dans lequel elle remonte.

Bibliographie

[1] Helmholtz, H., *Über den Ursprung, Sinn und Bedeutung der geometrischen Sätze*, Wiss.AGh., Band II, p. 214–224, 1878, Trad. *On the Origin and Signifiance of Geometrical Axioms*, in the Work of Mathematics, J.R. Newman (Ed.), vol. I, Simon and Schuster, New York, 1956, p. 647–668.

[2] Kant, I., *Critique de la raison pure*, Trad. fr. Tremeseygues A., Pacaud B., PUF Paris 1944.

[3] Lie, S., Engel, F., *Theorie der Transformationsgruppen*, 3 vol., Teubner Leipzig (vol. 1, 1888, 2, 1890, 3, 1893).

[4] Poincaré, H., Les fondements de la Géométrie, *Bull. Sciences mathématiques,* 2^{e} série, tome 26, p. 249–272 (sept. 1902).

[5] Poincaré, H., *Science et hypothèse*, Flammarion, Paris, 1902.

[6] Poincaré, H., *Les fondements de la géométrie*, *in* Œuvres, 11 vol., Gauthier-Villars, 1916-1956.

[7] Riemann, B., *Gesammelte Mathematische Werke und Wissenschaflicher Nachlass*, herausgegeben unter Mitwirkung von R. Dedekind und H. Weber, zu Aufl. von H. Weber, Nachträge herausgegeben von M. Nœther und W. Wirtinger, Dover, New York, 1953.

[8] Riemann, B., *Œuvres Mathématiques*, Trad. L. Laugel, Gauthiers-Villars, Paris, 1808.

[9] Toretti, R., *Philosophy of Geometry from Riemann to Poincaré*, D. Reidel Publishing Company, Dordrecht Holland / Boston / USA London ; England, 1978.

[10] Vuillemin, J., *Physique et métaphysique kantienne,* PUF, Paris, 1955.

[11] Vuillemin, J., *La philosophie de l’algèbre*, PUF, 1962.

10 Fluctuations du vide quantique

Serge Reynaud, Astrid Lambrecht, Marc-Thierry Jaekel

1. Introduction

La question de la relativité du mouvement a joué un rôle essentiel dans la naissance de la physique moderne au XVII^{e} siècle. Galilée a souligné que le mouvement à vitesse uniforme ne pouvait être distingué du repos mais, également, que cette propriété n'était vraie que lorsque la résistance opposée par l'air au mouvement pouvait être ignorée. Dans les *Discorsi*, il discute longuement les effets de trainée dans l'air, phénomène essentiel dans toute expérience faisable à son époque. C'est seulement en considérant, par la pensée, le cas idéal où la trainée ne se fait plus sentir que Galilée parvient à définir l'espace physique où doivent être analysées les lois de la chute des corps [1]. Newton insiste encore sur cette problématique en montrant dans les *Principia* que le mouvement des planètes ne serait pas ce qu'il est si la force de trainée avait un effet sensible [2]. Ceci lui permet d'affirmer que l'espace interplanétaire est vide, dans le même sens que les enceintes vidées par pompage dans les expériences faites alors par Pascal, von Guericke ou Boyle dans la foulée des pionniers, Torricelli et, là encore, Galilée.

La physique moderne pose donc une problématique du vide qui reprend les notions logiques défendues par les atomistes anciens et critiquées par leurs contradicteurs [3]. Le mouvement existe et ce simple fait d'évidence nous invite à penser un espace dans lequel se déroule ce mouvement. Cet espace a pour propriété essentielle de ne pas s'opposer au mouvement, d'où la notion de vide. Mais ce raisonnement repose paradoxalement sur des propriétés négatives, ce qui pose la question du statut même du vide. Les débats engendrés par cette question sont nombreux et ils recoupent d'autres questions portant sur la nature de l'espace avec de multiples alternatives [4]. On connait bien la critique de l'espace absolu de Newton [5] et son aboutissement dans la théorie einsteinienne de la relativité [6, 7]. Parallèlement se déploie la critique d'une autre idéalisation classique de l'espace, tout aussi essentielle dans la physique newtonienne, celle d'un espace absolument vide (références dans [8]).

Après le développement de l'électrodynamique, et en particulier après les équations de Maxwell, l'espace classique est le cadre dans lequel on comprend non seulement les lois du mouvement mécanique mais aussi les lois de propagation du champ électromagnétique. Dans la physique classique, on peut considérer la limite idéale d'un espace absolument vide aussi bien de matière que de champ. Cette idéalisation est remise en cause par la découverte du rayonnement du corps noir, un phénomène essentiel dans le développement de la physique statistique classique [9]. Ce rayonnement électromagnétique, présent dans tout l'espace, a des effets observables par la pression qu'il exerce sur les parois réfléchissantes ou la force de friction qu'il oppose à une surface en mouvement.

C'est précisément pour expliquer ces propriétés que Planck propose en 1900 la première loi quantique [10]. Dans les termes utilisés aujourd'hui, qui ne sont pas exactement ceux de Planck [11], cette loi associe un nombre entier n de photons à tout mode du champ électromagnétique, c'est-à-dire à toute solution libre des équations de Maxwell. Utilisant la combinatoire de Boltzmann, Planck calcule l'entropie associée à la distribution de ce nombre et en déduit l'énergie moyenne $\overline{E}$ dans un mode de fréquence ω comme le produit d'une énergie élémentaire $\hbar\omega$ par un nombre moyen $\overline{n}$ de photons

$$\overline{E} = \overline{n}\hbar\omega \qquad\qquad \overline{n} = \frac{1}{\exp\left(\frac{\hbar\omega}{k_B T}\right) - 1}\cdot \tag{10.1}$$

Cette loi est valable à l'équilibre thermodynamique à une température T, k_B étant la constante de Boltzmann et $\hbar$ la constante de Planck ; $\overline{n}$ doit être compris comme la moyenne statistique d'une variable entière positive n ; cette variable présente aussi des fluctuations dont nous reparlerons plus loin.

Le nombre moyen $\overline{n}$ tend vers zéro à la limite d'une température nulle, quellle que soit la valeur de la fréquence. Ceci signifie qu'on peut encore donner une définition opératoire du vide en considérant une enceinte débarrassée de toute matière par pompage, puis du rayonnement du corps noir par refroidissement au zéro absolu de température. Mais cette définition ne résiste pas longtemps au développement des conceptions quantiques. Insatisfait de la dérivation de sa première loi, Planck reprend son travail en 1912 et il obtient une expression différente avec un terme supplémentaire [12]

$$\overline{E} = \left(\frac{1}{2} + \overline{n}\right)\hbar\omega. \tag{10.2}$$

La différence entre les deux lois de Planck est justement ce que nous appelons aujourd'hui les « fluctuations de point zéro » qui subsistent à la limite d'une température nulle.

Les raisonnements de Planck ne sont plus considérés comme corrects aujourd'hui [13]. Mais ses résultats ont été pris au sérieux par de nombreux physiciens dès 1912. Einstein et Stern notent dès 1913 que la deuxième loi, contrairement à la première, tend vers le résultat classique attendu à haute température [14]

$$\left(\frac{1}{2} + \overline{n}\right)\hbar\omega = k_{\mathrm{B}}T + O\left(\frac{1}{T}\right) \qquad\qquad T \to \infty. \tag{10.3}$$

Debye insiste dès 1914 sur de possibles conséquences observables des fluctuations de point zéro dans le mouvement atomique, en discutant leur effet sur les intensités des pics de diffraction [15] tandis que Mulliken a fourni en 1924 les premières preuves expérimentales de ces fluctuations en étudiant des spectres vibrationnels de molécules [16].

Signalons que Planck, comme la plupart des physiciens qui s'intéressaient à ces problèmes à cette époque, préférait attribuer les fluctuations quantiques aux oscillateurs matériels plutôt qu'au rayonnement. Une exception bien connue est celle d'Einstein affirmant dès 1905 la quantification de la lumière [17] et exploitant dès 1907 l'analogie entre oscillateurs matériels et rayonnement pour expliquer les chaleurs spécifiques « anormales » [18]. Apparemment, Nernst est le premier à avoir affirmé clairement que les fluctuations de point zéro existaient également pour les modes du champ électromagnétique, avec la conséquence inéluctable que l'espace classique absolument vide n'était plus atteignable, même à température nulle [19]. Alors que la première loi décrit une cavité entièrement vide à température nulle, la deuxième loi indique qu'il reste à cette limite des fluctuations de champ qui correspondent à la moitié de l'énergie d'un photon par mode. Nernst en déduisait d'ailleurs que la théorie quantique devait affronter une difficulté grave, que nous discuterons plus loin sous le nom de « catastrophe du vide ».

2. Le vide quantique

À ce point, nous pouvons souligner que toutes ces discussions ont eu lieu avant 1925, c'est-à-dire avant que le statut des fluctuations quantiques ne soit assuré par les premiers calculs quantiques entièrement cohérents [20]. Résumons brièvement ce statut dans les termes modernes de l'optique quantique [21]. Nous pouvons représenter le champ électromagnétique $\mathcal{E}$ dans n'importe quel mode comme une somme de deux composantes de quadrature, c'est-à-dire des composantes en cosinus et sinus de l'onde de fréquence ω

$$\mathcal{E} = \mathcal{E}_1 \cos \omega t + \mathcal{E}_2 \sin \omega t. \tag{10.4}$$

Les deux quadratures $\mathcal{E}_1$ et $\mathcal{E}_2$ obéissent à une inégalité de Heisenberg similaire à celle bien connue pour la position et l'impulsion d'un oscillateur matériel. Autrement dit, si on définit la variance Δx d'une variable statistique x par la relation habituelle

$$\Delta x \equiv \sqrt{\overline{x^2} - \overline{x}^2}, \tag{10.5}$$

alors les variances $\Delta\mathcal{E}_1$ et $\Delta\mathcal{E}_2$ des deux quadratures ont un produit nécessairement supérieur à une certaine constante $\mathcal{E}_0^2$:

$$\Delta\mathcal{E}_1 \Delta\mathcal{E}_2 \geq \mathcal{E}_0^2. \tag{10.6}$$

Cette constante $\mathcal{E}_0$ est calculable à partir des constantes fondamentales et de la fréquence ω et elle caractérise le niveau des fluctuations quantiques. En fait, l'énergie E dans le mode étant proportionnelle à $\mathcal{E}_1^2 + \mathcal{E}_2^2$, l'inégalité de Heisenberg implique que cette énergie est supérieure à un minimum

$$E \geq \frac{1}{2}\hbar\omega. \tag{10.7}$$

Le vide est alors défini comme l'état où l'énergie est minimale, ce qui suffit pour déterminer toutes les propriétés de cet état

$$E = \frac{1}{2}\hbar\omega \qquad \overline{\mathcal{E}}_1 = \overline{\mathcal{E}}_2 = 0 \qquad \Delta\mathcal{E}_1 = \Delta\mathcal{E}_2 = \mathcal{E}_0. \tag{10.8}$$

Les valeurs moyennes de $\mathcal{E}_1$ et $\mathcal{E}_2$ sont nulles et les variances sont égales entre elles tout en minimisant le produit qui apparaît dans l'inégalité (10.6).

Ce sont ces fluctuations que nous appelerons « fluctuations du vide » dans la suite. Du point de vue de la terminologie, il est clair que le « vide quantique » est ainsi défini comme une réalité positive forcément distincte du

« rien » ou du « néant ». L'espace est de manière permanente rempli de fluctuations électromagnétiques de champ qui se propagent à la vitesse de la lumière, comme n'importe quel champ libre. Elles correspondent en énergie à la moitié d'un photon par mode et sont aussi « réelles » que les photons. Il faut souligner toutefois que, puisque le vide quantique est l'état du champ où l'énergie est minimale, on ne peut utiliser cette énergie pour fabriquer un mouvement perpétuel qui violerait les lois de la thermodynamique.

Dans un tel point de vue, on ne se contente pas des interprétations habituelles des relations de Heisenberg qui insistent sur les limites que celles-ci imposent à la précision des mesures ou à la détermination des grandeurs. Cette insistance présente en effet le grave inconvénient de fixer les ambitions du physicien par référence à la théorie classique. Mais celle ci s'est révélée totalement insuffisante pour rendre compte de phénomènes physiques dont le rayonnement du corps noir n'est qu'un exemple parmi de nombreux autres. C'est pour cette raison que la théorie quantique a été développée et les difficultés d'interprétation qu'elle a suscitées ne doivent pas faire oublier les immenses succès qu'elle a remportés. Il semble aujourd'hui nécessaire d'en tirer toutes les leçons.

En ce qui concerne les inégalités de Heisenberg, le point de vue moderne est plutôt que ces inégalités impliquent l'existence de fluctuations qui sont des propriétés physiques intrinsèques des objets quantiques. En termes moins techniques, les relations de Heisenberg établissent une identité nécessaire entre la potentialité du mouvement et son actualité, sa réalité. L'espace permet le mouvement, c'est une de ses propriétés essentielles, et la mécanique quantique impose alors un mouvement minimal, le mouvement de point zéro. De même, l'espace permet la propagation du champ, c'est là encore une de ses propriétés fondamentales, et la mécanique quantique impose l'existence de fluctuations minimales du champ, les fluctuations du vide quantique.

Ces fluctuations sont présentes dans l'espace, même à température nulle, et elles ont des conséquences observables bien connues dans le monde microscopique [22]. Un atome couplé seulement aux champs du vide subit des processus d'émission spontanée induits par ces champs. Quand il est tombé dans son état fondamental, l'atome ne peut plus émettre de photons mais son couplage au vide a encore des effets mesurables comme l'effet Lamb de décalage des fréquences d'absorption. Les fluctuations jouent un rôle important en physique subatomique en donnant naissance aux corrections radiatives [23], mais également en physico-chimie des édifices moléculaires ou supra-moléculaires, par l'intermédiaire des forces de Van der Waals. C'est en travaillant sur la dépendance de ces forces vis-à-vis de la distance que Casimir découvre en 1948 la force qui porte aujourd'hui son nom [24] et dont nous reparlerons plus loin.

3. Le bruit de photons

Les fluctuations quantiques du champ électromagnétique sont directement étudiées dans certaines expériences d'optique quantique. L'idée générale est simplement que le bruit de photon, c'est-à-dire les fluctuations du nombre de photons, reflète les fluctuations du champ. Cette idée a été formulée la première fois par Einstein lors de ses efforts pour développer une théorie statistique de la lumière [25]. En partant de la formule de Planck, Einstein a montré que le carré de la variance Δn^2 associée au nombre n de photons dans un mode s'écrivait comme une somme de deux termes

$$\Delta n^2 = \overline{n}^2 + \overline{n}. \tag{10.9}$$

Einstein a alors expliqué le premier terme comme un terme « ondulatoire » induit par les fluctuations statistiques classiques du champ et le second comme un terme « corpusculaire » associé à la quantification de la lumière. La coexistence de ces deux termes dans la même formule lui a permis d'affirmer qu'un changement profond des vues sur la constitution de la lumière était nécessaire pour fusionner les points de vue ondulatoire et corpusculaire.

Nous disposons aujourd'hui de cette approche unifiée, la théorie quantique de la lumière. Elle permet effectivement de comprendre les fluctuations de photons à partir des fluctuations de champ à condition de prendre en compte le fait que celles-ci sont quantiques et non pas classiques. Pour illustrer ce point, considérons une expérience de pensée où un faisceau lumineux frappe une lame semi-réfléchissante (figure 10.1).

Cette expérience peut être analysée en termes « corpusculaires », chaque photon ayant une probabilité de 50 % d'être transmis ou refléchi. Le caractère aléatoire du processus de répartition est observé comme un bruit de photon derrière la lame : les nombres de photons dans les deux voies de sortie sont égaux en moyenne $\overline{n}_\mathrm{c} = \overline{n}_\mathrm{d}$ mais pas au niveau des fluctuations. Comme les probabilités sont indépendantes pour chaque photon, la différence $n = n_\mathrm{c} - n_\mathrm{d}$ a des fluctuations caractérisées par une loi de Poisson $\Delta n^2 = \overline{n}_\mathrm{a}$ où n_a est le nombre de photons entrés par la voie « a ». Si nous analysons maintenant la même expérience en termes « ondulatoires », la variable n apparaît déterminée

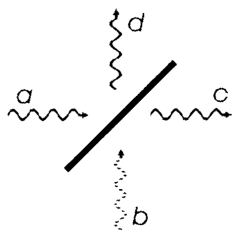


Fig. 10.1 – Description schématique d'une expérience simple d'optique quantique : un faisceau de photons arrive sur une lame séparatrice par la voie d'entrée « a » ; les fluctuations des nombres de photons sont analysées dans les deux voies de sortie « c » et « d » ; la statistique correspondante peut être comprise soit comme un bruit de répartition des photons sur la lame, soit comme le résultat des fluctuations du champ arrivant par la voie d'entrée « b » habituellement ignorée.

par les fluctuations du champ arrivant par la voie d'entrée « b » habituellement ignorée, et plus précisément par les fluctuations d'une composante de quadrature, disons $\mathcal{B}_1$, de ce champ [21]

$$\Delta n^2 = \overline{n}_{\mathrm{a}} \frac{\Delta \mathcal{B}_1^2}{\mathcal{E}_0^2}. \tag{10.10}$$

Habituellement, ce sont les fluctuations du vide qui entrent dans la voie « b », de sorte que $\Delta \mathcal{B}_1^2 = \mathcal{E}_0^2$, ce qui redonne la statistique de Poisson.

Ce raisonnement fournit donc une nouvelle interprétation du bruit de photons, qui révèle les fluctuations de champ. Mais il va au-delà de cette réinterprétation en montrant comment manipuler les statistiques de photons. En effet, l'inégalité de Heisenberg (10.6) implique l'existence des fluctuations quantiques, en fixant un minimum pour le produit des variances des deux quantités impliquées. Mais elle n'interdit pas de jouer avec ces fluctuations et de définir des états « comprimés » pour lesquels le bruit sur une des quadratures est plus petit que dans le vide, par exemple $\Delta \mathcal{E}_1^2 < \mathcal{E}_0^2$. Il faut alors que $\Delta \mathcal{E}_2^2$ soit assez grand pour que l'inégalité de Heisenberg soit respectée [21]. De nombreux dispositifs ont été étudiés en optique quantique avec l'objectif de fabriquer de tels états comprimés. Si un tel dispositif est interposé dans la voie d'entrée « b » dans l'expérience de pensée de la figure 10.1, la statistique de photons va être modifiée. En particulier, si on comprime les fluctuations de la composante de quadrature $\mathcal{B}_1$ responsable du bruit de répartition sur la lame, alors on peut obtenir un bruit sur la variable n inférieur au bruit de Poisson [26].

Ceci montre bien que le bruit de photon reflète les fluctuations des champs quantiques qui entrent dans le système optique et qu'il peut donc être contrôlé en agissant sur les composantes de quadrature impliquées. D'un point de vue fondamental, ceci signifie que le processus de répartition aléatoire des photons sur une lame est déterminé par les fluctuations du champ entrant dans la voie « b ». Comme ces fluctuations sont quantiques, ce mode de détermination ne peut pas se réduire à un schéma classique de déterminisme. Du point de vue des applications, cette idée peut être employée pour discuter la sensibilité des instruments optiques de mesure tels que les interféromètres conçus pour la détection des ondes gravitationnelles. Comme dans le cas simple discuté ci-dessus, l'analyse du bruit permet de proposer des techniques pour améliorer la sensibilité ultime de ces instruments [27, 28].

4. La catastrophe du vide

Nous en venons maintenant à la discussion d'une difficulté grave associée aux fluctuations du vide, difficulté qui avait été notée dès 1916 par Nernst [19]. Quand on calcule l'énergie du vide en ajoutant les énergies $\frac{1}{2}\hbar\omega$ de tous les modes de champ, on obtient une valeur infinie. Cette « catastrophe du vide » annule donc le succès obtenu en 1900 par Planck résolvant ce qui a été appelé depuis « la catastrophe ultraviolette ». En introduisant formellement une coupure à haute fréquence $\omega_{\max}$ sur le spectre des fluctuations du vide, on obtient pour la densité moyenne d'énergie

$$\overline{e} = \frac{\hbar\omega_{\max}^4}{8\pi^2c^3} + \frac{\hbar\theta^4}{160\pi^2c^3} \qquad\qquad \theta = \frac{2\pi k_B T}{\hbar}. \tag{10.11}$$

Le second terme, proportionnel à θ^4, où θ est la température mesurée comme une fréquence, est la densité d'énergie par unité de volume associée aux fluctuations thermiques dans la première loi de Planck. Cette expression est finie et correspond à la loi de Stefan-Boltzmann. Mais le premier terme, proportionnel à $\omega_{\max}^4$, c'est-à-dire la densité d'énergie du vide par unité de volume, diverge quand $\omega_{\max} \to \infty$. Ce problème n'est pas seulement formel, puisque la valeur calculée est énormément plus grande que l'énergie moyenne du vide observée dans le monde autour de nous par les phénomènes de la gravité, pour n'importe quel choix de la fréquence de coupure $\omega_{\max}$ qui préserve les lois de la théorie quantique aux énergies où elles sont bien vérifiées [29].

Cette catastrophe constitue depuis 1916 un conflit majeur entre la théorie classique de la relativité générale d'un côté et la théorie quantique de l'autre.

Il a souvent conduit à considérer que les fluctuations du vide quantique ne correspondaient pas vraiment à une énergie réelle. Cette opinion a par exemple été crûment énoncée par Pauli [30] : « À ce point il convient de noter qu'il est plus conséquent ici, contrairement au cas de l'oscillateur matériel, de ne pas attribuer une énergie de point zéro $\frac{1}{2}h\nu$ par degré de liberté. Parce que, d'une part celle-ci conduirait à une énergie infiniment grande par unité de volume, par suite du nombre infini de degrés de liberté, d'autre part elle serait essentiellement inobservable puisque elle ne peut ni être émise, absorbée ou dispersée ni par conséquent, contenue dans une enceinte et, comme il est évident par l'expérience, elle ne produit pas non plus de champ de gravité. »

Bien sûr, nous devons reconnaître avec Pauli que la valeur moyenne de l'énergie du vide ne contribue pas à la gravitation comme une énergie ordinaire. Ce point est tout simplement une évidence puisque l'univers autour de nous aurait une apparence très différente sinon. En d'autres termes, la référence pour le zéro d'énergie dans la théorie de la gravitation semble être finement accordée pour s'adapter à la valeur moyenne de l'énergie du vide. Cependant, ceci ne peut pas conduire à écarter tous les effets de l'énergie du vide. Même pour une compensation exacte de la contribution de l'énergie du vide, les différences d'énergie ou les fluctuations d'énergie doivent encore contribuer à la gravitation. De plus, il n'est plus question aujourd'hui de penser, comme le faisait Pauli, que l'énergie du vide ne peut « ni être émise, absorbée, dispersée... ni contenue dans une enceinte ». Nous l'avons déjà discuté, les fluctuations du vide électromagnétique ont des conséquences observables multiples et bien connues. De plus l'effet Casimir, discuté de manière plus détaillée ci-dessous, montre justement que le vide se manifeste par des effets physiques observables quand il est enfermé dans une enceinte, en l'occurence une cavité optique.

Le problème de l'énergie du vide est également connu comme « le problème de la constante cosmologique » en raison de son rapport direct avec l'introduction d'une constante cosmologique dans les équations de la gravitation d'Einstein [31–33]. Malgré des efforts considérables, il est resté non résolu pendant tout le vingtième siècle [34–36]. L'opinion la plus courante aujourd'hui est sans doute que ce problème est « mal posé » et que sa solution demandera des reformulations substantielles du formalisme théorique. Ceci ne veut pas dire qu'il faille le négliger en attendant. En fait, ce problème se trouve à un interface crucial entre théorie quantique et gravité, et peut de ce fait indiquer des pistes ou des indices permettant de faire avancer les questions difficiles qui subsistent à cet interface. Dans la suite de ce papier, nous présenterons un certain nombre d'indices qui correspondent à des questions bien définies avec des réponses adéquates dans le formalisme actuel et, simultanément, des conséquences observables.

5. La force de Casimir

Nous discutons maintenant la force de Casimir, c'est-à-dire la force mécanique exercée par les fluctuations du vide sur des miroirs formant une cavité optique [37]. Casimir a calculé cette force dans une configuration géométrique où deux miroirs plans sont parallèles, à une distance L l'un de l'autre, la surface A des miroirs étant beaucoup plus grande que le carré de la distance. Considérant le cas idéal de miroirs parfaitement réfléchissants, Casimir a obtenu les expressions suivantes pour la force F_{Cas} et l'énergie E_{Cas}

$$F_{\mathrm{Cas}} = \frac{\hbar c \pi^2 A}{240 L^4} \qquad E_{\mathrm{Cas}} = \frac{\hbar c \pi^2 A}{720 L^3} \qquad \left(A \gg L^2\right). \tag{10.12}$$

Cette force attractive a une amplitude faible – environ 0,1 μN pour $A = 1$ cm^2 et $L = 1$ μm – mais elle a été observée dans un certain nombre d'expériences historiques [24], avec une précision de l'ordre de 100 %. Elle a été récemment mesurée avec une précision nettement améliorée (références dans [38]), ce qui est important pour au moins deux raisons.

D'abord, la force de Casimir est la conséquence expérimentale la plus accessible dans le monde macroscopique des fluctuations du vide. En raison des difficultés évoquées ci-dessus, il est important d'examiner avec grand soin les prévisions que fait la théorie des champs quantiques à propos de l'énergie du vide. N'importe quelle définition pragmatique du vide implique une région de l'espace limitée par une enceinte et la force de Casimir n'est autre que la manifestation physique des fluctuations du vide enfermées dans cette cavité. Ces fluctuations étant modifiées par la cavité, l'énergie du vide dépend de la distance L entre les miroirs et il en résulte une force attractive. Dans le cas idéal de miroirs parfaitement réfléchissants dans le vide, la force dépend seulement de la distance et de deux constantes fondamentales, la vitesse de la lumière c et la constante de Planck $\hbar$. C'est une propriété universelle remarquable, en particulier parce que la force de Casimir est indépendante de la charge de l'électron, contrairement aux forces de Van der Waals. Autrement dit, la force de Casimir correspond à une réponse saturée des miroirs qui refléchissent au plus 100 % de la lumière. Cependant la plupart des expériences sont faites à température ambiante, avec des miroirs qui ne refléchissent pas parfaitement toutes les fréquences de champ, et ceci doit être pris en considération dans les évaluations théoriques [39, 40].

Ensuite, l'évaluation de la force de Casimir est un point crucial pour beaucoup de mesures très précises de force pour des distances entre le nanomètre et le millimètre. Ces expériences sont motivées par des tests de gravité newtonienne aux distances courtes [41] ou par des recherches de nouvelles faibles forces à courte portée prévues par les modèles théoriques d'unification (références dans [38]). Fondamentalement, elles visent à mettre des contraintes sur les déviations de la théorie standard actuelle par comparaison des résultats expérimentaux aux prédictions théoriques. La force de Casimir étant la force dominante entre deux objets non magnétiques neutres dans cette gamme de distance, il est important de tenir compte des différences entre le cas idéal considéré par Casimir et les situations réelles étudiées dans les expériences. Pour des comparaisons théorie-expérience de cette sorte, l'exactitude des calculs théoriques devient aussi cruciale que la précision des expériences [42].

En ce qui concerne l'évaluation de la force dans la configuration géométrique considérée par Casimir – force entre deux surfaces planes et parallèles à la limite d'une surface grande – on dispose désormais de méthodes théoriques fiables pour traiter les effets de réflexion imparfaite et de température non nulle. Dans ces méthodes, on décrit les miroirs par des amplitudes de diffusion qui dépendent de la fréquence tout en respectant des propriétés générales d'unitarité, de causalité et de transparence à haute fréquence. Avec ces hypothèses qui consistent à décrire un miroir réel comme on le fait en optique, la force de Casimir est simplement la différence de pression entre les côtés externe et interne des miroirs. L'expression obtenue ainsi est finie, contrairement à l'énergie du vide discutée plus haut, et elle redonne l'expression idéale de Casimir dans la limite des miroirs parfaits [43]. Pour les métaux par exemple, la force réelle dévie du résultat idéal (10.12) lorsque la distance L est plus courte que quelques longueurs plasma, typiquement $L \sim 0{,}3\ \mu$m pour l'or ou le cuivre [44]. La correction due à la température non nulle est obtenue en ajoutant la pression de rayonnement des fluctuations thermiques à celle des fluctuations du vide. La modification est importante pour les grandes distances, typiquement $L \sim 3\ \mu$m à température ambiante [45].

Ce genre d'évaluations théoriques devrait permettre une comparaison précise entre expérience et théorie, si d'autres déviations de la situation idéale sont également maîtrisées. Cette remarque concerne la géométrie – les expériences récentes sont faites avec une sphère et un plan alors que les calculs précis sont faits entre deux plans – et l'effet de rugosité des surfaces qui joue un rôle important pour les mesures à très courte distance. Il est clair que l'existence de la force de Casimir, son signe et sa grandeur, sont maintenant démontrés expérimentalement, à quelques % près pour la grandeur. Par conséquent, il n'est plus possible d'écarter les effets mécaniques des fluctuations du vide. Comme nous l'avons déjà dit, la question se pose alors de l'accord de ces effets avec les principes de la théorie de la relativité. Dans la suite, nous discutons plusieurs questions intéressantes de ce point de vue.

6. L'énergie de Casimir et le principe d'équivalence

La première de ces questions est directement reliée au principe d'équivalence d'Einstein. Nous l'avons déjà dit, alors que la valeur moyenne de l'énergie du vide ne contribue certainement pas à la gravité, par contre les différences d'énergie doivent contribuer. C'est en particulier le cas pour l'énergie de Casimir, une variation de l'énergie du vide avec la longueur de la cavité. Si le principe d'équivalence d'Einstein est vérifié, alors cette énergie doit également contribuer à l'inertie de la cavité de Fabry-Pérot. Bien sûr, cet effet aura une influence quantitativement faible sur le mouvement d'une cavité macroscopique. Néanmoins, il a une importance fondamentale en tant que version quantique, au niveau des fluctuations du vide, de l'argument d'Einstein sur l'inertie d'une boîte contenant un photon [46, 47]. Dans le cas de la cavité, qui est un corps composé soumis à une contrainte interne, la loi d'inertie de l'énergie doit être écrite sous la forme suivante [48] :

$$F_{\mathrm{mot}} = -\mu a \qquad \mu = \frac{E_{\mathrm{Cas}} - F_{\mathrm{Cas}} L}{c^2}. \tag{10.13}$$

Une analyse détaillée de cette question exige une évaluation des forces de Casimir quand les miroirs de la cavité bougent. Pour un mouvement global de la cavité avec une accélération uniforme, on trouve une force qui correspond exactement à la contribution inertielle de l'énergie de Casimir [49]. Ceci confirme que les variations de l'énergie du vide contribuent effectivement à la gravitation et à l'inertie comme le prévoient les principes généraux de la relativité.

En continuant selon les mêmes lignes, les fluctuations de l'énergie du vide doivent également contribuer à la gravitation. Ceci implique que les fluctuations quantiques des tenseurs d'énergie et des courbures d'espace-temps sont couplées entre elles [50–53]. Ce problème peut être traité par des techniques de réponse linéaire [54], par analogie avec l'étude des fluctuations du mouvement d'objets couplés à des forces fluctuantes [55] et il est ainsi lié à la question de la relativité du mouvement [56].

7. La relativité du mouvement dans le vide

Même un miroir seul et immobile dans le vide est soumis à une pression de rayonnement [57, 58]. La force résultante a une valeur moyenne nulle en raison d'une compensation entre les contributions des côtés opposés, mais elle présente des fluctuations puisque les champs arrivant des deux côtés ne sont pas corrélés. Quand le miroir est en mouvement, la compensation n'est plus exacte même pour la force moyenne, ce qui implique que le vide exerce une réaction au mouvement. Cette force dissipative est décrite par une susceptibilité qui permet d'exprimer la force F_{mot} en fonction du mouvement q, dans l'approximation linéaire et dans le domaine de Fourier,

$$F_{\mathrm{mot}}[\Omega] = \chi[\Omega] q[\Omega]. \tag{10.14}$$

La susceptibilité motionnelle $\chi[\Omega]$ est directement liée aux fluctuations de force évaluées pour un miroir au repos par des relations quantiques de type « fluctuation-dissipation ». La dissipation de l'énergie mécanique est associée à une émission de rayonnement et la force de réaction de rayonnement est juste le résultat de l'échange d'impulsion.

Considérons d'abord le cas simple d'un miroir se déplaçant à vitesse uniforme. Quand ce mouvement a lieu dans le vide, la force de réaction de rayonnement s'annule, de sorte que le mouvement inertiel ne peut pas être distingué du repos, en plein accord avec le principe de relativité du mouvement. Mais une force de frottement surgit quand un miroir se déplace dans un champ thermique, et elle est analogue à l'amortissement du mouvement par des molécules d'air [59, 60]. Avec les notations de cet article, cette force s'écrit, pour un miroir parfaitement réfléchissant dans les limites d'une grande surface plane et d'une grande température,

$$\chi[\Omega] \simeq \frac{i\hbar A\theta^4\Omega}{240\pi^2 c^4} \quad F_{\mathrm{mot}} \simeq \frac{\hbar A\theta^4}{240\pi^2 c^4} q'(t) \quad \left(A \gg \frac{c^2}{\Omega^2}\ ,\ \theta \gg \Omega\right). \tag{10.15}$$

Cette expression s'annule quand θ vaut zéro, mais elle constitue seulement une approximation classique qui ignore les fluctuations du vide. Pour un miroir parfait avec un mouvement arbitraire dans le vide quantique électromagnétique, nous obtenons une susceptibilité proportionnelle à la cinquième puissance de la fréquence soit, de manière équivalente, une force dissipative proportionnelle à la dérivée cinquième de la position par rapport au temps

$$\chi[\Omega] \simeq \frac{i\hbar A\Omega^5}{60\pi^2 c^4} \quad F_{\text{mot}} \simeq -\frac{\hbar A}{60\pi^2 c^4} q'''''(t) \quad \left(A \gg \frac{c^2}{\Omega^2}\ ,\ \theta = 0\right). \quad (10.16)$$

Ce résultat peut être deviné à partir du précédent par une simple analyse dimensionnelle : le facteur θ^4 apparaissant dans (10.15) est remplacé par Ω^4 dans (10.16). Cet argument jouait un rôle semblable dans la discussion de l'expression (10.11) de la densité d'énergie du vide avec cependant une différence importante : l'expression est maintenant parfaitement régulière pour le vide et elle correspond à un effet physique bien caractérisé du point de vue théorique et, au moins du point de vue des principes, observable.

Cet effet a été discuté pour la première fois en 1976 par Fulling et Davies dans le cas simple d'un miroir parfaitement réfléchissant se déplaçant dans le vide d'un champ scalaire dans un espace-temps à 2 dimensions – c'est-à-dire pour un espace à une dimension [61, 62]. Dans ce cas, la force dissipative est proportionnelle à la dérivée troisième de la position par rapport au temps q''', puisque $\frac{A}{c^2}\Omega^5$ est remplacé par Ω^3 dans l'argument dimensionnel précédent. La force de réaction du vide a ainsi la même forme que pour un électron dans le vide électromagnétique et elle soulève les mêmes problèmes de causalité et de stabilité [63]. Cette difficulté est cependant résolue pour n'importe quel miroir réel en remarquant que celui-ci est certainement transparent aux hautes fréquences, ce qui mène à un traitement satisfaisant du mouvement dans le vide [64]. En outre, l'équilibre stable atteint par le miroir couplé aux fluctuations de pression de rayonnement du vide contient une description cohérente des fluctuations quantiques du miroir, qui généralise l'équation de Schrödinger habituelle [65].

La force motionnelle ne soulève pas d'objection vis-à-vis du principe de relativité restreinte. En effet, la réaction du vide (10.16) s'annule dans le cas particulier d'une vitesse uniforme. Le formalisme quantique donne une interprétation intéressante de cette propriété : les fluctuations du vide apparaissent exactement identiques à un observateur inertiel et à un observateur au repos. Par conséquent l'invariance du vide sous les transformations de Lorentz est une propriété essentielle pour la validité du principe de relativité du mouvement, et elle établit une relation précise entre ce principe et les symétries du vide. Plus généralement, le vide ne s'oppose pas aux mouvements uniformément accélérés et cette propriété correspond à la symétrie conforme du vide quantique [66].

Il serait extrêmement intéressant de mettre en évidence expérimentalement les effets dissipatifs liés au mouvement dans le vide [67]. Ces effets sont excessivement petits pour n'importe quel mouvement qui pourrait être réalisé dans la pratique pour un miroir unique. Cependant, une observation expérimentale est imaginable avec une cavité oscillant dans le vide. Dans ce cas, l'émission de rayonnement induit par le mouvement est amplifiée de manière résonnante [68]. Des signatures spécifiques sont disponibles pour distinguer ce rayonnement des effets parasites [69], de sorte qu'une démonstration expérimentale semble être réalisable avec des cavités de très haute finesse [70].

8. Discussion

Les résultats évoqués ci-dessus démontrent que la théorie quantique a permis de mettre en place un cadre conceptuel tout à fait nouveau où les questions soulevées dans l'introduction peuvent maintenant trouver des réponses satisfaisantes.

L'espace dans lequel le mouvement a lieu ne peut plus être considéré comme absolument vide puisque des fluctuations de champ quantique y sont toujours présentes. Ces fluctuations provoquent des effets observables. En particulier, elles provoquent des effets mécaniques dissipatifs dans le cas du mouvement dans le vide d'un objet diffusant. Ces effets constituent un défi pour le principe de relativité du mouvement dans son acception la plus générale : le mouvement produit des effets observables, en l'occurence la résistance du vide au mouvement et l'émission de rayonnement par le miroir mobile, bien qu'il n'y ait aucune autre référence pour ce mouvement que les fluctuations du vide elles-mêmes. Ce défi ne constitue pas en lui-même une contradiction : en effet, la relativité générale est une théorie classique et il est tout à fait normal qu'elle ne rende pas compte d'effets quantiques liés aux fluctuations du vide. Toutefois, il signale une évolution importante du cadre conceptuel.

Les fluctuations du vide quantique sont une réalité positive parfaitement caractérisée du point de vue théorique. Quand un objet bouge dans l'espace, il bouge au moins par rapport à ces fluctuations qui ne peuvent être retirées de l'espace, alors que la matière peut en principe en être retirée, du moins par la pensée quand on considère une région délimitée de l'espace. Et le mouvement par rapport à ces fluctuations a des conséquences physiques observables. Comme ces effets disparaissent pour une grande catégorie de mouvements spécifiques, y compris les cas de vitesse uniforme et d’accélération uniforme, les fluctuations du vide quantique définissent une classe de systèmes de référence privilégiés pour la description des mouvements mécaniques.

Dans ce contexte, la catastrophe liée à l’énergie du vide est un paradoxe central, un point nodal des réflexions à l’interface entre théorie quantique d’une part, mouvement, inertie et gravité d’autre part. Il est probable que cette catastrophe ne fasse que mettre en évidence une question mal posée dans le formalisme actuel, mais tout aussi vraisemblable qu’elle constitue un signal qui, bien interprété, puisse nous suggérer des solutions aux difficiles problèmes qui ne sont pas encore résolus.

Bibliographie

[1] Galileo, G., *Discorsi e Dimostrazioni Matematiche intorno a Due Nuove Scienze*, 1638.

[2] Newton, I., *Philosophiae Naturalis Principia Mathematica*, 1687.

[3] Russell, B., *History of Western Philosophy*, George Allen & Unwin, 1961.

[4] Vilain, C., *Acta Cosmologica*, XXIV (1998) 153.

[5] Koyré, A., *From the Closed World to the Open Universe*, John Hopkins Press, 1957.

[6] Balibar, F., *Galilée, Newton lus par Einstein : Espace et relativité*, Presses Universitaires de France, 1984.

[7] Balibar, F., *Einstein 1905 : De l’éther aux quanta*, Presses Universitaires de France, 1992.

[8] Gunzig, E., Diner, S. (Eds), *Le Vide*, Revue de l’Université de Bruxelles et Éditions Complexe, 1997.

[9] Wolf, E., *Opt. News*, Winter issue (1979) 24.

[10] Planck, M., *Verh. Deutsch. Phys. Ges.* 2 (1900) 237.

[11] Darrigol, O., *Ann. Physik* 9 (2000) 951.

[12] Planck, M., *Verh. Deutsch. Phys. Ges.* 13 (1911) 138.

[13] Milonni, P.W., Shih, M.-L., *Am J. Phys.* 59 (1991) 684.

[14] Einstein, A., Stern, O., *Ann. Physik* 40 (1913) 551.

[15] Debye, P., *Ann. Physik* 43 (1914) 49.

[16] Mulliken, R.S., *Nature* 114 (1924) 349.

[17] Einstein, A., *Ann. Physik* 17 (1905) 132.

[18] Einstein, A., *Ann. Physik* 22 (1907) 180.
[19] Nernst, W., *Verh. Deutsch. Phys. Ges.* 18 (1916) 83.
[20] Dirac, P.A.M., *The Principles of Quantum Mechanics*, Oxford University Press, 1958.
[21] Reynaud, S., *Ann. Physique* 15 (1990) 63.
[22] Cohen-Tannoudji, C., Dupont-Roc, J., Grynberg, G., *Processus d'interaction entre photons et atomes*, InterÉditions, 1988 (Atom-Photon Interactions, Wiley, 1992).
[23] Itzykson, C., Zuber, J.-B., *Quantum Field Theory*, McGraw Hill, 1985.
[24] Sparnaay, M.J., *in* Sarlemijn A., Sparnaay M.J. (Éds), *Physics in the Making*, North-Holland, 1989, p. 235.
[25] Einstein, A., *Phys. Z.* 10 (1909) 185.
[26] Reynaud, S., Heidmann, A., Giacobino, E., Fabre, C., *in* Wolf E. (Ed), *Progress in Optics* XXX, North Holland, 1992, p. 1 et références.
[27] Reynaud, S., Heidmann, A., *Ann. Physique* 10 (1985) 227.
[28] Jaekel, M.T., Reynaud, S., *EuroPhys. Lett.* 13 (1990) 301, et références.
[29] Adler, R.J., Casey, B., Jacob, O.C., *Am. J. Phys.* 63 (1995) 620.
[30] Pauli, W., Die Allgemeinen Prinzipien der Wellenmechanik, *in* Geiger H. and Scheel K. (Eds), *Handbuch der Physik*, Springer, 1933, vol. 24, p.1.
[31] Abbott, L., *Sci. Am.*, 258 (1988) n.5 p. 106.
[32] Weinberg, S., *Rev. Mod. Phys.* 61 (1989) 1.
[33] Adler, R.J., in [8] (1997) p. 321.
[34] Demianski, M., *Ann. Physik* 9 (2000) 278.
[35] Weinberg, S., *Lecture at Dark Matter*, 2000 [arXiv astro-ph/0005265].
[36] Witten, E., *Lecture at Dark Matter* 2000 [arXiv hep-ph/0002297].
[37] Casimir, H.B.G., *Proc. K. Ned. Akad. Wet.* 51 (1948) 793.
[38] Reynaud, S., Lambrecht, A., Genet, C., Jaekel, M.T., *C. R. Acad. Sci. Paris*, 2-IV (2001) 1287 [arXiv quant-ph/0105053] et références.
[39] Lifshitz, E.M., *Sov. Phys. JETP* 2 (1956) 73.
[40] Schwinger, J., de Raad, L.L. Jr., Milton, K.A., *Ann. Physics* 115 (1978) 1.
[41] Fischbach, E., Talmadge, C., *The Search for Non Newtonian Gravity*, AIP Press/Springer Verlag, 1998.
[42] Lambrecht, A., Reynaud, S., *Phys. Rev. Lett.* 84 (2000) 5672.
[43] Jaekel, M.T., Reynaud, S., *J. Physique* I-1 (1991) 1395.
[44] Lambrecht, A., Reynaud, S., *Eur. Phys. J.* D8 (2000) 309.
[45] Genet, C., Lambrecht, A., Reynaud, S., *Phys. Rev.* A62 (2000) 0121110.
[46] Einstein, A., *Ann. Physik* 18 (1905) 639.
[47] Einstein, A., *Ann. Physik* 20 (1906) 627.

[48] Einstein, A., *Jahrb. Radioakt. Elektron.* 4 (1907) 411 and 5 (1908) 98.
[49] Jaekel, M.T., Reynaud, S., *J. Physique* I-3 (1993) 1093 et références.
[50] De Witt, B.S., *J. Math. Phys.* 3 (1962) 619.
[51] Feynman, R.P., *Acta Phys. Polonica* 24 (1963) 711.
[52] Weinberg, S., *Phys. Rev.* 138 (1965) B988.
[53] Zeldovich, Ya.B., Grishchuk, L.P., *Sov. Phys. Usp.* 29 (1986) 780.
[54] Jaekel, M.-T., Reynaud, S., *Ann. Physik* 4 (1995) 68, et références.
[55] Jaekel M.-T., Reynaud S., *Rep. Progr. Phys.* 60 (1997) 863, et références.
[56] Jaekel, M.-T., Lambrecht, A., Reynaud, S., *in* Gunzig E. and Diner S. (Eds), *Vacuum*, Plenum (in press) [arXiv quant-ph/9801071], et références.
[57] Jaekel, M.T., Reynaud, S., *Quantum Opt.* 4 (1992) 39.
[58] Barton, G., *in* Berman P. (Ed), Cavity Quantum Electrodynamics, Supplement, *Advances in Atomic, Molecular and Optical Physics*, Academic Press, 1994 et références.
[59] Einstein, A., *Ann. Physik* 17 (1905) 549.
[60] Einstein, A., *Phys. Z.* 18 (1917) 121.
[61] Fulling, S.A., Davies, P.C.W., *Proc. R. Soc.* A348 (1976) 393.
[62] Birrell, N.D., Davies, P.C.W., *Quantum Fields in Curved Space*, Cambridge, 1982.
[63] Rohrlich, F., *Classical Charged Particles*, Addison Wesley, 1965.
[64] Jaekel, M.T., Reynaud, S., *Phys. Lett.* A167 (1992) 227.
[65] Jaekel, M.T., Reynaud, S., *J. Physique* I-3 (1993) 1.
[66] Jaekel, M.T., Reynaud, S., *Quant. Semiclassical Opt.* 7 (1995) 499.
[67] Davies, P., *Nature* 382 (1996) 761.
[68] Lambrecht, A., Jaekel, M.T., Reynaud, S., *Phys. Rev. Lett.* 77 (1996) 615.
[69] Lambrecht A., Jaekel M.T., Reynaud S., *Euro. Phys. J.* D3 (1998) 95.
[70] Jaekel, M.T., Lambrecht, A., Reynaud, S., *in* Ruffini R. (Ed), *Marcel Grossmann IX*, World Scientific (in press) [arXiv quant-ph/0105050].

11 Causalité et localisation en Mécanique Quantique Relativiste

André Heslot

1. Introduction

La bonne théorie quantique relativiste de l'électron est la théorie de Dirac, complétée par l'interprétation des états d'énergie négative (« théorie des trous », dans sa forme primitive).

Malheureusement, la raison historiquement avancée par Dirac comme justification de sa théorie (linéarisation en $\hat{\boldsymbol{p}}$ de l'équation des ondes de Klein-Gordon, afin qu'espace et temps soient traités de façon symétrique) ne tient pas (c'est une formulation primitive, en réalité incorrecte, des exigences d'invariance relativiste).

En l'absence de raisons théoriques précises, on a alors tendance à légitimer la théorie de Dirac par le succès de ses prévisions (Feynman, par exemple).

Cette attitude est peu satisfaisante. Aussi, le travail qui suit a pour but d'élucider les raisons profondes qui font de la théorie de Dirac une bonne théorie.

Plus précisément, on va essayer de trouver ces raisons du côté de l'exigence de causalité.

Pour simplifier, on ne traitera ici que d'une seule dimension d'espace (l'espace-temps aura donc $1 + 1$ dimensions). Il n'y aura donc pas de considération de spin. Mais, en même temps, il n'est évidemment pas du tout certain que nos résultats puissent s'étendre à la dimension trois.

2. Causalité et localisation : le problème

2.1. Espace des états d'une particule

On considère, en dimension 1+1 d'espace-temps, une particule élémentaire libre de masse $m \neq 0$ et d'énergie > 0 (ceci peut être défini précisément *via* les représentations du groupe de Poincaré en dimension 1+1. L'élémentarité se traduit par l'irréductibilité de la représentation).

L'espace des états est engendré par les kets $|\,\boldsymbol{p}\,\rangle$, $\boldsymbol{p} \in \mathbb{R}$, avec

$$\hat{\boldsymbol{p}}\,|\,\boldsymbol{p}\,\rangle = \boldsymbol{p}\,|\,\boldsymbol{p}\,\rangle, \qquad \hat{H}\,|\,\boldsymbol{p}\,\rangle = \mathcal{E}\,|\,\boldsymbol{p}\,\rangle, \tag{11.1}$$

$$\langle\,\boldsymbol{p}\;|\,\boldsymbol{p}'\,\rangle = \delta(\boldsymbol{p}-\boldsymbol{p}'), \tag{11.2}$$

avec $\mathcal{E} = \sqrt{\boldsymbol{p}^2c^2 + m^2c^4}$. On ne parlera pas ici du générateur $\hat{\boldsymbol{K}}$ des boosts.

Un état quelconque $|\,\psi\,\rangle = \int d\boldsymbol{p}\,\psi(\boldsymbol{p})\,|\,\boldsymbol{p}\,\rangle$ est décrit par $\psi(\boldsymbol{p})$, et on prouve aisément que

$$(\hat{\boldsymbol{p}}\psi)(\boldsymbol{p}) = \boldsymbol{p}\,\psi(\boldsymbol{p}), \qquad (\hat{H}\psi)(\boldsymbol{p}) = \mathcal{E}\psi(\boldsymbol{p}), \tag{11.3}$$

$$\langle\,\phi\;|\;\psi\,\rangle = \int d\boldsymbol{p}\,\phi^\star(\boldsymbol{p})\psi(\boldsymbol{p}). \tag{11.4}$$

Particule non libre : une expression différente de $\hat{H}$:
$\hat{H}_0 \to \hat{H} = \hat{H}_0 + \hat{H}_{\text{int}}$.

2.2. Observable de position

Pas d'espace-temps dans ce qui précède (sinon comme espace sur lequel agit le groupe de Poincaré). En particulier, pas de position.

Une position est pourtant nécessaire :

- notion usuelle de particule (petit objet localisable en un point),
- plus précisément, nécessité d'une position pour décrire ultérieurement les interactions.

On cherche à construire une position $\hat{\boldsymbol{p}}$. On impose $[\hat{\boldsymbol{x}}, \hat{\boldsymbol{p}}] = i\hbar$ (transformation correcte par translation). On n'impose rien d'autre (à noter : en particulier, on ne fait aucune hypothèse concernant la transformation de $\hat{\boldsymbol{x}}$ par les boosts $\hat{\boldsymbol{K}}$). Il est facile de vérifier que la forme la plus générale de $\hat{\boldsymbol{x}}$ est alors

$$(\hat{\boldsymbol{x}}\psi)(\boldsymbol{p}) = i\hbar\psi'(\boldsymbol{p}) + \lambda(\boldsymbol{p})\psi(\boldsymbol{p}), \tag{11.5}$$

avec $\lambda(\boldsymbol{p})$ réel, $\hat{\boldsymbol{x}}$ devant être autoadjoint.

En fait, en redéfinissant les $|\,\boldsymbol{p}\,\rangle$ ($|\,\boldsymbol{p}\,\rangle \to e^{i\chi(\boldsymbol{p})}\,|\,\boldsymbol{p}\,\rangle$), et donc les $\psi(\boldsymbol{p})$, on fait disparaître $\lambda(\boldsymbol{p})$, d'où :

$$(\hat{\boldsymbol{x}}\psi)(\boldsymbol{p}) = i\hbar\psi'(\boldsymbol{p}). \tag{11.6}$$

Propriétés *a priori* satisfaisantes de $\hat{\boldsymbol{x}}$:

$$\hat{\boldsymbol{v}} = \frac{1}{i\hbar}[\hat{\boldsymbol{x}}, \hat{H}] = \frac{\hat{\boldsymbol{p}}c^2}{\hat{H}}, \tag{11.7}$$

d'où $\hat{\boldsymbol{v}}^2 < c^2$. Inversement,

$$\hat{\boldsymbol{p}} = \frac{m\hat{\boldsymbol{v}}}{\sqrt{1-\hat{\boldsymbol{v}}^2/c^2}},\ \hat{H} = \frac{mc^2}{\sqrt{1-\hat{\boldsymbol{v}}^2/c^2}}. \tag{11.8}$$

La diagonalisation de $\hat{\boldsymbol{x}}$ conduit à une « représentation $\boldsymbol{x}$ » :
$\hat{\boldsymbol{x}}\,|\,\boldsymbol{x}\,\rangle = \boldsymbol{x}\,|\,\boldsymbol{x}\,\rangle$ a pour solution (« normalisée »)

$$|\,\boldsymbol{x}\,\rangle = \int d\boldsymbol{p}\,\frac{e^{-i\boldsymbol{px}/\hbar}}{\sqrt{2\pi\hbar}}\,|\,\boldsymbol{p}\,\rangle. \tag{11.9}$$

On peut alors écrire

$$|\,\psi\,\rangle = \int d\boldsymbol{x}\,\psi(\boldsymbol{x})\,|\,\boldsymbol{x}\,\rangle,\quad \langle\,\phi\,|\,\psi\,\rangle = \int d\boldsymbol{x}\,\phi^\star(\boldsymbol{x})\psi(\boldsymbol{x}), \tag{11.10}$$

avec l'interprétation probabiliste usuelle de $\psi(\boldsymbol{x})$. On retiendra les formules de Fourier établies avec la convention :

$$\psi(\boldsymbol{p}) = \frac{1}{\sqrt{2\pi\hbar}}\int d\boldsymbol{x}\,e^{-i\boldsymbol{px}/\hbar}\psi(\boldsymbol{x}),\quad \psi(\boldsymbol{x}) = \frac{1}{\sqrt{2\pi\hbar}}\int d\boldsymbol{p}\,e^{i\boldsymbol{px}/\hbar}\psi(\boldsymbol{p}). \tag{11.11}$$

2.3. Violation de la causalité

$\psi(\boldsymbol{x}, t)$: fonction d'onde en représentation $\boldsymbol{x}$ à l'instant t, sachant que, pour $t = 0$, la particule est avec certitude au point *0*, *i.e.*, $\psi(\boldsymbol{x}, 0) = \delta(\boldsymbol{x})$.

On passe en représentation $\boldsymbol{p}$, où $\hat{H}$ s'exprime plus simplement (pas de racine de dérivation !) : $\psi(\boldsymbol{x}, 0) = \delta(\boldsymbol{x}) \leftrightarrow \psi(\boldsymbol{p}, 0) = 1/\sqrt{2\pi\hbar}$. L'évolution de cette dernière se déduit aisément de

$$|\,\psi(t)\,\rangle = \exp\left(-i\hat{H}t/\hbar\right)|\,\psi(0)\,\rangle,$$

soit

$$\psi(\boldsymbol{p}, t) = \frac{1}{\sqrt{2\pi\hbar}}e^{-i\mathcal{E}t/\hbar}, \tag{11.12}$$

et donc

$$\psi(\boldsymbol{x}, t) = \frac{1}{\sqrt{2\pi\hbar}} \int d\boldsymbol{p}\, e^{i\boldsymbol{px}/\hbar} e^{-i\mathcal{E}t/\hbar}. \tag{11.13}$$

On peut écrire,

$$\psi(\boldsymbol{x}, t) = \frac{1}{\sqrt{2\pi}} \frac{\partial}{\partial t} \int d\boldsymbol{p}\, e^{i\boldsymbol{px}/\hbar} \frac{e^{-i\mathcal{E}t/\hbar}}{\mathcal{E}} \equiv \frac{1}{\sqrt{2\pi}} \frac{\partial}{\partial t} \Delta^+(\boldsymbol{x}, t). \tag{11.14}$$

L'intérêt en est que $\Delta^+(\boldsymbol{x}, t)$ est invariant par transformation de Lorentz (orthochrone propre), ce qui en simplifie le calcul. On trouve :

$$\Delta^+(\boldsymbol{x}, t) = \begin{cases} -\frac{\pi}{c} N_0\left(\frac{mc}{\hbar}\sqrt{c^2t^2 - \boldsymbol{x}^2}\right) - i\frac{\pi}{c}\epsilon(t) J_0\left(\frac{mc}{\hbar}\sqrt{c^2t^2 - \boldsymbol{x}^2}\right) \\ \qquad\qquad\qquad\qquad\qquad\qquad \text{pour } \boldsymbol{x}^2 < c^2t^2, \\ \frac{\epsilon}{c} K_0\left(\frac{mc}{\hbar}\sqrt{\boldsymbol{x}^2 - c^2t^2}\right) \ \text{pour } \boldsymbol{x}^2 > c^2t^2, \end{cases} \tag{11.15}$$

où N_0, J_0 et K_0 sont des fonctions de Bessel, et $\epsilon(t)$ = signe de t. Il en résulte :

$$\psi(\boldsymbol{x}, t) = \begin{cases} -\frac{mc^2}{2\hbar} \frac{|t|}{\sqrt{c^2t^2 - \boldsymbol{x}^2}} J_0'\left(\frac{mc}{\hbar}\sqrt{c^2t^2 - \boldsymbol{x}^2}\right) \\ -i\frac{mc^2}{2\hbar} \frac{t}{\sqrt{c^2t^2 - \boldsymbol{x}^2}} N_0'\left(\frac{mc}{\hbar}\sqrt{c^2t^2 - \boldsymbol{x}^2}\right) \quad \text{pour } \boldsymbol{x}^2 < c^2t^2, \\ -i\frac{mc^2}{\pi\hbar} \frac{t}{\sqrt{\boldsymbol{x}^2 - c^2t^2}} K_0'\left(\frac{mc}{\hbar}\sqrt{\boldsymbol{x}^2 - c^2t^2}\right) \ \text{pour } \boldsymbol{x}^2 > c^2t^2. \end{cases} \tag{11.16}$$

Sur la figure 11.1 sont représentées les parties réelle et imaginaire de $\psi(\boldsymbol{x}, t)$. On observe que $\mathfrak{Im}\,\psi(\boldsymbol{x}, t)$ est non nul pour $\boldsymbol{x}^2 > c^2t^2$. Il y a donc *violation de la causalité*, et l'on peut conclure par le :

Théorème 11.1. *La mécanique quantique d'une particule élémentaire relativiste et localisable viole la causalité.*

3. Construction d'une théorie avec position causale

3.1. Nouvel espace des états

Idée n° 1 : on veut toujours une théorie quantique relativiste, une particule localisable (raison déjà exposée : décrire les intéractions). Il faut donc renoncer au caractère élémentaire, ou, tout au moins, à la description de

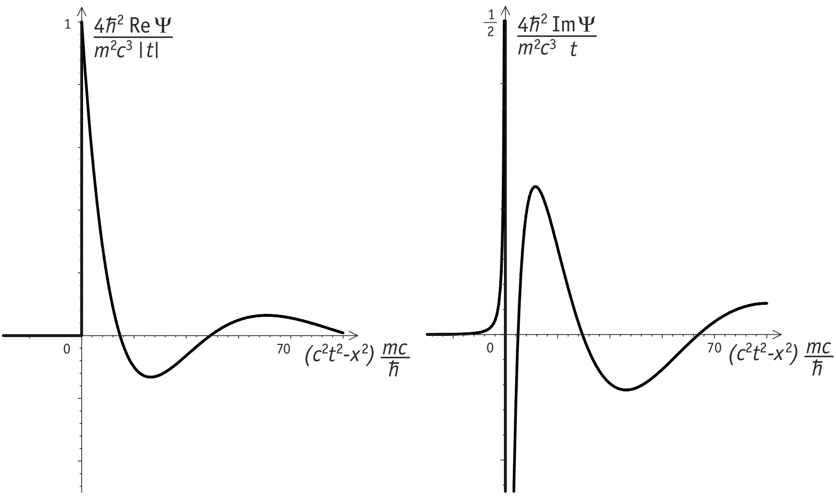


FIG. 11.1 – Parties rélles et imaginaires de $\psi(\boldsymbol{x}, t)$.

ce caractère élémentaire par une représentation *irréductible* du groupe de Poincaré.

Idée nº 2 : la violation de la causalité ne se manifeste que sur $\Im\mathrm{m}\, \psi(\boldsymbol{x}, t)$. Cela suggère d'essayer de construire une théorie dans laquelle $\psi(\boldsymbol{x}, t)$ serait remplacée par $\Re\mathrm{e}\, \psi(\boldsymbol{x}, t)$, *i.e.*,

$$\begin{aligned}
&\frac{1}{\sqrt{2\pi\hbar}} \int d\boldsymbol{p}\, e^{i\boldsymbol{px}/\hbar} e^{-i\mathcal{E}t/\hbar} \rightarrow \\
&\frac{1}{\sqrt{2\pi\hbar}} \int d\boldsymbol{p}\, \frac{e^{i\boldsymbol{px}/\hbar} e^{-i\mathcal{E}t/\hbar} + e^{-i\boldsymbol{px}/\hbar} e^{+i\mathcal{E}t/\hbar}}{2} \\
= &\frac{1}{\sqrt{2\pi\hbar}} \int d\boldsymbol{p}\, e^{i\boldsymbol{px}/\hbar} \frac{e^{-i\mathcal{E}t/\hbar} + e^{+i\mathcal{E}t/\hbar}}{2},
\end{aligned} \tag{11.17}$$

ce qui revient à considérer, à côté des états d'énergie positive $\mathcal{E}$, des états d'énergie négative $-\mathcal{E}$.

Développons précisément cette idée : pour décrire une particule de masse m, on considèrera une représentation réductible, somme directe d'une représentation irréductible de masse m à énergie positive et de la représentation iréductible de même masse m et d'énergie négative. En comparaison de (11.1), les états de base sont maintenant doublés :

$$\begin{aligned}
&\{|\,\boldsymbol{p},\epsilon\,\rangle\},\ \boldsymbol{p}\in\mathbb{R}, \qquad \epsilon=\pm= \text{ signe de l'énergie},\\
&\hat{\boldsymbol{p}}\,|\,\boldsymbol{p},\epsilon\,\rangle=\boldsymbol{p}\,|\,\boldsymbol{p},\epsilon\,\rangle, \qquad \hat{H}\,|\,\boldsymbol{p},\epsilon\,\rangle=\epsilon\mathcal{E}\,|\,\boldsymbol{p},\epsilon\,\rangle,\\
&\langle\,\boldsymbol{p},\epsilon\ |\,\boldsymbol{p}',\epsilon'\,\rangle=\delta_{\epsilon\epsilon'}\,\delta(\boldsymbol{p}-\boldsymbol{p}'),
\end{aligned} \tag{11.18}$$

(on ne dira rien du générateur $\hat{\boldsymbol{K}}$ des boosts).

Un état quelconque s'écrit :

$$|\ \psi\ \rangle=\sum_{\epsilon}\int d\boldsymbol{p}\ \psi_{\epsilon}(\boldsymbol{p})\,|\,\boldsymbol{p},\epsilon\,\rangle. \tag{11.19}$$

Il est donc décrit par le couple $\begin{pmatrix}\psi_+(\boldsymbol{p})\\ \psi_-(\boldsymbol{p})\end{pmatrix}$ (une fonction d'onde à deux composantes). On prouve facilement

$$\hat{\boldsymbol{p}}\begin{pmatrix}\psi_+(\boldsymbol{p})\\ \psi_-(\boldsymbol{p})\end{pmatrix}=\begin{pmatrix}\boldsymbol{p}\,\psi_+(\boldsymbol{p})\\ \boldsymbol{p}\,\psi_-(\boldsymbol{p})\end{pmatrix}, \quad \hat{H}\begin{pmatrix}\psi_+(\boldsymbol{p})\\ \psi_-(\boldsymbol{p})\end{pmatrix}=\begin{pmatrix}\mathcal{E}\psi_+(\boldsymbol{p})\\ -\mathcal{E}\psi_-(\boldsymbol{p})\end{pmatrix}, \tag{11.20}$$

et le produit scalaire de deux états est donné par

$$\langle\,\phi\ |\ \psi\,\rangle=\int d\boldsymbol{p}\,(\phi_+^{\star}(\boldsymbol{p})\,\psi_+(\boldsymbol{p})+\phi_-^{\star}(\boldsymbol{p})\,\psi_-(\boldsymbol{p})). \tag{11.21}$$

3.2. Construction d'une position

On impose à la position $\hat{\boldsymbol{x}}$ que l'on cherche à construire la relation $[\hat{\boldsymbol{x}},\hat{\boldsymbol{p}}]=i\hbar$, avec les mêmes justifications qu'auparavant. Il est facile de vérifier que la forme la plus générale de $\hat{\boldsymbol{x}}$ est alors

$$\hat{\boldsymbol{x}}\begin{pmatrix}\psi_+(\boldsymbol{p})\\ \psi_-(\boldsymbol{p})\end{pmatrix}=\begin{pmatrix}i\hbar\psi_+'(\boldsymbol{p})+\lambda(\boldsymbol{p})\,\psi_+(\boldsymbol{p})+\alpha(\boldsymbol{p})\,\psi_-(\boldsymbol{p})\\ i\hbar\psi_-'(\boldsymbol{p})+\mu(\boldsymbol{p})\,\psi_-(\boldsymbol{p})+\alpha^{\star}(\boldsymbol{p})\,\psi_+(\boldsymbol{p})\end{pmatrix}, \tag{11.22}$$

avec $\lambda(\boldsymbol{p})$ et $\mu(\boldsymbol{p})$ réels, $\hat{\boldsymbol{x}}$ devant être autoadjoint.

En fait, en redéfinissant les $|\,\boldsymbol{p},\epsilon\,\rangle$ ($|\,\boldsymbol{p},\epsilon\,\rangle\to\exp(i\chi_{\epsilon}(\boldsymbol{p})\,|\,\boldsymbol{p},\epsilon\,\rangle)$), on peut faire disparaître $\lambda(\boldsymbol{p})$ ou $\mu(\boldsymbol{p})$, d'où

$$\hat{\boldsymbol{x}}\begin{pmatrix}\psi_+(\boldsymbol{p})\\ \psi_-(\boldsymbol{p})\end{pmatrix}=\begin{pmatrix}i\hbar\psi_+'(\boldsymbol{p})+\alpha(\boldsymbol{p})\,\psi_-(\boldsymbol{p})\\ i\hbar\psi_-'(\boldsymbol{p})+\alpha^{\star}(\boldsymbol{p})\,\psi_+(\boldsymbol{p})\end{pmatrix}. \tag{11.23}$$

Par contre, le $\alpha(\boldsymbol{p})$ ne peut être modifié, sauf d'un facteur de phase constant.

Le problème maintenant est de trouver un $\alpha(\boldsymbol{p})$ tel que le $\hat{\boldsymbol{x}}$ correspondant soit causal. Bien sûr, $\alpha(\boldsymbol{p})=0$ est exclu, sinon on aurait simplement en deux exemplaires (en >0 et en <0) le système de la section précédente. On va voir qu'il existe un choix qui semble assez naturel (mais on ne peut dire que c'est le seul).

Calculons l'opérateur vitesse correspondant

$$\hat{\boldsymbol{v}} = \frac{1}{i\hbar}[\hat{\boldsymbol{x}}, \hat{H}],$$

$$\hat{\boldsymbol{v}} \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} = \begin{pmatrix} \dfrac{\boldsymbol{p}c^2}{\mathcal{E}}\psi_+ + \dfrac{2i\alpha}{\hbar}\mathcal{E}\psi_- \\ -\dfrac{\boldsymbol{p}c^2}{\mathcal{E}}\psi_- - \dfrac{2i\alpha^\star}{\hbar}\mathcal{E}\psi_+ \end{pmatrix}. \tag{11.24}$$

On a

$$i\hbar[\hat{\boldsymbol{v}}, \hat{\boldsymbol{p}}] = [[\hat{\boldsymbol{x}}, \hat{H}], \hat{\boldsymbol{p}}] = [[\hat{\boldsymbol{p}}, \hat{H}], \hat{\boldsymbol{x}}] + [[\hat{\boldsymbol{x}}, \hat{\boldsymbol{p}}], \hat{H}] = [0, \hat{\boldsymbol{x}}] + [i\hbar, \hat{H}] = 0.$$

Par conséquent, $\hat{\boldsymbol{v}}$ et $\hat{\boldsymbol{p}}$ commutent et peuvent donc être diagonalisés simultanément. Or, à $\boldsymbol{p}$ donné, les valeurs possibles de $\hat{\boldsymbol{v}}$ sont les valeurs propres de la matrice

$$\begin{pmatrix} \dfrac{\boldsymbol{p}c^2}{\mathcal{E}} & \dfrac{2i\alpha}{\hbar}\mathcal{E} \\ -\dfrac{2i\alpha^\star}{\hbar}\mathcal{E} & -\dfrac{\boldsymbol{p}c^2}{\sqrt{\mathcal{E}}} \end{pmatrix},$$

soit

$$\boldsymbol{v} = \pm\sqrt{\frac{\boldsymbol{p}^2c^4}{\mathcal{E}^2} + \frac{4|\alpha|^2}{\hbar^2}\mathcal{E}^2}. \tag{11.25}$$

Les deux signes viennent du degré de liberté correspondant aux deux signes possibles pour l'énergie.

On a donc, à un signe près, une relation $\boldsymbol{v} = \boldsymbol{v}(\boldsymbol{p})$. Quelle peut-elle être ? L'idée générale est d'invoquer des arguments de relativité (covariance). Ceci n'est pas simple à formuler très précisément (problème des commutateurs avec $\hat{\boldsymbol{K}}$), mais deux solutions sont évidentes :

- $\boldsymbol{v} = \pm\frac{\boldsymbol{p}c^2}{\mathcal{E}}$ correspondant à $\alpha = 0$. Elle doit être exclue car on sait que cela fournira un $\boldsymbol{x}$ non-causal.
- $\boldsymbol{v} = \pm c$! Ce choix élimine tout problème de covariance.

Étudions cette deuxième solution. Elle correspond à

$$\alpha(\boldsymbol{p}) = \frac{\hbar}{2}\frac{mc^3}{\mathcal{E}^2} \times \text{facteur de phase.}$$

On va prendre le cas particulier où le facteur de phase est constant (on pourra alors le faire disparaître). Ainsi,

$$\alpha(\boldsymbol{p}) = \frac{\hbar}{2}\frac{mc^3}{\mathcal{E}^2} \tag{11.26}$$

i.e.,

$$\hat{\boldsymbol{x}}\begin{pmatrix}\psi_+(\boldsymbol{p})\\ \psi_-(\boldsymbol{p})\end{pmatrix}=\begin{pmatrix} i\hbar\psi_+'(\boldsymbol{p})+\dfrac{\hbar}{2}\dfrac{mc^3}{\mathcal{E}^2}\psi_-(\boldsymbol{p})\\ i\hbar\psi_-'(\boldsymbol{p})+\dfrac{\hbar}{2}\dfrac{mc^3}{\mathcal{E}^2}(\boldsymbol{p})\psi_-(\boldsymbol{p})\end{pmatrix}. \tag{11.27}$$

3.3. Preuve du caractère causal de $\hat{\boldsymbol{x}}$

On commence par déterminer les valeurs propres et vecteurs propres de $\hat{\boldsymbol{x}}$. Pour chaque $\boldsymbol{x}\in\mathbb{R}$, il y a deux vecteurs propres, $|\,\boldsymbol{x},1\,\rangle$ et $|\,\boldsymbol{x},2\,\rangle$, représentés par les fonctions d'onde respectives

$$\frac{e^{\frac{i}{2}\mathrm{Arctg}\frac{\boldsymbol{p}}{mc}}\,e^{-i\boldsymbol{p}\boldsymbol{x}/\hbar}}{\sqrt{2\pi\hbar}}\begin{pmatrix}\dfrac{1}{\sqrt{2}}\\ \dfrac{1}{\sqrt{2}}\end{pmatrix}\ \text{et}\ \frac{e^{-\frac{i}{2}\mathrm{Arctg}\frac{\boldsymbol{p}}{mc}}\,e^{-i\boldsymbol{p}\boldsymbol{x}/\hbar}}{\sqrt{2\pi\hbar}}\begin{pmatrix}\dfrac{1}{\sqrt{2}}\\ -\dfrac{1}{\sqrt{2}}\end{pmatrix}. \tag{11.28}$$

Les facteurs $\sqrt{2}$ et $1/\sqrt{2\pi\hbar}$ sont introduits pour avoir les normalisations usuelles du spectre continu :

$$\langle\,\boldsymbol{x},\alpha\ |\,\boldsymbol{x}',\alpha'\,\rangle=\delta_{\alpha\alpha'}\delta(\boldsymbol{x}-\boldsymbol{x}'). \tag{11.29}$$

Supposons la particule dans un état de position déterminée $\boldsymbol{x}=0$ à l'instant $t=0$. Cela signifie que

$$\begin{pmatrix}\psi_+\\ \psi_-\end{pmatrix}_{t=0}=A\frac{e^{\frac{i}{2}\mathrm{Arctg}\frac{\boldsymbol{p}}{mc}}}{\sqrt{2\pi\hbar}}\begin{pmatrix}\dfrac{1}{\sqrt{2}}\\ \dfrac{1}{\sqrt{2}}\end{pmatrix}+B\frac{e^{-\frac{i}{2}\mathrm{Arctg}\frac{\boldsymbol{p}}{mc}}}{\sqrt{2\pi\hbar}}\begin{pmatrix}\dfrac{1}{\sqrt{2}}\\ -\dfrac{1}{\sqrt{2}}\end{pmatrix},$$

d'où à un instant ultérieur, compte tenu de l'expression de $\hat{H}$,

$$\begin{pmatrix}\psi_+\\ \psi_-\end{pmatrix}=\begin{pmatrix}\left(\dfrac{A}{\sqrt{2\pi\hbar}}\dfrac{e^{\frac{i}{2}\mathrm{Arctg}\frac{\boldsymbol{p}}{mc}}}{\sqrt{2}}+\dfrac{B}{\sqrt{2\pi\hbar}}\dfrac{e^{-\frac{i}{2}\mathrm{Arctg}\frac{\boldsymbol{p}}{mc}}}{\sqrt{2}}\right)e^{-i\mathcal{E}t/\hbar}\\ \left(\dfrac{A}{\sqrt{2\pi\hbar}}\dfrac{e^{\frac{i}{2}\mathrm{Arctg}\frac{\boldsymbol{p}}{mc}}}{\sqrt{2}}-\dfrac{B}{\sqrt{2\pi\hbar}}\dfrac{e^{-\frac{i}{2}\mathrm{Arctg}\frac{\boldsymbol{p}}{mc}}}{\sqrt{2}}\right)e^{+i\mathcal{E}t/\hbar}\end{pmatrix}.$$

L'amplitude de probabilité de présence au point $\boldsymbol{x}$ est alors obtenue en faisant le produit scalaire de ce vecteur avec un état localisé au point $\boldsymbol{x}$, *i.e.*, avec

$$\begin{pmatrix}\left(\dfrac{ae^{\frac{i}{2}\mathrm{Arctg}\frac{\boldsymbol{p}}{mc}}+be^{-\frac{i}{2}\mathrm{Arctg}\frac{\boldsymbol{p}}{mc}}}{\sqrt{2}}\right)\dfrac{e^{-i\boldsymbol{p}\boldsymbol{x}/\hbar}}{\sqrt{2\pi\hbar}}\\ \left(\dfrac{ae^{\frac{i}{2}\mathrm{Arctg}\frac{\boldsymbol{p}}{mc}}-be^{-\frac{i}{2}\mathrm{Arctg}\frac{\boldsymbol{p}}{mc}}}{\sqrt{2}}\right)\dfrac{e^{-i\boldsymbol{p}\boldsymbol{x}/\hbar}}{\sqrt{2\pi\hbar}}\end{pmatrix}.$$

En utilisant

$$\left(e^{\frac{i}{2}\mathrm{Arctg}\frac{p}{mc}}\right)^2 = e^{i\mathrm{Arctg}\frac{p}{mc}} = \frac{mc^2 + i\boldsymbol{p}c}{\mathcal{E}},$$

on calcule que cette amplitude vaut

$$\begin{aligned}
&(a^\star A + b^\star B) \ \ \frac{1}{2\pi\hbar}\int d\boldsymbol{p}\, e^{i\boldsymbol{px}/\hbar}\ \left(\frac{e^{-i\mathcal{E}t/\hbar} + e^{i\mathcal{E}t/\hbar}}{2}\right)\\
&+(a^\star B + b^\star A) \ \ \frac{1}{2\pi\hbar}\int d\boldsymbol{p}\, \frac{mc^2}{\mathcal{E}}\, e^{i\boldsymbol{px}/\hbar}\ \left(\frac{e^{-i\mathcal{E}t/\hbar} - e^{i\mathcal{E}t/\hbar}}{2}\right)\\
&+(b^\star A - a^\star B) \ \ \frac{1}{2\pi\hbar}\int d\boldsymbol{p}\, \frac{i\boldsymbol{p}c}{\mathcal{E}}\, e^{i\boldsymbol{px}/\hbar}\ \left(\frac{e^{-i\mathcal{E}t/\hbar} - e^{i\mathcal{E}t/\hbar}}{2}\right),
\end{aligned}$$

ou encore,

$$\left[(a^\star A + b^\star B)\frac{i}{2\pi}\frac{\partial}{\partial t} + (a^\star B + b^\star A)\frac{mc^2}{2\pi\hbar} + (b^\star A - a^\star B)\frac{c}{2\pi}\frac{\partial}{\partial \boldsymbol{x}}\right]\Delta(\boldsymbol{x},t),$$

où

$$\begin{aligned}
\Delta(\boldsymbol{x},t) \quad &= \int d\boldsymbol{p}\, e^{i\boldsymbol{px}/\hbar}\ \left(\frac{e^{-i\mathcal{E}t/\hbar} - e^{i\mathcal{E}t/\hbar}}{2\mathcal{E}}\right) =\\
\Im\mathrm{m}\,\Delta^+(\boldsymbol{x},t) \ &= \begin{cases} -i\dfrac{\pi}{c}\epsilon(t)J_0\left(\dfrac{mc}{\hbar}\sqrt{c^2t^2-\boldsymbol{x}^2}\right) & \text{pour } \boldsymbol{x}^2 < c^2t^2,\\ 0 & \text{pour } \boldsymbol{x}^2 > c^2t^2. \end{cases}
\end{aligned} \tag{11.30}$$

L'amplitude est donc causale.

4. Interprétation et discussion

4.1. Propriétés de la position

Interprétation des états d'énergie négative : *via* une théorie à la Dirac, *i.e.*, ils sont complètement remplis dans l'état du vide. Un état à une particule est de ce fait un état à énergie positive, *i.e.* $\begin{pmatrix}\psi(\boldsymbol{p})\\ 0\end{pmatrix} \equiv |\,\psi_>\,\rangle$.

On a

$$\hat{\boldsymbol{x}}\begin{pmatrix}\psi(\boldsymbol{p})\\ 0\end{pmatrix} = \begin{pmatrix} i\hbar\psi'(\boldsymbol{p})\\ \dfrac{\hbar}{2}\dfrac{mc^3}{\mathcal{E}^2}\psi(\boldsymbol{p})\end{pmatrix}. \tag{11.31}$$

Donc, un état physique n'a pas de position déterminée. On note cependant que, pour un tel état,

$$\langle\,\psi_>\,|\,\hat{\boldsymbol{x}}\,|\,\psi_>\,\rangle = \int d\boldsymbol{p}\,\psi^\star(\boldsymbol{p})i\hbar\psi'(\boldsymbol{p}),$$

i.e.,

$$\langle \psi_> \mid \hat{\boldsymbol{x}} \mid \psi_> \rangle = \langle \psi_> \mid \hat{\boldsymbol{q}} \mid \psi_> \rangle, \text{ où } \hat{\boldsymbol{q}} \begin{pmatrix} \psi_+ \\ \psi_- \end{pmatrix} = \begin{pmatrix} i\hbar\psi'_+ \\ i\hbar\psi'_- \end{pmatrix}$$

et cela donne une bonne limite non-relativiste :

$$\lim_{c\to\infty} \hat{\boldsymbol{x}} = \hat{\boldsymbol{q}}.$$

Calculons l'incertitude sur un état d'énergie positive.

$$\hat{\boldsymbol{x}}^2 \begin{pmatrix} \psi(\boldsymbol{p}) \\ 0 \end{pmatrix} = \begin{pmatrix} -\hbar^2\psi''(\boldsymbol{p}) + \dfrac{\hbar^2}{4}\left(\dfrac{mc^3}{\mathcal{E}^2}\right)^2 \psi(\boldsymbol{p}) \\ \text{quelque chose} \end{pmatrix}.$$

Ceci implique

$$\langle \boldsymbol{x}^2 \rangle = -\hbar^2 \int d\boldsymbol{p}\, \psi^\star(\boldsymbol{p})\psi''(\boldsymbol{p}) + \int d\boldsymbol{p}\, \psi^\star(\boldsymbol{p})\psi(\boldsymbol{p})\,\frac{\hbar^2}{4}\left(\frac{mc^3}{\mathcal{E}^2}\right)^2 = \langle \boldsymbol{q}^2 \rangle + \cdots.$$

Comme $\langle \boldsymbol{x} \rangle = i\hbar \int d\boldsymbol{p}\, \psi^\star(\boldsymbol{p})\psi'(\boldsymbol{p}) = \langle \boldsymbol{q} \rangle$, on obtient

$$\begin{aligned}(\Delta \boldsymbol{x})^2 &= \langle \boldsymbol{x}^2 \rangle - \langle \boldsymbol{x} \rangle^2 \\ &= \langle \boldsymbol{q}^2 \rangle - \langle \boldsymbol{q} \rangle^2 + \int d\boldsymbol{p}\, \psi^\star(\boldsymbol{p})\psi(\boldsymbol{p})\,\frac{\hbar^2}{4}\left(\frac{mc^3}{\mathcal{E}^2}\right)^2 \\ &\geq \frac{\hbar^2}{4\Delta \boldsymbol{p}^2} + \int d\boldsymbol{p}\, |\psi(\boldsymbol{p})|^2\,\frac{\hbar^2}{4}\left(\frac{mc^3}{\mathcal{E}^2}\right)^2. \end{aligned} \tag{11.32}$$

Par rapport à la théorie non-relativiste, il y a un $\Delta\boldsymbol{x}$ supplémentaire. Par exemple, si $|\phi(\boldsymbol{p})|^2 \approx \delta(\boldsymbol{p})$ (état dont l'impulsion est voisine de 0), on a

$$\Delta \boldsymbol{x}_{\text{suppl}} \approx \frac{\hbar}{2mc}. \tag{11.33}$$

Une autre façon de présenter cela serait considérer le terme de Darwin dans une intéraction avec potentiel extérieur.

4.2. Propriété de la vitesse

Avec notre choix de $\alpha(\boldsymbol{p})$,

$$\hat{\boldsymbol{v}} \begin{pmatrix} \phi_+ \\ \phi_- \end{pmatrix} = \begin{pmatrix} \dfrac{\boldsymbol{p}c^2\phi_+ + imc^3\phi_-}{\mathcal{E}} \\ \dfrac{-\boldsymbol{p}c^2\phi_- - imc^3\phi_+}{\mathcal{E}} \end{pmatrix}, \tag{11.34}$$

et

$$\hat{\boldsymbol{v}}^2 \begin{pmatrix} \phi_+ \\ \phi_- \end{pmatrix} = c^2 \begin{pmatrix} \phi_+ \\ \phi_- \end{pmatrix}. \tag{11.35}$$

Les valeurs propres de $\hat{\boldsymbol{v}}$ sont c et $-c$. Déterminons les vecteurs propres correspondants (dans le cas général, *i.e.*, on ne cherche pas ceux de ces vecteurs propres qui sont aussi vecteurs propres de $\hat{\boldsymbol{p}}$). On trouve facilement

$$\text{valeur } \; c \; : \; \phi_- = -i\frac{mc^2}{\boldsymbol{p}c + \mathcal{E}}\phi_+ \tag{11.36}$$

$$\text{valeur } \; -c \; : \; \phi_- = \frac{i}{mc^2}(\boldsymbol{p}c + \mathcal{E})\phi_+. \tag{11.37}$$

Ici, un état de vitesse déterminée est toujours une superposition d'états d'énergies des deux signes.

Un état physique à énergie positive ne peut pas avoir une vitesse déterminée :

$$\hat{\boldsymbol{v}} \begin{pmatrix} \phi \\ 0 \end{pmatrix} = \begin{pmatrix} \dfrac{\boldsymbol{p}c^2}{\mathcal{E}}\phi \\ \\ -i\dfrac{mc^3}{\mathcal{E}}\phi \end{pmatrix}, \tag{11.38}$$

ce qui implique

$$\langle \boldsymbol{v} \rangle = \int d\boldsymbol{p}\, \phi^\star(\boldsymbol{p})\phi(\boldsymbol{p})\, \frac{\boldsymbol{p}c^2}{\mathcal{E}} = \left\langle \frac{\boldsymbol{p}c^2}{\mathcal{E}} \right\rangle. \tag{11.39}$$

Ceci est satisfaisant puisqu'il y a complet accord avec les théories à un seul signe pour l'énergie. Enfin,

$$(\Delta \boldsymbol{v})^2 = \langle \boldsymbol{v}^2 \rangle - \langle \boldsymbol{v} \rangle^2 = c^2 - \langle \boldsymbol{v} \rangle^2. \tag{11.40}$$

Le fait qu'un état de $\boldsymbol{v}$ déterminée ne peut être un état d'énergie déterminée signifie aussi que $[\hat{H}, \hat{\boldsymbol{v}}] \neq 0$ et donc, que $\boldsymbol{v}$ n'est pas conservée au cours du temps, même pour une particule libre. On va étudier cette évolution non-triviale de $\boldsymbol{v}$, qu'on appelle le *Zitterbewegung*. Soit $\begin{pmatrix} \phi_+ \\ \phi_- \end{pmatrix}$ l'état du sytème à l'instant $t = 0$. La valeur moyenne de $\boldsymbol{v}$ dans cet état est donnée par

$$\langle \boldsymbol{v} \rangle_0 = \int d\boldsymbol{p}\, \frac{\boldsymbol{p}c^2}{\mathcal{E}} \left[\phi_+^\star\phi_+ - \phi_-^\star\phi_-\right] + i\frac{mc^3}{\mathcal{E}} \left[\phi_+^\star\phi_- - \phi_-^\star\phi_+\right]. \tag{11.41}$$

Pour obtenir $\langle \boldsymbol{v} \rangle_t$, il suffit de remplacer $\phi_\pm$ par $\phi_\pm e^{\mp i\mathcal{E}t/\hbar}$, ce qui donne

$$\langle \boldsymbol{v} \rangle_t = \int d\boldsymbol{p}\, \frac{\boldsymbol{p}c^2}{\mathcal{E}} \left[\phi_+^\star\phi_+ - \phi_-^\star\phi_-\right] + i\frac{mc^3}{\mathcal{E}} \left[e^{2i\mathcal{E}t/\hbar}\phi_+^\star\phi_- - e^{-2i\mathcal{E}t/\hbar}\phi_-^\star\phi_+\right]. \tag{11.42}$$

Autrement dit, $\langle \boldsymbol{v} \rangle_t$ est la somme d'un terme constant (comme si $\hat{\boldsymbol{x}}$ était $\hat{\boldsymbol{q}}$) et d'un terme oscillant. Plus précisément, si on prend un état d'impulsion voisine de $\boldsymbol{p}_i$, *i.e.*,

$$\begin{pmatrix} \phi_+ \\ \phi_- \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} \times \Delta(\boldsymbol{p} - \boldsymbol{p}_i), \text{ avec } \Delta^2(\boldsymbol{p} - \boldsymbol{p}_i) = \delta(\boldsymbol{p} - \boldsymbol{p}_i).$$

On aura

$$\langle \boldsymbol{v} \rangle_t = \underbrace{\frac{\boldsymbol{p}_i c^2}{\mathcal{E}_i}(|a|^2 - |b|^2)}_{\textit{terme constant}} + \underbrace{i\frac{mc^3}{\mathcal{E}_i}\left(e^{2i\mathcal{E}_i t/\hbar} a^\star b - e^{-2i\mathcal{E}_i t/\hbar} b^\star a\right)}_{\textit{terme oscillant (Zitterbewegung)}}. \tag{11.43}$$

Le terme oscillant est en $\sin(2\mathcal{E}t/\hbar)$ et $\cos(2\mathcal{E}t/\hbar)$. Il s'explique bien sûr par interférence entre les états d'énergie positive et ceux d'énergie négative correspondant à la valeur donnée $\boldsymbol{p}_i$ de $\boldsymbol{p}$ et cela produit le facteur de Bohr correspondant (bien sûr, le *Zitterbewegung* disparaît pour un état à énergie positive). On a rapporté en appendice les calculs similaires portant sur l'opérateur position.

Une dernière propriété de la vitesse et de la position (on y reviendra) : on calcule

$$[\hat{\boldsymbol{x}}, \hat{\boldsymbol{v}}] = 0, \tag{11.44}$$

ce qui est curieux *a priori* ...

4.3. Représentation *x* en théorie de Dirac

On a déjà vu les vecteurs propres de $\hat{\boldsymbol{x}}$ correspondant à une valeur propre $\boldsymbol{x}$:

$$|\,\boldsymbol{x}, 1\,\rangle \to e^{\frac{i}{2}\mathrm{Arctg}\frac{\boldsymbol{p}}{mc}} \frac{e^{-i\boldsymbol{px}/\hbar}}{\sqrt{2\pi\hbar}} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \tag{11.45}$$

$$|\,\boldsymbol{x}, 2\,\rangle \to e^{-\frac{i}{2}\mathrm{Arctg}\frac{\boldsymbol{p}}{mc}} \frac{e^{-i\boldsymbol{px}/\hbar}}{\sqrt{2\pi\hbar}} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}, \tag{11.46}$$

qui vérifient $\langle\, \boldsymbol{x}, \alpha \,|\, \boldsymbol{x}', \alpha'\,\rangle = \delta_{\alpha\alpha'}\,\delta(\boldsymbol{x} - \boldsymbol{x}')$.

Un état $|\,\psi\,\rangle$ peut s'écrire

$$|\,\psi\,\rangle = \int d\boldsymbol{x}\ (\psi_1(\boldsymbol{x})\,|\,\boldsymbol{x}, 1\,\rangle + \psi_2(\boldsymbol{x})\,|\,\boldsymbol{x}, 2\,\rangle) = \sum_\alpha \int d\boldsymbol{x}\ \psi_\alpha(\boldsymbol{x})\,|\,\boldsymbol{x}, \alpha\,\rangle.$$

Ainsi, $|\,\psi\,\rangle$ est décrit par $\begin{pmatrix}\psi_1(\boldsymbol{x})\\ \psi_2(\boldsymbol{x})\end{pmatrix}$. On calcule immédiatement :

$$\langle\,\phi\;|\,\psi\,\rangle = \int d\boldsymbol{x}\;\left(\phi_1^\star(\boldsymbol{x})\,\psi_1(\boldsymbol{x}) + \phi_2^\star(\boldsymbol{x})\,\psi_2(\boldsymbol{x})\right). \tag{11.47}$$

Donnons l'expression des divers opérateurs en représentation $\boldsymbol{x}$. On trouve successivement :

$$\hat{\boldsymbol{x}}\begin{pmatrix}\psi_1(\boldsymbol{x})\\ \psi_2(\boldsymbol{x})\end{pmatrix} = \begin{pmatrix}\boldsymbol{x}\,\psi_1(\boldsymbol{x})\\ \boldsymbol{x}\,\psi_2(\boldsymbol{x})\end{pmatrix}, \tag{11.48}$$

$$\hat{\boldsymbol{p}}\begin{pmatrix}\psi_1(\boldsymbol{x})\\ \psi_2(\boldsymbol{x})\end{pmatrix} = \begin{pmatrix}-i\hbar\psi_1'(\boldsymbol{x})\\ -i\hbar\psi_2'(\boldsymbol{x})\end{pmatrix}, \tag{11.49}$$

$$\hat{H}\begin{pmatrix}\psi_1(\boldsymbol{x})\\ \psi_2(\boldsymbol{x})\end{pmatrix} = \begin{pmatrix}mc^2\,\psi_2(\boldsymbol{x}) + \hbar c\,\psi_2'(\boldsymbol{x})\\ mc^2\,\psi_1(\boldsymbol{x}) - \hbar c\,\psi_1'(\boldsymbol{x})\end{pmatrix}, \tag{11.50}$$

ou encore, si on introduit

$$\hat{\beta}\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} = \begin{pmatrix}\psi_2\\ \psi_1\end{pmatrix}, \tag{11.51}$$

$$\hat{\boldsymbol{\alpha}}\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} = \begin{pmatrix}0 & -ic\\ ic & 0\end{pmatrix}\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix} = \begin{pmatrix}-ic\,\psi_2\\ ic\,\psi_1\end{pmatrix}, \tag{11.52}$$

alors

$$\hat{H} = \hat{\boldsymbol{\alpha}}\hat{\boldsymbol{p}} + \hat{\beta}mc^2, \tag{11.53}$$

et l'on retrouve la théorie de Dirac,

$$\hat{\boldsymbol{\alpha}}^2 = c^2,\; \hat{\beta}^2 = 1,\; \hat{\boldsymbol{\alpha}}\hat{\beta} + \hat{\beta}\hat{\boldsymbol{\alpha}} = 0. \tag{11.54}$$

5. Appendice : évolution de la position moyenne

Soit $\begin{pmatrix}\phi_+\\ \phi_-\end{pmatrix}$ l'état du sytème à l'instant $t = 0$. La valeur moyenne de $\boldsymbol{x}$ dans cet état est donnée par

$$\langle\boldsymbol{x}\rangle_0 = \int d\boldsymbol{p}\,\left[i\hbar\left(\phi_+^\star\phi_+' + \phi_-^\star\phi_-'\right) + \frac{\hbar}{2}\frac{mc^3}{\boldsymbol{p}^2c^2 + m^2c^4}\left(\phi_+^\star\phi_- + \phi_-^\star\phi_+\right)\right]. \tag{11.55}$$

Pour obtenir $\langle \boldsymbol{x} \rangle_t$, il suffit de remplacer $\phi_\pm$ par $\phi_\pm e^{\mp i\mathcal{E}t/\hbar}$, ce qui donne après quelques calculs :

$$\begin{aligned}
\langle \boldsymbol{x} \rangle_t &= \int d\boldsymbol{p}\, i\hbar \left(\phi_+^\star \phi_+' + \phi_-^\star \phi_-'\right) \\
&\quad + \int d\boldsymbol{p}\, \frac{\boldsymbol{p}c^2 t}{\mathcal{E}} \left(\phi_+^\star \phi_+' - \phi_-^\star \phi_-'\right) \\
&\quad + \int d\boldsymbol{p}\, \frac{\hbar}{2} \frac{mc^3}{\mathcal{E}^2} \left(e^{2i\mathcal{E}t/\hbar} \phi_+^\star \phi_- + e^{-2i\mathcal{E}t/\hbar} \phi_-^\star \phi_+\right) \\
&= a + bt + \text{ Zitterbewegung.}
\end{aligned} \tag{11.56}$$

Si on prend un état d'impulsion voisine de $\boldsymbol{p}_i$, *i.e.*,

$$\begin{pmatrix} \phi_+ \\ \phi_- \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} \times \Delta(\boldsymbol{p} - \boldsymbol{p}_i), \text{ avec } \Delta^2(\boldsymbol{p} - \boldsymbol{p}_i) = \delta(\boldsymbol{p} - \boldsymbol{p}_i),$$

le terme de Zitterbewegung se comporte comme

$$\frac{\hbar}{2} \frac{mc^3}{\mathcal{E}^2} \times \left[\cos(2\mathcal{E}t/\hbar) \text{ ou } \sin(2\mathcal{E}t/\hbar)\right] \approx \frac{\hbar}{2mc} \times \text{ facteur oscillant.}$$

Bibliographie

[1] Dirac, P.A.M., *The Principles of Quantum Mechanics*, Oxford University Press, Oxford, 1958.

[2] Hegerfeldt, G.C., Remarks on causality and particle localization, *Phys. Rev. D* 10 (1974), 3320–3321.

[3] Hegerfeldt, G.C., Ruijsenaars N.M., Remarks on causality, localization, and spreading of wave packets, *Phys. Rev. D* 22 (1980), 377–384.

[4] Kálnay, A.J., The Localization Problem, *Problems in the Foundations of Physics*, Ed. M. Bunge, Studies in the Foundations Methodology and Philosophy of Science, IV, Springer-Verlag, Berlin, 1971.

[5] Lévy-Leblond, J.-M., The pedagogical Role and Epistemological Significance of Group Theory in Quantum Mechanics *Riv. Nuovo Cimento* 4 (1974), 99–143.

[6] Newton, T.D., Wigner, E.P., Localized states for elementary systems, *Rev. Mod. Phys.* 21 (1949), 400–406.

[7] Schweber, S., *Relativistic Quantum Field Theory*, Harper & Row, New York, 1961.

[8] Wightman, A.S., On the localizability of quantum mechanical states, *Rev. Mod. Phys.* 34 (1962), 845–872.

12 Il y a différentes manières de prendre position[1]

Jean-Pierre Gazeau

1. Introduction

Le formalisme de la mécanique quantique et celui de l'analyse du signal ont beaucoup d'aspects communs. On trouve dans l'un et l'autre un ensemble *brut* X de paramètres ou *données* de base : il peut être l'espace des phases pour l'un, le temps, le plan temps-fréquence pour l'autre. En fait, il peut être bien plus. Cet ensemble a une structure minimale : il est doté d'une mesure $\mu(dx)$ ce qui lui confère le statut d'*espace mesurable*. Cela permet une lecture statistique de l'ensemble des fonctions mesurables $f(x)$, à valeurs réelles ou complexes : calculer par exemple des valeurs moyennes sur un sous-ensemble de mesure bornée. L'analyse du signal comme la mécanique quantique envisagent en fait des moyennes quadratiques. Leur objet d'étude devient l'espace de Hilbert $L^2(X, \mu)$ des fonctions $f(x)$ de carré intégrable sur X : $\int_X |f(x)|\, \mu(dx) < \infty$. On parlera de signal d'*énergie finie* dans un cas, et d'état quantique dans l'autre. C'est toutefois à ce point que la lecture *quantique* de X diffère à plusieurs titres de la lecture signaliste :

1. Toutes les fonctions de carré intégrable ne sont pas éligibles comme états quantiques.
2. Un état quantique est défini à un facteur de phase près.
3. Celles parmi les fonctions $f(x)$ qui sont éligibles comme états quantiques et qui sont normalisées à l'unité, $\int_X |f(x)|\, \mu(dx) = 1$, donnent lieu à une interprétation probabiliste : $|f(x)|^2 \mu(dx)$ est une densité de

1 – *Un dit de Jean-Jacques à la Rencontre de Cargese*, 29 Janvier 2001, 11 h 55.

probabilité interprétable en terme de *localisation* sur X et inhérente au calcul de valeurs moyennes de grandeurs dites *observables quantiques*.

Le premier point est au cœur du problème de *quantification*. Par quel procédé plus ou moins *canonique* sélectionner le bon grain quantique de l'ivraie signaliste, c'est-à-dire trouver le bon sous-espace de Hilbert, disons $\mathcal{H}$ formé des états quantiques, ou encore, trouver le bon projecteur $P_{\mathcal{H}}$ orthogonal qui lui correspond ?

Ce problème peut avoir une solution lorsque l'on trouve une application de X dans $\mathcal{H}$, $X \ni x \to \eta_x \in \mathcal{H}$, définissant une famille d'états $\{\eta_x\}$, normalisés à l'unité, indexés par *tous* les points de X, et *résolvant* le projecteur $P_{\mathcal{H}}$ de la manière suivante :

$$P_{\mathcal{H}} = \int_X P_{\{\eta_x\}}\, \mu(dx),$$

où $P_{\{\eta_x\}}$ est le projecteur orthogonal sur η_x. La quantification d'une *observable classique*, c'est-à-dire d'une fonction $f(x)$ sur X possédant des propriétés en rapport avec telle ou telle structure supplémentaire, topologie, géométrie et autres, imposée à X, consiste simplement à lui faire correspondre l'opérateur défini par :

$$F = \int_X f(x)\; P_{\{\eta_x\}}\, \mu(dx).$$

La construction explicite de la famille $\{\eta_x\}$ est évidemment cruciale. Il est remarquable que les deux formalismes, signaliste et quantique, se retrouvent à nouveau sur ce terrain : c'est ce qu'on appelle *analyse en ondelettes* pour le premier, et construction d'*états cohérents* (au sens large) pour le second.

Nous verrons dans l'exposé qui suit comment la quête d'une définition cohérente de l'espace en mécanique quantique trouve naturellement sa voie dans ce formalisme d'états cohérents (*sic*). Les outils mathématiques nécessaires sont d'abord introduits dans la section 2 à l'aide de l'un des exemples les plus simples que l'on puisse imaginer : X est le cercle unité et $\mathcal{H}$ est le plan euclidien, un formalisme quantique pour collège et lycée ! Les sections 3 et 4 esquissent un portrait, que l'on espère accessible à des non-spécialistes, des mécaniques quantiques galiléenne et einsteinienne, et des problématiques de localisation spatiale qui leur sont propres. On montre plus précisément dans les sections 5 et 6 qu'une localisation spatiale « taillée au couteau » est incompatible avec la causalité einsteinienne, alors qu'une localisation compatible mais moins abrupte existe grâce justement à une quantification « cohérente ».

2. Repères discrets, continus du plan

2.1. Le $\langle$ LE | GO $\rangle$ des $\langle$ bra | et des | ket $\rangle$

Considérons le plan $\mathbb{R}^2$ muni de sa structure euclidienne, produit scalaire, métrique, angle ..., structure que nous décrirons ici avec les notations des *bras* et des *kets* de Dirac [9]. Ainsi, un vecteur $\boldsymbol{u}$, de coordonnées x et y par rapport à un repère orthonormé est désigné par

$$| \, \boldsymbol{u} \, \rangle = \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{vecteur colonne,}$$

tandis que son adjoint s'écrit $\langle \, \boldsymbol{u} \, | = (x, y)$ (vecteur ligne). En conséquence, le produit scalaire $\boldsymbol{u} \cdot \boldsymbol{u}' = xx' + yy'$ de deux vecteurs $\boldsymbol{u}$ et $\boldsymbol{u}'$ s'écrit aussi $\langle \, \boldsymbol{u} \, | \, \boldsymbol{u}' \, \rangle$, et le projecteur orthogonal sur $\boldsymbol{u}$ est donné par

$$P_{\boldsymbol{u}} = \frac{| \, \boldsymbol{u} \, \rangle \langle \, \boldsymbol{u} \, |}{\langle \, \boldsymbol{u} \, | \, \boldsymbol{u} \, \rangle} = \frac{1}{x^2 + y^2} \begin{pmatrix} x^2 & xy \\ xy & y^2 \end{pmatrix}.$$

Un vecteur de longueur unité, d'angle polaire θ, $0 \leq \theta < 2\pi$, sera simplement désigné par $| \, \theta \, \rangle$.

2.2. Repères du plan

Le repère orthonormé habituel, cartésien, est constitué des deux vecteurs $| \, 0 \, \rangle$ et $| \, \frac{\pi}{2} \, \rangle$,

$$\langle \, 0 \, | \, 0 \, \rangle = 1 = \left\langle \frac{\pi}{2} \, \middle| \, \frac{\pi}{2} \right\rangle,$$

tels que la somme de leurs projecteurs *résout l'identité*,

$$\mathbb{1} = | \, 0 \, \rangle \langle \, 0 \, | + \left| \frac{\pi}{2} \right\rangle \left\langle \frac{\pi}{2} \right|, \tag{12.1}$$

simple traduction de l'identité matricielle

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Mais que se passe-t-il si nous remplaçons $| \, \frac{\pi}{2} \, \rangle \langle \, \frac{\pi}{2} \, |$ par $| \, \theta \, \rangle \langle \, \theta \, |$, $\theta \neq k\pi/2$? Nous obtenons un repère *exact* (formé de deux vecteurs linéairement indépendants), mais *oblique* (non-orthogonal) :

$$A_\theta = | \, 0 \, \rangle \langle \, 0 \, | + | \, \theta \, \rangle \langle \, \theta \, |, \tag{12.2}$$

Maintenant A_θ ne se réduit plus à l'identité. En notation matricielle, il s'écrit

$$A_\theta = \begin{pmatrix} 1+\cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{pmatrix}. \tag{12.3}$$

Cet opérateur est symétrique, positif, inversible, de valeurs propres $\lambda_c = 1 + \cos\theta$ et $\lambda_s = 1 - \cos\theta$. Il n'est égal à l'identité que pour $\theta = \pi/2, 3\pi/2$. La *largeur* ou degré d'oblicité du repère est définie par $|\lambda_c - \lambda_s|/(\lambda_c - \lambda_s) = |\cos\theta|$. Il est ce qu'on appelle en analyse hilbertienne ou en analyse du signal [21] un repère exact non strict pour $\theta = \pi/2, 3\pi/2$. Si nous ajoutons à A_θ un autre projecteur $|\,\phi\,\rangle\langle\,\phi\,|$, $\phi \neq 0, \neq \theta$, nous obtenons un repère inexact ou *surcomplet*, et certainement non strict en général, c'est-à-dire de largeur non nulle, *sauf* dans la circonstance particulière où $\theta = 2\pi/3$ et $\phi = 4\pi/3$. En effet, un calcul immédiat [7] montre que ce repère *équilatéral* vérifie

$$|\,0\,\rangle\langle\,0\,| + \left|\,\frac{2\pi}{3}\right\rangle\left\langle\frac{2\pi}{3}\,\right| + \left|\,\frac{4\pi}{3}\right\rangle\left\langle\frac{4\pi}{3}\,\right| = \frac{3}{2}\,\mathbb{1}. \tag{12.4}$$

Cette relation remarquable est aisément généralisable au repère étoilé à n branches ou *cyclotomique* d'ordre n :

$$\frac{2}{n}\sum_{q=0}^{n-1}\left|\,\frac{2\pi q}{n}\right\rangle\left\langle\frac{2\pi q}{n}\,\right| = \mathbb{1}. \tag{12.5}$$

2.3. Un repère continu du plan

La relation (12.5) a une limite continue pour $n \to \infty$ après avoir effectué les substitutions $(2/n \to (2\pi/n)(d\theta/\pi)$, $(2\pi q)/n \to \theta$. Cette limite s'écrit

$$\frac{1}{\pi}\int_0^{2\pi} d\theta\,|\,\theta\,\rangle\langle\,\theta\,| = \mathbb{1}, \tag{12.6}$$

une résolution de l'identité facile à vérifier par intégration directe, élément par élément, de la matrice du projecteur $P_\theta = |\,\theta\,\rangle\langle\,\theta\,|$. Nous avons ainsi obtenu un repère continu strict pour le plan, c'est-à-dire un ensemble continu de vecteurs unitaires pour repérer, avec une redondance extrême, un espace vectoriel réel à deux dimensions : le cercle unité comme repère du plan ! En fait, on trouve déjà dans les premiers travaux de Klauder [16] de telles résolutions continues dans les espaces euclidiens ou hermitiens, ou même hilbertiens. La relation opératorielle (12.6) doit aussi bien se lire au travers de son action sur un vecteur $|\, u \,\rangle = \|u\| \,|\, \phi \,\rangle$:

$$|\, u \,\rangle = \frac{\|u\|}{\pi} \int_0^{2\pi} d\theta \cos(\phi - \theta) \,|\, \theta \,\rangle, \tag{12.7}$$

relation qui illustre bien la *surcomplétude* du système $\{|\, \theta \,\rangle\}$. Les vecteurs de ce système ne sont pas linéairement indépendants, et leur recouvrement respectifs sont donnés par les produits scalaires $\langle\, \phi \,|\, \theta \,\rangle = \cos(\phi - \theta)$.

Cet exemple est certainement le plus simple que l'on puisse exhiber pour illustrer un formalisme de « bases continues » connu en mécanique quantique sous le nom *d'états cohérents* [3, 15, 16] et en analyse du signal sous le nom de *base continue d'ondelettes* [3, 7, 21].

2.4. La sonde transportée comme repère

La parenté avec les états cohérents et l'analyse continue en ondelettes peut être rendue encore plus flagrante si on adopte une démarche de type théorie des groupes dans la construction du repère

$$[0, 2\pi) \ni \theta \to P_\theta = |\, \theta \,\rangle\langle\, \theta \,|. \tag{12.8}$$

En effet, le plan $\mathbb{R}^2$ réalise une représentation unitaire, c'est-à-dire conservant le produit scalaire, du groupe $SO(2) = \{g(\phi), \phi \in [0, 2\pi)\}$:

$$\begin{aligned} |\, u \,\rangle &\xrightarrow{g(\phi)} g(\phi) \,|\, u \,\rangle \equiv \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\ \Leftrightarrow \langle\, u \,| &\xrightarrow{g(\phi)} \langle\, u \,|\, g(-\phi), \end{aligned} \tag{12.9}$$

action qui s'écrit plus simplement quand elle s'applique à un vecteur unitaire :

$$g(\phi) \,|\, \theta \,\rangle = |\, \theta + \phi \,\rangle, \tag{12.10}$$

ou encore au projecteur associé à ce dernier,

$$P_\theta \to g(\phi) P_\theta \, g(-\phi) = P_{\theta+\phi}. \tag{12.11}$$

Ainsi, la famille discrète finie $\{|\, (2\pi q)/n \,\rangle\}$ (resp. continue $\{|\, \theta \,\rangle\}$) de vecteurs peut aussi s'envisager comme l'orbite du vecteur « origine » $|\, 0 \,\rangle$ sous l'action de tous les éléments du groupe fini *cyclotomique* $\{g((2\pi q)/n)\}$ (resp. continu $SO(2)$). Autrement dit, *le repère* $\{|\theta\,\rangle\}$ *ou* $\{P_\theta\}$ *est obtenu par transport, ici sous* *forme de rotation, de la sonde* $|\,0\,\rangle$ (ou $|\,0\,\rangle\langle\,0\,|$). Cette sonde nous permet une exploration directionnelle (sur-)complète du plan. Notons enfin que la résolution de l'identité est en fait une simple conséquence du lemme de Schur : en appliquant (12.11) et en effectuant le changement de variable $\theta + \phi \to \theta$ dans l'intégrale, on voit bien que l'opérateur $A = \int_0^{2\pi} P_\theta \, d\theta$ commute avec les opérateurs $g(\phi)$ pour tout ϕ. Or ces derniers agissent irréductiblement sur le tore, et donc A est multiple de l'identité.

2.5. Le calcul symbolique

L'existence du repère continu $\{|\,\theta\,\rangle\}$ offre aussi la possibilité d'un calcul symbolique à la Berezin-Lieb [6, 10]. Cela signifie ici qu'à tout opérateur $\mathcal{O}$ linéaire *auto-adjoint*, c'est-à-dire donné par une matrice symétrique dans une base orthonormée, on peut associer deux types de *symboles*, fonctions $\check{\mathcal{O}}(\theta)$ et $\hat{\mathcal{O}}(\theta)$ définies sur le cercle unité respectivement par

$$\check{\mathcal{O}}(\theta) = \langle\,\theta\,|\,\mathcal{O}\,|\,\theta\,\rangle \ : \ \textit{symbole inférieur ou covariant}, \tag{12.12}$$

(en mécanique quantique, on parlerait de *valeur moyenne de l'observable $\mathcal{O}$ dans l'état $|\,\theta\,\rangle$*),

$$\mathcal{O} = \frac{1}{\pi}\int_0^{2\pi} d\theta\,\hat{\mathcal{O}}(\theta)\,|\,\theta\,\rangle\langle\,\theta\,|\,. \tag{12.13}$$

La fonction $\hat{\mathcal{O}}(\theta)$ apparaissant dans cette intégrale à valeurs opérateurs est appelée *symbole supérieur ou contravariant*. Trois matrices de base engendrent l'algèbre de Jordan de toutes les matrices symétriques réelles d'ordre 2 : la matrice identité, dont les symboles sont la fonction égale à 1, et les deux matrices de Pauli.

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{12.14}$$

Un élément quelconque $\mathcal{O}$ de l' algèbre se décompose alors comme

$$\mathcal{O} \equiv \begin{pmatrix} a & b \\ b & d \end{pmatrix} = \frac{a+d}{2}\mathbb{1} + \frac{a-d}{2}\sigma_3 + b\sigma_1 \equiv \alpha\,\mathbb{1} + \delta\sigma_3 + \beta\sigma_1. \tag{12.15}$$

Le produit dans cette algèbre est défini par

$$\mathcal{O} \odot \mathcal{O}' = \frac{1}{2}\,(\mathcal{O}\mathcal{O}' + \mathcal{O}'\mathcal{O}), \tag{12.16}$$

ce qui se traduit au niveau des composantes (α, δ, β) par

$$\alpha'' = \alpha\alpha' + \delta\delta' + \beta\beta',\ \ \delta'' = \alpha\delta' + \alpha'\delta,\ \ \beta'' = \alpha\beta' + \alpha'\beta. \tag{12.17}$$

Les symboles des trois éléments de base sont respectivement donnés par

$$\cos 2\theta = \check{\sigma}_3(\theta) = \frac{1}{2}\hat{\sigma}_3(\theta), \ \sin 2\theta = \check{\sigma}_1(\theta) = \frac{1}{2}\hat{\sigma}_1(\theta). \tag{12.18}$$

Il en résulte pour la matrice symétrique (12.15) les deux symboles

$$\begin{aligned}\check{O}(\theta) &= \frac{a+d}{2} + \frac{a-d}{2}\cos 2\theta + b\sin 2\theta = \alpha + \delta\cos 2\theta + \beta\sin 2\theta,\\ \hat{O}(\theta) &= \alpha + 2\delta\cos 2\theta + 2\beta\sin 2\theta = 2\check{O}(\theta) - \frac{1}{2}\operatorname{Tr} O.\end{aligned} \tag{12.19}$$

On remarquera que ces symboles sont tous des éléments du sous-espace des séries de Fourier réelles engendré par la constante et les cosinus et sinus de la deuxième harmonique. De plus, la loi multiplicative de Jordan (12.16), commutative mais non associative, génère un produit « $\star$ » (on dira *star produit*), de même nature, au niveau des symboles. Par exemple, nous avons pour les symboles supérieurs :

$$\widehat{O \odot O'}(\theta) \equiv \hat{O}(\theta) \star \hat{O}'(\theta) = \alpha\hat{O}'(\theta) + \alpha'\hat{O}(\theta) + \delta\delta' + \beta\beta' - \alpha\alpha', \tag{12.20}$$

et la formule est identique pour les symboles inférieurs.

La terminologie *inférieur/supérieur* est justifiée par deux inégalités, dites de Berezin-Lieb, qui se déduisent du formalisme symbolique. Soit g une fonction convexe. Désignant par $\lambda_\pm$ les valeurs propres de la matrice symétrique O, on a

$$\frac{1}{\pi}\int_0^{2\pi} g(\check{O}(\theta))\,d\theta \leq \operatorname{Tr} g(O) = g(\lambda_+) + g(\lambda_-) \leq \frac{1}{\pi}\int_0^{2\pi} g(\hat{O}(\theta))\,d\theta. \tag{12.21}$$

Cette double inégalité est non triviale. Dégagée du contexte euclidien, elle peut se lire ainsi :

$$\langle g(t + r\cos\theta)\rangle \leq \frac{1}{2}\,]g(t+r) + g(t-r)] \leq \langle g(t + 2r\cos\theta)\rangle, \tag{12.22}$$

où $t \in \mathbb{R}, r \geq 0$, et où $\langle\cdot\rangle$ désigne la valeur moyenne sur une période. Si on particularise à la fonction exponentielle $g(X) = e^X$, on obtient l'entrelacement intéressant entre fonction de Bessel du second genre et le cosinus hyperbolique :

$$\cdots \leq I_0(x) \leq \cosh x \leq I_0(2x) \leq \cosh 2x \leq \cdots \quad \forall x \in \mathbb{R}. \tag{12.23}$$

2.6. Le plan euclidien « vu » dans l'espace de Hilbert $L^2(S^1)$

Un des aspects du repère continu $\{P_\theta\}$ est aussi de nous donner la possibilité d'un plongement isométrique du plan euclidien dans l'espace de Hilbert $L^2(S^1)$ des fonctions réelles de carré intégrable sur le cercle, espace muni du produit scalaire

$$\langle f \mid g \rangle = \frac{1}{\pi} \int_0^{2\pi} f(\theta) g(\theta)\, d\theta, \tag{12.24}$$

connu encore comme l'espace des séries de Fourier réelles. Soit un vecteur $\mid u \rangle$, de coordonnées polaires r, ϕ. À θ donné, il lui correspond la fonction

$$u(\theta) = \langle \theta \mid u \rangle = r \cos(\phi - \theta), \tag{12.25}$$

qui est clairement un élément de $L^2(S^1)$. Nous avons là une isométrie puisqu'une application directe de la résolution de l'identité (12.6) nous amène à

$$\frac{1}{\pi} \int_0^{2\pi} u(\theta) v(\theta)\, d\theta = \frac{1}{\pi} \int_0^{2\pi} \langle u \mid \theta \rangle \langle \theta \mid v \rangle = \langle u \mid v \rangle. \tag{12.26}$$

Où l'on voit comment le plan euclidien s'injecte isométriquement dans l'espace des séries de Fourier, en ayant pour image le sous-espace de dimension deux engendré par le sinus et le cosinus de l'harmonique fondamentale. Où l'on remarque aussi que le sous-espace des symboles est le carré algébrique de cette image du plan euclidien.

2.7. Aspects probabilistes

Il y a derrière la résolution de l'unité (12.6) une interprétation intéressante en termes de probabilité géométrique. Considérons à cet effet un sous-ensemble Δ de Borel de l'intervalle $[0, 2\pi)$ et la restriction à Δ de l'intégrale (12.6)

$$a(\Delta) = \frac{1}{\pi} \int_\Delta d\theta \mid \theta \rangle \langle \theta \mid. \tag{12.27}$$

On vérifie aisément les propriétés suivantes

$$\begin{aligned} &a(\emptyset) = 0, \;\; a([0, 2\pi)) = \mathbb{1}, \\ &a(\cup_{i\in J} \Delta_i) = \sum_{i \in J} a(\Delta_i), \;\; \Delta_i \cap \Delta_j \text{ pour tous } i \neq j. \end{aligned} \tag{12.28}$$

L'application $\Delta \to a(\Delta)$ définit donc une mesure normalisée à l'unité sur la σ-algèbre des boréliens de l'intervalle $[0, 2\pi)$, à valeurs dans l'ensemble des opérateurs linéaires positifs sur le plan euclidien (l'acronyme anglais d'un tel objet est POV). D'ailleurs, nous pourrions aussi bien noter la « densité » $(1/\pi)P_\theta\, d\theta$ par $a(d\theta)$ de telle sorte que

$$a(\Delta) = \int_\Delta a(d\theta), \tag{12.29}$$

et il nous arrivera par la suite d'utiliser une telle notation. Notons de plus que cette mesure possède une propriété de covariance par rapport aux rotations du plan. En effet,

$$g(\phi)a(\Delta)g(-\phi) = a(\Delta + \phi \bmod(2\pi)). \tag{12.30}$$

La mise en évidence du caractère probabiliste de la mesure $a(\Delta)$ s'opère de la manière suivante. Considérons un vecteur unitaire $|\,\phi\,\rangle$. L'application

$$\Delta \to \langle\,\phi \mid a(\Delta) \mid \phi\,\rangle = \frac{1}{\pi}\int_\Delta \cos^2(\theta-\phi)\, d\theta \tag{12.31}$$

est bien une mesure de probabilité. Elle est positive, normalisée à l'unité, et elle hérite de $a(\Delta)$ ses propriétés de σ-additivité. Or la quantité $\langle\,\phi \mid a(\Delta) \mid \phi\,\rangle$ signifie que l'on examine la direction $|\,\phi\,\rangle$ du point de vue de la famille de vecteurs $|\,\theta\,\rangle$, $\theta \in \Delta$. En fait elle a signification de *probabilité géométrique* dans le plan [8]. On ne perdra point en généralité en posant $\phi = 0$. Rappelons qu'une droite quelconque $D_{\theta,p}$ du plan est décrite par son équation canonique

$$\langle\,\theta \mid u\,\rangle \equiv \cos\theta\, x + \sin\theta\, y = p, \tag{12.32}$$

où $|\theta\,\rangle$ est la direction perpendiculaire à $D_{\theta,p}$ tandis que p est égal à la distance de cette dernière à l'origine. Il en résulte que $dp\, d\theta$ est l'élément de mesure de probabilité (non normalisée) sur l'ensemble $\{D_{\theta,p}\}$ des droites prises au hasard dans le plan. Considérons alors, pour un choix de θ, l'ensemble $\{D_{\theta,p}\}$ des droites sécantes du segment d'origine O et de longueur $|\cos\theta|$ issu de la projection de $|\,\theta\,\rangle$ sur $|\,0\,\rangle$. Cet ensemble là a pour mesure

$$d\theta \left(\int_0^{\cos^2\theta} dp\right) d\theta = \cos^2\theta\, d\theta, \tag{12.33}$$

qui, une fois intégrée sur toutes les directions, redonne la surface du cercle unité. On peut donc penser que $\langle\,\phi \mid a(\Delta) \mid \phi\,\rangle$ est la probabilité pour une droite du plan d'être part de l'ensemble des sécantes aux segments issus de l'ensemble des projections $\langle\,\phi \mid \theta\,\rangle$ des vecteurs unitaires $|\,\theta\,\rangle$, $\theta \in \Delta$ sur le vecteur unitaire $|\,\phi\,\rangle$, vecteurs tous rapportés à l'origine du plan. Il y a aussi, en filigrane de cette interprétation plutôt alambiquée l'image de *polariseur* $\langle\,\theta\,|$ et *analyseur* $|\,\theta\,\rangle$ appliqués en « circuit fermé » au « signal directionnel » $|\,\phi\,\rangle$.

Cette longue disgression initiatrice autour d'aspects inattendus du plan euclidien nous aura finalement servi à mieux nous familiariser avec les notations de Dirac et avec des outils comme la théorie des groupes, les états cohérents, les systèmes de covariance, avant d'aborder des problèmes propres à la physique quantique.

3. Système élémentaire galiléen

L'objet mathématique sous-tendant le concept de système élémentaire galiléen (à $1+1$ dimensions) est le groupe *étendu* de Galilée :

$$\tilde{\mathcal{G}}(1,1) = \{g = (\theta, b, \boldsymbol{a}, \boldsymbol{v})\}\cdot \tag{12.34}$$

Ce groupe est structuré en produit semi-direct :

$$\tilde{\mathcal{G}}(1,1) = (\Theta \times \mathcal{T} \times \mathcal{S}) \rtimes \mathcal{V}, \tag{12.35}$$

où Θ est le sous-groupe des phases de l'extension, $\mathcal{T}$ (resp. $\mathcal{S}$) est le sous-groupe des translations de temps (resp. d'espace), et $\mathcal{V}$ est le sous-groupe des transformations pures de Galilée (*boosts galiléens*).

La loi de groupe est donnée par

$$g_1 g_2 = (\theta_1 + \theta_2 + \xi(g_1, g_2), b_1 + b_2, \boldsymbol{a}_1 + \boldsymbol{a}_2 + \boldsymbol{v}_1 b_2, \boldsymbol{v}_1 + \boldsymbol{v}_2), \tag{12.36}$$

où le *multiplicateur* ou *cocycle* $\xi : \tilde{\mathcal{G}} \times \tilde{\mathcal{G}} \to \mathbb{R}$ est défini par $\xi(g_1, g_2) = m[(1/2)\boldsymbol{v}_1^2 b_2 + \boldsymbol{v}_1 \cdot \boldsymbol{a}_2]$. $m > 0$ est une masse, celle justement du système élémentaire considéré. L'extension du groupe de Galilée *stricto sensu* $\mathcal{G}(1,1) = \{(b, \boldsymbol{a}, \boldsymbol{v})\}$ par $\Theta = \{\theta\}$ est requise dès que l'on considère le système élémentaire dans un contexte quantique [17] : c'est précisément l'adjonction de Θ qui, avec l'existence d'un paramètre masse, confère au sous-groupe de Galilée *isochrone* $\{(\theta, 0, \boldsymbol{a}, \boldsymbol{v})\} \simeq \tilde{\mathcal{G}}(1,1)/\mathcal{T}$ son statut de groupe de la mécanique quantique, ou groupe de Weyl-Heisenberg : même à temps fixé, les translations d'espace ne commutent pas avec les boosts comme on peut le constater dans (12.36), traduction, au niveau du groupe, des relations canoniques de commutation.

Le groupe de relativité détermine les deux géométries fondamentales à la base des mécaniques classique et quantique : l'espace-temps (point de vue lagrangien) et l'espace des phases (point de vue hamiltonien). Ces deux espaces sont chacun quotient du groupe par un sous-groupe bien identifié :

(i) Espace-temps galiléen

$$\tilde{\mathcal{G}}(1,1)/(\Theta\times\mathcal{V}) \simeq \{(0,b,\boldsymbol{a},\boldsymbol{0})\} \simeq \{(t,\boldsymbol{x})\}. \tag{12.37}$$

(ii) Espace des phases du système élémentaire (ou espace des états classiques)

$$\tilde{\mathcal{G}}(1,1)/(\Theta\times\mathcal{T}) \simeq \{(0,0,\boldsymbol{a},\boldsymbol{v})\} \simeq \{(\boldsymbol{q},\boldsymbol{p})\}\cdot \tag{12.38}$$

L'action du groupe sur chacune de ces deux géométries, l'une euclidienne, l'autre symplectique, se déduit de ces structures quotients. Rappelons ici que toute paire (G,H), où G est un groupe localement compact et H est un sous-groupe fermé de G, détermine un espace $X = G/H$ sur lequel G agit d'une manière transitive (on dit que X est un espace homogène). Les éléments de X sont de la forme $X = \{gH,\ g\in G\}$ et l'action *à gauche* de G sur X s'écrit

$$gH \overset{g'}{\to} (g'g)H. \tag{12.39}$$

Le système élémentaire galiléen quantique de masse m est celui associé à la représentation unitaire irréductible $R_m(g)$ de $\tilde{\mathcal{G}}$, définie par l'action suivante sur l'espace de Hilbert $L^2(\mathbb{R}, d\boldsymbol{k})$ des fonctions $\psi(\boldsymbol{k}) \equiv \langle \boldsymbol{k} \mid \psi\rangle$ de la variable quantité de mouvement $\boldsymbol{k}$ et de carré intégrable sur la droite :

$$(R_m(\theta,b,\boldsymbol{a},\boldsymbol{v})\psi)(\boldsymbol{k}) = e^{i(\theta+(\boldsymbol{k}^2/2m)b-\boldsymbol{k}\cdot\boldsymbol{a})}\,\psi(\boldsymbol{k}-m\boldsymbol{v}). \tag{12.40}$$

On pose ici et pour (presque) tout le reste de l'article $\hbar = 1$. Cette action du groupe de relativité traduit au niveau quantique l'équivalence entre tous les observateurs galiléens. Les générateurs infinitésimaux pour chacune des actions à un paramètre dans (12.40) sont réalisés comme opérateurs (essentiellement) autoadjoints dénotés, dans l'ordre d'apparition des paramètres, par $\mathbb{1}$, H, $\boldsymbol{P}$, $\boldsymbol{K}$. Ils forment les *observables* de base de la mécanique quantique galiléenne. Celles non nulles parmi leurs relations de commutation s'écrivent :

$$[\boldsymbol{K},H] = i\boldsymbol{P}, \quad [\boldsymbol{K},\boldsymbol{P}] = im\mathbb{1}. \tag{12.41}$$

La représentation, dite impulsion, des états quantiques avec laquelle vient d'être définie l'action du groupe de relativité est précisément celle où l'opérateur $\boldsymbol{P}$ se réalise comme opérateur de multiplication, c'est-à-dire celle où il prend une forme diagonale si l'on choisit comme « base » d'états les pics de Dirac $|\,\boldsymbol{\kappa}\,\rangle$ centrés en l'impulsion $\boldsymbol{\kappa}$:

$$\langle\,\boldsymbol{k}\,|\,\boldsymbol{\kappa}\,\rangle = \delta(\boldsymbol{k}-\boldsymbol{\kappa}), \quad \boldsymbol{P}\,|\,\boldsymbol{\kappa}\,\rangle = \boldsymbol{\kappa}\,|\,\boldsymbol{\kappa}\,\rangle. \tag{12.42}$$

Les générateurs H et $\boldsymbol{K}$ se réalisent, quant à eux, comme

$$H = \frac{\boldsymbol{P}^2}{2m} + \mathcal{U}, \tag{12.43}$$

où $\mathcal{U}$ est une constante communément appelée *énergie interne*, et

$$\boldsymbol{K} = m\,i\frac{\partial}{\partial \boldsymbol{k}}\cdot \tag{12.44}$$

Bien que les objets de « base » dans (12.42) soient des distributions et non des éléments de l'espace de Hilbert des états, il existe un formalisme, dit d'*espace nucléaire*, élaboré par l'école de Gel'fand [11] et légitimant leur caractère « normalisé à l'unité », $\langle\,\boldsymbol{\kappa}'\,|\,\boldsymbol{\kappa}\,\rangle = \delta(\boldsymbol{\kappa}'-\boldsymbol{\kappa})$, et *total* ou *surcomplet* en ce sens qu'ils résolvent l'unité :

$$\int_{-\infty}^{+\infty} |\,\boldsymbol{\kappa}\,\rangle\langle\,\boldsymbol{\kappa}\,|\,d\boldsymbol{\kappa} = \mathbb{1}. \tag{12.45}$$

Cette résolution assure la cohérence de l'égalité $\langle\,\boldsymbol{k}\,|\,\psi\,\rangle = \psi(\boldsymbol{k})$, trivialement vérifiée en intercalant (12.45) dans le premier membre et en utilisant la définition même de la distribution δ issue de (12.42).

En vérifiant par ailleurs que $|\,\boldsymbol{\kappa}\,\rangle = \exp\left(-i\boldsymbol{\kappa}\cdot\boldsymbol{K}/m\right)|\,\boldsymbol{0}_{\boldsymbol{k}}\,\rangle$, on notera aussi l'analogie formelle avec les états cohérents de $SO(2)$ introduits dans la section précédente. Ici, $|\,\boldsymbol{\kappa}\,\rangle$ est obtenu par transfert d'impulsion accordé à l'état ou *sonde* au repos $|\,\boldsymbol{0}_{\boldsymbol{k}}\,\rangle$. Ce transfert est assuré grâce à l'action unitaire du sous-groupe des boosts galiléens.

La première relation de commutation dans (12.41) impliquant le générateur $\boldsymbol{K}$ des boosts et le générateur $\boldsymbol{P}$ des translations est à la base de la localisation quantique galiléenne. En effet, l'opérateur essentiellement autoadjoint

$$\boldsymbol{Q} = \frac{\boldsymbol{K}}{m}, \quad [\boldsymbol{Q},\boldsymbol{P}] = i\mathbb{1} \tag{12.46}$$

possède toutes les propriétés idéalement requises pour être l'observable position. Mais avant de développer ce point, introduisons la représentation *position* des états du système, c'est-à-dire celle où l'opérateur $\boldsymbol{Q}$ prend sa forme diagonale et devient un opérateur de multiplication. D'une manière symétrique à la représentation impulsion, les pics de Dirac $|\,\xi\,\rangle$ normalisés et résolvant l'identité,

$$\langle\, \boldsymbol{x} \mid \xi\,\rangle = \delta(\boldsymbol{x}-\xi), \quad \int_{-\infty}^{+\infty} |\,\xi\,\rangle\langle\,\xi\mid d\xi = \mathbb{1}, \tag{12.47}$$

sont les distributions propres de $\boldsymbol{Q}$,

$$\boldsymbol{Q}\,|\,\xi\,\rangle = \xi\,|\,\xi\,\rangle, \tag{12.48}$$

et sont obtenus par translation du pic localisé à l'origine,

$$|\,\xi\,\rangle = \exp\left(-i\xi\cdot\boldsymbol{P}\right)|\,\boldsymbol{0}_{\boldsymbol{x}}\,\rangle. \tag{12.49}$$

Le passage entre les deux représentations (12.42) et (12.47) résulte d'un changement unitaire de base :

$$|\,\xi\,\rangle = \int_{-\infty}^{+\infty} \langle\,\boldsymbol{\kappa}\mid\xi\,\rangle\,|\,\boldsymbol{\kappa}\,\rangle\, d\xi = \mathbb{1}, \tag{12.50}$$

où

$$\langle\,\boldsymbol{\kappa}\mid\xi\,\rangle = \frac{1}{\sqrt{2\pi}}e^{i\boldsymbol{\kappa}\cdot\xi} = \overline{\langle\,\xi\mid\boldsymbol{\kappa}\rangle}. \tag{12.51}$$

Cela n'est rien d'autre qu'une transformation de Fourier reliant $\langle\,\boldsymbol{k}\mid\psi\,\rangle = \psi(\boldsymbol{k})$ à $\langle\,\boldsymbol{x}\mid\psi\,\rangle \equiv \phi(\boldsymbol{x})$, avec un nécessaire changement de notation pour désigner l'état en représentation position afin d'éviter toute confusion :

$$\phi(\boldsymbol{x}) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\psi(\boldsymbol{k})\,d\boldsymbol{k}, \quad \psi(\boldsymbol{k}) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}\phi(\boldsymbol{x})\,d\boldsymbol{x}. \tag{12.52}$$

En représentation position, les générateurs H et $\boldsymbol{P}$ s'expriment par

$$H = -\frac{1}{2m}\frac{d^2}{d\boldsymbol{x}^2} + \mathcal{U}, \quad \boldsymbol{P} = -i\frac{\partial}{\partial\boldsymbol{x}}, \tag{12.53}$$

et la représentation unitaire R_m opère maintenant comme suit :

$$\left(\hat{R}_m(\theta, b, \boldsymbol{a}, \boldsymbol{v})\phi\right)(\boldsymbol{x}) = e^{i(\theta+(\boldsymbol{P}^2/2m)b+m\boldsymbol{v}\cdot(\boldsymbol{x}-\boldsymbol{a}))}\phi(\boldsymbol{x}-\boldsymbol{a}). \tag{12.54}$$

La présence de l'opérateur $\boldsymbol{P}^2/2m$ dans le facteur exponentiel traduit l'intervention de l'opérateur d'évolution temporelle $\exp\left(-iH(-b)\right)$ agissant sur un état $\phi(\boldsymbol{x})$ a priori *intemporel*, un point de vue *à la Heisenberg*. Il est donc logique d'adopter le point de vue *à la Schrödinger* et d'accorder à $|\,\phi\,\rangle$ un développement temporel en le considérant comme l'état du système à l'instant

$t = 0$, par exemple, et en faisant agir l'opérateur d'évolution $\exp(-iHt)$ (à un facteur de phase près) afin de définir l'état à l'instant t :

$$\Phi(\boldsymbol{x}, t) = e^{-i(\boldsymbol{P}^2/2m)t}\Phi(\boldsymbol{x}, 0), \;\; \Phi(\boldsymbol{x}, 0) \equiv \phi(\boldsymbol{x}). \tag{12.55}$$

La représentation du groupe de Galilée sur les états $\Phi(\boldsymbol{x}, t)$ s'exprime alors d'une manière plus naturelle :

$$\left(\hat{R}_m(\theta, b, \boldsymbol{a}, \boldsymbol{v})\Phi\right)(\boldsymbol{x}, t) = e^{i(\theta-(m\boldsymbol{v}^2/2)(t-b)+m\boldsymbol{v}\cdot(\boldsymbol{x}-\boldsymbol{a}))}\Phi(\boldsymbol{x} - \boldsymbol{a} - \boldsymbol{v}(t-b), t-b). \tag{12.56}$$

Les résolutions (12.45) et (12.47) donnent un sens précis à la localisation respectivement en quantité de mouvement et en position dans le cadre de la mécanique quantique galiléenne. Considérons par exemple (12.47) et restreignons l'intégration à un ensemble de Borel $\Delta \subset \mathbb{R}$. L'opérateur

$$P(\Delta) = \int_\Delta |\, \boldsymbol{\xi} \,\rangle\langle\, \boldsymbol{\xi} \,|\, d\boldsymbol{\xi} \tag{12.57}$$

est un projecteur. En effet, on vérifie aisément que

$$P(\Delta)\phi(\boldsymbol{x}) = \chi_\Delta(\boldsymbol{x})\phi(\boldsymbol{x}), \tag{12.58}$$

où $\chi_\Delta(\boldsymbol{x})$ est la fonction caractéristique de Δ. C'est en fait la mesure à valeur projecteur (on dira *PV measure* en anglais) qui réalise la résolution spectrale de l'opérateur position $\boldsymbol{Q}$:

$$\boldsymbol{Q} = \int_{-\infty}^{+\infty} \boldsymbol{\xi} \,|\, \boldsymbol{\xi} \,\rangle\langle\, \boldsymbol{\xi} \,|\, d\boldsymbol{\xi} \equiv \int_{-\infty}^{+\infty} \boldsymbol{\xi}\, P(d\boldsymbol{\xi}). \tag{12.59}$$

Ainsi, $P(\Delta)$ est un opérateur de localisation spatiale pour l'espace des états du système. Sa valeur moyenne dans un état $|\, \psi \,\rangle$,

$$\langle\, \psi \,|\, P(\Delta) \,|\, \psi \,\rangle = \mathcal{P}_\psi(\Delta), \tag{12.60}$$

s'interprète comme la probabilité de localiser le système dans Δ si lui-même se trouve dans l'état ψ, et cette localisation est *nette* ou *binaire* en ce sens que $P(\Delta)$ est un projecteur et donc implique une logique binaire. De plus, la famille d'opérateurs $P(\Delta)$ possède une propriété de covariance sous l'action du sous-groupe des translations d'espace :

$$R_m(0, 0, \boldsymbol{a}, \boldsymbol{0})P(\Delta)R_m(0, 0, -\boldsymbol{a}, \boldsymbol{0}) = P(\Delta + \boldsymbol{a}). \tag{12.61}$$

Une caractéristique importante de cette localisation binaire est sa parfaite instabilité du point de vue de son évolution temporelle : un état localisé à un instant donné dans une région bornée du plan s'étale à l'infini immédiatement après. Il suffit pour admettre cela de considérer l'évolution temporelle d'un état localisé, à l'instant $t = 0$, dans l'intervalle $[0, 1]$, et dont la fonction d'onde serait justement à cet instant : $\Phi(\boldsymbol{x}, 0) = \chi_{[0,1]}(\boldsymbol{x})$. On montre grâce à des résultats classiques sur la transformation de Fourier que $\Phi(\boldsymbol{x}, t)$ est donné pour $t > 0$ par

$$\Phi(\boldsymbol{x}, t) = \sqrt{\frac{im}{2\pi t}} \int_0^1 \exp\left(-im\frac{(\boldsymbol{x} - \boldsymbol{x}')^2}{t}\right) d\boldsymbol{x}' \tag{12.62}$$

et ne peut être à support borné.

Une mise au point s'impose ici au regard de confusions rencontrées dans certaines interprétations de la mécanique quantique. Cet étalement « brutal » de la fonction d'onde ne viole en rien la causalité, puisque, dans le contexte galiléen, il n'y a aucune limite supérieure imposée à la vitesse de propagation d'une information : le cône du futur est le demi-plan supérieur tout entier ! La situation (et les difficultés) sont évidemment tout autres dans le contexte de la mécanique quantique einsteinienne, comme on le verra dans les sections suivantes.

La paire (R_m, P) est dite *système d'imprimitivité* [18], dans la mesure où la représentation R_m est elle-même induite à partir de la représentation unitaire irréductible du sous-groupe des translations (que l'on devrait compléter avec les rotations dans le cas tridimensionnel).

Toujours dans ce contexte de mécanique quantique galiléenne, il existe aussi un concept de localisation dans l'espace des phases X (12.38). Cette localisation est toutefois non binaire, car assujettie à la contrainte à la Heisenberg $\Delta\boldsymbol{Q}\,\Delta\boldsymbol{P} \geq 1/2$ imposée par la relation canonique de commutation (12.46). Ce concept se bâtit à l'aide d'états cohérents pour le groupe de Galilée étendu [1, 4] et conduit à la notion de *système de covariance*, plus générale que celle de système d'imprimitivité.

Comme suggéré dans (12.38), l'espace des phases $X = \tilde{\mathcal{G}}(1, 1)/(\Theta \times \mathcal{T})$ possède un système de coordonnées globales $(\boldsymbol{q}, \boldsymbol{p}) \equiv (\boldsymbol{a} - \boldsymbol{v}b, m\boldsymbol{v}) \in \mathbb{R}^2$ pour lesquelles la mesure de Lebesgue $d\boldsymbol{q}\, d\boldsymbol{p}$ est invariante sous l'action de $\tilde{\mathcal{G}}(1, 1)$. Considérons alors la *section* $\sigma_0 : X \to \tilde{\mathcal{G}}$ définie par

$$\sigma_0(\boldsymbol{q}, \boldsymbol{p}) = \left(0, 0, \boldsymbol{q}, \frac{\boldsymbol{p}}{m}\right). \tag{12.63}$$

Choisissons un état *sonde* $|\,\eta\,\rangle$ normalisé, $\|\eta\| = 1$, du système, et transportons le à l'aide de cette section en faisant varier $(\boldsymbol{q}, \boldsymbol{p})$ dans $\mathbb{R}^2$. Nous obtenons ainsi une famille de vecteurs, dits états cohérents pour le groupe de Galilée (étendu),

$$|\,\eta_{\sigma_0(\boldsymbol{q},\boldsymbol{p})}\,\rangle = R_m(\sigma_0(\boldsymbol{q},\boldsymbol{p})\,|\,\eta\,\rangle, \tag{12.64}$$

dont on montre, grâce à l'unitarité et à l'irréductibilité de la représentation R_m, qu'ils résolvent l'unité [1, 4],

$$\frac{1}{2\pi}\int_X |\,\eta_{\sigma_0(\boldsymbol{q},\boldsymbol{p})}\,\rangle\langle\,\eta_{\sigma_0(\boldsymbol{q},\boldsymbol{p})}\,|\,d\boldsymbol{q}\,d\boldsymbol{p} = \mathbb{1}. \tag{12.65}$$

D'une manière analogue à (12.104), nous pouvons maintenant définir sur la σ-algèbre $\mathcal{B}(X)$ des boréliens de l'espace des phases une mesure à valeurs opérateur positif :

$$\mathcal{B}(X) \ni \Delta \to a_{\sigma_0}(\Delta) = \frac{1}{2\pi}\int_\delta |\,\eta_{\sigma_0(\boldsymbol{q},\boldsymbol{p})}\,\rangle\langle\,\eta_{\sigma_0(\boldsymbol{q},\boldsymbol{p})}\,|\,d\boldsymbol{q}\,d\boldsymbol{p} \equiv \int_\Delta a_{\sigma_0}(d\boldsymbol{q}\,d\boldsymbol{p}). \tag{12.66}$$

Nous avons bien ici toutes les propriétés d'une mesure :

$$\begin{aligned} &a_{\sigma_0}(\emptyset) = 0, \quad a_{\sigma_0}(X) = \mathbb{1}, \\ &a_{\sigma_0}(\cup_{i=0}^{\infty}\Delta_i) = \sum_{i=0}^{\infty} a_{\sigma_0}(\Delta_i), \quad \Delta_i \cap \Delta_j = 0,\ i \neq j, \end{aligned} \tag{12.67}$$

où la convergence des séries doit se comprendre au sens faible dans l'espace de Hilbert des états. Examinons alors la propriété de covariance de ces opérateurs sous l'action de $\tilde{\mathcal{G}}(1,1)$. Elle résulte de la covariance des états cohérents :

$$\begin{aligned} &R_m(\theta, b, \boldsymbol{a}, \boldsymbol{v})\,|\,\eta_{\sigma_0(\boldsymbol{q},\boldsymbol{p})}\,\rangle = e^{i(\theta+(\boldsymbol{k}^2/2m)b)}\,|\,\eta_{\sigma_0(\boldsymbol{q}+\boldsymbol{a},\boldsymbol{p}+m\boldsymbol{v}s)}\,\rangle \\ &R_m(\theta, b, \boldsymbol{a}, \boldsymbol{v})a_{\sigma_0}(\Delta)R_m^{\dagger}(\theta, b, \boldsymbol{a}, \boldsymbol{v}) = a_{\sigma_0}(\Delta + (\boldsymbol{a}, m\boldsymbol{v})). \end{aligned} \tag{12.68}$$

On dira en conséquence que les couples $(R_m, a_{\sigma_0}(\Delta))$, pour Δ variant dans $\mathcal{B}(X)$, forment un système de covariance pour le groupe de Galilée étendu.

Du fait de la résolution de l'identité, la famille des états cohérents (12.64) forment un repère continu strict pour l'espace $L^2(\mathbb{R})$ des états du système. mais ce repère nous offre de plus une méthode de quantification *canonique* des plus naturelles. Rappelons brièvement ce qu'on entend par quantification canonique. Partant de l'espace des phases X de la mécanique classique, variété symplectique munie de sa 2-forme $d\boldsymbol{q} \wedge d\boldsymbol{p}$, il s'agit d'établir une correspondance entre *observables classiques*, fonctions $f(\boldsymbol{q},\boldsymbol{p})$ ayant un certain degré de régularité, et *observables quantiques*, opérateurs F (essentiellement) auto-adjoints agissant dans un espace de Hilbert (projectif) $\mathcal{H}$ dit espace des états quantiques du système. Cette correspondance doit obéir aux règles suivantes.

(i) À l'observable $f(\boldsymbol{q},\boldsymbol{p}) = 1$ correspond l'opérateur identité $\mathbb{1}$ sur $\mathcal{H}$.
(ii) La correspondance $f \mapsto F$ est linéaire.
(iii) Au crochet de Poisson classique correspond, au moins à l'ordre $\hbar$, le commutateur quantique multiplié par $i\hbar$

$$\{f_1, f_2\} = f_3 \mapsto [F_1, F_2] = i\hbar F_3 + O(\hbar^2). \tag{12.69}$$

(iv) Certaines conditions de minimalité imposées à l'algèbre des observables quantiques images des observables classiques, en fait remplies par et seulement par la représentation habituelle de Schrödinger pour la mécanique quantique galiléenne !

Examinons si la correspondance

$$f \mapsto F = \frac{1}{2\pi}\int_X f(\boldsymbol{q},\boldsymbol{p}) \mid \eta_{\sigma_0(\boldsymbol{q},\boldsymbol{p})}\,\rangle\langle\, \eta_{\sigma_0(\boldsymbol{q},\boldsymbol{p})} \mid d\boldsymbol{q}\,d\boldsymbol{p} \equiv \int_\Delta f(\boldsymbol{q},\boldsymbol{p})\, a_{\sigma_0}(d\boldsymbol{q}\,d\boldsymbol{p}) \tag{12.70}$$

satisfait justement à ces règles. Les deux premières conditions sont trivialement remplies. Une indication de preuve pour (iii) est fournie par la considération de ces deux cas fondamentaux que sont les observables $\boldsymbol{q}$ et $\boldsymbol{p}$. Si l'on suppose réelle, $\bar{\eta} = \eta$, ou encore à parité définie, $\eta(-\boldsymbol{k}) = \pm\eta(\boldsymbol{k})$, la fonction sonde $\eta(\boldsymbol{k})$ choisie pour construire le repère, alors les quantités

$$\boldsymbol{Q} = \int_\Delta \boldsymbol{q}\, a_{\sigma_0}(d\boldsymbol{q}\,d\boldsymbol{p}),\ \ \boldsymbol{P} = \int_\Delta \boldsymbol{p}\, a_{\sigma_0}(d\boldsymbol{q}\,d\boldsymbol{p}), \tag{12.71}$$

sont précisément les opérateurs position et quantité de mouvement étudiés ci-dessus. On peut ainsi affirmer que le point de vue sur l'espace des phases X offert par le repère des états cohérents **est** le point de vue quantique.

En fait, on peut deviner que la valeur moyenne $\langle\, \psi \mid a_{\sigma_0}(\Delta) \mid \psi\,\rangle$ s'interprète comme la probabilité de localisation du système dans la région $\Delta \subset X$ lorsque son état est ψ. Le repère (12.65) nous offre ainsi la possibilité de formuler la mécanique quantique galiléenne sur l'espace des phases X. Il suffit d'introduire le plongement isométrique $W\,:\ L^2(\mathbb{R}, d\boldsymbol{k}) \rightarrow L^2(X, d\boldsymbol{q}\,d\boldsymbol{p})$ donné par

$$(W\psi)(\boldsymbol{q},\boldsymbol{p}) = \langle\, \eta_{\sigma_0(\boldsymbol{q},\boldsymbol{p})} \mid \psi\,\rangle \equiv \Psi(\boldsymbol{q},\boldsymbol{p}). \tag{12.72}$$

On parlera alors de fonction d'onde $\Psi(\boldsymbol{q},\boldsymbol{p})$ et de densité de probabilité $\|\Psi(\boldsymbol{q},\boldsymbol{p})\|^2$ du système dans sa représentation *espace des phases*. L'évolution temporelle en sera donnée par

$$\Psi(\boldsymbol{q},\boldsymbol{p},t) = \langle\, \eta_{\sigma(\boldsymbol{q},\boldsymbol{p})} \mid e^{-iHt}\psi\,\rangle. \tag{12.73}$$

Nous verrons en fin d'article comment ce formalisme passe « en douceur » les obstacles imposés par la relativité einsteinienne.

4. Système élémentaire einsteinien

Les éléments du groupe cinématique de l'espace-temps de Minkowski à $1+1$ dimensions, c'est-à-dire du groupe de Poincaré $\mathcal{P}_+^{\uparrow}(1,1)$, $+$ pour *propre*, $\uparrow$ pour *orthochrone*, s'écrivent

$$g = (a, h(p)) \in \mathbb{R}^2 \rtimes SO_0(1,1). \tag{12.74}$$

Ils obéissent à la loi de multiplication $gg' = (a + h(p)a', h(p)h(p'))$ héritée de la structure en produit semi-direct du sous-groupe $\mathbb{R}^2$ des translations temps-espace $a = (a_0, \boldsymbol{a})$ avec le sous-groupe de Lorentz $SO_0(1,1)$ qui se réduit en $1+1$ dimensions aux *boosts* :

$$h(p) = \begin{pmatrix} p_0/m & \boldsymbol{p}/m \\ \boldsymbol{p}/m & p_0/m \end{pmatrix}, \ p_0 = \sqrt{m^2 + \boldsymbol{p}^2}. \tag{12.75}$$

On a posé ici et pour le reste de l'article $c = 1$. On notera que $h^{-1}(p_0, \boldsymbol{p}) = h(p_0, -\boldsymbol{p})$. Un paramètre de masse a été introduit dans (12.75), et nous considèrerons le bivecteur $(p_0, \boldsymbol{p})$ comme un paramètre de groupe évoluant sur l'hyperbole de masse $\mathcal{V}_m^+ = \{(p_0, \boldsymbol{p}) \in \mathbb{R}^2, \mid \ p_0^2 - \boldsymbol{p}^2 = m^2\}$. Ainsi, la matrice (12.75) agit sur l'espace des bivecteurs $k = (k_0, \boldsymbol{k}) \in \mathcal{V}_m^+$ écrits sous forme de vecteurs colonne. Nous ferons par la suite principalement usage de cette représentation, de préférence à celle bâtie sur le paramètre de rapidité,

$$h(p) \equiv \Lambda(\phi) = \begin{pmatrix} \cosh\phi & \sinh\phi \\ \sinh\phi & \cosh\phi \end{pmatrix}, \ \tanh\phi = \frac{\boldsymbol{p}}{p_0}. \tag{12.76}$$

Les états quantiques d'un système élémentaire einsteinien de masse m forment l'espace de Hilbert $L^2(\mathcal{V}_m^+, d\boldsymbol{k}/k_0)$ des fonctions $\psi(k)$ de carré intégrable sur l'hyperbole de masse munie de la mesure invariante de Lorentz $d\boldsymbol{k}/k_0$. Cet espace porte la représentation unitaire irréductible U_m, dite de Wigner, définie par

$$(U_m(a, h(p))\psi)(k) = e^{ik\cdot a}\psi(h^{-1}(p)k), \tag{12.77}$$

où $k \cdot a = k_0 a_0 - \boldsymbol{k} \cdot \boldsymbol{a}$. Dans cette représentation, les trois générateurs infinitésimaux respectivement pour les translations de temps, d'espace, et pour les boosts se réalisent de la manière suivante :

$$P_0\psi(k) = k_0\psi(k), \ \boldsymbol{P}\psi(k) = \boldsymbol{k}\psi(k), \ \boldsymbol{K}\psi(k) = ik_0\frac{d}{d\boldsymbol{k}}\psi(k), \tag{12.78}$$

et obéissent aux relations de commutation,

$$[P_0, \boldsymbol{P}] = \boldsymbol{0}, \; [\boldsymbol{K}, P_0,] = i\boldsymbol{P}, \tag{12.79}$$

$$[\boldsymbol{K}, \boldsymbol{P}] = iP_0. \tag{12.80}$$

La troisième de ces relations nous laisse deviner ce que pourrait être un opérateur position de type « galiléen ». En effet, en utilisant le fait que P_0 est un opérateur positif, inversible, et qui commute avec $\boldsymbol{P}$, nous pouvons réécrire (12.80) sous la forme d'une relation canonique de commutation :

$$[\boldsymbol{Q}_{NW}, \boldsymbol{P}] = i\mathbb{1}, \text{ où } \boldsymbol{Q}_{NW} \equiv P_0^{-1/2}\boldsymbol{K}P_0^{-1/2}. \tag{12.81}$$

$\boldsymbol{Q}_{NW}$ est l'opérateur position que Newton et Wigner ont proposé dans [19] pour définir et asseoir leur axiomatique de localisation spatiale en mécanique quantique einsteinienne, un concept approfondi par Wightman dans [22]. Revoyons plus en détail cette localisation. On remarquera tout d'abord la manière dont $\boldsymbol{Q}_{NW}$ se transforme sous l'action des translations spatio-temporelles du groupe de Poincaré :

(i) Translation d'espace,

$$e^{i\boldsymbol{a}\cdot\boldsymbol{P}}\boldsymbol{Q}_{NW}e^{-i\boldsymbol{a}\cdot\boldsymbol{P}} = \boldsymbol{Q}_{NW} + \boldsymbol{a}\mathbb{1}. \tag{12.82}$$

(ii) Translation de temps,

$$e^{ia_0P_0}\boldsymbol{Q}_{NW}e^{-ia_0P_0} = \boldsymbol{Q}_{NW} - \frac{\boldsymbol{P}}{P_0}a_0. \tag{12.83}$$

Ainsi, un état « localisé » en $\boldsymbol{0}$ le serait en $\boldsymbol{a}$ après translation spatiale de $\boldsymbol{a}$, et le serait, après translation temporelle de a_0, de $\boldsymbol{V}a_0$ où $\boldsymbol{V} = \boldsymbol{P}/P_0$ représente l'opérateur vitesse. On constate un accord formel avec la situation galiléenne décrite dans la section précédente. Il est alors naturel dans le présent contexte de définir l'espace comme l'ensemble des valeurs spectrales de l'opérateur (essentiellement) auto-adjoint $\boldsymbol{Q}_{NW}$, et d'introduire l'espace des états sur lesquels cet opérateur se réalise comme opérateur de multiplication. Cela s'effectue grâce à l'isométrie hilbertienne,

$$\mathcal{F} : L^2\left(\mathcal{V}_m^+, \frac{d\boldsymbol{k}}{k_0}\right) \ni \psi \to \phi \in L^2(\mathbb{R}, d\boldsymbol{x}), \tag{12.84}$$

définie par

$$\phi(\boldsymbol{x}) = (\mathcal{F}\psi)(\boldsymbol{x}) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\psi(k)\,\frac{d\boldsymbol{k}}{k_0}, \tag{12.85}$$

et qui n'est rien d'autre qu'une transformation de Fourier « pondérée ». On vérifie aisément l'action de l'opérateur transformé $\hat{\boldsymbol{Q}} = \mathcal{F}\boldsymbol{Q}_{NW}\mathcal{F}^{-1}$:

$$\hat{\boldsymbol{Q}}\phi(\boldsymbol{x}) = \boldsymbol{x}\phi(\boldsymbol{x}). \tag{12.86}$$

L'action transformée de la représentation de Wigner est elle-même donnée sur l'espace des fonctions $\Phi(x)$ de la variable minkowskienne $x = (x_0, \boldsymbol{x})$ construit par évolution temporelle des fonctions $\phi(\boldsymbol{x})$:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-ik\cdot x}\psi(k)\,\frac{d\boldsymbol{k}}{k_0} = e^{-ix_0P_0}\phi(\boldsymbol{x}) \tag{12.87}$$

$$\hat{U}_m = \mathcal{F}U_m\mathcal{F}^{-1},\ \ (\hat{U}_m(a, h(p))\Phi)(x) = \Phi(h(p)^{-1}(x-a)). \tag{12.88}$$

Désignons par $P(d\boldsymbol{x})$ la mesure ou *résolution* spectrale de l'opérateur $\hat{\boldsymbol{Q}}$:

$$\hat{\boldsymbol{Q}} = \int_{-\infty}^{+\infty} \boldsymbol{x}\, P(d\boldsymbol{x}). \tag{12.89}$$

Cette mesure définit la famille d'opérateurs à valeur projecteur $\{P(\Delta)\}$, où Δ varie dans l'ensemble des boréliens de $\mathbb{R}$:

$$\left(\int_\Delta P(d\boldsymbol{y})\right)\phi(\boldsymbol{x}) \equiv P(\Delta)\phi(\boldsymbol{x}) = \chi_\Delta(\boldsymbol{x})\phi(\boldsymbol{x}). \tag{12.90}$$

Cette mesure à valeur opérateur hérite de (12.88) la propriété de covariance sous l'action des translations d'espace $\hat{U}_m((0,\boldsymbol{a}), \mathbb{1}_2) = \exp(-i\boldsymbol{a}.\hat{\boldsymbol{P}})$

$$\exp(-i\boldsymbol{a}.\hat{\boldsymbol{P}})\, P(\Delta)\, \exp(i\boldsymbol{a}.\hat{\boldsymbol{P}}) = P(\Delta - \boldsymbol{a}). \tag{12.91}$$

Ainsi, comme dans le cas galiléen, la famille des $P(\Delta)$ avec la restriction de la représentation de Wigner $\hat{U}_m$ au sous-groupe des translations spatiales constituent un sytème d'imprimitivité.

Ce schéma de localisation à la Newton-Wigner, certes séduisant par sa simplicité toute galiléenne, ne résiste néanmoins pas à une analyse plus approfondie de toutes ses implications au plan d'une théorie à la fois causale et quantique. Par exemple, la valeur moyenne $\langle\,\phi \mid P(\Delta) \mid \phi\,\rangle$ ne peut être interprétée comme la probabilité de trouver le système dans la région Δ. Ou encore, $\|\Phi(t,\boldsymbol{x})\|^2$ n'est en aucune manière la composante temporelle d'un courant de probabilité conservé. La section suivante en apporte une preuve rigoureuse.

5. Les difficultés de la localisation à la Newton-Wigner

La problématique, en mécanique quantique einsteinienne, d'une localisation spatiale qui soit exempte d'incohérences ou au mieux d'ambiguïtés a une longue histoire, bien illustrée par la bibliographie donnée par Kálnay dans son article de revue publié en 1971 [14]. Il ressort de ce dernier travail que, si l'on souhaite retenir dans le cadre einsteinien tous ces critères qui font de l'opérateur position un objet inattaquable dans le cadre galiléen, l'on est inévitablement amené à assumer des contradictions. Parmi les plus sérieuses, on doit retenir l'incompatibilité entre une localisation stricte *à la* Newton-Wigner dans une région finie de l'espace et le principe de causalité. C'est justement le contenu du théorème de Hegerfeldt [12], un résultat dont la simplicité égale l'importance et que nous allons maintenant exposer. Avec les notations de la section précédente, posons

$$\psi_t = e^{-itP_0}\ \psi_{t=0}. \tag{12.92}$$

Considérons une région Δ bornée de l'« espace » $\mathbb{R}$ constitué par l'ensemble des valeurs spectrales de l'opérateur de Newton-Wigner $\mathbf{Q}_{NW}$ et partons des hypothèses minimales suivantes.

(i) **Localisation**

Il existe un opérateur $\mathcal{N}(\Delta)$ tel que la valeur moyenne

$$\langle\, \psi_t \mid \mathcal{N}(\Delta) \mid \psi_t \,\rangle \equiv \mathcal{P}_{\psi,t}(\Delta) \tag{12.93}$$

représente la probabilité de présence du système dans la région Δ à l'instant t.

Donnons ici deux commentaires.

1. Cette interprétation entraîne $0 \leq \mathcal{P}_{\psi,t}(\Delta) \leq 1$ et donc que $\mathcal{N}(\Delta)$ est autoadjoint et borné : $0 \leq \mathcal{N}(\Delta) \leq 1$.
2. Elle donne lieu à une définition de localisation binaire : le système, ou plutôt son état, est dit localisé dans Δ à l'instant t si $\mathcal{P}_{\psi,t}(\Delta) = 1$, et il est dit non localisé dans Δ à l'instant t si $\mathcal{P}_{\psi,t}(\Delta) = 0$.

(ii) **Causalité**

Si, à l'instant $t_0 = 0$, un état est localisé dans Δ, alors il existe un nombre r_t tel que, à l'instant $t > 0$, le translaté du système par tout $\boldsymbol{a}$, $|\boldsymbol{a}| \geq r_t$, est *non*-localisé dans Δ.

Traduisons cela en termes plus mathématiques.

Si $\langle\, \psi_{t=0} \mid \mathcal{N}(\Delta) \mid \psi_{t=0} \,\rangle = 1$, alors $\exists r_t$ tel que $\forall \boldsymbol{a}$, $|\boldsymbol{a}| \geq r_t$, on a $\langle\, U(-\boldsymbol{a})\psi_t \mid \mathcal{N}(\Delta) \mid U(-\boldsymbol{a})\psi_t \,\rangle = 0$, où $U(-\boldsymbol{a})$ est une notation condensée pour $U((0, -\boldsymbol{a}), I_2)$.

Nous sommes maintenant en mesure d'énoncer le résultat d'Hegerfeldt :

Théorème 12.1. *Supposons que l'assertion de localisation (i) soit vraie. Alors, il n'y a aucun état de système élémentaire einsteinien localisé à l'instant $t = 0$ dans la région finie Δ qui puisse satisfaire à la condition de causalité (ii).*

Démonstration. Supposons l'état $|\, \psi_{t=0} \,\rangle$, $\|\psi_{t=0}\| = 1$, localisé dans Δ : $\langle\, \psi_{t=0} \mid \mathcal{N}(\Delta) \mid \psi_{t=0} \,\rangle = 1$. Du fait de $0 \leq \mathcal{N}(\Delta) \leq 1$ et de l'inégalité de Schwarz, l'état $|\, \psi_{t=0} \,\rangle$ est vecteur propre de $\mathcal{N}(\Delta)$ pour la valeur propre 1 :

$$\mathcal{N}(\Delta) \mid \psi_{t=0} \,\rangle = \mid \psi_{t=0} \,\rangle. \tag{12.94}$$

Si $|\boldsymbol{a}| \geq r_t$ et $|\, \psi_t \,\rangle = U(t) \mid \psi_{t=0} \,\rangle$, alors la condition de causalité (ii) implique

$$\langle\, U(-\boldsymbol{a})U(t)\psi_{t=0} \mid \mathcal{N}(\Delta) \mid U(-\boldsymbol{a})U(t)\psi_{t=0} \,\rangle = 0.$$

Comme $(\mathcal{N}(\Delta))^{1/2}$ est aussi autoadjoint, cela entraîne :

$$(\mathcal{N}(\Delta))^{1/2}\, U(-\boldsymbol{a})U(t) \mid \psi_{t=0} \,\rangle = 0,$$

et donc $\mathcal{N}(\Delta)U(-\boldsymbol{a})U(t) \mid \psi_{t=0} \,\rangle = 0$, et donc encore $\langle\, \psi_{t=0} \mid \mathcal{N}(\Delta)U(-\boldsymbol{a})U(t) \mid \psi_{t=0} \,\rangle = 0$. En utilisant (12.94), cela nous donne finalement la condition :

$$\begin{aligned} 0 &= \langle\, \psi_{t=0} \mid U(-\boldsymbol{a})U(t) \mid \psi_{t=0} \,\rangle \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-i\boldsymbol{k}\boldsymbol{a}} e^{-ik_0 t} |\psi_{t=0}(k)|^2 \, \frac{d\boldsymbol{k}}{k_0}, \ \forall \boldsymbol{a},\ |\boldsymbol{a}| \geq r_t. \end{aligned} \tag{12.95}$$

Autrement dit, la transformée de Fourier $G(\boldsymbol{a})$ de la fonction

$$f(\boldsymbol{k}) \equiv \frac{e^{-ik_0 t}|\psi_{t=0}(k)|^2}{k_0}$$

est nulle pour $|\boldsymbol{a}| \geq r_t$. Mais alors, $f(\boldsymbol{k}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{i\boldsymbol{k}\boldsymbol{a}} G(\boldsymbol{a})\, d\boldsymbol{a}$, en tant que transformée de Fourier (inverse) de la distribution à support borné $G(\boldsymbol{a})$, doit pouvoir s'étendre à une fonction analytique en la variable $\boldsymbol{k}$ complexe. Or ceci est impossible à cause de la présence de $k_0 = \sqrt{\boldsymbol{k}^2 + m^2}$ dans le facteur exponentiel, le dénominateur pouvant être quant à lui une singularité éliminable.

□

Des résultats similaires sont obtenus pour des régions Δ infinies dont la distance à leurs translatées respectives devient arbitrairement grande, et aussi pour des régions d'espace-temps et non plus seulement d'espace [12, 13].

En conclusion, il est impossible de préparer un état à *un* système élémentaire qui soit strictement localisé dans une région finie Δ. En particulier, $\mathcal{N}(\Delta)$ ne peut être un projecteur, à la différence de la mécanique quantique galiléenne. Toute localisation binaire à la Newton-Wigner porte en elle une violation de la causalité.

6. L'alternative : localisation dans l'espace des phases

D'une manière analogue à la théorie galiléenne, l'espace des phases du système élémentaire einsteinien de masse m peut être considéré comme l'espace quotient

$$X = \mathcal{P}_+^{\uparrow}(1,1)/T, \tag{12.96}$$

où T est le sous-groupe des translations de temps. Cet espace possède une paramétrisation globale $(\boldsymbol{q}, \boldsymbol{p}) \in \mathbb{R}^2$ de type « position-quantité de mouvement » pour laquelle $d\boldsymbol{q}\, d\boldsymbol{p}$ est une mesure invariante. La première étape d'une quantification ou plutôt d'une analyse quantique de l'ensemble X *via* un repère d'états cohérents consiste dans le choix d'une section (borélienne) $\sigma : X \to \mathcal{P}_+^{\uparrow}(1,1)$ *affine* :

$$\sigma(\boldsymbol{q}, \boldsymbol{p}) = \sigma_0(\boldsymbol{q}, \boldsymbol{p})((f(\boldsymbol{q}, \boldsymbol{p}), \mathbf{0}), I_2), \tag{12.97}$$

où $\sigma_0(\boldsymbol{q}, \boldsymbol{p}) \equiv ((0, \boldsymbol{q}), h(p))$ est la section à temps nul ou encore *galiléenne* et $f(\boldsymbol{q}, \boldsymbol{p}) = \boldsymbol{q}\vartheta(\boldsymbol{p}) + \varphi(\boldsymbol{p})$. La fonction $\varphi(\boldsymbol{p})$ est arbitraire tandis que $\vartheta(\boldsymbol{p})$ est contraint à satisfaire l'inégalité

$$|\beta(\boldsymbol{p})| < 1, \text{ où } \beta(\boldsymbol{p}) = \frac{p_0\, \vartheta(\boldsymbol{p})\, \mathrm{sgn}(\boldsymbol{p})}{m + \boldsymbol{p}\, \vartheta(\boldsymbol{p})}. \tag{12.98}$$

La quantité $\beta(\boldsymbol{p})$ a une signification profonde en terme de *chronogéométrie* minkowskienne. Écrivant (12.97) sous la forme d'un élément de groupe, $\sigma(\boldsymbol{q}, \boldsymbol{p}) = (\hat{q}, h(p))$, et dans la « jauge » où $\varphi(\boldsymbol{p})$ est nulle, on a la relation $\hat{q}_0 = \beta(\boldsymbol{p})\, \mathrm{sgn}(\boldsymbol{p})\, \hat{\boldsymbol{q}}$. Ainsi, la contrainte (12.98) indique que l'axe des $\hat{\boldsymbol{q}}$ est de genre espace pour tout $\boldsymbol{p}$. On parlera aussi bien pour σ d'une section de genre espace. Une discussion plus approfondie de ces aspects se trouve dans [4] dans le cas à 1 + 1 dimensions) et dans [2, 5] dans le cas à 3 + 1 dimensions.

Choisissons dans l'espace de Hilbert de représentation de Wigner un état *à énergie finie* :

$$\eta \in L^2(\mathcal{V}_m^+, d\boldsymbol{k}/k_0), \ \langle\, \eta \mid P_0 \mid \eta\, \rangle = \int_{-\infty}^{\infty} |\eta(k)|^2 \, d\boldsymbol{k} < \infty. \tag{12.99}$$

Il nous servira de sonde dont le transport à l'aide de la section σ nous fournira une famille d'états :

$$\eta_\sigma(\boldsymbol{q},\boldsymbol{p}) \equiv U_m(\sigma(\boldsymbol{q},\boldsymbol{p}))\eta, \quad (\boldsymbol{q},\boldsymbol{p}) \in X. \tag{12.100}$$

On montre, grâce au caractère spatial de la section, que ces états résolvent en général non pas l'identité, mais un opérateur A_σ positif, borné, à inverse borné :

$$\int_X \mid \eta_{\sigma(\boldsymbol{q},\boldsymbol{p})}\, \rangle\langle\, \eta_{\sigma(\boldsymbol{q},\boldsymbol{p})}\, d\boldsymbol{q}\, d\boldsymbol{p} = A_\sigma. \tag{12.101}$$

On obtient donc un repère continu et en général non strict. On parlera malgré tout d'états cohérents pour le groupe de Poincaré. Il existe des sections pour lesquelles le repère est strict, c'est-à-dire pour lesquelles A_σ est proportionnel à l'identité ; c'est le cas, par exemple, de la section galiléenne σ_0. Par contre, il en existe d'autres pour lesquelles le repère n'est jamais strict, quel que soit l'état sonde η choisi.

Sélectionnons une section σ pour laquelle le repère est strict. En renormalisant au besoin η, on pourra donc écrire :

$$\int_X \mid \eta_{\sigma(\boldsymbol{q},\boldsymbol{p})}\, \rangle\langle\, \eta_{\sigma(\boldsymbol{q},\boldsymbol{p})}\, d\boldsymbol{q}\, d\boldsymbol{p} = \mathbb{1}. \tag{12.102}$$

Pour un élément arbitraire $(a, h(k))$ du groupe de Poincaré, la section transformée,

$$\sigma_{(a,h(k)}(\boldsymbol{q},\boldsymbol{p}) \equiv (a, h(k)\sigma((a, h(k))^{-1}(\boldsymbol{q},\boldsymbol{p})), \tag{12.103}$$

fournit aussi un repère strict vérifiant (12.102). Si l'on restreint l'intégration dans (12.102) à un sous-ensemble de Borel $\Delta \subset X$, on obtient l'opérateur borné positif

$$a_\sigma(\Delta) = \int_\Delta \mid \eta_{\sigma(\boldsymbol{q},\boldsymbol{p})}\, \rangle\langle\, \eta_{\sigma(\boldsymbol{q},\boldsymbol{p})}\, d\boldsymbol{q}\, d\boldsymbol{p} \equiv \int_\Delta a_\sigma(d\boldsymbol{q}\, d\boldsymbol{p}), \tag{12.104}$$

et l'application $\mathcal{B}(X) \ni \Delta \to a_\sigma(\Delta)$ définit bien une mesure à valeurs opérateurs positifs. On vérifiera aisément la manière dont cette mesure se transforme sous l'action de la représentation de Wigner :

$$U_m(a, h(k))a_\sigma(\Delta)(U_m(a, h(k)))^\dagger = a_{\sigma_{(a,h(k)}}((a, h(k))\Delta). \tag{12.105}$$

On dira, dans le prolongement du système de covariance galiléen (12.68), que les opérateurs (12.104) forment un système de covariance généralisé.

Portons notre attention sur la section galiléenne $\sigma_0(\boldsymbol{q}, \boldsymbol{p}) = ((0, \boldsymbol{q}), h(p))$. Lorsque $(\boldsymbol{q}, \boldsymbol{p})$ varie dans $\mathbb{R}^2$, l'image de cette section dans $\mathcal{P}_+^\uparrow(1,1)$ s'identifie au produit cartésien $\mathbb{R} \times \mathcal{V}_m^+$. Dans l'hypothèse où la sonde η possède une parité définie, $\eta(-\boldsymbol{k}) = \pm\eta(\boldsymbol{k})$, un calcul simple montre que les opérateurs

$$\boldsymbol{Q}_{\sigma_0}^\eta \equiv \int_X \boldsymbol{q}\, a_{\sigma_0}(d\boldsymbol{q}\, d\boldsymbol{p}), \quad \boldsymbol{P}_{\sigma_0}^\eta = \int_X \boldsymbol{p}\, a_{\sigma_0}(d\boldsymbol{q}\, d\boldsymbol{p}), \qquad (12.106)$$

sont égaux respectivement à l'opérateur de Newton-Wigner $\boldsymbol{Q}_{NW}$ donné par (12.81) et à l'opérateur $\boldsymbol{P}$ donné par (12.78). Ils obéissent donc à la relation canonique de commutation et on peut accorder aux opérateurs $a_{\sigma_0}(\Delta)$ un statut d'opérateur de localisation au sens de Newton-Wigner non plus sur l'« espace » à temps nul $\mathbb{R}$, mais sur l'espace des phases $\mathbb{R} \times \mathcal{V}_m^+$. De plus, on a une propriété de covariance intéressante pour cette position sous l'action d'une transformation de Lorentz :

$$U_m(h(k))\boldsymbol{Q}_{\sigma_0}^\eta(U_m(h(k)))^\dagger = \int_X \boldsymbol{q}'\, a_{\sigma_{(0,h(k))}}(d\boldsymbol{q}\, d\boldsymbol{p}), \qquad (12.107)$$

avec $(\boldsymbol{q}', \boldsymbol{p}') = (0, h^{-1}(k))(\boldsymbol{q}, \boldsymbol{p})$.

Comme nous l'avions annoncé dans la section 2, les opérateurs $a_{\sigma_{(0,h(k))}}(\Delta)$ nous offrent une localisation sur l'espace des phases $\Sigma_n \times \mathcal{V}_m^+$ où Σ_n est l'axe de genre espace perpendiculaire, pour la métrique minskowskienne, au vecteur unitaire $n = h(k)(1, \mathbf{0})$. En effet, pour un borélien $\Delta \subset X$, posons

$$\hat{\Delta} = \left\{(\hat{q}, p) \in \Sigma_n \times \mathcal{V}_m^+ \mid \hat{q} = h^{-1}(k)(0, \boldsymbol{q}),\ p = \left(\sqrt{\boldsymbol{p}^2 + m^2}, \boldsymbol{p}\right),\ (\boldsymbol{q}, \boldsymbol{p}) \in \Delta\right\}.$$

Alors, la valeur moyenne $\langle\, \psi \mid a_{\sigma_{(0,h(k))}}(\Delta) \mid \psi\,\rangle$ s'interprètera tout naturellement et en accord avec (12.105) comme la probabilité de localisation du système $\hat{\Delta}$ lorsque son état est ψ.

En réalité, cette localisation dans l'espace des phases prend tout son sens lorsque les états du système sont représentés par des fonctions de carré intégrable sur X. La famille des états cohérents $\eta_{\sigma(\boldsymbol{q},\boldsymbol{p})}$ nous permet de réaliser cette représentation grâce au plongement isométrique $W : L^2(\mathcal{V}_m^+, d\boldsymbol{k}/k_0) \to L^2(X, d\boldsymbol{q}\, d\boldsymbol{p})$ donné par

$$(W\psi)(\boldsymbol{q}, \boldsymbol{p}) = \langle\, \eta_{\sigma_0(\boldsymbol{q},\boldsymbol{p})} \mid \psi\,\rangle \equiv \Psi(\boldsymbol{q}, \boldsymbol{p}). \qquad (12.108)$$

On parlera alors de fonction d'onde $\Psi(\boldsymbol{q}, \boldsymbol{p})$ et de densité de probabilité $\|\Psi(\boldsymbol{q}, \boldsymbol{p})\|^2$ du système dans sa représentation *espace des phases*. L'évolution temporelle en sera donnée par

$$\Psi(\boldsymbol{q}, \boldsymbol{p}, t) = \langle\, \eta_{\sigma(\boldsymbol{q},\boldsymbol{p})} \mid e^{-iP_0 t}\psi \,\rangle. \tag{12.109}$$

Finalement, pour un choix de sonde η à valeurs réelles, il existe un courant covariant de Lorentz [20] :

$$j_\mu(q) = \int_{\mathcal{V}_m^+} \frac{p_\mu}{m} \|\Psi(\boldsymbol{q}, \boldsymbol{p},)\|^2 \frac{d\boldsymbol{p}}{p_0}, \; q = (q_0, \boldsymbol{q}) = (t, \boldsymbol{q}), \tag{12.110}$$

et ce courant est conservé,

$$\partial^\mu j_\mu(q) = 0. \tag{12.111}$$

La localisation stricte, à la Newton-Wigner-Wightman, s'appuyant sur le concept de mesure à valeur projecteur, posait problème dans le contexte de la physique quantique einsteinienne. L'alternative que nous venons de présenter, principalement initiée et développée dans [2, 20], s'appuie, quant à elle, sur une notion plus souple, celle de mesure à valeur opérateur positif, dotée de propriétés de covariance et de signification probabiliste, indissociable de la notion d'état cohérent et de repère continu. Cette alternative semble être pour l'instant la seule réponse raisonnable à apporter à une question déjà bien ancienne.

Bibliographie

[1] Ali, S.T., Stochastic localization, quantum mechanics on phase space and quantum space-time *Riv. Nuovo Cimento* 8 (1985), 1–128.

[2] Ali, S.T., Systems of Covariance in Relativistic Quantum Mechanics, *Int. Journ. Theor. Phys.* 37 (1998), 365–373.

[3] Ali, S.T., Antoine, J.-P., Gazeau, J.-P., *Coherent states, wavelets and their generalizations*, Graduate Texts in Contemporary Physics, Springer-Verlag, New York, 2000.

[4] Ali, S.T., Antoine, J.-P., Gazeau, J.-P., Relativistic Quantum Frames, *Ann. Phys.* 222 (1993), 38–88.

[5] Ali, S.T., Gazeau, J.-P., Karim, M.R., Frames, the β-duality in Minkowski space and spin coherent states, *J. Phys. A: Math. Gen.* 29 (1996), 5529–5549.

[6] Berezin, F.A., General concept of quantization, *Comm. Math. Phys.* 40 (1975), 153–174.

[7] Daubechies, I., *Ten lectures on wavelets*, SIAM-CBMS, 1992.

[8] Deltheil, R., *Probabilités géométriques*, Traité de Calcul des Probabilités et de ses Applications par Émile Borel, Tome II, Gauthiers-Villars, Paris, 1926.

[9] Dirac, P.A.M., *The Principles of Quantum Mechanics*, Oxford University Press, Oxford, 1958.

[10] Feng, D.H., Klauder, J.R., Strayer, M. (Eds.), *Coherent States: Past, Present and Future (Proc. Oak Ridge 1993)*, World Scientific, Singapore, 1994.

[11] Guelfand, I.M., Vilenkin, N.Y., *Les distributions*, tome 4, *applications de l'analyse harmonique*, Dunod, Paris, 1967.

[12] Hegerfeldt, G.C., Remarks on causality and particle localization, *Phys. Rev. D* 10 (1974), 3320–3321.

[13] Hegerfeldt, G.C., Ruijsenaars N.M., Remarks on causality, localization, and spreading of wave packets *Phys. Rev. D* 22 (1980), 377–384.

[14] Kálnay A.J., The Localization Problem, *in Problems in the Foundations of Physics*, M. Bunge (Ed.), Studies in the Foundations Methodology and Philosophy of Science, IV, Springer-Verlag, Berlin, 1971.

[15] Klauder, J.R., Continuous-Representation Theory I. Postulates of continuous-representation theory, *J. Math. Phys.* 4 (1963), 1055–1058.

[16] Klauder, J.R., Continuous-Representation Theory II. Generalized relation between quantum and classical dynamics, *J. Math. Phys.* 4 (1963), 1058–1073.

[17] Lévy-Leblond, J.-M., The pedagogical Role and Epistemological Significance of Group Theory in Quantum Mechanics *Riv. Nuovo Cimento* 4 (1974), 99–143.

[18] Mackey, G.W., *Induced Representations of Groups and Quantum Mechanics*, Benjamin, New York, 1968.

[19] Newton, T.D., Wigner, E.P., Localized states for elementary systems, *Rev. Mod. Phys.* 21 (1949), 400–406.

[20] Prugovečki, E., Consistent formulation of relativistic dynamics for massive spin-zero particles in external fields, *Phys. Rev. D* 18 (1978), 3655–3673 ; Relativistic quantum kinematics on stochastic phase space for massive particles *Journ. Math. Phys.* 19 (1978), 2261–2270.

[21] Torrésani, B., *Analyse continue par ondelettes*, Savoirs Actuels, CNRS Editions & Editions de Physique, Paris, 1995.

[22] Wightman, A.S., On the localizability of quantum mechanical states, *Rev. Mod. Phys.* 34 (1962), 845–872.

13

Quantification canonique et énergie du vide

Jacques Renaud

1. Introduction

Les mesures récentes indiquant que la constante cosmologique n'est probablement pas nulle relancent l'intérêt pour l'espace-temps de de Sitter. Considérons en effet l'équation d'Einstein

$$G(g) + \Lambda g = T$$

qui permet de déterminer la métrique g de l'univers. Dans cette équation, $G(g)$ est le tenseur d'Einstein qui ne dépend que de g, Λ est la constante cosmologique et le terme source T est déterminé par la répartition de la matière et plus généralement de l'énergie dans l'univers. Se placer dans le cadre de la relativité restreinte consiste à supposer que la matière ne contribue pas à la courbure de l'espace-temps c'est-à-dire que $T = 0$ dans l'équation d'Einstein (le champ de gravitation est alors, en cas de besoin, introduit comme un champ extérieur). Si la constante cosmologique est nulle, l'équation devient

$$G(g) = 0$$

et la relativité restreinte doit s'écrire sur l'espace-temps de Minkowski qui est la seule solution de cette équation. En revanche, si Λ est strictement positive, c'est l'espace-temps de de Sitter, seule solution de l'équation

$$G(g) + \Lambda g = 0$$

qui doit porter la relativité restreinte.

Ceci est évidement un argument de poids pour procéder à une étude détaillée de la théorie quantique des champs sur l'espace-temps de de Sitter. Mais il y a au moins une autre raison pour s'intéresser à cet espace, c'est la présence d'un groupe de symétrie de dimension maximale.

Dans le cas de l'espace-temps de Minkowski, la procédure de quantification fait l'objet d'un large consensus, ce qui n'est pas le cas pour d'autres espace-temps où l'on assiste à une prolifération des quantifications. Ceci est dû essentiellement à l'absence d'un groupe d'isométrie jouant un rôle comparable à celui, décisif, du groupe de Poincaré [12]. De ce point de vue, l'espace de de Sitter présente des qualités remarquables pour tester les différentes méthodes de quantification : l'existence d'un groupe de Lie d'isométries de dimension maximale et d'une procédure bien maîtrisée de passage à la limite de courbure nulle dans les représentations de ce groupe. Il semble alors raisonnable d'exiger de toute procédure de quantification qu'elle donne, quand elle est appliquée à l'espace de de Sitter, un résultat complètement covariant (comme c'est le cas dans l'espace de Minkowski) et que ce champ donne à la limite de courbure nulle les résultats connus pour Minkowski.

Malgré toutes ces bonnes raisons, la théorie quantique des champs sur l'espace de de Sitter est loin d'être totalement claire. En particulier, le champ minimalement couplé, qui est une étape naturelle vers la quantification de la gravitation dans son approximation linéaire, était considéré comme ne pouvant pas être quantifié de façon covariante. Nous montrerons dans ce qui suit qu'une nouvelle méthode de quantification conduit à une quantification parfaitement covariante de ce champ et, de surcroît, cette nouvelle quantification présente la propriété intéressante de supprimer la divergence ultraviolette dans le calcul de l'énergie du champ et de donner une énergie nulle pour l'état du vide [10, 11].

L'espace-temps de de Sitter peut être représenté comme un hyperboloïde plongé dans l'espace de Minkowski à cinq dimensions :

$$M_H = \left\{X \in \mathbb{R}^5 |\ X^2 = \eta_{\alpha\beta} X^\alpha X^\beta = X_\alpha X^\alpha = -H^{-2}\right\},$$

où $\eta^{\alpha\beta} = \mathrm{diag}(1, -1, -1, -1, -1)$. La (pseudo-)sphere M_H est clairement invariante sous l'action des transformations de Lorentz en dimension 5. Donc cet espace-temps admet un groupe d'isométries à 10 paramètres, le groupe de de Sitter : $O(1, 4)$ dont on considèrera seulement la composante connexe de l'identité $SO_0(1, 4)$.

Le champ minimalement couplé que nous allons considérer dérive du lagrangien ($\hbar = 1$)

$$S(\phi) = \frac{1}{2}\int_{M_H} g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi \, d\mu, \tag{13.1}$$

dont dérive l'équation du champ :

$$\Box\phi = 0, \tag{13.2}$$

où $\Box$ est l'opérateur de Laplace-Beltrami défini sur la variété lorentzienne M_H. Le produit scalaire de Klein-Gordon est défini pour tout ϕ, ψ solutions de (13.2) par

$$\langle\phi, \psi\rangle = i\int_\Sigma \phi^*(\overrightarrow{\partial}_\mu - \overleftarrow{\partial}_\mu)\psi d\sigma^\mu \equiv i\int_\Sigma \phi^* \overleftrightarrow{\partial_\mu} \psi d\sigma^\mu, \tag{13.3}$$

où Σ est une surface de Cauchy, c'est-à-dire une hypersurface de genre espace telle que des données de Cauchy sur Σ déterminent de façon unique une solution de (13.2), et $d\sigma^\mu$ est l'élément de surface sur Σ. Ce produit est de Sitter invariant et indépendant du choix de Σ.

Le champ minimallement couplé a fait l'objet de nombreux travaux [1, 2, 13] parce qu'il est une étape vers la quantification du champ de gravitation et que d'autre part la quantification canonique usuelle ne fonctionne pas dans ce cas. Ceci a amené de nombreux auteurs à proposer des quantifications non covariantes. Dans ce qui suit, nous montrerons qu'il est possible de quantifier ce champ de façon covariante en choisissant une nouvelle représentation des relations canoniques de commutation. Dans la section suivante, nous présentons cette nouvelle quantification en nous plaçant dans le cadre plus général des espaces globalement hyperboliques.

2. Champs sur espace-temps courbe

Dans tout ce qui suit, on se place sur un espace-temps M, c'est-à-dire une variété pseudo-riemannienne dont la métrique $g_{\mu\nu}(x)$ est de signature $(-1, +1, +1, +1)$. Nous nous limiterons au champ scalaire ; la densité de lagrangien du système est donné par

$$\mathcal{L} = g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - \zeta^2\phi,$$

où ζ est un paramètre réel de couplage que nous ne discuterons pas ici. L'équation d'Euler correspondante est donnée par

$$\Box\phi + \zeta^2\phi = 0, \tag{13.4}$$

qui est l'équation du champ.

L'espace-temps sera toujours supposé globalement hyperbolique [8, 12], ce qui signifie que l'équation du champ admet des hypersurfaces de Cauchy de genre espace Σ et une fonction à deux points $\tilde{G}$ telles que

$$\phi(x) = \int_\Sigma \tilde{G}(x, x') \overleftrightarrow{\partial_\mu} \phi(x')\, d\sigma^\mu(x'),$$

pour toute solution régulière ϕ de l'équation du champ. La fonction $\tilde{G}$ est appelée le commutateur du champ, bien que ce soit un objet purement géométrique. Elle est solution de l'équation du champ en chacune de ses variables, et, pour x fixé, son support en x' est inclus dans le cône causal de x. On voit que le champ classique ϕ est déterminé par ses conditions initiales sur l'hypersurface Σ. L'ensemble des solutions de l'équation du champ est muni d'un produit scalaire non nécessairement positif, appelé produit de Klein-Gordon :

$$\langle \phi_1, \phi_2 \rangle = i \int_\Sigma \phi_1^* \overleftrightarrow{\partial_\mu} \phi_2\, d\sigma^\mu, \tag{13.5}$$

qui est indépendant du choix de l'hypersurface de Cauchy Σ. Cette dernière est souvent choisie comme une surface x^0 = constante. Notons que, dans le cas plat, ce produit scalaire n'est rien d'autre que le produit scalaire naturel entre les fonctions définies sur l'hyperboloïde de masse dans l'espace réciproque.

Il existe aussi sur la variété le produit scalaire habituel (produit L^2) :

$$(f_1, f_2) = \int_M f_1^*(x) f_2(x) d\lambda(x), \tag{13.6}$$

où $d\lambda$ est la mesure naturelle $d\lambda(x) = \sqrt{|g(x)|}\, d^4x$. Contrairement au précédent, ce produit scalaire est défini positif et s'applique à des fonctions qui ne sont pas nécessairement solutions de l'équation du champ.

Dans les exemples qui nous concernent ici (espace-temps de Minkowski et de de Sitter), il existe de plus un groupe G d'isométries de l'espace-temps. Ce groupe agit naturellement sur les solutions de l'équation par $U_g\phi(x) = \phi(g^{-1}x)$ pour $g \in G$ et le produit de Klein-Gordon est invariant par cette action :

$$\langle U_g\phi_1, U_g\phi_2 \rangle = \langle \phi_1, \phi_2 \rangle.$$

Il en est de même du produit L^2 :

$$(U_g f_1, U_g f_2) = (f_1, f_2).$$

Le problème posé est celui de la quantification de ce champ. Il s'agit donc de construire une distribution à valeurs opérateurs φ. En toute rigueur, c'est $\varphi(f)$, où f est une fonction test, c'est-à-dire une fonction C^∞ à support compact sur M, qui est un opérateur bien défini, mais nous ferons le plus souvent l'abus de langage consistant à écrire $\varphi(x)$. Il y a un certain nombre de propriétés qu'un tel champ doit raisonnablement vérifier pour être admissible.

Équation du champ

Le champ quantique doit vérifier l'équation du champ :

$$(\Box + \zeta^2)\varphi(x) = 0,$$

les dérivées étant prises au sens des distributions.

Microcausalité

Si x et x' sont causalement disjoints, c'est-à-dire que x' n'est pas dans le cône causal de x, les champs en ces points sont indépendants et doivent donc commuter :

$$[\varphi(x), \varphi(x')] = 0$$

dès que x et x' sont causalement disjoints.

Covariance

À l'action du groupe d'isométries sur l'espace-temps, correspond une représentation unitaire $\underline{U}$ du groupe sur l'espace des états et donc une action sur les observables quantiques, ce qui conduit à la condition de covariance :

$$\underline{U}_g^{-1}\varphi(x)\underline{U}_g = \varphi(g \cdot x), \quad \forall x \in M,\ \forall g \in G.$$

On a montré [1] que dans le cas très important du champ minimalement couplé ($\zeta = 0$ dans (13.4)) sur l'espace temps de de Sitter, on ne pouvait pas construire un tel objet, du moins dans le cadre habituel des espaces de Hilbert. La raison profonde de cette impossibilité est que le Lagrangien correspondant à ce cas présente une invariance analogue à l'invariance de jauge (voir section 5). Il est bien connu que, dans le cas de l'électrodynamique, la seule façon de faire cohabiter la covariance avec cette invariance de jauge dans le cadre d'une quantification canonique, est de quantifier à la façon de Gupta et Bleuler ; le prix à payer étant l'apparition d'états non physiques pouvant être de norme négative. C'est ainsi que nous procèderons ; mais, auparavant, nous allons rappeler le principe de la quantification canonique pour montrer que notre quantification est également une quantification canonique ne différant de la quantification usuelle que par une nouvelle représentation des relations canoniques de commutation.

3. Quantification canonique usuelle

La méthode de quantification canonique est très heuristique mais il est facile de la rendre rigoureuse ; elle fonctionne de la façon suivante. On suppose construite une famille ϕ_k (les modes) de solutions de l'équation du champ vérifiant les conditions suivantes :

$$\begin{aligned}\langle\phi_k,\phi_{k'}\rangle &= \delta_{kk'}, &\text{(a)}\\ \langle\phi_k,\phi^*_{k'}\rangle &= 0, &\text{(b)}\\ \textstyle\sum_k \phi_k(x)\phi^*_k(x') - \phi^*_k(x)\phi_k(x') &= -i\tilde{G}(x,x'). &\text{(c)}\end{aligned}$$

Ces conditions appellent quelques remarques. Le choix des ϕ_k détermine le choix de l'espace $\mathcal{H}_p$ qu'ils engendrent comme base hilbertienne. Cet espace est appelé traditionnellement secteur à une particule, bien que l'interprétation en terme de particules prête à discussion. Quand l'espace-temps est stationnaire, c'est-à-dire quand il existe un champ de vecteur de Killing X qui est de genre temps en chaque point, il y a une notion naturelle de hamiltonien pour le secteur à une particule à savoir $-iX$, et l'espace $\mathcal{H}_p$ doit alors être choisi de telle sorte que, sur cet espace, ce hamiltonien admette un spectre positif ; les modes sont en général choisis comme vecteurs propres, à valeur propre positive, de cet opérateur. On dit alors que les ϕ_k sont à fréquence positive. Cependant les espace-temps que nous considérons ne sont pas nécessairement stationnaires et les conditions (a) et (b) sont là pour remplacer, de façon plus faible, la condition de fréquence positive. Mais ces conditions ne déterminent pas de façon unique l'espace à une particule : la famille $\tilde{\phi}_k = A\phi_k - B\phi^*_k$ avec $|A|^2 - |B|^2 = 1$ (transformation de Bogoliubov) vérifieront ces conditions dès que les ϕ_k les vérifient. De plus les fonctions à deux points, c'est-à-dire les valeurs moyennes sur le vide calculées pour le champ obtenu par quantification canonique vont dépendre de ce choix. Comme le vide est défini par référence aux modes, il est devenu traditionnel de dire que ce choix de mode revient à choisir un vide, nous verrons que notre point de vue est tout à fait différent. La condition (c) assure que la famille ϕ_k est complète, c'est-à-dire que l'espace engendré par les ϕ_k et les ϕ^*_k contient toutes les solutions régulières de l'équation du champ dont les conditions initiales sont à support compact.

Ceci étant posé, on définit alors le champ quantique par

$$\varphi(x) = \sum_k \phi_k(x)A_k + \phi^*_k(x)A^\dagger_k,$$

où les A_k et $A_k^\dagger$ sont des opérateurs vérifiant les relations canoniques de commutation (ccr) :

$$[A_k, A_{k'}^\dagger] = \delta_{kk'}, \quad [A_k, A_{k'}] = 0, \quad [A_k^\dagger, A_{k'}^\dagger] = 0.$$

Ces relations sont une conséquence directe des principes généraux de quantification d'une théorie lagrangienne : transformation des crochets de Poisson en commutateurs. La raison pour laquelle notre notation est différente de la notation usuelle est que ces relations sont le point de départ aussi bien de la quantification canonique traditionnelle que de notre quantification à la Gupta-Bleuler : ces deux quantifications ne diffèreront que par la représentation de ces relations, c'est-à-dire la réalisation des A_k comme opérateurs agissant sur un certain espace. Avant tout choix de représentation, il est clair que, au moins formellement, le champ ainsi construit vérifie l'équation et la condition de causalité puisque les relations canoniques de commutation impliquent que

$$[\varphi(x), \varphi(x')] = -i\tilde{G}(x, x').$$

La condition de covariance semble moins immédiate ; nous allons cependant montrer qu'il existe une façon de présenter la quantification canonique habituelle pour laquelle la covariance deviendra évidente, la seule condition à réaliser étant que l'espace $\mathcal{H}_p$ soit fermé sous l'action du groupe d'isométries. Rappelons que l'espace $\mathcal{H}_p$ est défini par

$$\mathcal{H}_p = \left\{ \sum_k c_k \phi_k ; \ \sum_k |c_k|^2 < \infty \right\}.$$

En ce qui concerne la quantification canonique usuelle, la représentation des relations de commutation canoniques se réalise dans l'espace de Fock symétrique [16] $\underline{\mathcal{H}_p}$ construit sur $\mathcal{H}_p$ en posant

$$A_k = a_k := a(\phi_k) \text{ et } A_k^\dagger = a_k^\dagger := a^\dagger(\phi_k),$$

où $a(\phi_k)$ (resp. $a^\dagger(\phi_k)$) est l'anihilateur (resp. le créateur) du mode ϕ_k. On a alors, utilisant la linéarité et l'anti-linéarité des opérateurs de création et d'anihilation

$$\begin{aligned} \varphi(x) &= \sum_k \phi_k(x) a_k + \phi_k^*(x) a_k^\dagger \\ &= a(p(x)) + a^\dagger(p(x)), \end{aligned}$$

où $p(x)$ est l'élément de $\mathcal{H}_p$ défini par

$$p(x) = \sum_k \phi_k^*(x)\phi_k.$$

Le point crucial est que l'on peut définir p de façon indépendante de la base des ϕ_k ; pour cela on se rappellera que p, tout comme φ, est en fait une distribution, et qu'il est nécessaire pour être rigoureux, de faire intervenir des fonctions test (que l'on choisira réelles) $f \in C_o^\infty(M)$. La définition correcte de φ est donc :

$$\begin{aligned}
\varphi(f) &= \int f(x)\varphi(x)d\mu(x) \\
&= \sum_k \int \phi_k(x)f(x)d\mu(x)a_k + \sum_k \int \phi_k^*(x)f(x)d\mu(x)a_k^\dagger \\
&= \sum_k (\phi_k^*, f)a_k + \sum_k (\phi_k, f)a_k^\dagger,
\end{aligned}$$

où les parenthèses désignent le produit L^2. Par linéarité on obtient :

$$\begin{aligned}
\varphi(f) &= \sum_k (\phi_k^*, f)a(\phi_k) + \sum_k (\phi_k, f)a^\dagger(\phi_k) \\
&= a\left(\sum_k (\phi_k, f)\phi_k\right) + a^\dagger\left(\sum_k (\phi_k, f)\phi_k\right).
\end{aligned}$$

Définissons

$$p(f) = \sum_k (\phi_k, f)\phi_k \in \mathcal{H}_p, \tag{13.7}$$

on a

$$\varphi(f) = a\left(p(f)\right) + a^\dagger\left(p(f)\right), \tag{13.8}$$

où p est une distribution à valeurs dans $\mathcal{H}_p$ qui peut être définie de la façon suivante : $p(f)$ est l'unique élément de l'espace de Hilbert $\mathcal{H}_p$ pour lequel

$$\langle p(f), \psi\rangle = (f, \psi), \qquad \forall \psi \in \mathcal{H}_p, \tag{13.9}$$

où les produits scalaires $\langle\ ,\ \rangle$ et $(\ ,\ )$ sont définis en (13.5) et (13.6) respectivement. Les formules (13.8) et (13.9) permettent alors une définition rigoureuse du champ qui dépend du choix de $\mathcal{H}_p$ mais pas de la base fournie par les modes. Cette écriture plus intrinsèque permet une démonstration très facile de la covariance du champ. Celle-ci découle facilement de la covariance de la distribution p, la seule hypothèse nécessaire étant que $\mathcal{H}_p$ soit fermé sous l'action du groupe G. Plus précisément, p entrelace les représentations naturelles U du groupe sur $\mathcal{C}_o^\infty(M)$ et $\mathcal{H}_p$, en fait on a

$$U_g p = pU_g. \tag{13.10}$$

En effet, pour tout $\psi \in \mathcal{H}_p$

$$\begin{aligned}\langle U_g p(f), \psi\rangle &= \langle p(f), U_g^{-1}\psi\rangle \\ &= (f, U_g^{-1}\psi) \\ &= (U_g f, \psi) \\ &= \langle p(U_g f), \psi\rangle.\end{aligned}$$

La représentation U s'étend en une representation $\underline{U}$ sur $\underline{\mathcal{H}_p}$ et

$$\begin{aligned}\underline{U}_g\varphi(f)\underline{U}_g^{-1} &= a\left(U_g p(f)\right) + a^\dagger\left(U_g p(f)\right) \\ &= a\left(p(U_g f)\right) + a^\dagger\left(p(U_g f)\right) \\ &= \varphi(U_g f).\end{aligned}$$

Ainsi la quantification canonique usuelle, et sa covariance, repose sur l'existence d'un espace de solutions $\mathcal{H}_p$ où le produit scalaire est positif, qui est fermé pour l'action du groupe et tel que $\mathcal{H}_p + \mathcal{H}_p^*$ contienne toutes les solutions régulières de l'équation. Il convient en fait de rajouter une condition technique : la définition (13.9) repose sur le fait que l'application $\psi \mapsto (f, \psi)$ est continue de $\mathcal{H}_p \to \mathbb{C}$. Cette condition n'est pas toujours réalisée et doit donc être vérifiée au cas par cas.

La théorie de la quantification canonique usuelle que l'on vient de décrire présente de sérieuses difficultés. Il est remarquable que, dès les premiers pas, des termes infinis apparaissent dans les calculs. Une des premières quantités à calculer est le tenseur énergie-moment, et sa valeur moyenne dans certains états du champ, en particulier dans le vide. Ce calcul fait apparaître les termes $\langle 0|\partial_\mu\varphi(x)\partial_\nu\varphi(x)|0\rangle$ ou $\langle 0|\varphi(x)\varphi(x)|0\rangle$ qui sont infinis. Reportons par exemple la définition du champ dans ce dernier terme. On sait, utilisant les ccr et $a_k|0\rangle = 0$, que

$$\langle 0|a_k a_{k'}|0\rangle = 0 = \langle 0|a_{k'}^\dagger a_k^\dagger|0\rangle, \text{ et pour } k \neq k' \ \langle 0|a_k^\dagger a_{k'}|0\rangle = 0 = \langle 0|a_{k'} a_k^\dagger|0\rangle.$$

Il reste donc

$$\begin{aligned}
\langle 0|\varphi(x)\varphi(x)|0\rangle &= \sum_k |\phi_k(x)|^2 \langle 0|(a_k a_k^\dagger + a_k^\dagger a_k)|0\rangle \\
&= 2\sum_k |\phi_k(x)|^2 \langle 0| \left(a_k^\dagger a_k + \frac{1}{2} \right) |0\rangle \\
&= 2\sum_k |\phi_k(x)|^2 \langle 0|a_k^\dagger a_k|0\rangle + \sum_k |\phi_k(x)|^2 \langle 0|0\rangle \\
&= 0 + \infty.
\end{aligned}$$

Ce calcul repose sur les ccr, mais aussi sur la définition du vide par $a_k|0\rangle = 0$. Dans le paragraphe suivant, nous présenterons une nouvelle quantification dans laquelle les ccr sont toujours valables mais avec un vide différent ne vérifiant pas $a_k|0\rangle = 0$.

4. Quantification à la Gupta-Bleuler

Une quantification du type Gupta-Bleuler suppose qu'on distinguera l'espace des états sur lequel sont définies les observables de l'espace des états physiques contenu dans le précédent, où seront prises les valeurs moyennes des observables. De plus le grand espace est muni d'une métrique indéfinie, c'est-à-dire que certains états (non physiques) ont une norme négative.

La véritable quantification de Gupta-Bleuler avait été inventée pour sauvegarder la covariance en présence d'une invariance de jauge. Il n'est pas surprenant qu'une telle méthode fonctionne (au contraire de la quantification canonique habituelle) pour le champ minimallement couplé où apparaît également une sorte d'invariance de jauge.

On se donne les modes vérifiant les mêmes conditions que ci-dessus. La seule différence est que notre construction n'exige plus que $\mathcal{H}_p$ soit fermé par l'action du groupe. Il suffira que le grand espace $\mathcal{H} = \mathcal{H}_p + \mathcal{H}_p^*$ le soit, ce qui est, bien sûr, une condition strictement plus faible. Là encore, le champ est défini par

$$\varphi(x) = \sum_k \phi_k(x) A_k + \phi_k^*(x) A_k^\dagger,$$

où les A_k et $A_k^\dagger$ sont des opérateurs vérifiant les relations canoniques de commutation :

$$[A_k, A_{k'}^\dagger] = \delta_{kk'}, \quad [A_k, A_{k'}] = 0, \quad [A_k^\dagger, A_{k'}^\dagger] = 0.$$

La représentation de ces relations sera obtenue de la façon suivante. On considère l'espace de Fock $\underline{\mathcal{H}}$ [14] construit sur l'espace de Krein [3] $\mathcal{H} = \mathcal{H}_p + \mathcal{H}_p^*$, où $\mathcal{H}_p^*$ est l'espace anti-Hilbertien engendré par les ϕ_k^*. On pose $a_k = a(\phi_k)$ et $b_k = a(\phi_k^*)$ puis

$$A_k = \frac{1}{\sqrt{2}}(a_k - b_k^\dagger) \text{ et } A_k^\dagger = \frac{1}{\sqrt{2}}(a_k^\dagger - b_k).$$

Ces opérateurs vérifient les ccr, on a donc bien une autre réalisation des ccr qui induit une nouvelle définition du champ. On peut maintenant reprendre point par point la discussion du paragraphe précédent. Les deux premières conditions (équation du champ et microcausalité) sont toujours vérifiées puisqu'elles ne dépendent pas de la représentation des ccr utilisée.

En vue de démontrer la covariance, on va tout d'abord réécrire ce champ en s'affranchissant de la base des ϕ_k, pour cela on procède comme dans la section précédente. On définit la distribution p à valeur dans $\mathcal{H}$ de la façon suivante. Pour toute fonction test f, $p(f)$ est, sous réserve d'une condition technique sur $\mathcal{H}$, l'unique élément de $\mathcal{H}$ tel que

$$\langle p(f), \psi\rangle = (f, \psi) \quad \forall\psi \in \mathcal{H}. \tag{13.11}$$

En posant formellement $p(f) = \int p(x)f(x)dx$, on obtient la définition formelle $\langle p(x), \psi\rangle = \psi(x)$. On vérifie que le noyau de p est la fonction $\tilde{G}$:

$$\langle p(x), p(x')\rangle = -i\tilde{G}(x, x').$$

Il est alors facile de vérifier, en développant sur la base des ϕ_k que

$$\varphi(x) = a(p(x)) + a^\dagger(p(x)). \tag{13.12}$$

L'avantage de l'écriture (13.12) est que la covariance du champ se montre comme ci-dessus (*cf.*, équation (13.10) et sa démonstration), la seule condition étant que l'espace $\mathcal{H}$ (et non plus $\mathcal{H}_p$) soit fermé sous l'action du groupe. Ceci est un des gros avantages de la méthode, car, comme c'est discuté en détail dans [6, 10], il existe des situations où l'espace $\mathcal{H}$ n'a pas de sous-espace de Hilbert invariant par le groupe, et il n'est alors pas possible de définir $\mathcal{H}_p$ de façon covariante.

Notons néanmoins un grosse différence avec la représentation usuelle : les A_k ne vérifient pas $A_k|0\rangle = 0$. Cela entraîne la disparition des termes infinis dans les calculs des $T_{\mu\nu}$, ceci est discuté en détail dans [6, 7, 10] et nous nous contenterons de montrer ici comment le calcul du paragraphe précédent se modifie.

Le point crucial est que

$$[b_k, b_k^\dagger] = [a(\phi_k^*), a^\dagger(\phi_k^*)] = \langle \phi_k^*, \phi_k^* \rangle = -1,$$

ce qui implique

$$a_k a_k^\dagger + a_k^\dagger a_k + b_k b_k^\dagger + b_k^\dagger b_k = 2a_k^\dagger a_k + 2b_k^\dagger b_k.$$

On en déduit que

$$(A_k A_k^\dagger + A_k^\dagger A_k) = a_k^\dagger a_k + b_k^\dagger b - a_k b_k - a_k^\dagger b_k^\dagger.$$

On considère maintenant $|0\rangle$, le vide de l'espace de Fock $\underline{\mathcal{H}}$. Ce vide est caractérisé par $a_k|0\rangle = b_k|0\rangle = 0$ mais $A_k|0\rangle \neq 0$. On a toujours

$$\langle 0|A_k A_{k'}|0\rangle = 0 = \langle 0|A_{k'}^\dagger A_k^\dagger|0\rangle, \text{ et pour } k \neq k' \ \langle 0|A_k^\dagger A_{k'}|0\rangle = 0 = \langle 0|A_{k'} A_k^\dagger|0\rangle.$$

Il reste donc

$$\langle 0|\varphi(x)\varphi(x)|0\rangle = \sum_k |\phi_k(x)|^2 \langle 0|(A_k A_k^\dagger + A_k^\dagger A_k)|0\rangle = 0,$$

au vu de ce qui précède. Et ceci montre que cette quantification fournit un renormalisation automatique des expressions intervenant dans le calcul des $T_{\mu\nu}$. On montre dans [7, 10] que la valeur moyenne du tenseur énergie moment dans le vide est non seulement finie mais nulle.

La construction que nous venons d'exposer permet donc, sous des conditions plus faibles qu'habituellement, de définir un champ causal et covariant présentant une propriété remarquable de renormalisation automatique sans brisure de covariance.

En revanche, cette construction fait apparaître des états non physiques. Ceci pose principalement deux questions. D'abord, comment sélectionner les états physiques ? Ensuite, l'existence d'états non physiques n'induit-elle pas des résultats inadmissibles comme la mesure d'énergies négatives ?

L'espace des états physiques « à une particule » doit être un sous-espace de $\mathcal{H}$, fermé sous l'action du groupe et sur lequel le produit scalaire est positif. Dans les cas massifs, le choix naturel pour l'espace des états physiques est l'espace $\mathcal{H}_p$, cependant ce choix n'est en général pas unique et on retrouve là les ambiguités habituelles en théorie quantique des champs sur espace-temps courbe (notons néanmoins la différence de point de vue : le vide considéré ici est le vide de l'espace de Fock, il est invariant par une transformation de Bogoliubov, celle-ci consiste simplement à changer d'espace des états physiques). Mais en présence d'invariance de jauge, l'espace $\mathcal{H}_p$ ne sera en général pas fermé sous l'action du groupe. Pour le cas de la masse nulle sur les espace-temps de de Sitter et Minkowski en dimension 2, nous avons utilisé une amélioration de la méthode des orbites qui, associée à un critère d'analyticité, nous a fourni un espace des états physiques conformément invariant [5]. Cette méthode qui fait apparaître la masse nulle plus comme des valeurs aux bords que comme une limite de la masse non nulle a prouvé son efficacité en permettant, pour la première fois, de construire un calcul symbolique conformément invariant pour la dimension deux sur de Sitter ou Minkowski [15]. En l'absence de groupe, la méthode dite de quantification géométrique semble être un bon candidat pour construire l'espace des états physiques.

En tout état de cause, il faut souligner qu'en présence d'invariance de jauge, cet espace ne sera pas un espace de Hilbert ce qui aura deux conséquences importantes. Premièrement, on n'échappera pas à une quantification du type Gupta-Bleuler, la quantification canonique usuelle ne fonctionnant pas dans ce cas. Deuxièmement, il semble difficile de caractériser ces espaces, et donc la quantification, par une fonction à deux points.

L'espace des états physiques « à une particule » étant choisi, on construit l'espace des états physiques par seconde quantification, c'est-à-dire par création d'états physiques à partir du vide. Il est alors facile de vérifier que les valeurs moyennes de T_{00} sur de tels états sont toujours positives, ceci est dû à la positivité du produit scalaire sur l'espace des états physiques.

La section suivante est consacrée à l'application de ce qui précède à l'espace-temps de de Sitter.

5. Retour au champ minimalement couplé

Le Lagrangien correspondant à cette équation est donné par :

$$\mathcal{L} = \sqrt{|g|}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi.$$

Ce lagrangien est invariant par $\phi \to \phi + \lambda$. De ce fait, la situation est très similaire à celle de l'électrodynamique pour laquelle la quantification de Gupta-Bleuler a été introduite afin de concilier la covariance avec la liberté de jauge. Cette construction repose en définitive sur l'existence, dans le secteur à une particule, d'un triplet, dit triplet de Gupta-Bleuler [4, 9] :

$$\mathcal{N} \subset \mathcal{K} \subset \mathcal{H},$$

où $\mathcal{H}$ est l'espace total où est défini le champ, il contient des solutions de l'équation non physiques de norme négative. L'espace $\mathcal{N}$ est constitué des états de jauge, solutions de l'équation et orthogonaux à tous les éléments de $\mathcal{K}$, y compris eux-mêmes. L'espace des états physiques est *stricto sensu* $\mathcal{K}/\mathcal{N}$, mais on fait souvent l'abus de langage de désigner les éléments de $\mathcal{K}$ comme états physiques.

Le problème de la détermination de chacun des trois éléments du triplet de Gupta-Bleuler pose des problèmes de nature différentes. L'espace $\mathcal{N}$ est lié à la forme du lagrangien. Dans le cas qui nous occupe, les états de jauge sont les fonctions constantes. Ils correspondent à ce que l'on pourrait appeler un changement de jauge global, mais dans ce cas il n'en existe pas d'autre, ceci est dû au caractère très simple du Lagrangien et pas à la nature de la quantification. Soit donc ϕ_g la fonction constante égale à 1, on a

$$\mathcal{N} = \mathbb{C}\phi_g.$$

Le calcul des modes s'obtient de façon standart par séparation des variables dans le système de coordonnées dit conforme [10]. Mais l'espace de Hilbert formé à partir de ces modes n'est pas fermé sous l'action du groupe de de Sitter. Un tel phénomène se reproduira pour tous les systèmes de coordonnées : c'est le contenu du théorème de Allen [1]. Si on élimine cependant le mode zéro, on obtient une famille de modes ϕ_k $k \in K' = K \setminus \{0\}$ qui, complétée par ϕ_g, engendre un espace fermé sous l'action du groupe. Cet espace

$$\mathcal{K} = \left\{ c_g\phi_g + \sum_{k\in K\setminus\{0\}} c_k\phi_k, \text{ avec } \sum_{k\in K\setminus\{0\}} |c_k|^2 < \infty \right\},$$

est le terme central du triplet de Gupta-Bleuler.

On a, pour $k, l \in \mathbb{Z} \setminus \{0\}$,

$$\langle \phi_k, \phi_l \rangle = \delta_{k,l}, \quad \langle \phi_k, \phi_l^* \rangle = 0, \tag{13.13}$$

mais pour tout $l \in \mathbb{Z}$

$$\langle \phi_g, \phi_l \rangle = 0 = \langle \phi_g, \phi_l^* \rangle = \langle \phi_g, \phi_g \rangle, \tag{13.14}$$

ce qui signifie que le produit scalaire sur $\mathcal{K}$ est dégénéré. C'est pourquoi toute tentative pour quantifier naïvement le champ en utilisant la base de $\mathcal{K}$ comme système de modes échoue sur la covariance, bien que l'espace $\mathcal{K}$ lui-même soit fermé sous l'action du groupe de de Sitter. Notre point de vue sur la quantification *via* (13.11) et (13.12) permet de comprendre ce problème facilement : l'espace $\mathcal{K}$ est dégénéré pour son produit scalaire naturel ce qui interdit de définir p et donc de montrer la covariance du résultat. De nombreux auteurs arguent de cette difficulté pour justifier une brisure de symétrie, au contraire notre construction « à la Gupta-Bleuler » permet d'obtenir un champ covariant car si $\mathcal{K}$ est dégénéré, il peut être plongé dans un grand espace $\mathcal{H}$ qui ne l'est pas tout en étant covariant, et c'est sur cet espace $\mathcal{H}$ que seront définis p et φ. La quantification de Gupta-Bleuler a été utilisée la première fois pour rendre compatibles la covariance et une invariance de jauge. Il n'est pas surprenant qu'elle joue à nouveau ce rôle dans les cas ci-dessus qui présentent la même difficulté.

L'espace total, $\mathcal{H}$ est défini par

$$\mathcal{H} = \mathcal{H}_p \oplus \mathcal{H}_p^*,$$

avec

$$\mathcal{H}_p = \left\{ \sum_{k \in K} c_k \phi_k, \;\text{ avec } \sum_{k \in K} |c_k|^2 < \infty \right\}.$$

La construction se poursuit alors comme indiqué dans la section précédente. On obtient ainsi un champ causal et covariant.

L'existence d'états de jauge pose alors un problème de définition des observables. En effet, un changement de jauge ne doit pas pouvoir être observé, voir [10] pour les détails. Il apparaît ainsi que le champ lui-même n'est pas une observable, ce qui n'est pas surprenant dans un contexte Gupta-Bleuler, mais $\partial_\mu \varphi$ est une observable ainsi que le tenseur $T_{\mu\nu}$. La discussion de la section précédente s'applique alors sans changement, et on constate que la présence d'états de norme négative n'entraîne pas l'apparition d'énergies négatives. De plus, la valeur moyenne de $T_{\mu\nu}$ sur le vide est non seulement finie mais nulle.

Le fait que le champ ne soit pas une observable explique aussi pourquoi les fonctions à deux points de Wightman et Hadamard ne jouent pas de rôle important dans notre construction : elles ne sont pas des invariants de jauge. En particulier, elles ne peuvent pas servir à caractériser le vide qui n'est lui-même unique qu'à un choix de jauge près.

Bibliographie

[1] Allen, B., Vacuum states in de Sitter space, *Phys. Rev. D* 32 12 (1985) 3136–3149.

[2] Allen, B., Folacci, A., Massless minimally coupled scalar field in de Sitter space, *Phys. Rev. D* 35 12 (1987) 3771–3778.

[3] Bognar, J., *Indefinite inner product spaces*, Springer-Verlag, Berlin, 1974.

[4] Binegar, B., Fronsdal, C., Heidenreich, W., Conformal QED, *J. Math. Phys.* 24 12 (1983) 2828–2846.

[5] De Bièvre, S., Renaud, J., Quantization of the nilpotent orbits in $so(1,2)^*$ and massless particles on (anti)-de Sitter spacetime, *J. Math. Phys.* 35 8 (1994) 3775–3793.

[6] De Bièvre, S., Renaud, J., The massless Gupta-Bleuler vacuum on the 1 + 1-dimensional de Sitter space-time, *Phys. Rev. D* 57 10 (1998) 6230–6241.

[7] De Bièvre, S., Renaud, J., A fully conformally covariant quantum field in 1 + 1 dimension, *J. Phys. A* 34 (2001) 10901–10919.

[8] Fulling, S.A., *Aspects of Quantum Field Theory in Curved Space-Time*, Cambridge University Press, Cambridge, 1989.

[9] Gazeau, J.-P., Gauge fixing and Gupta-Bleuler triplets in de Sitter QED, *J. Math. Phys.* 26 7 (1985) 1847–1854.

[10] Gazeau, J.-P., Renaud, J., Takook, M., Gupta-Bleuler quantization for massless minimally coupled scalar fields in de Sitter space, *Class. Quantum Grav.* 17 (2000) 1415–1434.

[11] Gazeau, J.-P., Garidi, T., Renaud, J., Rouhani, S., Takook, M., Linear covariant Quantum Gravity in de Sitter space, *In progress*.

[12] Isham, C.J., Quantum field theory in curved space-times: A general mathematical framework *in* Differential geometrical methods in mathematical physics II, *Lec. Notes in Math.,* vol. 676, K. Bleuler *et al.* (Eds.), Springer, Berlin, 1978.

[13] Kirsten, K., Garriga, J., Massless minimally coupled fields in de Sitter space: O(4)-symmetric states versus de Sitter-invariant vacuum, *Phys. Rev. D* 48 (1993) 567–577.

[14] Mintchev, M., Quantization in indefinite metric, *J. Phys. A* 13 (1990) 1841–1860.

[15] Renaud, J., Symbolic calculus on the nilpotent orbit of $SO_0(1,2)$, *J. Geom. Phys.* 19 (1996) 277–286.

[16] Reed, M., Simon, B., *Fourier Analysis and Self-Adjointness*, Academic Press, New York 1975.

14 Courbes elliptiques, homotopie et extensions de l'espace

Joseph Kouneiher

« Le développement de la physique théorique dans le dernier quart du XX^e siècle a été guidé par un véritable système de valeurs romantiques. Aspirant à décrire les processus fondamentaux à l'échelle de Planck, les physiciens sont obligés de perdre toutes connexions avec le monde observable. Dans ce contexte social, les mathématiques sophistiquées émergeant en théorie des cordes cessent de n'être que des outils techniques compliqués requis uniquement pour calculer des effets mesurables mais deviennent matière à principe. Aujourd'hui, certains d'entre nous sont encore nourris par l'ancien sentiment platonicien selon lequel les idées mathématiques sont prédestinées à décrire le monde physique, aussi loin que leurs origines semblent éloigner de la réalité. De ce point de vue, on peut espérer que la théorie des nombres devienne la branche des mathématiques la plus appliquée. » (Y. Manin [10])

1. Théorie des cordes et le statut de l'espace-temps

La philosophie de la physique s'intéresse habituellement à des questions métaphysiques/ontologiques (comme la nature de l'espace ou de la probabilité physique) ou à des questions épistémologiques/méthodologiques (comme l'indétermination d'une théorie ou le réalisme scientifique en référence spécifique à la physique). Dans les deux cas, les discussions se restreignent à une théorie raisonnablement bien définie et bien établie et/ou parfois à un petit ensemble homogène de théories rivales bien établies.

Cependant, dans le cas de la théorie des cordes, nous rencontrons d'emblée des questions fondamentales qui vont au-delà des questions posées en philosophie des sciences en général. En effet, une des particularités de cette théorie est l'absence de données empiriques : la théorie des cordes a une échelle naturelle qui est de l'ordre de la longueur de Planck $L_p = \sqrt{\frac{G\hbar}{c^3}}$, ce qui rend toute exploration expérimentale pour le moment impossible.

L'absence des data a des implications à la fois en physique et en philosophie. En physique la conséquence principale est simplement liée à la difficulté de construire une théorie : l'absence, après plus de quarante ans, d'une théorie satisfaisante de la gravitation quantique témoigne de cette difficulté. Cette dernière est reliée aussi à la philosophie de deux manières : premièrement aux problèmes conceptuels, à propos de l'espace, du temps et de la matière par exemple[1]. Deuxièmement, l'absence d'un consensus à propos de la nature des data qu'une théorie quantique de la gravitation peut fournir, rend la construction d'une telle théorie difficile. Plus précisément, l'argument dimensionnel évoqué plus haut suggère que seuls des phénomènes à l'échelle infiniment petite, ou à haute énergie, peuvent exhiber des effets gravito-quantiques ; les applications principales d'une théorie comme celle des cordes ou de la gravitation quantique impliquent principalement les mêmes conditions qu'au début de l'Univers.

Dans une telle situation, la construction d'une théorie devient inévitablement influencée par des considérations théoriques. Plus précisément, elle tend à se baser sur des *prima facie* du point de vue de la forme qu'une théorie peut avoir : en partie les présupposés philosophiques du chercheur concerné ; et en partie l'existence de certaines techniques mathématiques qui ont déjà montré leur validité dans des branches de la physique théorique, comme les théories de jauges non abéliennes. Un programme de recherche tend alors vers la construction de schémas théoriques abstraits qui soient compatibles avec certains cadres conceptuels préconçus, et qui possèdent une consistance interne mathématique.

Ainsi, pour le philosophe, la théorie des cordes et la théorie de la gravitation quantique ne présentent pas une branche de la physique méthodologiquement et conceptuellement unifiée, mais plutôt une gamme d'approches disparates.

Nous constatons aujourd'hui l'existence d'un projet alternatif à la position traditionnelle de la philosophie de la géométrie. En effet, les immenses développements depuis l'habilitation de Riemann en 1854 jusqu'à l'établissement de la relativité générale ont transformé la philosophie de la géométrie au-delà de toute reconnaissance. En particulier, l'apriorisme Kantien a été presque remplacé par un certain empirisme et conventionnalisme. L'idée selon laquelle l'espace puisse avoir, à une échelle infiniment petite, une structure

1 – Nous pensons ici aux problèmes conceptuels qui surgissent de la différence des bases de la relativité générale et de la théorie quantique, ainsi que des problèmes à propos de chacune des théories en elle-même.

autre que celle d'une variété, devient aisément acceptable, ce qui n'était pas le cas lorsque l'influence de la forme originale des positions kantiennes dominait [9]. En fait, faute de théorie physique valable qui l'admette, cette idée n'avait presque aucune influence sur la philosophie de la géométrie[2].

Aujourd'hui, plusieurs approches non conventionnelles, par exemple la gravitation quantique, postulent une structure de non variété à l'espace temps. Même dans les approches plutôt conventionnelles, nous remarquons des indications d'une structure discrète sous-jacente à l'image continue de départ. Par exemple, dans le programme d'Ashtekar [1], la surface et le volume deviennent discrets. En théorie des supercordes, il y a des indications fortes sur la présence d'une valeur minimale de la longueur.

Ces propositions soulèvent plusieurs questions pour les philosophes de la géométrie. La principale concerne leur accommodation avec les positions traditionnelles, versions diverses de l'empirisme et du conventionnalisme. Notre discussion compte-t-elle pour ou contre le réalisme, en particulier le réalisme scientifique ? Présuppose-t-elle le réalisme ou plutôt sa falsifiabilité ? Même en prenant les propositions de théorie des cordes de façon réaliste, elle ne compte ni en faveur ni à l'encontre du réalisme scientifique. Ce manque d'engagement n'est pas surprenant vu le statut problématique de la question de l'élaboration de test en théorie des cordes.

En effet, le réalisme scientifique présuppose que les prétentions d'une théorie scientifique, réussie ou mature, soient vraies ou approximativement vraies et décrivent une réalité indépendante de nous. Ainsi, c'est une thèse conjonctive où une ontologie, à propos de la notion de vérité comme correspondance, est associée à une réalité indépendante.

Évidemment, les discussions concernant la théorie des cordes ou celles de la gravitation quantique, ne suivent pas obligatoirement cette doctrine. En effet, la théorie des cordes, et ceci indépendamment de ce que « réussie » et « mature » signifient, ne nous fournit pas de telles théories. Plus précisément, même si nous approuvons la première conjonction de réalisme scientifique concernant la vraisemblance de la théorie, la seconde, *i.e.*, son association avec une réalité indépendante, ne s'applique pas à la théorie des cordes.

Néanmoins, nous pouvons rester soumis au réalisme scientifique à travers le traitement de deux « théories ingrédients », la théorie quantique et la relativité générale. Mais ceci ne signifie pas un engagement au réalisme scientifique

2 – Son rôle principal, et *via* le point de vue de Riemann endossé aujourd'hui par Grunbaum : l'idée est que dans un espace discret et non pas une variété, la métrique est ou peut être intrinsèque et par conséquent non conventionnelle (voir [12]).

pour deux raisons. La première concerne l'optimisme épistémologique d'un tel réalisme scientifique. En effet, élaborer ou évaluer les ontologies suggérées par les théories scientifiques est aussi compatible avec le fait de nier cet optimisme. La deuxième raison est peut être moins évidente car elle admet l'hypothèse qu'élaborer ou évaluer des ontologies n'implique pas un engagement à une notion correspondante de vérité, ou de vérité approximative, caractéristique du réalisme. Ceci peut être surprenant, car les considérations que les philosophes de la physique accordent aux sujets de la vérité des « électrons » et des « points de l'espace-temps » ont le même degré de réalisme que celles « du sens commun » : concernant par exemple la référence aux chaises et la véracité des propositions à propos de ces chaises. Cependant, nous pouvons voir que cette suggestion est fausse car, quel que soit l'argument général que nous pourrions avancer à l'encontre du « réalisme » des considérations concernant la « référence » et la « vérité » à l' égard des chaises, cet argument peut sans doute s'appliquer aux électrons et aux points de l'espace-temps. Bien entendu, il faut tenir compte de la différence liée à la difficulté notoire de comprendre une « chose » quantique d'une manière réaliste simple.

De plus, les philosophes de la physique, dans leur expérience, tendent à endosser des considérations réalistes aux « références » et à la « vérité ». Cela est dû fondamentalement à la tendance psychologique à accepter l'existence des objets physiques réels correspondant, par leurs propriétés et leurs relations, à des objets de certains modèles mathématiques, *a fortiori* quand ces modèles sont réussis. Mais cette tendance est la cause et non pas la raison, et ceci concerne la question de « concrétisation mal placée », selon Whitehead, en physique théorique.

Remarquons aussi que la physique prétend fournir une description complète de son sujet ou « subject-matter ». À ce niveau, nous n'insistons pas sur le degré de précision de cette idée : par exemple, quel sens peut avoir « la notion de description complète » ? Est-elle censée être une partie de ce qu'on appelle la physique ? L'idée générale est que la physique prétend être complète et suffisante. Puisqu'elle entraîne qu'en physique, ou au moins en physique théorique, un changement de doctrine à propos de « sujet » est construit plutôt comme un changement de sujet lui-même, comme dans le cas de la théorie de supercordes.

Finalement, nous voulons insister sur les propositions fournies par le programme de la théorie des cordes à propos des trois problèmes conceptuels suivants :

(i) Ce programme conserve les idées techniques de base de la théorie quantique standard tout en les adaptant pour la structure de jauge de la théorie.

(ii) La conception de l'espace-temps dans les théories des cordes perturbatives fait appel à la géométrie différentielle tout en conservant le point de vue classique de l'espace-temps M. Cependant, comme nous allons voir dans le reste de cette discussion, la dimension de M est plus grande que 4. Ainsi, un scénario type « Kaluza-Klein » est nécessaire dans lequel les extra-dimensions sont suffisamment courbées pour être sans effet sur la physique normale qui se déroule dans un espace-temps 4 dimensionnel.

(iii) Les théories des cordes montrent clairement comment la relativité générale se réalise comme fragment d'une structure plus large, éliminant du coup le rôle significatif imputé aux notions de l'espace et du temps. En effet, la limite basse énergie de ces théories est une forme de supergravité, mais dans laquelle les idées standard de l'espace-temps ne jouent pas un rôle central. Par exemple, le fait que le graviton soit une particule parmi le nombre infini d'autres particules est le reflet de cette idée. En particulier, le groupe de difféomorphisme D de l'espace-temps apparaît comme une partie d'une plus grande structure. Par conséquent, son importance technique pour le schéma de quantification est subsumé par ce grand groupe.

Le programme de recherche en théorie des cordes a connu deux phases. La première [15], qui a commencé tôt au milieu des années 1980 (en suivant d'autres travaux du milieu des années 1970), utilise une approche perturbative. La deuxième, qui a commencé au début de l'année 1990 et continue aujourd'hui, a fourni des indications concernant la théorie non perturbative sous-jacente. Mais la structure de cette dernière n'est pas encore claire, en particulier son traitement de l'espace-temps.

Depuis le début des années 1990, la plupart des travaux au sein du programme de supercordes essaient d'explorer la théorie non-perturbative sous-jacente. Ces développements semblent avoir des implications importantes sur notre conception de l'espace et du temps à l'échelle de Planck. Ils se basent sur divers types de dualités [8] ou symétries. Par exemple, une des formes la plus simple de dualité est la T-dualité qui survient lorsque l'espace-but[3]

3 – L'évolution d'une corde dans l'espace-temps peut être formalisée en considérant une application du segment (dans le cas d'une corde ouverte) ou de cercle (corde fermée) dans la variété d'espace-temps M appelée aussi espace-but.

ou d'arrivé est une variété à cinq dimensions de la forme $M \times S^1$ (S^1 est le cercle). Il ressort que les prédictions physiques de la théorie sont invariantes en échangeant le rayon R de la cinquième dimension par $\frac{2\alpha'}{R}$. Ainsi, dans la dimension supplémentaire, nous ne pouvons pas différencier physiquement un petit rayon d'un autre très grand. Plus précisément, les théories sont jauge-équivalentes. Implication importante, il existe une longueur minimale $R_{\min} = \sqrt{2\alpha'}$[4], une idée qui va avoir sûrement des retombées sur notre compréhension des implications conceptuelles de la théorie.

Un autre type de dualité, la S-dualité [11] évoque l'idée que, dans certaines théories, la physique à la limite du couplage *fort* est donnée par la limite de *faible* couplage d'une théorie *duale*, dans laquelle les entités fondamentales s'identifient aux excitations solitoniques de la théorie originale.

La croyance générale aujourd'hui est que plusieurs théories perturbatives consistantes des supercordes sont liées, d'une part entre-elles, et à et d'autre part à une autre théorie [17] contenant des objets étendus (« membranes ») de dimension plus grande que 1, par de telles dualités[5].

Finalement, nous remarquons aujourd'hui dans ce programme de recherche la dominance d'un ensemble particulier de thèmes et d'idées concernant la relation entre les champs de jauge et les cordes.

K. Wilson [16] a remarqué, en étudiant la limite couplage fort d'une théorie de jauge sur réseau, que les excitations élémentaires sont représentées par des cordes fermées décrivant un flux dont les charges sont de type *couleur*. En présence des quarks, ces cordes s'ouvrent et leurs extrémités s'attachent aux quarks, garantissant ainsi leur confinement. De plus, dans la théorie de jauge $SU(N)$ l'interaction des cordes est faible lorsque N est élevé. Ce fait montre que, dans la limite physique continue, la meilleure description de la théorie doit impliquer des lignes de flux (des cordes) et non pas des champs. Ceci nous ramène de Maxwell à Faraday. En d'autres termes, il est naturel de trouver une dualité exacte entre les champs de jauge et les cordes. Le défi est de construire, du point de vue des cordes, une théorie précise de cette dualité.

Nous pouvons regarder le problème autrement. Parmi les théories des cordes, existe la classe particulière des *supercordes*. Ces dernières peuvent être visualisées comme des lignes de flux d'une certaine théorie de jauge inconnue. Ceci nous donnerait un point de vue particulièrement nouveau sur

4 – Ce phénomène peut être généralisé pour davantage d'extra dimensions, avec une topologie plus complexe que celle du produit de cercles.

5 – Un troisième type de dualité, « symétrie miroir », joue aussi un rôle important à ce niveau. Voir [5, 7].

ce que nous appelons l'espace-temps. L'espace-temps, dans un sens profond, n'existe pas, mais il est une limite quasi-classique d'une certaine théorie de jauge abstraite qui « ne réside nulle part ». Nos observables sont ainsi formées à partir de produits d'autres plus élémentaires, et la théorie doit nous fournir les valeurs moyennes des différentes combinaisons.

Ces développements suggèrent fortement que la conception de l'espace-temps comme variété n'est pas applicable à la limite de la longueur de Planck et que l'espace-temps est une notion *émergente* valide approximativement à des échelles plus larges.

2. Théorie des cordes et les extensions de l'espace

2.1. Introduction

Pour traiter la plupart des problèmes en théorie quantique des champs, nous pouvons procéder de deux façons complémentaires : en utilisant soit l'approche *lagrangienne*, où nous avons recours à l'intégrale de chemin, soit l'approche *hamiltonienne*, appelée aussi la *quantification canonique*[6]. Par la suite, nous voulons décrire la théorie des cordes de ces deux points de vue. Nous nous intéresserons à la théorie des cordes *bosoniques* [6], et plus spécifiquement celles des cordes fermées. Notre but est d'esquisser pourquoi tout se passe mieux dans un espace-temps à 26 dimensions. Classiquement, ce type de corde est simplement une application d'une surface fermée dans l'espace temps.

Dans l'approche lagrangienne[7] de la quantification, nous commençons par choisir la forme de l'action. Nous utilisons la possibilité la plus simple, c'est-à-dire l'aire de la surface. Bien entendu, pour définir l'aire de la surface dans

6 – De ce point de vue, quantifier correspond à construire une algèbre A et une représentation unitaire, telle que la dynamique apparaît comme une symétrie particulière. La dynamique et la symétrie sont données en des formes adjointes, i.e. si les équations classiques sont de la forme $q' = F(q)$, la dynamique est réalisée par le crochet $F(q) = [H, q]$; où H est un élément de l'algèbre A.

7 – Selon Feynman, pour calculer l'amplitude de diffusion, on intègre sur tous les arrangements possibles de branchements des particules. De plus, pour une particule voyageant entre deux événements x et y de l'espace-temps, nous devons admettre au niveau quantique toutes les trajectoires classiques possibles : pour évaluer la propagation d'une particule de x à y, nous intégrons sur tous les chemins possibles entre x et y, tout en utilisant un facteur de poids qui dérive de l'action classique pour le chemin. Une propriété importante des diagrammes de Feynman vus comme des graphes à une dimension est le fait qu'ils soient singuliers au point de branchement. Cette propriété induit deux difficultés centrales : la présence des infinis en théorie quantique des champs et le nombre arbitraire des théories quantiques des champs lié à l'arbitraire dans le choix des coefficients de couplage.

l'espace-temps, il faut que l'espace-temps ait une métrique. Le plus simple est alors de travailler avec un espace de Minkowski de dimension n.

Nous cherchons les équations du mouvement de la corde en extrêmisant l'action. Ces équations impliquent que l'évolution de la corde au cours du temps est décrit comme une collection de lacets constitués d'un matériel parfaitement élastique. En effet, nous remplaçons la trajectoire unidimensionnelle de la particule dans l'espace-temps par l'orbite à deux dimensions d'une corde. Ces cordes peuvent être de taille quelconque mais, sous certaines conditions imposées par les prédictions de la théorie de la constante de Newton et de la constante de la structure fine, cette taille est de l'ordre de 10^{-32} cm. Ainsi, parfois, le remplacement des particules par des cordes n'est pas si crucial, mais dans d'autres cas il est fondamental. La situation est comparable à l'introduction de la constante de Planck en passant de la physique classique à celle quantique : la théorie des cordes introduit une constante fondamentale α qui contrôle la tension de la corde.

Une conséquence de l'introduction du concept de corde est le fait que les diagrammes de Feynman deviennent lisses : alors que des lignes d'Univers se joignent singulièrement aux points d'intersection, les tubes d'Univers se joignent d'une façon lisse[8]. Ce fait résout les difficultés liées aux divergences. De plus, il suffit de comprendre la propagation des cordes *libres* pour étudier

8 – Ajoutons que la théorie des cordes traite aussi d'autre type de singularité. Mais la définition d'une singularité en théorie des cordes est différente de celle en relativité générale, même au niveau classique. En relativité générale, on définit habituellement une singularité en termes d'incomplétude des géodésiques fondée sur le mouvement de particules test. En théorie des cordes, il faut utiliser des cordes test. Ainsi un espace-temps est considéré comme singulier si les cordes test ne se comportent pas bien. (Plus précisément, il faut également que les autres objets étendus en théorie des branes aient une propagation bien définie.) Il est donc possible que certains espace-temps qui en relativité générale sont singuliers ne le soient pas en théorie des cordes.
Un exemple simple serait le quotient de l'espace euclidien par un sous groupe discret du groupe de rotation. L'espace résultant, appelé *orbifold*, a une singularité conique à l'origine. En relativité générale, ceci conduit à une incomplétude des géodésiques, tandis que la théorie des cordes est totalement exempte de ce type de singularité. La raison principale en est que les cordes sont des objets étendus. Cet orbifold a une singularité très douce, mais même les singularités de courbure peuvent disparaître en théorie des cordes. La théorie des cordes a comme solutions exactes le produit d'un espace de Minkowski à 4 dimensions et d'un espace de Calabi-Yau compact. (Un espace de Calabi-Yau est un espace vers lequel les extra dimensions exigées par la théorie des cordes tendent après enroulement. Cet espace est consistent avec les équations de la théorie.) Une variété de Calabi-Yau admet généralement une famille complète de métriques de Ricci plates. On peut donc construire une solution dans laquelle les quatre grandes dimensions restent approximativement plates et la variété de Calabi-Yau change lentement d'une métrique de Ricci plate à une autre. Dans ce processus, l'espace de Calabi-Yau peut avoir une singularité de la courbure. Dans plusieurs cas, ceci peut être vu comme résultant d'un rétrécissement topologique non trivial de S^2 ou S^3 à une surface nulle. On montre ainsi que la théorie reste complétement bien définie. Par conséquent, l'évolution continue à travers la singularité géométrique vers un espace de Calabi-Yau non singulier présent sur l'autre facette.

les interactions. Ainsi, nous n'avons plus besoin de décrire les propriétés des points d'interaction, d'où la non-nécessité de l'espace-temps.

Par conséquent, nous avons un nouveau paradigme : l'espace-temps est un concept dérivé à partir d'une théorie de champs 2-dimensionnelle. En effet, un espace-temps qui obéit aux équations des champs classiques correspond à une théorie des champs à 2 dimensions qui est invariante conforme (*i.e.*, invariante par le changement de mesure de la distance le long de la corde). Si nous calculons les conditions exigées par cette invariance conforme de la théorie quantique dérivée d'un lagrangien, tout en supposant que les champs varient lentement à l'échelle de la corde, nous obtenons généralement des équations covariantes qui sont simplement les équations d'Einstein plus d'autres termes.

Ces lacets vibrent, se séparent et se rejoignent au cours du temps. Il semble peut être plus facile de comprendre comment la corde vibre si l'on se place dans le cadre de l'approche hamiltonienne. Ceci est pourtant subtil puisque la théorie des cordes contient un grand nombre de symétries de jauge[9]. Or, il existe une technique pour obtenir les symétries de jauge de toute théorie à partir de l'action. En l'appliquant à la théorie des cordes, on trouve que deux applications de la surface dans l'espace temps sont physiquement équivalentes si elles diffèrent simplement d'une reparamétrisation de cette surface (décrite dans l'espace temps).

Plusieurs techniques permettent de rendre compte des symétries de jauge dans l'approche hamiltonienne. Une d'elles est l'approche appelée « jauge fixée sur le cône de lumière ». Cela revient à choisir une paramétrisation de notre surface telle que ses deux coordonnées soient reliées de façon simple à deux des coordonnées de l'espace de Minkowski de dimension n. Ceci résulte de l'invariance par reparamétrisation. Mais une fois la paramétrisation fixée, nous n'avons plus la possibilité de la modifier : d'où le terme *fixation de jauge*. L'application de la surface S vers l'espace temps de dimension n se représente comme un champ sur S à n composantes. Dans le choix de la jauge du cône de lumière, deux composantes de ce champ s'expriment simplement en fonction des autres. Ceci nous permet de voir notre corde comme un champ à $n-2$ composantes satisfaisant l'équation d'onde

$$\left(\frac{d^2}{dt^2} - \frac{d^2}{dx^2}\right) X(t,x) = 0,$$

9 – Nom par lequel les physiciens désignent une symétrie qui permet de passer entre différentes descriptions mathématiques qui rendent compte de la même situation physique.

la même que celle qui décrit une corde de violon idéalisée. L'unique différence est que, à la place d'un segment de corde de violon, on a un ensemble de lacets fermés de corde. L'énergie, ou hamiltonien, est alors donnée par l'équation d'onde hamiltonienne usuelle :

$$H = (1/2) \int \left[\left(\frac{dX}{dt} \right)^2 + \left(\frac{dX}{dx} \right)^2 \right] dx.$$

Le premier terme représente l'énergie cinétique de la corde, le deuxième son énergie potentielle due à son étirement. Par la suite, nous considérons les vibrations d'un unique lacet de la corde, sans tenir compte du fait que la corde peut se séparer ou se joindre. La linéarité de l'équation d'onde permet de décomposer toute solution de l'équation d'onde en des ondes sinusoïdales se déplaçant dans des directions opposées, appelées mouvements à droite (*right-movers*) et mouvements à gauche (*left-movers*), ainsi qu'en une solution de la forme :

$$X(t, x) = A + Bt$$

qui décrit le mouvement du centre de masse de la corde. Les mouvements à droite et à gauche n'interagissent pas entre eux ni avec le mouvement du centre de masse. Ainsi il nous suffit d'étudier un de ces mouvements, par exemple celui de droite.

Supposons que le champ X ait une seule composante : les modes de vibrations des mouvements sur la droite deviennent

$$X(t, x) = A \sin(ik(t - x)) + B \cos(ik(t - x))$$

de fréquences $k = 1, 2, 3, \ldots$ Abstraitement, chacun de ces modes de vibrations apparaît comme un oscillateur harmonique de fréquence k, si bien que l'on peut considérer la corde comme une collection d'oscillateurs harmoniques. Par conséquent la quantification[10] de la corde (ou plus précisément des modes des mouvements à droite) consiste à quantifier un groupe d'oscillateurs harmoniques, un oscillateur de fréquence k étant associé à chaque nombre naturel k. Ceci est d'autant plus simple que l'oscillateur harmonique est un des systèmes physiques les plus simples à quantifier. L'oscillateur harmonique quantique de fréquence k a des niveaux d'énergie discrets $k/2, 3k/2, 5k/2\ldots$[11].

10 – En réalité, dans la théorie des cordes on a une triple quantification : (i) remplacer une particule par une corde ; (ii) quantifier la corde, ce qui fait apparaître les équations d'Einstein comme les conditions de renormalisation ; (iii) réaliser les symétries classiques, ce qui est le sens propre de la quantification.

11 – Nous travaillons ici en unités où $\hbar = 1$; sans quoi le facteur $\hbar$ serait requis.

En particulier, l'énergie du plus bas niveau est appelée *point zéro d'énergie* ou *énergie du vide*. En général, il n'est pas trop difficile de le soustraire en redéfinissant le hamiltonien ; mais il est parfois important. Pour obtenir l'expression du point zéro d'énergie totale de tout les modes des mouvements à droite, on additionne les points zéro d'énergie $k/2$ pour toutes les fréquences $k = 1, 2, 3, \ldots$:

$$\frac{(1+2+3+\ldots)}{2}.$$

Bien que cette série soit divergente, nous pouvons, comme nous verrons plus bas, lui assigner des valeurs. En effet, Euler a associé[12] à la série 1 + 2 + 3... la valeur −1/12. Ainsi, l'énergie totale du point zéro est : −1/24. Plus généralement, si nous avons une corde dans un espace temps de Minkowski de dimension n, le champ X a $n-2$ composantes, et l'énergie totale du point zéro est

$$\frac{-(n-2)}{24}.$$

D'autre part, la validité de la théorie des cordes exige que le point zéro d'énergie soit égal à −1[13]. En effet, la subsistance de l'invariance lorentzienne de

12 – La formule d'Euler

$$1+2+3+\ldots = -\frac{1}{12}$$

est un exemple de régularisation de la fonction zéta. La fonction de Riemann zéta est définie par

$$\zeta(s) = \frac{1}{1^s} + \frac{1}{2^s} + \frac{1}{3^s} + \ldots,$$

lorsque la somme converge, et par sa continuation analytique aux valeurs de s pour lesquelles cette somme ne converge plus. Par continuation analytique, on obtient :

$$\zeta(s) = -\frac{1}{12}.$$

On peut utiliser l'équation fonctionnelle pour la fonction de Riemann zéta selon laquelle :

$$F(s) = F(1-s)$$

où

$$F(s) = \pi^{-\frac{s}{2}}\,\Gamma\left(\frac{s}{2}\right)\zeta(s)$$

et Γ est la célèbre fonction telle que $\Gamma(n) = (n-1)!$ pour $n = 1, 2, 3, \ldots$ et $\Gamma(s+1) = s\,\Gamma(s)$ pour tout s. En utilisant

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

et

$$\zeta(2) = \frac{\pi^2}{6}$$

l'équation fonctionnelle implique que $\zeta(-1) = -\frac{1}{12}$. Mais il faudrait , bien sûr pour cela avoir avant tout prouvé l'équation fonctionnelle.

13 – Cela a un rapport avec les subtilités de la fixation de jauge en théorie quantique des champs. Par exemple, certaines symétries au niveau classique peuvent parfaitement être perdues au niveau quantique, à cause de la présence de ces anomalies.

la théorie des cordes au niveau quantique[14], dans le choix de la jauge du cône de lumière, exige cette valeur. Ainsi la théorie des cordes bosoniques est satisfaisante pour

$$\frac{(n-2)}{24} = 1$$

ou, en d'autres termes, si $n = 26$.

2.2. Courbes elliptiques et nombre 24

En plus de la fonction zéta de Riemann, un grand nombre de fonctions spéciales apparaissent dans l'étude des courbes elliptiques. De façon non surprenante, nombre d'entre elles jouent aussi un rôle important en théorie des cordes.

2.2.1. Courbes elliptiques

La fonction sinus est une fonction analytique sur le plan complexe et vérifie la propriété

$$\sin(z + 2\pi) = \sin z.$$

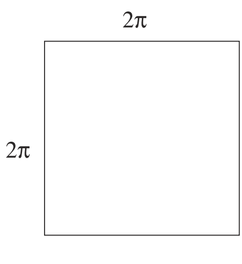
Elle satisfait l'équation différentielle

$$(\sin' z)^2 = 1 - (\sin z)^2.$$

Les fonctions elliptiques la généralisent. Soit $P(z)$ une fonction analytique doublement périodique sur le plan complexe avec

$$P(z + 2\pi) = P(z)$$
$$P(z + 2i\pi) = P(z).$$

Cette fonction reste la même sur le carré $(2\pi\,; 2\pi)$:



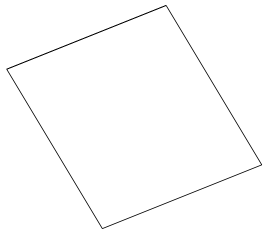
14 – Plus présicément, maintenir les symétries (symétrie conforme et le difféomorphisme) à la source et l'invariance Lorentzienne au but.

Ainsi, nous pouvons la traiter comme une fonction sur le tore résultant des identifications des côtés deux par deux, haut et bas, gauche et droite.

Plus généralement, soit $P(z)$ une fonction doublement périodique sur le plan complexe, de périodes ω_1 et ω_2, *i.e.* :

$$\begin{aligned} P(z+\omega_1) &= P(z) \\ P(z+\omega_2) &= P(z). \end{aligned}$$

Ainsi, $P(z)$ reste invariante en tous les points du réseau $L = n\omega_1 + m\omega_2$, qui peut soit avoir la forme de carré vu plus haut ou comme ceci :



Si $P(z)$ est analytique sur tout le plan (et sans pôle), elle sera bornée sur chaque petit parallélogramme. Comme elle est doublement périodique, elle sera une fonction analytique bornée sur tout le plan complexe, donc constante d'après le théorème de Liouville.

Intéressons-nous au cas non trivial, c'est-à-dire une fonction qui a des pôles. Supposons d'abord ces pôles localisés en les points de réseau :

$$L = n\omega_1 + m\omega_2.$$

Nous pouvons montrer que le cas où $P(z)$ a des pôles de type $\frac{1}{z-w}$ en chacun des points de réseau est en fait trivial. En revanche, si les pôles sont d'ordre 2, *i.e.* de type $\frac{1}{(z-w)^2}$, nous obtenons une fonction périodique. Mais la somme ne converge pas. Néanmoins, après renormalisation, cette somme converge vers une limite $P(z)$ qui est une fonction analytique partout sauf aux pôles d'ordre 2. Ceci n'est autre que la fonction elliptique de Weierstrass qui dépend de ω_1, ω_2 et z. De plus, il est facile de voir que $P(z)$ est une généralisation de la fonction *sinus* et vérifie l'équation différentielle :

$$(P'(z))^2 = 4(P(z))^2 - g_2 P(z) - g_3$$

où g_2 et g_3 sont des constantes qui dépendent de ω_1 et ω_2.

Cependant, nous pouvons voir cette construction selon un autre point de vue. Rappelons nous que $P(z)$ est une fonction sur le tore. Ainsi, $P(z)$ et $P'(z)$ peuvent fournir un système de coordonnées sur le tore. En effet, si $x = P(z)$ et $y = P'(z)$, nous pouvons trouver z sur le tore. Bien entendu, x et y doivent vérifier ainsi l'équation différentielle

$$y^2 = 4x^3 - g_2x - g_3$$

où x et y sont des nombres complexes. L'ensemble des paires (x, y) qui satisfont l'équation cubique forment quelque chose comme un tore, appelé courbe elliptique[15]. De plus le tore peut être vu comme un groupe muni d'une loi additive en additionnant les nombres complexes modulo les éléments de réseau L. Par conséquent, les solutions de l'équation

$$y^2 = 4x^3 - g_2x - g_3$$

forment un groupe.

2.2.2. Fonction de Dedekind

Considérons la fonction de partition de la théorie des cordes bosoniques. La fonction de partition d'un système quantique avec un hamiltonien est définie par :

$$Z(b) = \text{trace}\left(e^{-b\mathcal{H}}\right)$$

où $b > 0$ est l'inverse de la température. Cette fonction est fondamentale en mécanique statistique.

Avant de considérer les cordes bosoniques, examinons tout d'abord la fonction de partition pour un *oscillateur harmonique quantique*. Pour simplifier, on soustrait l'énergie du point zéro de sorte que les niveaux d'énergie deviennent $0, k, 2k, \ldots$ Mathématiquement, ces niveaux d'énergie sont les valeurs propres du hamiltonien de l'oscillateur harmonique $\mathcal{H}$. Ainsi, les valeurs propres de $e^{-b\mathcal{H}}$ sont $1, e^{-bk}, e^{-2bk}, \ldots$ La trace de cet opérateur est simplement la somme de ses valeurs propres :

$$Z(b) = 1 + e^{-bk} + e^{-2bk} + \ldots = \frac{1}{(1 - e^{-bk})}.$$

Ce résultat a été obtenu pour la première fois par Planck 1899. Planck a supposé que l'oscillateur harmonique a des niveaux discrets d'énergie séparés

15 – Généralement toutes les courbes elliptiques peuvent être mises, après changement de variables, sous la forme $y^2 = 4x^3 - g_2x - g_3$.

régulièrement et a ainsi calculé sa fonction de partition afin de comprendre la thermodynamique des champs électromagnétiques.

Considérons le cas des cordes bosoniques[16], où l'on ne prend en compte que les modes de mouvement à droite et où le champ X décrivant les vibrations de la corde n'a qu'une composante. Comme nous l'avons vu auparavant, la corde devient alors équivalente à une collection d'oscillateurs harmoniques quantiques de fonction de partition $\frac{1}{\left(1-e^{-bk}\right)}$. Pour obtenir la fonction de partition d'un système quantique constitué d'un ensemble de parties sans interaction, on multiplie les fonctions de partitions de chaque partie[17]. Donc la fonction de partition de notre corde est :

$$\Pi_{k=1}^{\infty} \frac{1}{\left(1-e^{-bk}\right)}.$$

On considère maintenant l'énergie du point zéro. Pour cela, il faut soustraire 1/24 du hamiltonien de la corde, ce qui revient à multiplier sa fonction de partition par $e^{\frac{b}{24}}$. On parvient alors à

$$Z(b) = e^{\frac{b}{24}} \Pi_{k=1}^{\infty} \frac{1}{\left(1-e^{-bk}\right)},$$

c'est-à-dire la réciproque de la fonction de Dedekind. Cette fonction joue un rôle important dans la théorie des courbes elliptiques. On l'écrit généralement comme une fonction de $q = e^{-b}$, de sorte que :

$$\eta(q) = q^{\frac{1}{24}} \Pi_{k=1}^{\infty} \left(1 - q^k\right).$$

Afin de voir la relation avec les courbes elliptiques, il faut à nouveau changer de variables et définir $q = e^{(2\pi i \tau)}$. En physique, la courbe elliptique est une possibilité pour former une surface à partir d'une corde. Le nombre 1 exprime la distance parcourue par la surface dans la direction spatiale avant de s'enrouler, et le nombre τ exprime le temps écoulé avant qu'elle ne s'enroule[18].

16 – On suppose là encore que le point zéro d'énergie a été soustrait.

17 – Rappelons que la trace d'un produit tensoriel d'opérateurs est le produit de leurs traces.

18 – L'idée d'enroulement dans le temps peut sembler étonnante mais elle est importante en physique. En effet, étudier en mécanique statistique un système à une température donnée est équivalent à étudier la théorie quantique des champs euclidienne d'un système de temps périodique. Cette idée est ce qui relie b et τ.

2.3. Homotopie et dimensions 24

2.3.1. Introduction

Rappelons nous certaines notions sur les groupes d'homotopie de la sphère. Il existe différentes façons topologiques d'enrouler une sphère à m dimensions sur une sphère à k dimensions. Par exemple, pour $m = k = 1$, il s'agit de la manière d'enrouler un cercle sur un cercle. Elles sont classifiées par un nombre entier appelé nombre de nœuds. Pour rendre cette notion plus concrète, imaginons par exemple le cercle unité du plan complexe (l'ensemble des nombres complexes tels que $|z| = 1$). Soit d un entier et f la fonction

$$f(z) = z^d$$

du cercle unité dans lui-même. Si d est positif, cette fonction enroule le cercle unité sur lui-même d fois dans le sens des aiguilles d'une montre. Si d est nul, $f(z) = 1$: une fonction constante sans enroulement.

Nous pouvons montrer que toute fonction continue du cercle dans lui-même peut être déformée continuement pour donner exactement une fonction de la forme $f(z) = z^d$. La théorie des homotopies concerne des déformations continues. Selon cette théorie, nous dirons que deux fonctions d'un espace dans un autre sont homomorphiques si on peut déformer continuement l'une d'elle pour donner l'autre, si elles appartiennent à la même *classe d'homotopie*. Dans le cadre de la topologie, à la place du terme fonction continue, c'est celui d'application qui est plutôt utilisé. Selon ces termes, on peut dire que l'on connaît la classe d'homotopie d'une application d'un cercle dans lui-même, si on connaît son nombre de nœuds.

En dimensions supérieures, nous ne connaissons les classes d'homotopie des applications d'une m-sphère dans une k-sphère que pour certaines valeurs spécifiques de m et k. Pour mieux appréhender cela, on prend les notations standards, utilisant $\Pi_m(X)$ pour désigner l'ensemble des classes d'homotopies des applications de la m-sphère dans l'espace X. Quand $m > 0$, cet ensemble est un groupe appelé m-ième *groupe d'homotopie* de X. Ces groupes sont d'une importance fondamentale en topologie algébrique. Nous nous intéressons à $\pi_m(S^k)$: l'ensemble des classes d'homotopie des façons d'enrouler une m-sphère sur une k-sphère. Dans le cas du cercle, vu plus haut

$$\pi_1(S^1) = \mathbb{Z}$$

où $\mathbb{Z}$ représente des entiers, puisque le nombre de nœuds est un entier. En augmentant le nombre de dimensions, on obtient :

$$\pi_2(S^2) = \mathbb{Z}.$$

En d'autres termes, on peut enrouler une 2-sphère (ou sphère ordinaire) d fois autour d'elle-même pour tout d entier. Soient ϕ et θ les coordonnées sphériques décrivant un point d'une sphère à partir du pôle nord et la latitude respectivement, il suffit d'introduire la fonction :

$$f(\phi, \theta) = (\phi, d, \theta).$$

Toute application de S^2 dans elle-même est homotopique à exactement une de ces fonctions. La même idée de base peut être généralisée à toute dimension :

$$\pi_k(S^k) = \mathbb{Z} \quad \text{pour tout} \quad k \geq 1.$$

Ainsi, il existe toujours un entier k qui joue le rôle de nombre de nœuds d'une application d'une n-sphère dans elle-même. En mathématiques ce nombre est appelé degré.

Dans le cas où m est différente de k,

$$\pi_m(S^k) = 0 \quad \text{pour tout} \quad m < k,$$

où 0 désigne habituellement l'ensemble ne contenant qu'un élément. Cela signifie qu'il existe seulement une classe d'homotopie pour appliquer une sphère dans une sphère de dimension supérieure.

Dans le cas où on applique une sphère dans une sphère de dimension inférieure, nous remarquons par exemple, qu'il existe seulement une classe d'homotopie d'applications d'une 2-sphère dans un cercle :

$$\pi_2(S^1) = 0.$$

Cela semble évident mais si on augmente le nombre de dimensions de 1, on obtient[19]

$$\pi_3(S^2) = \mathbb{Z}.$$

Il existe un moyen de déterminer un entier appelé invariant de Hopf qui conserve la trace du nombre de classes d'homotopie d'une 3-sphère dans une

19 – Ce résultat fut un choc important quand Heinz Hopf le découvrit dans les années 1930. En effet, auparavant personne ne savait combien de classes d'homotopie il y avait dans ce cas.

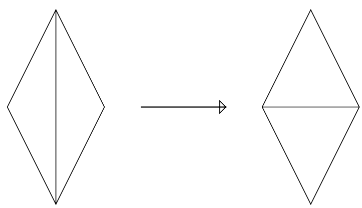
2-sphère. Il existe différentes façons de l'obtenir. Soit f l'application régulière qui envoie $f : S^3 \to S^2$. Ainsi la plupart des points p de S^2 ont la propriété que les points $x \in S^3$ tels que $f(x) = p$ forment un lien : une collection de noeuds dans S^3. Si on prend deux points différents de S^2 avec cette propriété, alors on obtient deux liens.

À partir de ces deux liens, on peut calculer un entier appelé nombre de liens : par exemple, on peut simplement tracer ces deux liens et compter le nombre de fois que l'un passe au-dessus et an-dessous de l'autre (avec le signe plus ou moins approprié). On peut montrer que ce nombre ne dépend pas du choix des deux points. De plus, il dépend seulement de la classe d'homotopie de f. Il est appelé *nombre invariant de Hopf* de f.

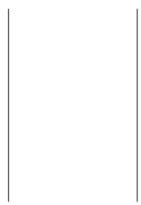
En considérant une dimension supplémentaire, on montre :

$$\pi_4(S^3) = \mathbb{Z}/2$$

où $\mathbb{Z}/2$ est le groupe à deux éléments, habituellement notés 0 et 1, avec l'addition modulo 2. Ceci résulte de fait qu'en dimension 4 les liens sont faciles à dénouer. On peut dénouer quelque chose comme :



et le mettre sous la forme



Ainsi, le nombre de liens en dimension 4 est défini modulo 2 et l'invariant de Hopf est donc lui-même défini modulo 2. La même situation se reproduit en dimension supérieure :

$$\pi_{k+1}(S^k) = \mathbb{Z}/2 \quad \text{pour tout} \quad k \geq 3.$$

Ceci illustre un fait important, quand le nombre de dimensions devient grand, nous avons toujours de « la place » pour dénouer et par conséquent les groupes d'homotopie tout en se simplifiant, finissent par être stabilisés. C'est le point important qui sous-tend la théorie d'homotopie stable.

Considérons d'autres exemples. On a :

$$\begin{aligned}\pi_3(S^1) &= 0\\ \pi_4(S^2) &= \mathbb{Z}/2\\ \pi_5(S^3) &= \mathbb{Z}/2\\ \pi_6(S^4) &= \mathbb{Z}/2\end{aligned}$$

et ainsi

$$\pi_{k+2}(S^k) = \mathbb{Z}/2 \quad \text{pour tout} \quad k \geq 2.$$

Enrouler une 4-sphère sur une 2-sphère est un problème intéréssant. En composant des éléments des groupes d'homotopie des sphères, on en déduit de nouveaux. Par exemple, pour obtenir un élément non trivial de $\pi_4(S^2)$. On considère l'application $f : S^3 \to S^2$ qui génère $\pi_3(S^2)$ et on le compose avec une application[20] $g : S^4 \to S^3$ générant $\pi_4(S^3)$ pour déduire l'application désirée de S^4 dans S^2. De même :

$$\begin{aligned}\pi_4(S^1) &= 0\\ \pi_5(S^2) &= \mathbb{Z}/2\\ \pi_6(S^3) &= \mathbb{Z}/12\\ \pi_7(S^4) &= \mathbb{Z}/2+\mathbb{Z}/12\\ \pi_8(S^5) &= \mathbb{Z}/24\\ \pi_9(S^6) &= \mathbb{Z}/24\end{aligned}$$

qui peuvent se réécrire :

$$\pi_{k+3} = \mathbb{Z}/24 \quad \text{pour tout} \quad k \geq 5.$$

Notez la présence du nombre 24. La précédente équation stipule que pour k suffisamment grand, il y a exactement 24 classes d'homotopie d'applications d'une $(k+3)$-sphère dans une k-sphère. Il faut noter que

$$\pi_m = 0 \quad \text{pour tout} \quad m \geq 2.$$

Ensuite, il existe une formule quand les groupes d'homotopie sont fixés. Ainsi $\pi_{k+^n}(S^k)$ est indépendant de k aussi longtemps que $k \geq n+2$.

Les groupes d'homotopie peuvent être stabilisés avant, comme nous l'avons vu, pour $n = 2$ mais pas après. En général, ils se stabilisent pour

20 – g est obtenue à partir de f par suspension.

$k = n + 2$. Il y a une raison simple à ce fait. On remarque que $\pi_{k+1}(S^k)$ se stabilise pour $k = 3$ parce qu'il est simple de dénouer en dimension 4 et plus. De même, $\pi_{k+n}(S^k)$ doit se stabiliser à $k = n + 2$ car il est facile de dénouer les noeuds des surfaces de dimension n dans un espace de $2n + 2$ dimensions ou plus.

2.3.2. Relation avec la théorie des cordes

La surface d'univers de la corde est une surface de Riemann. On note $\mathcal{M}(g, n)$ l'espace de module[21] des surfaces de Riemann de genre g et ayant n points marqués (ou une perforation) et $\mathcal{M}(g, n)$ le groupe de classes

21 – L'espace de Teichmueller et l'espace de module sont deux éléments mathématiques simples et nécessaires à la physique d'espace temps de dimension 2. Deux est un nombre de dimensions suffisamment bas pour comprendre ce qui se passe dans des problèmes qui deviennent véritablement plus compliqués à un nombre supérieur de dimensions. Il existe de nombreuses façons de décrire l'espace de Teichmueller et l'espace de module. Peut-être la plus simple est la suivante. Partant d'un tore à g anses, ou surface de genre g, on peut le transformer en surface de Riemann si on le recouvre d'un ensemble de cartes, chacune d'elles s'apparentant à une partie du plan complexe et telle que la fonction de changement de coordonnées reliant deux cartes qui se superposent est analytique, dans le sens usuel des variables complexes. La surface de Riemann la plus simple est la sphère de Riemann qui est de genre zéro. On l'obtient en partant du plan complexe et en identifiant l'un de ses points à l'infini. Si on a une surface de Riemann, on peut déterminer si une fonction à valeurs complexes définie sur elle est analytique en examinant simplement son comportement sur les cartes. Ainsi on peut appliquer l'analyse complexe usuelle sur les surfaces de genre g aussi longtemps qu'il est possible de la transformer en surface de Riemann. Transformer une surface de genre g en surface de Riemann revient à lui donner une structure complexe.

Nous disons que deux surfaces de Riemann sont les mêmes s'il existe entre elles une application bijective et inversible qui soit partout analytique et d'inverse analytique. Une telle application s'appelle application *biholomorphe*, holomorphe étant simplement une autre désignation pour analytique. Il existe un théorème de Riemann célèbre, selon lequel pour une surface de genre 0, il y a une seule manière de lui donner une structure complexe ; toute surface de Riemann de genre 0 étant biholomorphiquement équivalente à la sphère de Riemann. Cependant, pour des genres supérieurs, il y a un nombre infini de façons de donner une structure complexe à une surface de Riemann. On peut imaginer l'espace de toutes ces manières : c'est précisément l'espace de module de genre g. Le premier problème clé dans cette théorie est de décrire concrètement l'espace de module. La théorie des cordes donne un grand nombre de motivations pour cela puisque la surface d'univers d'une corde, c'est-à-dire la corde évoluant dans l'espace temps, est une surface, et l'intégrale de chemin de Feynman en théorie des cordes impose l'intégration sur toutes les structures complexes sur cette surface.

On remarque que l'espace de module peut être vu comme l'espace des classes d'équivalence des structures complexes sur une surface donnée de genre g où deux structures complexes sont les mêmes si elles sont biholomorphiquement équivalentes. L'espace de Teichmueller est quand à lui défini en utilisant une notion plus fine pour l'équivalence. On remarque que toute application biholomorphe est un *difféomorphisme*, c'est-à-dire une application lisse d'inverse lisse. Ainsi il préserve l'orientation puisqu'une application qui renverse l'orientation, telle que la conjugaison complexe, n'est pas holomorphe. Par conséquent, un difféomorphisme de surface est un difféomorphisme qui préserve l'orientation. On dit qu'un difféomorphisme f est connecté à l'identité s'il existe une famille de difféomorphismes à un paramètre partant de f et finissant au difféomorphisme identité. En d'autres termes, un difféomorphisme est connecté à l'identité si on peut passer à l'identité graduellement sans avoir à jamais couper la surface.

Un espace de Teichmueller peut être défini comme l'espace des classes d'équivalence des structures complexes sur une surface donnée de genre g, où deux structures complexes sont considérées comme égales si elles sont biholomorphiquement équivalentes par un difféomorphisme connecté à l'identité.

d'applications correspondant. Puisque $\mathcal{M}(g, n)$ est le quotient de l'espace de Teichmueller par $\mathcal{M}(g, n)$ et que l'espace de Teichmueller est contractile, on a $\mathcal{M}(g, n) = \mathcal{BM}(g, n)$ où $\mathcal{B}$ représente l'espace classifié. Il y a une inclusion naturelle

$$\mathcal{M}(g, n) \to \mathcal{M}(g + 1, n),$$

définie en cousant un tore à deux perforations sur une surface de genre g avec n perforations, ce qui augmente le genre de 1. On introduit $\mathcal{M}(\infty, n)$ comme la limite directe pour $g \to \infty$ et

$$\mathcal{M}(\infty, n) \to \mathcal{BM}(\infty, n).$$

On remarque maintenant que $\mathcal{M}(\infty, 1)$ est muni d'une sorte de produit. La raison en est que les produits

$$\mathcal{M}(g, 1) \times \mathcal{M}(h, 1) \to \mathcal{M}(g + h, 1)$$

permettent de coudre deux surfaces ensemble avec une sphère à trois perforations. En utilisant ce produit, on peut définir le complété du groupe :

$$\mathcal{M}(\infty, 1)^+,$$

qui donne :

$$\pi_3(\mathcal{M}(\infty, 1)^+) = \mathbb{Z}/24 + H$$

pour un groupe inconnu H. Comme ce résultat concerne les groupes de classes d'applications de surfaces, il doit être relié à la façon dont la théorie des champs conformes donne toujours des représentations projectives de ces groupes de classe d'applications avec l'ambiguïté de phase de la forme $e^{2\pi ci/24}$, où c est la charge centrale.

La démonstration est directement liée aux groupes stables d'homotopies de la sphère. On peut utiliser les applications explicites entre le groupe stable d'homotopies des sphères :

$$\pi_{k+3}(S^k) = \mathbb{Z}/24 \quad \text{pour tout} \quad k \leq 5$$

et $\pi_3(\mathcal{M}(\infty, 1)^+)$. La véritable énigme reste le rapport entre $\pi_3(\mathcal{M}(\infty, n)^+)$ et les extensions centrales de $\mathcal{M}(g, n)$ pour g fini.

2.4. Supercordes et la conjecture « Monstre de Moonshine »

Le problème est apparu en 1978, quand John McKay de l'Université de Concordia a remarqué une coïncidence étrange : « Je lisais un livre du

XIX^e siècle sur les fonctions elliptiques modulaires et j’ai remarqué quelque chose d’étrange dans le développement de la fonction j [22]. » (McKay.)

Cette coïncidence apparaît à McKay quand il examine les coefficients de la fonction j, écrite sous la forme d’une somme infinie. Le troisième coefficient était 196,884. Or, 196,883 est le second nombre dans la table du *Monstre*, le premier étant 1. Le groupe monstre[23] simple est le plus grand groupe isolé, simple, fini connu. Il représente les symétries de quelque chose sur lequel les mathématiciens n’ont pas d’indice. Quelque chose de trop compliqué pour le nommer objet géométrique parce que le Monstre vit non pas à 3 dimensions mais à 196,883. Et en 21, 296, 876 dimensions ainsi que toutes les autres dimensions de la colonne de la table. Quel que soit l’objet qui donne jour au groupe monstrueux, il doit être excessivement symétrique.

Chaque coefficient de la fonction modulaire est une simple somme de nombres de cette liste de dimensions dans lesquels le Monstre vit.

Conway et d’autres suggèrent [4] que ces connexions ne sont pas une coïncidence mais l’effet d’une unité plus profonde. Appelant cette conjecture Moonshine, une nouvelle spécialité mathématique naît pour essayer de la prouver.

Borcherds [2, 3], après huit années de travail sur ce problème, trouva en 1989 la troisième pièce de cette énigme : la connexion était la théorie des cordes par l’intermédiaire du nombre 26. Comme nous l’avons déjà vu, en théorie des cordes, l’idée de base est que les particules élémentaires ne sont plus fondamentales mais sont composées de lacets de cordes de dimension 1. Pour comprendre comment les lois de la nature opèrent dans différentes théories, les physiciens utilisent des diagrammes de lignes. Chaque ligne représente la trace de la particule et les interactions, ou vortex, interviennent quand les particules entrent en collision ou interagissent. En théorie des cordes, les lacets remplacent les points, de sorte que maintenant les diagrammes ne sont plus réalisés à partir de lignes mais de tubes. Les mathématiques utilisées en théorie des cordes décrivent ce qui se passe quand ces tubes se rencontrent par l’intermédiaire de l’algèbre de vertex.

22 – Les fonctions modulaires elliptiques apparaissent quand on commence à étudier les surfaces de tores qui résultent de l’enroulement du plan complexe. Sur une feuille de papier, on peut dénombrer les colonnes avec les nombres entiers (1, 2, 3, ...) et les lignes par les nombres imaginaires ($1i, 2i, 3i, \ldots$). On peut replier la feuille de telle sorte à joindre la fin des tubes du tore pour faire des tores de tailles différentes. En d’autres termes, si on se donne une taille pour le tore, on peut alors utiliser la fonction j pour convertir cette taille en un nombre complexe. Bien que la fonction j semble obscure, c’est un outil très utile en mathématique et physique.

23 – Les symétries des objets géométriques et autres constructions mathématiques forment les éléments des groupes finis. Une théorie des cordes particulières appliquée à un tore replié en dimensions 26 a plus de 1054 symétries et produit le groupe monstre.

Quand on essaie de réaliser ces calculs en théorie des cordes, il faut que certains éléments s'annulent entre eux, ce qui peut seulement arriver quand il y 24 extra dimensions, soit au total 26 (en comptant le temps et la corde elle-même), exactement ce qui est nécessaire au Monstre. « Si la dimension critique de la théorie des corde était autre chose que 26, je n'aurais pas pu prouver les conjectures moonshine. » (Borcherds.)

En fait, Borcherds en inventant l'algèbre de vertex, a essentiellement donné les lois de la théorie des cordes. L'algèbre de vertex décrit une corde se déplaçant dans un espace à 26 dimensions qui a la caractéristique unique que toutes les dimensions sont enroulées. Cela s'apparente à un tore enroulé sur lui-même de façon simple, en utilisant une technique qui fonctionne uniquement à 26 dimensions.

Dans le cas de tores plus compliqués, c'est la fonction j qui agit. Et Borcherds a montré que le Monstre est simplement le groupe de toutes les symétries d'une théorie des cordes particulière, théorie qui n'a certainement rien à voir avec le monde dans lequel nous évoluons.

Remerciements : Je remercie Marc Lachièze-Rey qui, par ses remarques, a contribué à enrichir ce texte.

Bibliographie

[1] Ashtekar, A., New available for classical and quantum gravity, *Phys. Rev. Lett*. 57 (1986) 2444–2247.

[2] Borcherds, R., Vertex algebra, Kac-Moody algebras and the monster, *Proc. Natl. Acad. Sci. USA* 83 (1986) 3068–3071.

[3] Borcherds, R., *Sporadic groups and string theory*, *in* Proceedings of the first European congress of Mathematics, Paris, July 1982, Joseph, A., *et al.* (Eds.), vol. 1, Birkhauser, 411–421, 1994.

[4] Borcherds, R., Conway, J.H., Queen, L., Sloane, M.J.A., A monster lie algebra? *Adv. Math*. 53 (1984) 75–79.

[5] Giveon, A., Witten, E., Mirror symmetry as a Gauge symmetry, *Phys. Lett*. B332 (1990) 44–50, `hep-th/9404154`.

[6] Green M.B., Schwarz, J.H., Witten, E., *Superstring theory*, vol. 1, Cambridge University Press, 1987.

[7] Greene, B.R., Plesser, M.R., *Mirror manifolds: a brief review and progress report*, `hep-th/91110014`.

[8] Hull, C.M., Townsend, P.K., Unity of superstrigs dualities, *Nucl. Phys.* B438 (1995) 109–137.

[9] Kaiser, D., More roots of complementarity: Kantian aspects and influence, *Studies in history and philosophy of science* 23 (1992) 213–239.

[10] Manin, Y.I., *Reflections on arithmetical physics*, *in* Conformal invariance and string theory, Dita, P., Georgescu, V., Academic Press, 1989.

[11] Montonen C., Olive, D.I., Magnetic monopoles as Gauge particles? *Phys. Lett.* B72 (1977) 117–120.

[12] Petitot, J., Actuality of transcendantal æsthetics for modern physics, 1830-1930, a ccentury of geometry, *Lectures Notes in Physics* (1992) 402.

[13] Polchinski, J., Recent results in string duality, *Prog. Theor. Phys. Suppl.* 123 (1996) 9–18, `hep-th/9511157`.

[14] Rovelli, C., Smolin, L., Loop space representation of quantum general relativity, *Nucl. Phys.* B331 (1990) 80–132.

[15] Schwarz, J.H. (Ed.), *Superstrings, the first 15 years of superstring theory*, vols. 1 et 2, World Scientific, 1985.

[16] Wilson, K.G., *Phys. Rev.* D10 (1974) 2445.

[17] Witten, E., String theory in various dimensions, *Nucl. Phys.* B443 (1995) 85–126, `hep-th/9503124`.

15

Espace et observateurs en cosmologie

Marc Lachièze-Rey

1. Introduction

La relativité générale et la cosmologie relativiste considèrent l'*espace-temps* comme arène de la physique. Cependant, les physiciens essaient depuis longtemps d'y définir séparément *espace* et *temps*. Si les problèmes de relativité générale et de cosmologie peuvent se passer d'une telle prescription, qui apparaît ainsi plutôt académique (il s'agit du choix d'un *système de référence global*), la littérature abonde de références à l'espace, par exemple lorque l'on déclare que l'espace est plat (ou non), homogène (ou non) dans un modèle cosmologique donné. Par ailleurs, la physique quantique, ou plutôt son interprétation, requiert le plus souvent un découpage de l'espace-temps en espace et temps. D'où la nécessité de définir un système de référence convenable. Le simple exemple de deux observateurs de vitesses différentes, dans l'espace-temps de Minkowski, montre qu'*une définition de l'espace doit dépendre de l'observateur*.

Un observateur a besoin d'un *repère* pour faire de la physique dans son environment. En relativité générale, l'espace associé à un observateur est défini *localement* sans ambigüité, par orthogonalité à sa ligne d'univers, c'est-à-dire à sa vitesse u. Mais il y a de multiples manières d'étendre cette définition au-delà d'un voisinage local. Si l'on veut envisager des procédures de synchronisation, qui mettent en jeu l'observateur à différents moments de son histoire, il sera nécessaire de définir un repère *global*, aux propriétés convenables tout le long de sa ligne d'univers *O*, ou au moins d'une partie de celle-ci : par exemple d'y être de type minkowskien. Cela est loin d'être suffisant pour déterminer le repère. Cet article présente une prescription qui associe à tout observateur, défini par sa ligne d'univers, un unique système de référence muni de propriétés particulières.

Un système de référence implique en premier lieu une famille d'hypersurfaces spatiales, orthogonales à *O*. Mais une telle famille peut être choisie de différentes manières. Par exemple, parmi les prescriptions les plus populaires, les *coordonnées de Fermi* impliquent les hypersurfaces engendrées par les géodésiques de genre espace, orthogonales à *O*. Mais il est bien connu que les surfaces spatiales ainsi définies se coupent, même dans une situation aussi simple que pour l'observateur de Langevin (voir ci-dessous). Ceci interdit une définition du temps valide loin de *O* (différentes valeurs du temps seraient assignées à un même événement). En cosmologie, un autre choix populaire consiste à sélectionner les hypersurfaces spatiales orthogonales d'homogénéité. Mais de telles hypersurfaces n'existent pas dans tous les espaces-temps, et un tel choix n'a guère de sens lorsque l'observateur lui-même brise les symétries spatiales, par son accélération ou sa rotation. La prescription proposée ici ne souffre pas de ces inconvénients. En outre, elle est la seule à obéir à un « critère de simultanéité », absent de toutes les autres prescriptions (sauf dans le voisinage immédiat de *O*) : les différents points de l'espace sont vus comme *simultanés* (définition précise ci-dessous) par l'observateur. Elle a encore pour avantage de ne dépendre que de la structure conforme de la métrique (à l'exception du temps propre de l'observateur). Enfin, sa validité est plus étendue que celle des autres propositions (elle s'applique, par exemple, en l'absence d'homogénéité spatiale).

Cette prescription a déjà été appliquée *de facto* dans certaines circonstances, comme pour l'observateur de Rindler (voir ci-dessous), mais sa validité s'étend bien au-delà, sans conditions trop restrictives : pour tout observateur (inertiel ou non), dans tout espace-temps (avec certaines conditions, voir ci-dessous) elle définit uniquement un « espace de simultanéité » (en bref, espace) Σ_τ à chaque moment τ. Les Σ_τ ne se coupent pas, même dans les circonstances où les hypersurfaces de Fermi le font, comme par exemple pour l'observateur de Langevin (voir ci-dessous), et elles sont définies même en l'absence d'homogénéité spatiale. Cela permet d'étendre la validité du temps propre de l'observateur à tout l'espace-temps. Dans les modèles de Friedmann-Lemaître, l'espace ainsi défini ne coincide pas avec l'idée intuitive que l'on peut s'en faire.

De nombreuses tentatives de définir une procédure de quantification dans l'espace-temps courbe, et/ou pour un observateur non inertiel (voir par exemple Birrel et Davis [2]), impliquent plus ou moins explicitement un découpage espace-temps. Cela est spécialement important pour donner une interprétation physique des états quantiques en termes de fréquences ou de particules. Je montre par exemple que l'interprétation courante de l'effet Unruh, basée sur les coordonnées de Rindler, correspond en fait à la présente prescription.

Toutes les quantités introduites ici sont *covariantes*, y compris celles qui dépendent de l'observateur comme sa vitesse, accélération, ligne d'univers. Une quantité observable (par exemple l'énergie) est une combinaison de quantités covariantes associées à l'observateur (sa vitesse u) avec une quantité covariante associée au système observé (le tenseur d'énergie-impulsion). La définition d'un système *global* associé à un observateur doit permettre de définir convenablement des quantités non locales qui lui sont associées.

Le but de ce travail est de construire un système de référence pour un observateur en situation cosmologique. Sa validité concerne donc la partie de l'espace-temps (les événements) qui lui est causalement liée dans les deux directions temporelles, son *enveloppe causale*. Par abus de langage, j'appellerai « espace-temps » l'ensemble $\mathcal{M}_0$ des événements à l'intérieur de l'horizon des particules et de l'horizon des événements de l'observateur (s'ils existent). Cette restriction s'applique en fait à toute construction d'un système de référence global. En second lieu, les arguments de synchronicité requièrent que les géodésiques nulles n'admettent pas de points conjugués (bien que cette restriction puisse sans doute être levée). Cela exclut la présence de lentilles gravitationnelles (incompatible également avec les autres prescriptions). Puisque cette analyse concerne un observateur en situation cosmologique (adaptée à la quantification), ou l'étude du développement perturbatif d'irrégularités, cette condition ne semble pas trop restrictive. (Je souligne encore que la validité de cette approche est plus large que celle des autres propositions.)

Dans la section 2, j'introduis la définition de l'espace de simultanéité, et les notions reliées. Je montre la possibilié de définir un système de référence global adapté à un observateur, ainsi qu'une congruence d'observateurs associés. En section 3, j'applique ces résultats aux observateurs inertiel et de Langevin dans l'espace-temps de Minkowski. La section 4 considère l'observateur de Rindler dans l'espace-temps de Minkowski, et l'effet Unruh associé. La section 5 considère les modèles cosmologiques de Friedmann-Lemaître.

2. Système global de référence pour un observateur

2.1. L'observateur accéléré

L'observateur le plus général est défini par sa ligne de temps O, paramétrisée par son temps propre τ. Ceci définit sa vitesse $u(\tau)$ et son accélération $a(\tau) := \nabla_\tau u$ partout sur O. J'appelle $\hat{a}$ le vecteur unitaire parallèle (et dans la même direction) à l'accélération. Je ne considèrerai ici que des observateurs *sans rotation*, pour lesquels $\nabla_\tau \hat{a} := \nabla_\tau \hat{e}_1 = |a|\ u$ (toute la procédure s'étend sans difficulté au cas avec rotation). Le transport le long de la ligne d'univers correspond à une rotation de Lorentz dans le plan u, a.

En définissant $\hat{e}_0(\tau) = u(\tau)$ et $\hat{e}_1(\tau) = \hat{a}(\tau)$, le transport s'exprime par

$$\nabla_\tau \hat{e}_A = \Omega\ \hat{e}_A \quad (A = 1, 2), \tag{15.1}$$

où $(\Omega\ V)^\mu := a^{[\mu}\ u^{\nu]}\ V_\nu$ pour l'observateur sans rotation (spatiale).

On peut associer deux repères naturels orthonormés à l'observateur sans rotation, le long de sa ligne d'univers. Le premier, E, est transporté parallèlement mais, pour un observateur non inertiel, aucun vecteur de ce repère ne coïncide avec la vitesse. Le second, le repère Fermi-transporté f, est tel que $f_0 := u$. En l'absence de rotation spatiale, $f_1 = a/|a|$ et la loi $\nabla_\tau \hat{e}_A = \Omega.\hat{e}_A$ s'applique pour $A = 0, 1, 2, 3$.

Pour les *observateurs sans rotation*, les sections spatiales (ou espaces) conservent la symétrie axiale autour de l'unique direction spatiale définie par le vecteur accélération. Pour un observateur inertiel (que l'on considère le plus souvent ici, sauf en sections 3.1 et 4), le problème est purement bidimensionnel, à symétrie sphérique spatiale. La plupart des calculs présentés ici restent valables dans le cas général à 4-dimensions, y compris pour l'observateur en rotation.

2.2. Définition de l'espace

Un espace-temps admet en général plusieurs foliations temporelles (avec des sections spatiales associées) compatibles avec la ligne d'univers d'un observateur donné O. Je vais montrer que l'imposition de la synchronicité permet d'en sélectionner une seule, ce qui fournit une définition globale de l'*espace*. Ceci nécessite de considérer l'histoire complète de l'observateur, y compris son futur. Par la suite, je supposerai que la structure causale de l'espace-temps n'admet que des cônes de lumières sans replis et sans points conjugués. Ces restrictions sont appropriées à la cosmologie, et caractérisent un espace-temps approprié pour la quantification. Cela exclut un espace multi-connexe. Dans tout l'article, je noterai $\tilde{v}$ la une-forme métrique-duale à un vecteur v : $\langle \tilde{v}, v \rangle = g(v, v) := |v|^2$.

Il est bien connu qu'il est impossible de définir la simultanéité absolue en relativité restreinte ou en relativité générale. Cependant, une prescription *locale* de simultanéité ou, mieux, de *synchronicité, du point de vue d'un observateur* est largement utilisée (voir, *e.g.*, [6]). Cela définit un découpage *local* espace-temps pour cet observateur. La construction présentée ici se fonde sur l'extension de cette prescription au-delà d'un voisinage local, avec des arguments de synchronicité parfaitement opérationnels.

Pour un observateur O, je définis Σ_τ, son *hypersurface de synchronicité* (HS) au temps propre τ, comme l'ensemble des événements reliés par des géodésiques nulles, à la fois à $O(\tau + \delta)$ et à $O(\tau - \delta)$, où δ est un intervalle arbitraire de temps propre pour O :

$$\Sigma_\tau = \cup_\delta \left[I^{\text{future}}(\tau - \delta) \cap I^{\text{past}}(\tau + \delta) \right], \tag{15.2}$$

où $I^{\text{past}}(\tau)$ [resp. $I^{\text{future}}(\tau)$] est le cône nul passé [futur] de l'observateur au temps propre τ.

Avec les restrictions énoncées, les surfaces Σ_τ, pour les différentes valeurs de τ, remplissent totalement $\mathcal{M}_0$. Cela permet d'étendre le champ u à la totalité de $\mathcal{M}_0$, en lui imposant d'être partout unitaire ($u.u = 1$) et orthogonal à $\Sigma(\tau)$ (figure 15.1). Le vecteur u constitue une foliation de $\mathcal{M}_0$, avec les Σ_τ comme surfaces transverses (orthogonales). Chaque ligne intégrale de u sera nommée d'après son intersection $\mathbf{R}$ avec $\Sigma(\tau = 1)$ (par exemple), si bien que chaque point de $\mathcal{M}_0$ peut être écrit comme $(\tau, \mathbf{R})$. En pratique, $R := |\mathbf{R}|$ (évalué dans la surface spatiale Σ_τ, voir ci-dessous) peut être vu comme la moitié de l'intervalle de temps propre entre l'émission d'un flash de lumière qui illumine un objet cosmique situé en x, etl'observation de l'image qui en réulte ; ou, de manière équivalente, la moitié de l'intervalle de temps propre entre l'émission d'un rayon lumineux (ou signal radar) par un observateur, et sa réception après réflexion sur l'objet en x. La synchronicité impose que « l'espace pour $O(\tau)$ » est Σ_τ.

Les lignes intégrales de u définissent une famille unique d'observateurs, les *observateurs canoniques* associés à O. Ceux-ci ne partagent pas nécessairement les propriétés de O. Par exemple, ils ne sont pas géodésiques, même si

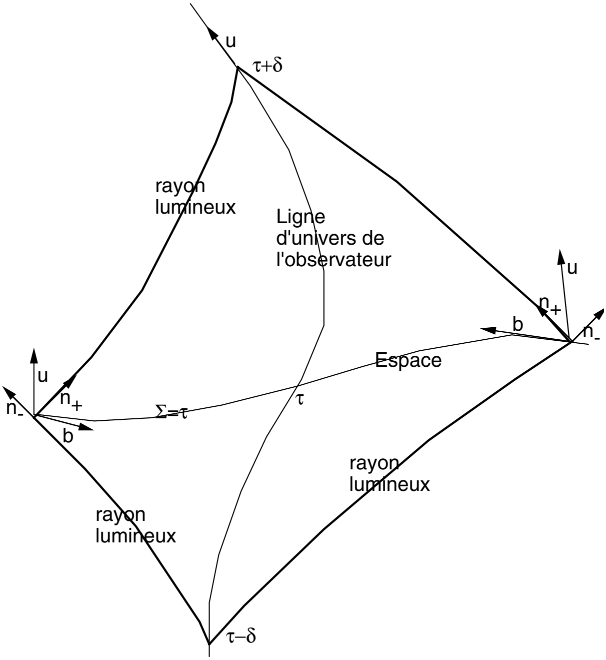


FIG. 15.1 – En chaque point x, les rayons lumineux issus de l'observateur (dans le passé comme dans le futur) définissent les vecteurs n^{+} et n^{-}. La vitesse u de l'observateur est transportée (non parallèlement) pour donner les vecteurs u et b en x.

O l'est, dans le cas avec expansion. Le long des lignes d'univers de u, le temps propre est appelé t, avec $t = \tau$ le long de O.

2.3. Transport le long des rayons lumineux

Les propriétés de l'espace sont définies à partir de celles de la ligne d'univers de l'observateur, transportées par les rayons lumineux passés et futurs. Pour les explorer, nous définissons les deux fonctions nulles $\mathcal{N}_{-}(x)$ et $\mathcal{N}_{+}(x)$, telles que la valeur $\mathcal{N}_{-}(x)$ [resp. $\mathcal{N}_{+}(x)$] soit le temps propre $\tau - \delta$ de l'observateur O, lorsqu'il émet un rayon lumineux qui atteint x [resp. $\tau + \delta$, lorsqu'il reçoit un rayon lumineux émis de x]. Autrement dit, l'hypersurface nulle $\mathcal{N}_{-}(x) = \tau$ [resp. $\mathcal{N}_{+}(x) = \tau$] est le cône du futur [resp. du passé] de l'observateur à son temps propre τ. On définit ses générateurs (nuls) (passé et futur) comme $\tilde{n}_{\pm} = \nabla \mathcal{N}_{\pm} = d\mathcal{N}_{\pm}$. Tous deux pointent vers le futur et sont normalisés de manière à ce que la fréquence émise ou reçue par l'observateur soit l'unité (voir ci-dessous).

On montre facilement que Σ_τ est l'ensemble des points x vérifiant l'équation

$$T(x) := [\mathcal{N}_-(x) + \mathcal{N}_+(x)]/2 = \tau\,. \tag{15.3}$$

Pour tout point, T constitue une coordonnée de temps naturelle. En outre, on définit l'hypersurface cylindrique déformée d'équation

$$R(x) := [\mathcal{N}_+(x) - \mathcal{N}_-(x)]/2 = \delta \tag{15.4}$$

comme l'ensemble des événements situés à un intervalle de temps propre constant δ de l'observateur parcourant sa ligne d'univers. On a défini « l'intervalle de temps propre » (PT-intervalle) δ d'un événement x comme la moitié de l'intervalle de temps propre de l'observateur, entre les instants $\tau_1 = \tau + \delta$ et $\tau_2 = \tau - \delta$, lorsqu'il reçoit et émet les rayons lumineux émis et reçu par l'événement x. Pour tout point, $R = |\mathbf{R}|$ définit naturellement une coordonnée spatiale radiale.

Compte tenu de la normalisation,

$$dT = (\tilde{n}_- + \tilde{n}_+)/2 \text{ et } dR = (\tilde{n}_+ - \tilde{n}_-)/2. \tag{15.5}$$

On vérifie facilement que $dT \cdot dR = 0$, et

$$dT \cdot dT = -dR \cdot dR := N^{-2} = n_+ \cdot n_-/2, \tag{15.6}$$

ce qui définit la *fonction lapse* N associée à cette foliation. Comme dT est orthogonal à Σ, $\tilde{u} = N\, dT$. Et $u^2 = 1$ implique $N\, u \cdot dT = 1$. Puisque $\tau = T$ le long de $\mathcal{O}$, $N = 1$ sur $\mathcal{O}$.

Partout (sauf sur O), définissons

$$\tilde{b} := N\, dR = \tilde{u} - N\, \tilde{n}_- = -\tilde{u} + N\, \tilde{n}_+. \tag{15.7}$$

Alors $b^2 = -1$, $u \cdot b = 0$, $u + b = N\, n_+$ et $u - b = N\, n_-$. Ainsi, b est un vecteur spatial unitaire, tangent à Σ_τ et pointant vers l'observateur O. En géneral, le vecteur b n'est pas géodésique mais *chorodésique*, à cause de la propriété de synchronicité et la congruence des observateurs associés est *quasi-rigide* (voir [1]).

2.4. Observateurs associés

Le champ vectoriel u, défini partout, caractérise la famille des observateurs canoniques. Celle-ci définit une « cinématique » dans le sens de Smarr et York [10], et tout le formalisme des projecteurs, fonctions lapse et shift (ce dernier est ici nul), courbure intrinsèque, ..., s'applique parfaitement.

2.4.1. Décalages

Un objet, au point (événement) x dans l'espace-temps, et de vitesse $V(x)$, est vu avec un décalage $z^{+} = (n^{+}(x) \cdot V(x))^{-1}$ par l'observateur (dans le futur). Il voit l'observateur (dans son passé) avec un décalage $z^{-} = n^{-}(x) \cdot V(x)$. Pour la congruence d'observateurs canoniques, $z^{+} = N(x)$ et $z^{-} = N(x)^{-1}$. Ces observateurs sont comobiles par rapport à la coordonnée R : ils conservent une valeur constante de R.

2.4.2. Référentiels

En supposant l'observateur sans rotation, le problème possède la symétrie spatiale axiale autour de la direction de son accélération (et même sphérique si l'observateur est inertiel).

En chaque point $x \in \mathcal{M}$ (sauf sur O), les deux vecteurs $n_{+}(x)$ et $n_{-}(x)$, ou $u(x)$ et $b(x)$, définissent le même plan D_x dans l'espace tangent T_x. Ces plans engendrent une foliation de surfaces intégrales $\mathcal{D}$ (figure 15.2). Chacune est engendrée, avec le temps, par une direction spatiale constante (α, β) du point de vue de l'observateur. Dans chaque $\mathcal{D}$ (elle-même paramétrée par (α, β)), nous utiliserons R comme coordonnée spatiale (radiale) et T comme coordonnée temporelle : $x = (T, R, \alpha, \beta)$. Par définition, un rayon lumineux issu de [vers] un point (T, R, α, β) atteint [resp. part de] l'observateur au temps propre $T + R$ [resp. $T - R$]. Le coréférentiel (dT, dR) a pour dual $(e_T := \partial_T, e_R := \partial_R)$. Ces deux vecteurs engendrent la surface $\mathcal{D}$. De

$$\langle dT, e_T \rangle = \langle dR, e_R \rangle = 1, \ \langle dR, e_T \rangle = \langle dT, e_R \rangle = 0, \tag{15.8}$$

on montre $e_T = N^2 \ (n_{+} + n_{-})/2 = N \ u$ et $e_R = N^2 \ (n_{+} - n_{-})/2 = -N \ b$. Les deux vecteurs commutent et sont mutuellement Lie-transportés : $e_T \cdot \nabla \ e_R = e_R \cdot \nabla \ e_T$. Il en est de même des deux vecteurs $N^2 \ n^{+}$ et $N^2 \ n^{-}$.

Le plan spatial C_x orthogonal à ces vecteurs est tangent aux cônes de lumière passé et futur à ce point, ainsi qu'à l'hypersurface Σ en ce point. C'est le plan tangent à la surface

$$C = I^{+} \cap I^{-} = I^{+} \cap \Sigma. \tag{15.9}$$

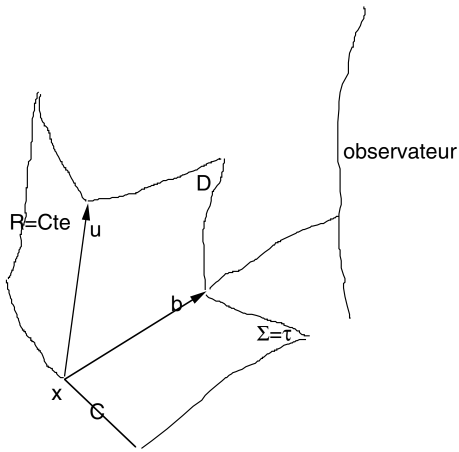


FIG. 15.2 – En chaque point x, u et b sont tangents à la surface $\mathcal{D}$. La courbe C (tangente à b est contenue dans l'espace Σ_τ et dans le « cylindre » $R = C^{te}$.

Cette 2-sphère déformée $\mathcal{C}(\tau, \delta)$ est l'ensemble de tous les points à la même valeur de δ au moment τ, *i.e.*, l'ensemble de tous les points de l'hypersurface Σ_τ à même PT-intervalle. u et b sont tous deux orthogonaux à $\mathcal{C}$.

Nous pouvons choisir deux vecteurs spatiaux unitaires orthogonaux e_3 et e_4 dans C (plusieurs choix sont possibles ; en l'absence de rotation, le problème est à 2 dimensions et ce choix importe peu). e_3 et e_4 forment une base de C, tangent à la 2-surface $\mathcal{C}$. chacune des paires précédentes forme avec ces deux vecteurs une base (pseudo-ON ou ON) de l'espace tangent à l'espace-temps. En outre, n^+ or n^-, e_3 et e_4 forment une base pseudo-ON pour l'espace tangent aux cônes de lumière.

2.4.3. La métrique

Puisque u et b sont orthogonaux et unitaires, la métrique s'écrit

$$ds^2 = \tilde{u}^2 - \tilde{b}^2 - (e^3)^2 - (e^4)^2 = N^2\ (dT^2 - dR^2) - (e^3)^2 - (e^4)^2, \qquad (15.10)$$

où $-(e^3)^2 = g_{\alpha\alpha}\ d\alpha^2$ et $-(e^4)^2 = g_{\beta\beta}\ d\beta^2$. Ainsi, N apparaît comme la *fonction lapse* associée à la foliation, ou à la congruence des observateurs associés (le vecteur shift est zéro). Puisque $N = 1$ en O, la métrique est localement minkowskienne pour lui, sur sa ligne d'univers. Le formalisme ADM usuel permet de définir les projecteurs temporel et spatial, ainsi que les formes fondamentales (métrique et courbure extrinsèque) des surfaces Σ_τ (voir, *e.g.*, Smarr et York [10]).

Les fonctions T, R (accompagnées de coordonnées angulaires appropriées) constituent un système de coordonnées, et un repère global dans $\mathcal{M}_0$. Dans ces coordonnées, les coefficients de la métrique se réduisent à $\pm N^2$, avec $N = 1$ pour l'observateur inertiel. Pour lui, T et R sont des *coordonnées normales* basées sur un point quelconque de sa ligne d'univers. En général, T et R ne sont *pas* des coordonnées *gaussiennes normales* puisque T n'est pas un temps propre, sauf sur O (il est en général impossible de construire des coordonnées gaussiennes normales pour cette foliation). Elles ont cependant l'avantage que les surfaces à « temps » T constant sont des surfaces globales de simultanéité, alors que les surfaces de coordonnées gaussiennes normales constantes ne le sont pas (elles ne vérifient que la simultanéité *locale*).

3. L'observateur inertiel dans l'espace-temps de Minkowski

L'observateur inertiel O (zéro accélération) dans l'espace-temps de Minkowski (pas d'expansion) a pour vitesse

$$u^0 = c,\ u^1 = s,\ u^2 = u^3 = 0,$$

où $c := \cosh \psi$ et $s := \sinh \psi$, la *rapidité* ψ étant constante. Sa ligne d'univers est donnée par $x^0 = c\ \tau,\ x^1 = s\ \tau,\ x^2 = x^3 = 0$. Un rayon lumineux passant par x atteint O au temps τ tel que

$$\left(x^0 - c\ \tau\right)^2 = \left(x^1 - s\ \tau\right)^2 + \left(x^2\right)^2 + \left(x^3\right)^2 . \tag{15.11}$$

Les calculs [5] donnent pour la surface Σ_τ le plan d'équation $c\,x^0 - s\,x^1 = \tau$, incliné de ψ par rapport à la verticale, orthogonal à la ligne d'univers de l'observateur inertiel : l'espace diffère d'un observateur inertiel à l'autre. Finalement, $u = c\ dx^0 - s\ dx^1$ et

$$\tilde{b} = \frac{1}{R}\ \left[\left(s\,x^0 - c\,x^1\right)\ \left(s\ dx^0 - c\ dx^1\right) + x^2\ dx^2 + x^3\ dx^3\right]$$

pointe vers l'observateur au temps $\tau = T(x)$. Dans l'espace Σ_τ (figure 15.3), on définit naturellement la coordonnée $y := -s\,x^0 + c\,x^1 = y/c$ de manière à ce que $R^2 = y^2 + (x^2)^2 + (x^3)^2$ et la métrique spatiale $-ds^2 = dy^2 + (dx^2)^2 + (dx^3)^2$.

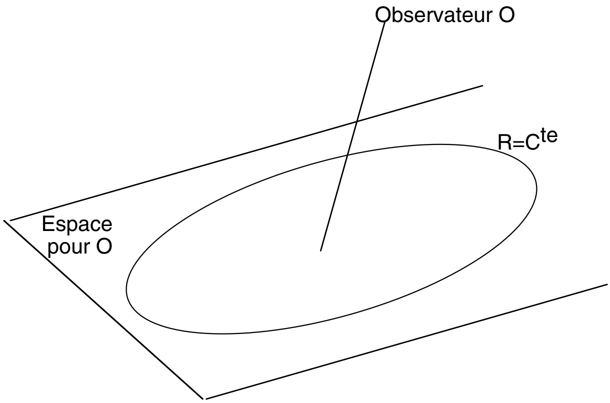


FIG. 15.3 – Pour l'observateur inertiel dans l'espace-temps de Minkowski, de vitesse arbitraire (rapidité ψ), l'espace est l'hyperplan d'inclinaison ψ. Nous avons dessiné une courbe $R = C^{te}$ dans ce plan.

3.1. L'observateur de Langevin dans l'espace-temps de Minkowski

La solution du célèbre « paradoxe des jumeaux de Langevin » se trouve dans la géométrie de l'espace-temps. J'appelle *observateur de Langevin* un observateur initialement inertiel, qui subit une accélération instantanée, et qui reste ensuite de nouveau inertiel (figure 15.4). Un tel observateur peut rencontrer son jumeau, lui demeuré rigoureusement inertiel, après une durée différente de temps propre. Il est souvent écrit (voir par exemple Misner *et al.* [8]) qu'il est impossible de définir globalement l'espace pour un tel observateur. Je montre que la prescription de synchronicité s'applique parfaitement et fournit une définition non ambigüe de l'espace pour cet observateur, de trajectoire spatio-temporelle

$$x^0 = \tau , x^1 = x^2 = x^3 = 0, \qquad \text{pour} \qquad t < 0,$$

$$x^0 = c\ \tau , x^1 = s\ \tau , x^2 = x^3 = 0, \qquad \text{pour} \qquad t > 0,$$

avec $c := \cosh \psi$ et $s := \sinh \psi$.

Le cône de lumière de l'observateur à $t = 0$, $\mathcal{L}^0$, divise l'espace-temps en trois parties I, II et III (voir figure 15.4), correspondant aux régions passée, future et spatialement reliée au point d'accélération. L'étude des rayons lumineux du [vers le] point x vers [de] l'observateur fournit les fonctions $\mathcal{N}^{\varepsilon}(x)$, $T(x)$ et $R(x)$ dans les trois régions.

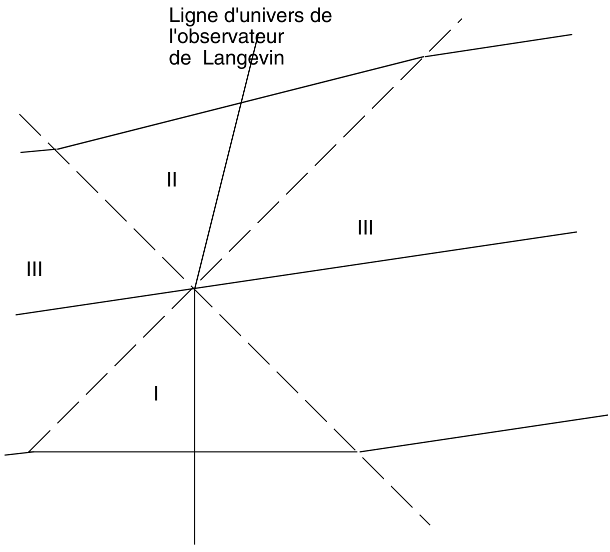


FIG. 15.4 – Ligne d'univers et (morceaux de l')espace à différents instants pour l'observateur de Langevin. Les tirets indiquent son cône de lumière.

La surface Σ_τ d'équation $T(x) = \tau$ définit l'espace pour l'observateur au temps propre τ. Nous nous intéressons à sa projection sur le plan (x^0, x^1) (figure 15.4) :

- **Région I** : $x^0 = \tau$. C'est une droite horizontale, et $R(x) = x^1$.
- **Région II** :

$$c\, x^0 - s\, x^1 = \tau, \quad \text{ou} \quad \tau / \cosh \psi = x^0 - \tanh \psi\, x^1. \tag{15.12}$$

  C'est une droite inclinée de ψ par rapport à la verticale, orthogonale à la ligne d'univers de l'observateur dans cette région. Le long de cette droite,

$$R(x) = \sqrt{(c\, x^0 - s\, x^1)^2 - (x^0)^2 + (x^1)^2} = c\, x^1 - s\, x^0.$$

- **Region III** :

$$\begin{aligned} 2\tau &= x^0 + c\, x^0 - s\, x^1 + \sqrt{(c\, x^0 - s\, x^1)^2 - (x^0)^2 + (x^1)^2} - x^1 \\ &= x^0 + c\, x^0 - s\, x^1 + c\, x^1 - s\, x^0 - x^1, \end{aligned} \tag{15.13}$$

  ou

$$\frac{\tau \ \exp(\psi/2)}{\cosh(\psi/2)} = -x^0 + x^1 \tanh(\psi/2) : \tag{15.14}$$

c'est une droite inclinée de $\psi/2$ par rapport à la verticale, bissectrice (en angle hyperbolique) des deux droites précédentes (figure 15.4). Le long de cette droite,

$$\begin{aligned}2R(x) &= -x^0 + c\,x^0 - s\,x^1 + \sqrt{(c\,x^0 - s\,x^1)^2 - (x^0)^2 + (x^1)^2} + x^1\\ &= -x^0 + c\,x^0 - s\,x^1 + c\,x^1 - s\,x^0 + x^1.\end{aligned} \tag{15.15}$$

Pour l'observateur, l'espace se constitue d'un disque plan S^I [ou S^{II}] qui s'étend jusqu'au cône de lumière $\mathcal{L}^0$. Au delà, il se continue par une surface composite S^{III}. En dehors du seul instant d'accélération, l'espace n'est ni plat ni homogène.

Cet exemple est le plus simple pour lequel notre prescription diffère des autres. On sait que, dans ce cas, il est impossible d'étendre les coordonnées de Fermi en dehors des régions coniques. Par ailleurs, aucune surface homogène ne peut faire l'affaire. Notre prescription est donc ici *la seule* à fournir un système de référence associé à l'observateur et valide dans la totalité de l'espace-temps, à permettre l'extension de son temps propre, et à permettre des procédures de synchronisation (conclusion similaire par Dolby et Gull [3]).

4. L'observateur de Rindler

L'observateur de Rindler dans l'espace-temps de Minkowski est défini par son accélération constante. Il a pour vitesse

$$u^0 = \cosh(a\,\tau),\ u^1 = \sinh(a\,\tau),\ u^2 = u^3 = 0, \tag{15.16}$$

pour accélération $a^0 = a\sinh(a\,\tau)$, $a^1 = a\cosh(a\,\tau)$, $a^2 = a^3 = 0$ (sans perdre de généralité, l'accélération est parallèle à la direction x^2). La ligne d'univers est une hyperbole dans l'espace-temps, d'équation $x^0 = a^{-1}\ \sinh(a\tau)$, $x^1 = a^{-1}\ \cosh(a\tau)$, $x^2 = x^3 = 0$. Un rayon de (vers) x atteint l'observateur au temps propre τ tel que

$$\left[x^0 - a^{-1}\ \sinh(a\tau)\right]^2 = \left[x^1 - a^{-1}\ \cosh(a\tau)\right]^2 + \left(x^2\right)^2 + \left(x^3\right)^2. \tag{15.17}$$

Nous verrons que $(x^1)^2 - (x^0)^2 > 0$, ce qui permet d'introduire les coordonnées de Rindler

$$x^0 := a^{-1}\exp(a\,\xi)\sinh(a\,\eta)\ \text{ et }\ x^1 := a^{-1}\exp(a\,\xi)\cosh(a\,\eta). \tag{15.18}$$

4.1. Le problème à 2 dimensions

Le problème est le plus souvent traité à deux dimensions seulement : il est particulièrement simple et pédagogique (cas à 4 dimensions ci-dessous). On trouve (détails dans [5])

$$\mathcal{N}_\varepsilon = \eta + \varepsilon|\xi|,$$

$$T = \eta, R = |\xi|$$

et $N^2 = a^{-2}\ \exp(2a\,\xi) = [x^2 - t^2]$. L'hypersurface (la ligne) Σ_τ a pour équation $\eta = \tau$, soit $x^0 = \tanh\tau\ x^1$: c'est une droite passant par l'origine. Les surfaces de PT-intervalle δ constant sont les hyperboles d'équation $\xi = \delta$, ou $(x^1)^2 - (x^0)^2 = a^{-2}\ \exp(2a\ \delta)$.

4.2. Le problème à 4 dimensions

Le calculs (Lachièze-Rey, 2001) donnent $\mathcal{N}_\varepsilon = \eta + \varepsilon\ \xi'$, $T = \eta$, $R = \xi'$. Les formes $dT = d\eta = \frac{x^1\ dx^0 - x^0\ dx^1}{(x^1)^2 - (x^0)^2}$ et $dR = d\xi' = d\xi = \frac{-x^0\ dx^0 + x^1\ dx^1}{(x^1)^2 - (x^0)^2}$, si bien que $n^\varepsilon = \frac{\varepsilon\ dx^0 + dx^1}{x^1 + \varepsilon\ x^0}$ et $N^2 = a^{-2}\ \exp(2a\ \xi) = [(x^1)^2 - (x^0)^2]$.

L'hypersurface Σ_τ est l'hyperplan plat passant par l'origine, d'équation $x^0 = \tanh\tau\, x^1$ (il se projette selon la droite vue à la section précédente). La métrique spatiale sur Σ_τ est donnée par

$$d\sigma^2 = N^2\, dR^2 + \left(dx^2\right)^2 + \left(dx^3\right)^2 = a^{-2}\,[d\exp(a\ \xi)]^2 + \left(dx^2\right)^2 + \left(dx^3\right)^2$$

$$= \left(dx^1/\cosh\eta\right)^2 + \left(dx^2\right)^2 + \left(dx^3\right)^2 \ : \qquad (15.19)$$

l'hypersurface est plate et homogène.

Les surfaces de PT-intervalle δ constant sont données par $\xi' = \delta$, ou

$$2\,a^{-1}\cosh(a\ \delta)\ \sqrt{(x^1)^2 - (x^0)^2} - a^{-2} = \left(x^1\right)^2 - \left(x^0\right)^2 + \left(x^2\right)^2 + \left(x^3\right)^2. \quad (15.20)$$

Ces calculs montrent que les coordonnées de Rindler, largement utilisées, correspondent en fait à la définition de l'espace et du temps ici introduites. Ceci justifie *a posteriori* leur utilisation, et apporte un éclairage nouveau sur l'effet Unruh : il apparaît comme conséquence directe des différentes séparations temps – espace pour les deux observateurs (inertiel et Rindler) : ils associent différentes fréquences temporelles au même état, le vide inertiel de Minkowski. Dans cet esprit, Pauri et Vallisneri [9] ont pu invoquer une origine classique (non quantique) à cet effet.

5. L'observateur cosmologique

Les modèles cosmologiques de Friedmann-Lemaître possèdent des sections spatiales à symétrie maximale. Avec des coordonnées convenables, la métrique s'écrit

$$ds^2 = A(\eta)^2 \left(d\eta - \left[d\sigma^2 - S(\sigma)^2 \left(d\alpha^2 + \sin^2\alpha\ d\beta^2\right)\right]\right), \qquad (15.21)$$

où A est le facteur d'échelle. L'expression entre crochets est la métrique des sections spatiales à symétrie maximale ($\mathbb{R}^3$, S^3 or H^3) et η est le *temps conforme*.

Je considère l'observateur cosmologique *inertiel* (CIO) O_I décrivant la ligne $\sigma = 0$. Le problème est à symétrie sphérique. Les rayons consiérés sont radiaux, α, β ne joueront aucun rôle. Le temps propre τ est donné par $d\tau = A\ d\eta$. Les fonctions $\eta(\tau)$, et sa réciproque f telle que $f[\eta(\tau)] := \tau$, joueront un rôle important. Puisque $\eta > 0$, le CIO a un horizon des particules d'intérieur $\sigma < \eta$.

Pour le CIO,

$$\mathcal{N}_\varepsilon(\eta,\sigma) = f[\eta + \varepsilon\ \sigma],\quad 2T(\eta,\sigma) = f[\eta+\sigma] + f[\eta-\sigma], 2R(\eta,\sigma) = f[\eta+\sigma] - f[\eta-\sigma]. \qquad (15.22)$$

Par dérivation, $n_\varepsilon = A^\varepsilon\ (d\eta + \varepsilon\ d\sigma)$, en définissant $A^\varepsilon(\eta,\sigma) := A(\eta + \varepsilon\ \sigma)$. Par somme et différence,

$$dT = (A^+ + A^-)/2\ d\eta + (A^+ - A^-)/2\ d\sigma \qquad (15.23)$$

$$dR = (A^+ - A^-)/2\ d\eta + (A^+ + A^-)/2\ d\sigma, \qquad (15.24)$$

$$\text{et}\quad N^2(\eta,\sigma) = \frac{A(\eta)^2}{A^+\ A^-}. \qquad (15.25)$$

5.1. L'espace pour l'observateur cosmologique inertiel

L'espace Σ_τ a pour équation

$$f[\eta+\sigma] + f[\eta-\sigma] = 2\tau,\ \text{with}\ \sigma < \eta, \qquad (15.26)$$

qui diffère de $\eta = C^{te}$ (sauf dans le cas sans expansion, $\eta = \tau$) : l'espace pour le CIO *n'est pas* une section spatiale à symétrie maximale. L'expansion cosmique brise l'homogénéité spatiale (mais pas l'isotropie si l'observateur est inertiel). Une section spatiale $\eta = C^{te}$, quelquefois appelée « espace », ne vérifie pas la condition de synchronicité (elle vérifie une condition en temps conforme, sans pertinence physique pour l'observateur).

En même temps, l'expansion cosmique imprime une courbure sur l'espace : même si l'espace-temps admet des sections spatiales de courbure constante (comme par exemple dans le cas Einstein-de Sitter, où elles sont plates), ce n'est pas le cas de l'espace ici défini.

L'espace est limité par l'horizon $\sigma = \eta$, ou $T = R$. Vers l'horizon, $A^{-} \to 0$ et $N \to \infty$: l'espace tend à devenir nul (de genre lumière).

5.1.1. Le modèle Einstein-de Sitter

Le modèle *Einstein-de Sitter* correspond à $A(\tau) = \tau^{2/3} = (\eta/3)^2$, avec des sections spatiales plates. On a $\eta(\tau) = 3\,\tau^{1/3}$ et $f(y) = (y/3)^3$. D'où, pour Σ_τ, l'équation (figure 15.5)

$$27\,\tau = \eta\,(\eta^2 + 3\,\sigma^2), \tag{15.27}$$

et non pas $\tau = \eta^3/27$, l'équation d'une section spatiale plate. Nous avons $N(\eta,\sigma) = \frac{\eta^2}{\eta^2-\sigma^2} = \left[1 + \sqrt[3]{\frac{\tau+R}{\tau-R}} + \sqrt[3]{\frac{\tau-R}{\tau+R}}\right]/4$. L'espace est bien orthogonal à la ligne d'univers de l'observateur.

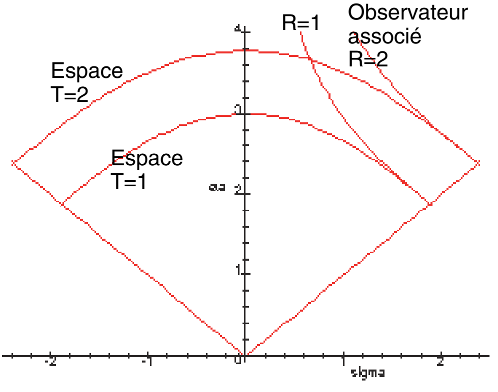


FIG. 15.5 – Espace ($T = Ct$) et lignes d'univers des observateurs associés ($R = Ct$) CIO dans l'espace-temps d'Einstein-de Sitter.

5.1.2. Observateurs associés

Tout observateur a une classe d'observateurs associés. Ce ne sont pas les observateurs comobiles définis par $\sigma = C^{te}$. Ceux-ci obéissent à $dT = \frac{A^{+}+A^{-}}{A^{+}-A^{-}}\, dR$. Mais les observateurs associés conservent par définition un PT-intervalle constant, et obéissent à

$$dR = (A^{+} - A^{-})/2\ d\eta + (A^{+} + A^{-})/2\ d|\sigma| = 0. \tag{15.28}$$

À l'horizon, un observateur associé est vu par le CIO avec un décalage $z^{+} \to \infty$.

5.1.3. Temps et distances

Toute mesure effectuée par un observateur, locale ou non, se réfère à son temps propre. Pour un CIO, le temps physique, permettant de mesurer des durées, ou de dater les événements, n'est ni t ni η, mais T (T et t coïncident sur la ligne d'univers du CIO).

Par ailleurs, la *distance propre* sert à mesurer un intervalle entre deux objets considérés simultanément, c'est-à-dire à une valeur commune du temps. Mais aucun observateur n'a accès au temps conforme η. La simultanéité (non pas absolue, mais relative à l'observateur) doit donc être définie non pas à partir de η mais de T. Ceci conduit à utiliser une *distance-temps-propre* (PT-distance) entre deux objets, calculées par intégration de la métrique le long de Σ_T, et non d'une section spatiale $t = C^{te}$ (or $\eta = C^{te}$) :

$$d_{PT}(g) = \int_{\Sigma_T} ds = \int_{\Sigma_T} N\ dR. \tag{15.29}$$

C'est réellement la distance entre deux objets dans *l'espace*, à un instant bien défini et unique pour le CIO. Pour lui, T et R apparaissent ainsi comme de bonnes coordonnées pour mesurer espace et temps.

5.2. L'observateur inertiel dans l'espace-temps de de Sitter

Ce cas présente un intérêt particulier à cause de sa symétrie maximale. Il est souvent considéré comme une base pour la quantification. Il est bien connu qu'il admet différents découpages temps-espace, les sections spatiales pouvant avoir une courbure positive, zéro ou negative. À chaque découpage correspond un système de coordonnées. Je choisis ici celui où les sections spatiales ont une courbure positive constante, qui couvre la totalité de l'espace-temps (l'hyperboloïde de Sitter).

La métrique

$$\begin{aligned} ds^2 &= dt^2 - \rho^2 \left(\cosh \rho^{-1} t\right)^2 \left[d\sigma^2 + \sin^2\sigma \; d\Omega^2\right] \\ &= A^2(\eta) \left[d\eta^2 - d\sigma^2 - \sin^2\sigma \; d\Omega^2\right], \end{aligned} \tag{15.30}$$

où le temps conforme $\eta = 2\tan^{-1}[e^{t/\rho}]$ et $A(\eta) = (\rho/2)[\tan(\eta/2) + 1/\tan(\eta/2)]$ est la constante qui caractérise la courbure de l'espace-temps).

Le CIO a pour temps propre $t = \rho \ln[\tan(\eta/2)]$, si bien que $f(y) := \rho \ln[\tan(y/2)]$. Donc, $\mathcal{N}_\varepsilon(t, \sigma) = \rho \ln\tan(\frac{\eta+\varepsilon\ \sigma}{2})$. D'où

$$2T(\eta, \sigma) = \rho \left[\ln\tan\left(\frac{\eta+\sigma}{2}\right) + \ln\tan\left(\frac{\eta-\sigma}{2}\right)\right], \tag{15.31}$$

$$2R(\eta, \sigma) = \rho \left[\ln\tan\left(\frac{\eta+\sigma}{2}\right) - \ln\tan\left(\frac{\eta-\sigma}{2}\right)\right]. \tag{15.32}$$

L'espace au temps propre τ, pour le CIO, est donné par

$$\tan\frac{\eta+\sigma}{2}\ \tan\frac{\eta-\sigma}{2} = e^{2\tau/\rho}, \tag{15.33}$$

ou

$$\frac{\cosh(t-S)}{\cosh(t+S)} = e^{2\tau/\rho}, \tag{15.34}$$

où l'on a défini $e^{S/\rho} := \tan(\sigma/2)$. Ce n'est pas la surface $t = C^{te}$ de courbure constante. Il semble donc pertinent d'étudier la quantification sur cette surface (avec le temps orthogonal).

6. Discussion

La prescription de synchronicité définit espace et temps globalement, sans ambigüité, pour tout observateur, inertiel ou non, dans un espace-temps arbitraire (sans croisement de géodésiques nulles), en particulier celui de Minkowski ou les modèles Friedmann-Lemaître. La définition de l'espace est relative à l'observateur (il diffère pour un autre), à chaque instant de sa ligne d'univers. Cela constitue une foliation de l'espace-temps, relative à cet observateur, que l'on peut interprèter comme une classe d'observateurs *canoniquement associés*, ou une « cinématique » de l'espace-temps [10]. Cela fournit un système naturel de référence, qui reste minkowskien le long de la ligne d'univers de l'observateur (le temps y coïncide avec son temps propre) et pertinent pour des mesures physiques. Pour l'observateur de Rindler, ce système coincide avec celui généralement utilisé sans autre justification que d'être « naturel », et le justifie *a posteriori*.

L'application à l'espace-temps de Minkowski confirme que l'espace et le temps diffèrent pour les observateurs inertiels avec des vitesses différentes. Elle fournit une définition globale de l'espace et du temps pour l'observateur de Langevin, à qui aucune autre prescription ne s'applique. Dans le cas de l'observateur de Rindler (accélération constante), les coordonnées synchrones ici définies sont les coordonnées de Rindler usuelles. L'effet Unruh s'interprète alors comme dû à la dépendance de l'espace et temps de l'observateur.

En cosmologie, la prescription définit l'espace pour un observateur inertiel. Ce n'est pas une section spatiale à symétrie maximale : dans un modèle Friedmann-Lemaître, aucun observateur inertiel ne « voit » un espace homogène, cette absence d'homogénéité étant engendrée par la courbure spatio-temporelle correspondant à l'expansion cosmique. En particulier, l'espace n'est ni plat ni homogène (cependant le caractère inertiel de l'observateur préserve l'isotropie) dans le modèle Einstein-de Sitter, pourtant souvent qualifié « d'univers plat » ! Dans l'espace-temps de de Sitter, l'espace n'est pas, non plus, une hypersurface à symétrie maximale.

Ces résultats ne modifient aucune formule cosmologique, pourvu qu'elle soit écrite sous forme covariante, et qu'elle n'implique aucune définition de l'espace. Ils modifient cependant l'*interprétation* de résultats observationnels se réfèrant à l'espace (qu'il soit homogène, plat, etc.). Ils modifient également l'interprétation de la *distance propre* : celle-ci n'apparaît plus comme la mesure d'un intervalle purement spatial, entre deux événements se déroulant au même moment ; mais plutôt comme celle d'un intervalle mixte entre deux événements non synchrones du point de vue de l'observateur (ils seraient synchrones si la montre de l'observateur indiquait le temps conforme). On peut introduire une « distance-temps-propre », un intervalle véritablement spatial (dans le sens introduit ici) entre deux événements synchrones pour l'observateur. Elle correspond au résultat d'une mesure que l'observateur peut effectivement mener avec sa montre (qui indique par définition son temps propre).

Cette prescription concerne l'interprétation d'effets quantiques en espace-temps courbe, et/ou du point de vue d'observateurs non inertiels. Pour l'observateur de Rindler, cela confirme les résultats usuels à propos de l'effet Unruh, et en fournit une compréhension plus claire. Mais, dans la plupart des cas, cette prescription diffère des tentatives de quantification déjà effectuées, qui utilisent les sections spatiales à symétrie maximale, plutôt que celles définies ici.

Bibliographie

[1] Bel, L., 1998, gr-qc/9812062.
[2] Birrel, N.D., Davis, P.C.W., *Quantum fields in curved espace*, Cambridge University Press, 1982.
[3] Dolby, C.E., Gul, S.F., 2001, gr-qc/0104077.
[4] Ali, S.T., Antoine, J.-P., Gazeau, J.-P., *Ann. Inst. Henri Poincaré*, vol. 52, n° 1, pp. 83–111, 1990.
[5] Lachièze-Rey, M., *A&A* 376 (2001) 17–27 (arXiv:gr-qc/0107010).
[6] Landau, L., Lifshitz, E., *Field Theory*, MIR (URSS), 1966.
[7] Marzlin, K.-P., 1994, gr-qc/9402010 v2.
[8] Misner, C.W., Thorne, K.S., Wheeler, J.A., *Gravitation*, Freeman et co., 1973.
[9] Pauri, M., Vallisneri, M., 1999, gr-qc/9903052 v2.
[10] Smarr, L., York, J.W. Jr, *Phys. Rev. D* 17, 10 (1978) 2329.
[11] Sriramkumar, L., Padmanabhan, T., 1999, gr-qc/9903054 v2.

16 La machine électromagnétique à remonter le temps

Mario Novello

1. Introduction

La force électromagnétique subie par un photon dans un régime non linéaire peut être géometrisée. C'est un résultat certainement inattendu, et en même temps très intéressant, de l'analyse récente du comportement des discontinuités des champs électromagnétiques non-linéaires. Nous allons montrer comment une telle géometrisation est possible. Cette propriété n'est pas restreinte aux champs de spin 1, mais elle est très générale, valable pour des théories non linéaires de spin quelconque. Cependant nous nous bornerons ici aux cas de spin 1.

2. Commentaire général sur l'électrodynamique non-linéaire

Les modifications de la propagation de la lumière ont récemment beaucoup attiré l'attention des scientifiques, surtout dans des situations relatives à des vides non triviaux (produits par des effets divers comme la température, des conditions particulières de contour, une polarisation du vide, etc.). Des telles investigations ont démontré que la propagation de la lumière peut être traitée comme celle d'ondes électromagnétiques dans un milieu dispersif classique. Le milieu induit des modifications sur les équations de mouvement, décrites par des termes non linéaires du champ. Nous allons nous concentrer ici dans le cas où les termes non-linéaires peuvent être reproduits par un Lagrangien dépendant de deux invariants F et G.

Le résultat le plus important et inattendu de ce travail consiste en la reconnaissance que, dans un régime non linéaire, les photons se propagent suivant des géodésiques nulles, non dans une géometrie de Minkowski, mais comme si la structure métrique de l'espace-temps avait une autre forme.

3. Définitions

Le champ électromagnétique est noté par le tenseur $F_{\mu\nu}$, et par son dual

$$F^*_{\alpha\beta} \doteq \frac{1}{2}\eta_{\alpha\beta}{}^{\mu\nu}F_{\mu\nu}, \tag{16.1}$$

où $\eta_{\alpha\beta\mu\nu}$ représente le tenseur antisymétrique de Levi-Civita ; le tenseur métrique de Minkowski est représenté par sa forme standard $\eta^{\mu\nu}$. Les deux invariants que l'on peut construire avec ces deux tenseurs sont :

$$F \doteq F^{\mu\nu}\, F_{\mu\nu}, \tag{16.2}$$

$$G \doteq F^{\mu\nu}\, F^*_{\mu\nu}. \tag{16.3}$$

Étant donné que les modifications du vide ne changent pas l'invariance de jauge de la théorie, la forme la plus générale du Lagrangien peut s'écrire en fonction de ces deux invariants, c'est-à-dire

$$L = L(F,\, G).$$

On va noter L_F et L_G les dérivées du Lagrangien L par rapport aux invariants F et G, respectivement ; et de même pour les dérivées d'ordre supérieur. Nous sommes particulièrement intéressés par la dérivation des surfaces caractéristiques qui guident la propagation des discontinuités du champ.

Soit Σ une surface de discontinuité pour le champ électromagnétique. Selon Hadamard, nous considérons le champ comme continu à travers Σ, et sa dérivée première ayant une discontinuité finie. Nous écrivons

$$[F_{\mu\nu}]_\Sigma = 0, \tag{16.4}$$

et

$$[\partial_\lambda\, F_{\mu\nu}]_\Sigma = f_{\mu\nu}k_\lambda, \tag{16.5}$$

dans lequel le symbole

$$[J]_\Sigma \equiv \lim_{\delta\to 0^+} (J|_{\Sigma+\delta} - J|_{\Sigma-\delta})$$

représente la discontinuité d'une fonction arbitraire J à travers la surface Σ caractérisée par l'équation $\Sigma(x^\mu) = constant$. On appelle le tenseur $f_{\mu\nu}$ pour la discontinuité du champ, et

$$k_\lambda = \partial_\lambda \Sigma \tag{16.6}$$

le vecteur propagation.

4. La méthode de la géometrie effective

4.1. Lagrangien à un paramètre

Dans cette section nous nous bornerons à analyser les conséquences des non-linéarités dans l'équation d'evolution des ondes électromagnétiques pour les cas des Lagrangiens du genre

$$L = L(F). \tag{16.7}$$

Le principe de moindre action donne les équations du champ :

$$\partial_\mu \left(L_F F^{\mu\nu}\right) = 0. \tag{16.8}$$

Les conditions (16.4) et (16.5) pour la discontinuité du champ (16.8) à travers Σ donnent

$$L_F f^{\mu\nu} k_\nu + 2L_{FF}\, \xi F^{\mu\nu} k_\nu = 0, \tag{16.9}$$

où ξ est défini par

$$\xi \doteq F^{\alpha\beta} f_{\alpha\beta}. \tag{16.10}$$

La conséquence d'une telle discontinuité dans l'identité cyclique donne

$$f_{\mu\nu} k_\lambda + f_{\nu\lambda} k_\mu + f_{\lambda\mu} k_\nu = 0. \tag{16.11}$$

Pour obtenir une relation scalaire, contractons cette équation avec $k_\alpha \eta^{\alpha\lambda} F^{\mu\nu}$. Nous obtenons ainsi

$$\xi k_\nu k_\mu \eta^{\mu\nu} + 2F^{\mu\nu} f_\nu{}^\lambda k_\lambda k_\mu = 0. \tag{16.12}$$

Dans le cas où ξ n'est pas nul[1], les équations (16.9) and (16.12) donnent la propagation des discontinuités du champ :

$$\left(L_F \eta^{\mu\nu} - 4L_{FF} F^{\mu\alpha} F_\alpha{}^\nu\right) k_\mu k_\nu = 0. \tag{16.13}$$

1 – Dans le cas $\xi = 0$, la quantité $f_{\mu\nu}$ est une deux-forme singulière. Selon Lichnerowicz, on peut la décomposer en termes du vecteur de propagation k_μ et un vecteur du genre espace $a_\mu = a\,\epsilon_\mu$, orthogonal à k_μ, dans lequel ϵ_μ est le vecteur de polarization normalisé. On peut alors écrire $f_{\mu\nu} = k_\mu a_\nu - k_\nu a_\mu$ sur Σ. L'équation (16.9) donne $f^{\mu\nu} k_\nu = 0$, et en contractant (16.11) avec h $\eta^{\lambda\rho} k_\rho$, on obtient $f_{\mu\nu}\eta^{\alpha\beta} k_\alpha k_\beta = 0$. On obtient alors le résultat conventionnel : ces modes se propagent selon des géodésiques nulles dans l'espace de Minkowski.

L'expression (16.13) suggère que l'on peut interpréter l'auto-interaction du champ $F^{\mu\nu}$, par rapport à la propagation des discontinuités électromagnétiques (16.5), comme s'il y avait une modification de la géometrie de l'espace-temps $\eta_{\mu\nu}$, selon une géometrie effective :

$$g^{\mu\nu}_{\mathrm{eff}} = L_F \, \eta^{\mu\nu} - 4 \, L_{FF} \, F^{\mu}{}_{\alpha} \, F^{\alpha\nu}. \tag{16.14}$$

Un simple examen de cette équation démontre que, dans le cas linéaire de l'électrodynamique de Maxwell, les discontinuités du champ électromagnétique se propagent comme des courbes nulles dans la géometrie de Minkowski.

L'expression générale de la géométrie effective peut, d'une façon équivalente, être écrite en faisant appel aux *vacuum expectation value* (VEV) du tenseur d'énergie défini par

$$T_{\mu\nu} \equiv \frac{2}{\sqrt{-\gamma}} \frac{\delta \Gamma}{\delta \gamma^{\mu\nu}}, \tag{16.15}$$

où Γ représente l'action effective

$$\Gamma \doteq \int d^4x \sqrt{-\gamma} \, L, \tag{16.16}$$

et $\gamma_{\mu\nu}$ la métrique de Minkowski écrite dans un système arbitraire de coordonnées ; γ est le déterminant de cette métrique. Dans le cas où le Lagrangien s'écrit $L = L(F)$, on obtient

$$T_{\mu\nu} = -4 L_F \, F_{\mu}{}^{\alpha} \, F_{\alpha\nu} - L \, \eta_{\mu\nu}. \tag{16.17}$$

Nous avons choisi un système de coordonnées tel que $\gamma_{\mu\nu}$ se réduit à $\eta_{\mu\nu}$. En termes de ce tenseur, la géométrie effective (16.14) se réduit à[2]

$$g^{\mu\nu} = \left(L_F + \frac{L \, L_{FF}}{L_F} \right) \eta^{\mu\nu} + \frac{L_{FF}}{L_F} T^{\mu\nu}. \tag{16.18}$$

5. Géodésiques nulles

L'importance de la géometrie effective va bien au-delà de cette simple caractérisation. En effet, une structure dynamique intervient aussi, comme nous allons le démontrer : les courbes intégrales du vecteur k_ν (*i.e.*, les trajectoires

2 – Par simplification on va noter la métrique effective $g^{\mu\nu}$ et non pas $g^{\mu\nu}_{\mathrm{eff}}$.

des photons) sont des géodésiques de la métrique effective. Nous définissons la connexion de Levi-Civita $\Gamma^{\alpha}{}_{\mu\nu} = \Gamma^{\alpha}{}_{\nu\mu}$, et l'opérateur de dérivation covariante ∇_{λ}, tels que

$$\nabla_{\lambda} g^{\mu\nu} \equiv g^{\mu\nu}{}_{;\lambda} \equiv g^{\mu\nu}{}_{,\lambda} + \Gamma^{\mu}{}_{\sigma\lambda} g^{\sigma\nu} + \Gamma^{\nu}{}_{\sigma\lambda} g^{\sigma\mu} = 0. \tag{16.19}$$

Cela revient à dire que la structure de la géométrie de Riemann est complètement donnée par la métrique effective

En contractant (16.19) avec $k_{\mu} k_{\nu}$ on obtient

$$k_{\mu} k_{\nu} g^{\mu\nu}{}_{,\lambda} = -2 k_{\mu} k_{\nu} \Gamma^{\mu}{}_{\sigma\lambda} g^{\sigma\nu}. \tag{16.20}$$

Une différentiation donne

$$2 k_{\mu,\lambda} k_{\nu} g^{\mu\nu} + k_{\mu} k_{\nu} g^{\mu\nu}{}_{,\lambda} = 0. \tag{16.21}$$

ce qui entraîne

$$g^{\mu\nu} k_{\mu;\lambda} k_{\nu} \equiv g^{\mu\nu} \left(k_{\mu,\lambda} - \Gamma^{\sigma}{}_{\mu\lambda} k_{\sigma} \right) k_{\nu} = 0. \tag{16.22}$$

Comme le vecteur de propagation $k_{\mu} = \Sigma_{,\mu}$ est un gradient, on a $k_{\mu;\lambda} = k_{\lambda;\mu}$. Et, en définissant $k^{\mu} \doteq g^{\mu\nu} k_{\nu}$, on peut écrire

$$k_{\mu;\lambda} k^{\lambda} = 0, \tag{16.23}$$

ce qui montre que k_{μ} est un vecteur géodésique. Comme c'est un vecteur nul (par rapport à la géometrie effective $g^{\mu\nu}$), il s'en suit que les photons se propagent comme des géodésiques nulles dans la métrique effective.

6. Lagrangien à deux paramètres

Nous donnerons ici le résultat de la propagation des photons dans le cas où le Lagrangien depend des deux invariants

$$L = L(F, G) \tag{16.24}$$

et l'équation de mouvement a la forme

$$\partial_{\nu} \left(L_F F^{\mu\nu} + L_G F^{*\,\mu\nu} \right) = 0. \tag{16.25}$$

En utilisant la même procédure (Hadamard) qu'auparavant, on obtient que le photon se propage en vérifiant

$$g^{\mu\nu} k_{\mu} k_{\nu} = 0, \tag{16.26}$$

où la métrique effective est

$$
\begin{aligned}
g^{\mu\nu} = {} & L_F \eta^{\mu\nu} - 4 \left[(L_{FF} + \Omega L_{FG})\, F^{\mu}{}_{\lambda} F^{\lambda\nu} \right. \\
& \left. + (L_{FG} + \Omega L_{GG})\, F^{\mu}{}_{\lambda} F^{*\,\lambda\nu} \right].
\end{aligned} \tag{16.27}
$$

Nous avons utilisé les définitions

$$
\begin{aligned}
\Omega_1 = {} & -L_F L_{FG} + 2F L_{FG} L_{GG} \\
& + G\left(L_{GG}^2 - L_{FG}^2\right),
\end{aligned} \tag{16.28}
$$

$$
\begin{aligned}
\Omega_2 = {} & (L_F + 2G L_{FG})(L_{GG} - L_{FF}) \\
& + 2F\left(L_{FF} L_{GG} + L_{FG}^2\right),
\end{aligned} \tag{16.29}
$$

$$
\begin{aligned}
\Omega_3 = {} & L_F L_{FG} + 2F L_{FF} L_{FG} \\
& + G\left(L_{FG}^2 - L_{FF}^2\right),
\end{aligned} \tag{16.30}
$$

$$
\Omega = \frac{-\Omega_2 \pm \sqrt{\Delta}}{2\Omega_1}, \tag{16.31}
$$

avec

$$
\Delta \doteq (\Omega_2)^2 - 4\Omega_1\Omega_3.
$$

7. Propagation dans un milieu diélectrique

Nous allons considérer maintenant la situation où le régime non linéaire est dû à un milieu diélectrique. Le champ électromagnétique est représenté par deux tenseurs antisymétriques, le champ électromagnétique $F_{\mu\nu}$ et la polarization $P_{\mu\nu}$. On les décompose, d'une façon générale, en des parties électrique et magnétique par rapport à un observateur de vitesse v_μ. Nous écrirons :

$$
F_{\mu\nu} = E_\mu v_\nu - E_\nu v_\mu + \eta^{\rho\sigma}{}_{\mu\nu}\, v_\rho H_\sigma, \tag{16.32}
$$

$$
P_{\mu\nu} = D_\mu v_\nu - D_\nu v_\mu + \eta^{\rho\sigma}{}_{\mu\nu}\, v_\rho B_\sigma. \tag{16.33}
$$

En suivant Hadamard, nous prenons les discontinuités comme données par

$$
\begin{aligned}
\left[\nabla_\lambda E_\mu\right]_\Sigma &= k_\lambda e_\mu, \\
\left[\nabla_\lambda D_\mu\right]_\Sigma &= k_\lambda d_\mu, \\
\left[\nabla_\lambda H_\mu\right]_\Sigma &= k_\lambda h_\mu, \\
\left[\nabla_\lambda B_\mu\right]_\Sigma &= k_\lambda b_\mu.
\end{aligned} \tag{16.34}
$$

Dans le cas linéaire, nous avons

$$D_\alpha = \epsilon\, E_\alpha, \tag{16.35}$$

$$B_\alpha = \frac{H_\alpha}{\mu}. \tag{16.36}$$

Il en découle

$$d_\alpha = \epsilon\, e_\alpha, \tag{16.37}$$

$$b_\alpha = \frac{h_\alpha}{\mu}. \tag{16.38}$$

Le résultat de la procedure ci-dessus donne

$$k_\mu\, k_\nu\, [\gamma^{\,\mu\nu} + (\epsilon\, \mu - 1) v^\mu\, v^\nu] = 0. \tag{16.39}$$

Et pour le cas non linéaire

$$P_{\mu\nu} = \epsilon(E) F_{\mu\nu}, \tag{16.40}$$

où ϵ représente la susceptibilité électrique. On obtient alors, pour la métrique effective,

$$g^{\mu\nu} = \epsilon\, \eta^{\mu\nu} + \frac{\epsilon'}{E} F^\mu{}_\alpha F^{\alpha\nu}. \tag{16.41}$$

Ou, d'une façon équivalente

$$g^{\mu\nu} = \epsilon\, \eta^{\mu\nu} - \frac{\epsilon'}{E} \left(E^\mu\, E^\nu - E^2\, \delta^\mu_t\, \delta^\nu_t \right), \tag{16.42}$$

où $E^2 \equiv -E_\alpha\, E^\alpha > 0$. Cela revient à dire

$$g^{tt} = \epsilon + \epsilon' E \tag{16.43}$$

$$g^{ij} = -\epsilon\, \delta^{ij} - \frac{\epsilon'}{E} E^i\, E^j. \tag{16.44}$$

Ce qui montre que la discontinuité du champ électromagnétique dans un milieu diélectrique non linéaire se propage suivant des courbes nulles dans la métrique effective donnée ci-dessus.

8. Diélectrique en mouvement

Dans le cas où le diélectrique a un mouvement par rapport à un observateur de vitesse v^μ, un calcul semblable donne pour résultat une métrique effective de la forme

$$g^{\mu\nu} = \eta^{\mu\nu} + v^\mu\, v^\nu\, (\mu\, \epsilon - 1 + \mu\, \epsilon' E) - \frac{\epsilon' E}{\epsilon} \ell^\mu\, \ell^\nu, \tag{16.45}$$

où ℓ^μ est le vecteur unitaire dans la direction du champ électrique.

Le vide d'Euler-Heisenberg :

L'action effective pour l'électrodynamique due à des corrections quantiques a été calculée par Heisenberg. Pour des basses fréquences, $\nu \ll m_e c^2/h$, le Lagrangien effectif prend la forme

$$L = -\frac{1}{4}F + \frac{\mu}{4}\left(F^2 + \frac{7}{4}G^2\right), \tag{16.46}$$

où

$$\mu \doteq \frac{8}{45}\,\alpha^2 \left(\frac{\hbar}{m_e\, c}\right)^3 \frac{1}{m_e\, c^2}, \tag{16.47}$$

et α représente la constante de structure fine.

La trace du tenseur d'énergie corespondant est

$$T = \mu \left(F^2 + \frac{7}{4}\,G^2\right). \tag{16.48}$$

Les coefficients $\mathcal{M}$ et $\mathcal{N}$ de la métrique associée sont donnés par

$$\mathcal{M} = -\frac{1}{4} + \mu F, \tag{16.49}$$

$$\mathcal{N} = -2\mu. \tag{16.50}$$

9. Exemple-I

Nous allons montrer l'efficacité de cette méthode par deux exemples. Dans les deux cas, le photon subit des effets électromagnétiques qui lui font suivre une trajectoire fermée dans l'espace-temps. Nous allons ainsi montrer que la géometrie effective peut avoir des propriétés du même type que la géometrie découverte par Gödel dans le cas de la Relativité génerale, où existent des courbes fermées de genre temps.

La géométrie de fond est celle de Minkowski, décrite dans un système de coordonnées (t, r, φ, z) c'est-à-dire

$$ds^2 = dt^2 - dr^2 - r^2\, d\varphi^2 - dz^2. \tag{16.51}$$

Le système physique examiné se constitue d'un champ magnétique qui induit un champ électrique de telle façon que la composante F^{02} n'est pas nulle (il n'est pas nécessaire ici de préciser cette composante). Nous choisissons un observateur en co-mouvement avec le milieu diélectrique et avec une vitesse normalisée $v^{\mu} = \gamma \left(1, 0, \frac{\dot{\varphi}}{c}, 0\right)$, où $\gamma \equiv \left(1 - \frac{v^2}{c^2}\right)^{-\frac{1}{2}}$ et $v \equiv r\dot{\phi}$. Il en résulte que la composante électrique non-nulle est $E^{(0)} = -\frac{E\gamma r\dot{\varphi}}{c}$ et $E^{(2)} = -\frac{E\gamma}{r}$. Le photon se propage suivant des géodésiques nulles dans la géometrie effective. En utilisant nos résultats précédents, nous obtenons les composantes non-nulles de la métrique,

$$g^{00} = 1 + (\mu\,\epsilon - 1 + \mu\,\epsilon' E)\,\gamma^2 - \frac{\epsilon' E}{\epsilon}\gamma^2 v, \tag{16.52}$$

$$g^{11} = -1, \tag{16.53}$$

$$g^{22} = -1 + (\mu\,\epsilon - 1 + \mu\,\epsilon' E)\,(\gamma\, v)^2 - \frac{\epsilon' E}{\epsilon}\gamma^2, \tag{16.54}$$

$$g^{02} = (\mu\,\epsilon - 1 + \mu\,\epsilon' E)\,(\gamma\, v)^2 - \frac{\epsilon' E}{\epsilon}\gamma^2 v, \tag{16.55}$$

$$g^{33} = -1. \tag{16.56}$$

Considérons maintenant la courbe définie par les équations $t = constant$, $r = constant$, et $z = constant$. L'élément de longueur fondamentale de cette courbe a la forme

$$ds^2_{\mathrm{eff}} = g_{22}\, d\varphi^2. \tag{16.57}$$

Cette courbe est un chemin pour le photon si la condition $g_{22} = 0$ est satisfaite, c'est-à-dire si la vitesse du diélectrique est donnée par

$$\frac{v}{c^2} = \sqrt{\epsilon\,\mu}. \tag{16.58}$$

Une telle vitesse est possible pour des matériaux réels (*e.g.*, $v < c$) lorsque la permitivité est plus petite que 1, c'est-à-dire, lorsque la susceptibilité électrique χ, définie par $\epsilon = \epsilon_0\,(1 + \chi)$ est négative.

10. Exemple-II

Le système physique consiste en un fil chargé à travers un solénoide. La géométrie de fond (Minkowski) est écrite dans un système de coordonnées (t, r, φ, z)

$$ds^2 = dt^2 - dr^2 - r^2\, d\varphi^2 - dz^2. \tag{16.59}$$

Les composantes non-nulles du tenseur de Maxwell $F^{\mu\nu}$, compatibles avec les symétries et les propriétés du système, sont $F^{01} = E(r)$ et $F^{12} = B(r)$. Dans ce cas, les équations de mouvement se réduisent à

$$r\,L_F\,F^{01} = Q \tag{16.60}$$

$$r\,L_F\,F^{12} = \mu, \tag{16.61}$$

où les constantes Q et μ déterminent la densité de charge du fil et le courant à travers le solénoide, respectivement. Nous sommes intéressés par l'analyse de la propagation des ondes électromagnétiques dans la région intérieure au solénoide. Le même traitement qu'auparavant montre que les photons se propagent comme dans un espace-temps de métrique effective riemannienne. De l'équation (16.27), nous obtenons les composantes de la métrique effective :

$$g^{00} = 1 - \psi\,E^2 \tag{16.62}$$

$$g^{11} = -1 - \psi\,\left(B^2\,r^2 - E^2\right) \tag{16.63}$$

$$g^{02} = \psi\,E\,B \tag{16.64}$$

$$g^{22} = -\left(\frac{1}{r^2} + \psi\,B^2\right) \tag{16.65}$$

$$g^{33} = -1, \tag{16.66}$$

où Ψ est donné par

$$\Psi = 4\,\frac{L_{FF}}{L_F}\cdot \tag{16.67}$$

Les trajectoires des photons sont des géodésiques nulles dans cette métrique. Procédons comme ci-dessus. On considère une courbe définie par les équations $t = constant$, $r = constant$, et $z = constant$. L'élément infinitésimal de longueur est

$$ds^2_{\text{eff}} = g_{22}\,d\varphi^2. \tag{16.68}$$

La courbe peut être une trajectoire du photon s'il existe une valeur $r = r_c$ telle que $g_{22}(r_c) = 0$. En termes des composantes contravariantes de la métrique effective, cette condition s' écrit

$$\left(1 - \Psi E^2\right) = 0. \tag{16.69}$$

La solution de cette équation est

$$r_c = \frac{2Q}{L_F}\left(\frac{L_{FF}}{L_F}\right)^{1/2} \tag{16.70}$$

et par conséquent

$$\left.\frac{L_{FF}}{L_F}\right|_{r_c} > 0. \tag{16.71}$$

Un exemple simple d'une telle situation est donné par un lagrangien du genre Born-Infeld

$$L = \frac{\beta^2}{2}\left(\sqrt{1-\frac{F}{\beta^2}}-1\right). \tag{16.72}$$

11. Commentaires

Ces deux exemples montrent que la possibilité d'avoir une trajectoire fermée pour des particules n'est pas une exclusivité de la gravitation dans le schéma de la théorie de la relativité générale. Nous avons montré qu'il peut exister une courbe fermée pour le photon dans un régime non-linéaire de l'électromagnetisme. On s'aperçoit alors que le point important n'est pas la force gravitationnelle mais son caractère non-linéaire. Une situation semblable peut exister en électromagnetisme, dans un régime non-linéaire.

Cela entraîne des situations tout à fait étranges et non-conventionnelles. Cela remet en cause la question ancienne : quelle est la vraie nature de la géométrie de l'espace-temps ? D'après ce que nous avons exposé ici, cela dépend de l'objet que l'on utilise pour la mesurer. Pour des photons dans un milieu non-linéaire, la réponse peut être bien différente de celle que nous aurions obtenue en utilisant de la matière comme des électrons ou des protons. Cela nous ramène à une situation semblable à celle du commencement du XX^{e} siècle : il n'existe pas une seule géométrie de l'espace-temps.

17 Équations (F.R.W.) de la cosmologie et cosmologie quantique

Edgar Elbaz

Le modèle du Big Bang de la cosmologie, basé sur l'utilisation des équations fondamentales de la Relativité Générale, utilise de manière extensive, les équations (F.R.W.) de la cosmologie [1].

Pour en rappeler brièvement la formulation et le contenu, nous indiquerons que ces équations utilisent en premier lieu l'hypothèse du *Principe Cosmologique*, posant que la densité d'énergie dépend de la seule coordonnée de temps. La définition de coordonnées co-mobiles, pour un observateur en mouvement avec le fluide cosmique, introduit le concept de **temps cosmologique**, indispensable à la description globale de notre univers. Cette hypothèse permet en effet de sortir de l'aspect local, inhérent à la Relativité Générale, pour conduire à une description globale de notre univers depuis sa création.

La métrique (F.R.W.) de Friedmann-Robertson-Walker

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + R^2(t)\left[\mathrm{d}r^2/(1-kr^2) + r^2\mathrm{d}\Omega^2\right] \qquad (17.1)$$

montre en effet que le coefficient $g_{00}(x)$ est indépendant de la position d'un observateur et conduit ainsi à un temps t valable partout, temps global, temps cosmologique. La fonction $R(t)$ représente un facteur d'échelle, improprement appelé « rayon d'univers », le paramètre k, enfin, désigne la courbure spatiale de l'Univers, de valeur +1 dans un modèle elliptique, −1 hyperbolique, et 0 pour un modèle euclidien.

Les équations fondamentales de la cosmologie [1], s'écrivent souvent avec le facteur d'expansion θ, dérivée par rapport au temps du logarithme du facteur d'échelle élevé au cube :

$$\theta = 3(\partial_t R)/R = \mathrm{d}/\mathrm{d}t(\log R^3). \qquad (17.2)$$

L' équation d'état-γ du fluide cosmique

$$P(t) = (\gamma - 1)\rho(t) \tag{17.3}$$

lie la densité d'énergie ρ, à la pression P (dans un système d'unités naturel dans lequel $G = c = 1$), et les différents états du fluide cosmique sont associées aux valeurs du coefficient :

$$\begin{array}{lll} \gamma = 0 & P = -\rho & \text{vide quantique} \\ \gamma = 4/3 & P = 1/3\rho & \text{ère radiative} \\ \gamma = 1 & P = 0 & \text{ère de matière (actuelle).} \end{array} \tag{17.4}$$

Les équations d'Einstein de la Relativité générale permettent ensuite d'introduire les trois équations fondamentales (F.R.W.), de la cosmologie [1] :

Conservation de l'énergie :

$$\partial_t \rho + \gamma\, \rho\theta = 0. \tag{17.5}$$

Équation de Raychaudhuri :

$$\partial_t \theta + 1/3\theta^2 + 1/2(3\gamma - 2) = 0. \tag{17.6}$$

Équation F.R.W. :

$$1/3\theta^2 = \rho - 3k/R^2. \tag{17.7}$$

Il faut d'ailleurs noter que cette dernière équation peut être considérée comme la traduction de la conservation de l'énergie totale E d'une particule de masse m, soumise à un potentiel gravitationnel newtonien classique non-relativiste $\mathrm{V}(R) = -G\,\mathrm{M}\,m/R$, dans laquelle on pose $k = -2E/m$, pour un fluide de densité uniforme ρ, de pression P nulle [1].

Reprenant l'essentiel d'un article déjà publié [2], déterminons, une forme hamiltonienne des équations précédentes en posant :

$$\mathrm{H} = -p^2 + A\,q^{\mathrm{n}}. \tag{17.8}$$

Les coefficients A et n, ainsi qu'une relation entre les variables (p, q) et (θ, ρ) restent à établir. On définit les variables p et q par les formes monomiales :

$$p = a\theta^{\alpha}\rho^{\beta} \tag{17.9}$$

$$q = b\rho^{\sigma}. \tag{17.10}$$

La première équation de Hamilton $\partial_t q = \partial_p$ H donne avec les formes précédentes de p et q la valeur de

$$\partial_t \rho = a/\sigma b\theta^{\alpha} \rho^{\beta-\sigma+1}, \tag{17.11}$$

ce qui, par comparaison avec (18.5) fixe les valeurs des coefficients

$$\alpha = 1 \quad a/\sigma b = -\gamma \,. \tag{17.12}$$

La seconde équation de Hamilton $\partial_t p = -\partial_q$ H donne avec la forme choisie de $V(q)$ et les coefficients précédents

$$\partial_t \theta + a/b\theta^2 + \mathrm{n}Ab^{\mathrm{n}-1}/a\rho^{\beta(\mathrm{n}-2)} = 0. \tag{17.13}$$

La comparaison avec (18.6) fixe donc les paramètres A et n :

$$\begin{aligned} &a/b = 1/3 = -\gamma\,\beta, \\ &\beta(\mathrm{n}-2) = 1 \quad \text{et} \quad \mathrm{n} = (2-3\gamma), \\ &A = -1/6b^{2-\mathrm{n}}. \end{aligned} \tag{17.14}$$

On a ainsi obtenu une forme hamiltonienne [2] des équations F.R.W. de la cosmologie puisque :

$$\mathrm{H} = -p^2 - 1/6b^{3\gamma} q^{2-3\gamma} \tag{17.15}$$

avec les variables conjuguées :

$$p = 1/3b\theta\rho^{-1/3\gamma} \tag{17.16}$$

$$q = b\rho^{-1/3\gamma}\,. \tag{17.17}$$

On obtient en particulier le rapport

$$p/q = 1/3\theta = \partial_t \log R. \tag{17.18}$$

Pour que q ait la dimension d'une longueur et p celle d'un moment, il faut affecter une dimension au paramètre libre b. L'intégration de l'équation (17.5) de conservation de l'énergie donne $\rho(t)$ en fonction du facteur d'échelle $R(t)$:

$$\rho(t) = \mathrm{Const}[R(t)] - 3\gamma = M\gamma\,[R(t)] - 3\gamma\,. \tag{17.19}$$

Dans l'ère quantique ($\gamma = 0$), la densité d'énergie constante ρ_0 est la constante M_0 elle-même, tandis que dans chacune des ères qui ont suivi,

$$M_\gamma = (\rho R^{3\gamma}), \tag{17.20}$$

constante durant la période considérée, alors même que le facteur d'échelle augmente. En portant $\rho(t)$ dans la définition (17.17) de la variable q, il vient

$$q = (bM_\gamma^{-1/3\gamma})R. \tag{17.21}$$

La constante b étant arbitraire, on peut la fixer à la valeur $M_\gamma^{1/3\gamma}$, de sorte que la variable q soit égale au facteur d'échelle, tandis que la relation (17.18) montre ensuite facilement que

$$q = R \quad \text{entraîne} \quad p = \partial_t q = \partial_t R. \tag{17.22}$$

Cette approche hamiltonienne du fluide cosmique, montre en particulier que l'on peut modéliser la force agissant sur ce fluide dans ses différents états-γ. Le potentiel effectif d'interaction $V_\gamma(q)$ s'écrit en effet

$$V_\gamma(R) = -1/6M_\gamma R^{2-3\gamma}. \tag{17.23}$$

Cela conduit à une force newtonienne effective $F_\gamma = (1/3 - 1/2\gamma)M_\gamma R^{1-3\gamma}$, **attractive** : $F_1(R) = -1/6M_1R^{-2}$ pour un fluide cosmique sans pression ; $F_{4/3}(R) = -1/3M_{4/3}R^{-3}$ pour un fluide constitué de pur rayonnement.

Le vide quantique est modélisé classiquement par une **force répulsive** proportionnelle au facteur d'échelle $F_0(R) = 1/3M_0R$. Est-ce la traduction de l'existence de la constante cosmologique ? Il faut alors noter que cette force répulsive n'est pas constante, car le facteur d'échelle R dépend du temps cosmologique. Cela est assez étonnant quand on pense à l'hypothèse de la **quintessence**, récemment introduite par les cosmologistes [3], pour expliquer l'accélération de l'expansion, mesurée en 1998 par deux équipes d'astrophysiciens [4].

Réécrivons maintenant le hamiltonien H, avec les variables θ et ρ, en utilisant (17.15), (17.16), (17.17), puis la dernière des équations F.R.W. :

$$H = 1/6b^2\rho^{-2/3\gamma}[1/3\theta^2 - \rho] = -1/2(bM_\gamma^{-1/3\gamma})^2k = -k/2. \tag{17.24}$$

La comparaison avec le cas classique, non-relativiste mentionné au début montre que la constante apparaissant au second membre de l'équation précédente, est l'énergie totale du fluide cosmique dans son état-γ, modélisé maintenant par une particule de masse unité, de moment p, dans un potentiel $V_\gamma(q)$. On peut donc noter

$$\underline{H} = E_\gamma. \tag{17.25}$$

Cela montre en particulier que ce hamiltonien est nul lorsque le fluide cosmique décrit le vide quantique, ère pendant laquelle $\gamma = 0$ et $P = -\rho = -\rho_0$, hypothèse habituellement utilisée en Cosmologie Quantique, tandis que la courbure spatiale k est de valeur nulle, d'après (17.24), en cette période. Cela implique que la densité d'énergie du vide ρ_0 est la densité critique $\rho_c = 3\theta^2/8\pi G$. On peut en déduire l'évolution du facteur d'échelle en fonction du temps cosmologique durant cette ère, puisque par une simple intégration, on obtient $R(t) = R(0)\exp[(\sqrt{8\pi G}/3\rho_0)t]$. La valeur initiale $R(0)$ peut être quelconque, différente de zéro, excluant ainsi le problème de la singularité initiale. En supposant que le vide quantique prend place pendant l'ère de Planck, le facteur d'échelle au début de l'ère radiative, ne sera pas nul, mais égal à $R_P(t) = R(0)\exp[(\sqrt{8\pi G}/3\rho_0)t_P]$, où t_P est le temps de Planck, de l'ordre de 10^{-44} seconde, les paramètres $R(0)$ et ρ_0 restant indéterminés.

Pendant les ères radiatives, puis de matière qui se sont succédé, l'énergie E_γ, correspondant à une valeur constante du précédent hamiltonien, est proportionnelle à la courbure spatiale, et inversement proportionnelle à une énergie de référence M_γ.

La forme hamiltonienne des équations F.R.W. d'un fluide cosmique d'équation d'état-γ est donc :

$$H = -p^2 - 1/6b^{3\gamma} q^{2-3\gamma} = E_\gamma . \qquad (17.26)$$

Cela conduit à une méthode nouvelle d'approche de l'univers quantique, complètement différente,et combien plus simple que celle de Wheeler et De Witt !

Il est en effet facile de déterminer la fonction d'onde de ce fluide en considérant le hamiltonien comme un opérateur agissant sur la fonction d'onde $\psi_\gamma(q)$ du fluide cosmique dans l'état-γ, et en introduisant le principe de correspondance de la Mécanique Ondulatoire $p \Rightarrow -i\partial/\partial q$ ce qui donne une équation de type Schrödinger :

$$[\partial^2/\partial q^2 + 1/3b^{3\gamma} q^{2-3\gamma}]\psi_\gamma(q) = E_\gamma \psi_\gamma(q). \qquad (17.27)$$

En revenant à la variable R, facteur d'échelle, avec la relation (17.21) et le choix précédent de la constante b, on obtient **l'équation fondamentale de la Cosmologie Quantique, pour un fluide cosmique dans l'état-γ** :

$$[\partial^2/\partial R^2 + 1/3M_\gamma R^{2-3\gamma}]\psi_\gamma(R) = E_\gamma \psi_\gamma(R). \qquad (17.28)$$

L'état de vide quantique ($\gamma = 0$), domaine d'application naturel de la Cosmologie Quantique, conduit à l'équation simple :

$$[\partial^2/\partial R^2 + 1/3\rho_0 R^2]\psi_0(R) = E_0\psi_0(R). \tag{17.29}$$

Il faut noter que ces équations d'onde ne dépendent pas explicitement du temps cosmologique ; celui-ci apparaît par le biais du facteur d'échelle $R(t)$.

L'équation (17.28) a été utilisée par T.R. Mongan [5], dans l'ère purement radiative, montrant que la solution est du genre Hankel sphérique,

$$\psi_{4/3} = ARh_{ip-1/2}(i\kappa R) \text{ avec } \kappa^2 = E_{4/3},$$

dans laquelle $p = (1/3M_{4/3} - 1/4)^{1/2}$ et A est une constante de normalisation. Il en déduit que la fonction d'onde s'annule pour $R = 0$, ne prenant pas en compte le fait que l'ère radiative ne prend place qu'après l'ère quantique décrite par l'équation (17.29).

L'intérêt de notre méthode est qu'elle permet de suivre la fonction d'onde du fluide cosmique dans ses différents états : vide quantique, puis ère purement radiative, puis ère de matière enfin, conduisant à la description classique de notre univers. En raccordant les fonctions d'onde et leur dérivée première à chaque changement d'ère, donc d'équation d'état, on peut espérer mieux comprendre l'évolution de notre univers depuis sa création, dans le modèle du Big-Bang. Le problème de l'interprétation de la fonction d'onde, et du passage de l'ère quantique à l'ère radiative classique est le même qu'avec la détermination de la fonction d'onde par la méthode de Wheeler–De Witt.

Bibliographie

[1] Elbaz, E., *Cosmologie*, Ellipses, Paris, 1992.

[2] Elbaz, E., Novello, M., Salim, J.M., Motta da Silva, M.C., Klippert, R., *General Relativity* 29 (1997) 481.

[3] Caldwell, R.R., Dave, R., Steinhardt, P.J., *Phys. Rev. Lett.* 377 (1995) 600.

[4] Turner, M.S., *Type la Supernovae: Theory and Cosmology*, Cambridge University Press, 2000.

[5] Mongan, T.R., *General Relativity and Gravitation* 31 (1999) 9 ; *Intern. J. Theor. Phys.* 38 (1999) 5.

Dans la même collection...

Sur la science cosmologique
Conditions de possibilité et problèmes philosophiques

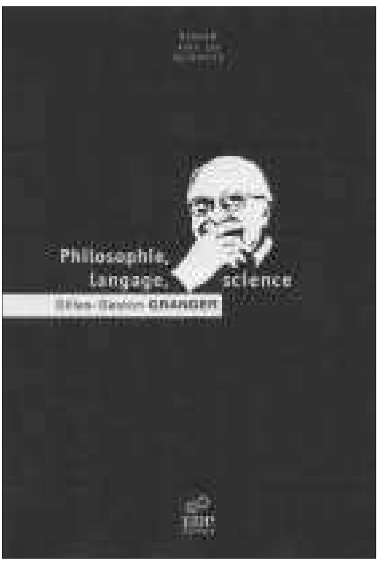
Jacques Merleau-Ponty

Comment s'est formé l'Univers ? Quelle est sa nature ? Qu'est-ce que le « big-bang » ? Jacques Merleau-Ponty a consacré la majeure partie de son œuvre à l'analyse des théories cosmologiques du XX^{e} siècle, sur lesquelles il a porté son regard de philosophe et d'historien des sciences.

Les textes réunis dans cet ouvrage, dont certains sont inédits, retracent les réflexions de leur auteur sur la définition de la cosmologie, les problèmes philosophiques posés par cette discipline et sur les conceptions des idées cosmologiques qui ont eu cours depuis le XVIII^{e} siècle, de Pierre Simon de Laplace à Albert Einstein.

Jacques Merleau-Ponty, disparu le 7 juin 2002 à l'âge de 86 ans, a consacré la plus grande partie de son œuvre à l'analyse des théories cosmologiques du XX^{e} siècle. Il était professeur de philosophie et d'histoire des sciences à l'Université Paris X-Nanterre.

• 2003 • 2-86883-579-1 • 352 pages • 39,50 €



Philosophie, langage, science

Gilles-Gaston Granger

Les ouvrages et l'enseignement de Gilles-Gaston Granger ont marqué plusieurs générations de philosophes. Portées par l'exigence d'une réflexion sur les sciences nourrie par l'étude des œuvres scientifiques, ses recherches d'« épistémologie comparative » se tiennent à la jonction des courants de l'épistémologie historique et de l'analyse logique et structurale de ces œuvres.

Dans cet ouvrage, Philosophie, langage, science, *qui réunit des textes soit inédits, soit dispersés dans des publications d'accès difficile aujourd'hui, Gilles-Gaston Granger examine le type particulier de connaissance qu'est la connaissance philosophique dans son rapport au langage, à l'histoire et aux sciences.*

Gilles-Gaston Granger, professeur émérite au Collège de France, a fondé et dirigé la chaire « Épistémologie comparée » à l'Université d'Aix-en-Provence.

• 2003 • 2-86883-578-3 • 344 pages • 39,50 €