

Introduction à la pensée scientifique moderne

Pierre Sagaut

Institut Jean Le Rond d'Alembert

Université Pierre et Marie Curie – Paris 6

pierre.sagaut@upmc.fr
<http://www.lmm.jussieu.fr/~sagaut>

Cours de culture générale, Licence

Année 2008-2009

Table des matières

Table des matières

1	INTRODUCTION	8
1.1	OBJECTIF DU COURS	8
1.2	LA SCIENCE NE SE RESUME PAS A DES EQUATIONS !	8
1.3	QU'EST-CE QUE L'EPISTEMOLOGIE ?	10
1.3.1	PRESENTATION GENERALE	10
1.3.2	UNE VUE ALTERNATIVE : L'EPISTEMOLOGIE COMME SCIENCE EMPIRIQUE	13
1.4	SOUS QUELLES CONDITIONS SAIT-ON QUELQUE CHOSE ?	14
1.4.1	DISCUSSION LIMINAIRE : QU'EST-CE QU'UNE DEFINITION ?	14
1.4.2	LA CONNAISSANCE SELON PLATON	15
1.4.3	LE PROBLEME DE GETTIER	15
1.4.4	REPNSES AU PROBLEME DE LA JUSTIFICATION : INTERNALISME, EXTERNALISME	17
1.4.5	LA CONNAISSANCE SELON NOZICK ET DRETSKE	19
1.4.6	LE DEFI DU SCEPTICISME	19
1.4.7	LES REPNSES AU SCEPTICISME	21
2	QU'EST-CE QUE LA SCIENCE ?	22
2.1	COMMENT DEFINIR LA SCIENCE ?	22
2.1.1	DEFINITION(S) DE LA SCIENCE ET CRITERES DE SCIENTIFICITE	22
2.1.2	L'OPPOSITION RATIONALISME/RELATIVISME	24
2.2	CLASSIFICATION(S) DES SCIENCES	24
2.2.1	QUELQUES ELEMENTS DE CLASSIFICATION	25
2.2.2	PRESENTATION HISTORIQUE DES CLASSIFICATIONS	26
2.3	LA SCIENCE COMME PHENOMENE SOCIAL	30
2.3.1	LA SCIENCE ET LE LANGAGE	30
2.3.2	LA SCIENCE AU-DELA DES INDIVIDUS ? OBJECTIVISME ET INDIVIDUALISME	31
3	ELEMENTS CONSTITUTIFS D'UNE SCIENCE EMPIRIQUE	33
3.1	FAITS, LOIS, PRINCIPES, MODELES ET THEORIES	33
3.1.1	FAITS, PHENOMENES ET EVENEMENTS	33
3.1.2	QU'EST-CE QU'UNE LOI SCIENTIFIQUE ? ET UN PRINCIPE ?	33
3.1.3	QU'EST-CE QU'UN MODELE ?	38
3.1.4	QU'EST-CE QU'UNE THEORIE SCIENTIFIQUE ?	41
3.2	APPROCHES GLOBALES DE LA SCIENCE ET DES THEORIES	42
3.2.1	PARADIGME ET MATRICE DISCIPLINAIRE : KUHN	42
3.2.2	LES PROGRAMMES DE RECHERCHE : LAKATOS	44
3.3	DEFINIR UN OBJET	45
3.3.1	GENERALITES : DEFINITION EXTENSIONNELLE ET DEFINITION ANALYTIQUE/INTENSIONNELLE	46
3.3.2	UNE MODE DEFINITOIRE PROPRE AUX SCIENCES EMPIRIQUES : LA DEFINITION OPERATOIRE	49
3.3.3	EXEMPLE 1 : QU'EST-CE QUE LA MASSE ? QU'EST-CE QU'UNE FORCE ?	50
3.3.4	EXEMPLE 2 : QU'EST-CE QUE LE MOUVEMENT ? QU'EST-CE QU'UNE DEFORMATION ?	52
3.4	DEFINIR UN OBJET (SUITE) : COMPLEXITE ET FRONTIERES FLOUES	54
3.4.1	CONCEPTS VAGUES ET PARADOXES SORITES	55
3.4.2	EXEMPLE 1 : VIVANT OU INERTE ? MIMIVIRUS ET ATTV	56
3.4.3	EXEMPLE 2 : LA DEFINITION DE LA MORT ET LA CRYPTOBIOSE DU TARDIGRADE	60

3.4.4	EXEMPLE 3 : ANIMAL, CHAMPIGNON OU VEGETAL ?	62
3.5	DEFINIR UN OBJET (FIN) : LA CRITIQUE DE LAKATOS	64
4	QU'EST-CE QUE FAIRE UNE EXPERIENCE ?	66
4.1	DE L'OBSERVATION A L'EXPERIMENTATION	66
4.2	LA MESURE DES PHENOMENES PHYSIQUES	66
4.2.1	PERCEPTION, MESURE ET TEMOIGNAGE DES DONNEES	66
4.2.2	MESURE, PRECISION ET ETALONS	67
4.2.3	REPERES : HISTOIRE DU SYSTEME METRIQUE	69
4.2.4	LES INSTRUMENTS DE MESURE	71
4.2.5	EXEMPLE 1 : LA MACHINE D'ATTWOOD : MESURER LA MASSE ET LES FORCES	72
4.2.6	EXEMPLE 2 : LA CHAMBRE A BULLES : « VOIR » LES PARTICULES EN MOUVEMENT	78
4.2.7	EXEMPLE 3 : LE MICROSCOPE A EFFET TUNNEL : « VOIR » LES ATOMES	80
4.2.8	EXEMPLE 4 : LES LENTILLES GRAVITATIONNELLES : « VOIR » LES OBJETS LOINTAINS	82
4.3	L'OBSERVATION N'EST PAS NEUTRE	85
4.4	PERCEPTION ET FIABILITE DES OBSERVATIONS	87
4.4.1	RECEPTION ET PERCEPTION	87
4.4.2	PERCEPTION DE L'ESPACE ET GEOMETRIE(S)	92
5	LA VALEUR DES SCIENCES EMPIRIQUES	97
5.1	PRELIMINAIRE 1: QU'EST-CE QUE LA VERITE ?	97
5.2	PRELIMINAIRE 2 : UN PEU DE LOGIQUE	98
5.2.1	LES 3 AXIOMES D'ARISTOTE	98
5.2.2	TARSKI : CORRESPONDANCE, VERITE, LANGAGE ET META-LANGAGE	99
5.2.3	GÖDEL : VERITE, DEMONSTRABILITE ET INCOMPLETUDE	100
5.2.4	INTERLUDE : INFORMATIQUE, CALCULABILITE ET THEORIE DES NOMBRES	101
5.3	EST-IL POSSIBLE DE FONDER LA SCIENCE SUR LES FAITS ?	105
5.3.1	CARNAP ET L'EMPIRISME LOGIQUE	105
5.3.2	QUINE ET LE RELATIVISME ONTOLOGIQUE	107
5.4	PEUT-ON VERIFIER LES ENONCES SCIENTIFIQUES ?	108
5.4.1	VERIFICATION DES ENONCES D'OBSERVATION	108
5.4.2	VERIFICATION DES ENONCES THEORIQUES	109
5.5	VALIDER LES THEORIES SCIENTIFIQUES : LES METHODES	111
5.5.1	QU'EST-CE QU'UNE METHODE ?	111
5.5.2	LE CAS DES SCIENCES FORMELLES : LA DEMONSTRATION	114
5.5.3	LE CAS DES SCIENCES EMPIRIQUES : LA METHODE EXPERIMENTALE	116
5.5.4	RETOUR SUR LA THEORIE DE L'INDUCTION	119
5.6	LE FALSIFICATIONISME : POPPER	121
5.6.1	LA THEORIE FALSIFICATIONISTE	121
5.6.2	LE CONCEPT DE VERISIMILARITE DE POPPER	123
5.7	LE HOLISME EPISTEMOLOGIQUE : DUHEM, QUINE	123
5.8	EFFICACITE ET PRODUCTION DE TECHNOLOGIES	125
5.8.1	RELATIVITE GENERALE ET LOCALISATION PAR SYSTEME GPS	126
5.8.2	MANIPULER LES MOLECULES : PINCES ET CISEAUX OPTIQUES	127
6	LA SCIENCE EXPLIQUE-T-ELLE QUELQUE CHOSE ?	130
6.1	DECRIRE OU EXPLIQUER ?	130
6.1.1	QUE VEUT DIRE EXPLIQUER ?	130
6.1.2	EXPLICATION NOMOLOGICO-DEDUCTIVE (N-D)	131
6.1.3	EXPLICATION INDUCTIVE-STATISTIQUE (I-S)	133
6.1.4	CLASSIFICATION DE HALBWACHS	133
6.1.5	EXPLICATIONS NON-CAUSALES	135

6.2	EXPLIQUER OU COMPRENDRE ?	136
6.1	INTERPRETER, EXPLIQUER ET COMPRENDRE	137
6.2	SCIENCES PREDICTIVES ET SCIENCES ANTICIPATIVES	137
6.2.1	PREDICTION, ANTICIPATION ET SCIENTIFICITE	137
6.2.2	PREVISION = VALIDATION ? LA DECOUVERTE DE NEPTUNE	138
6.2.3	ET LORSQUE CELA NE MARCHE PAS ? LES TRAJECTOIRES DE MERCURE ET DE LA LUNE	138
7	<u>LA SCIENCE NOUS APPREND-ELLE QUELQUE CHOSE SUR LE MONDE ?</u>	140
7.1	L'OPPOSITION REALISME/ANTIREALISME	140
7.1.1	REALISME ET IDEALISME	140
7.1.2	RETOUR SUR LA QUESTION DE LA VERITE	142
7.1.3	LA THEORIE SCIENTIFIQUE EST-ELLE UNIQUE ?	142
7.1.4	PEUT-ON TRANCHER LE DEBAT ?	142
7.2	LE PHENOMENISME	143
7.3	LE CONVENTIONNALISME	145
7.4	LE REDUCTIONNISME	145
7.4.1	ISOLATIONNISME ET CONSTRUCTIONNISME	145
7.4.2	UN EXEMPLE ISSU DE LA PHYSIQUE MODERNE : LE MODELE STANDARD	149
8	<u>LES MATHEMATIQUES DANS LES SCIENCES EMPIRIQUES</u>	155
8.1	UN EXEMPLE : LA MECANIQUE	155
8.1.1	EXEMPLE 1 : LE CALCUL QUANTITATIF	155
8.1.2	EXEMPLE 2 : LA MANIPULATION SYMBOLIQUE	155
8.1.3	DEDUIRE LES LOIS DE LA PHYSIQUE : SYMETRIES ET PRINCIPES VARIATIONNELS	157
8.2	LES APPORTS DE LA MATHEMATISATION	159
8.3	POSER ET RESOUDRE UNE EQUATION : HISTOIRE D'UN LONG PROCESSUS	161
8.3.1	UNE BREVE HISTOIRE DU CALCUL	162
8.3.2	UNE « INVENTION » REVOLUTIONNAIRE : L'INCONNUE	165
9	<u>LA QUESTION DU PROGRES SCIENTIFIQUE</u>	168
9.1	CONSTRAINTES SUR L'ACCEPTATION DES THEORIES	169
9.1.1	LES TYPES DE CONTRAINTE	169
9.1.2	UNE EVOLUTION STRICTEMENT OBJECTIVE ? L'EPISTEMOLOGIE NAÏVE	170
9.1.3	UNE EVOLUTION PARTIELLEMENT SUBJECTIVE ? INTERNALISME ET EXTERNALISME	170
9.1.4	LA NOTION D'OBSTACLE EPISTEMOLOGIQUE SELON BACHELARD	171
9.1.5	UN CAS D'ECOLE : L'AFFAIRE LYSENKO	174
9.2	CONCEPTIONS DU PROGRES SCIENTIFIQUE	174
9.2.1	LA CONCEPTION TELEOLOGIQUE DU PROGRES	175
9.2.2	LA CONCEPTION EVOLUTIONNISTE DU PROGRES	175
9.3	LA DYNAMIQUE DES SCIENCES : CONTINUITE OU REVOLUTIONS ?	176
9.3.1	EVOLUTION DISCONTINUE ET REVOLUTIONS SCIENTIFIQUES : KUHN	176
9.3.2	LA THESE CONTINUISTE	179
9.4	FALSIFICATIONISME, PROGRES ET EVOLUTION DES THEORIES SCIENTIFIQUES	180
9.5	L'EVOLUTION DES PROGRAMMES DE RECHERCHE ET DES THEORIES SELON LAKATOS	182
9.6	LE PROGRES SCIENTIFIQUE EST-IL INELUCTABLE ?	184
10	<u>CAUSE, CAUSALITE, DETERMINISME ET HASARD</u>	186
10.1	CAUSE ET CAUSALITE : DE ARISTOTE AU XXE SIECLE	186
10.1.1	LES QUATRE CAUSES SELON ARISTOTE	187
10.1.2	LA CAUSALITE SELON HUME	188
10.1.3	LA CAUSALITE CHEZ KANT	190
10.1.4	MACH : LA THEORIE DE LA RELATION FONCTIONNELLE	192

10.1.5	LA CAUSALITE PERCEPTIBLE : WHITEHEAD ET PIAGET	193
10.1.6	LA CAUSALITE SELON RUSSELL : LES LIGNES CAUSALES	194
10.1.7	CAUSALITE ET COMPLEXITE : LA CAUSE INUS DE MACKIE	195
10.1.8	LA CAUSALITE SELON SALMON : LA THEORIE MT	198
10.1.9	LA THEORIE DE LA QUANTITE CONSERVEE : SKYRMS	201
10.2	CAUSALITE ET LOIS SCIENTIFIQUES	201
10.3	LE DETERMINISME	205
10.3.1	LE DETERMINISME SELON LAPLACE	205
10.3.2	L'ANALYSE DE KOJEVE	208
10.3.3	LEGALISME, DETERMINISME, INCOMPLETUDE ET INDETERMINISME	211
10.4	LE HASARD	212
10.4.1	LE HASARD SELON ARISTOTE	213
10.4.2	LE HASARD SELON COURNOT	215
11	BREVE INCURSION DANS LE MONDE QUANTIQUE	216
11.1	EXEMPLE INTRODUCTIF : L'EXPERIENCE DES FENTES D'YOUNG	216
11.2	PAQUET D'ONDES ET D'ETATS SUPERPOSES	219
11.3	MESURE ET REDUCTION DU PAQUET D'ONDES	221
11.4	INTRICATION QUANTIQUE ET NON-SEPARABILITE	222
11.5	CONSEQUENCES EPISTEMOLOGIQUES	224
11.6	VERS DE NOUVELLES APPLICATIONS : CRYPTOGRAPHIE QUANTIQUE, INFORMATIQUE QUANTIQUE ET TELEPORTATION QUANTIQUE	225
11.6.1	ORDINATEUR QUANTIQUE	226
11.6.2	CRYPTOGRAPHIE QUANTIQUE	228
11.6.3	TELEPORTATION ET FAX QUANTIQUE	230
12	HISTOIRE DE LA SCIENCE DU MOUVEMENT	232
12.1	LA PHYSIQUE D'ARISTOTE	232
12.2	LA PHYSIQUE MEDIEVALE	235
12.2.1	JEAN PHILOPON D'ALEXANDRIE : L'ENERGEIA	235
12.2.2	LA NAISSANCE DE LA CINEMATIQUE : LE MERTON COLLEGE D'OXFORD	236
12.2.3	LA DYNAMIQUE DE L'IMPETUS : BURIDAN	237
12.3	LA GEOMETRISATION DU MOUVEMENT	239
12.3.1	GALILEE : LA CHUTE DES GRAVES	239
12.3.2	DESCARTES : LES LOIS DU MOUVEMENT	244
12.3.3	HUYGENS : L'ORGANISATION DEDUCTIVE DES SCIENCES DU MOUVEMENT ET LES CHOCS	246
12.4	LA SCIENCE NEWTONIENNE	249
12.5	L'ALGORITHMISATION DE LA SCIENCE DU MOUVEMENT	253
12.5.1	LEIBNIZ : LE CALCUL DIFFERENTIEL ET INTEGRAL	253
12.5.2	VARIGNON : L'ALGORITHME DE LA CINEMATIQUE	254
12.6	REORGANISATIONS DE LA MECANIQUE	256
12.6.1	LE PRINCIPE DE MOINDRE ACTION	256
12.6.2	PRINCIPE D'INERTIE ET RELATIVITE DU MOUVEMENT : DES PREMISSES AU PRINCIPE DE MACH	258
12.6.3	LA RELATIVITE DU MOUVEMENT ET LE VIDE QUANTIQUE	265

Avertissement

Le présent document est un support de cours destiné aux étudiants de l'Université Pierre et Marie Curie. Il ne s'agit pas d'une œuvre originale : la rédaction est largement et librement inspirée d'œuvres originales qui sont dûment référencées, et dont la lecture est vivement conseillée. Les principales sources ayant servi à la préparation de ces notes de cours pour les chapitres 1 à 10 sont :

- « La science et l'hypothèse », H. Poincaré, Flammarion, 1968
- « Les épistémologies constructivistes », J.L. Le Moigne, Que sais-je 2969, PUF, 1995
- « L'idée du déterminisme dans la physique classique et dans la physique moderne », A. Kojève, Le livre de poche, 1990
- « La pensée scientifique moderne », J. Ullmo, Flammarion, 1969
- « En quête de science », M. Gagnon et D. Hébert, FIDES, 2000
- « Histoire des méthodes scientifiques », J.M. Nicolle, Bréal, 1994
- « Eléments d'épistémologie », C. Hempel, Armand Colin, 1972
- « Introduction à l'épistémologie », L. Soler, Ellipses, 2000
- « La philosophie des sciences », D. Lecourt, Que sais-je 3624, PUF, 2001
- « L'épistémologie », H. Barreau, Que sais-je 1475, PUF, 1990
- « Les théories de la causalité », M. Bunge et al. éditeurs, PUF, 1971
- « Enquête sur le concept de causalité », L. Viennot et C. Debru éditeurs, PUF, 2003
- « La querelle du déterminisme », K. Pomian éditeur, Gallimard, 1990
- « Philosophie de la connaissance », J. Dutant et P. Engel éditeurs, Vrin, 2005
- « Philosophie des sciences », t. I et II, S.Laugier et P. Wagner éditeurs, Vrin, 2004
- « Philosophie des sciences », t. I et II, D. Andler, A. Fagot-Largeault, B. Saint-Sernin, Folio, collection « Essais », Gallimard, 2002
- « Qu'est-ce-que la science ? », A.F. Chalmers, Livre de poche, collection « Biblio essais », 1987
- « La philosophie des sciences au XXe siècle », A. Barberousse, M. Kistler et P. Ludwig, collection « Champs Universités », Flammarion, 2000

pour le chapitre 11

- « Petit voyage dans le monde des quanta », E. Klein, « Champs », Flammarion, 2004
- « Le cantique des quantiques », S. Ortoli, J.P. Pharabod, La découverte, 2007

pour le chapitre 12

- « Et pourtant, elle tourne », J. Gapaillard, Seuil, 1993
- « Chute des corps et mouvement de la Terre », A. Koyré, Vrin, 1973
- « Mécanique, une introduction par l'histoire de l'astronomie », E. Lindemann, De Boeck Université, 1999
- « La science du mouvement », M. Blay, Belin, 2002

- « Petite histoire de la physique », J.P. Maury, Larousse, 1992
- « Histoire de la physique », J. Rosmorduc et al., Lavoisier, 1987
- « La physique au Moyen Age », E. Grant, PUF, 1995
- « Une histoire de la mécanique », R. Dugas, réédition Jacques Gabay, 1996
- « La mécanique », E. Mach, réédition Jacques Gabay, 1987

1 Introduction

1.1 Objectif du cours

La formation scientifique passe par l'apprentissage d'un certain nombre de savoirs et de savoir-faire relatifs aux disciplines considérées. Au-delà de l'acquisition de ces connaissances, il est souhaitable que le scientifique, au cours de sa formation, reçoive les moyens de développer une analyse critique sur la nature de ce qu'il apprend. Il ne s'agit pas ici de statuer sur la véracité ou l'intérêt du contenu des cours dispensés, mais de comprendre ce qu'est la science et de voir en quoi cette forme de connaissance du monde qui nous environne est particulière.

L'objectif premier de ce cours est de présenter les bases de l'épistémologie, comprise comme la théorie de la connaissance scientifique, et plus particulièrement l'épistémologie des sciences dites empiriques (c'est-à-dire celles qui décrivent le monde en se basant sur des données sensibles fournies par l'expérience: mécanique, physique, chimie, ...). L'accent sera porté sur les bases déjà acquises, c'est-à-dire sur la mécanique de Newton. La théorie de la relativité, la mécanique quantique et la mécanique des milieux continus ne seront pas discutées en profondeur car elles ne sont abordées que plus tard dans le cursus universitaire. Seules leur conséquences majeures dans la manière dont les scientifiques perçoivent le monde et conçoivent la science seront abordées. Pour illustrer certains points, des exemples seront empruntés à plusieurs autres disciplines : mathématiques, biologie, informatique.

Outre les bases de la réflexion relative à l'épistémologie de la physique et de la mécanique, ce cours a pour objectif second de présenter, à travers quelques exemples, l'histoire de résultats fondamentaux de la mécanique, comme le principe fondamental de la dynamique (le produit de l'accélération par la masse est égal à la résultante des forces exercées sur un corps solide rigide), la conservation de l'énergie cinétique et de la quantité de mouvement lors du choc élastique de deux solides, ou encore la nature du mouvement des planètes au sein du système solaire. Ces résultats font partie du bagage de tout bachelier se destinant à une formation scientifique supérieure, et paraissent « vrais » et « évidents », c'est-à-dire ne pas prêter à discussion. Or, un aperçu, même bref, de l'histoire du développement de ces résultats, montre que, très loin d'être des choses simples et intuitives, ces résultats sont d'une grande complexité et qu'ils n'ont été obtenus qu'après des siècles, voire des millénaires, de réflexion, de travail et de controverses d'une grande intensité. Ceci illustre parfaitement le besoin, pour tout scientifique, chercheur ou ingénieur, de maîtriser les bases de la réflexion sur la nature de la science et donc de comprendre la portée et le domaine de validité de son savoir, au-delà de la simple érudition dans son domaine de spécialisation.

1.2 La science ne se résume pas à des équations !

Pour illustrer notre propos, prenons un exemple : la mécanique de Newton. Elle est apprise dès le lycée, et semble a priori très simple : elle est « ancienne » (donc,

intuitivement, on serait tenté de croire qu'elle ne traite que de choses élémentaires en regard de théories plus récentes comme la théorie de la relativité et la mécanique quantique), elle s'illustre avec des exemples pratiques nécessitant des technologies élémentaires lors de travaux pratiques, et elle fait appel à des concepts qui semblent être facilement appréhendables : force, accélération, masse, mouvement, repos. De plus, elle s'exprime au moyen de trois lois d'expression mathématique simple, connues sous le nom de lois de Newton (valables dans un référentiel galiléen) :

1. **Première loi de Newton** (encore appelée loi ou principe d'inertie)

$$\sum \vec{F} = \vec{0} \iff \text{Mouvement Rectiligne Uniforme ou Repos}$$

2. **Deuxième loi de Newton** (encore appelée loi ou principe fondamental de la dynamique)

$$\sum \vec{F} = m\vec{\gamma}$$

3. **Troisième loi de Newton** (encore appelée loi ou principe de l'action et de la réaction)

$$\vec{F}_{A \rightarrow B} = -\vec{F}_{B \rightarrow A}$$

Mais cela est-il aussi simple ? Non.

Tout d'abord, sommes-nous capables de définir exactement ce que sont les entités (masse, force, ...) qui apparaissent dans ces équations ? Comment, par quel procédé théorique ou pratique, peut-on **définir** puis **évaluer** ces quantités ?

Ensuite, les relations rappelées ci-dessus sont appelées des lois. Mais qu'est-ce qu'une loi ? Comment sont-elles obtenues, et quels sont leur place, leur rôle et leur statut au sein des sciences ?

Plus problématique encore : ces lois sont elles vraies ? Comment les vérifier ? Et que signifie pour une loi scientifique être vraie ou fausse ?

Généralisons encore le champ de la réflexion. On parle de mécanique newtonienne. L'adjectif « newtonienne » indique qu'il y a plusieurs mécaniques possibles : effectivement, il existe une mécanique hamiltonienne, une mécanique relativiste associée au nom d'Einstein, une mécanique quantique, ... Puisqu'elles sont différentes de la mécanique de Newton, se pose la question intuitive et naïve de savoir si la mécanique de Newton est fausse ? Y aurait-il des mécaniques plus vraies (ou fausses) que d'autres ? Et d'abord, cette question a-t-elle un sens ?

La mécanique est la première science de la nature à s'être constituée historiquement, en Grèce, il y a environ 2500 ans. Comment, en quoi, pourquoi, sous quelles impulsions a-t-elle évolué pour former la science que nous connaissons aujourd'hui ? Cette évolution a-t-elle été continue, ou existe-t-il des « révolutions scientifiques » qui marquent des tournants décisifs. Peut-on dire qu'il y a eu progrès de la mécanique ?

L'exemple de la mécanique permet de poser, de manière plus générale, la question de la genèse et de l'évolution des sciences.

Mais, bien évidemment, il se pose une question fondamentale : qu'est-ce qu'une science ?

L'ensemble de ces questions relève de l'**épistémologie**, qui fait l'objet de ce cours. Pour chacune des grandes questions qui viennent d'être soulevées, nous allons nous efforcer de dresser un panorama synthétique des différentes réponses apportées et de présenter les points de vue spécifiques de grands courants modernes de l'épistémologie.

1.3 Qu'est-ce que l'épistémologie ?

1.3.1 Présentation générale

Le terme **épistémologie** (ou plus précisément le terme anglais **epistemology**) a été forgé par le métaphysicien James Frederick Ferrier (1808-1864) pour désigner une théorie de la connaissance. Ce néologisme est construit à partir des termes grecs **épistémé** (connaissance théorique, savoir) et **logos** (discours rationnel, langage, jugement). Il apparaît pour la première fois dans un ouvrage de langue française en 1901, dans la traduction d'un ouvrage de Bertrand Russell¹. Il est ensuite popularisé et est aujourd'hui très répandu.

Le terme épistémologie est employé pour désigner deux choses différentes :

1. Une **théorie générale de la connaissance humaine**, scientifique et non scientifique. Dans cette acception, qui est celle la plus courante pour le terme anglais *epistemology*, l'épistémologie peut être considérée comme une branche de la philosophie qui traite de la nature, de la valeur et des limites de la connaissance humaine.
2. Une **théorie de la connaissance scientifique**, ou encore comme la philosophie des sciences. C'est cette seconde définition, plus restreinte et généralement retenue par les auteurs de langue française, qui fait l'objet de ce cours. La théorie générale de la connaissance est désignée par le terme (rare) **gnoséologie**².

La définition correspondante donnée par Léna Soler³ est : « *L'épistémologie vise fondamentalement à caractériser les sciences existantes, en vue de juger de leur valeur et notamment de décider si elles peuvent prétendre se rapprocher de l'idéal d'une connaissance certaine et authentiquement justifiée.*

Elle s'emploie pour atteindre cet objectif, à décrire la manière dont procède telle ou telle discipline dite scientifique pour élaborer et tester ses théories, à spécifier la

¹ Bertrand Russell (1872-1970) philosophe et logicien

² Gnoséologie est formé à partir des termes grecs **gnôsis** (terme théologique signifiant « connaissance », lui-même dérivé de **gignôskein** « connaître ») et **logos**.

³ « Introduction à l'épistémologie », L. Soler, éditions Ellipses, 2000

physionomie de ces théories elles-mêmes, enfin à estimer la valeur logique et cognitive de telles théories. »

Les trois questions fondatrices de l'épistémologie sont donc :

Questions fondatrices de l'épistémologie

1. Qu'est-ce que la science (ou les sciences) ? Qu'est-ce qui distingue ce type de savoir des autres ? Comment la définir ?
2. Comment la science (ou une science) s'est-elle constituée ? Quels ont été les facteurs (technologiques, mathématiques, sociologiques, philosophiques, religieux, ...) qui ont influé sur son évolution ? Quelles méthodes de travail et de réflexion ont été employées pour la construire ?
3. Comment juger de sa validité ou de sa valeur ? Que veut dire qu'une théorie scientifique est vraie ? Comment vérifier la validité d'une théorie scientifique ?

Cette définition met en évidence deux points importants. Tout d'abord, l'épistémologie est un discours sur la science, et donc présuppose la science. En ce sens, elle vient donc après cette dernière. Ensuite, l'épistémologie est un discours critique sur la connaissance scientifique, son élaboration et son évolution. Elle doit donc définir son objet (qu'est-ce que la science ? voir le chapitre 2) et ses méthodes.

L'identification d'un savoir authentiquement scientifique et d'un savoir non-scientifique, souvent appelé savoir commun ou sens commun, suppose l'existence d'une démarcation nette entre ces deux types de connaissance. Ce problème est évoqué au chapitre 2. La restriction du champ d'analyse de l'épistémologie au seul savoir scientifique permet de la distinguer de la **philosophie de la connaissance**.

On peut également (ce que font bon nombre d'épistémologues contemporains) la distinguer de la **philosophie des sciences** proprement dit, cette dernière subordonnant la réflexion sur la science à des préoccupations de philosophie générale.

Si le terme épistémologie est de création récente, la philosophie des sciences remonte à l'antiquité, puisque la science (même si le contenu exact de ce mot a considérablement évolué depuis) est prise comme objet de réflexion par Platon (428-346 av. J.-C.) et Aristote (384-322 av. J.-C.), et que cette réflexion est présente chez de nombreux philosophes et scientifiques au cours des vingt-cinq derniers siècles. La philosophie des sciences proprement dite est fondée par le philosophe allemand Emmanuel Kant (1724-1804), qui le premier distingue la science comme un objet autonome soumis à l'analyse philosophique.

Toutefois, de par sa nature, l'épistémologie appartient au champ de la philosophie, et ne peut pas être considérée comme complètement étrangère à la philosophie des sciences. Elle en fait même partie. La question d'une distinction claire entre philosophie des sciences et épistémologie est encore aujourd'hui une question ouverte. Il est notamment utile de remarquer que, comme tous les travaux philosophiques, les textes traitant d'épistémologie sont profondément marqués par les opinions philosophiques de leurs auteurs, alors que la science a pour idéal de réduire autant que faire se peut toute part de subjectivité dans son contenu. L'épistémologie n'est donc pas une « science de la science ».

Il est important de retenir que l'épistémologie a deux aspects : un **aspect normatif** qui vise à définir ce qu'est la science et donc à délimiter le champ de la science (qu'est-ce que la science ? une discipline est-elle scientifique ?) et un **aspect descriptif** (quelles sont la structure et la dynamique interne d'une discipline scientifique ?).

Pour analyser son objet, la connaissance scientifique, l'épistémologue peut procéder selon deux méthodes :

- La **méthode synchronique**, qui consiste à considérer les disciplines scientifiques à un instant donné, indépendamment de leur développement historique. On procède alors à une **analyse directe** de ces disciplines, de leurs contenus, de leurs méthodes.
- La **méthode diachronique**, qui consiste à analyser les disciplines scientifiques dans la perspective de leur genèse, de leur développement et de leur maturation. On procède alors à une **analyse génétique** (par opposition à l'analyse directe). Cette méthode permet notamment de poser la question du progrès scientifique.

L'épistémologie recourt à l'**histoire des sciences** (par construction lorsque la méthode diachronique est employée) pour y puiser les « faits bruts » nécessaires à son analyse. L'étude de l'histoire des disciplines scientifiques, des modèles et des théories auxquels elles font appel permet d'une part de comprendre les disciplines scientifiques dans leur état de développement actuel et d'analyser la nature de l'évolution scientifique (y a-t-il progrès scientifique ? Quel est le processus d'évolution des sciences ?) mais également de faire la part entre ce qui est imposé par les faits expérimentaux (donc *a priori*) et ce qui est rajouté par les scientifiques (donc *a posteriori*). Il existe une longue tradition d'étude de l'histoire des sciences et d'histoire philosophique des sciences⁴. Comme nous le verrons par la suite, la méthode diachronique occupe une place centrale dans le développement des théories épistémologiques modernes comme le falsificationisme de Popper, la théorie des révolutions scientifiques de Kuhn ou encore la théorie des programmes de recherche de Lakatos.

L'épistémologie se tourne également vers la **sociologie des sciences** pour y puiser des renseignements sur les interactions entre la science (ou une science) et la société, et ceci afin de répondre aux deux questions suivantes :

1. En quoi une science (ou la science) a-t-elle influé sur l'organisation politique, économique et sociale d'une société (ou d'un sous-groupe identifié), ou encore sur l'évolution de la pensée philosophique et religieuse, de la littérature, ...
2. Quelles ont été ou sont les contraintes sociologiques exercées par la société (ou un sous-groupe) sur le développement d'une science ou d'une théorie scientifique ? Les opinions philosophiques, religieuses et politiques des chercheurs ont-elles une influence sur le développement et le contenu des théories scientifiques, et si oui laquelle ?

⁴ Pour une introduction, le lecteur intéressé pourra se référer à « La philosophie des sciences », Dominique Lecourt, Que sais-je ? No 3624, PUF, 2001

1.3.2 Une vue alternative : l'épistémologie comme science empirique

Il a été dit plus haut que l'épistémologie, telle qu'elle est conçue par la majorité des épistémologues contemporains appartient au domaine de la philosophie et n'est pas une science de la science. Toutefois, certains philosophes ont proposé de voir l'épistémologie comme une science empirique (au même titre que la physique, la mécanique, la chimie, ...) dont l'objet d'étude serait les théories scientifiques. Ce point de vue est illustré par la démarche du philosophe et logicien américain Willard Van Orman Quine (1908-2000) qui défendit cette thèse dans un article daté de 1951 et introduisit le terme **naturalized epistemology** (traduit en français par **épistémologie naturalisée** ou **naturelle**) en 1969. Dans cette nouvelle perspective, le but de l'épistémologie est de décrire la structure interne des sciences et leur évolution (ce qui correspond à l'aspect descriptif évoqué plus haut) et d'approfondir la signification de leurs énoncés, et celle-ci ne peut prétendre fonder la connaissance scientifique. Une conséquence est donc qu'il faut renoncer à tout caractère normatif absolu. En effet, l'épistémologie étant conçue comme une science parmi les autres, il y aurait un cercle vicieux si celle-ci définissait ce qu'est la science. Cet abandon de la recherche d'une justification absolue des connaissances place la philosophie de Quine en opposition avec la **démarche fondationnaliste**, qui vise à élaborer une théorie générale de la vérité et des moyens pour l'atteindre. Etant une science empirique, l'épistémologie doit, selon Quine, utiliser à la fois les résultats et les méthodes des autres sciences. Etant une science empirique, elle doit utiliser la méthode expérimentale pour confronter ses théories aux faits. Par ailleurs, il est important de retenir que selon Quine il n'y a pas d'opposition entre science et philosophie, et que l'épistémologie naturalisée appartient également au domaine de cette dernière : « *la philosophie, [...] comme effort en vue de se faire une idée plus claire des choses, ne doit pas, pour ce qui est l'essentiel de ses objectifs et de sa méthode, être distinguée de la science.* »⁵

Cette vision naturalisée de l'épistémologie repose sur l'idée que les sciences sont des objets qui peuvent être soumis à l'analyse scientifique. Un pas supplémentaire est franchi par les défenseurs d'un genre particulier d'épistémologie naturalisée connue sous le nom d'**épistémologie évolutionnaire**, qui fondent leur analyse du progrès scientifique sur la théorie biologique de l'évolution initiée par le naturaliste anglais Charles Darwin (1809-1882) dans son livre *L'origine des espèces* paru et épuisé le 24 novembre 1859.



Figure 1: W.O. Quine

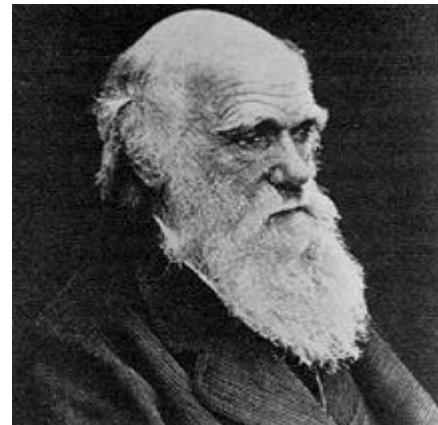


Figure 2: C. Darwin

⁵ W. Quine, « Le mot et la chose », 1960. Cité par L. Soler, Ibid.

Dans sa version dite **littérale**, cette école de pensée interprète les capacités intellectuelles de l'homme et ses produits (dont la science) comme des éléments participant à la sélection naturelle des espèces vivantes. La science, permise par le développement des facultés intellectuelles de l'homo sapiens, est alors analysée comme une retombée de la concurrence entre les espèces qui assure la domination et la survie du genre humain, au même titre que la bipédie ou l'existence de mains avec des pouces en opposition.

Dans sa version **analogique**, l'épistémologie évolutionnaire utilise le paradigme de la sélection naturelle (basé sur la notion de compétition, de mutation aléatoire transmise aux descendants et de survie du plus apte) pour expliquer l'évolution des théories scientifiques : les théories les plus adaptées sont sélectionnées après un processus de variation des contenus des théories et de concurrence. Il est important de noter que cette utilisation analogique de la théorie de la sélection naturelle n'implique pas l'acceptation que cette théorie soit vraie dans le cas de la biologie : il ne s'agit ici que d'un schéma explicatif.

1.4 Sous quelles conditions sait-on quelque chose ?

Avant de discuter plus avant les différentes questions qui relèvent de l'épistémologie, donc de la connaissance scientifique, il convient de s'intéresser tout d'abord au problème de la **définition de la connaissance** : quand peut-on dire que l'on sait quelque chose ? Ici, savoir est opposé à croire, le savoir n'étant qu'un type de croyance, et le savoir scientifique (donc la science) un cas particulier de savoir.

1.4.1 Discussion liminaire : qu'est-ce qu'une définition ?

Pour discuter de manière adéquate de la définition de la connaissance, il convient de préciser ce que l'on appelle une définition⁶. Ainsi, on pourra s'assurer de la validité des définitions de la connaissance proposées par différents auteurs et, le cas échéant, détecter leurs faiblesses.

On appelle **définition substantielle** ou **analyse** une proposition qui met en relation une expression qui est définie (le *definiendum*) et une expression qui définit (appelée *definiens*) et qui remplit les trois conditions suivantes :

Critères de validité d'une définition substantielle

1. **propriété d'adéquation extensionnelle** : le *definiens* doit s'appliquer à tout ce que désigne le *definiendum*. Cela est traduit par l'emploi de l'expression « si et seulement si », familière en mathématiques.
2. **propriété d'adéquation intensionnelle ou conceptuelle** : le *definiens* doit donner des renseignements sur la nature de la chose définie.
3. **propriété de non-circularité** : le *definiens* ne doit pas faire appel à lui-même, directement ou indirectement.

Donnons quelques exemples.

⁶ La question de la définition sera reprise au début du chapitre 3, où d'autres types de définitions seront présentés.

La proposition « n est un nombre premier si et seulement si est un entier naturel qui a exactement deux diviseurs, 1 et lui-même » vérifie bien la propriété d'adéquation extensionnelle. Par contre, « X est un homme si et seulement si X appartient au genre humain » ne nous apprend rien sur ce qu'est un homme : cette proposition n'est pas conceptuellement adéquate, et est donc une **pseudo-définition** et non pas une définition. Enfin « X est un homme si et seulement si X est un homme » est circulaire.

1.4.2 La connaissance selon Platon

Une première définition de la connaissance est fournie par Platon⁷ :

Définition de la connaissance chez Platon

La connaissance est une *croyance vraie et justifiée*.

Cette définition implique deux conditions : pour accéder au statut de connaissance, une croyance doit non seulement être vraie (c'est-à-dire correspondre à une réalité), mais le sujet doit également être fonder à croire (c'est-à-dire que l'on ne peut pas savoir « par hasard » ou par erreur).

Avant d'aller plus loin, notons que cette définition de la connaissance soulève de nombreuses autres questions, qui pour certaines seront reprises dans les chapitres suivants (le plus souvent en considérant le cadre restreint de la connaissance scientifique):

- Qu'est-ce qu'une connaissance vraie ? Comment peut-on s'assurer de la véracité de quelque chose ?
- Quel accès pouvons-nous avoir à la réalité ? Celle-ci existe-t-elle, et comment vérité et réalité sont-elles liées ?
- Qu'est-ce qu'une connaissance justifiée ? Que peut-on dire de la théorie de la justification ?

La question de la définition de la connaissance n'a reçu que peu d'attention entre Platon et le XX^e siècle. Tout au long des 2500 ans qui séparent ces deux périodes, plusieurs philosophes et théologiens ont donné des définitions de la connaissance, mais ce n'est que récemment (à partir de 1950) qu'un effort de d'analyse systématique et de remise en cause de la définition platonicienne a été réalisé. Le chapitre suivant présente quelques contributions marquantes dans ce domaine.

1.4.3 Le problème de Gettier

Avant de présenter les objections formulées par Edmund Gettier⁸ en 1963, reformulons tout d'abord la définition donnée par Platon de manière plus formelle comme suit :

⁷ Théétète, 201, 207-208 ; Ménon, 98.

⁸ L'article de Gettier est disponible en français dans « Philosophie de la connaissance – croyance, connaissance, justification », édité par J. Dutant et P. Engel, Vrin, 2005

Ecriture formelle de la définition de Platon

X sait que P si et seulement si :

1. X croit que P
2. P est vrai
3. X est justifié à croire que P

Dans son article fondateur, Gettier montre que la définition donnée par Platon n'est pas satisfaisante, car il existe des croyances vraies et justifiées qui ne sont pas des connaissances.

Prenons l'exemple suivant, qui traite d'une croyance vraie injustifiée. X met ses clés dans sa poche de son manteau. Si on lui demande où elles sont, X sait qu'elles sont dans sa poche. Cette proposition est vraie, et X est justifié à la croire puisqu'il les y a mises lui-même. Mais imaginons maintenant que, sans s'en apercevoir, X les ait perdues et qu'un ami les ait remises dans la poche du manteau de X, toujours sans que X s'en aperçoive. Alors, X n'est plus justifié à croire qu'elles sont dans sa poche puisqu'elles n'y sont que par hasard : il les a perdues, les clés ont été remises par hasard et sans que X le sache. Selon la définition de Platon, on ne peut plus dire que X sait où sont ses clés !

Passons maintenant à un cas de croyance vraie et justifiée qui n'est pas une connaissance. Cet exemple est donné par Gettier : « *Supposez que Smith et Jones soient tous les deux candidats à un certain emploi. Et supposez que Smith a de bonnes raisons en faveur de la proposition conjonctive suivante :*

(d) Jones est celui qui sera embauché, et Jones a dix pièces dans sa poche.

Les raisons que Smith a de croire en (d) peuvent être que le directeur de l'entreprise lui a assuré que Jones serait finalement embauché, et qu'il y a dix minutes Smith a compté les pièces que Jones a dans sa poche. La proposition (d) implique que :

(e) celui qui sera embauché a dix pièces dans sa poche.

Supposons que Smith voit cette implication (e) par (d), et accepte (e) sur la base de (d) en faveur de laquelle il a de très bonnes raisons. Dans ce cas, il est clair que Smith est justifié à croire que (e) est vraie.

Mais imaginons qu'en outre, à l'insu de Smith, c'est lui, et non Jones, qui obtiendra l'emploi. Et qu'aussi, toujours à l'insu de Smith, il a lui-même dix pièces dans sa poche. La proposition (e) est donc vraie, bien que la proposition (d), de laquelle Smith a inféré (e), soit fausse. Dans notre exemple, tout ceci est vrai :

1. *que (e) est vraie*
2. *que Smith croit que (e) est vraie*
3. *que Smith est justifié à croire que (e) est vraie*

Mais il est tout aussi clair que Smith ne sait pas que (e) est vraie ; car (e) est vraie en vertu du nombre de pièces qu'il y a dans la poche de Smith, alors que Smith ne sait pas combien de pièces il y a dans la poche de Smith, et fonde sa croyance en (e) sur le décompte des pièces dans la poche de Jones, qu'il croit à tort être celui qui sera embauché. »

La définition de Platon n'est pas extensionnellement exacte. En effet, elle est de la forme « A si et seulement si B », et comprend donc deux affirmations

1. A si B : B est une condition suffisante de A
2. A seulement si B : B est une condition nécessaire de A

Pour montrer qu'une telle proposition est fautive, il suffit de trouver un cas où l'on a B sans avoir A (violation de 1) ou un cas où l'on a A sans B (violation de 2). Le contre-exemple donné plus haut viole la relation « condition suffisante ». Cela montre que le *definiens* donné par Platon n'est pas adéquat.

Depuis 1963, de très nombreux travaux ont visé à donner une définition « parfaite » de la connaissance qui ne souffre pas de contre-exemple, et qui respecte les trois critères d'adéquation donnés plus haut. Deux points particulièrement durs sont d'éviter la circularité et d'arriver à une notion satisfaisante de justification. De très nombreux contre-exemples de plus en plus sophistiqués ont été produits en regard de définitions de plus en plus complexes, donnant naissance à ce qui est parfois appelé « gettierologie » (avec une connotation péjorative). Une discussion approfondie de ces exemples dépasse le cadre de ce cours. La section suivante a pour but de dresser un bref panorama des différentes réponses apportées au problème de la justification.

1.4.4 Réponses au problème de la justification : internalisme, externalisme

Le problème de Gettier met en lumière un problème fondamental de la théorie platonicienne de la connaissance : celui de la justification. En effet, selon cette théorie, le sujet ne peut connaître que si sa croyance est justifiée. Dans quels cas peut-on considérer que la croyance est justifiée ? On distingue ici deux familles de théories philosophiques.

La première famille est celle de l'**internalisme**, pour laquelle le sujet a, au moins en théorie, accès aux justifications de ses croyances. Se pose alors la question de ce qui fonde la justification : quelle est la justification de la justification ? Et la justification de la justification de la justification ? Cette question peut être répétée à l'infini, donnant naissance à un problème de fermeture de cette hiérarchie infinie de questions emboîtées.

On rencontre ici un problème connu sous le nom **problème de la régression épistémique** dans la théorie de la justification qui conduit au **trilemme d'Agrippa**⁹. Parmi les cinq « modes de suspension du jugement » formulés par Agrippa¹⁰, les trois premiers forment le trilemme suivant :

⁹ Ce trilemme est aussi connu sous le nom de « trilemme de Fries » d'après J.F. Fries ou encore comme le « trilemme de Münchhausen », du nom du baron légendaire qui disait s'être arraché à un marécage en se tirant par les cheveux.

¹⁰ On suppose qu'Agrippa vivait au premier siècle avant J.C. Il n'existe pas de trace directe de cet auteur. C'est Diogène Laërce (dans son livre *Vies et Doctrines des philosophes illustres*) qui lui attribue cet argument. Mais on trouve des traces antérieures, par exemple chez Aristote (*Analytiques postérieures*). La meilleure présentation antique est fournie par Sextus Empiricus au deuxième siècle après J.C., dans ses *Esquisses Pyrrhoniennes*.

Trilemme d'Agrippa

1. **L'arrêt dogmatique** : on décide d'arrêter la suite à un niveau donné (pour une raison qui nous appartient).
2. **La régression à l'infini** : on poursuit l'enchaînement des propositions sans arrêt.
3. **Le cercle auto-référentiel** : une proposition fait appel à une autre déjà donnée. On retrouve ici le problème de la circularité.

Les deux autres solutions proposées par Agrippa sont le *désaccord* (l'opinion inverse est défendable) et la *relativité* (une chose peut paraître différente à une autre personne).

On distingue deux types de choix, qui fondent des écoles internalistes différentes :

- Le **cohérentisme** (basé sur la troisième réponse d'Agrippa) consiste à supposer que les connaissances se fondent mutuellement les unes les autres, et que l'ensemble des connaissances forme un bloc cohérent auto-justifié. La définition cohérentiste de la connaissance possède donc un certain degré de circularité.
- Pour les partisans du **fondationnalisme** (basé sur la première réponse d'Agrippa), les connaissances forment un édifice dont le socle est composé de connaissances fondées sur des principes ou des jugements qui ne sont eux-mêmes fondés sur rien d'autre. Le problème est alors bien sûr de savoir quelles sont ces connaissances fondamentales qui n'ont pas besoin d'être justifiées.

La seconde famille est celle de l'**externalisme**. Selon cette théorie, il n'est pas nécessaire que le sujet ait lui-même accès à la justification de sa croyance pour qu'il y est connaissance : il suffit qu'un observateur extérieur possède cette justification. Les théories externalistes sont souvent formulées comme des théories causales¹¹ de la connaissance qui s'expriment comme suit :

Définition externaliste causale de la connaissance

X sait que P si X est relativement à P dans une relation causale appropriée.

Sans entrer dans le détail (cela sera fait plus loin dans le cadre de l'analyse de la connaissance scientifique), disons qu'une relation causale appropriée pour un être humain est par exemple une perception sensorielle directe (vue, ouïe, toucher,...). Le problème qui apparaît alors est celui de la fiabilité de cette relation causale.

¹¹ La notion de causalité et les théories afférentes sont discutées au chapitre 10.

1.4.5 La connaissance selon Nozick et Dretske

On s'intéresse ici à la définition de la connaissance trouvée indépendamment par F. Dretske en 1970 et par Robert Nozick¹² en 1981. Cette définition a ceci de nouveau qu'elle repose sur l'idée que

Définition de la connaissance de Dretske et Nozick

Connaître, c'est suivre la vérité à la trace

Ce qui se traduit de manière formelle comme suit :

Définition de la connaissance de Dretske et Nozick

X sait que P si et seulement si :

1. P est vrai
2. X croit P
3. si P était faux, X ne croirait pas P
4. si P était vrai, X croirait P

Cette définition appartient à la famille des analyses dites **alternatives pertinentes** : elle contient des propositions de la forme « si ... alors ... » (**propositions contrefactuelles**) qui permettent d'envisager des situations différentes de la réalité (ceci est traduit par la présence du mot si). Selon ces théories, connaître revient à éliminer les alternatives pertinentes à la croyance de X. Un point important, qui fait la force mais aussi la faiblesse de cette définition de la connaissance, est qu'elle ne requiert pas d'éliminer toutes les alternatives : il suffit de conserver toutes les situations qui sont au moins aussi semblables à la « vraie » situation actuelle que ne l'est la situation la plus semblable dans laquelle P est faux.

Par exemple, si X voit une voiture à l'arrêt, il saura que c'est une voiture (même s'il n'en connaît pas le modèle, ou même si il s'agit d'une maquette ayant toutes les apparences extérieures d'une vraie voiture).

La faiblesse de cette approche réside dans le fait qu'il faut pouvoir distinguer les alternatives pertinentes de celles qui ne le sont pas.

1.4.6 Le défi du scepticisme

La connaissance existe-t-elle ? Comme on vient de la voir, aucune théorie n'est pleinement satisfaisante et n'emporte une adhésion complète des philosophes de la connaissance. Aussi, cette impossibilité d'arriver à une définition de la connaissance ouvre la porte au **scepticisme**. Pour le sceptique, il est impossible d'arriver à une croyance vraie et véritablement justifiée, car rien ne prouvera jamais de manière irréfutable que nos perceptions sont justes. On rencontre ici les deux exemples

¹² L'article de Nozick est disponible en français dans « Philosophie de la connaissance – croyance, connaissance, justification », édité par J. Dutant et P. Engel, Vrin, 2005

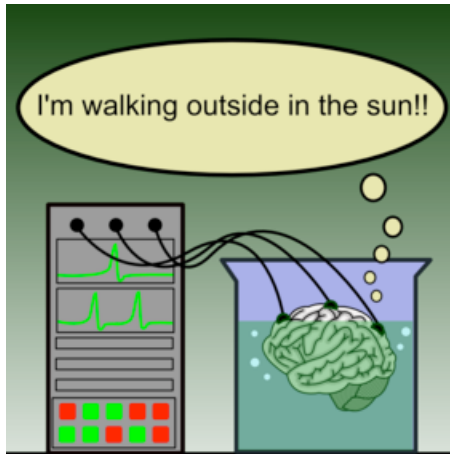


Figure 3 Illustration de "cerveau dans la cuve"

fameux du « mauvais génie »¹³ et du « cerveau dans la cuve »¹⁴, qui tous les deux remettent en question la possibilité de percevoir la réalité. Qu'est-ce qui me prouve que l'ensemble de nos perceptions ne sont pas le fruit des manipulations d'un « mauvais génie » qui prendrait plaisir à nous voiler la réalité ? La variante du « cerveau dans la cuve » est la suivante : Qu'est-ce qui me prouve que mon corps et le monde existent, et que je ne suis pas en réalité un cerveau flottant dans une cuve en train de rêver ?

Les théories basées sur la notion de croyance justifiée ne permettent pas de contrer l'objection sceptique, puisque celle-ci interdit toute justification (du moins toute justification se basant sur un rapport avec la réalité). La forme logique de l'objection sceptique est la suivante :

Forme logique de l'argument sceptique

1. X ne sait pas que non-P
2. si X ne sait pas que non-P, alors X ne sait pas que Q
3. donc, X ne sait pas que Q

Illustrons cela avec l'exemple du « cerveau dans la cuve ». La formulation sceptique est de la forme :

1. je ne sais que je ne suis pas un cerveau dans une cuve
2. si je ne sais pas que je ne suis pas un cerveau dans une cuve, alors je ne sais pas que j'ai deux mains
3. donc, je ne sais pas que j'ai deux mains

A l'opposé, la formulation contraire au scepticisme est :

1. Je sais que j'ai deux mains
2. Si je sais que j'ai deux mains, alors je sais que je ne suis pas un cerveau dans une cuve
3. Donc je sais que je ne suis pas un cerveau dans une cuve

¹³ L'exemple du mauvais génie est dû à René Descartes (1596-1650) dans sa *Première méditation* : « Je supposerai donc qu'il y a, non point un vrai Dieu, qui est la souveraine source de vérité, mais un certain mauvais génie, non moins rusé et trompeur que puissant qui a employé toute son industrie à me tromper. Je penserai que le ciel, l'air, la terre, les couleurs, les figures, les sons et toutes les choses extérieures que nous voyons, ne sont que des illusions et tromperies, dont il se sert pour surprendre ma crédulité. Je me considérerai moi-même comme n'ayant point de mains, point d'yeux, point de chair, point de sang, comme n'ayant aucun sens, mais croyant faussement avoir toutes ces choses. Je demeurerai obstinément attaché à cette pensée ; et si, par ce moyen, il n'est pas en mon pouvoir de parvenir à la connaissance d'aucune vérité, à tout le moins il est en ma puissance de suspendre mon jugement. »

¹⁴ Cet exemple a été introduit par le philosophe américain Hilary Putnam en 1981, dans son livre « Raison, Vérité et Histoire »

1.4.7 Les réponses au scepticisme

Plusieurs tentatives pour redéfinir la connaissance et contrer l'argument de l'ignorance sceptique ont été formulées. Nous allons maintenant en voir quelques unes.

Tout d'abord apparaît ici un autre intérêt de la définition de la connaissance de Nozick. Les deux formes données ci-dessus du problème du cerveau dans la cuve (la sceptique et son opposée) reposent toutes les deux sur le **principe de clôture épistémique** qui s'énonce comme suit :

Principe de clôture épistémique

Si X sait que P, et si X sait que P implique Q, alors X sait que Q.

Selon Nozick, ce principe n'est pas valable, car il existe des cas où l'on peut suivre à la trace la vérité de P et celle de « P implique Q » sans pour autant suivre à la trace la vérité de Q. Dans l'exemple précédent, cela se traduit par le fait que si j'avais été un cerveau dans une cuve, j'aurais toujours cru avoir deux mains (par construction de l'exemple : le cerveau dans la cuve est persuadé d'avoir deux mains). Ceci contredit le point 3 de la définition de la connaissance donnée par Nozick, qui apporte donc une forme de réponse à l'objection sceptique.

Une autre réponse est celle de **l'externalisme du mental**. Selon cette théorie, la connaissance résulte d'un état mental et, comme tout état mental, résulte (en partie au moins) de notre interaction avec le monde qui nous environne. Cette nécessaire action du monde extérieur dans la définition de la connaissance (qui n'est plus conçue comme une croyance vraie et justifiée) interdit l'existence du « mauvais génie » ou de l'expérience du cerveau dans la cuve. Plutôt qu'une définition, cette approche conduit à un schéma de la connaissance qui peut être énoncé comme suit :

Schéma de la connaissance suivant l'externalisme du mental

Si X sait que P, alors X ne pourrait pas facilement se tromper au sujet de P.

La dernière réponse au défi du scepticisme discutée ici est celle du **contextualisme**, défendue par exemple par Ludwig Wittgenstein (1889-1951). Pour les contextualistes, c'est la tentative de trouver une définition universelle de la connaissance qui est vouée à l'échec. Mais il est possible de s'accorder sur des définitions « au cas par cas » de ce qu'est une connaissance dans différents domaines : les mathématiques, la physique, en histoire, etc. La connaissance est donc définie dans un contexte particulier, et sa définition peut donc varier avec celui-ci.

2 Qu'est-ce que la science ?

2.1 Comment définir la science ?

2.1.1 Définition(s) de la science et critères de scientificité

Deux démarches sont possibles pour définir ce qu'est une science.

La première est une **démarche normative**, qui consiste à édicter **a priori** une norme de **scientificité**, c'est-à-dire de donner les critères qui permettent de statuer sur le caractère scientifique d'une discipline. Cette approche tend à concevoir les différentes disciplines scientifiques comme des cas particuliers d'une Science idéale, qui n'est jamais incarnée dans sa totalité.

La seconde **démarche** est **descriptive** : elle consiste à analyser les différentes disciplines reconnues comme scientifiques, et à en dégager **a posteriori** les points communs, qui seront ensuite pris comme des critères de scientificité.

Le mot **science** apparaît en 1080 dans la première des chansons de geste françaises, *La chanson de Roland*. Il est dérivé du latin classique **scientia** (connaissance, et plus particulièrement connaissance scientifique, rationnelle), qui prend très tôt le même sens que le terme grec **épistémé**. *Scientia* vient de **sciens**, **scientis** qui signifie « qui sait », « instruit », « habile ». Il faut noter que l'emploi de l'appellation « **scientifiques** » pour désigner ceux qui pratiquent la science est beaucoup plus récent. Le terme « **scientist** » a été introduit dans la langue anglaise par William Whewell vers la moitié du XIX^e siècle. En français, « **scientifiques** » ne sera couramment employé qu'à partir du XX^e siècle, remplaçant « **savants** ». Comme le note Ken Adler¹⁵, c'est le révolutionnaire Jean-Paul Marat, qui, en 1792, fut le premier à appliquer l'étiquette « **scientifiques** » aux savants de l'époque, lorsqu'il raillait le projet de l'Académie des sciences de mesurer la longueur de méridien terrestre pour définir un mètre étalon dans le cadre d'un programme d'unification universelle des poids et des mesures (voir le chapitre 4.2.3).

Regardons maintenant les définitions actuelles du mot « science » trouvées dans quelques dictionnaires couramment utilisés :

- **Définition 1 (Petit Robert)** : « *ensemble de connaissances, d'études d'une valeur universelle, caractérisées par un objet et une méthode déterminés, et fondées sur des relations objectives vérifiables.* »
- **Définition 2 (Larousse)** : « *ensemble cohérent de connaissances relatives à une certaine catégorie de faits, d'objets ou de phénomènes.* »

auxquelles on peut ajouter d'autres définitions, comme par exemple pour les sciences empiriques (ou sciences de la nature)

¹⁵ « Mesurer le monde », K. Adler, coll. Cahamps Histoire, Flammarion, 2005

- **Définition 3¹⁶** : « *La science est une connaissance objective qui établit entre les phénomènes des rapports universels et nécessaires autorisant la prévision de résultats (effets) dont on est capable de maîtriser expérimentalement ou de dégager par l'observation la cause.* »

Ces définitions illustrent le contenu attribué par le sens commun au terme science. Au terme de science sont généralement associés par la plupart des gens des caractères positifs, valorisants, l'adjectif « scientifique » étant souvent employé dans le sens de « vrai », « rigoureux », « sûr ». Mais une analyse plus approfondie est nécessaire pour dégager ce qu'est la science (ou une science), donc d'identifier les critères de scientificité qui permettront de statuer sur la nature scientifique d'une théorie ou d'un ensemble de connaissances.

Les définitions présentées dans ce qui précède mettent en lumière plusieurs points clés.

Tout d'abord, l'objet dont traite une science doit être clairement identifié. Se pose ici le problème de la définition des frontières du domaine couvert par une discipline scientifique, et de son possible chevauchement avec d'autres disciplines. Ceci est de plus en plus vrai, avec l'émergence de sciences dites multidisciplinaires. Par exemple, la chimie, dans certaines de ses branches, comme la chimie quantique, possède des recouvrements avec la physique. De même, la mécanique admet de nombreux recouvrements avec la physique la chimie, mais aussi ... la psychologie (la psycho-acoustique, par exemple) et la biologie (la biomécanique). Enfin, il existe des sciences plus « exotiques » comme la bioinformatique, qui est à la frontière entre la biologie et l'informatique et l'informatique quantique, qui allie informatique et mécanique quantique, ou encore l'éconophysique, qui se situe à la frontière de la physique et de l'économie. Si l'objet principal dont traite une science est souvent relativement simple à expliquer, une définition exacte et exhaustive est souvent hors d'atteinte et les frontières d'une discipline demeurent floues.

Ensuite, une science doit apporter des connaissances sur son objet, c'est-à-dire avoir un contenu. Pour accéder au statut de connaissance et non de simple croyance, son contenu doit être justifiable, c'est-à-dire vérifiable ou validable. Cette étape de validation, centrale dans la définition de la science dans l'acceptation moderne du terme, implique que la vérification doit pouvoir être faite par toute personne le désirant (si elle possède le bagage théorique et technique requis pour mener à bien cette opération). C'est en cela que le **savoir scientifique** est dit **objectif** : il est (idéalement) indépendant de la personne menant l'opération de vérification. Par exemple, la masse d'un objet est supposée être indépendante de la personne qui le pose sur une balance. De plus, pour que la vérification soit acceptable, il faut être capable de dire en quoi le processus de vérification mis en œuvre justifie l'énoncé scientifique que l'on cherche à justifier.

Un autre caractère central de la science est le caractère universel de son contenu : la loi de la gravité est supposée s'appliquer partout dans l'univers, aussi bien hier, aujourd'hui que demain. L'**espace** et le **temps** scientifiques sont en cela **homogènes**

¹⁶ Philosophie critique. Tome 3 « La science – épistémologie générale », M.C. Bartholy, J.P. Despin, G. Grandpierre, Magnard, 1978.

et isotropes (c'est ce que l'on appelle le **principe cosmologique**): tous les lieux de l'espace et tous les temps sont équivalents du point de vue de l'applicabilité des énoncés scientifiques. Il s'agit là d'une rupture profonde avec les conceptions magiques ou mystiques qui confèrent des propriétés particulières à certains lieux ou certaines périodes.

Enfin, et il ne s'agit pas là du point le moins important, une discipline scientifique se doit d'être capable d'explicitier ses moyens d'investigation et sa méthode de construction de la connaissance.

Comme il a été vu plus haut, un problème central de l'épistémologie est celui de la définition des **critères de scientificité**, qui sont les critères qui doivent permettre de départager science et non-science. Ces critères, et même leur existence, font l'objet de débats entre plusieurs écoles de pensées. Dans la suite de ce cours, nous verrons comment la science est définie par plusieurs des principaux courants de pensée en épistémologie.

2.1.2 L'opposition rationalisme/relativisme

Existe-t-il des critères de scientificité qui soient universels et qui soient valables à toutes les époques du développement des sciences ? Ce point fait l'objet d'un débat.

Pour les partisans de l'**école rationaliste**, la réponse est positive. On trouve ici une démarche strictement normative. Pour un épistémologue rationaliste « radical », une définition de la science doit pouvoir être formulée sous la forme d'un critère universel radical. Ce critère de scientificité est applicable à toutes les disciplines, et cela à tous les stades de leur développement historique. La thèse rationaliste a été défendue par exemple par Imre Lakatos qui écrit que « *le problème central en philosophie des sciences est celui d'établir des conditions universelles déterminant qu'une théorie est scientifique.* »

A l'inverse, les défenseurs de **relativisme** soutiennent qu'il n'existe pas de critère de scientificité universel. Les critères définitoires sont variables d'une discipline à l'autre, et peuvent évoluer au cours du temps et varier d'une communauté humaine à l'autre. Dans cette perspective, les facteurs psychologiques, sociaux, philosophiques ou religieux acquièrent une grande importance, qu'ils n'ont pas dans la perspective rationaliste. Dans sa version la plus radicale, le relativisme ne reconnaît pas l'existence d'un corpus global de connaissances que l'on peut appeler « science », mais seulement l'existence de plusieurs domaines séparés que l'on peut qualifier individuellement de science.

2.2 Classification(s) des sciences

La multiplicité et la diversité des disciplines scientifiques rendent très difficile une présentation globale du domaine de connaissance et des méthodes couverts par celles-ci. Aussi, pour rendre cette tâche plus aisée, et de mettre en lumière les

interconnexions, les échanges entre les différentes disciplines, il a de tout temps été essayé d'opérer une classification des sciences.

Une telle classification est par nature subjective, en ce sens qu'elle est faite en se basant sur des critères dont le choix n'a rien d'évident ni d'automatique, comme le prouve le fait que de nombreuses classifications aient été proposées au cours des vingt cinq derniers siècles. Par exemple, le philosophe français Auguste Comte (1798-1857), fondateur du positivisme et auteur d'une classification fameuse présentée dans la section suivante, indiquait qu'il avait sélectionné sa classification parmi ... 720 autres choix possibles ! Des critères classiques sont des regroupements par type de problèmes étudiés ou encore par type de méthodes employées.

Un autre point qui réduit la portée de telles classifications est le caractère « flou » déjà mentionné du périmètre de chaque discipline scientifique, qui rend parfois très difficile de classer une discipline parmi telle ou telle catégorie. Ce problème est également renforcé par le fait que les sciences évoluent dans le temps, et qu'une classification proposée à un instant donné peut devenir obsolète dans un futur plus ou moins lointain.

Enfin, il faut noter qu'aucune des grandes classifications proposées n'est réellement satisfaisante, puisque l'on trouve des cas particuliers qui ne rentrent pas (ou très difficilement) dans le cadre d'analyse proposé.

2.2.1 Quelques éléments de classification

Voici quelques critères courants de classification :

- **Sciences formelles et sciences empiriques.**

Cette distinction est centrale dans ce cours, puisqu'il est principalement dédié à la présentation de l'épistémologie des sciences empiriques.

Les sciences empiriques sont les sciences qui font appel à l'expérience sensible (c'est-à-dire aux perceptions que nous avons du monde par nos cinq sens) pour bâtir leurs théories. L'objet des sciences empiriques est donc supposé ne pas être une pure création de l'esprit humain et, d'une manière ou d'une autre, être lié à une réalité extérieure. Parmi les sciences empiriques, on compte : mécanique, physique, chimie, biologie, sociologie, économie, ...

Les sciences formelles se distinguent des sciences empiriques en ce sens qu'elles ne se réfèrent pas à notre perception du monde extérieur. Ces sciences font donc abstraction du contenu pour se focaliser sur la forme. Par exemple, lorsqu'un mathématicien définit l'addition dans le cadre de l'arithmétique, il ne spécifie pas ce qui est additionné (des choux, des carottes, ...) : il définit un cadre formel pour cette opération, qui pourra ensuite être utilisé pour chaque cas particulier. Le principal exemple de science formelle est donné par les mathématiques.

La première différence entre ces deux types de sciences est une différence concernant la nature de l'objet : objet matériel pour les sciences empiriques, objet conceptuel pour les sciences formelles.

La seconde différence est une différence de méthode de construction de la connaissance : les sciences formelles ne se basent que sur la méthode dite hypothético-déductive, alors que les sciences empiriques ont de plus recours à la méthode expérimentale pour vérifier leurs énoncés. Ces méthodes sont discutées au chapitre 5.5.

- **Sciences de la nature et sciences humaines et sociales.**

Ces sciences sont toutes des sciences empiriques. Les sciences de la nature (mécanique, physique, chimie, biologie, ...) ont pour objet le fonctionnement interne de la nature. Les sciences humaines et sociales (économie, sociologie, psychologie, ...) étudient le comportement humain et les structures sociales, c'est-à-dire ce qui est spécifiquement humain (bien que la notion de spécifiquement humain soit de plus en plus affaiblie par les résultats récents provenant de l'étude de certains animaux comme les grands singes¹⁷).

- **Sciences dures et sciences molles.**

Ce critère, souvent retenu par le grand public, est un critère très flou souvent associé à un jugement de valeur : les sciences les plus dures seraient les plus « scientifiques », les plus sérieuses, les plus rigoureuses. Les sciences dites dures sont les sciences formelles et les sciences de la nature, les sciences molles étant les sciences humaines et sociales.

2.2.2 Présentation historique des classifications

La liste des classifications présentée ici n'a nullement l'objectif d'être exhaustive. Elle n'a pour but que de donner des exemples de telles classifications, et, à travers ces classifications, de montrer l'évolution à travers les âges de la manière dont la science a été conçue.

Une des premières classifications connues est celle proposée par Aristote (384-322 av. J.-C.), qui distingue :

1. Les sciences théoriques ou de pure connaissance : mathématiques, physique, métaphysique.
2. Les sciences pratiques ou de l'action : morale, économie, politique.
3. Les sciences poétiques ou de la création : rhétorique, dialectique, poétique.

Comme on peut le voir, la conception des sciences d'Aristote déborde très largement la définition acceptée de nos jours. Elle couvre en effet l'ensemble des domaines de

¹⁷ Voir « Aux origines de l'humanité », 2 tomes, sous la direction de P. Picq et Y. Coppens, Fayard

l'activité intellectuelle du monde à son époque, et serait de nos jours considérée comme plus proche d'une théorie générale de la connaissance.



Figure 4: F. Bacon

Deux mille ans plus tard, le philosophe anglais Francis Bacon (1561-1626) proposait la classification suivante :

1. Les sciences de la mémoire : histoire naturelle, histoire civile.
2. Les sciences de l'imagination : poésie.
3. Les sciences de la raison : philosophie, conçue comme l'étude de Dieu, de la nature et de l'homme.

Cette classification est fondée sur les « facultés de l'âme » (mémoire, imagination, raison) telles qu'elles étaient imaginées par Bacon. On remarque que, dans ce cas également, le champ des disciplines scientifiques est plus étendu que celui reconnu aujourd'hui. Il est à noter que la classification proposée par le philosophe et mathématicien Jean le Rond d'Alembert (1717-1783), dans le *Discours préliminaire de l'Encyclopédie* (1751), s'inspire de celle de Bacon.

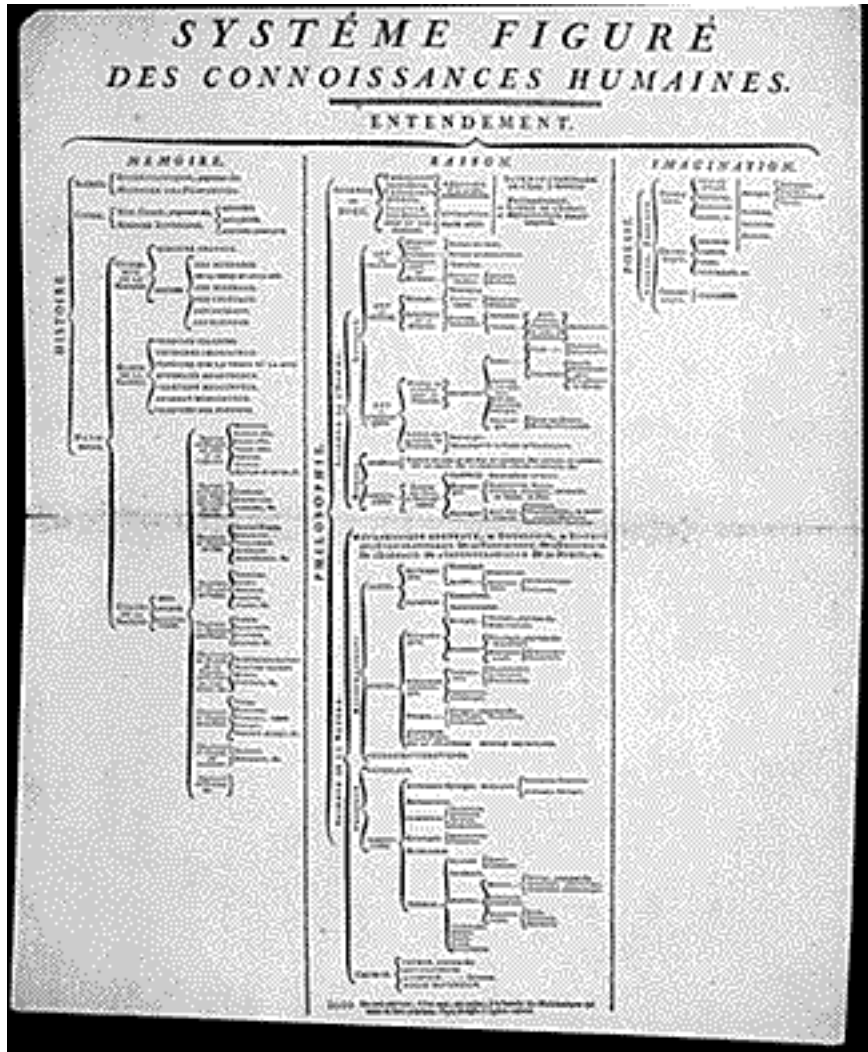


Figure 5: Arbre des connaissances de Francis Bacon

Gustave Ampère (1775-1836) a établi une classification hiérarchique basée sur le principe de dichotomie : chaque classe se divise en deux, chacune des nouvelles classes se divisant à son tour en deux, et ainsi de suite. Au total, la classification d'Ampère compte 128 sciences. Le premier niveau de division est

1. Les sciences cosmologiques ou de la matière.
2. Les sciences noologiques ou de l'esprit.

Le désir de symétrie rend cette classification caduque de nos jours, car elle conduit à la définition artificielle de disciplines scientifiques.

Le philosophe et sociologue anglais Herbert Spencer (1820-1903) a mis au point une classification plus proche de la conception moderne de la science :

1. Les sciences abstraites, qui ont pour objet les



Figure 6: G. Ampère

- formes générales des phénomènes : logique, mathématiques.
2. Les sciences abstraites-concrètes, qui étudient les phénomènes dans leurs éléments fondamentaux : mécanique, physique, chimie.
 3. Les sciences concrètes, qui traitent des phénomènes dans leur ensemble : astronomie, géologie, biologie, psychologie, sociologie.

Comme on peut le voir, toutes les sciences retenues par Spencer vérifient les critères de scientificité retenus de nos jours.

La dernière classification évoquée ici est celle proposée (indépendamment) par les philosophes et scientifiques français Auguste Comte (1798-1857) et Augustin Cournot (1801-1877).

Pour établir sa classification, Comte opère une séparation entre les sciences théoriques et les techniques qu'elles fondent (exemple : l'aérodynamique se fonde sur la mécanique), puis il sépare les sciences abstraites et générales (celles qui ont pour objet la découverte des lois générales : physique, chimie, mécanique) des sciences concrètes, particulières et descriptives (celles qui analysent comment les lois générales sont réalisées dans des cas concrets : zoologie, botanique, minéralogie).

La première classification proposée par Auguste Comte pour les sciences théoriques/abstraites et générales est :

1. Mathématiques
2. Astronomie.
3. Physique.
4. Chimie.
5. Biologie.
6. Sociologie.

Cette classification suit un double critère logique et chronologique.



Figure 8: A. Comte

Le critère logique est un ordre de généralité décroissante (dans l'esprit de Comte, chaque science est tributaire de la précédente dans la classification, mais ne s'en déduit pas) et de complexité croissante des phénomènes étudiés. Ici apparaît le problème lié au caractère évolutif des disciplines scientifiques : l'astronomie connue du temps de Comte ne prend pas en compte l'astrophysique moderne. De même, la physique de cette époque ne comprend pas les grandes avancées que furent la théorie de la relativité et la mécanique quantique. Ainsi, l'ordre de complexité croissante des objets étudiés peut-il être jugé comme largement révisable aujourd'hui. C'est également le cas pour le critère de généralité décroissante : la physique moderne n'est sans doute



Figure 7: A. Cournot

pas moins « générale » que l'astronomie ou l'astrophysique.

Le critère chronologique est celui de la naissance de chaque discipline scientifique. Ce critère est bien évidemment extrêmement difficile à manier, car il faut être en mesure d'identifier et de dater la naissance d'une science. Outre le problème de la disponibilité des documents historiques nécessaires, se pose celui de fixer une date de naissance pour une science. Ce dernier point est particulièrement complexe, car les idées et les méthodes connaissent des maturations lentes, s'étendant souvent sur plusieurs générations, et la détermination d'une date est souvent arbitraire. On détermine souvent la « date de naissance » en repérant, dans l'histoire des idées, le moment où apparaissent des résultats fondateurs ou encore des méthodes de travail jugées fondatrices pour les développements futurs. Un tel travail est bien évidemment éminemment subjectif, et l'établissement d'une date ne signifie en rien que les objets et les problèmes considérés par une science n'étaient pas étudiés avant cette date. Les repères chronologiques donnés dans ce qui suit sont ceux qui correspondent à l'établissement de la liste de Comte.

Ainsi, il est considéré que les mathématiques avec Euclide (~325- ~270 av. J.C.) et la mécanique avec Archimède (~287-~212 av. J.C.) se constituèrent comme sciences autonomes en Grèce au III^e siècle avant notre ère. La naissance de l'astronomie au XVI^e siècle est associée aux travaux de Nicolas Copernic (1473-1543) et celle de la physique à ceux de Galilée (1564-1642) au XVII^e siècle. La chimie « naît » sous sa forme moderne avec Antoine Lavoisier (1743-1794) au XVIII^e siècle. Le XIX^e siècle voit la naissance de la biologie avec Claude Bernard (1813-1878), de la sociologie avec Auguste Comte (1798-1857) et de la psychologie avec Wundt et Ribot.

Une liste rectifiée peut être produite en insérant la psychologie, connue au XIX^e siècle, et en remplaçant l'astronomie par la mécanique :

1. Mathématiques
2. Mécanique.
3. Physique.
4. Chimie.
5. Biologie.
6. Psychologie.
7. Sociologie.

Augustin Cournot (1801-1877) a proposé une liste similaire à peu près au même moment, mais en émettant des critères différents, puisqu'il divisait chaque science en théorique, pratique et historique. Alors que les sciences théoriques ont pour objet les lois naturelles et leurs applications, les sciences historiques ont pour objet les données de faits irréductibles aux lois et à la théorie.

2.3 La science comme phénomène social

2.3.1 La science et le langage

La science est un fait humain collectif qui nécessite la communication, et donc la traduction de ces objets (concepts, lois, énoncés) au moyen d'un langage. C'est l'emploi d'un langage qui permet à la science de découper ses objets, de les

individualiser. En ce sens, on peut dire que sans langage il n'y aurait pas de science, si on se réfère au sens communément admis de ce mot. Cette mise en forme des théories scientifiques incite à se pencher sur les différences que l'on peut faire entre les différents types d'énoncés qui les expriment.

On distingue tout d'abord les **énoncés d'observation** et les **énoncés théoriques**.

Les énoncés d'observation décrivent des objets et des faits directement observables (exemple : il y a un objet posé sur la table). Les énoncés théoriques se réfèrent à des choses non directement perçues (exemple : cet objet est radioactif). Mais un énoncé théorique peut être indirectement connecté à des perceptions, au sens où l'on peut parfois en déduire des énoncés d'observation (exemple : un compteur Geiger, mis en présence de cet objet, indique la présence d'une source radioactive). L'ensemble des énoncés d'observation qui peuvent être rattachés à un énoncé théorique forme son **contenu empirique**, encore appelé son **contenu factuel**. Enfin, on appelle **énoncé métaphysique** tout énoncé théorique sans contenu factuel.

Parmi les énoncés d'observation, on fait la différence entre les **énoncés singuliers** qui se réfèrent à un unique fait observable à un endroit et un instant donnés (exemple : il y a une pomme sur cette table) et les **énoncés universels**, qui portent sur la totalité des événements des d'un type particulier (exemple : les objets pesants tendent à tomber vers le sol).

Une seconde distinction porte sur les **énoncés analytiques** et les **énoncés synthétiques**. Les énoncés synthétiques sont vrais ou faux en vertu des rapports qu'ils entretiennent avec l'expérience (exemple : un objet est posé sur la table). Les énoncés analytiques sont vrais ou faux en raison de leur cohérence interne (exemple d'énoncé faux : il y a et il n'y a pas d'objet sur la table).

2.3.2 La science au-delà des individus ? Objectivisme et individualisme

Une théorie scientifique existe-t-elle indépendamment des croyances individuelles des scientifiques qui la développe ? Dépasse-t-elle les connaissances immédiates des scientifiques, acquérant par une certaine « autonomie », en dehors des champs individuels et sociaux ? Cette question fait débat et, comme il sera discuté au chapitre 9, a une importance certaine dans l'analyse du progrès scientifique.

Pour les partisans de l'**individualisme**, une théorie scientifique n'est rien d'autre que la somme des connaissances (au sens discuté au chapitre précédent : des croyances vraies et justifiées) des scientifiques qui la développent. La question de la connaissance scientifique est alors subordonnée à celle de la connaissance par un individu, telle qu'elle a été évoquée au chapitre 1.

Pour les défenseurs de l'**objectivisme** (Popper, Lakatos, Chalmers, ...), une théorie scientifique ne se résume pas à la somme des croyances individuelles. Pour reprendre Popper, la science est une connaissance objective et, en tant que telle, est « *indépendante de la croyance d'une quelconque personne, ou de sa disposition à admettre, ou à affirmer, ou à agir. La connaissance au sens objectif est connaissance sans connaisseur ; elle est connaissance sans sujet connaissant* ». Les arguments en faveur de cette position sont multiples. Tout d'abord, il convient de remarquer que le

champ couvert par une discipline scientifique moderne comme la physique est si vaste qu'aucun individu ne peut le maîtriser intégralement. Les chercheurs sont aujourd'hui regroupés en communautés spécialisées, et aucune de celles-ci n'est capable d'identifier l'ensemble des liens ou des contradictions contenus par la physique. C'est l'**argument de la ramification** : puisqu'elle est si complexe, elle existe au-delà des connaissances individuelles. Ensuite, une théorie scientifique a des conséquences, appelées **opportunités objectives**, qui ne sont pas explicitement connues et formulées par les chercheurs qui la développent. Par exemple, James Clerk Maxwell n'a jamais imaginé l'existence des ondes radio, découvertes après sa mort, bien que leur existence fût contenue dans sa théorie de l'électromagnétisme. C'est l'**argument de l'existence des opportunités objectives**.

3 Éléments constitutifs d'une science empirique

3.1 Faits, lois, principes, modèles et théories

Nous allons maintenant nous attacher à définir, par ordre croissant de généralité et de profondeur représentative et explicative, les différents niveaux d'abstraction rencontrés dans les sciences empiriques.

3.1.1 Faits, phénomènes et évènements

Les **faits** sont les données immédiatement accessibles par l'observation. On distingue parfois deux types de faits : le fait brut, et le fait scientifique. Le **fait brut** est l'observation non analysée : la chute d'une pomme, le mouvement de la lune sont des faits bruts. Le **fait scientifique** est que tous les corps chutent sous l'effet de l'attraction terrestre (c'est ce que montrent tous les faits bruts relatifs à ce sujet).

On identifie ensuite, parmi les faits, les phénomènes et les évènements. Un **phénomène** est un fait associé à un changement, et qui est répétable : la chute de la pomme est un phénomène. Tout fait n'est pas un phénomène, soit parce qu'il n'implique pas de changement (exemple : la (quasi-)rotondité de la terre est un fait, qui n'est pas associé à un changement ou une évolution), soit parce qu'il n'est pas répétable (on parle alors d'un **évènement**). Un exemple d'évènement est la désintégration d'un noyau atomique, qui ne se produit qu'une seule fois.

3.1.2 Qu'est-ce qu'une loi scientifique ? Et un principe ?

Le terme loi est à prendre avec des significations différentes dans son usage commun et son usage scientifique.

Dans le sens commun, il a une connotation juridique et désigne une règle en vigueur dans un groupe humain, dont la violation sera sanctionnée. Il s'agit d'une convention sociale arbitraire, valable durant un certain temps dans un domaine de l'espace défini. La loi, dans cette acceptation, est contraignante : elle tend à gouverner ou à régir le comportement des individus qui lui sont soumis. Le respect ou la violation de la loi est de la part de ceux-ci un acte intentionnel. Le point important à retenir ici est que le respect de la loi n'est ni nécessaire ni obligatoire.

Le sens du terme loi, tel qu'il est employé dans le cadre de sciences empiriques, est très différent ; on emploie parfois le terme de **loi naturelle** ou **loi de la nature**. Une loi scientifique est l'expression mathématisée d'une corrélation répétable, d'un comportement constant ou d'une fréquence statistique observée parmi un ensemble de faits. Elle est déduite d'un certain nombre d'observations et généralise celles-ci, en en retenant le caractère stable. Il est donc inexact de dire que les faits sont régis par des lois : il faut dire que les faits comportent des lois. A la différence du sens juridique usuel, la loi scientifique est constatative et non normative.

Les lois scientifiques partagent toutes la même structure logique, et peuvent s'énoncer sous la forme :

Forme générale d'une loi

Quel que soit x , si x est A , alors, x est B

On peut distinguer les **lois qualitatives** des **lois quantitatives**. Les premières portent sur des variables qualitatives (dites encore **variables ordinales**) qui peuvent être mises en ordre d'intensité, mais ne se prêtent pas aux opérations arithmétiques. Un exemple de loi qualitative est

Loi qualitative : le cuivre est conducteur d'électricité.

Cette loi peut être réécrite sous la forme logique générale donnée plus haut :

Loi qualitative (forme logique générale) : Quelque soit un matériau, si ce matériau est du cuivre, alors il conduit l'électricité.

Les lois quantitatives portent à l'inverse sur des variables quantitatives (appelées également **variables cardinales**) sujettes aux opérations arithmétiques et qui peuvent être mesurées lors d'expériences. Certaines lois font apparaître des constantes, qui gardent des valeurs fixes. La loi indiquant qu'il existe une relation constante entre deux ou plusieurs variables, elle permet en un sens de prédire la valeur d'une de ces variables si toutes les autres sont données. La structure d'une telle prédiction s'exprime sous la forme de ce qui est appelée une **contra-factuelle** ou encore une **conditionnelle irréal** : *si ... alors ...*. Le caractère conditionnel est porté par le mot *si*, et irréal vient de ce que l'évènement n'est pas réalisé au moment où l'on émet la prédiction.

Illustrons maintenant les lois quantitatives au moyen de la loi des gaz parfaits (loi de Mariotte ou de Boyle-Mariotte). L'expression la plus simple, dans le cas de transformations à température constante (cas isotherme) est :

Loi quantitative de Mariotte : $PV = \text{constante}$

où P et V désignent respectivement la pression et le volume. Cette loi peut être réécrite sous la forme générale :

Loi quantitative (forme générale) de Mariotte : Quel que soit x , si x est un gaz parfait, alors le produit de sa pression par le volume occupé est constant lors d'une transformation isotherme.

La valeur prédictive est immédiate. Connaissant la valeur du produit P_0V_0 pour un gaz parfait x pour une pression P_0 et un volume V_0 , on prédit en se servant de la loi la valeur de la pression P_1 si le volume est changé en V_1 (la température étant maintenue constante) :

$$P_1 = P_0 \frac{V_0}{V_1}$$

Revenons à ce qui a été dit plus haut concernant la validité des lois, qui est conditionnée aux observations. La loi de Mariotte, telle qu'elle est exprimée plus haut, repose sur l'hypothèse que la température est constante, et peut être proposée à partir d'une série d'observations vérifiant cette hypothèse. Que se passe-t-il maintenant si la température ne peut être maintenue constante ? La loi ne permet plus de représenter les faits. Il faut la rendre plus générale pour tenir compte des relations qui peuvent exister entre température, pression et volume :

$$PV = nRT$$

où T désigne la température, n le nombre de molécules et R est la constante des gaz parfaits (1,985 calories/degé Celsius). Cette nouvelle loi, qui fait apparaître une constante R valable pour tous les gaz parfaits, contient la précédente et en étend le domaine de validité. Cette loi peut encore être généralisée en prenant en compte des interactions physiques plus complexes. Par exemple, la loi de Van der Waals¹⁸ rend compte de l'attraction mutuelle des molécules et de leur taille :

$$P\left(1 + \frac{a}{V^2}\right)(V - b) = RT$$

où a et b sont des constantes qui dépendent du gaz.

Nous allons voir maintenant que la notion de loi est plus compliquée qu'il n'y paraît, et soulève des problèmes qui demeurent ouverts.

Une loi est-elle universelle et définitive ? Non, car, déduite de l'expérience, elle est conditionnée par les hypothèses implicites qui ont guidé la collecte et l'analyse des faits : la précision des mesures, le choix des objets (masse, vitesse, énergie, ...) à conserver pour exprimer la loi, ... Une loi est donc une hypothèse jugée très probable car corroborée par de nombreuses observations. Mais elle peut être améliorée, généralisée si de nouvelles observations indiquent que la corrélation attendue n'est pas observée. En ce sens, on ne peut pas dire qu'une loi scientifique est violée par les faits (les faits sont ce qu'ils sont et ils n'ont pas d'intention), mais plutôt que la loi est caduque dans telle ou telle situation. Mais elle possède toutefois un caractère universel, car elle est sensée être valide pour tous les cas qui satisfont ses hypothèses de dérivation.

Un second problème consiste à faire la distinction entre les lois et les **généralisations accidentelles**. Les généralisations accidentelles sont des expressions qui possèdent la même forme logique universelle que les lois (quel que soit x , ...) mais qui, à l'évidence, ne peuvent pas être retenues comme des lois scientifiques. L'exemple fameux suivant a été proposé par Nelson Goodman en 1955 : « *Toutes les pièces que j'ai dans ma poche à l'instant t sont en argent.* » Comme on peut le voir, la structure logique de cette proposition correspond bien à celle d'une loi, puisqu'elle vérifie les critères de vérité et d'universalité. Comment faire la différence ? On ne peut pas

¹⁸ Johannes van der Waals (1837-1923), physicien néerlandais, prix Nobel de Physique en 1910

utiliser le fait que la généralisation accidentelle fait référence à des entités individuelles précises (instant particulier, poche déterminée, ...) car il est possible de reformuler les énoncés des généralisations accidentelles sans faire référence à de telles entités, et certaines lois font appel à de telles entités¹⁹. Une autre objection pourrait être que les généralisations accidentelles ne sont pas vraiment universelles, et portent sur un nombre fini de cas singuliers. Ce critère fonctionne bien dans le cas de l'exemple de Goodman, mais il n'est pas pleinement satisfaisant. La raison est que toutes les lois sont bâties à partir d'un nombre fini d'énoncés d'observation, et que rien ne permet donc de les distinguer a priori des généralisations accidentelles avec ce critère. Ceci est illustré par un exemple suggéré par Reichenbach en 1954 :

1. Tous les corps en uranium enrichi pur ont un volume inférieur à 1 km^3 .
2. Tous les corps en or pur ont un volume inférieur à 1 km^3 .

Aucune de ces deux propositions ne fait référence explicitement à des entités individuelles précises, mais elles portent bien sur un nombre fini d'objets réels (le nombre d'atomes d'uranium et d'or présents dans l'univers étant fini). Mais elles ont des statuts différents : la première est une loi, car un corps d'un volume de 1 km^3 d'uranium enrichi aurait une masse très supérieure à la masse critique à partir de laquelle une réaction nucléaire se déclenche spontanément²⁰ (environ 50 kg pour une sphère nue d'uranium enrichi) ; la seconde est une généralisation accidentelle, un tel corps en or pur pouvant exister.

Cette difficulté à définir ce qu'est une loi a conduit au développement de plusieurs écoles de pensées.

Pour les philosophes de tendance **empiriste**, il n'existe pas de critère parfaitement objectif, et le fait de considérer un énoncé comme une loi est le résultat d'un *consensus* parmi une communauté scientifique. Il s'agit donc ici d'un jugement, emporté par la conviction des scientifiques, et comme tel, reconnaître qu'un énoncé est une loi est un fait subjectif. Un *critère jugé psychologiquement important* est celui de l'optimalité (le plus simple, le plus efficace, ...) du système des lois retenues, le problème étant reporté sur la définition de ces critères et la possibilité de les mesurer en pratique.

Pour les philosophes de tendance **réaliste** (c'est-à-dire ceux qui considèrent que les énoncés scientifiques portent sur des entités qui existent réellement dans le monde, indépendamment de l'activité scientifique et humaine), la distinction entre lois et généralisations accidentelles vient de ce que les premières portent sur des propriétés objectives, considérées comme des **universaux**, alors que les dernières ne font référence qu'à des réalisations contingentes particulières. Le problème des approches réalistes (il en existe de nombreuses variantes) réside alors dans la définition de la nature de ces universaux (par exemple, certains philosophes admettent qu'il peut exister des universaux non-instanciés, c'est-à-dire sans aucune concrétisation dans l'univers, alors que d'autres non) et des relations qui peuvent exister entre eux (une question importante étant de savoir si les relations qui existent entre universaux et qui sont décrites par les lois sont nécessaires ou pas).

¹⁹ Un exemple est donné par la première loi de Kepler, qui stipule que les orbites des planètes sont des ellipses dont le Soleil occupe un foyer. Cette loi fait apparaître une entité précise, le Soleil !

²⁰ Le phénomène de fission spontanée fut découvert par les physiciens G.N Flerov et K.A. Petrzhak en 1940 à partir de noyaux d'uranium 238.

Les lois discutées dans ce qui précède sont des lois dites **déterministes**, car à un système déterminé et parfaitement connu, elles associent de manière *certaine* un état ou une propriété. Depuis l'avènement de la physique statistique au XIX^e siècle avec la naissance de la théorie cinétique des gaz et de la thermodynamique statistique, il existe une autre sorte de lois, les **lois statistiques**. Elles peuvent s'énoncer de la manière suivante :

Forme générale d'une loi statistique

Quel que soit x , si x est A , alors, x est B avec la probabilité $P(B/A)$

où la probabilité $P(B/A)$ est connue par ailleurs. De telles lois sont employées pour décrire l'évolution de grands ensembles d'entités individuelles, dont le suivi individuel est impossible ou sans intérêt. Un exemple courant est celui de la description de la radio-activité, qui provient de la désintégration spontanée du noyau atomique. On ne s'intéresse pas au comportement de chaque atome, mais au nombre moyen d'atomes qui se désintègre par unité de temps. On ne sait donc pas dire ce que devient chaque atome, mais on peut déduire, à chaque instant, quelle est la probabilité qu'il se soit désintégré. Une telle description probabiliste de la nature est à la base de la physique quantique, comme cela est discuté au chapitre 11.

Notons que le problème de la généralisation accidentelle existe aussi pour les lois statistiques : des phénomènes a priori sans rapport les uns avec les autres peuvent avoir des comportements semblables, laissant à penser qu'ils sont corrélés. Ceci est illustré par la figure ci-dessous, qui, sur le plan statistique, suggère qu'il existe une corrélation forte entre le nombre de tâches solaires et le nombre de d'élus républicains au Sénat des Etats Unis !

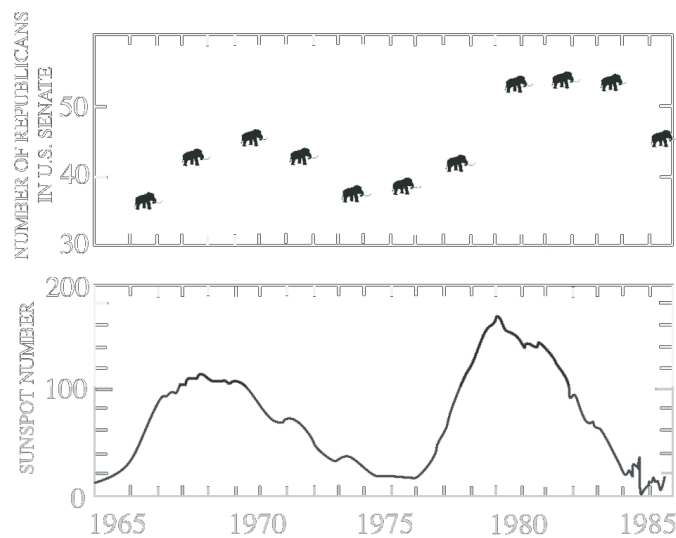


Figure 9 Exemple de corrélation accidentelle. En haut : le nombre de sénateur républicains aux Etats Unis ; en bas : le nombre de tâche solaires observées.

Les textes scientifiques font également parfois mention de **principes**. Dans les sciences expérimentales, les principes sont des sortes de loi d'un haut degré de généralité, qui dépassent de très loin les faits qui les ont suggérés. Chez certains auteurs, les principes acquièrent un statut proche de celui des postulats mathématiques, en ce sens que leur niveau de généralité ne permet pas une validation

directe et complète. Ainsi, la seconde loi de Newton est parfois nommée **principe fondamental de la dynamique**, et l'hypothèse d'égalité entre masse inertielle et masse gravitationnelle est désignée sous le nom de **principe d'équivalence**. Un autre principe important est le **principe copernicien**, qui stipule que l'espace et le temps sont homogènes au sens où les lois de la physique sont les mêmes en tout temps et à tout point dans l'univers.

3.1.3 Qu'est-ce qu'un modèle ?

Les faits permettent de dégager des lois. Pour obtenir un cadre de réflexion qui permette de synthétiser ces lois, de les regrouper et de trouver une certaine unité entre elles, on utilise des modèles. Un modèle peut être défini²¹ comme « *un cadre représentatif, idéalisé et ouvert, reconnu approximatif et schématique mais jugé fécond par rapport à un but donné : prévoir, agir sur la nature, la connaître mieux, etc.* »

Exemple 1 : Le solide rigide.

Dans la discussion sur la définition du mouvement (voir le paragraphe 3.3.4), on introduit la notion de solide rigide, c'est-à-dire celui d'un corps solide idéal qui ne subit aucune déformation durant son mouvement. Considérons un corps réel, par exemple une boule de billard, que l'on suit pendant son déplacement et les chocs avec les autres boules ou les bandes de la table de billard. La boule est « dure » (d'après le sens commun), et, durant la partie, elle semble à l'œil nu conserver sa sphéricité. De plus, son mouvement sur la table peut aisément, pour un spectateur, être représenté par une composition de translations et de rotations. Pour les besoins des joueurs et des spectateurs, l'assimilation des boules à des solides rigides est donc un bon modèle, en ce sens que, couplé aux trois lois de Newton, il permet de décrire et de prévoir le déplacement des boules. Il est efficace et utile en cela, et permet par exemple aux joueurs de prévoir leurs coups. Cela veut-il dire que les boules ne subissent aucune déformation ? Non, à l'échelle microscopique, un observateur disposant des moyens d'observation adéquats verrait des déformations de la surface de la boule lors des impacts et la création d'ondes à l'intérieur du solide, qui sont décrites dans un cadre théorique plus complexe : celui de la mécanique des milieux continus. Le modèle de solide rigide est donc une idéalisation, une vue de l'esprit, qui, conservant certaines propriétés des objets, permet, grâce aux lois adéquates, d'obtenir une certaine maîtrise des faits.

Exemple 2 : Le milieu continu.

Dans l'exemple qui précède (le solide rigide), on considère la matière comme continue, puisque la définition du solide rigide est basée sur l'invariance de la distance entre les points situés dans le solide. Or, nous le savons, la matière n'est pas continue, mais faite d'atomes séparés par du vide.

Pour caractériser les matériaux à l'échelle macroscopique et étudier les objets dont la taille est « grande » devant celle des atomes ou des molécules qui le composent, il n'apparaît pas utile d'utiliser une description à l'échelle de l'atome, et l'on introduit un modèle très puissant en mécanique : le modèle du milieu continu. Les propriétés macroscopiques de ce milieu (la rigidité d'un solide, la viscosité d'un fluide, ...)

²¹ L. Soler, « Introduction à l'épistémologie », Ellipses, 2000

représentent l'effet, à grande échelle, des interactions entre les atomes ou les molécules.

Exemple 3 : Le modèle de Maxwell en cinétique des gaz



Figure 10 D. Bernoulli

Examinons maintenant un modèle célèbre : celui introduit en 1866 par le physicien anglais James Clerk Maxwell (1831-1879) pour décrire les gaz. Des travaux sur le même sujet ont été également réalisés indépendamment par le physicien allemand Ludwig Boltzmann (1844-1906).

C'est en 1730 que Daniel Bernoulli (1700-1782) fait état d'une « marche au hasard » des molécules d'un gaz avec chocs entre molécules et parois du récipient qui les contient. Cette idée sera ensuite par, entre autres, James Prescott Joule (1818-1889) et Rudolf Clausius (1822-1888). En 1827, la preuve expérimentale de cette agitation est apportée par le botaniste écossais Robert Brown (1773-1858), non pas sur les gaz mais sur les liquides ; il observe à cette époque au microscope que les particules de pollen bougent de manière désordonnée dans l'eau (mais il attribue cela à une activité vitale des grains de pollen, et on a l'action des molécules d'eau !). Le *mouvement brownien* des molécules est mis en évidence. Il s'agit d'un déplacement désordonné des molécules, qui dépend de la température, appelé *mouvement d'agitation thermique*. En ce qui concerne les gaz, l'agitation thermique n'a été mise en évidence expérimentalement qu'en 1908, par Maurice de Broglie (1875-1960), et confirmée par Louis Dunoyer (1880-1963) en 1911.



Figure 11 J.C. Maxwell

L'idée de Maxwell qui est à la base du modèle théorique qui porte son nom consiste à assimiler les atomes ou les molécules qui composent le gaz à des petites sphères rigides, qui se conduisent comme des boules de billard : elles s'entrechoquent, elles rebondissent sur les parois (solides) du récipient qui contient le gaz, et leur mouvement peut être décrit au moyen des lois de Newton. Il s'agit bien entendu d'un modèle : les

molécules ne sont pas sphériques, et ne sont pas rigides. Mais l'approximation de Maxwell, *son modèle*, permet de rendre compte de certaines propriétés essentielles de leur évolution, et permet notamment, en faisant une moyenne sur un grand nombre de molécules, de retrouver les caractéristiques macroscopiques d'un gaz. Elle permet également de lier des grandeurs associées au gaz aux caractéristiques des molécules idéalisées.



Figure 12: M. de Broglie

Ainsi, la pression exercée par un gaz sur la paroi d'un récipient est associée à l'énergie cinétique transmise à la paroi par les molécules lorsque celles-ci viennent rebondir dessus. La température du gaz est liée à l'énergie cinétique des molécules : plus celles-ci vont vite, plus la température est élevée.

Outre le fait qu'il permet de lier des grandeurs macroscopiques à des phénomènes microscopiques, le modèle de Maxwell possède également des vertus explicatives : il permet, par exemple, de lier la loi de Mariotte à des comportements à l'échelle microscopique. Considérons la forme la plus simple de la loi de Mariotte : $PV = \text{constante}$, pour une transformation à température constante, et voyons ce qui se passe lorsqu'on comprime un gaz dans un récipient clos en réduisant son volume (avec un piston, par exemple). La constance de la température indique que la vitesse des molécules reste constante (en norme, pas en direction). Si l'on réduit le volume du récipient, on en rapproche les parois. Le temps moyen mis par une molécule pour rebondir d'une paroi à l'autre est donc plus court, et, par unité de temps et par unité de surface, il y a plus de molécules qui viennent frapper la paroi, ce qui correspond bien à une augmentation de la pression.

Il faut noter qu'il existe une hiérarchie de modèles en théorie cinétique des gaz, qui donnent accès à des descriptions physiques de plus en plus riches :

- Dans sa version initiale (la plus simple), la taille des molécules et leurs interactions sont négligées. A l'échelle macroscopique, on retrouve la loi des gaz parfaits.
- Si on prend en compte la taille des molécules et leurs collisions, on peut décrire les phénomènes de transports macroscopiques : diffusion, conductibilité thermique. Ces propriétés sont associées, à l'échelle macroscopique, à de nouvelles propriétés des fluides, comme la viscosité et la diffusivité (qui s'ajoutent donc à la pression et la température)
- En prenant en compte les interactions entre molécules (cela est dû au physicien Van der Waals (1837-1923, prix Nobel de physique en 1910), en 1873), on peut également tenir compte de la transition liquide-gaz (phénomènes d'évaporation, de condensation et d'ébullition).

Ces exemples permettent de dégager les caractéristiques générales des modèles bâtis par les scientifiques :

- Les modèles sont basés sur des simplifications des faits. Cette étape de simplification nécessite d'identifier les aspects importants pour le but considéré (rendre compte de telle ou telle caractéristique observée). Elle est le point-clé de la modélisation. Par exemple, le modèle de Maxwell ne considère que des interactions mécaniques simples (collisions), et néglige de possibles interactions de nature chimique ou électromagnétique. Cette simplification implique que le modèle possède un domaine de validité doublement limité : il n'est construit que pour rendre compte de certains phénomènes (exemple : le modèle de Maxwell n'est d'aucune utilité pour étudier la chimie des gaz) et n'est valable que pour un certain domaine d'étude (exemple : le modèle de Maxwell n'est plus efficace aux très hautes températures, où la ionisation des molécules devient importante).
- Un point important est qu'il n'y a pas forcément correspondance exacte entre les éléments constitutifs du modèle et les faits observés : les molécules idéales de Maxwell sont très différentes des molécules réelles d'un gaz.

- Un modèle, lorsqu'il est analysé au moyen des mathématiques et de la logique, permet de prédire des faits virtuels (ce que Popper appelle des falsificateurs virtuels). L'emploi des mathématiques permet des prédictions quantitatives, qui peuvent être confrontées aux résultats des observations disponibles. Par exemple, le modèle de Maxwell peut être employé pour prédire l'évolution de la pression et de la température d'un gaz lorsqu'on le comprime. Ces prédictions peuvent être comparées aux données expérimentales.

3.1.4 Qu'est-ce qu'une théorie scientifique ?

Le dernier niveau, le plus général, est celui des théories scientifiques. Une théorie est un système cohérent qui coordonne, relie et unifie des lois, des hypothèses, des principes et des modèles, les uns apparaissant comme complémentaires des autres. Elle est plus générale que les modèles qu'elle utilise. Une théorie peut utiliser un ou plusieurs modèles, et replace ceux-ci dans un contexte conceptuel plus général. Il est à noter qu'un modèle peut être employé par des théories différentes.

De manière plus précise, on distingue généralement quatre fonctions d'une théorie scientifique :

4 fonctions de la théorie

- La **fonction explicative-prédictive** : en employant la théorie, le calcul et des hypothèses, on peut déduire des lois et des faits (qui devront, bien sûr, être confrontés avec l'expérience) alors que ceux-ci n'ont pas encore été observés.
- La **fonction unificatrice** : la théorie permet de rassembler, d'unifier un grand nombre de faits au sein d'un même cadre conceptuel, et de les englober dans un nombre réduit de modèles et de lois.
- La **fonction heuristique** : les théories guident les recherches en suggérant de nouvelles voies ou au contraire en conduisant à l'abandon de certains développements jugés moins féconds.
- La **fonction de représentation** : les modèles et le cadre fournis par une théorie offrent une représentation de l'univers sensible, qui permet à chacun d'ordonner sa vision du monde.

Exemple 1 : La théorie cinétique des gaz

La théorie cinétique des gaz est basée sur le modèle de Maxwell, qui traite de la dynamique des gaz. Au moyen de l'analyse mathématique, elle permet de déduire entre autres la loi des gaz parfaits et la loi de diffusion de Graham.

Exemple 2 : Les théories ondulatoire et corpusculaire de la lumière

Deux autres exemples sont les théories ondulatoire et corpusculaire de la lumière.

Dans le cadre de la théorie corpusculaire de Newton, la lumière est composée de particules (les photons) dont la masse varie avec les couleurs du spectre. Selon la théorie ondulatoire dérivée de l'électromagnétisme de Maxwell, la lumière se propage

sous forme d'une onde transversale. Ces deux théories rendent compte de différents phénomènes : diffraction, réfraction, réflexion.

Il faut noter que l'effet photo-électrique n'est pas pris en compte dans le cadre de la théorie ondulatoire. Ces deux théories ne sont ni plus vraies ni plus fausses l'une que l'autre : elles sont plus ou moins efficaces pour traiter tel ou tel problème.

3.2 Approches globales de la science et des théories

3.2.1 Paradigme et matrice disciplinaire : Kuhn

Nous allons maintenant voir le concept de **paradigme** (encore appelé **matrice disciplinaire**) introduit par l'épistémologue Thomas Kuhn (1922-1996).

Pour Kuhn, un paradigme est ce qui fait l'objet d'un consensus au sein d'une communauté scientifique (la mécanique de Newton, la théorie corpusculaire de la lumière, ...). Le paradigme est un ensemble de plusieurs choses différentes :

Contenu du paradigme selon Kuhn

- Des **contenus théoriques**, que personne ne songe à remettre en question
- Des **normes de la recherche scientifique**, qui représentent un ensemble de valeurs de la communauté scientifique, et qui dérivent de ses critères de scientificité.
- Un **savoir-faire théorique et pratique**.

Cette structure multiple fait que, pour les sciences empiriques dures, on peut identifier dans un paradigme des **généralisations symboliques** (les équations qui traduisent des idées, comme les lois de Newton ou la loi des gaz parfaits) et un **contenu métaphysique** associé aux modèles et aux images employés par les scientifiques (exemple : les molécules sphériques et élastiques de Maxwell dans son modèle de théorie cinétique des gaz). Ce contenu métaphysique est soit **ontologique**, si l'on défend l'idée de la théorie-reflet, soit **heuristique** dans le cadre de la théorie-outil²². Enfin, selon Kuhn, on doit également y inclure les **exemples communs** employés par la communauté pour former ses membres et faire comprendre les lois. Outre leur fonction pédagogique, ces exemples ont également un rôle dans la recherche puisqu'ils permettent de raisonner par analogie.

Ainsi, le paradigme de la mécanique newtonienne est composé des éléments suivants : les trois lois de la dynamique et la loi de la gravitation de Newton, des méthodes pour appliquer les lois de Newton à divers problèmes (mouvement des corps célestes, chocs élastiques des corps rigides, mouvement des pendules). Le paradigme contient

²² Les notions de théorie-reflet et théorie-outil sont discutées au chapitre 7, dans la section consacrée à l'opposition réalisme/antiréalisme. En résumé, contentons nous de dire pour le moment que si une théorie scientifique décrit vraiment le monde tel qu'il est elle sera qualifiée de théorie-reflet (elle reflète le monde), alors que si elle n'est qu'une simple convention passée entre les scientifiques pour rendre compte d'une manière simple des observations effectuées, elle sera une théorie-outil. Dans ce dernier cas, les concepts auxquels elle fait appel (masse, énergie, onde, ...) ne sont pas sensés avoir d'existence de le « monde réel ».

également les instruments de mesures et les techniques expérimentales adéquates, comme l'usage du télescope pour l'observation du mouvement des planètes, dans le cadre de l'application de la mécanique newtonienne à l'astronomie.

Les paradigmes de Kuhn ne se résument pas à des lois explicitement formulées et des prescriptions méthodologiques : la plus grande partie de la connaissance associée à un paradigme est tacite, c'est-à-dire qu'elle n'est jamais formulée explicitement. Cette **connaissance tacite** est très large, et est acquise au cours de la formation scientifique. Un autre point important est qu'un paradigme n'est jamais parfait, au sens où chaque paradigme admet des problèmes (ou **énigmes**) non résolus appelés **anomalies**. Un exemple d'anomalie du paradigme de la mécanique newtonienne est la trajectoire de Mercure (qui est explicable au moyen de la théorie de la relativité d'Einstein).

Ces caractéristiques font des paradigmes de Kuhn des structures assez floues, qu'il est impossible de définir exactement avec précision. Pour expliquer cela, Kuhn établit une analogie avec le problème de la définition d'un « jeu » étudié par Ludwig Wittgenstein (1889-1951). Dans sa théorie du jeu, Wittgenstein a montré qu'il est impossible de définir un ce qu'est un jeu, au sens où dès qu'une définition exacte est donnée, elle conduit à définir comme des jeux choses qui n'en sont pas alors qu'elle ne reconnaît pas comme tel des choses qui en sont. Mais, pour Kuhn et ses partisans, cela n'interdit pas pour autant d'utiliser le concept de paradigme.

Quand un paradigme peut-il être qualifié de scientifique ? Ou, pour reformuler la question, quels sont les critères de scientificité associés à la théorie de Kuhn ? Pour Kuhn,

Critère de scientificité de Kuhn

un paradigme est scientifique s'il mène au développement d'une tradition de **science normale**.

Le concept de science normale sera développé au chapitre 9. Pour le moment, retenons qu'une science normale est une science mûre guidée par un paradigme unique. Dans une science normale, le nombre et la gravité des anomalies sont suffisamment faibles pour que les scientifiques puissent travailler de manière féconde et sereine sans remettre en cause les fondements du paradigme. On voit que le critère de Kuhn est basé sur une analyse historique du développement d'une science. Il s'agit donc d'une analyse a posteriori. En conséquence, Kuhn et ses partisans font un grand emploi de la méthode diachronique.

La position de Kuhn sur l'universalité de son critère de scientificité est-elle rationaliste ou relativiste ? Kuhn se considérait comme un rationaliste très attaché au progrès scientifique. Il défendait la thèse selon laquelle les théories récentes sont meilleures que celles qu'elles ont remplacées, mais il déniait que ce progrès soit associé à un progrès vers la vérité. Sa théorie est également analysée comme étant une théorie relativiste par certains commentateurs, du fait de la grande importance qu'il donne aux facteurs sociaux et psychologiques dans le processus d'évolution des théories scientifiques et à sa théorie de **l'incommensurabilité des paradigmes**²³. Notons que la définition de la science de Kuhn a été critiquée par d'autres penseurs contemporains, comme Popper et Lakatos.

²³ ce concept est développé au chapitre 9. Pour le moment, il suffit de retenir que, pour Kuhn, deux paradigmes différents sont incommensurables en ce sens qu'ils ne peuvent pas être comparés.

3.2.2 Les programmes de recherche : Lakatos

La réflexion de l'épistémologue Imre Lakatos est fondée sur le concept de **programme de recherche**. Tout comme la théorie de de Kuhn dont elle représente une évolution, celle de Lakatos est construite de manière à rendre compte du caractère évolutif des théories scientifiques en fonction du temps. Un programme de recherche lakatosien est une structure qui permet (du moins selon Lakatos) de rendre compte du développement des théories scientifiques et d'expliquer leur évolution. Un programme de recherche est composé de quatre éléments :

Structure d'un programme de recherche selon Lakatos

- Un **noyau dur**, qui est formé d'hypothèses théoriques très générales. C'est ce noyau qui caractérise le mieux le programme de recherche. Le noyau dur est arbitrairement considéré comme infalsifiable (au sens de Popper) par « *décision méthodologique de ses protagonistes* »²⁴, c'est-à-dire qu'il ne peut pas être remis en question.
- Une **ceinture protectrice**, qui est composée d'hypothèses auxiliaires explicites qui complètent le noyau dur, d'autres hypothèses sous-jacentes à la description des conditions initiales et en des énoncés d'observation.
- Une **heuristique négative**, qui consiste en ce que le noyau dur est considéré comme infalsifiable.
- Une **heuristique positive**, qui selon Lakatos²⁵ « *consiste en une série partiellement formulée de propositions ou d'indications sur la façon d'opérer des transformations, de développer la ceinture protectrice réfutable* ».

Un point important ici que le caractère irréfutable du noyau dur du programme. Toute inadéquation entre un énoncé d'observation et la théorie doit donc mener à une modification de la ceinture protectrice, c'est-à-dire par l'ajout de nouvelles hypothèses ou le remplacement d'anciennes.

Prenons l'exemple du programme de développement de la mécanique de Newton : son noyau dur est composé des trois lois de Newton et de sa loi de la gravitation. Le but de ce programme était d'expliquer l'ensemble des phénomènes connus (chute des corps, chocs de corps solides, mouvement des planètes, ...). La ceinture protectrice (à un instant donné, puisque celle-ci évolue constamment) est l'ensemble des hypothèses auxiliaires et des observations disponibles qui lui sont relatives (par exemple, les observations astronomiques sur le déplacement des planètes).

Il faut également noter le caractère arbitraire de la définition de ce noyau pour un programme donné, puisqu'il résulte d'une décision. Il a donc un caractère subjectif. Changer ou modifier le noyau dur, c'est quitter le programme de recherche et en définir un nouveau. Des exemples de changement de programme de recherche sont le passage de la mécanique médiévale à la mécanique newtonienne, puis de la

²⁴ I. Lakatos « Falsification and the methodology of scientific research programmes », 1974, cité par A. Chalmers, « Qu'est-ce que la science ? », Coll. Biblio essais, Livre de poche, p.136

²⁵ I. Lakatos, Ibid.

mécanique newtonienne à la mécanique relativiste d'Einstein, et enfin le passage à la mécanique quantique. Un autre exemple tiré de l'histoire de l'astronomie est le passage du modèle de Ptolémée (modèle géocentrique, dont le noyau dur est « le Soleil et les planètes gravitent autour de la Terre ») au modèle de Copernic (modèle héliocentrique, qui a pour noyau dur « toutes les planètes gravitent autour du Soleil »).

Quand peut-on qualifier un programme de recherche de scientifique ? Ou, pour reformuler la question, quels sont les critères de scientificité associés à la théorie de Lakatos ? Selon Lakatos, un programme doit vérifier les deux conditions suivantes pour être scientifique :

Critères de scientificité de Lakatos

- Le programme de recherche doit avoir un degré de cohérence suffisant pour lui permettre d'inclure un programme pour la recherche future
- Un programme de recherche doit être fécond, c'est-à-dire mener de temps en temps à la découverte de phénomènes nouveaux

Tout programme de recherche n'est pas scientifique au sens donné ci-dessus. Lakatos cite le marxisme et la psychologie freudienne comme exemples de programmes de recherche qui vérifient la première condition mais pas la seconde, et la sociologie moderne comme exemple d'un programme qui vérifie la seconde condition mais pas la première.

La position de Lakatos concernant le caractère universel des critères de scientificité est clairement proche du rationalisme (au sens défini au chapitre 2). Lakatos rejette vigoureusement l'approche relativiste de la science. Pour lui, le problème central de la philosophie des sciences est celui de l'établissement de critères universels de scientificité. Autrement, il n'y aurait « *pas de moyen de juger une théorie autrement qu'en évaluant le nombre, la foi et la puissance vocale de ses partisans* ». La conséquence serait qu'alors la « *vérité se trouverait dans le pouvoir* » et l'évolution de la science se réduirait à un phénomène de psychologie des foules. Il faut toutefois noter, et Lakatos le reconnaissait lui-même, que ce critère de scientificité est un critère a posteriori, au sens où il permet seulement de dire si un programme de recherche déjà mature est scientifique ou non. Il n'offre pas de guide aux participants qui leur permette d'assurer le caractère scientifique de leurs travaux au démarrage du programme.

3.3 Définir un objet

Les sciences manipulent des objets théoriques qui sont à la base de leurs énoncés. Ainsi, la mécanique newtonienne utilise des objets comme la masse, la force, l'énergie, la quantité de mouvement. La physique moderne fait intervenir des ondes, des atomes, pendant que la chimie introduit d'autres objets comme le pH. Ces notions sont familières pour toute personne possédant un bagage scientifique élémentaire dans ces disciplines scientifiques. Mais que représentent-elles ? Et comment les sciences empiriques définissent-elles leurs objets ? Nous allons voir dans ce qui suit que la réponse à ces questions est loin d'être triviale.

Le mot **définir** apparaît à la fin du XIIe siècle. Il est dérivé du latin **definire** qui signifie « déterminer », « délimiter », « fixer ». Il est originellement utilisé au sens de

« préciser, faire connaître ». Il ne prend le sens de « déterminer exactement » qu'à la fin du XV^e siècle.

L'action de définir consiste donc à exprimer le plus clairement possible ce que le mot défini signifie par rapport à ce qu'il ne signifie pas. Nous allons voir que la nature de la définition varie suivant le type d'objet que l'on cherche à identifier. Nous mettrons ici l'accent sur la définition des groupes et des familles d'objets, et sur celui des objets employés pour bâtir les théories physiques à partir des observations expérimentales.

3.3.1 Généralités : définition extensionnelle et définition analytique/intensionnelle

Lorsque l'on cherche à définir une famille ou un groupe d'objets, ceci peut être réalisé au moyen de deux types de définitions :

1. La **définition extensionnelle**, qui consiste à énumérer tous les objets auxquels se rapportent la définition. Par exemple, la famille des planètes du système solaire est définie comme

$$\text{planètes} = \{\text{Mercure, Vénus, Terre, Mars, Jupiter, Saturne, Uranus, Neptune}\}$$

Comme on peut le voir, ce type de définition est parfaitement univoque : on peut vérifier de façon simple si un astre fait oui ou non partie des planètes du système solaire. Il suffit pour cela de balayer l'ensemble de la liste, et de voir si l'élément est contenu dans cette liste ou non. Mais cette définition est toutefois peu intéressante, car elle ne nous dit rien de ce qu'ont en commun les éléments de cette liste : nous connaissons les planètes, mais nous ne savons pas leur point commun qui a fait qu'il a été jugé utile de les regrouper au sein d'une même famille. Elle ne nous indique pas pourquoi Pluton, découverte en 1930 par l'astronome américain Clyde Tombaugh, a été exclue de la liste des planètes du système solaire par l'Union Astronomique Internationale le 24 août 2006. Un autre handicap de cette approche est qu'elle est très peu commode pour les familles comportant un grand nombre, voire une infinité d'objets. C'est par exemple le cas de nombreux ensembles rencontrés en mathématiques. Considérons le cas de l'ensemble des entiers naturels noté \mathbb{N} , qui contient une infinité dénombrable d'éléments. Une tentative de définition extensionnelle serait de la forme

$$\mathbb{N} = \{0, 1, 2, 3, 4, 5, \dots\}$$

Les trois points indiquent ici qu'ils existent d'autres éléments qui, pour des raisons pratiques, ne peuvent pas être écrits dans cette liste : puisqu'il y a une infinité d'entiers naturels, on ne peut en dresser une liste nécessairement finie. Cette impossibilité, outre son aspect théorique, est facilement illustrée par le fait suivant : on estime de 10^{80} à 10^{87} le nombre total de particules élémentaires (protons, neutrons, électrons) dans l'univers observable, et à

environ 10^{79} le nombre d'atomes²⁶. En associant une particule à chaque nombre, on voit que la limite atteignable est le nombre de particules. Il est par exemple impossible de poursuivre la liste jusqu'au Googol, qui est le nom inventé à sa demande en 1938 par le neveu alors âgé de neuf ans du mathématicien américain Edward Kasner (1878-1955) pour le nombre 10^{100} . Ceci sans parler du Googolplex, défini comme 10^{Googol} !

Pour palier ce problème, les mathématiques (il n'existe pas d'équivalent pour les objets des sciences empiriques) offrent la possibilité de définir de tels ensembles en partant de leur règle de construction. Le mathématicien Guiseppe Peano (1858-1932) a ainsi donné la définition de l'ensemble des entiers naturels à l'aide d'une suite d'axiomes :

Axiomes de Peano

Axiome 1 : 0 est un entier naturel

Axiome 2 : Tout entier naturel possède un successeur qui est lui aussi un entier naturel

Axiome 3 : 0 n'est le successeur d'aucun entier naturel

Axiome 4 (dit axiome de récurrence) : Tout sous-ensemble de \mathbb{N} qui contient 0 et qui est tel que le successeur de chaque élément de ce sous-ensemble appartient également à ce sous-ensemble est identique à \mathbb{N} .

On le voit, cette nouvelle forme de définition extensionnelle est différente de la précédente en ce sens qu'elle n'offre pas une liste exhaustive explicite (au sens où tous les éléments sont écrits individuellement sur la page) mais une liste implicite, puisqu'elle donne le mode de construction de l'ensemble.

2. La **définition intensionnelle**²⁷, encore appelée **définition analytique** ou encore **définition substantielle**, qui consiste à donner la liste des critères (appelés **critères définitoires**) que doit vérifier un objet pour correspondre à la définition. Ce type de définition possède le double avantage de donner des informations sur les points communs entre les différents membres d'une même famille, mais permet également de définir des familles possédant une infinité de membres. De plus, elle permet d'augmenter les listes en y incorporant de nouveaux objets qui étaient inconnus lorsque la famille a été définie. Une telle définition met en relation une expression qui est définie (*definiendum*) et une expression qui définit (*definiens*). Pour qu'elle soit valide, il faut qu'elle remplisse les trois conditions suivantes

Rappel : critères de validité d'une définition intensionnelle

1. **Adéquation extensionnelle** : le *definiens* doit s'appliquer à tout ce que le *definiendum* désigne, et rien d'autre. Cela se traduit par des relations du type « si et seulement si » : A est B ssi ...
2. **Adéquation intensionnelle ou conceptuelle** : le *definiens* doit donner la nature de la chose définie
3. **Non-circularité** : une définition ne peut pas faire appel à elle-même.

²⁶ Le calcul est le suivant. L'univers est essentiellement composé d'hydrogène. Pour obtenir le nombre total d'atomes, on effectue le produit suivant : nombre d'atomes d'hydrogène par gramme (nombre d'Avogadro, à peu près égal à 10^{24}) x masse du Soleil (environ 10^{33} grammes) x nombre moyen d'étoiles dans une galaxie (10^{11}) x nombre de galaxies (10^{11})

²⁷ Cette notion a déjà été discutée dans le chapitre 1. Pour simplifier la lecture, on reprend ici les principaux éléments associés à cette notion.

Illustrons ces critères. La définition « A est un homme si et seulement si A est capable de rire » est-elle valide ? Non, car elle n'est pas conceptuellement adéquate puisqu'elle ne nous dit pas ce qu'est un homme. De plus, les travaux sur les grands singes ayant montré que le rire existe par exemple chez le chimpanzé, elle ne vérifie pas non plus le critère d'adéquation extensionnelle. Passons maintenant à la proposition « A est un homme si et seulement si A est un homme ». Cette définition vérifie les deux premiers critères, mais pas le troisième : elle est circulaire, et donc ne nous apprend rien.

Reprenons l'exemple de la définition des planètes. La définition analytique de celles-ci adoptée le 24 août 2006 par l'Union Astronomique Internationale est :

planète = « un corps céleste, qui est en orbite autour du Soleil, a une masse suffisante pour que sa gravité l'emporte sur les forces de cohésion du corps solide et le maintienne en équilibre hydrostatique, sous une forme presque sphérique, a éliminé tout corps susceptible de se déplacer sur une orbite proche. »

Cette définition illustre parfaitement les avantages de la définition analytique. Il permet aussi de voir que, dans certains cas, il est difficile de trouver une définition parfaitement stricte des membres de la famille : certains critères sont parfois flous (forme « presque sphérique »). Certaines définitions analytiques admettent donc une marge d'incertitude, et se basent sur un « air de famille ». Ceci est d'autant plus vrai que les objets considérés sont complexes. Il est par exemple très difficile de définir exactement ce qu'est un « être vivant ». Mais cela explique la disparition de Pluton de la liste extensionnelle dressée depuis 1930, puisqu'elle ne vérifie pas le dernier critère. Pluton est maintenant classée parmi les planètes naines²⁸.

La différence avec la construction axiomatique de l'ensemble des entiers naturels est subtile : dans le cas de la proposition de Peano, les entiers naturels ne sont pas décrits par un ensemble de caractéristiques, mais leur ensemble est construit, ce qui revient à une liste exhaustive définie de manière itérative à partir de l'élément zéro et de la règle qui fait passer d'un entier à son successeur.

Les critères pour obtenir une bonne définition analytique sont les suivants : la définition et l'objet qu'elle définit doivent être substituables dans les énoncés sans modifier le contenu ou la qualité de ceux-ci et la définition ne doit pas être circulaire, c'est-à-dire faire appel à un autre terme dont la définition renverrait à la première.

²⁸ La définition analytique des planètes naines retenue par l'UAI le même jour est : « *une planète naine un corps céleste, qui est en orbite autour du Soleil, a une masse suffisante pour que sa gravité l'emporte sur les forces de cohésion du corps solide et le maintienne en équilibre hydrostatique, sous une forme presque sphérique, n'a pas éliminé tout corps susceptible de se déplacer sur une orbite proche, n'est pas un satellite.* » On connaît 3 planètes naines dans le système solaire : Eris (diamètre = 2400 ±100 km, masse = 1,67²² kg), Pluton (diamètre = 2306 ±20 km, masse = 1,30²² kg) et Cérés (diamètre ≈1000 km, masse = 9,5²⁰ kg) . Des dizaines d'autres corps célestes pourraient rejoindre cette famille.

3.3.2 Une mode définitoire propre aux sciences empiriques : la définition opératoire

Les sciences empiriques ne trouvent pas leurs objets « tout faits » dans la nature : la définition est une opération de construction intellectuelle. Partant de cette constatation, et du fait que c'est par la mesure que les sciences empiriques accèdent au monde, le physicien et philosophe Percy Williams Bridgman (1882-1961, prix Nobel de Physique 1946) proposa, dans ses ouvrages *The logic of modern physics* (1927) et *Nature of physical theory* (1936) le concept de **définition opératoire** (qui traduit le terme anglais **operational definition**). Selon Bridgman, dans les sciences empiriques, c'est la mesure, c'est-à-dire l'ensemble des opérations qu'il faut effectuer pour obtenir la valeur de la quantité mesurée, qui définit l'objet de la mesure : « *La vraie définition d'un concept ne se fait pas en termes de propriétés, mais en termes d'opérations effectives.* » En conséquence, Bridgman proposait pour objectif pour les sciences empiriques de n'utiliser que des objets définis de manière opératoire. Cette vision est à rapprocher de celle défendue par le philosophe français Gaston Bachelard (1884-1962), qui, dans son livre *Le nouvel esprit scientifique* (1934) disait « *l'expérience fait donc corps avec la définition de l'Être. Toute définition est une expérience* » et résumait cette conception par la formule « *Dis-moi comment l'on te cherche, je te dirais qui tu es.* »



Figure 13: P.W. Bridgman

En reprenant les termes de Jean Ullmo dans son ouvrage *La pensée scientifique moderne* (1969), on peut dire que « *une définition opératoire est une définition qui comporte la description d'un procédé régulier pour repérer, mesurer, plus généralement atteindre et identifier le concept défini.* » Cette définition met en lumière une caractéristique importante du processus de mesure : il doit être régulier, c'est-à-dire pouvoir être répété par toute personne le désirant et conduire au même résultat. En anticipant sur le chapitre consacré à la mesure, on peut dire que le procédé à la base d'une définition opératoire doit être fondé sur des corrélations répétables. Cette régularité est nécessaire pour pouvoir définir des objets stables, utilisables pour les théories scientifiques. Ce besoin de répétition, de stabilité, entraîne une sélection des objets des sciences empiriques. Toute séquence d'opérations de mesure ne conduit pas à la définition d'un objet acceptable pour le scientifique. Le physicien et philosophe allemand Ernst Mach (1838-1916) a exprimé cela en écrivant qu'il ne s'agit pas de « *classer et de rassembler des faits qui sont individuellement donnés ; le savant doit avant tout trouver les caractères dont il faut tenir compte.* »

Le programme de Bridgman (bâtir les sciences empiriques sur les seules définitions opératoires) se heurte toutefois aux limitations intrinsèques à cette approche. Notamment, on peut noter que

- L'objet étant défini par un procédé de mesure, il se pose la question de sa persistance si l'on modifie le procédé de mesure (par exemple en employant des instruments différents).
- La répétabilité stricte est une abstraction de l'esprit : une expérience est toujours unique et ne peut pas être exactement reproduite, car l'ensemble des

interactions des composants du système de mesure avec le reste univers est toujours changeante. Par exemple, une pesée effectuée avec une balance de Roberval de donnera pas exactement la même valeur si elle est répétée à deux jours d'intervalle, car les forces d'attraction de la lune et du soleil auront changé (mais rassurons-nous, ces différences seront infimes et les variations de la mesure tout à fait négligeables pour la vie pratique et la plupart des besoins des scientifiques). Une certaine marge d'incertitude doit être incorporée à la définition opératoire.

- Quel est le degré d'identité de deux objets définis par deux procédés différents ? On peut par exemple définir la masse de plusieurs manières différentes, tout en espérant parler de la même chose. A quel point cela est-il possible en utilisant des définitions opératoires ?
- La définition opératoire indique comment il faut définir les objets et donne certains critères de recevabilité, mais ne dit pas quels sont les objets à définir. C'est sur d'autres considérations que le scientifique doit se baser pour trouver ses objets.

3.3.3 Exemple 1 : Qu'est-ce que la masse ? Qu'est-ce qu'une force ?

Pour illustrer les propos précédents, considérons tout d'abord le problème de la définition de la masse inertielle (c'est la masse qui apparaît dans la mécanique de Newton). Comment définir ce concept pourtant familier à tout scientifique ? Nous allons voir que cela est loin d'être évident, car la masse ne peut pas être définie de manière isolée.

Ce problème est abordé par le grand mathématicien et épistémologue français Henri Poincaré (1854-1912), dans son livre *La science et l'hypothèse* (1902). A propos de la possibilité de valider expérimentalement la seconde loi de Newton :

$$\sum \vec{F} = m\vec{\gamma}$$

où m désigne la masse inertielle, il écrit : « Cette loi peut-elle être vérifiée par l'expérience ? Pour cela, il faudrait mesurer les trois grandeurs qui figurent dans l'énoncé : accélération, force et masse.

J'admets qu'on puisse mesurer l'accélération, parce que je passe sur la difficulté provenant de la mesure du temps. Mais comment mesurer la force, ou la masse ? Nous ne savons même pas ce que c'est.

Qu'est-ce que la masse ? C'est, répond Newton, le produit du volume par la densité. – Il vaudrait mieux dire, répondent Thomson et Tait, que la densité est le quotient de la masse par le volume. – Qu'est-ce que la force ? C'est, répond Lagrange, une cause qui produit le mouvement d'un corps ou qui tend à le produire. – C'est, dira Kirchhoff, le produit de la masse par l'accélération. Mais alors, pourquoi ne pas dire que la masse est le quotient de la force par l'accélération ?

Ces difficultés sont inextricables. »

Et de conclure que « les masses sont des coefficients qu'il est commode d'introduire dans les calculs » et que « c'est par définition que la force est égale au produit de la masse par l'accélération ». Outre les implications profondes sur la possibilité de

vérifier la seconde loi de Newton par l'expérience (Poincaré conclut négativement, puisqu'il s'agit ici pour lui d'une définition de la force – ou de la masse – qui échappe à la vérification), ce texte fait clairement ressortir le problème de la circularité évoqué plus haut concernant la définition de la masse et de la force : on ne peut définir ces deux concepts isolément.

A cette conclusion, Poincaré ajoute que ce qui est important, ce n'est pas de proposer des définitions en soi de la masse ou de la force, mais de donner des définitions qui permettent de les mesurer, se rapprochant ainsi du concept (formalisé plus tard par Bridgman) de définition opératoire.

La définition opératoire de la masse au moyen de la machine d'Attwood²⁹ est décrite par J. Ullmo (*La pensée scientifique moderne*, 1969). La machine d'Attwood est un dispositif simple qui permet de mesurer la masse d'un solide. Elle est composée (voir sur la figure) d'un fil et d'une poulie. A un bout du fil est attaché le solide dont on cherche à déterminer la masse, et à l'autre un solide de référence qui assure que le fil reste tendu. Le procédé de mesure est le suivant. On se munit d'un grand nombre (en théorie, un nombre infini) de solides (même taille, même composition, même forme que le solide dont on désire connaître la masse). On attache ensuite le solide au fil, et on le maintient en position basse. A l'instant initial, on lâche le solide, et on mesure le temps qu'il lui faut pour atteindre la position haute de référence. En faisant l'hypothèse que le solide est soumis à une accélération constante, on déduit celle-ci du temps de parcours. Puis on répète l'expérience en ajoutant un, puis deux, puis trois solides identiques au premier considéré, et l'on mesure les accélérations pour ces solides composés. On constate alors qu'il existe une relation répétable : le produit de l'accélération par le nombre de solides garde une valeur constante, si la différence de poids entre le solide de référence celui que l'on veut peser est très petite (c'est ce que l'on appelle une *hypothèse de linéarisation*). La constance de ce produit indique l'existence d'une propriété permanente de notre système, qui sera nommée (c'est une définition opératoire !) tension du fil ou encore la force qu'il exerce sur les solides et dont la valeur est mesurée par la valeur de la constante. Cette valeur est un paramètre, car elle dépend du solide choisi initialement : si l'on change sa composition, ou son volume, on constatera encore une fois l'existence d'une valeur constante, mais différente de la précédente.

²⁹ George Attwood (ou Atwood) (1746-1807), mathématicien anglais. La première description de sa machine date de 1784.

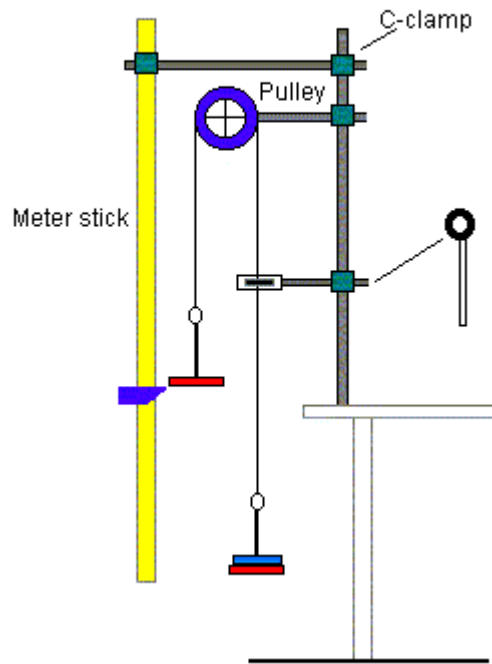


Figure 14: principe de la machine d'Atwood

Pour atteindre la masse du solide, il faut procéder à un second jeu d'expériences. On considère maintenant un seul exemplaire du solide dont on veut connaître la masse, et on fait varier le nombre de solides de référence utilisés pour faire contrepoids (ici aussi, on considère une collection de solides identiques). On mesure, pour chaque cas (un, deux, ... solides de référence) l'accélération du solide. On observe l'existence d'une nouvelle relation répétable : l'accélération du solide est directement proportionnelle au nombre de solides de référence employés (toujours en faisant appel à l'hypothèse de linéarisation). Cette répétabilité indique l'existence d'une propriété permanente de ce qui est resté inchangé, à savoir le solide soumis à l'analyse. Cette propriété sera définie comme étant la masse du solide, et sa valeur sera donnée par le rapport de proportionnalité entre accélération et nombre de solides de référence.

On le voit, cette définition (opératoire) est entièrement basée sur le procédé expérimental (et souffre des problèmes liés à toute expérience concrète), et ne dit rien sur une possible nature intrinsèque de la masse ou de la tension. Ces définitions évitent le caractère circulaire de définitions non-opératoires basées sur la formule de Newton, mais elles impliquent, de manière sous-jacente, un certain nombre d'hypothèses, qui seront détaillées dans le chapitre consacré à la mesure. On peut, en utilisant ce procédé, mesurer la masse de tout corps solide, si l'expérience est réalisable en pratique. Pour mesurer la masse d'une planète, ou d'une étoile, il faut bien évidemment recourir à d'autres définitions.

3.3.4 Exemple 2 : Qu'est-ce que le mouvement ? Qu'est-ce qu'une déformation ?

Considérons maintenant le problème de la définition du mouvement. Comment définir le mouvement ou, de manière équivalente, quand peut-on dire qu'un corps solide c'est déplacé ? Nous allons voir que, pour un solide de dimensions finies (donc, pas un

point matériel dont le cas sera évoqué en fin de paragraphe), la notion de déplacement est indissociable de celle de déformation.

Un premier écueil est que le mouvement (du moins tel qu'il est accepté aujourd'hui) ne se conçoit que comme relatif à un repère de référence : c'est le relativisme galiléen. Il n'y a pas de mouvement absolu, pas plus qu'il n'y a pour le mécanicien de repère absolu. Dans ce qui suit, on choisit un repère, pris comme référence. Dans ce repère, quand peut-on dire qu'un solide identifié, pris pour simplifier de volume constant, s'est déplacé ?

Considérons le cas d'un solide sphérique, de masse volumique uniforme (le centre de gravité du solide est donc confondu avec le centre géométrique de la sphère), initialement au repos dans le repère considéré, et appliquons-lui une force non-nulle à l'instant initial. Que se passe-t-il ?

Dans le cas d'un solide indéformable (encore appelé solide rigide), la déformation est par définition nulle, et le solide se mettra en mouvement, conformément à la seconde loi de Newton, qui indique que l'accélération initiale sera non-nulle. C'est-à-dire qu'il existera au moins un temps après le temps initial pour lequel au moins un point du solide n'occupera pas sa position initiale. Il faut prendre garde au fait que mouvement du solide n'est pas égal à déplacement de son centre de gravité : dans le cas d'une rotation autour de son centre de gravité, il y a bien mouvement de la sphère alors que son centre de gravité est fixe. Voici donc une définition qui semble satisfaisante, valable pour les solides indéformables et les points matériels (ces derniers étant de dimension nulle, il ne peuvent se déformer !).

Mais qu'en est-il de la définition de la déformation, qui est implicitement utilisée dans celle du mouvement donnée précédemment ? Considérons maintenant un solide déformable, et appliquons lui une force : il va se déformer, c'est-à-dire qu'il existe au moins un temps situé après le temps initial auquel la distance entre au moins deux points du solide aura varié par rapport au temps initial. Pour que la distance varie, il faut qu'au moins un des points se soit déplacé par rapport aux autres : on fait donc appel ici à la notion de déplacement, et apparaît le problème de circularité. De plus on constate que la définition « naïve » donnée plus haut du déplacement pour les solides rigides ne permet plus de différencier déplacement et déformation pour un solide déformable.

On peut rendre les définitions plus rigoureuses en procédant comme suit. Tout d'abord, il convient de distinguer les solides rigides et leurs mouvements.

On dira qu'un corps évolue comme un solide rigide si, à tout instant, la distance entre chacun de ses points demeure inchangée. Mathématiquement, cela se traduit par le fait que les positions de chacun de ses points à un instant donné peuvent être déduites de celles occupées à un autre instant au moyen de la combinaison d'une rotation et d'une translation. On dira que ce solide rigide s'est déplacé si cette rotation ou cette translation est non nulle.

Un solide sera dit déformable si il n'est pas rigide. Dans ce cas, déformation et déplacement ne sont pas dissociables, car la déformation implique le déplacement d'au moins certaines parties du solide. Il est possible alors de décomposer arbitrairement (c'est-à-dire que cette décomposition n'est pas unique) l'évolution à un

instant donné comme la somme d'un déplacement de corps rigide et d'une déformation. Si il existe deux points dont la distance relative n'a pas changée par rapport à l'instant initial, on peut identifier la rotation et la translation nécessaires pour établir la correspondance avec leurs positions initiales. La déformation est alors définie comme l'opération qu'il faut ajouter à cette rotation et cette translation pour obtenir la position de l'ensemble des points du solide.

Dans le cas où il n'existe pas deux points dont la distance reste fixe, on peut se contenter de choisir un point de référence (par exemple : le centre de gravité), de trouver la translation qui permet d'associer sa position à sa position de référence, et de définir la déformation comme la transformation complémentaire à cette translation.

Cet exemple montre que, même pour des notions qui paraissent intuitives et qui sont d'emploi courant en mécanique et en physique, il est très difficile de parvenir à des définitions objectives qui ne posent pas de problèmes. Le mouvement est parfaitement défini pour des solides rigides strictement indéformables, ou pour l'abstraction que représente le point matériel. Mais pour des corps « réels », la définition garde une part de subjectivité.

Notons enfin que la notion de mouvement est une notion intrinsèquement relative : un corps se déplace par rapport à quelque chose. C'est le **principe de relativité**, parfois appelé **principe de relativité de Galilée**. Un problème fondamental de la mécanique newtonienne est de savoir si il existe des référentiels privilégiés pour définir le mouvement, et, si oui, de savoir comment les identifier. Cette question, qui a été très longuement débattue à la fin du XIXe siècle pour déboucher sur la théorie de la relativité d'Einstein et ses prolongements, est abordée dans le dernier chapitre.

La relativité du mouvement est aisément illustrée par l'exemple suivant. Considérons un lecteur de ce paragraphe, assis sur une chaise. A la fin de sa lecture (qui dure environ une minute), sur quelle distance se sera-t-il déplacé ? Une distance nulle par rapport à la chaise. Mais la Terre tourne autour de son axe à la vitesse de 0,5 km/s (soit 1800 km/h) à l'équateur³⁰. En 1 minute, notre lecteur aura parcouru 30 km. De plus, la Terre tourne autour du Soleil à la vitesse moyenne³¹ de 29,7 km/s (soit 106 920 km/h). En 1 minute, notre lecteur s'est donc déplacé de 1782 km ! On peut poursuivre le raisonnement, en prenant en compte que notre système solaire tourne comme le reste de notre galaxie (la Voie Lactée) à la vitesse³² de 217 km/s (soit 781 200 km/h). Donc, en une minute, le trajet parcouru dans le repère galactique par notre lecteur est égal à 13020 km. On pourrait y ajouter encore le mouvement propre du système solaire par rapport à la galaxie : 19,5 km/s (soit 70 200 km/h) en direction de la constellation d'Hercule. Notre lecteur est-il immobile ?

3.4 Définir un objet (suite) : complexité et frontières floues

Les exemples précédents portaient sur des notions qui pouvaient apparaître comme « simples » et couramment utilisées en mécanique et en physique, c'est-à-dire par les

³⁰ A l'équateur, la distance circonférence terrestre est égal à 40 000 km et la période de la rotation est de 24h, soit 86400 secondes.

³¹ L'orbite terrestre mesure 924 375 700 km, et la période de rotation est de 365,25 jours (soit 31 557 600 secondes)

³² le système solaire est à environ $2,5 \cdot 10^{17}$ km du centre de la galaxie, et effectue une rotation complète en $2,26 \cdot 10^8$ années

deux sciences empiriques parmi les plus anciennes et les plus mathématisées qui soient. Pourtant, comme nous l'avons vu, il est très difficile de parvenir à des définitions parfaitement satisfaisantes et objectives.

La science moderne fait apparaître des notions et des objets **complexes**, qui rendent parfois impossible la production de définitions exactes et satisfaisantes. Ce type de problème est particulièrement présent en biologie, qui, par définition, traite d'objets extrêmement complexes : les êtres vivants³³. Comme nous allons le voir par la suite, on ne possède plus de définition exacte pour des concepts fondamentaux comme vivant/inerte, mort/vivant ou encore animal/végétal. Le problème est que pour chacune de ces distinctions, qui sont a priori intuitives et fondent la réflexion en biologie, des « contre-exemples » ont été trouvés, qui sont tels qu'ils ne peuvent plus être rangés dans une seule catégorie.

Avant de discuter ces exemples, notons que d'autres disciplines, comme la physique, connaissent le même problème : il existe des « états indécis » de la matière, qui n'entrent pas dans les catégories classiques. Par exemple, le verre obtenu lors d'une transition liquide-solide n'est pas un solide classique, et représente un état intermédiaire entre un liquide et un cristal. Ceci est dû au fait que lorsqu'on refroidit (rapidement) un liquide surfondu visqueux, les molécules sont figées avant d'avoir pu atteindre un état d'équilibre : le matériau est donc hors d'équilibre. Il existe d'autres exemples : les cristaux liquides ou encore les gels.

Par ailleurs, les mathématiciens ont développé des **logiques floues**, dans lesquelles une propriété n'est plus ni vraie ni fausse, mais possède un **degré de véracité** compris entre 0 et 1 (la réponse booléenne vrai/faux, transcrite par une réponse binaire 0/1, est remplacée par une fonction continue qui varie entre 0 et 1).

3.4.1 Concepts vagues et paradoxes sorites

Une difficulté particulière est associée aux **concepts** dits **vagues** ou **flous**, qui sont, comme leur nom l'indique, des concepts qui échappent à une définition précise. Le langage commun comprend un nombre important de concepts de ce type, qui ne semblent pas poser de problème au premier abord. Toutefois une analyse détaillée montre que de tels concepts induit l'existence de paradoxes appelés **paradoxes sorites**³⁴ lorsqu'on leur applique l'induction logique classique.

Le plus connu est le **paradoxe du tas de sable**. Comment définir un tas de sable ? Pour cela, partons des deux prémisses ci-dessous :

1. Un tas de sable est composé d'un grand nombre de grains de sable
2. Un tas de sable auquel on enlève un grain de sable reste un tas de sable

³³ Notons que cette complexité est d'ordre fonctionnel et structurel. Le corps humain comprend environ 10^{28} atomes, qui forment environ 10^{24} molécules. Celles-ci composent 10^{13} cellules vivantes. La complexité est aussi anatomique : 1 cm² de peau est formé par environ 3 millions de cellules, et comprend 3 vaisseaux sanguins, une dizaine de pores, une douzaines de nerfs, une quinzaine de glandes sébacées et une centaine de glandes sudoripares ; le cerveau est formé par environ 10^{11} neurones.

³⁴ sorite vient de la racine grecque *soros*, qui veut dire « tas », en l'honneur du paradoxe du tas de sable déjà discuté par les philosophes de la Grèce antique.

Considérons maintenant un tas de sable. Celui-ci contient un nombre grand mais fini de grains, que nous noterons N . Appliquons ensuite $(N-1)$ fois le prémisses 2 : on obtient la conclusion qu'un unique grain de sable forme un tas de sable, ce qui est en contradiction avec le prémisses 1 ! La contradiction vient du fait que la notion de « grand nombre » est-elle même vague, ce qui ne permet pas d'invalider l'application répétée du prémisses 2. Comment lever ce paradoxe ? Plusieurs solutions ont été proposées durant le deux derniers millénaires. Parmi celles-ci :

- Les **solutions triviales**. Une première consiste à dire que le concept « tas de sable » n'a pas de sens, puisqu'on ne peut pas le définir, et qu'il ne doit donc pas être employé. Une seconde solution, préconisée par Bertrand Russell, est d'interdire l'emploi de la logique pour manipuler les concepts vagues.
- **Fixer un seuil arbitraire**. Par exemple, on peut décider qu'un tas de sable est fixé lorsque 100 000 grains sont réunis. Cette solution, qui est souvent employée en pratique, n'est toutefois pas satisfaisante sur le plan philosophique. En effet, quelle est la différence entre un ensemble de 99 999 grains de sable et un tas de 100 000 grains ?
- L'emploi de la **logique floue**, qui permet de passer de « n'est pas un tas de sable » à « est un tas de sable » de manière continue, et donc de définir une infinité de réponses intermédiaires du type « est presque un tas de sable » ou encore « est probablement un tas de sable ». En levant l'axiome aristotélicien du tiers exclu (voir la section 5.2.1), on résout le problème des paradoxes sorites.
- Le **consensus au sein du groupe** des personnes engagées dans la réflexion sur le problème posé. La probabilité qu'un ensemble de grains de sable forme un tas de sable dépend donc de la répartition des avis au sein du groupe. On retrouve ici une notion de vérité-consensus. Il est important de noter que chaque individu peut avoir sa propre définition de ce qu'est un tas de sable : seul le résultat du « vote » compte. Ceci illustre comment des groupes humains peuvent s'accorder sur l'emploi de termes vagues (justice, démocratie, ...).

3.4.2 Exemple 1 : Vivant ou inerte ? Mimivirus et ATTV³⁵

La biologie étudie, par définition, les êtres vivants. Une question fondamentale est donc celle de la définition du vivant. Qu'est-ce qu'un être vivant ? Ou, autrement formulé, quelle est la différence entre le vivant et l'inerte ? Nous allons voir que cette question est beaucoup plus problématique que les précédentes. Outre son caractère strictement scientifique, cette question porte une très forte charge affective et émotionnelle. Elle a de très fortes implications philosophiques, religieuses, éthiques ou encore sociales et politiques. La réflexion sur la définition de la vie est donc, par construction, beaucoup plus subjective et sujette aux débats extra-scientifiques que celles évoquées plus haut dans le domaine de la physique. La distinction entre le vivant et l'inerte semble a priori simple : un éléphant ou un arbre sont des êtres vivants, un caillou ne l'est pas. Malheureusement, le problème est beaucoup plus complexe lorsqu'il s'agit de donner une définition scientifique du vivant, qui

³⁵ Source : « Les virus, inertes ou vivants ? », A. Saïb, Pour la science No 350, 2006

s'applique à tous les êtres rencontrés. Si, le problème ne se pose pas pour les grands organismes (les êtres multicellulaires) pour lesquelles l'identification est simple, il est rencontré à l'échelle microscopique. Rappelons tout d'abord qu'il n'existe pas aujourd'hui de définition unique de ce qu'est un être vivant. Toutefois, les différentes définitions ont des points communs, rappelés ci-dessous :

Critères de définition du vivant

Une entité est vivante si :

- Elle peut utiliser ou créer de l'énergie
- Elle se reproduit
- Elle échange avec le milieu extérieur
- Elle répond à des stimuli
- Elle croît
- Il peut exister une variabilité génétique qui distingue un parent de ses descendants

Ceci est à rapprocher de la première esquisse d'une définition générale de la vie, qui est due à Aristote (384-322) dans son *Traité de l'Âme* (II,1) : « *Nous entendons par vie le fait de se nourrir, de croître et de dépérir par soi-même* ».

Depuis la fin des années 1970, à la suite des travaux du microbiologiste américain Carl Woese, l'ensemble des êtres vivants est regroupé selon trois règnes (au lieu de deux précédemment):

Les 3 règnes du vivant

- Les **eucaryotes**, qui possèdent une structure intracellulaire complexe, qui sépare le cytoplasme du nucléoplasme où baigne l'ADN
- Les **procaryotes** (les bactéries), qui ne possèdent pas de noyau
- Les **archéobactéries** ou **archées** (depuis la fin des années 1970), qui ne possèdent pas de noyau mais qui présentent des différences de structure de leur membrane telles qu'elles ne peuvent être regroupées avec les procaryotes.

Les **virus** (du latin *virus* : poison) sont des entités complètement différentes. Découverts chez les plantes à la fin du XIX^e siècle, puis chez les animaux et les humains au début du XX^e siècle, ils sont constitués d'un acide nucléique (ADN ou ARN) protégé par une structure protéique appelée capsid. Jusqu'à très récemment, les virus connus ne pouvaient pas être classés comme des êtres vivants car :

- Ce sont des parasites obligatoires, qui ne se multiplient qu'au sein d'une cellule vivante hôte, à laquelle ils imposent une série de tâches qui la condamne à fabriquer des virus. Certains virus, dits endogènes, s'insèrent dans le patrimoine génétique de l'hôte et sont transmis à chaque génération de descendants de l'hôte. Notons que le parasitisme obligatoire n'est pas un argument suffisant, puisque certaines bactéries sont totalement dépendantes d'une cellule hôte.

- Ils sont incapables de synthèse protéique
- Ils ne peuvent croître en dehors du contexte cellulaire de l'hôte
- Ils ne contiennent qu'un seul type d'acide nucléique (ADN ou ARN), alors que les deux types sont présents chez tous les êtres vivants
- Ils n'ont aucun métabolisme propre, l'énergie requise pour leur multiplication étant fournie par l'hôte.
- Mais ils sont susceptibles de variabilité génétique (ce qui fait que certains médicaments deviennent inefficaces)

Des découvertes récentes ont placé les virus dans une sorte de *no man's land* entre vivant et inerte, ce qui rend floue la définition du vivant.

En 1992, un virus géant nommé Mimicking Microbe Virus (*Mimivirus*) a été isolé. Ses caractéristiques hors normes ont fait qu'il a d'abord été considéré comme une nouvelle bactérie aux caractéristiques étranges. Ce n'est que beaucoup plus récemment, en 2003, qu'une équipe de chercheurs français (Didier Raoult et ses collègues à Marseille) l'a enfin identifié comme étant un virus. Mimivirus possède de nombreuses propriétés qui remettent en cause la séparation entre les virus et les êtres vivants. Ce virus est d'une taille (400 nanomètres de diamètre) supérieure à celle de certaines bactéries, et possède un génome plus grand et plus complexe que plusieurs dizaines d'entre elles. Plusieurs gènes du Mimivirus semblent participer à la synthèse protéique et à la réparation de son propre patrimoine génétique. Enfin, cotoyant l'ADN du virus, des ARN ont été détectés. Plus surprenant encore : alors que l'on pensait que les gènes viraux avaient une origine cellulaire (une sorte de « vol génétique » chez les hôtes au cours de l'évolution), de nombreux gènes du Mimivirus n'ont aucun homologue chez les êtres vivants connus. Ce virus géant serait donc le premier exemple découvert d'une nouvelle famille qui serait apparue sur Terre en même temps que la vie elle-même, il y a environ quatre milliards d'années. Ce qui pousse certains biologistes à proposer la création d'une quatrième branche dans l'arbre du vivant : les Girus (pour Giant Virus).

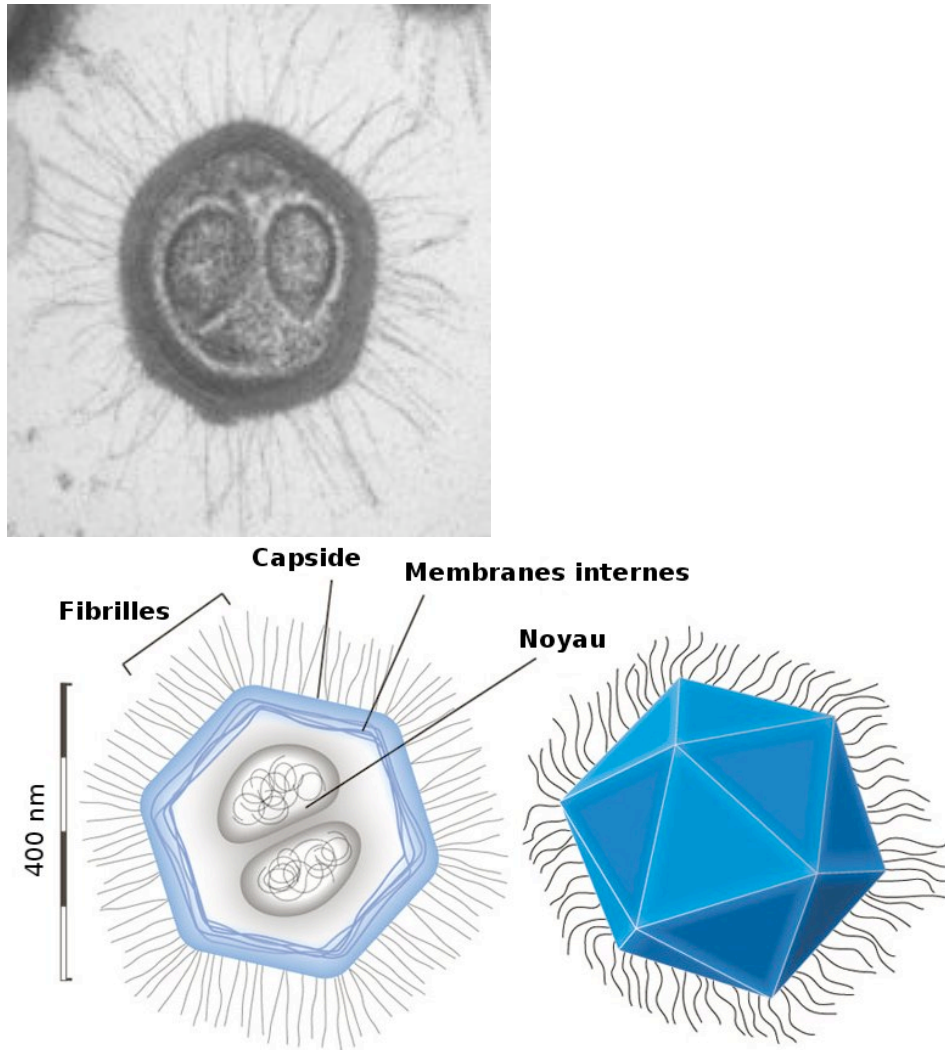


Figure 15: Mimivirus. Haut: photo; bas: schéma

Plus récemment encore, on a découvert un virus capable de changer de forme et de se développer par un processus actif en dehors de tout contexte cellulaire : le virus ATTV (*Acidianus Two-Tailed Virus*). Ce virus infecte des archéobactéries qui vivent près des sources hydrothermales chaudes. Il a une forme de citron, qui s'allonge à chaque extrémité lorsque la température est suffisamment élevée (environ 80 degrés Celsius).

Ces deux exemples montrent la difficulté de trouver des définitions précises qui résistent aux nouvelles découvertes. Et il faut garder en mémoire que l'on ne connaît aujourd'hui qu'environ 10 000 virus, alors qu'on estime le nombre de particules virales différentes à environ 10^{31} sur notre planète (une diversité infiniment supérieure à celle cumulée de l'ensemble des trois règnes du vivants : eucaryotes, procaryotes et archées).

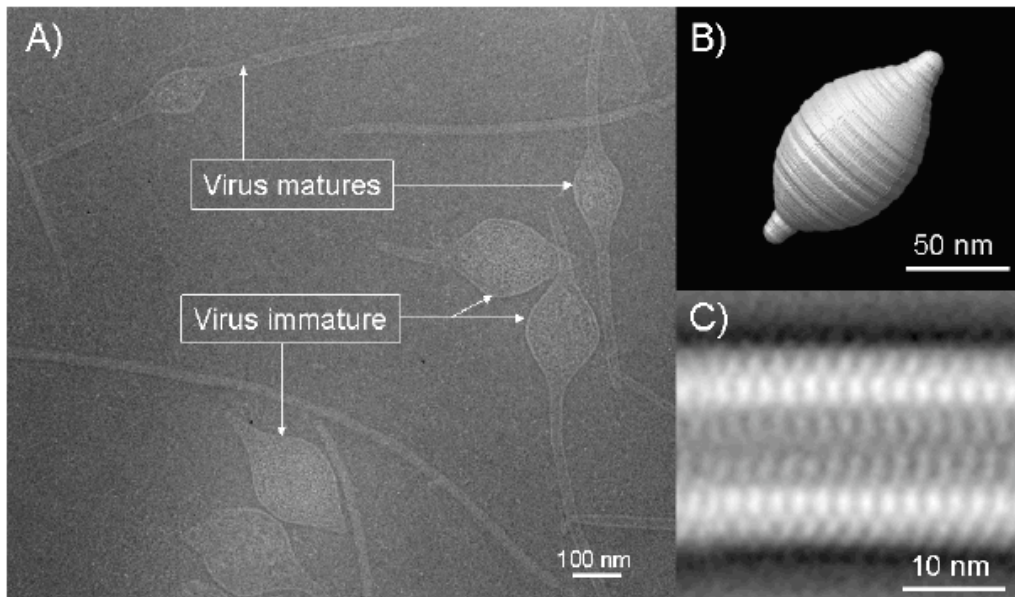


Figure 16: Image de Ricardo Aramayo. Collaboration avec le Dr David Prangishvili (Institut Pasteur, Paris). Etude structurale du virus ATTV (acidianus two-tailed virus) de la famille nouvellement découverte des *bicaudaviridae*. A) Cryomicroscopie électronique du virus à l'état congelé-hydraté. B) Reconstruction de l'enveloppe d'un virus immature permettant d'évaluer son volume moyen. C) Image moyenne d'une extension tubulaire d'un virus mature.

3.4.3 Exemple 2 : la définition de la mort et la cryptobiose du Tardigrade

Les êtres vivants multicellulaires sont tous marqués par un processus de vieillissement, qui s'achève par un arrêt des fonctions vitales : la mort. Quand peut-on dire qu'un organisme vivant a cessé de vivre, c'est-à-dire qu'il est mort ? Nous allons voir que même pour les organismes multicellulaires, cette question est très complexe. La notion de vie est généralement associée à celle de métabolisme (voir plus haut), c'est-à-dire d'une activité physico-chimique globale du système vivant, la **mort biologique** étant définie comme l'arrêt du métabolisme. L'importance psychologique de la définition de la mort est illustrée par le développement de la notion, reconnue légalement pour les êtres humains, de **mort cérébrale**³⁶, qui est née dans les années 1960 après les avancées médicales sur le coma dépassé.

³⁶ La notion de mort cérébrale est apparue légalement en France avec l'arrêt du Conseil d'Etat du 2 juillet 1993, dit « arrêt Milhaud », puis fixée par un décret du 2 décembre 1996.

Si la personne présente un arrêt cardiaque et respiratoire persistant, le constat de la mort ne peut être établi que si les trois critères cliniques suivants sont simultanément présents (article 671-1-1 du code de la Santé Publique) :

- Absence totale de conscience et d'activité motrice spontanée
- Abolition de tous les réflexes du tronc
- Absence totale de ventilation spontanée

De plus, en compléments de ces trois critères cliniques, il doit être recouru pour attester du caractère irréversible de la destruction encéphalique (article 671-1-2 du code de la Santé Publique):

- Soit à deux électroencéphalogrammes nuls et aréactifs effectués à un intervalle minimal de quatre heures, réalisés avec amplification maximale sur une durée d'enregistrement de trente minutes et dont le résultat doit être immédiatement consigné par le médecin qui en fait l'interprétation.

Intéressons-nous à la mort biologique, pour des animaux simples. Peut-on dire que tout animal, dont le métabolisme est nul (au sens ou plus aucune activité physique ou chimique n'est décelable dans son corps), est mort ? La réponse 'oui' semble s'imposer comme une évidence. Pourtant, il existe un contre-exemple célèbre : le tardigrade (*Echiniscus trisetosus*). Ce petit animal (il mesure moins de 2 mm) possède une propriété rarissime : celle d'entrer en état de **cryptobiose** lorsque l'environnement lui devient hostile. La cryptobiose est un état de métabolisme réduit, que l'on observe par exemple chez les animaux qui hibernent. Chez le tardigrade, la cryptobiose atteint un état unique, puisque aucun moyen de mesure actuellement disponible ne permet de détecter la moindre trace d'activité physico-chimique. Le tardigrade est donc alors totalement 'inerte', et ne rentre plus dans la définition des systèmes vivants. Pourtant, mis dans un environnement suffisamment humide, le tardigrade sort de son état de cryptobiose et revient à la vie dans un délai de une à deux heures. Ce phénomène, unique en son genre, repousse la définition classique de la mort.

En état de cryptobiose, le tardigrade prend une forme de tonnelet qui lui permet de réduire sa surface d'évaporation et de limiter les pertes en eau qui pourraient être fatales aux organites très sensibles à la dissécatation. L'animal produit également un sucre qui remplace l'eau dans son corps, ce qui permet de limiter l'altération des membranes cellulaires pour qu'elles puissent être réparées lors de la réhydratation. Il joue également le rôle d'antigel, et évite l'endommagement des cellules dû à la formation de cristaux de glace à très basse température.

Dans cet état, le tardigrade possède des capacités de résistance uniques dans le monde du vivant et qui ont fait sa célébrité. Les expériences menées en laboratoire ont montré que le tardigrade pouvait sortir de cryptobiose et revenir à une vie normale après avoir été soumis aux conditions suivantes :

- -272,9 degrés Celsius pendant 20 heures
- -253 degrés Celsius pendant 26 heures
- -190 degrés Celsius (dans l'air liquide) pendant 25 mois
- plusieurs siècles vers -20 degrés Celsius (découvert dans les glaces aux pôles)
- 30 minutes à 360 degrés Celsius
- le vide absolu
- des pressions de 6000 atmosphères (la limite normale d'altération des cellules et de l'ADN est de l'ordre de 300 atmosphères)
- l'alcool pur (alcool à 100°)
- le sulfure d'hydrogène (H₂S)
- l'anhydride carbonique (CO₂)
- la lumière ultraviolette
- les rayons X, dans des doses normalement mortelles

-
- Soit à une angiographie objectivant l'arrêt de la circulation encéphalique dont le résultat doit être immédiatement consigné par le radiologue qui en fait l'interprétation.

Si la personne dont le décès est constaté cliniquement est assistée par ventilation mécanique et conserve une fonction hémodynamique, l'absence de ventilation spontanée est vérifiée par une épreuve d'hypercapnie : l'augmentation du gaz carbonique dissous dans le plasma sanguin.

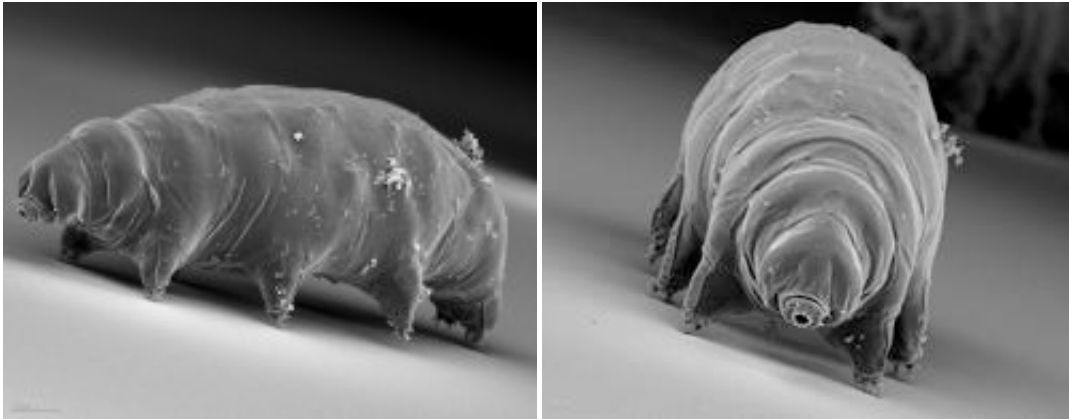


Figure 17: Tardigrade en état de métabolisme actif

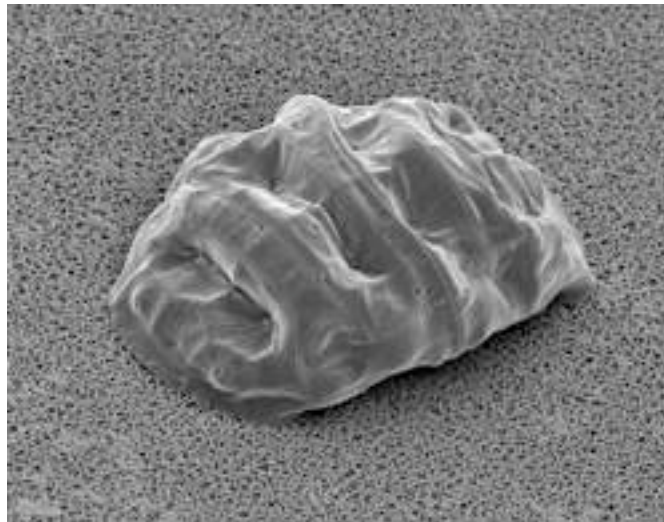


Figure 18: Tardigrade en état de cryptobiose

3.4.4 Exemple 3 : animal, champignon ou végétal ?³⁷

Venons-en maintenant à une autre question, symboliquement moins chargée que les précédentes, et beaucoup plus simple au premier abord : qu'est-ce qu'un animal, un champignon et un végétal ?

Les classifications modernes des eucaryotes sont basées sur les **analyses phylogénétiques** : par analyse du patrimoine génétique de chaque organisme vivant, on retrace le processus d'évolution, ce qui permet de distinguer des **groupes** dits **monophylétiques** (il en existe 8 en tout), qui réunissent chacun un ancêtre commun et toute sa descendance. L'ancêtre unique commun à tous les eucaryotes reste inconnu, de même que l'ancêtre commun aux trois règnes du vivant (eucaryotes, procaryotes et

³⁷ Source : « Animal ou végétal ? une distinction obsolète », M.A. Selosse, Pour la Science No 350, 2006

archées). Ce dernier est nommé LUCA (pour *Last Unknown Common Ancestor*) par les biologistes spécialistes de l'évolution. En observant l'arbre phylogénétique des eucaryotes, on peut voir que la définition des termes animal, champignon ou végétal est problématique, puisque ces familles ne correspondent pas à des groupes monophylétiques.

- Les **animaux** regroupent les métazoaires (multicellulaires, en bleu sur la figure) et les protozoaires (en orange sur la figure). Rappelons que la séquence ADN de l'être humain est identique à environ 98% à celle du chimpanzé, à 80% de celle de la souris et à 60% de celle de la mouche drosophile.
- Les **végétaux** sont les organismes capables de photosynthèse, ce qui regroupe tous les organismes dotés de plastides (des organites qui contiennent la chlorophylle et sont doués de photosynthèse). Cette capacité est acquise au cours de l'évolution lorsque une cellule photosynthétique (cyanobactérie pour la lignée verte, algue unicellulaire (rouge ou verte) pour les autres) est incorporée à l'intérieur d'une autre. On nomme ce processus l'**endosymbiose**.
- Les **champignons** ne sont pas des animaux, mais ne sont pas doués de photosynthèse. Ils se nourrissent de molécules organiques (on dit qu'ils sont hétérotrophes), qu'ils absorbent à travers leur paroi rigide qui interdit la phagocytose (ils sont absorbotrophes). Ils sont formés de fins filaments (appelés hyphes) qui permettent l'exploration du milieu environnant et la recherche de nourriture. Certains champignons sécrètent des enzymes qui digèrent les molécules complexes hors de la cellule. Pour protéger cette digestion extérieure contre des organismes « concurrents », les champignons sécrètent des souvent des antibiotiques (comme la pénicilline).

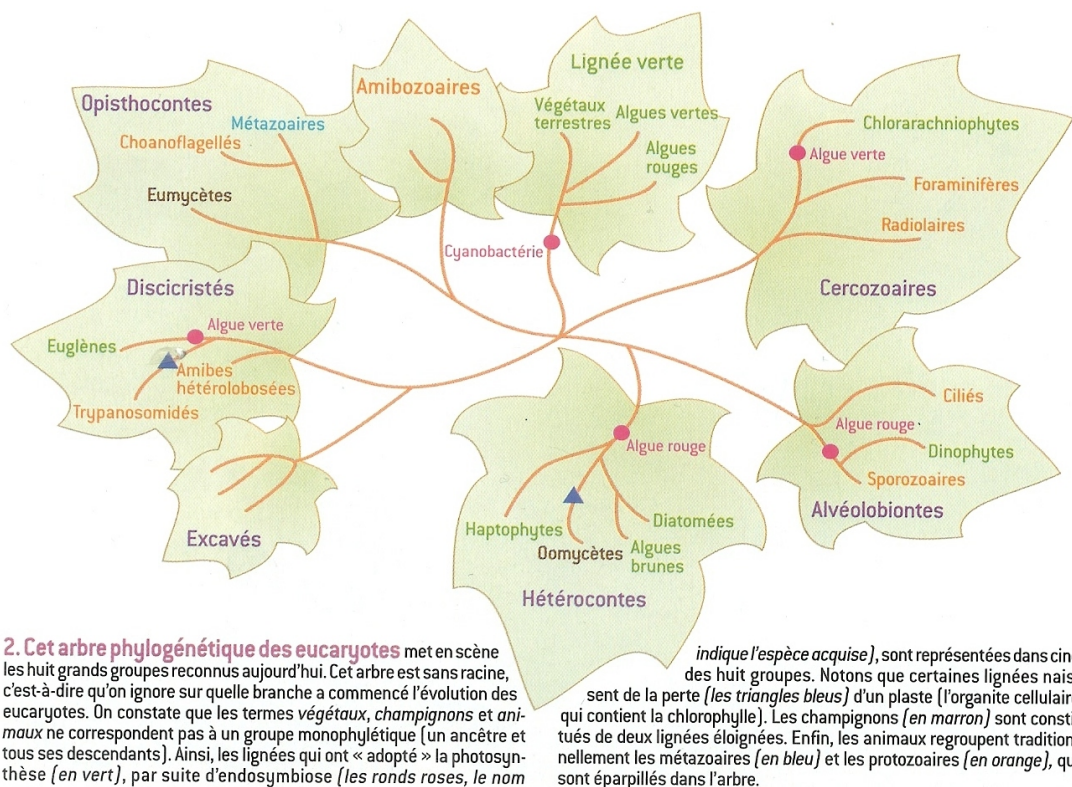


Figure 19: arbre phylogénétique des eucaryotes

L'observation des groupes monophylétiques montre que :

- aucune famille (animal, végétal, champignon) n'est entièrement contenue au sein d'un même groupe. Ainsi, il existe des végétaux qui n'ont aucun ancêtre commun, mis à part l'ancêtre commun à tous les eucaryotes. Ceci résulte du fait que l'acquisition d'un plaste par endosymbiose s'est produite plusieurs fois au cours de l'évolution au sein de lignées différentes. On observe aussi que la perte du plaste donne naissance à de nouveaux champignons sans liens de parenté. Enfin, les animaux sont présents dans 5 des groupes monophylétiques.
- Un groupe monophylétique peut regrouper des éléments trois familles en son sein. On constate ainsi que les êtres humains (des métazoaires), sont plus proches des cèpes des forêts et des morilles (champignons de la famille des eumycètes) que de certains autres animaux !

Ces constatations montrent que la définition des animaux, des végétaux et des champignons n'est pas triviale et regroupe des organismes qui n'ont aucun lien de parenté génétique (mis à part l'ancêtre unique inconnu). Ceci est rendu encore plus complexe par le fait que des organismes différents peuvent développer des stratégies semblables. Ainsi, certaines bactéries se comportent comme des champignons : les bactéries actinomycètes sont filamenteuses, hétérotrophes, dotées d'une paroi et fabriquent des enzymes et des antibiotiques (par exemple, la streptomycine). On trouve aussi des plantes parasites semblables aux champignons, qui ont perdu leur chlorophylle (par exemple, les plus grosses fleurs connues sont celles de la *Rafflesia*, une plante parasite tropicale, dotée de paroi et absorbotrophe).

3.5 Définir un objet (fin) : la critique de Lakatos

Les modes de définition qui viennent d'être discutés ne sont pas pleinement satisfaisants pour décrire la manière dont les concepts sont définis dans les théories scientifiques. Les critiques que l'on peut formuler sont les suivantes :

- La définition extensionnelle n'est pas utilisable en pratique, car elle conduit à l'emploi de listes infinies
- La définition intensionnelle d'un concept fait généralement appel à d'autres concepts, et ainsi de suite. On se trouve de nouveau face à un problème de régression infinie qui n'admet pas de solution satisfaisante (circularité ou clôture arbitraire). C'est par exemple l'analyse de Poincaré de la mécanique newtonienne.
- La définition opératoire n'est pas utilisable pour tous les concepts de la physique moderne (par exemple en mécanique quantique), et elle est affaiblie par l'incertitude intrinsèque aux énoncés d'observation.

Un constat est qu'une définition, même si elle est basée sur un ensemble auto-consistant de concepts (une sorte de dictionnaire, où chaque concept est défini en fonction d'autres concepts également définis) fait appel à des concepts qui appartiennent à la même théorie (pour reprendre notre analogie, un mot est défini dans le dictionnaire à partir d'autres mots du même dictionnaire). Pour avoir accès à la

signification d'un concept, il faut connaître la théorie au sein de laquelle il est défini (toujours avec le même exemple, il faut déjà posséder une certaine maîtrise du langage pour pouvoir se servir d'un dictionnaire). Un point important est le caractère incommensurable des théories. Ce point de vue, défendu par plusieurs épistémologues, implique que les concepts d'une théorie ne sont pas exactement exprimables au moyen des concepts d'une autre théorie (un dictionnaire d'anglais ne sera d'aucune utilité pour quelqu'un ne parlant que le français et qui cherche à se servir d'un dictionnaire de français). Par exemple, le concept newtonien de masse n'est pas formulable en utilisant la mécanique pré-newtonienne, de même le concept de masse dans la théorie de la relativité d'Einstein lui est très différent.

Ce constat a conduit Lakatos à formuler la proposition selon laquelle les définitions doivent être rejetées en tant que procédures fondamentales pour donner du sens à un concept. Pour lui, c'est dans les théories que les concepts puisent leur sens. Cette réflexion est étayée par l'exemple précédent de la masse. Le caractère historique et évolutif des théories conduit également Lakatos à faire des programmes de recherche une source de signification pour les concepts. Ceci est dû que de nombreux concepts (masse, énergie, atome, onde, ...) ont initialement été définis de manière floue, et on vu leur définition évoluer et se préciser au fur et à mesure que les théories ont été développées.

4 Qu'est-ce que faire une expérience ?

Les chapitres précédents ont mis en évidence l'importance des faits et de leur mesure pour les théories dans les sciences empiriques. Il convient donc de s'interroger sur la nature et la valeur des données obtenues de l'expérience.

4.1 De l'observation à l'expérimentation

Commençons d'abord par distinguer l'**observation** de l'**expérimentation**. Dans les deux cas, le scientifique interroge la nature pour obtenir des données qui vont alimenter sa réflexion.

Dans le cas de l'observation, le scientifique est *idéalement neutre*, passif en ce sens qu'il ne cherche pas à modifier le déroulement naturel des faits. C'est par exemple le cas de l'astronomie, de l'astrophysique et de la climatologie : il est impossible de perturber une étoile pour valider un modèle de son fonctionnement. On ne peut que mesurer les signaux émis par les étoiles observables. Il en est de même pour le climat : le climatologue ne dispose que d'un seul exemple qu'il ne peut modifier à volonté. C'est pourquoi les résultats provenant de la paléoclimatologie sont si importants : à défaut d'avoir plusieurs planètes semblables à la Terre à observer, on cherche à collecter le plus possible d'informations sur les états passés du climat.

Dans le cas de l'expérimentation, le scientifique influe sur le cours des événements pour tenter d'isoler un phénomène qu'il veut analyser. Il est en cela *actif*. C'est l'expérience répétée qui permet de collecter les faits et, par exemple, d'accéder à une définition opératoire.

Les sciences recourant à l'expérimentation sont appelées **sciences expérimentales** (mécanique, physique, chimie, ...), celles uniquement basées sur l'observation, **sciences d'observation**.

Avant d'aller plus loin, il est important de noter que la neutralité de l'observation est toute relative, ce qui conduit à rendre parfois peu utile la distinction entre observation et expérimentation. En effet, l'observation partage avec l'expérimentation la nécessité d'effectuer des mesures, donc d'utiliser des appareils de mesure et de recourir à des théories scientifiques (avant la mesure pour déterminer quelles grandeurs doivent être mesurées, et après pour tirer des conclusions), problèmes qui font l'objet de ce chapitre et qui font que l'observation totalement neutre, c'est-à-dire sans aucun préalable, est impossible.

4.2 La mesure des phénomènes physiques

4.2.1 Perception, mesure et témoignage des données

La pratique scientifique moderne dans le champ des sciences empiriques fait un large usage des données issues des expériences. Il faut noter que l'observation par l'expérimentateur au moyen de perceptions sensorielles directes (la vue dans la très grande majorité des cas) n'est plus que très marginale. Ce qui compte, en revanche,

c'est l'acquisition de **données**, c'est-à-dire de valeurs mesurées associées à des champs et des grandeurs qui apparaissent dans les expressions mathématisées que sont les lois. L'acquisition de ces données est le résultat de la **mesure**, qui sera discutée plus loin. Les données brutes, c'est-à-dire les valeurs issues de la mesure, ne sont pas directement exploitables : pour acquérir un sens et servir à la production de connaissance, elles doivent d'abord être mises en forme dans le cadre d'un **modèle de données** puis être interprétées dans un cadre théorique. On parle alors de **témoignage des données**.

Enfin, il faut noter que dans certaines expériences, par exemple en physique nucléaire, le rôle de la perception directe par les scientifiques est pratiquement nul : les mesures sont automatisées et les données sont ensuite transmises à des ordinateurs qui les exploitent. Les intervenants humains n'accèdent qu'à des résultats provenant de modèles théoriques, et non plus aux « faits bruts ». Dans de tels cas, la perception humaine est réduite à la lecture des graphiques, tableaux et textes produits par les ordinateurs, et est donc très différente de l'activité perceptive telle qu'elle était conçue lors des siècles précédents.

4.2.2 Mesure, précision et étalons

L'opération de mesure a pour but de quantifier, c'est-à-dire de ramener sur une échelle de nombres réels dont la valeur unité est associée à une référence connue de tous, certaines grandeurs associées aux objets et phénomènes étudiés.

Si l'échelle de mesure est basée sur les nombres réels, les valeurs mesurées seront approchées en ce sens qu'elles seront tronquées : le nombre de décimales exprimables est par nature fini. Sur une telle base, une mesure exacte n'est pas réalisable. Mais, en pratique, cette source d'inexactitude n'est pas un problème pour la vie courante et bon nombre de mesures réalisées dans le cadre d'expériences de laboratoire. C'est la limite de résolution ou de précision des instruments de mesure qui pose problème.

Pour parler un langage commun, les scientifiques partagent les mêmes unités de mesure³⁸. Pour la mécanique et la physique, les unités aujourd'hui standards sont le mètre pour les longueurs, le kilogramme pour les masses et la seconde pour les temps. L'emploi d'unités de mesure universelles est également considéré comme une simplification pour les échanges économiques, la fiscalité ... L'harmonisation des poids et des mesures était par exemple une revendication très forte lors des états généraux convoqués en 1788 qui ont précédé la Révolution Française. On estime³⁹ que la France comptait alors près de 250 000 unités de mesure différentes ! Chaque unité est associée à un étalon, qui est un objet ou un processus naturel. Les propriétés requises pour un **étalon** sont la **stabilité** (l'objet doit conserver ses caractéristiques sur un temps théoriquement infini, en pratique « très long ») et la **disponibilité** : toute personne suffisamment équipée doit être en mesure de se servir de l'étalon pour garantir la valeur de ses mesures. Les définitions du mètre, de la seconde et du kilogramme sont rappelées dans ce qui suit.

³⁸ Seuls trois pays n'ont pas adopté officiellement le système métrique à ce jour : les Etats-Unis d'Amérique, l'Union du Myanmar et le Libéria.

³⁹ Voir par exemple « Mesurer le monde – L'incroyable histoire de l'invention du mètre », K. Adler, collection « Champs histoire », Flammarion, 2005

Qu'est-ce qu'un mètre ?

La définition du mètre a évolué au cours du temps, au sens où l'étalon utilisé pour le définir a varié (c'est donc sa définition opératoire qui a évolué).

A l'origine, le mètre était défini comme une fraction de la longueur du périmètre de la terre à l'équateur (voir plus loin pour un historique plus précis).

Cette définition fut ensuite remplacée en 1889 par une barre de platine-iridium de section en forme de X déposée au musée des Poids et Mesures, à Sèvres. On définit le mètre-étalon comme la distance entre deux marques gravées sur cette barre quand la barre est à la température de la glace fondante et qu'elle est posée sur deux rouleaux horizontaux symétriques perpendiculaires à son axe, distants l'un de l'autre de 0,571 mètre. La section en X garantit une grande rigidité de la barre et sa composition assure une faible variabilité en fonction de la température.

En octobre 1960, une nouvelle définition est adoptée par le Comité international des poids et mesures dans la résolution 7 de la 11^e Conférence générale des poids et mesures: le mètre est défini comme 1 650 763,73 longueurs d'onde dans le vide de la radiation correspondant à la transition entre les niveaux $2p^{10}$ et $5d^5$ de l'isotope 86/36 du krypton.

Pour assurer la répétabilité, le Comité « recommande que la radiation du krypton 86 adoptée comme étalon fondamental de longueur soit réalisée au moyen d'une lampe à décharge à cathode chaude contenant du krypton 86 d'une pureté non inférieure à 99 pour cent, en quantité suffisante pour assurer la présence de krypton solide à la température de 64 °K, cette lampe étant munie d'un capillaire ayant les caractéristiques suivantes : diamètre intérieur 2 à 4 millimètres, épaisseur des parois 1 millimètre environ. »

La définition actuelle est celle adoptée en 1983 dans la résolution 1 de la 17^e Conférence générale des poids et mesures: « Le mètre est la longueur du trajet parcouru dans le vide par la lumière pendant une durée de $1/299\,792\,458$ de seconde ». Cette définition fixe la vitesse de la lumière dans le vide à $299\,792\,458$ m.s⁻¹.

Qu'est-ce qu'une seconde ?

La durée de la seconde fut fixée en octobre 1960 par la résolution 9 de la même Conférence qui fixa la longueur du mètre : « La seconde est la fraction $1/31\,556\,925,9747$ de l'année tropique pour 1900 janvier 0 à 12 heures de temps des éphémérides. »

Cette définition a été changée comme suit en 1968 pour arriver à la définition actuelle: « la seconde est la durée de $9\,192\,631\,770$ périodes de la radiation correspondant à la transition entre les deux niveaux hyperfins de l'état fondamental de l'atome de césium 133. » Il a été confirmé récemment, en 1997, que cette définition fait référence à un atome de césium au repos, à une température de 0 degré Kelvin.

Qu'est-ce qu'un kilogramme ?

La définition du kilogramme n'a, elle, pas changé depuis 1889 : il est défini comme la masse d'un prototype en platine iridié déposé au musée de Sèvres.

On remarque que ce sont des définitions opératoires !

De très nombreuses unités dérivées sont déduites de ses unités de base. Les phénomènes physiques rencontrés dans la nature font apparaître des ordres de

grandeur très différents, et il est nécessaire d'employer des unités adaptées pour la mesure, au sens où les nombres seront plus commodes à manipuler. Les conventions internationales, qui reprennent la nomenclature fixée par les révolutionnaires français en 1795, sont rappelées dans le tableau ci-dessous.

Facteur	Préfixe	Symbole	Facteur	Préfixe	Symbole
10^{24}	yotta	Y	10^{-1}	déci	d
10^{21}	zetta	Z	10^{-2}	centi	c
10^{18}	exa	E	10^{-3}	milli	m
10^{15}	peta	P	10^{-6}	micro	μ
10^{12}	téra	T	10^{-9}	nano	n
10^9	giga	G	10^{-12}	pico	p
10^6	méga	M	10^{-15}	femto	f
10^3	kilo	k	10^{-18}	atto	a
10^2	hecto	h	10^{-21}	zepto	z
10^1	déca	da	10^{-24}	yocto	y

Tableau 1: unités dérivées suivant les conventions internationales

4.2.3 Repères : histoire du système métrique

Avant de poursuivre, intéressons-nous à l'histoire du système métrique. Ce système d'unités, nous allons le voir, est intimement lié à l'histoire et à la politique. Il est le produit de la Révolution française, qui voulut offrir « *à tous les temps et à tous les peuples, pour leur plus grand avantage* », un système métrologique unifié et stable.



Figure 20: C. Borda

- **Ancien régime**
 - **Vers 1660** : Dans un souci d'unification, la Royal Society of London propose de choisir comme unité de longueur celle d'un pendule battant la seconde. Cette idée est reprise par le français Jean Picard en 1671, par le néerlandais Christian Huygens en 1673, puis par le français La Condamine, l'anglais John Miller et l'américain Thomas Jefferson.
 - **1670** : l'abbé Gabriel Moulton propose de choisir le millième de la minute sexagésimale du méridien (soit 1/1000 de mille marin, environ 1,85 mètre) comme unité de longueur. Cette unité est peu commode, car trop grande. Elle sera donc abandonnée.
 - **1672** : Richer découvre que le pendule battant la seconde est plus court à Cayenne qu'à Paris. Ce manque de stabilité de l'étalon conduit à son abandon.
 - **1758** : Louis Dupuy suggère d'unifier les mesures de longueur et de poids sur la base du pesage d'un volume d'eau mesuré en unités de longueur
- **Révolution Française et 1^{er} Empire**

- **8 mai 1790** : Sur proposition de Talleyrand, la Constituante se prononce pour la création d'un système de mesure stable et unifié. Une commission est constituée à l'Académie des Sciences (Lagrange, Laplace, Monge, Borda, Lavoisier). L'unité de mesure retenue est la longueur d'un pendule battant la seconde.
- **26 mars 1791** : Le principe du pendule est abandonné, et remplacé par le choix du **mètre**⁴⁰ (le nom est proposé par Jean-Charles de Borda (1733-1799)), défini comme la dix-millionième partie du quart du méridien terrestre. Il faudra alors s'atteler à la tâche de la mesure du méridien. On uniformise alors les différentes unités de mesure : l'unité de surface (are) est définie comme celle d'un carré ayant pour côté un décamètre ; l'unité de masse, le kilogramme, correspond à celle d'un volume unitaire (le litre) d'eau pure à la température de la glace fondante, corrigé des effets de latitude et de pression.
- **1792** : Début de l'«*expédition de la méridienne*», destinée à mesurer la longueur du méridien terrestre. Elle est dirigée par Jean-Baptiste Joseph Delambre (1749-1822) et Pierre François André Méchain (1744-1804). On décide de relier Dunkerque à Barcelone par une suite de triangles géodésiques repérés au moyen de visées angulaires. Cette tâche s'achèvera en 1799. La précision obtenue est de 0,2 millimètres.
- **1^{er} août 1793** : Décret instituant la décimalisation des évaluations monétaires, des mesures de longueur, de superficie, de volume et de poids. On force ainsi l'abandon du système hérité de l'ancien régime, basé sur une division par 12.
- **7 avril 1795 : (loi du 18 germinal, an III)** organisation du système métrique et fixation de la nomenclature actuelle des unités (décimètre, millimètre, gramme, décigramme, kilogramme, ...)
- **9 juin 1795** : Lenoir construit le premier étalon métrique légal sur la base des mesures effectuées par La Caille de la distance entre le pôle et l'équateur.
- **25 juin 1795** : création du Bureau des Longitudes à Paris.
- **1799** : Réunion à Paris de la première conférence internationale pour analyser les résultats de Delambre et Méchain et débattre de la question de l'adoption universelle du système métrique. La proposition est rejetée, car trop « révolutionnaire » et « française ».
- **22 juin 1799** : dépôt aux Archives Nationales des prototypes en platine du mètre construit par Lenoir sur la base des résultats de l'expédition de la méridienne et du kilogramme
- **10 décembre 1799 : (loi du 19 frimaire, an VIII)** : fixation des étalons définitifs, deuxième définition du mètre (celui de l'étalon des Archives Nationales), et obligation théorique de l'usage du système métrique.
- **4 novembre 1800** : Loi faisant du mètre la seule unité de mesure officielle ; sur injonction de Napoléon Bonaparte, la nomenclature basée sur les préfixes grecs est abandonnée.

⁴⁰ mètre est construit à partir du grec **metron** qui signifie « mesure, toute quantité mesurée, limite, juste mesure »

- **12 février 1812** : suivant la volonté de Napoléon, de nouvelles unités de mesure portant les noms de celles de l'ancien régime sont introduites « pour les opérations ordinaires ».
- **Après la Révolution Française**
 - **1^{er} janvier 1840** : entrée en vigueur de la loi du 4 avril 1837, qui institue définitivement l'obligation d'utiliser le système métrique
 - **1875** : Création du Bureau international des poids et mesures à Sèvres
 - **22 avril 1876** : dépôt de l'étalon métrique international au Pavillon de Breteuil, à Sèvres, que la France cède au Comité international des poids et mesures en le dotant du privilège d'extra-territorialité.

4.2.4 Les instruments de mesure

La mesure est effectuée au moyen d'instruments, qui ont le double avantage de quantifier les phénomènes (alors que nos cinq sens – vue, ouïe, toucher, goût, odorat – ne fournissent que des indications qualitatives) et de permettre d'accéder à des objets qui se trouvent au-delà des limites de la perception humaine : le microscope permet de voir des objets très petits, le compteur Geiger de détecter les radiations, etc.

La nature de l'instrument de mesure est double : c'est un instrument concret, construit, qui est la matérialisation d'un instrument conçu, virtuel. Les propriétés de ces deux instruments diffèrent du fait des contraintes technologiques (un miroir de télescope n'est jamais parfaitement lisse, la masse du plateau d'une balance n'est jamais exactement connue, ...) et des simplifications théoriques employées pour concevoir l'instrument virtuel. Par exemple, une balance de Roberval utilisée sur Terre est soumise à l'attraction terrestre, et ces éléments peuvent se déformer très légèrement. Le scientifique raisonne sur l'instrument conçu, qui est une idéalisation. Le procédé de mesure est basé sur des concepts théoriques (force, masse, énergie, courant électrique, pression d'un gaz, ...) et le résultat de la mesure, la valeur rendue par l'instrument réel, est également interprété dans le cadre théorique par l'observateur. Les différences de fonctionnement entre l'instrument virtuel et l'instrument construit étant généralement faibles, les mesures effectuées sont utiles pour le travail théorique.

Ceci conduit à la réflexion suivante : les instruments de mesure sont des « théories matérialisées », selon le mot de Gaston Bachelard. En effet, la donnée brute fournie par l'instrument est interprétée par l'expérimentateur pour être ensuite utilisée : la variation de l'aiguille d'un manomètre est automatiquement interprétée comme le signe d'un changement de la pression d'un gaz. Il y a donc établissement d'un lien entre deux niveaux lors de la mesure : le niveau de la perception directe (on voit l'aiguille du manomètre bouger) et celui de la théorie (la pression du gaz varie). Ce lien est assuré par une élaboration théorique qui repose sur les théories scientifiques qui ont servi à concevoir l'appareil de mesure. La fiabilité de la mesure est donc gouvernée par la confiance que l'on a dans les théories sous-jacentes à la conception de l'instrument de mesure. Plus les problèmes théoriques sont nombreux, moins la mesure est exploitable.

Dans la pratique quotidienne, lorsque des instruments d'utilisation courante et reconnus fiables sont utilisés, cette conversion des données brutes en résultats théoriques est faite automatiquement, presque inconsciemment, par l'expérimentateur.

Toutefois, l'histoire des sciences fournit de multiples exemples où le manque de précision des instruments de mesure a conduit à la formulation de conclusions erronées. Par exemple, l'astronome danois Tycho Brahé déclara avoir prouvé la fausseté de la théorie héliocentrique de Nicolas Copernic. En effet, si la Terre tourne autour du Soleil, les étoiles fixes doivent avoir une parallaxe (ce qui veut dire que leur position dans le ciel doit varier au cours de l'année). Or ses observations ne lui permirent pas de constater ce phénomène, ce qui le poussa à rejeter le modèle de Copernic pour en proposer un autre, dans lequel les planètes tournent autour du Soleil, et le Soleil autour de la Terre qui elle, est immobile. Le problème rencontré par Brahé était qu'il avait surestimé l'effet de parallaxe, et que l'effet réel était beaucoup trop faible pour pouvoir être détecté par les lunettes astronomiques dont il disposait. Il faudra attendre plusieurs siècles pour que des instruments optiques suffisamment précis pour mesurer ce phénomène soient mis au point. Mais les mesures de Tycho Brahé sur les trajectoires des planètes ont été suffisamment précises pour permettre à Johannes Kepler (1571-1630) de formuler ses trois fameuses lois sur les orbites planétaires. Un autre exemple d'imperfection des lunettes astronomiques est celui des aberrations optiques, qui conduisirent Kepler à noter que « *les étoiles sont carrées et vivement colorées* ».

Nous allons maintenant illustrer ces réflexions en analysant quelques appareils de mesure.

4.2.5 Exemple 1 : La machine d'Attwood : mesurer la masse et les forces

Comme premier exemple, reprenons la machine d'Attwood déjà mentionnée dans la discussion sur la définition de la force et de la masse. Nous allons voir que ce dispositif, extrêmement simple sur le plan technologique et qui ne fait appel qu'à une théorie scientifique « simple » (la mécanique de Newton), se révèle être en fait très complexe sur le plan de l'analyse théorique.

Traçons tout d'abord le schéma de la machine, et analysons le fonctionnement de la machine idéale, virtuelle.

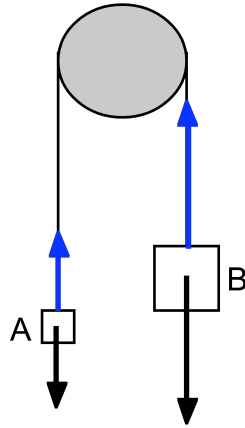


Figure 21: machine d'Atwood virtuelle. En noir : les poids des solides A et B. En bleu : les forces de réaction.

Les deux solides A (de masse m_A) et B (de masse m_B) sont reliés par le fil f , de masse nulle, qui reste tendu grâce à l'emploi de la poulie, et qui est supposé transmettre les efforts sans en modifier la norme (il n'est pas élastique). Qu'observerons-nous ? Un observateur « naïf » dira qu'il voit les deux solides se mettre en mouvement, le plus lourd tirant le plus léger vers le haut.

Cet énoncé est-il vraiment satisfaisant ? Non. Établissons le bilan des forces (on néglige la force d'Archimède due à l'air) : le solide A est soumis à la force de pesanteur et à la force exercée par le fil. De même, le solide B est soumis à son poids et à la force exercée par le fil. Donc, on ne peut pas dire que le poids d'un solide est appliqué à l'autre : c'est la tension du fil, déduite de la loi de l'action et de la réaction appliquée à l'autre extrémité du fil, qui est appliquée au solide. Nous avons donc implicitement fait usage de la troisième loi de Newton pour interpréter notre expérience. La fiabilité de nos conclusions dépend donc de notre confiance dans cette loi. Est-elle vérifiable ? Non, répond Henri Poincaré (1854-1912) (*La science et l'hypothèse*, 1902) : il n'est pas possible de vérifier cette loi, qui doit être prise pour une définition de la réaction (et comme telle, ne peut être violée). Donc, notre interprétation repose sur la confiance que nous mettons dans l'articulation des concepts de la mécanique newtonienne.

Continuons l'analyse : l'utilisation de la machine d'Atwood est basée sur la seconde loi de Newton, qui permet de lier force, masse et accélération. Ce qui est mesuré, c'est le temps que met le solide qui monte pour parcourir une distance donnée. On en déduit l'accélération en faisant l'hypothèse que celle-ci est constante au cours du mouvement. Et ensuite, on utilise la loi de Newton pour interpréter le résultat.

Comme on le voit, ici aussi, l'appel à la théorie est constitutif du procédé de mesure.

Venons-en maintenant aux différences entre machine réelle et machine virtuelle. Ces différences sont très nombreuses, et non allons nous contenter d'en lister quelques unes :

- Les propriétés mécaniques des matériaux employés ne sont pas prises en compte : un fil n'est jamais parfaitement inextensible et sans masse, et de plus il peut avoir une certaine rigidité ; il peut y avoir des frottements sur la poulie ; lors de leur mouvement, les solides peuvent être freinés par la résistance de l'air ; la force d'Archimède n'est pas nulle ...
- Le poids d'un solide à la surface de la Terre dépend de la pesanteur g à la surface de celle-ci. Or la pesanteur n'est pas constante à la surface de la Terre. Tout d'abord, il faut prendre en compte le fait que la Terre n'est pas

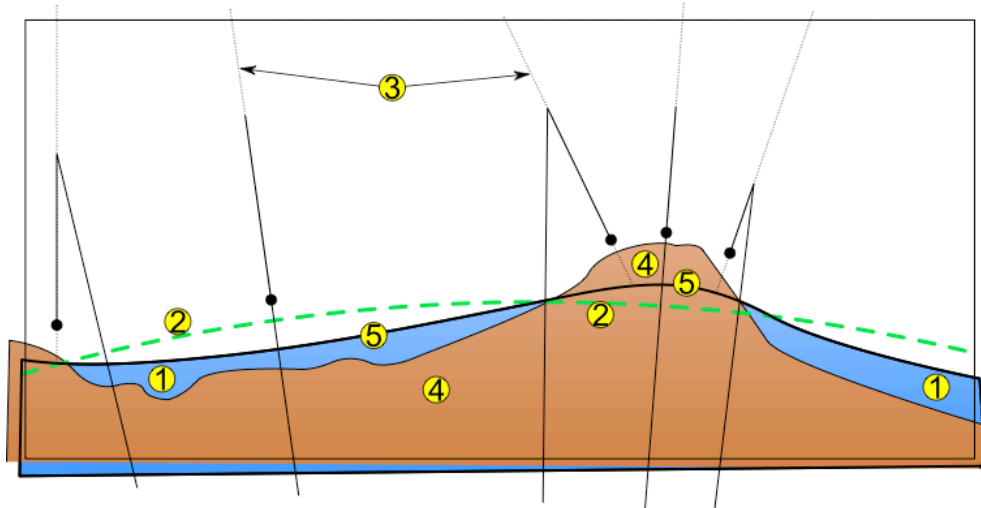
sphérique : sous l'effet de la rotation, elle prend la forme d'un ellipsoïde aplati avec une différence de rayon entre pôle et équateur de l'ordre de 21 kilomètres. La distance d'un point situé à la surface de la Terre par rapport au centre est donc variable, ce qui induit une variation de la pesanteur en fonction de la latitude⁴¹: $9,78 \text{ m.s}^{-2}$ à l'équateur contre $9,83 \text{ m.s}^{-2}$ aux pôles (soit une variation de l'ordre de $5 \cdot 10^{-3} \text{ g}$). La pesanteur ressentie par un corps pesant varie bien sûr également avec l'altitude par rapport à la surface terrestre. Donc, lorsqu'un solide monte ou descend, son poids change, ce qui n'est pas pris en compte dans l'analyse théorique. De plus, le champ de force produit par l'attraction terrestre n'est pas uniforme, car les masses ne sont pas réparties de manière symétrique au sein de la croûte et du manteau du globe terrestre⁴². Il faut également prendre en compte l'attraction gravitationnelle due aux autres objets célestes, comme la Lune et le Soleil qui induisent le phénomène de marée (y compris la marée terrestre, moins connue que la marée pour les océans, qui fait monter et descendre le sol d'environ 40 cm). Des variations permanentes de l'ordre de $5 \cdot 10^{-4} \text{ g}$ sont ainsi engendrées, ainsi que des variations cycliques de l'ordre de $3 \cdot 10^{-7} \text{ g}$. On a également enregistré des variations sur des échelles de temps plus grandes allant de 10^{-8} à 10^{-9} g . Ces variations locales de la pesanteur sont représentées par les géophysiciens par la hauteur du *géοïde*, qui est la surface virtuelle de la Terre associée à la différence locale de la pesanteur mesurée par rapport à celle que l'on obtiendrait si la Terre était un ellipsoïde liquide parfait de densité uniforme⁴³. Les variations de la surface du géοïde vont de -107 mètres à $+85,4$ mètres.

- La mesure est une mesure d'accélération, et sa précision est donc limitée par celle de notre dispositif de mesure de l'accélération, dans notre cas de celui du temps.

⁴¹ L'augmentation de la pesanteur lorsqu'on se déplace de l'équateur vers le pôle en raison du carré du sinus de la latitude a été établie par Pierre Louis de Maupertuis en 1737, à la suite des mesures qu'il a effectuées en Laponie en 1736 et 1737. Ces observations seront complétées quelques années plus tard par l'expédition de Pierre Bouguer et Charles-Marie de la Condamine au Pérou en 1735-1744. On prouva ainsi que la Terre était aplatie, comme le soutenait Newton, et non allongée, comme le défendait Cassini.

⁴² L'une des premières mesure précise locale des variations locales de la pesanteur est due au physicien hongrois Loránd von Eötvös (1848-1919), qui utilisa une balance à torsion pour mesurer la pesanteur près de Budapest en 1901. Les premières mesures à bord d'un sous-marin furent effectuées par le néerlandais Felix Andries Vening Meinesz entre 1923 et 1939.

⁴³ La densité moyenne de la Terre est aujourd'hui estimée à $5,527$. Cette valeur est très supérieure à celle de la croûte terrestre (environ 2,5 plus grande). Ce qui rend d'autant plus remarquable la valeur de $5,48$ annoncée par Henry Cavendish (1731-1810) en 1798, à une époque où l'on ignorait la composition interne de la Terre ! La masse de la Terre est actuellement estimée à $5,9736 \cdot 10^{24} \text{ kg}$. On la décompose comme suit : atmosphère = $5,1 \cdot 10^{18} \text{ kg}$, océans = $1,4 \cdot 10^{21} \text{ kg}$, croûte = $2,6 \cdot 10^{22} \text{ kg}$, manteau = $4,043 \cdot 10^{24} \text{ kg}$, noyau externe = $1,835 \cdot 10^{24} \text{ kg}$, noyau interne (graine) = $9,675 \cdot 10^{22} \text{ kg}$.

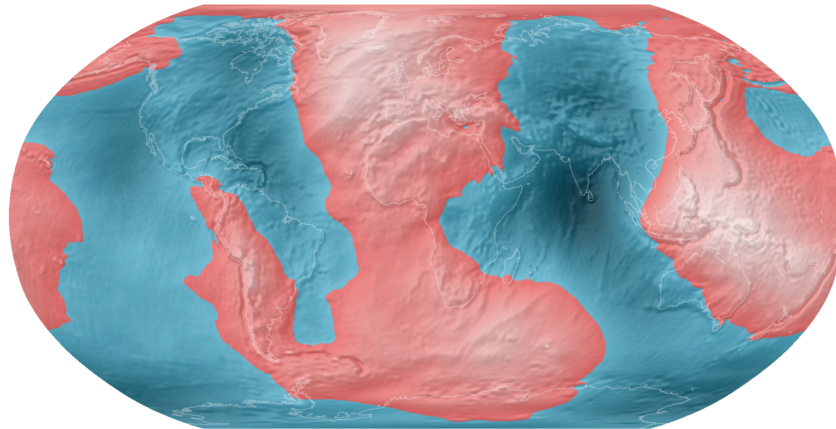


- 1 : océan
- 2: ellipsoïde de référence
- 3: direction normale à la surface terrestre
- 4: croûte terrestre
- 5: géoïde

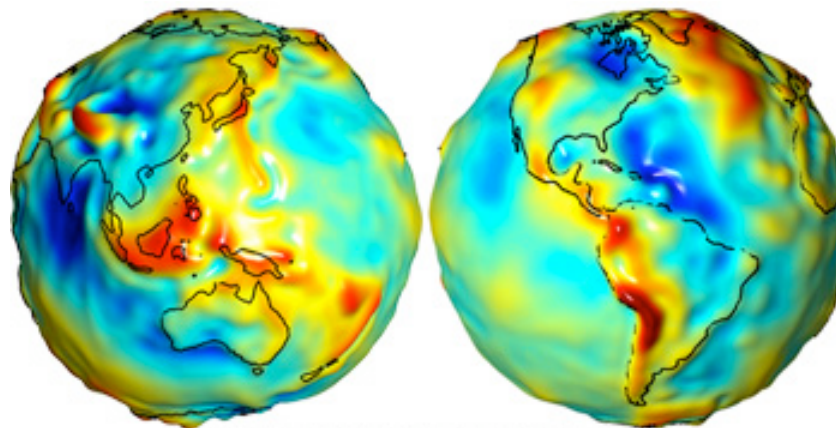
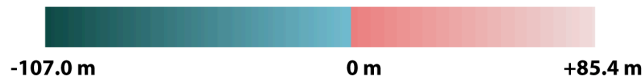
Figure 22: définition du géoïde terrestre

Deviation of the Geoid from the idealized figure of the Earth

(difference between the EGM96 geoid and the WGS84 reference ellipsoid)



Red areas are above the idealized ellipsoid; blue areas are below.



Earth's Gravity Field Anomalies (milligals)

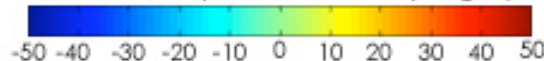


Figure 23: représentation du géoïde. Haut: isovaleurs de la hauteur du géoïde. Bas: représentation 3D des variations de la pesanteur terrestre

Notons enfin que, lors de l'analyse du fonctionnement de la machine d'Attwood, nous avons fait appel à ce qu'Einstein nomme le **principe d'équivalence**. Ce principe stipule que la masse inertielle (c'est-à-dire la masse m du corps qui apparaît dans l'énoncé de la deuxième loi de Newton) et la masse gravitationnelle (le coefficient m associé au corps dans la formule newtonienne de la force d'attraction gravitationnelle) sont égales. Cette égalité est déjà supposée par Galilée, puisqu'elle est à l'origine de son résultat sur l'universalité de la chute libre⁴⁴. Mais rien, dans la mécanique de Newton, n'indique cette égalité doit être obligatoirement vérifiée. Aussi, depuis trois siècles, de nombreuses expériences ont été menées pour corroborer expérimentalement cette hypothèse. Cet intérêt provient de l'importance de ce

⁴⁴ L'universalité de la chute libre signifie que tous les corps pesants tomberaient selon le même mouvement si ils étaient soumis à l'attraction terrestre en l'absence d'autre force, quelque soit leur masse.

principe : il est une des bases sur lesquelles repose la théorie de la relativité générale et il est un test pour les modèles d'unification des quatre forces fondamentales (gravité, force électromagnétique, interaction forte, interaction faible) développés par les physiciens théoriciens. De nombreuses techniques ont été employées pour valider cette égalité avec une précision remarquable : elle est aujourd'hui testée avec succès avec une précision meilleure qu'une partie sur 10^{12} . Cinq types de procédé technologique ont été mis en œuvre (les résultats sont regroupés dans le tableau ci-dessous):

- *Les expériences de chute libre.* Si il est aujourd'hui certain que Galilée n'a jamais pratiqué lui-même cette expérience, elle a depuis été réalisée à plusieurs reprises. Les plus récentes ont été pratiquées dans des tours de chute libre dans lesquelles on réalise un vide presque parfait. On met en œuvre des dispositifs de compensation de traînée pour réduire les effets du gaz résiduel.
- *Les pendules à fil.* Ce procédé, réellement employé par Galilée, repose sur la propriété d'isochronisme des pendules : la période de battement d'un pendule ne dépend que de la longueur du fil et de la pesanteur⁴⁵, et pas de la masse du corps pesant qui y est attaché. En posant le pendule sur un plan incliné très lisse, on réduit considérablement les effets de résistance aérodynamique due à l'air (car la vitesse atteinte est beaucoup plus faible que dans le cas de la chute libre).
- *L'observation du mouvement des planètes.* Il s'agit ici de vérifier que le mouvement observé des planètes et satellites correspond bien à celui prédit par la mécanique newtonienne. Depuis les années 1970, cette technique a été améliorée par des mesures très précises de la distance Terre-Lune au moyen de rayons Laser réfléchis par des miroirs disposés à la surface de la Lune.
- *Les balances de torsion.* Cet appareil a été promu comme un des procédés mécaniques les plus précis pour la mesure des forces faibles par Coulomb dans son *Mémoire* de 1784. Il est employé en 1798 par le physicien anglais Henry Cavendish (1731-1810) pour effectuer la première mesure de la constante G de la loi de gravitation publiée par Newton en 1686. Les balances de torsion sont toujours employées de nos jours pour mesurer la constante de gravitation.
- *Les mesures de chute libre effectuées dans l'espace* dans des laboratoires embarqués sur des satellites.

⁴⁵ La période du pendule est proportionnelle à la racine carrée du quotient de la longueur du fil par la pesanteur

Tableau 2: Récapitulatifs des expériences de validation du principe d'équivalence entre la masse inertielle et lamasse gravitationnelle. D'après "L'universalité de la chute libre", S. Reynaud, Dossier Pour la Science No 38, 2003

Procédé	Auteur	date	Niveau d'erreur
Chute libre	Galilée (virtuel)	Vers 1600	
	Faller	1980-1990	10^{-10}
	Dittus	2000	10^{-12}
Pendules à fil	Galilée	Vers 1600	10^{-2}
	Newton	Vers 1680	10^{-3}
	Bessel	Vers 1830	10^{-5}
	Potter	1920	10^{-6}
Mouvements planétaires	Newton	Vers 1680	
	Laplace	Vers 1780	10^{-7}
	Laser-Lune	2000	$2 \cdot 10^{-13}$
Balance de torsion	Eötvös	Vers 1920	10^{-8}
	Dicke	Vers 1960	10^{-11}
	Adelberger	Vers 1990	$3 \cdot 10^{-13}$
satellites	Projet MICROSCOPE	Vers 2005	10^{-15}

4.2.6 Exemple 2 : La chambre à bulles : « voir » les particules en mouvement

Le second exemple est celui de la chambre à bulles (détecteur mis au point en 1953 par le physicien américain Donald Arthur Glaser, prix Nobel de physique en 1960), qui permet de visualiser les trajectoires des particules de la physique atomique : atomes, électrons, ... Son principe de fonctionnement est le suivant : un liquide (souvent de l'hydrogène) placé dans une chambre, est comprimé par un piston de manière que sa température soit supérieure à sa température d'ébullition sous la pression atmosphérique, mais inférieure à sa température d'ébullition sous la pression à laquelle il se trouve placé. On fait ensuite passer les particules dans la chambre. Lors de leur mouvement, elles vont interagir avec les molécules proches d'elles et les ioniser. Juste après le passage de la particule à détecter, on détend le piston (pendant 1 ms environ), et le liquide revient à la pression atmosphérique. L'ébullition débute autour des ions créés par la particule. Si on éclaire par un flash et que l'on photographie, la trajectoire est matérialisée par un chapelet de petites bulles.



Figure 24: D. Glaser

Une illustration est donnée dans la figure ci-dessous.

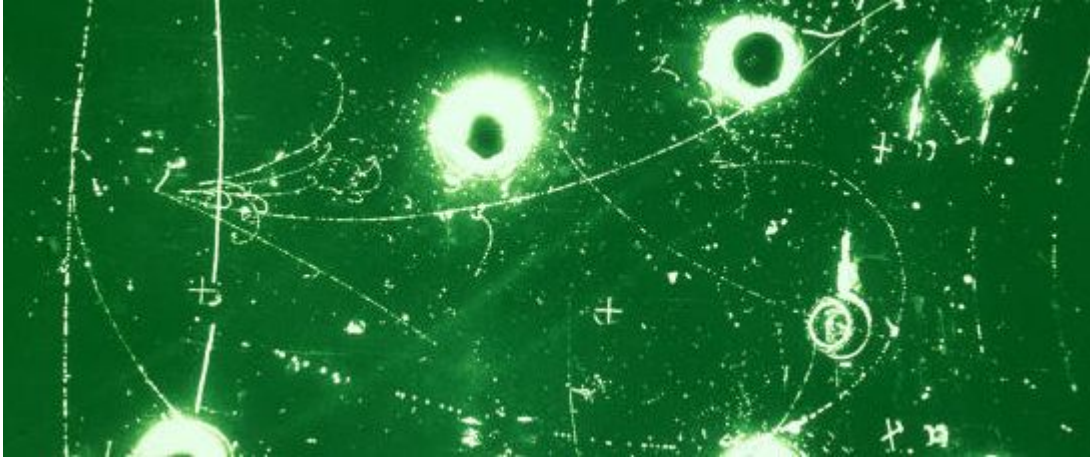


Figure 25: visualisation de trajectoires de particules élémentaires dans une chambre à bulles

Qu'observons nous ? Des lignes blanches, plus ou moins identifiables, qui sont interprétées comme étant les trajectoires des particules élémentaires qui ont été projetées dans la chambre.

Mais, de manière plus exacte, que voyons-nous ? Les particules en mouvement ? Non. Nous observons un cliché photographique du liquide contenu dans la chambre à bulles qui, convenablement éclairée, montre la présence de bulles. Analysons les différentes étapes qui permettent d'établir le lien entre le fait brut (nous voyons une photo en couleur qui présente des formes reconnaissables) et son interprétation (une particule a suivi telle trajectoire).

La première étape consiste à faire le lien entre les différences de couleurs sur la photo et des phénomènes physiques. Les différences de couleur du cliché sont causées par des variations d'exposition des molécules photo-sensibles du film. Ces variations indiquent que des radiations électromagnétiques de longueurs d'onde variables ont été émises par le liquide contenu dans la chambre. Ces variations sont interprétées comme étant le résultat de fluctuations des propriétés optiques de la matière dans la chambre. Les zones blanches sont associées à des zones contenant du gaz (les bulles) le reste aux régions occupées par le liquide.

La seconde étape consiste à relier les bulles et les particules. Comme dit plus haut, les bulles résultent de la vaporisation du liquide, vaporisation qui commence autour des ions. On fait donc ici appel à deux contenus théoriques : la physique du changement de phase (pourquoi il y a vaporisation et pourquoi autour des ions) et la physique de l'interaction entre la particule étudiée et les molécules du liquide contenu dans la chambre. L'utilisation de ces deux corpus théoriques est nécessaire, et n'est pas justifiée par la photo : on ne « voit » pas les molécules d'hydrogène passer à l'état d'ions lors d'une interaction, ni la vaporisation s'effectuer (pas plus qu'on ne voyait une force se transmettre du fil aux solides dans le cas de la machine d'Atwood) : on prend comme hypothèse, pour concevoir l'instrument virtuel, que ces phénomènes seront présents.

La chambre à bulles combine donc la chimie, l'électromagnétisme et la physique atomique. Elle mobilise donc un ensemble de théories, de modèles et d'hypothèses plus vaste que la machine d'Atwood. Son utilisation ne permet pas d'observer

directement les particules en mouvement, mais seulement une trace de leur passage : il s'agit là d'une **mesure indirecte**.

4.2.7 Exemple 3 : Le microscope à effet tunnel : « voir » les atomes

Peut-on « voir » les atomes, même si ceux-ci ne bougent pas ? Oui, avec un microscope à effet tunnel, dont l'invention a valu le prix Nobel de physique à Gerd Binnig et Heinrich Rohrer en 1986. Les bases théoriques impliquées sont très avancées (l'effet tunnel relève de la mécanique quantique). Nous nous bornerons donc ici à donner les grands principes de fonctionnement de cet instrument, pour illustrer la très grande complexité théorique qui peut être enfouie dans la mesure.



Figure 26: G. Binnig

Le microscope à effet tunnel (*STM: Scanning Tunneling Microscope*) utilise l'émission par effet tunnel d'électrons entre l'échantillon et une sonde en forme de pointe métallique lorsque celle-ci parcourt la surface de l'échantillon à très faible distance.



Figure 27: H. Rohrer

L'effet tunnel est une des prédictions de la mécanique quantique. En mécanique classique, une particule, qui n'a pas assez d'énergie pour surmonter une barrière de potentiel de hauteur V_0 , reste indéfiniment bloquée du côté de la barrière où elle se trouve initialement. En mécanique quantique, par contre, la fonction d'onde associée à cette même particule n'est pas nulle de l'autre côté de la barrière. Il y a donc une certaine probabilité pour que cette particule franchisse la barrière, et ce même si son énergie est inférieure à V_0 ; c'est ce que l'on appelle l'effet tunnel. Tout se passe comme si la particule empruntait un " tunnel " au travers de la barrière de potentiel.

Lorsque la pointe de la sonde parcourt la surface, des électrons sont échangés par effet tunnel entre celle-ci et les atomes de la surface. L'intensité du courant induit est fonction de la distance entre l'atome de la pointe et celui de l'échantillon. En mesurant ce courant, on accède donc à une mesure de cette distance. En balayant la surface, on établit une cartographie de la surface de l'échantillon à l'échelle atomique. Ce principe est illustré sur la figure ci-dessous :

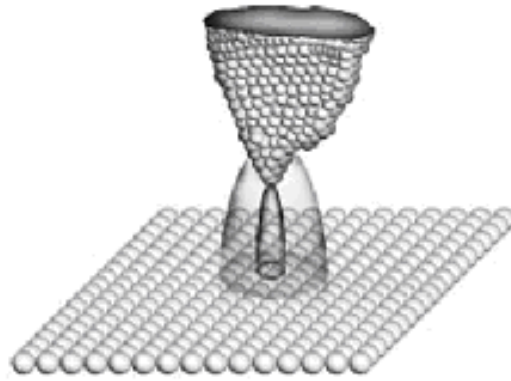


Figure 28: principe du microscope à effet tunnel - un courant électrique s'établit grâce à l'effet tunnel

Le schéma complet d'un microscope à effet tunnel est donné dans la figure suivante

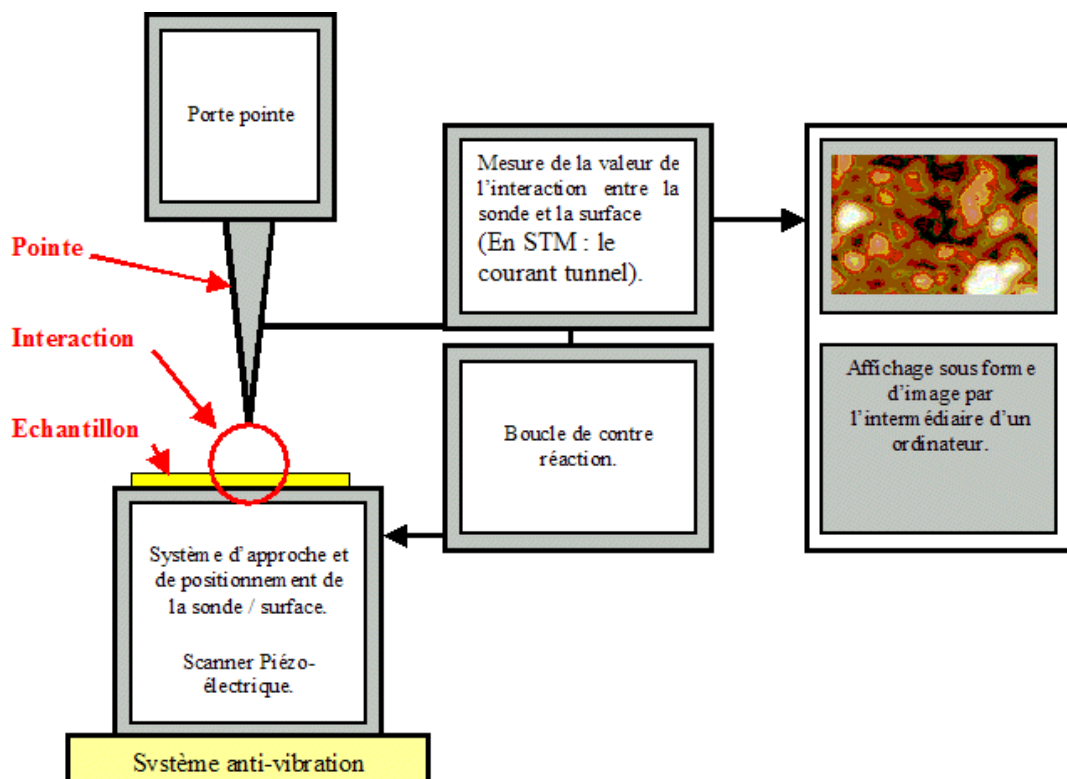


Figure 29: schéma du montage d'un microscope à effet tunnel

Un exemple de visualisation (atomes d'or sur une surface plane) est montré ci-dessous. Les régions jaunes indiquent les régions où la surface est élevée (présence d'un atome) et les zones rouges les régions de faible altitude (vide entre les atomes). On retrouve bien l'organisation des atomes telle qu'elle est attendue (réseau organisé).

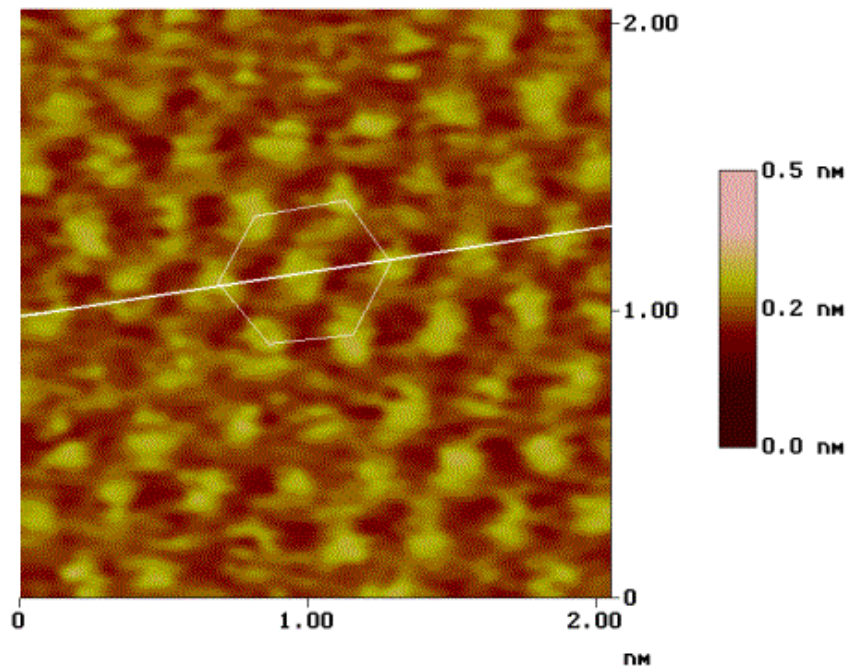


Figure 30: visualisation des atomes d'or par microscopie à effet tunnel

Comme dans les exemples précédents, la mesure est indirecte (on ne voit pas les atomes au sens commun du terme, on reconstruit une représentation visuelle par ordinateur à partir de mesures quantifiées de courants électriques) et présuppose, de la part de l'observateur, l'emploi d'un domaine scientifique très large (mécanique quantique, électronique, informatique, mathématiques, ...).

4.2.8 Exemple 4 : Les lentilles gravitationnelles : « voir » les objets lointains⁴⁶

Les lunettes optiques sont employées depuis des siècles par les astronomes pour observer les objets cosmiques (planètes, étoiles, galaxies, amas). Le principe de base est simple : une lentille de verre est utilisée pour faire converger les rayons lumineux parallèles issus de sources très lointaines en un point unique bien défini, appelé point focal. La forme convexe de la lentille est conçue de manière à obtenir une efficacité maximale.

La théorie de la relativité générale publiée par Albert Einstein en 1915 prédit (ce qu'Einstein fera explicitement en 1916) que les rayons lumineux sont défléchis par le champ gravitationnel des objets massifs. Cette prédiction sera vérifiée en mai 1919 par deux expéditions (une à Soral au Brésil, l'autre sur l'île de Principe au large de l'Afrique) dirigées par Arthur Eddington lors d'une éclipse totale du Soleil : lorsque le Soleil passe devant les étoiles, leurs images se déplacent de 1,75 secondes d'angle. Cette propriété fait que certains objets très massifs (galaxies, étoiles) peuvent se comporter comme des lentilles optiques « naturelles » qui grossissent la taille des objets observés et amplifient leur éclat. Les astronomes et les astrophysiciens les utilisent pour voir des objets qu'ils ne pourraient observer avec les moyens optiques

⁴⁶ Voir « Les mirages gravitationnels », J. Wambsganss, Dossier Pour La Science No 38, 2003

classiques. En effet, les lentilles gravitationnelles sont l'une des rares méthodes permettant de cartographier la matière noire⁴⁷ dans l'Univers. Elles permettent également de détecter des trous noirs et de sonder la structure interne des quasars⁴⁸.

L'idée que la lumière puisse être sensible au champ de gravitation n'est pas neuve. En se basant sur la théorie corpusculaire de la lumière de Newton, et donc en attribuant une masse aux particules de lumière qui les rend sensible à la gravité, le physicien anglais John Mitchell avait imaginé en 1783 le concept d' « étoile noire ». Ce concept fût ensuite repris par Pierre Simon de Laplace sous le nom d'astre « occlus ». En 1801, l'astronome et géographe allemand Johann Georg von Soldner émit l'idée que les particules de lumière adopteraient la même trajectoire que des petites billes massives lancées à la même vitesse en passant à proximité du Soleil. Toutefois, après 1808, lorsque Thomas Young eut démontré que la lumière se comportait comme une onde, le modèle ondulatoire initié par Christiaan Huygens gagna en popularité et l'idée que la lumière puisse être sensible à la gravité cessa d'être explorée jusqu'aux travaux d'Einstein.

Les lentilles gravitationnelles présentent quelques différences avec leurs homologues classiques. En effet, pour une lentille classique, les rayons sont d'autant plus déviés qu'ils sont éloignés de l'axe optique, ce qui permet la focalisation en un point unique. Pour une lentille gravitationnelle, les rayons les plus éloignés de l'axe sont les moins déviés. En conséquence, la focalisation se produit sur une ligne ou une surface et non un point. Dans une lentille classique en verre, la déviation dépend de la longueur d'onde de l'onde lumineuse. Ceci est dû au fait que la déviation provient du ralentissement de la lumière dans le verre, et que ce ralentissement dépend de la longueur d'onde. Avec une lentille gravitationnelle, toutes les longueurs d'onde sont déviées de la même manière, puisque la lumière se propage dans le vide. Les quatre principaux effets induits par les lentilles gravitationnelles, parfois appelés *mirages gravitationnels* par analogie avec les mirages optiques « ordinaires », sont :

- Le changement position apparente des objets
- Le grossissement : on a observé des grossissements allant jusqu'à 100 (mais également des grossissement négatifs !)
- La déformation des objets, qui sont souvent étirés le long d'un cercle centré sur la masse défectrice
- La multiplication de l'image des objets (pour les lentilles gravitationnelles les plus puissantes)

⁴⁷ On estime aujourd'hui que la matière noire représente plus de 90 % de la masse de l'Univers.

⁴⁸ Les quasars sont des noyaux de galaxies primitives qui peuplaient l'Univers il y a 10 milliards d'années. Ce sont donc des objets très lointains (parmi les plus éloignés jamais observés !) qui apparaissent comme des sources ponctuelles très puissantes.

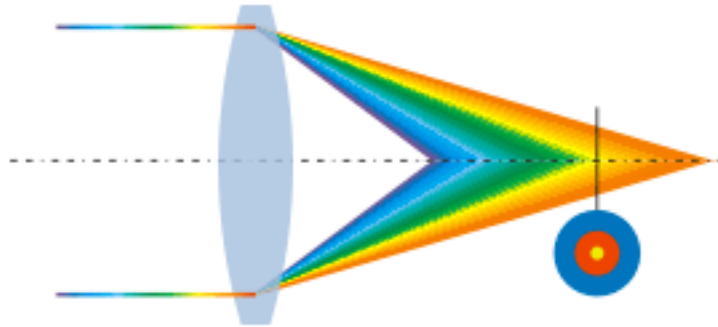


Figure 31: principe de la lentille optique classique

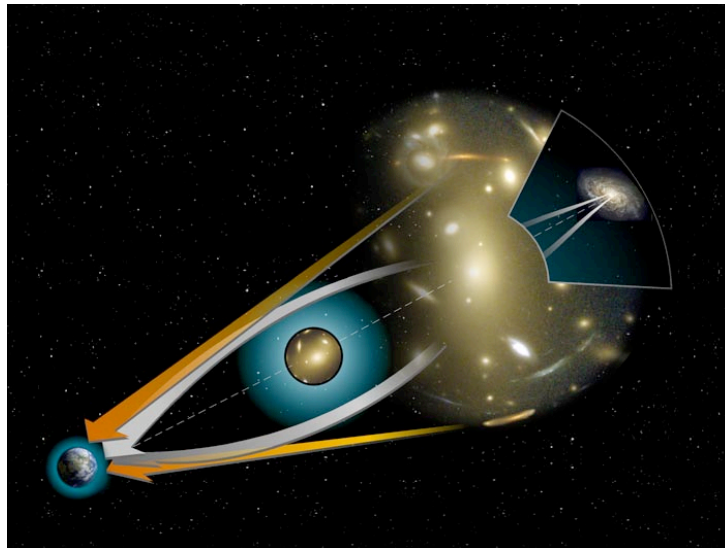


Figure 32: principe de la lentille gravitationnelle

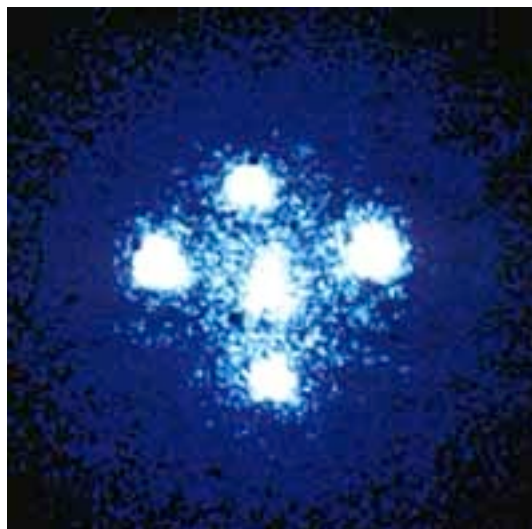


Figure 33: exemple de mirage gravitationnelle sur un Quasar. L'image originale est multipliée, formant la "croix d'Einstein"

4.3 L'observation n'est pas neutre

Les discussions et les exemples précédents montrent que l'observation n'est pas neutre, et cela pour deux raisons.

Tout d'abord, on peut dire que les faits sont « chargés de théorie » : l'observateur ne peut appréhender dans les faits bruts que ce que son cadre conceptuel théorique lui permet d'identifier, de reconnaître et d'interpréter: la machine d'Attwood ne nous apprend rien si nous ne savons rien de la mécanique newtonienne et que les concepts de force, de masse et d'accélération nous sont inconnus. Tout ce que nous verrions, c'est un assemblage de fils, de poulies et de solides qui bougerait dans certaines conditions. C'est ce que dit Auguste Comte (1798-1857) lorsqu'il écrit dans son *Cours de philosophie positive* : « Si d'un côté toute théorie positive doit nécessairement être fondée sur des observations, il est également sensible, d'un autre côté, que, pour se livrer à l'observation, notre esprit a besoin d'une théorie quelconque. Si, en contemplant les phénomènes, nous ne les rattachions point immédiatement à quelques principes, non seulement il nous serait impossible de combiner ces observations isolées et, par conséquent, d'en tirer aucun fruit, mais nous serions même entièrement incapables de les retenir ; et, le plus souvent, les faits resteraient inaperçus sous nos yeux. » Plus loin, il ajoute : « En quelque ordre de phénomènes que ce puisse être, même envers les plus simples, aucune véritable observation n'est possible autant qu'elle est primitivement dirigée et finalement interprétée par une théorie quelconque. » Dans ce passage, il faut comprendre observation neutre lorsque Comte écrit observation.

L'observation et l'expérimentation nécessitent l'existence d'un cadre théorique pour pouvoir être menées à bien, c'est-à-dire pour que le lien puisse être établi entre des sensations (visuelles dans les cas étudiés plus haut) et des concepts scientifiques. De plus, les observations se font toujours dans le cadre de recherches scientifiques, et sont donc orientées : d'une certaine manière, on peut dire que l'on ne trouve que ce que l'on cherche.

Les exemples de découverte accidentelle, comme celle des rayons X par le physicien allemand Wilhelm Conrad Röntgen (1845-1923), qui reçut le prix Nobel de physique pour cela en 1901, montrent que découverte accidentelle ne veut pas dire « découverte sans cadre théorique ». En novembre 1895, Röntgen réalisait des expériences sur les tubes cathodiques. Pour ne pas être gêné par la lumière produite par le tube, il le recouvre de papier noir opaque. Le 8 novembre, en alimentant son circuit à haute tension, il observe par hasard qu'un écran mis à proximité du tube devient fluorescent. Il entame alors une série d'expériences. Ne connaissant pas la nature de ces rayons mystérieux, il les nomme « rayons X ». Le 28 novembre, il réalise la première radiographie : celle de la main de sa femme (figure ci-dessous).



Figure 34: W. Röntgen



Figure 35: première radiographie de l'histoire : la main de la femme de Röntgen (28/11/1895)

Outre son aspect ludique, l'anecdote montre que, si la découverte est accidentelle, elle n'a pu avoir lieu que parce que Röntgen possédait les outils théoriques nécessaires pour identifier qu'il se passait quelque chose qui méritait d'être analysé. Dans ce cas, on ne trouve pas ce que l'on cherche, on observe ce que l'on pensait ne pas voir (le tube étant rendu opaque, il ne pouvait pas éclairer l'écran si seuls les rayons visibles étaient émis).

Les mesures quantitatives sont effectuées à partir d'instruments qui, eux aussi, sont « chargés de théories ». La mesure n'est donc possible que si un cadre théorique qui permet l'interprétation des faits est disponible. Faire une mesure, c'est donc mobiliser des théories au travers des instruments.

La seconde raison pour laquelle on peut dire que la mesure n'est pas neutre, c'est que mesurer, c'est perturber l'objet que l'on veut soumettre à la mesure. Toute mesure est basée sur la quantification d'une différence, d'une variation, et c'est d'ailleurs à cela que nos sens réagissent : différence d'exposition lumineuse sur les cellules de la rétine, différence de pression sur la peau pour le toucher, ... La machine d'Attwood fonctionne grâce aux différences des forces exercées, la chambre à bulles sur les différences d'exposition des molécules photosensibles du film, le microscope à effet tunnel sur des différences d'intensité électrique. Pour mesurer cette différence, les procédés de mesure produisent une perturbation de l'objet



Figure 36: W. Heisenberg

soumis à la mesure : on soumet les atomes de l'échantillon à l'effet tunnel, on fait interagir les particules avec des atomes d'hydrogène dans la chambre à bulles, on suspend le solide et on le met en mouvement avec la machine d'Attwood. Ces perturbations affectent toujours l'objet à mesurer. On ne mesure donc les propriétés que d'un objet en interaction, jamais d'un objet isolé, au repos. Ceci est intrinsèque, puisque mesurer, c'est mettre en interaction l'objet à mesurer et l'instrument de mesure. Mais peut-on alors espérer obtenir des informations sur l'objet « isolé » : oui, si la perturbation induite par la mesure est faible en intensité devant les propriétés de l'objet lorsqu'il n'est pas soumis à la

perturbation. Peut-on espérer, grâce aux progrès technologiques, développer un instrument parfaitement non-intrusif ? Non : la relation théorique due au physicien allemand Werner Karl Heisenberg (1901-1976, prix Nobel de physique 1932), appelée **principe d'incertitude** ou **principe d'indétermination** montre qu'il existe un seuil intrinsèque de résolution de la mesure, dû à cette perturbation. On ne peut pas accéder au même moment aux valeurs exactes de toutes les grandeurs physiques associées à un système. Il ne peut donc pas exister de mesure neutre (voir le chapitre 11).

4.4 Perception et fiabilité des observations

Les expériences, mêmes menées à l'aide d'instruments qui servent à quantifier et mesurer, reposent sur notre perception sensorielle des signaux reçus et sur l'interprétation qu'en fait notre cerveau. Il convient donc de s'interroger sur les possibles sources d'erreur ou de biais introduites par nos organes sensoriels et notre cerveau.

4.4.1 Réception et perception⁴⁹

Il faut tout d'abord distinguer entre la **réception** des signaux extérieurs (lumière, ondes de pression associées au son, molécules correspondant au goût, ...), qui est un mécanisme strictement physique et physiologique, de la **perception**, qui est la manière dont notre esprit interprète les messages reçus. L'observation scientifique est bien sûr basée sur notre perception du monde extérieur, et nous allons donc voir brièvement les principaux résultats admis concernant la perception humaine.

Avant cela, rappelons les limites physiologiques de la réception pour la vue et l'ouïe. Les limitations (en terme de longueur d'onde et/ou de fréquence et d'intensité) pour la vue⁵⁰ et l'ouïe chez l'homme « normal » sont illustrées dans la figure et le tableau ci-dessous.

⁴⁹ Voir « L'espace et le temps aujourd'hui », E. Noël, Editions du Seuil, Coll. Points Sciences S36, 1983

⁵⁰ L'œil humain est un détecteur très performant capable de percevoir des signaux très faibles : on sait qu'il suffit d'un seul photon pour déclencher un signal optique percevable. Cette performance est due à une amplification biochimique en chaîne : l'arrivée d'un seul photon induit le blocage du passage d'environ 10^8 ions et la création d'un signal électrique d'une intensité d'un demi-picoampère. La rétine comprend environ 160 millions de capteurs photorécepteurs, et il envoie des signaux au cerveau par environ un million de canaux.

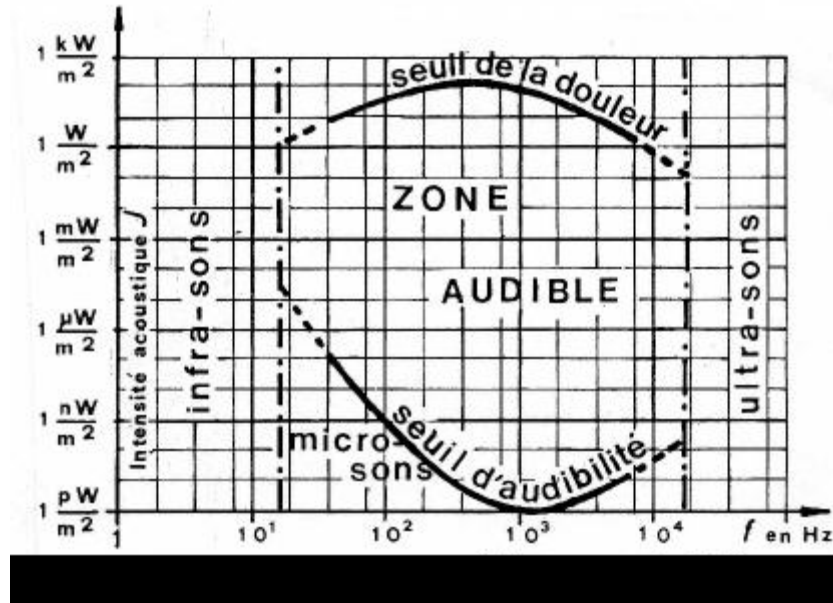


Figure 37 : limites de perception de l'oreille humaine

Tableau 3: spectre des ondes électromagnétiques

Longueur d'onde	Domaine
> 10 cm	Radio
De 1 mm à 10 cm	Micro-onde et radar
De 1 μm à 500 μm	Infrarouge
De 400 nm à 700 nm :	Lumière visible
Rouge (620-700 nm)	
Orange (592-620 nm)	
Jaune (578-592 nm)	
Vert (500-578 nm)	
Bleu (446-500 nm)	
Violet (400-446 nm)	
De 10 nm à 400 nm	Ultraviolet
De 10 ⁻⁸ m à 10 ⁻⁷ m	
De 10 ⁻¹¹ m à 10 ⁻⁸ m	Rayons X
De 10 ⁻¹⁴ m à 10 ⁻¹² m	Rayons Gamma

La perception est un phénomène complexe, dont l'étude fait appel à de nombreuses disciplines : psychologie, psychophysiologie, pratiques de l'esthétique, linguistique, ... Chaque discipline appréhende le phénomène de perception sous un angle différent.

De très nombreux travaux de psychologie expérimentale ont été consacrés à l'étude de la perception de l'espace et du temps. Un premier résultat important est qu'il est abusif de dire que nous percevons l'espace et le temps. En effet, nous ne percevons pas le temps, mais des événements qui ont une certaine durée et qui se succèdent dans un certain ordre. De même, nous ne percevons pas l'espace mais des objets qui ont tous une certaine étendue, et une relation de position et d'orientation entre eux et par

rapport à nous. Jusqu'aux années 1960, il était courant de traiter séparément la perception de l'espace et celle du temps. Depuis, les approches « modernes » visent une description globale centrée sur les processus mentaux qui permettent de traiter les informations transmises au cerveau, quelque soit leur origine, la perception étant comprise comme un phénomène essentiellement psychologique.

Concernant la perception du temps, il existe aujourd'hui de nombreux résultats sur la perception de la durée et sur celle de la succession. Par exemple, une stimulation auditive est estimée plus longue qu'une stimulation visuelle, et une stimulation intense est perçue comme plus longue qu'une stimulation d'intensité plus faible. Une conséquence est que les intervalles courts ont tendance à être sous-estimés, alors que les intervalles longs sont surestimés. Les études sur la perception de la succession ont porté sur plusieurs paramètres : seuils de fusion d'évènements, de simultanéité d'évènements et d'intervalles suffisants pour détecter l'ordre des évènements.

Les deux questions les plus abordées concernant la perception de l'espace sont la perception des distances et celle de la taille d'un objet. Pour percevoir les distances (ou plus précisément les estimer), notre cerveau utilise trois types d'indices :

- Les *indices monoculaires statiques*. Quelques exemples : la présence dans le champ visuel d'objets familiers dont la taille est connue, ce qui permet des comparaisons ; les ombres ; l'interposition des objets, l'objet masquant étant plus proche que l'objet masqué ; la perspective géométrique ; la texture est également très importante, puisque nous ne percevons que des surfaces (la rétine est une surface bidimensionnelle !) dont les variations de texture permettent d'apprécier l'inclinaison et la forme d'un objet.
- Les *indices monoculaires dynamiques*, comme la parallaxe du mouvement (la vitesse angulaire perçue d'un objet mobile dépend de sa distance)
- Les *indices binoculaires*, qui proviennent des différences enregistrées par les deux yeux.

La perception auditive est également un moyen de localisation dans l'espace. Le cerveau utilise alors les différences d'intensité et le décalage entre les signaux perçus par chaque oreille pour reconstruire un schéma tridimensionnel.

Une question encore ouverte est de savoir comment le cerveau combine les différents indices et les différentes perceptions pour construire une image globale du monde qui l'environne. Cette question des schémas mentaux qui sont à la base de cette reconstruction est cruciale, puisque ce sont ces schémas qui gouvernent notre perception du monde et, dans le cadre qui nous intéresse, de notre capacité à recueillir des informations par l'observation. Ces schémas sont actifs : outre la reconstruction à partir des informations, ils structurent l'information reçue en hiérarchisant les signaux. Ils agissent comme des filtres distordants. Ce phénomène est très connu en psychoacoustique pour l'étude de la perception des sons et des structures rythmiques.

De multiples questions se posent à propos de ces schémas mentaux. Une question connue est celle de leur fiabilité, qui n'est pas totale. Cela est illustré par les illusions sensorielles, les plus connues étant les illusions optiques (voir les figures ci-dessous).

Mais il existe aussi des illusions auditives⁵¹, des illusions olfactives, ... Chacun des cinq sens peut induire des illusions. Une autre question est celle de la nature de ces schémas mentaux : sont-ils innés, ou acquis par l'expérience ? Les résultats d'expériences menées sur des enfants âgés de quelques minutes et sur des mammifères montrent que dès la naissance, ces schémas existent, mais que l'expérience acquise au cours des années va permettre de les affiner et de les modifier. Les aspects physiologiques et anatomiques du cerveau, qui contient environ 10^{11} neurones, sont également étudiés. Par exemple, il semble acquis aujourd'hui que l'hémisphère gauche du cerveau est spécialisé pour l'analyse des durées et des successions, l'hémisphère droit étant principalement impliqué dans la perception de l'espace. L'hémisphère droit est également supérieur pour le maniement de concepts géométriques, cette supériorité allant croissant avec la complexité de la géométrie (la géométrie la plus « simple » pour le cerveau semblant être la géométrie euclidienne – voir le paragraphe 4.4.2).

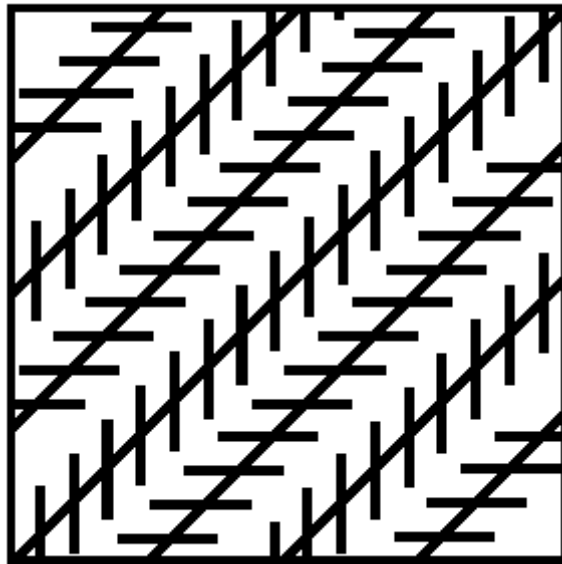
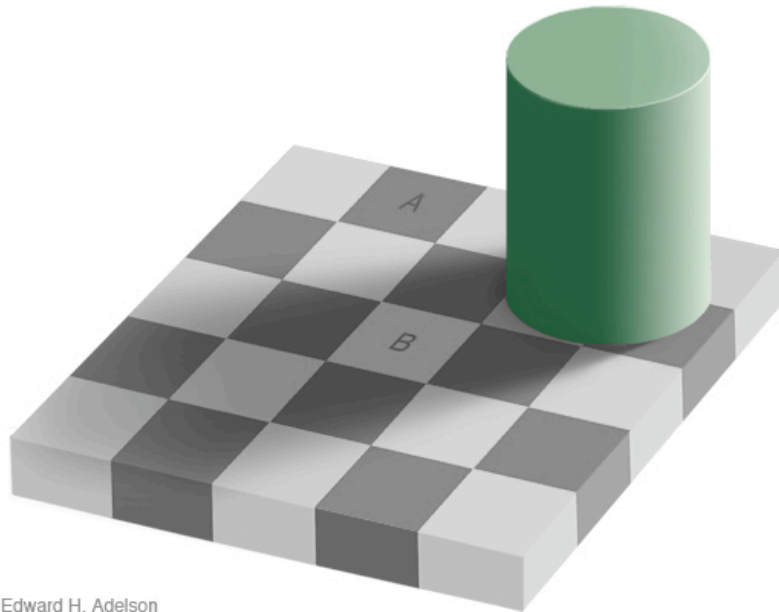


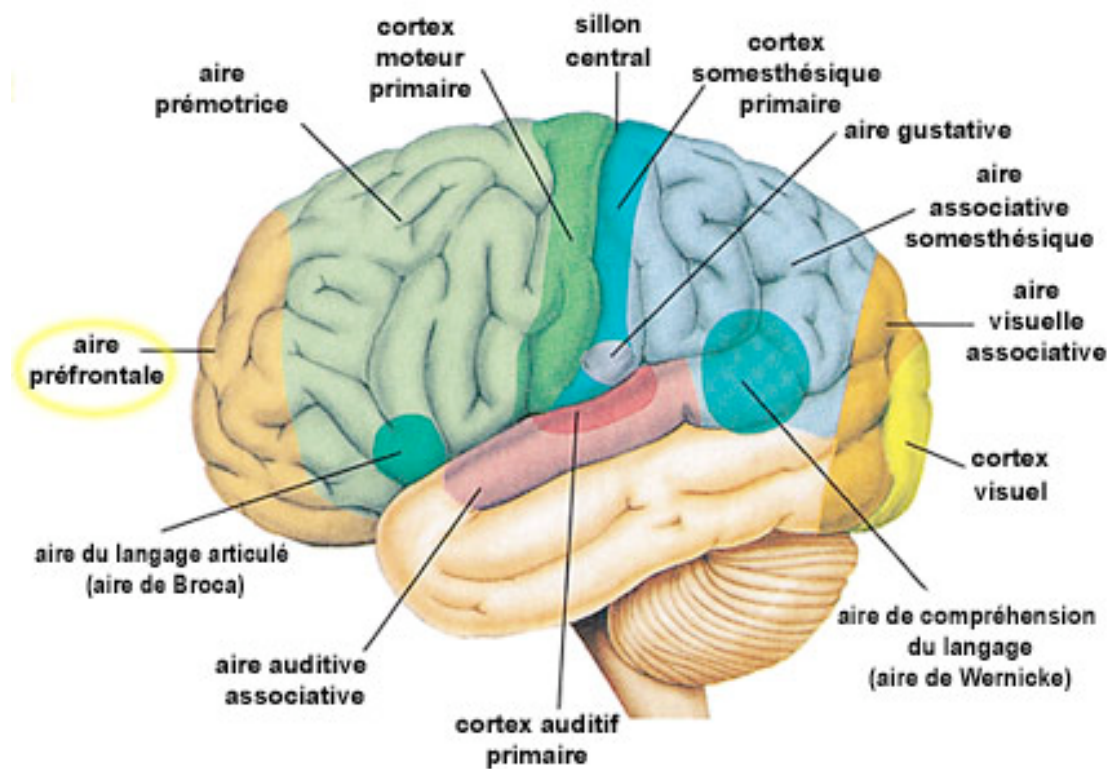
Figure 38: les droites sont-elles parallèles ? (Illusion de Zöllner)

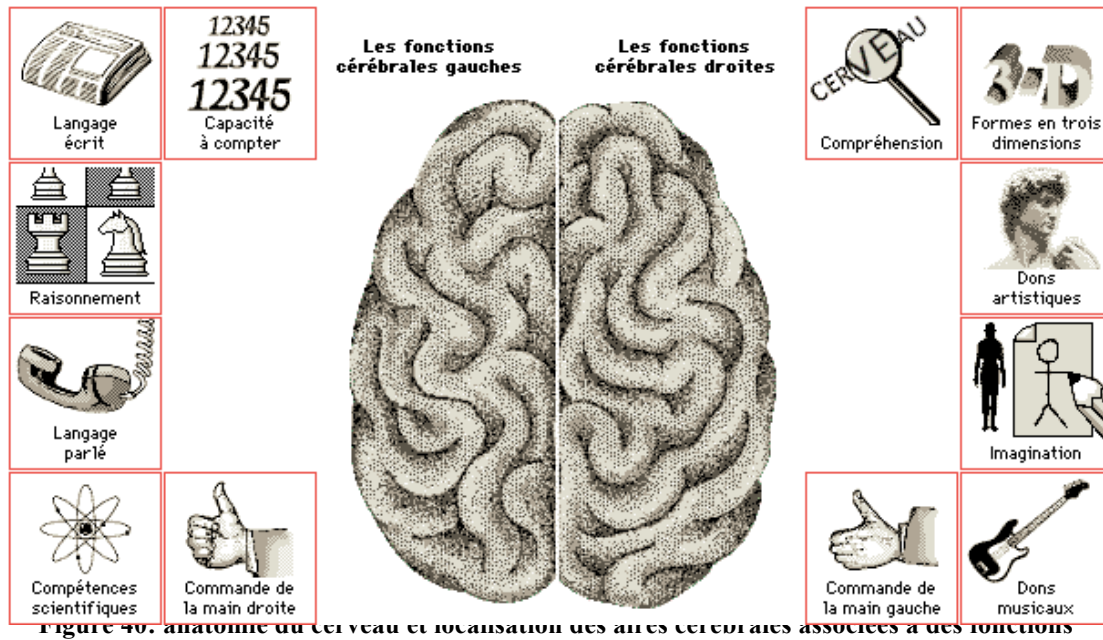
⁵¹ Deux exemples célèbres sont les gammes de Roger Shepard et de Jean-Claude Risset (disponibles sur <http://asa.aip.org/sound.html>)



Edward H. Adelson

Figure 39: les couleurs des cases A et B sont-elles différentes ? (échiqier de Adelson)





4.4.2 Perception de l'espace et géométrie(s)

On peut également voir des choses qui violent notre conception de la réalité, comme le montre l'image « impossible » ci-dessous. Bien que perçue, et ne relevant pas du domaine de l'illusion optique, cette image est « choquante » car elle ne correspond pas à notre intuition de ce qui est possible dans notre espace à trois dimensions.

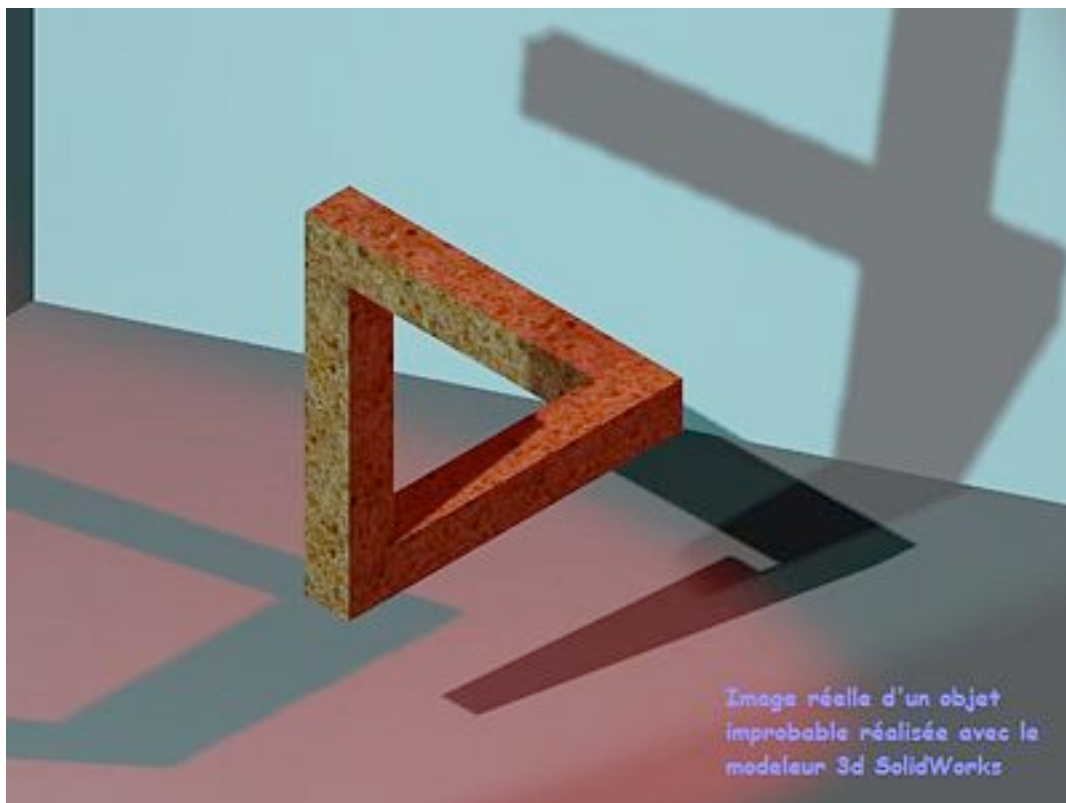


Figure 41: le triangle de Penrose est une image "absurde". Pourquoi ?

Notre perception intuitive de l'espace semble liée à la géométrie euclidienne⁵². Ou, plus précisément, la géométrie définie par Euclide dans ses *Eléments*⁵³ procède de notre perception visuelle du monde. C'est ce que le mathématicien allemand David Hilbert disait lorsqu'il écrivait, dans son article sur *Les fondements de la Géométrie* (1899): « *Ce problème [la géométrie] est celui de l'analyse de notre intuition de l'espace* ». Pour comprendre cela, il faut se rappeler que la démonstration mathématique, depuis l'essor des mathématiques grecques jusqu'au développement de l'analyse, donc pendant près de 2 000 ans, a été basé sur la construction de figures géométriques sur un plan à l'aide d'une règle et d'un compas. La démonstration était donc subordonnée à la possibilité de dessiner et de voir des objets sur une figure. La géométrie euclidienne est donc subordonnée à la structure de l'espace telle que nous la percevons directement sans instrument de mesure.

D'autres géométries, qui seraient non-euclidiennes, existent-elles et sont-elles utiles pour décrire l'univers? La réponse à ces deux questions est oui. Tout d'abord, rappelons que la géométrie euclidienne repose sur les cinq postulats suivants :

5 postulats de la géométrie euclidienne

1. Un segment de droite peut être tracé en joignant deux points quelconques.
2. Un segment de droite peut être prolongé indéfiniment en une ligne droite.
3. Etant donné un segment de droite quelconque, un cercle peut être tracé en prenant ce segment comme rayon et l'une de ses extrémités comme centre.
4. Tous les angles droits sont congruents.
5. Si deux lignes sont sécantes avec une troisième de telle façon que la somme des angles intérieurs d'un côté est strictement inférieure à deux angles droits, alors ces deux lignes sont forcément sécantes de ce côté.

Le cinquième postulat peut être reformulé sous la forme du célèbre postulat des parallèles :

Postulat des parallèles d'Euclide

Soient une droite d et un point p n'appartenant pas à d . Il existe une unique droite parallèle à d dans le plan engendré par d et p et passant par p .

C'est la violation du postulat des parallèles qui est sans aucun doute le caractère le plus spectaculaire des géométries non-euclidiennes, puisque le dessin sur une feuille de papier ne semble pas indiquer d'autre solution possible que celle d'Euclide.

Le développement de nouvelles géométries nécessite l'emploi d'outils mathématiques plus sophistiqués que la démonstration par construction de figures. Après le développement de ces outils, sous l'impulsion de savants tels que René Descartes et Pierre Fermat, il devient possible de s'abstraire des perceptions sensorielles pour construire des géométries. Ce sont ces nouveaux outils qui permettront de statuer sur

⁵² Cela semble vrai en première approximation. Les études récentes sur la perception montrent que la disposition des objets dans l'image reconstruite par nos schémas mentaux présente des distorsions qui ne sont pas conformes à la géométrie euclidienne

⁵³ Dans ses éléments, Euclide introduit 23 définitions, 5 notions communes et 5 postulats.

l'impossibilité de résoudre à l'aide d'une règle et d'un compas les trois grands problèmes de la géométrie euclidienne que sont : la **quadrature du cercle**, la **trisection de l'angle** et la **duplication du cube**. Mais il faudra longtemps pour que de nouvelles géométries voient le jour, bien que les géomètres aient été troublés par la complexité du 5^e postulat d'Euclide pendant des siècles. Il y eu plusieurs tentatives pour tenter de le déduire des autres postulats, c'est-à-dire d'en faire un théorème. Ainsi, en 1733, Giovanni Saccheri, en tentant de prouver la géométrie euclidienne, jeta les bases de ce qui serait plus tard la géométrie elliptique. Près d'un siècle plus tard, au début du XIX^e siècle, Carl Friedrich Gauss (1777-1855) s'interroge sur le postulat des parallèles. Il écrit en 1813: « *Pour la théorie des parallèles, nous ne sommes pas plus avancés qu'Euclide, c'est une honte pour les mathématiques* ». Il semble que Gauss ait acquis dès 1817 la conviction qu'il est possible de développer des géométries non-euclidiennes, mais il ne publiera pas de travaux sur le sujet à cette époque. Entre 1820 et 1823, le mathématicien hongrois Janos Bolyai (1802-1860) travaille sur ce sujet. Il finalisera ce travail en 1826, et rédige un mémoire sur le sujet qui paraîtra en 1832. Il n'y démontre pas formellement l'existence d'une géométrie non-euclidienne, mais une forte présomption est acquise. Indépendamment, le russe Nicolaï Lobatchevsky (1792-1856) devance Bolyai sur la description d'une géométrie analogue en publiant ses travaux dans le journal russe *Le messenger de Kazan* en 1829 (mais le tirage ne sera disponible qu'en 1831 !), puis en langue allemande en 1840. Les travaux de Bolyai et Lobatchevsky, qui ouvrent la voie vers les géométries dites hyperboliques, n'auront cependant que peu d'échos à l'époque.

De nouveaux développements seront ensuite réalisés par le mathématicien allemand Bernhard Riemann (1826-1866), qui établit vers 1854 l'existence d'une autre famille de géométries non euclidiennes, les géométries elliptiques. Ce résultat est issu de ses travaux de thèse, effectués sous la direction de Gauss. Ici encore, l'impact à l'époque est faible, la thèse n'étant publiée que deux ans après sa mort.

Mais la situation est devenue confuse, et la possibilité de violer le cinquième postulat d'Euclide fait l'objet de controverses. Une théorie unifiée des différentes géométries sera proposée en 1872 par Felix Klein (1848-1925) lors de son discours inaugural pour sa nomination comme professeur à l'Université d'Erlangen. Les travaux sur la généralisation de la géométrie feront ensuite l'objet d'un effort continu. Mais il faut noter que la géométrie euclidienne continue de susciter l'intérêt. C'est ainsi que David Hilbert publie en 1899 ses *Fondements de la géométrie*, dans lesquels il propose de refonder la géométrie euclidienne au moyen d'une nouvelle base axiomatique (qui contient 21 axiomes) et en respectant les trois critères ci-dessous⁵⁴ :

- **Etre minimale** : aucun axiome ne doit pouvoir être retranché sans que soient alors possibles d'autres géométries que celle d'Euclide.
- **Etre complète** : une démonstration doit pouvoir exister pour montrer la véracité ou non de toute proposition.
- **Etre intrinsèque** : comme celle d'Euclide, la base axiomatique ne doit pas faire appel à d'autres notions mathématiques, comme par exemple les nombres réels.

⁵⁴ Eliakim Moore démontrera en 1902 que la liste de Hilbert n'est pas minimale : le 21^e axiome est redondant. La liste révisée est minimale. Il sera démontré plus tard par Kurt Gödel que la complétude n'est pas possible (voir à ce sujet la discussion au paragraphe 5.2.3)

Les caractéristiques marquantes des géométries non-euclidiennes sont présentées dans le tableau ci-dessous. Elles sont associées à la notion d'espace courbe, et non pas à un espace plan comme la géométrie euclidienne.

Tableau 4: principales caractéristiques des géométries euclidienne et non-euclidiennes. La géométrie euclidienne est une géométrie parabolique.

Type	Géométrie elliptique (Riemann)	Géométrie parabolique (Euclide)	Géométrie hyperbolique (Bolyai-Lobatchevsky)
Date	≈ 1830	3 ^e siècle av. JC	≈ 1830
Surface de référence	sphère	plan	pseudo-sphère
Nombre de parallèles passant par un point	0	1	infini
Somme des angles d'un triangle	> 180°	= 180°	< 180°
Longueur d'une droite	finie	infinie	infinie
Nombre de droites définies par 2 points	infini	1	1

Les surfaces de référence sont illustrées sur la figure ci-dessous.

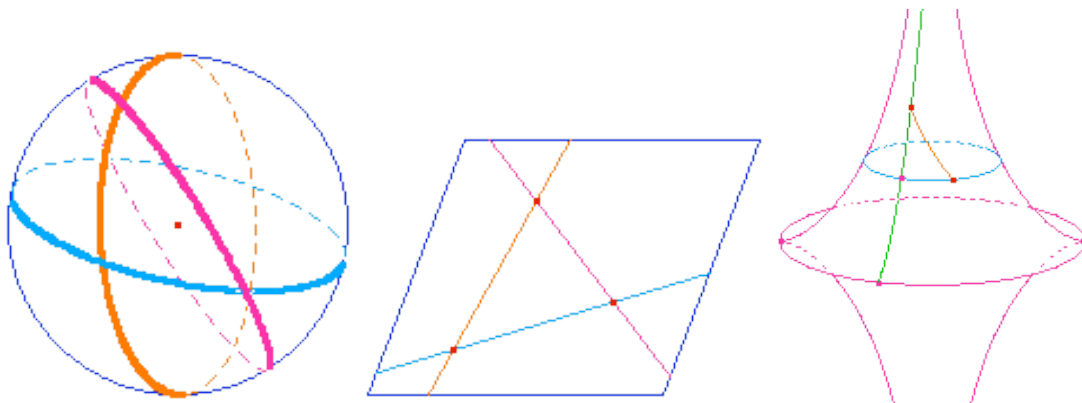


Figure 42: surfaces de référence pour la définition des géométries. Gauche: sphère pour les géométries elliptiques; Centre: plan pour les géométries paraboliques (comme la géométrie euclidienne); Droite: pseudo-sphère pour les géométries hyperboliques

Pourquoi rompre avec la géométrie euclidienne pour décrire le monde qui nous environne ? La géométrie euclidienne est intimement liée à la mécanique newtonienne, et nombre de ces développements sont inspirés par la mécanique du solide rigide. C'est ce qu'exprime le mathématicien Bela Kerékjarto lorsqu'il écrit : « *Les bases empiriques de la géométrie sont fournies par l'examen des mouvements des corps rigides.* » Durant toute la période durant laquelle la mécanique newtonienne fût le modèle dominant en physique (c'est-à-dire jusqu'à la fin du XIX^e siècle), la

géométrie euclidienne répondait aux besoins des physiciens. Le principal problème connu durant cette période était celui des géographes qui tentaient d'établir des cartes fidèles de la surface terrestre : passant d'une géométrie elliptique (la Terre réelle) à une géométrie euclidienne (le plan sur papier), il est impossible de conserver les distances et les angles⁵⁵.

L'emploi des géométries non-euclidiennes en physique prendra son essor avec la théorie de la relativité générale d'Einstein (1915), qui établit la correspondance entre un mouvement accéléré dans un champ de potentiel le long d'une trajectoire courbe dans un espace euclidien et un mouvement libre le long d'une ligne géodésique dans un espace non-euclidien. Ainsi, on associe le champ gravitationnel à une courbure de l'espace-temps (voir aussi le paragraphe 4.2.8). Près d'un objet massif, cette courbure est telle que l'univers est localement décrit par une géométrie elliptique. L'expansion de l'univers depuis le Big-Bang fait que dans les régions sans matière la géométrie de l'espace-temps correspond à un modèle hyperbolique. Enfin, l'invariance de la vitesse de la lumière dans le vide et la règle de composition de vitesse qui en découle (la **transformation de Lorentz**) requiert l'emploi d'une géométrie hyperbolique.

Pour la physique moderne, l'emploi d'un type particulier de géométrie est lié à l'échelle des phénomènes que l'on étudie. On distingue ainsi quatre niveaux⁵⁶ :

1. **échelle microscopique** ($< 10^{-18}$ mètre) : c'est le domaine des théories spéculatives d'unification et de la gravitation quantique. Ces échelles n'ont jamais été observées à ce jour (les moyens actuels ne permettent pas d'accéder à des échelles plus petites que 10^{-18} mètre). Les fluctuations quantiques de la géométrie de l'espace-temps sont estimées être de l'ordre de la longueur de Planck ($2 \cdot 10^{-35}$ mètre). Les différentes théories (cordes, supercordes, branes, théorie-M, géométrie quantique ...) font également apparaître jusqu'à plus de 10 dimensions spatiales supplémentaires. Ces dimensions resteraient invisibles à notre échelle car l'univers serait « replié » à très petite échelle.
2. **échelle locale** (entre 10^{-18} et 10^{11} mètres) : les effets de courbure de l'univers sont le plus souvent négligeables. La géométrie euclidienne est un bon modèle. C'est le domaine d'application de la mécanique, de la physique classique et de la relativité restreinte.
3. **échelle macroscopique** (entre 10^{11} et 10^{25} mètres) : les effets de courbure sont importants. Il faut donc employer une géométrie non-euclidienne qui décrive les variations locales de courbure de l'univers. Les champs d'applications sont la cosmologie et la théorie de la relativité générale.
4. **échelle globale** (l'univers entier) : c'est le domaine de la topologie cosmique. On s'intéresse non seulement à la courbure de l'univers, mais aussi à sa forme (c'est le domaine de la topologie), qui peut être très compliquée. On emploie la géométrie non-euclidienne.

⁵⁵ Cette difficulté a donné naissance à de nombreuses représentations différentes, qui sont utilisées en fonction de besoins particuliers.

⁵⁶ Voir « L'univers chiffonné », J.P. Luminet, Gallimard, coll. Folio Essais No 449, 2005

5 La valeur des sciences empiriques

Ce chapitre aborde la question de la valeur des sciences empiriques. On peut distinguer trois niveaux de réflexion.

Le premier porte sur la définition de la valeur : qu'est-ce qui fait que l'on accorde de la valeur aux sciences empiriques ? De multiples critères peuvent être envisagés : la véracité, l'efficacité prédictive, l'utilité, ... Il existe aujourd'hui un consensus sur l'idée qu'une théorie scientifique, pour être valable, doit être au moins capable de rendre compte des faits observés relevant de son domaine. La valeur accordée vient du lien intime que les théories sont supposées entretenir avec la réalité, lien garanti par la concordance avec les faits connus.

Le deuxième niveau est celui de la mise à l'épreuve des théories et des énoncés scientifiques. Comment, suivant quelle méthode et avec quels moyens pratiques, tester leur validité ?

Le troisième et dernier niveau, plus général, est celui des critères de scientificité. Qu'est-ce qu'une théorie scientifique ? Cette question fait appel aux deux précédentes, car, pour pouvoir dire d'une théorie qu'elle est scientifique, il faut avoir décidé ce qui fait la valeur d'une science et comment tester si la théorie considérée est éligible.

5.1 Préliminaire 1: qu'est-ce que la vérité ?

La notion de vérité est très valorisée dans la culture occidentale moderne et participe souvent de la valeur accordée aux sciences : la science est souvent considérée comme l'exemple type du savoir « vrai », et est donc précieuse.

Avant de discuter la question de la valeur des sciences empiriques, il convient donc de passer en revue les différentes significations du terme vérité.

Commençons par clarifier les différences qui existent entre le **vrai**, le **réel** et le **juste**. Le réel (par opposition à l'irréel) désigne les choses qui existent. Le vrai (par opposition au faux) concerne nos énoncés qui traitent des choses qui existent. Enfin, le juste (par opposition à l'injuste) qualifie nos énoncés portant sur ce qui doit être. Enfin, notons que le **bon** et le **mauvais**, qui sont des jugements de valeur morale, ne sont pas pertinents pour caractériser les énoncés scientifiques.

Mais qu'est-ce que la vérité ? On distingue trois principales conceptions de la vérité.

- La vérité comme correspondance entre un énoncé et une réalité extralinguistique (c'est-à-dire une réalité qui existe indépendamment de l'énoncé) : c'est la **vérité-correspondance**, appelée encore **vérité matérielle** ou **vérité factuelle**. Une pensée exprimée au travers d'un énoncé est vraie si le contenu de l'énoncé (objets, relations, ...) existe en réalité. Il s'agit ici de la conformité d'une affirmation avec la réalité qu'elle décrit. Cette définition fait que le

caractère vrai ou faux d'un énoncé est une propriété en soi de cet énoncé. Une expression de la scolastique médiévale la définit comme « *La vérité est l'adéquation de la chose et de l'intellect.* » Cette définition, qui est celle qui domine la pensée scientifique dans les sciences empiriques, peut être comprise de deux manières :

- « *La vérité est l'adéquation de l'intellect à la chose* » : l'intellect doit se conformer aux faits perçus qui sont le reflet de la réalité. La vérité sera donc établie par l'observation « neutre ».
- « *La vérité est l'adéquation de la chose à l'intellect* » : dans ce cas, l'esprit émet des hypothèses et interroge la réalité au moyen de l'expérimentation, qui est une manipulation de la réalité.
- La vérité comme cohérence : c'est la **vérité-cohérence**, ou **vérité formelle**. La vérité-cohérence ne fait pas référence à une réalité extra-linguistique : une théorie est vraie si elle est exempte de contradiction logique, c'est-à-dire si aucun de ses énoncés (lois, modèle, faits prédits) n'en contredit un autre. C'est la définition de la vérité employée dans le cadre des sciences formelles : mathématiques, logique.
- La vérité comme consensus : c'est la **vérité-consensus**. Dans cette dernière acceptation, une proposition est vraie si elle obtient le consensus au sein de la communauté scientifique. Un argument en sa faveur est que le consensus, si il existe, doit s'être imposé pour de bonnes raisons. Cette définition n'est toutefois pas reconnue comme acceptable par une très grande majorité des scientifiques, car la notion de consensus est très vague et difficile à vérifier (comment sonder l'ensemble d'une communauté ?).

Quelle est aujourd'hui la définition employée dans les sciences empiriques ? La vérité-correspondance est la définition dominante, mais elle n'exclut pas la vérité-cohérence, surtout pour les sciences empiriques fortement mathématisées (mécanique, physique). Ces sciences possèdent des cadres conceptuels basés sur les mathématiques (calcul différentiel, géométrie, ...) et leurs énoncés se prêtent souvent à la vérification formelle. Toutefois, cette cohérence interne est jugée nécessaire, mais pas suffisante.

5.2 Préliminaire 2 : un peu de logique

5.2.1 Les 3 axiomes d'Aristote

Vrai ou faux ? C'est la question qui vient à l'esprit lorsque nous réfléchissons à une hypothèse scientifique. Après l'emploi d'une procédure adéquate, nous nous attendons à pouvoir répondre à cette question, et conclure que l'hypothèse est soit vraie, soit fausse, et certainement pas les deux à la fois ou encore ni vraie ni fausse. Mais d'où nous vient cette idée qu'une hypothèse est soit vraie, soit fausse ? Serait-il possible qu'une proposition puisse être autre chose ? Oui, en théorie.

Notre mode de pensée quotidien est fondé sur une logique, qui est celle formalisée par Aristote il y a près de 2300 ans, et qui repose sur les trois axiomes ci-dessous :

Les 3 axiomes de la logique d'Aristote

1. **Axiome d'identité** : A est A, et cela éternellement.
2. **Axiome de non-contradiction** : Il n'est pas possible de nier et d'affirmer en même temps. Ce qui s'exprime symboliquement comme : B ne peut être à la fois A et non-A
3. **Axiome du tiers exclu** : Toute chose est ou affirmée ou niée. B est ou A ou non-A.

Ces trois axiomes forment la base de nos raisonnements, et cela depuis que le modèle conceptuel de la science occidentale s'est formé dans la Grèce antique. Sont-ils « vrais » : la question n'a pas de sens. Ce sont des axiomes, donc des choses posées a priori, à partir desquelles on fonde une logique. Existe-t-il d'autres solutions ? Oui. Les logiciens ont développé des logiques non standards, basées sur d'autres axiomes, par exemple en remplaçant l'axiome du tiers exclu par celui du tiers inclus. Mais, rassurons-nous, l'étude des sciences empiriques « simples » (mécanique newtonienne, ...) telles qu'elles existent aujourd'hui ne nécessite pas l'appel à ces autres logiques. C'est la logique aristotélicienne qui sera employée dans la suite de nos raisonnements. A l'inverse, une théorie physique radicalement différente comme la mécanique quantique, qui manipule des objets de base correspondant à des densités de probabilité et non des objets déterministes, a conduit à des analyses logiques très différentes. Ces développements ne seront pas traités ici. Certains de ses aspects sont évoqués au chapitre 11.

Notons enfin que des épistémologues contemporains comme le français Jean-Louis Le Moigne⁵⁷, pour fonder une épistémologie des sciences de la complexité, comme par exemple la biologie, proposent de réviser l'axiomatique de base d'Aristote.

5.2.2 Tarski : correspondance, vérité, langage et méta-langage

Une question centrale de ce chapitre est la correspondance de la vérité avec les faits. Selon la vérité-correspondance, un énoncé est vrai si son contenu empirique est en accord avec les faits. Se pose alors une difficulté, qui est celle des **paradoxes logiques auto-référentiels**. Un exemple courant est celui d'une feuille de papier sur laquelle on écrit au recto « la phrase écrite sur l'autre face est vraie » et au verso « la phrase écrite sur l'autre phase est fausse ». On arrive à la conclusion paradoxale (en tenant compte de l'axiome du tiers exclu) que chaque proposition est à la fois vraie et fausse. La correspondance avec les faits ne garantit pas la validité logique d'un énoncé d'observation. La vérité-correspondance ne semble pas induire la vérité-cohérence a priori. Comment garantir cela ? Le logicien polonais Alfred Tarski (1902-1983) a démontré qu'il est possible d'éviter de tels paradoxes dans un système linguistique donné. Pour cela, il convient de distinguer plusieurs niveaux logiques de langage :

⁵⁷ Voir ses ouvrages *Le constructivisme, tome 1* (éditions ESF, 1994), et *Les épistémologies constructivistes*, Que sais-je No 2969, PUF

Système d'analyse linguistique de Tarski

1. le **langage-objet**, qui est constitué des phrases du système linguistique soumis à l'analyse
2. le **méta-langage**, qui est constitué des phrases du système linguistique avec lequel on parle du langage-objet

On peut également identifier un méta-métalangage pour analyser le méta-langage, et ainsi de suite. Tarski a montré que le paradoxe n'existe que si on considère que les deux phrases écrites sur la feuille de papier sont situées au même niveau (le langage-objet). Le paradoxe disparaît si on les place à deux niveaux différents : une au niveau du langage-objet, l'autre au niveau du méta-langage. Il n'y a plus alors de boucle auto-référentielle.

Un résultat fondamental de la **théorie de la correspondance** de Tarski est que, pour parler de la vérité des énoncés de ce langage, il convient d'utiliser un méta-langage. Grâce à cette approche, Tarski a réussi à développer systématiquement une théorie de la correspondance de la vérité avec les faits pour tous les énoncés du langage-objet. Il a montré qu'il est possible d'éviter les paradoxes pour les énoncés simples du type « la neige est blanche » en donnant une signification mathématique de la **satisfaction** de ce type d'énoncé. Partant de cette théorie de la satisfaction pour le langage-objet, il a ensuite pu l'étendre au méta-langage (et aux niveaux supérieurs), obtenant ainsi une définition récursive de la vérité.

Alors que Tarski déniait à sa théorie toute implication épistémologique, l'épistémologue Karl Popper y voit un résultat très important. Il écrit en effet que Tarski a réhabilité « *la théorie de la vérité objective ou absolue comme correspondance, qui était devenue suspecte. Il a en effet revendiqué le libre usage de la notion intuitive de la vérité comme accord avec les faits.* »⁵⁸ Pour Popper, la théorie de Tarski permet d'utiliser la vérité-correspondance comme critère de vérité pour les sciences empiriques. Pour d'autres épistémologues, la conclusion de Popper est exagérément optimiste, car la démonstration de Tarski ne porte que sur des langages-objet beaucoup plus simples que les langages réels.

5.2.3 Gödel : vérité, démontrabilité et incomplétude

Toute proposition mathématique vraie est-elle démontrable dans un système axiomatique ? La réponse (négative) à cette question est la contribution majeure du logicien et mathématicien Kurt Gödel (1906-1978) aux mathématiques du XXe siècle. En inventant un moyen de surmonter le barrage entre les différents niveaux de langage de Tarski, il parvint à démontrer en 1931 son fameux **théorème d'incomplétude** qui peut être énoncé de manière non technique comme suit

Théorème d'incomplétude de Gödel (1931)

Toute théorie non contradictoire axiomatisée suffisamment puissante (c'est-à-dire en mesure d'exprimer l'arithmétique) est incomplète.

⁵⁸ K. Popper, « Conjectures et réfutations »

Cela signifie que toutes les propositions vraies de telles théories (y compris l'arithmétique) ne peuvent pas être démontrées dans un seul et unique système formel donné : outre les propositions que l'on peut démontrer être vraies ou fausses, il existe des **propositions indécidables** sur lesquelles on ne peut pas statuer. Il est important de retenir que l'**indécidabilité** n'est pas liée à la difficulté de la démonstration, mais au fait que cette démonstration est impossible avec la théorie. Ce résultat est considéré comme l'une des plus grandes avancées mathématiques jusqu'à ce jour. Sur la base des travaux de Gödel, le logicien américain Alonzo Church (1903-1995) démontra en 1936 le théorème qui porte son nom et qui donne une réponse négative au **problème de la décision** posé par David Hilbert (1862-1943) dans son fameux programme de refondation des mathématiques : « *un problème mathématique donné doit obligatoirement pouvoir admettre une solution exacte, soit sous la forme d'une réponse directe à la question posée, soit par la démonstration de son insolubilité et de l'échec inévitable qui y est lié* ». Hilbert s'était trompé : il existe des propositions vraies indécidables. Le théorème de Gödel peut également être énoncé sous les formes suivantes :

Formes alternatives du théorème d'incomplétude de Gödel

- Les mathématiques sont inépuisables : elles ne peuvent être complétées.
- Toute théorie formelle non-contradictoire des mathématiques contient obligatoirement des propositions indécidables
- Aucun système formel des mathématiques (englobant au minimum la théorie des nombres naturels) ne peut être à la fois non-contradictoire et complet.

Une conséquence de ce théorème, énoncée par Tarski en 1933, est qu'il est impossible de définir de façon objective le concept de proposition vraie en arithmétique, au sens où des propositions peuvent être vraies sans qu'on puisse le prouver.

La présence d'indécidable a également été illustrée de manière très frappante par le mathématicien Gregory Chaitin qui a écrit en 1987 une équation indécidable. Cette équation est à coefficients entiers et dépend d'un paramètre n et de ... 12 000 variables. Elle ne fait intervenir que les opérations arithmétiques et des puissances. Le problème consiste à savoir, pour chaque valeur de n , si cette équation possède un nombre fini ou infini de solutions entières. On montre que répondre à cette question échappe à tout système formel non-contradictoire.

Ces résultats montrent que la notion de vérité-cohérence est bien plus complexe que ne le laisse présager une analyse « naïve ». La théorie de Gödel n'a pas de répercussion directe sur la notion de vérité-correspondance pour les sciences empiriques, mais elle montre que la notion de vérité n'est pas forcément associée à la démontrabilité.

5.2.4 Interlude : informatique, calculabilité et théorie des nombres

Les résultats de Gödel ont également des répercussions profondes en informatique. Sur leur base, Alan Turing (1912-1954) démontra l'indécidabilité du **problème de l'arrêt d'un programme** : il n'existe pas de procédé permettant, pour tout programme, de savoir si celui-ci s'arrêtera au non.

Une autre conséquence de l'incomplétude est l'existence de **fonctions non-calculables**. Une fonction calculable est une fonction définie par un algorithme, c'est-à-dire une suite de règles qui sont exprimées comme des textes de **longueur finie** dans une langue quelconque (français, anglais, ...). Un exemple de fonction incalculable introduit par Turing est celui de la fonction qui vaut 1 si un programme s'arrête avec des données fixées ou 0 sinon.

Cette notion de calculabilité a également des conséquences importantes en théorie des nombres⁵⁹. A la classification classique des nombres (entiers naturels, entiers, rationnels, ...) illustrée sur la figure ci dessous

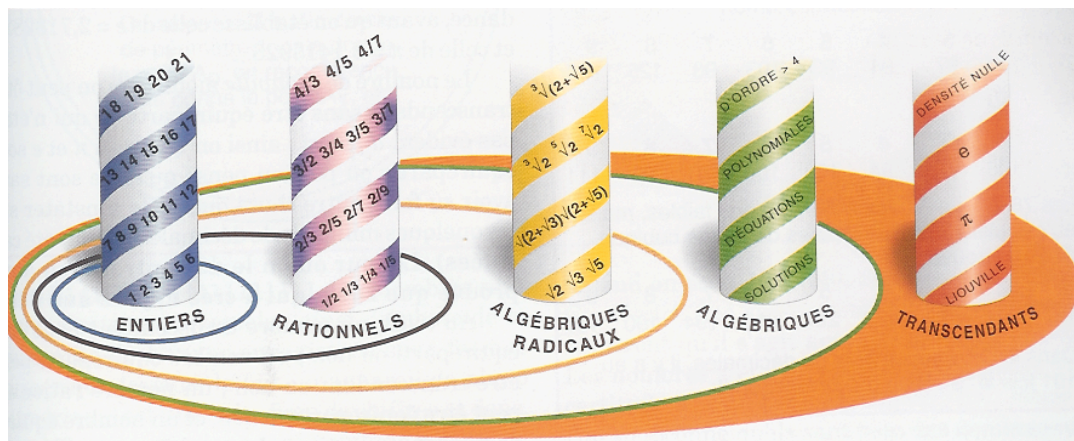


Figure 43 Classification usuelle des nombres réels ("L'intelligence et le calcul", J.P. Delahaye, Belin, 2003)

on peut substituer d'autres classifications basées sur la complexité ou la calculabilité des nombres. Ces travaux font suite à ceux du mathématicien russe Andreï Kolmogorov (1903-1987), qui est l'un des pères fondateurs de la théorie de la complexité⁶⁰.



Figure 44 A. Kolmogorov

Par exemple, on peut établir une classification en se basant sur les propriétés des décimales d'un nombre, après qu'il ait été écrit dans une base choisie. On distingue classiquement :

- Les **nombres-univers**, qui sont des nombres qui contiennent chaque séquence finie possible, quelle que soit la base choisie (binaire, ternaire, ...). Les nombres univers tiennent leur nom du fait qu'ils contiennent toutes les séquences possibles, donc (sous forme encodée) tous les livres, tous les films, toutes les musiques possibles, et le compte-rendu de toutes les existences !
- Les **nombres équirépartis**, qui sont les nombres

⁵⁹ Voir « L'intelligence et le calcul », J.P. Delahaye, Bibliothèque Pour La Science, Belin, 2003, qui a servi de base pour la rédaction de cette section.

⁶⁰ Kolmogorov a également proposé en 1941 un théorie statistique de la turbulence en mécanique des fluides qui est considérée comme l'une des plus féconde connue à ce jour !

tels que tous les chiffres ont la même fréquence d'apparition lorsqu'on les écrits (par exemple, en base 10, chaque chiffre 0,1 ... 9 doit avoir une fréquence d'apparition égale à 1/10).

- Les **nombre normaux**, qui sont les nombres équirépartis tels que tous les doublets, triplets, quadruplets, ... ont la même fréquence d'apparition (fréquence 1/100 pour chaque doublet de 00 à 99, fréquence 1/1000 pour chaque triplet de 000 à 999, etc.). Le mathématicien Emile Borel (1871-1956) a démontré que presque tous les nombres sont normaux dans toutes les bases. Mais on n'a toujours pas réussi à construire explicitement un tel nombre !

Les propriétés de quelques nombres célèbres sont présentées dans le tableau ci-dessous :

Nom	définition	transcendant	normal	Nombre-univers
π	3,141592653...	oui	?	?
e	2,718281828...	oui	?	?
Liouville	0,110001000... 1 en position n ! 0 ailleurs	oui	non	non
Champerpowne	0,12345678910... suite des entiers	oui	oui	Oui, en base 10
Pair	0,2468101214... suite des nombres pairs	oui	oui	Oui, en base 10
Erdős	0,2357111317... suite des nombres premiers	?	oui	Oui, en base 10
Copeland	0,46891012... suite des nombres composés	?	oui	Oui, en base 10

Enfin, on peut classer les nombres selon la possibilité qu'il y a de les calculer avec un ordinateur. Ici, on peut distinguer :

- Les **nombre calculables par automates finis**, c'est-à-dire par un ordinateur ayant une mémoire fixée.
- Les **nombre calculables en temps linéaires**, c'est-à-dire les nombres dont le temps de calcul des n premières décimales est directement proportionnel à n .
- Les **nombre calculables**, qui sont les nombres dont on peut calculer un nombre arbitraire de décimale, mais pas avec un coût linéaire. C'est par exemple le cas de π , pour lequel on possède de nombreux algorithmes de calcul, et dont on connaît déjà plus de 200 milliards de décimales.
- Les **nombre non calculables**, mais **finiment définissables** dans la théorie de Zermelo-Fraenkel. Un nombre finiment définissable est un nombre que l'on peut le définir avec un nombre fini de symboles dans un langage donné. Un exemple est le nombre Ω , introduit par G. Chaitin. Ce nombre, compris entre 0 et 1, est défini comme la probabilité qu'une machine de Turing à programme auto-délimité s'arrête. On a vu plus haut que ce problème est indécidable, ce qui se traduit par le fait que le nombre de Chaitin n'est pas calculable.
- Les **nombre non-finiment définissables** dans la théorie ZF.
- Les **nombre aléatoires** au sens de Martin-Löf. Un nombre est dit aléatoire dans cette théorie si ses décimales ne possèdent aucune propriété exceptionnelle que l'on puisse effectivement vérifier. Par propriété exceptionnelle, il faut entendre des propriétés comme « se terminer par une

infinité de 0 » ou « ne pas être normal » (puisque l'on démontre que presque tous les nombres sont normaux, ne pas l'être est exceptionnel). Les nombres aléatoires ne sont pas calculables, mais peuvent être finiment ou non-finiment définissables. Ainsi Ω est aléatoire et finiment définissable.

Ces classifications sont récapitulées sur la figure ci-dessous.

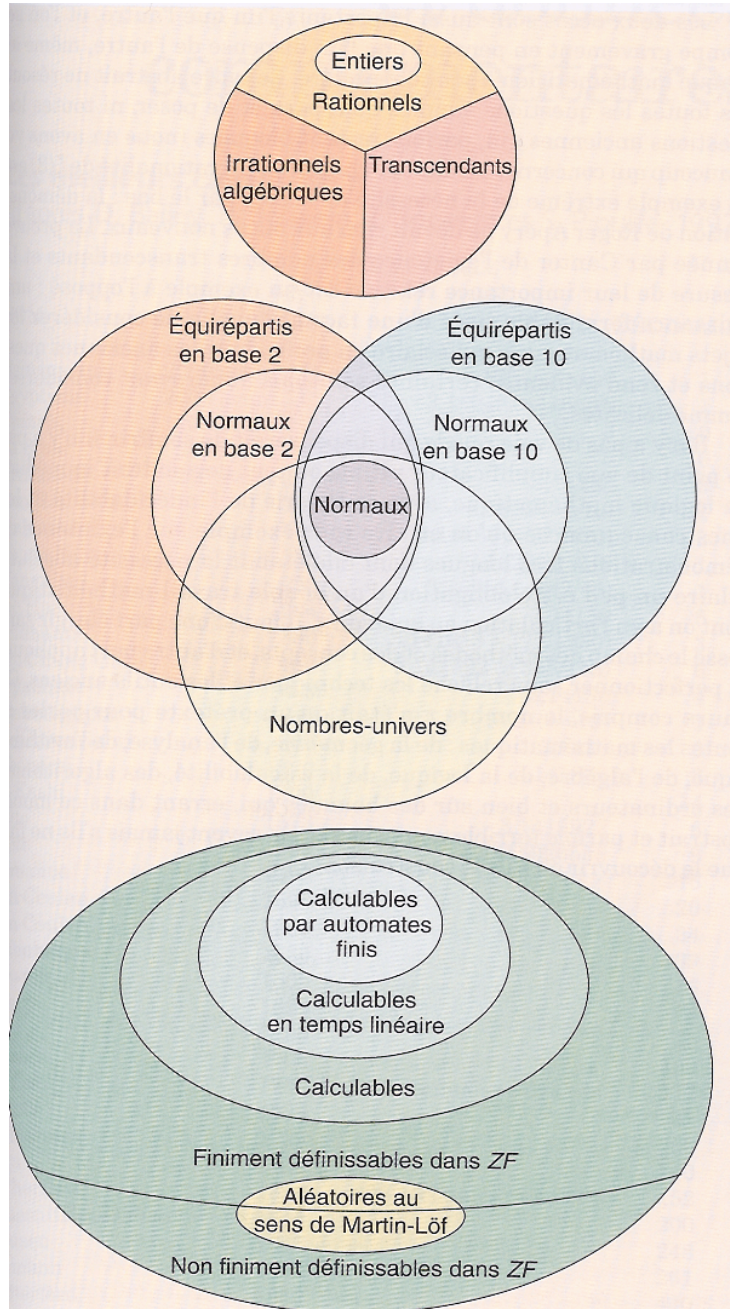


Figure 45 Récapitulatif des classifications des nombres ("Le fascinant nombre Pi", J.P. Delahaye, Bibliothèque Pour La Science, Belin, 1999)

5.3 Est-il possible de fonder la science sur les faits ?

Les sciences empiriques entretiennent un lien particulier avec les faits. En effet, ceux-ci apparaissent aux deux extrémités du processus de la réflexion : leur observation suggère des hypothèses qui sont ensuite validées par une étape de vérification basée sur un test de concordance entre faits et hypothèses. C'est ce lien qui fait qu'on accorde de la valeur aux sciences empiriques, car elles sont supposées refléter une réalité extra-linguistique et donc ne pas être de pures fictions. La question de la valeur de la science est donc subordonnée à deux questions fondamentales : quelle est la nature des faits ? Quel est le lien entre faits et théorie ?

5.3.1 Carnap et l'empirisme logique

La première position philosophique concernant la question de la fondation de la science sur les faits est celle appelée **empirisme logique**, ou encore **positivisme logique** et parfois **néopositivisme**. Cette thèse, née dans les années 1920-1930 en Europe et principalement illustrée par les travaux menés à Vienne (le cercle de Vienne⁶¹) et Berlin, a compté le logicien et épistémologue Rudolf Carnap (1891-1970) parmi ses principaux défenseurs. Pour les défenseurs de l'empirisme logique,

Critère de scientificité de l'empirisme logique

est scientifique ce qui est vérifiable (c'est ce que l'on appelle le **vérificationnisme**), et est vérifiable ce qui peut être mis en rapport avec des perceptions partagées et incontestables.



Figure 46: R. Carnap

Ce critère de scientificité, basé sur la vérité-correspondance, vise à établir une démarcation nette entre science et non-science. Les empiristes logiques vont plus loin, en soutenant que la seule connaissance qui mérite ce nom est la connaissance scientifique.

Pour comprendre les implications de cette définition, revenons sur la distinction faite entre énoncé d'observation, énoncé théorique et énoncé métaphysique. Par définition, les

énoncés d'observation sont directement vérifiables. Ils formeront donc la base de la théorie scientifique. Pour Carnap, c'est en vertu de cette concordance que les énoncés d'observation peuvent être tenus pour vrais et qu'ils acquièrent un sens, lorsqu'ils sont vérifiés. Les énoncés théoriques ne sont pas directement vérifiables : ils ne peuvent être mis à l'épreuve que par le biais des énoncés d'observation que l'on peut déduire d'eux. Un énoncé théorique ne sera vérifié que si tous les énoncés

⁶¹ Le Cercle de Vienne a été fondé en 1929 par le philosophe Moritz Schlick, sur la base de l'association Ernst Mach créée un an auparavant. Ce fut un lieu de réflexion et d'échanges intenses pour de nombreux philosophes et mathématiciens comme Ludwig Wittgenstein, Rudolph Carnap, Heinrich Gomperz, Karl Popper, Alfred Tarski, Karl Menger, Hans Hahn et Kurt Gödel. L'assassinat de Schlick le 22 juin 1936 par son étudiant et activiste nazi Hans Nelböck sonna le glas du premier cercle de Vienne.

d'observation qu'il induit sont vérifiés, sans exception. En conséquence, pour les empiristes logiques, une théorie scientifique doit pouvoir être ramenée à un ensemble d'énoncés d'observation, puisque ce sont eux qui permettent la vérification. Le contenu d'une science ne dépasse donc pas celui de l'ensemble des faits qu'elle considère pour vérifier les énoncés d'observation (d'où le terme empirisme dans empirisme logique). Pour réaliser un tel objectif, il est nécessaire, grâce à l'analyse logique du langage (d'où le terme logique dans empirisme logique), d'identifier le contenu factuel des énoncés synthétiques rencontrés dans la théorie.

L'empirisme logique est-il exempt de tout problème ? Non, car son examen soulève plusieurs problèmes.

Un des principaux problèmes est celui appelé **problème de la complétude de la base empirique**. La vérification, pour les empiristes logiques, est la pierre angulaire de la science. Or celle-ci est effectuée grâce à un ensemble de faits disponibles, qui forment ce que l'on appelle la **base empirique**, auxquels seront comparés les énoncés d'observation déduit de la théorie. Une telle vérification n'acquiert une valeur certaine qui si la base est complète, c'est-à-dire que si elle contient tous les faits possibles, sans quoi certains énoncés d'observation pourraient ne pas être vérifiés, faute de comparaison possible. Ce problème a été noté par l'épistémologue (empiriste logique) Carl Hempel (1905-1997) dans son livre *Eléments d'épistémologie* (1966) : « ... *une recherche scientifique ainsi conçue ne pourrait jamais débiter. Même sa première démarche ne pourrait être conduite à son terme, car, à la limite, il faudrait attendre la fin du monde pour constituer une collection de tous les faits ; et même tous les faits établis jusqu'à présent ne peuvent être rassemblés, car leur nombre et leur diversité sont infinis. Pouvons-nous examiner, par exemple, tous les grains de sable de tous les déserts et de toutes les plages, recenser leur formes, leur poids, leur composition chimique, leurs distances, leur température toujours changeante et leur distance au centre de la lune, qui change, elle aussi, tout le temps ? Allons-nous recenser toutes les pensées flottantes qui traversent notre esprit au cours de cette fastidieuse entreprise ? Et que dire de la forme des nuages et de la teinte changeante du ciel ? De la construction et de la marque de notre machine à écrire ? ... »*

La base empirique ne peut donc pas consigner tous les faits. Il ne faut retenir que les faits significatifs, pertinents. Mais qu'est-ce au juste qu'un **fait pertinent** ? Dire d'un fait qu'il est pertinent est un jugement de valeur, et comme tel possède une part de subjectivité. Un fait est pertinent pour un but poursuivi, et les faits pertinents pour deux buts distincts peuvent être très différents. Par exemple, pour la dynamique des gaz parfaits, ce sont les chocs entre molécules qui sont importants, alors que pour le chimiste, ce sont les propriétés chimiques des molécules du gaz qui seront intéressantes. Ensuite, le jugement de pertinence est fait par un être humain en fonction de ses connaissances antérieures et de ses croyances. Une modification de celles-ci peut conduire à revenir sur le jugement.

En conséquence, l'idée de constituer une base empirique complète doit être abandonnée. Ce qui mène à la conclusion suivante : puisque la base empirique est partielle, et qu'elle est constituée de manière subjective, elle n'est strictement ni invariante ni irrécusable.

Mais faut-il pour autant complètement rejeter l'idée centrale de l'empirisme logique ? Pas nécessairement, car, en pratique, la communauté scientifique procède le plus souvent par extension de la base empirique qu'elle emploie (de nouvelles expériences viennent s'ajouter aux anciennes). C'est principalement un processus cumulatif. De plus, le rythme d'évolution des bases empiriques retenues par chaque discipline évolue lentement.

Un second problème, plus fondamental, est abordé dans la section suivante.

5.3.2 Quine et le relativisme ontologique

Une position très différente est soutenue par les défenseurs du **relativisme ontologique**, dont Willard V.O. Quine (1908-2000) est l'un des plus illustres représentants.

Cette thèse a pour prémisse la proposition suivante : *tout langage est une théorie*. Détaillons-la. Pour Carnap et les empiristes logiques, le langage est un **langage-reflet** : le fait d'exprimer les énoncés d'observation dans une langue (français, anglais, ...) est parfaitement neutre, et n'intervient en rien dans la correspondance entre les faits énoncés et les faits « réels » tels qu'ils existent en dehors de la science.

Ceci est-il parfaitement exact ? Non, car une analyse linguistique simple montre que toutes les langues humaines n'ont pas les mêmes mots, ce qui fait que toute traduction est également une interprétation. Donc, utiliser un langage, utiliser les mots d'une langue, c'est découper et hiérarchiser les faits tels que nous les percevons : nous exprimons les énoncés d'observation (exemple : le solide léger monte lorsque la machine d'Attwood est mise en fonctionnement) en utilisant les concepts attachés aux mots de la langue que nous employons (ici, dans l'exemple : solide, léger, monter, machine, ..). Chaque langage est donc une grille de lecture, une grille de découpage conceptuel du monde. Or ce découpage n'est pas imposé par la réalité extra-linguistique, car il est variable suivant les langues. C'est ce que résume l'épistémologue autrichien Karl Popper (1902-1994) lorsqu'il dit, dans *La logique de la découverte scientifique* (1934) : « *tout énoncé a le caractère d'une théorie, d'une hypothèse* ».

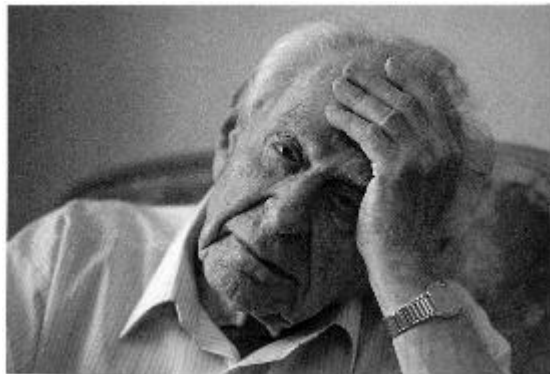


Figure 47: K. Popper

Cette absence de neutralité du langage affaiblit considérablement le programme des empiristes logiques de fonder la science sur une base empirique invariante et irrécusable : puisque le langage impose des concepts, il devient possible que les mêmes faits conduisent à des énoncés d'observation différents, peut-être même

incompatibles. Cela aurait pour conséquence de permettre que des communautés scientifiques développent des théories vérifiées mais irréconciliables.

Quine a poussé cette logique à sa limite, en affirmant que la question « qu'est-ce qui existe » n'a aucun sens absolu, et ne peut être conçue que dans le cadre d'un langage. C'est la thèse de la **relativité de l'ontologie**, qui donne son nom au **relativisme ontologique**. Selon Quine, le fait que les concepts inhérents à un langage soient efficaces en ce sens qu'il nous permettent d'agir sur le monde ne garantit en rien l'existence extra-linguistique des objets auxquels ils se réfèrent : il prouve seulement que les concepts sont bien adaptés à l'existence humaine. Une autre remarque (voir le paragraphe 4.4) va dans le même sens : les données sensibles que nous recevons sont transmises par nos organes sensoriels, qui ont des domaines de fonctionnement limités. Si nous possédions les organes sensoriels de la chauve-souris, ou la vision d'un insecte, aurions-nous identifié les mêmes objets et développé la même mécanique ? Un problème lié est celui de la fiabilité de nos perceptions brutes, qui est illustré par les illusions (voir le paragraphe 4.4).

5.4 Peut-on vérifier les énoncés scientifiques ?

Passons maintenant à la question de la vérification des énoncés. Pour simplifier le problème, considérons une communauté scientifique idéale, ne parlant qu'une seule langue et qui possède une base empirique reconnue par tous comme étant fiable. Lui est-il possible de vérifier, de manière immédiate, justifiée et irrécusable les énoncés des théories qu'elle considère ? Peut-elle déterminer la valeur de vérité et la signification des énoncés ?

5.4.1 Vérification des énoncés d'observation

Commençons par le cas qui semble le plus simple, à savoir celui des énoncés d'observation. Sont-ils, puisqu'ils expriment des propositions qui peuvent être directement confrontées à nos perceptions sensorielles, trivialement vérifiables ? Non !

Le premier argument est de nature philosophique, et vient de ce que la correspondance entre un énoncé d'observation et les perceptions n'est pas logiquement justifiable, puisque ces deux entités n'ont pas le même statut théorique. La correspondance entre un énoncé d'observation et des faits perçus est le fruit d'un **sentiment de certitude immédiate** (ce que j'énonce correspond à ce que je vois), ce qui ne suffit pas à justifier la conclusion de la vérification. En pratique, c'est le consensus entre les différents participants (vérité-consensus) qui est pris comme garantie que la correspondance est réelle. La vérification est donc subjective. Ceci est illustré par un exemple simple : Posons deux objets, un rouge et un vert, sur une table. Énonçons la proposition « les deux objets sont de la même couleur », et soumettons là à deux observateurs, dont l'un est daltonien. Ils concluront bien évidemment de manière opposée. Donc, pour reprendre Karl Popper « *aussi intense soit-il, un sentiment de conviction ne peut jamais justifier un énoncé* ». Si l'expérience ne peut justifier un énoncé, elle peut, toujours selon Popper, « *motiver une décision et par là l'acceptation ou le rejet d'un énoncé* ». Le rapport à l'expérience est en quelque sorte

affaibli, mais la décision n'est pas complètement arbitraire : elle retient quelque chose des faits perçus.

Le second argument tient de l'influence des théories sur l'acceptation ou le rejet des énoncés de la base empirique. En effet, tout énoncé peut varier dans le temps, dans sa signification ou dans sa valeur de vérité, au fil des évolutions des connaissances théoriques et techniques. Par exemple, la proposition « la longueur de cette règle est égale à un mètre » verra sa valeur de vérité changer si l'on change le mètre-étalon de manière significative. En principe, valeur de vérité et signification d'un énoncé sont toujours révisables et ne peuvent jamais être considérées comme définitives.

Que conclure de tout cela ? Que les énoncés d'observation entretiennent des liens étroits avec les faits perçus, mais qu'ils ne tirent pas leur signification et leur valeur de vérité de ce seul lien ; ils ont également un lien avec les énoncés théoriques. C'est cette dépendance qui fait que l'on dit que « les faits sont chargés de théorie ». On entend par là que l'esprit rajoute quelque chose à la perception, ce qui est logique dans la pensée occidentale qui se base sur une opposition sujet/objet (il faut noter qu'ici théorie est à prendre dans un sens très large). Par ailleurs, il n'y a que des faits énoncés pour la science : un fait qui n'est pas rapporté sous forme d'un énoncé n'existe pas, sur le plan scientifique. Et, comme nous l'avons vu, les faits énoncés retenus par la science sont des énoncés vérifiés, et la vérification implique la théorie.

Ceci implique que la conception de la base empirique, telle que l'imaginait Carnap, n'est pas soutenable : les faits énoncés qui compose la base empirique comprennent tous une part de subjectivité. Leur sélection vient d'une décision humaine, qui ne peut être justifiée strictement, mais seulement motivée. La base empirique ne peut donc pas être le fondement de la connaissance, car elle lui est relative. C'est ce que résume Karl Popper⁶² : « *La base empirique de la science objective ne comporte donc rien d'absolu. La science ne repose pas sur une base rocheuse. La structure audacieuse de ses théories s'édifie en quelque sorte sur un marécage. Elle est comme une construction bâtie sur pilotis. Les pilotis sont enfoncés dans le marécage mais pas jusqu'à la rencontre de quelque base naturelle ou donnée et, lorsque nous cessons d'essayer de les enfoncer davantage, ce n'est pas parce que nous avons atteint un terrain ferme. Nous nous arrêtons, tout simplement, parce que nous sommes convaincus qu'ils sont assez solides pour supporter l'édifice, du moins provisoirement.* »

Faut-il en déduire que la science telle que nous l'avions imaginée est impossible ? Non. Car, à une époque donnée et dans une discipline donnée, les scientifiques arrivent en pratique à se mettre d'accord sur le contenu d'une base empirique suffisante pour leurs besoins, qu'ils font évoluer dans le temps.

5.4.2 Vérification des énoncés théoriques

⁶² K. Popper, « La logique de la découverte scientifique », Payot, 1984, p.111

Passons maintenant à la vérification des énoncés théoriques. Comment, à partir des énoncés d'observation qui en découlent, peut-on les vérifier ? Là aussi, nous séparons les problèmes en admettant que les problèmes liés à l'existence de la base empirique sont résolus : nous considérons une communauté scientifique qui possède une base empirique qui ne pose pas de problème, et nous supposons que les énoncés d'observation sont vérifiables de manière satisfaisante. Dans ces conditions, est-il possible de vérifier un énoncé théorique ? Ici encore, la réponse est négative, si l'on reprend ce qui a été dit plus haut : un énoncé théorique est vérifié si tous les énoncés d'observation qui lui sont attachés le sont. Mais le problème est qu'il existe le plus souvent une infinité d'énoncés d'observation déductibles de l'énoncé théorique.

C'est notamment le cas des lois, dont on rappelle la structure logique :

Quel que soit x , si x est A , alors, x est B

Une loi a une portée universelle : elle énonce une propriété vraie pour tous les éléments qui « sont A », ce qui entraîne qu'un énoncé d'observation peut être déduit pour chaque x possible. Examinons le cas de la forme la plus simple de la loi de Boyle-Mariotte : $PV = \text{constante}$. Pour chaque gaz parfait, une observation peut être faite pour chaque valeur de la pression et pour chaque valeur du volume. Ces deux variables sont réelles, donc continues. Elles génèrent donc un ensemble infini (et indénombrable !) de possibilités. La loi de Boyle-Mariotte n'est donc pas vérifiable, au sens strict du terme. Mais comment s'assurer qu'elle n'est pas une pure fiction ? Revenons sur la démarche qui a conduit à sa proposition : elle a été formulée en se basant sur une cohérence, une stabilité observée sur un nombre fini et important d'observations (plus précisément un nombre suffisamment élevé pour emporter l'adhésion de la communauté scientifique).

Comment passer d'un nombre restreint d'exemples à une proposition universelle ? Cette démarche se nomme l'**induction amplifiante** (introduite par Francis Bacon, par opposition à l'**induction complète** d'Aristote qui est basée sur l'ensemble des faits possibles) : on suppose que ce qui a été vérifié sur tous les cas connus est vrai de tous les cas possibles. Cette démarche n'est pas logiquement valide : rien ne permet d'affirmer qu'il n'existe pas un contre-exemple (et il en suffit d'un et un seul pour invalider l'énoncé théorique) qui n'a pas encore été rencontré. De manière plus générale,

on nomme **induction** l'inférence qui permet de passer du particulier au général, par opposition à la **déduction**, qui est le passage du général au particulier.

Il existe donc une dissymétrie fondamentale entre la déduction et l'induction. L'induction est *nécessairement vraie* si ses prémisses le sont également, ce qui n'est pas le cas de l'induction amplifiante. L'induction amplifiante ne justifie pas, elle emporte la conviction. Plus le nombre d'exemples employés pour l'induction est grand, plus l'énoncé théorique nous semble plausible, moins la découverte d'un contre-exemple nous paraît probable.

La vérification de nouveaux énoncés d'observation relatifs à un énoncé théorique ne mène pas à sa vérification, mais à sa confirmation, à sa **corroboration**.

Ils « vont dans son sens », mais ne garantissent pas sa valeur de vérité. Carnap, dans la dernière phase de ses recherches, avait noté ce problème, et visait à développer une logique inductive (et non pas déductive, comme la logique classique), car il proposait d'identifier le degré de confirmation d'un énoncé théorique à la probabilité pour que celui-ci soit vrai.

La corroboration tient d'une conviction basée sur la régularité des faits observés. Mais cette régularité peut être trompeuse, comme le souligne Bertrand Russell avec l'exemple suivant : « *L'homme qui a nourri le poulet tous les jours de sa vie finit par lui tordre le cou, montrant par là qu'il eut été bien utile audit poulet d'avoir une vision plus subtile de l'uniformité de la nature.* » Le point souligné ici est la caractère psychologique et non logique de la corroboration. Comment expliquer notre croyance que les événements futurs ressembleront à ceux du passé ? Puisqu'il n'existe pas de démonstration logique, David Hume propose que notre conviction doit être basée sur deux autres principes psychologiquement aussi important : l'habitude et l'accoutumance engendrées par la régularité observées dans les expériences passées.

5.5 Valider les théories scientifiques : les méthodes

Nous venons de discuter les aspects philosophiques du problème de la vérification des énoncés. Nous allons maintenant examiner la pratique scientifique de la mise à l'épreuve des théories et des lois. Cette pratique se fonde sur l'existence de méthodes, qui sont différentes pour les sciences formelles et les sciences expérimentales. Avant de passer à la description et à la discussion, il est important de noter que c'est sur le gage de « sérieux » qu'apportent ces méthodes que se fonde le consensus autour de la valeur de la science. La science, dans l'esprit commun, est synonyme de « rigueur » ; L'emploi d'une méthode, un gage de scientificité.

5.5.1 Qu'est-ce qu'une méthode ?

Suivant le dictionnaire Larousse, une méthode est « *une démarche ordonnée, raisonnée ; une technique employée pour obtenir un résultat.* » Il vient des racines grecques **méta** (vers) et **hodos** (route). La méthode est donc conçue comme un moyen rationnel de parvenir à un but, sans commettre d'erreur (on cherche à garantir que l'objectif sera bien atteint) et avec efficacité (l'objectif sera atteint avec un effort minimal ; une activité menée au hasard demanderait beaucoup plus d'efforts pour parvenir au but).

La question de la méthode ne peut pas être discutée sans faire mention des travaux de René Descartes (1596-1650), qui a introduit la méthode cartésienne, à laquelle tout bachelier français a été formé. Selon Descartes, l'emploi d'une méthode est une obligation de l'esprit, qui sinon ne peut parvenir à des conclusions (*Discours de la méthode*, 1637) : « *Il est [...] bien préférable de ne jamais chercher la vérité sur aucune chose plutôt que de le faire sans méthode.* » Ne pas utiliser de méthode, c'est

ouvrir la porte à l'erreur : « *La diversité de nos opinions ne vient pas de ce que les uns sont plus raisonnables que les autres, mais seulement de ce que nous conduisons nos pensées par diverses voies, et nous ne considérons pas les mêmes choses.* » Ce besoin de méthode est également présent chez d'autres philosophes, comme par exemple Francis Bacon (1561-1626), qui, dans son *Novum Organum*, écrit : « *L'expérience vague et qui n'a d'autre guide qu'elle-même n'est qu'un pur tâtonnement et sert plutôt à étonner les hommes qu'à les éclairer ; mais lorsqu'elle ne marchera plus qu'à la lumière d'une méthode sûre et fixe, lorsqu'elle n'avancera que par degrés et ira pour ainsi dire pas à pas, ce sera alors véritablement qu'on pourra espérer faire d'utiles découvertes.* »



Figure 48 R. Descartes

Sa méthode a été inspirée à Descartes par ses travaux mathématiques et l'analyse de trois sciences formelles : la logique, l'algèbre et la géométrie :

« *J'avais un peu étudié, étant plus jeune, entre les parties de la philosophie, à la logique, et, entre les mathématiques, à l'analyse des géomètres et à l'algèbre, trois arts ou sciences qui semblaient devoir contribuer quelque chose à mon dessein. Mais, en les examinant, je pris garde que, pour la logique, les syllogismes et la plupart de ses autres instructions servent plutôt à expliquer à autrui les choses qu'on sait, ou même,*

comme l'art de Lulle, à parler sans jugement de celles qu'on ignore, qu'à les apprendre. Et bien qu'elle contienne, en effet, beaucoup de préceptes très vrais et très bons, il y en a toutefois tant d'autres mêlés, qui sont ou nuisibles ou superflus, qu'il est presque aussi malaisé de les en séparer que de tirer une Diane ou une Minerve hors d'un bloc de marbre qui n'est point encore ébauché. Puis, pour l'analyse des anciens et l'algèbre des modernes, outre qu'elles ne s'étendent qu'à des matières fort abstraites, et qui ne semblent d'aucun usage, la première est toujours si astreinte à la considération des figures, qu'elle ne peut exercer l'entendement sans fatiguer beaucoup l'imagination ; et on s'est tellement assujetti en la dernière à certaines règles et certains chiffres, qu'on en a fait un art confus et obscur qui embarrasse l'esprit, au lieu d'une science qui le cultive. Ce qui fut cause que je pensai qu'il fallait chercher quelque autre méthode qui, comprenant les avantages de ces trois, fût exempte de leurs défauts. Et, comme la multitude des lois fournit souvent des excuses aux vices, en sorte qu'un Etat est bien mieux réglé lorsque, n'en ayant que fort peu, elles y sont fort étroitement observées ; ainsi, au lieu de ce grand nombre de préceptes dont la logique est composée, je crus que j'aurais assez des quatre suivants, pourvu que je prisse une ferme et constante résolution de ne manquer pas une seule fois à les observer.

Le premier était de ne recevoir jamais aucune chose pour vraie que je ne la connusse évidemment être telle ; c'est-à-dire d'éviter soigneusement la précipitation et la prévention ; et de ne comprendre rien de plus en mes jugements que ce qui se

présenterait si clairement et si distinctement à mon esprit que je n'eusse aucune occasion de le mettre en doute.

Le second, de diviser chacune des difficultés que j'examinerais en autant de parcelles qu'il se pourrait et qu'il serait requis pour les mieux résoudre.

Le troisième, de conduire par ordre mes pensées, en commençant par les objets les plus simples et les plus aisés à connaître, pour monter peu à peu, comme par degrés, jusques à la connaissance des plus composés ; et supposant même de l'ordre entre ceux qui ne se précèdent point naturellement les uns les autres.

Et le dernier, de faire partout des dénombrements si entiers, et des revues si générales, que je fusse assuré de ne rien omettre.

Ces longues chaînes de raison, toutes simples et faciles, dont les géomètres ont coutume de se servir pour parvenir à leurs plus difficiles démonstrations, m'avaient donné occasion de m'imaginer que toutes les choses qui peuvent tomber sous la connaissance des hommes, s'entresuivent en même façon, et que, pourvu seulement qu'on s'abstienne d'en recevoir aucune pour vraie qui ne le soit, et qu'on garde toujours l'ordre qu'il faut pour les déduire les unes des autres, il n'y en peut avoir de si éloignées auxquelles enfin on ne parvienne, ni de si cachées qu'on ne découvre. »

On peut donc synthétiser la méthode cartésienne, qui inspire l'enseignement en France, comme la combinaison de quatre règles :

Les 4 règles de Descartes

1. **Règle de l'évidence** : n'accepter que des connaissances certaines.
2. **Règle de l'analyse** : décomposer un problème en sous-problèmes plus simples
3. **Règle de l'ordre** : commencer par les problèmes les plus simples
4. **Règle des dénombrements** : recenser toutes les parties pour ne rien oublier

La méthode cartésienne, basée sur deux étapes successives, la première de **décomposition** et la seconde de **recomposition**, est bien établie et a structuré la pensée scientifique occidentale. Est-elle unique, ou peut-on envisager des épistémologies basées sur d'autres préceptes ?

Une de ses règles, la seconde, possède une place centrale : décomposer les problèmes complexes en sous-problèmes plus simples, dont la résolution séquencée fournira la solution du problème global. Cette règle est le fondement du **réductionnisme de méthode** (voir aussi le chapitre 7.4), qui consiste à réduire un problème à des sous-problèmes solvables. Cette règle, souvent très féconde, induit deux problèmes. Tout d'abord, elle n'indique pas comment déterminer les sous-problèmes, ce qu'objectait déjà Leibniz⁶³ : « *Ce deuxième précepte de Monsieur Descartes nous est de bien peu d'utilité aussi longtemps qu'il ne nous dit pas comment décomposer en parties.* » Ensuite, certaines disciplines scientifiques qui étudient des objets complexes (exemple : la biologie et les êtres vivants) qui semblent ne pas se réduire à la somme de leurs parties et dont les propriétés de se déduisent pas simplement de celles de leurs parties (une souris est plus que la somme de ses composants chimiques ou même

⁶³ Cité par J.-L. Le Moigne, dans *Les épistémologies constructivistes*, Que sais-je No 2969

de ses organes) ; on parle alors de **propriétés émergentes**. C'est pourquoi les **épistémologies constructivistes** visent à éliminer le principe réductionniste pour l'étude de l'émergence.

5.5.2 Le cas des sciences formelles : la démonstration

Avant de discuter la méthode dominante dans les sciences expérimentales, il convient de présenter celle qui est utilisée dans les sciences formelles (logique, mathématiques). Une première raison pour cela est l'antériorité, puisque la méthode de démonstration que nous allons voir plus loin a été formalisée dans la Grèce antique au milieu du V^e siècle avant notre ère, et qu'elle a, comme nous l'avons vu, largement influé sur les réflexions postérieures, comme celles de Descartes. La seconde raison est qu'aujourd'hui, de nombreuses sciences empiriques sont très fortement mathématisées (mécanique et physique sont deux exemples parfaits), dont certaines branches théoriques exhibent des théorèmes. Dans ces cas, la méthode de démonstration des sciences formelles est également employée.

Tout d'abord, qu'est-ce que démontrer ? Démontrer, c'est prouver, c'est établir la vérité d'un énoncé théorique au moyen d'une déduction rigoureuse qui montre qu'une proposition est vraie parce qu'elle est la conséquence nécessaire d'une proposition antérieurement admise. Il n'y a donc pas de rapport à l'expérience sensible, mais seulement à la logique et à des propositions non-vérifiées prises comme point de départ (les axiomes). Le mot grec dont dérive démontrer, **deicnumi**, possède deux sens distincts : montrer au sens de faire voir, présenter à la vue, et montrer au sens d'expliquer.

L'invention de la démonstration fait partie de ce qui est appelé le « miracle grec ». D'autres civilisations antérieures, en Egypte ou en Mésopotamie, connaissaient le calcul, avaient résolu certains problèmes mathématiques et possédaient des connaissances astronomiques. Mais aucune n'avait rationalisé et codifié la démarche, et surtout, ne l'avait voulue universelle. C'est en Grèce que s'est opérée l'abstraction qui a fait passer de montrer à démontrer.

Les premières démarches visant à prouver une proposition étaient basées sur la constatation visuelle, sur une figure effectivement tracée sur un support. C'est la démarche dite pythagoricienne. Ce qui emporte la conviction, c'est le sentiment de certitude qu'éprouve le mathématicien devant le dessin qu'il observe. Il s'agit donc bien ici de « montrer ».

Le passage de montrer à démontrer est déjà effectué chez Euclide, dans les *Eléments*. Euclide codifie la méthode en six étapes, qui seront toujours les mêmes et ordonnées de la même manière :

La démonstration euclidienne

1. L'énoncé ou la *proposition* : il s'agit d'énoncer la proposition à vérifier.
2. L'exposition : il s'agit d'introduire un dessin où les points sont désignés par des lettres (on emploie encore ici la visualisation – les grecs ne connaissaient pas l'algèbre !).
3. La détermination : il s'agit de répéter l'énoncé à propos de la figure.
4. La construction, ou préparation : quand c'est nécessaire, on prépare la figure avec des constructions auxiliaires.
5. La démonstration : il s'agit de déduire le résultat.
6. La conclusion : il s'agit de reformuler la proposition comme étant le résultat de la démonstration. On y ajoute « ce qu'il fallait démontrer » (CQFD) pour un théorème ou « ce qu'il fallait faire » (CQFF) pour une construction.

Ce qui est important ici, c'est que ce n'est pas la vue de la figure qui emporte l'adhésion, mais le raisonnement qui a conduit à sa construction. C'est en construisant la figure à partir de concepts universels (triangle, cercle, ...) dont le dessin n'est qu'une illustration que l'on parvient à une conclusion de portée universelle.

La formalisation de ce type de démonstration, appelée **méthode hypothético-déductive**, est illustrée par l'exemple ci-dessous :

- *Prémises* :
 - Tous les hommes sont mortels (loi générale acceptée)
 - Socrate est un homme (cas particulier)
- *Conclusion* : Socrate est mortel

Soit, de manière un peu plus abstraite :

Méthode hypothético-déductive

Prémises :

1. Quelque soit x , si x est A, il est B
2. X est A

Conclusion : X est B

Ici, on passe du général (la loi générale acceptée) au particulier (Socrate) : on parle donc de déduction. Il s'agit de la méthode de vérification des sciences formelles. Ces résultats sont universels (cela est vrai pour tout x , et pour tous les temps) et nécessaires (la conclusion dérive logiquement des prémisses – accepter les prémisses comme vrais conduit nécessairement à accepter la conclusion comme également vraie – on peut accepter les prémisses et rejeter la conclusion).

Ce passage de montrer à démontrer n'est pas neutre sur le plan épistémologique. En effet, il consiste à donner la primauté aux constructions de l'esprit sur les perceptions, c'est-à-dire se baser sur des connaissances a priori et non sur des connaissances a posteriori. C'est donc s'inscrire dans la démarche du **rationalisme**, par opposition à l'**empirisme**.

5.5.3 Le cas des sciences empiriques : la méthode expérimentale

Venons-en maintenant au cas plus complexe des sciences empiriques. La méthode qui leur est associée est la **méthode expérimentale**, qui établit le lien entre les énoncés théoriques et les faits au travers des énoncés d'observation.

La méthode expérimentale, sous sa forme moderne, a été théorisée par le physiologiste français Claude Bernard (1813-1878), dans son *Introduction à l'étude de la médecine expérimentale* (1865) :

« Le savant complet est celui qui embrasse à la fois la théorie et la pratique expérimentale. 1^o il constate un fait ; 2^o à propos de ce fait, une idée naît dans son esprit ; 3^o en vue de cette idée, il raisonne, institue une expérience, en imagine et en réalise les conditions matérielles. 4^o De cette expérience résultent de nouveaux phénomènes qu'il faut observer, et ainsi de suite. L'esprit du savant se trouve en quelque sorte toujours placé entre deux observations : l'une qui sert de point de

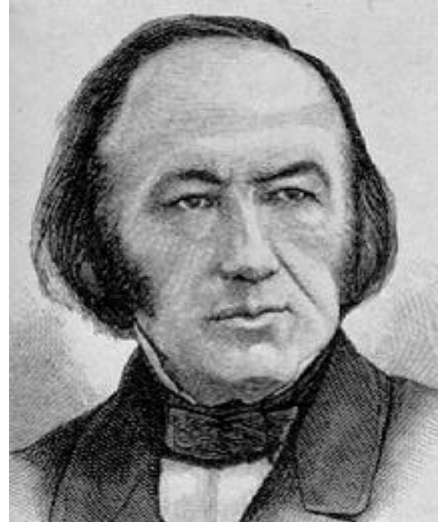


Figure 49: C. Bernard

départ au raisonnement, et l'autre qui lui sert de conclusion.

Pour être plus clair, je me suis efforcé de séparer les diverses opérations du raisonnement expérimental. Mais quand tout cela se passe à la fois dans la tête d'un savant qui se livre à l'investigation dans une science aussi confuse que l'est encore la médecine, alors il y a un enchevêtrement tel, entre ce qui résulte de l'observation et ce qui appartient à l'expérience, qu'il serait impossible et d'ailleurs inutile de vouloir analyser dans leur mélange inextricable chacun de ces termes. Il suffira de retenir en principe que l'idée a priori ou mieux l'hypothèse est le stimulus de l'expérience, et qu'on doit s'y laisser aller librement, pourvu qu'on observe les résultats de l'expérience d'une manière rigoureuse et complète. Si l'hypothèse ne se vérifie pas et disparaît, les faits qu'elle aura servi à trouver resteront néanmoins acquis comme des matériaux inébranlables de la science. »

La méthode expérimentale se décline donc en trois temps :

Les 3 phases de la méthode expérimentale

Observer – Conjecturer – Vérifier

Si elle a été codifiée par Claude Bernard au XIX^e siècle, elle est en revanche beaucoup plus ancienne. L'idée de donner une place importante, voire prépondérante aux observations dans les sciences empiriques se retrouve chez Bacon (1561-1626) et Galilée (1564-1642).

Chez Bacon, la science consiste en une collecte minutieuse des faits (qu'il appelle la « chasse de Pan »). Mais cette collecte n'est pas faite au hasard : on doit s'attacher à « d'abord découvrir les causes et les principes véritables, chercher les expériences lumineuses », appelées encore expériences « de la croix » (**expériences cruciales**, en français moderne), par référence aux poteaux indicateurs d'un carrefour. Ces

expériences cruciales, dont il propose l'idée en 1620, doivent permettre de trancher entre deux hypothèses concurrentes en réfutant de manière irrécusable l'un des deux énoncés. Donc, chez Bacon, l'expérience n'est pas exempte de recours à la raison. Un des grands apports de Bacon est l'induction amplifiante : comment, à partir d'un nombre restreint de faits, formuler un énoncé théorique de portée universelle ? Nous l'avons déjà vu, c'est l'induction amplifiante. Cela est nouveau, car avant Bacon, seule l'induction complète basée sur l'observation de tous les phénomènes possibles était acceptée, et cela depuis Aristote.

On attribue parfois à Nicolas Copernic (1473-1543) la paternité de la méthode expérimentale, en ce sens que c'est chez Copernic qu'apparaît la manipulation de l'objet par l'observateur pour obtenir la réponse recherchée, bref le passage de l'observation passive à l'expérimentation. Mais la contribution de Galilée (1564-1642) est également très importante, puisqu'il osa défendre la thèse de la primauté des faits devant les idées : ces dernières doivent se plier de manière à correspondre aux faits. C'est d'ailleurs ce qui lui a valu l'hostilité de l'Eglise, beaucoup plus que la thèse de l'héliocentrisme à laquelle elle n'était pas fondamentalement opposée. En effet, la thèse héliocentrique était rejetée par l'Eglise car elle semblait incompatible avec certains versets de la Bible, qui, dans une interprétation littérale, semblent indiquer que le soleil tourne autour de la Terre, un des plus célèbres étant

(Josué, 10, 12-13) Alors Josué parla au Seigneur [...] : « Soleil, arrête-toi sur Gabaon , lune, sur la vallée d'Ayyalôn ! » Et le soleil s'arrêta et la lune s'immobilisa [...]. Le soleil s'immobilisa au milieu des cieux et il ne se hâta pas de se coucher pendant près d'un jour entier.

Le caractère révolutionnaire de la position de Galilée était de rendre la primauté aux observations. Dans sa lettre du 21 décembre 1613 à son ami et disciple le Père Benedetto Castelli, il lance la polémique en écrivant : « ... étant évident que deux vérités ne peuvent se contredire, le devoir des interprètes sagaces est de se donner pour tâche de montrer que les véritables significations des textes sacrés s'accordent aux conclusions naturelles, aussitôt que nous ont rendus sûrs et certains le témoignage manifeste des sens ou d'irréfutables démonstrations. Je dirais plus : les Ecritures, encore qu'inspirées par l'Esprit Saint, admettant en bien des passages, pour les raisons que j'ai alléguées, des interprétations éloignées de leur sens littéral, et nous-mêmes ne pouvant affirmer en toute certitude que leurs interprètes parlent tous sous l'inspiration divine, j'estimerai prudent de ne permettre à personne d'engager les sentences de l'Ecriture et de ne les obliger en quelque sorte à garantir la vérité de telle conclusion naturelle dont il pourrait arriver que nos sens ou des démonstrations indubitables nous prouvent un jour le contraire. »

Revenons maintenant sur les trois étapes de la méthode expérimentale.

- **Observer** : au vu de ce qui a déjà été dit sur la mesure, nous savons que l'observation n'est pas neutre, qu'elle n'est pas séparable de la théorie. C'est ce que résume Auguste Comte (1798-1857) de manière catégorique, dans son *Cours de philosophie positive*: « En quelque ordre de phénomènes que ce puisse être, même envers les plus simples, aucune véritable observation n'est possible qu'autant qu'elle est primitivement dirigée et finalement interprétée par une théorie quelconque. »

- **Conjecturer** : conjecturer consiste à formuler une hypothèse, ce qui bien sûr requiert le cadre théorique adéquat. Mais l'hypothèse contient plus que le contenu factuel des énoncés d'observation à partir desquels elle est formulée, puisqu'elle procède de l'induction amplifiante. Les faits n'imposent pas l'hypothèse, ils ne peuvent que la suggérer. L'hypothèse procède donc également de l'imagination créatrice de son auteur. L'hypothèse, pour être vérifiable, doit, dit Claude Bernard, « *avoir toujours un point d'appui dans la réalité observée, c'est-à-dire dans la nature.* » Les théories et les principes guident le scientifique dans la formulation de ses hypothèses : ils sont donc des instruments de découverte pour la recherche expérimentale. Il n'existe pas de méthode automatique pour dégager les hypothèses des observations. Toutefois, le logicien et philosophe anglais John Stuart Mill (1806-1873) a codifié les méthodes couramment employées :



Figure 50: J.S. Mill

Les méthodes inductives selon John Stuart Mill

- **La méthode des résidus.** Pour traiter les phénomènes complexes, cette méthode consiste à retrancher d'un phénomène complexe et des paramètres qui agissent sur lui tous les éléments expliqués par des lois déjà connues ainsi que les paramètres qu'elles font intervenir. On peut ensuite conjecturer que les phénomènes inexpliqués restants sont associés aux paramètres déterminants restants.
 - **La méthode des concordances.** Quand deux phénomènes apparaissent toujours en même temps, on peut supposer qu'ils sont reliés, et donc qu'une loi les associant peut être formulée.
 - **La méthode des différences.** Idem, quand deux phénomènes disparaissent toujours en même temps.
 - **La méthode des variations concomitantes.** Idem, quand deux phénomènes varient toujours en même temps (exemple : la pression et le volume dans la loi de Mariotte à température constante).
- **Vérifier.** La vérification passe par l'expérimentation, c'est-à-dire une expérience contrôlée, dans laquelle les différents termes de l'hypothèse peuvent être isolés séparément, de manière à obtenir une base empirique adéquate. Le jugement de validité procède alors des questions traitées précédemment. Apparaît ici un nouveau problème concernant l'objectivité de la vérification : le **problème de l'arrêt du test expérimental**. Quand arrête-t-on une expérience destinée à valider une théorie, et donc de trancher entre plusieurs possibles ? Dans un cas idéal, c'est la nature même de l'expérience qui décide, et la question de savoir pourquoi on met fin à celle-ci n'a pas de sens. Mais l'analyse d'expériences complexes (par exemples en physique des particules) montre que la réalité est toute différente. L'arrêt de l'expérience procède d'un choix, basé sur les données acquises, mais aussi sur des jugements de valeur : l'expérience est arrêtée lorsque les résultats sont jugés

suffisamment convaincants pour permettre de corroborer de manière significative une des hypothèses que l'on cherche à tester (ou de les infirmer toutes).

5.5.4 Retour sur la théorie de l'induction

La méthode de l'induction a été évoquée à plusieurs reprises dans ce qui précède, et il a été montré qu'elle est au cœur de la méthode expérimentale qui fonde les sciences empiriques. Aussi, vu son importance, nous allons revenir sur la théorie de l'induction.

Comme il a déjà été dit au chapitre 5.4.2, l'induction totale, qui requiert que tous les énoncés d'observation possibles soient disponibles, n'est pas utilisable en pratique, et on a recourt à l'induction amplifiante introduite par Bacon. Pour pouvoir utiliser cette dernière, trois conditions doivent être satisfaites :

1. Le nombre d'énoncés d'observation disponibles doit être élevé.
2. Les observations doivent être répétables et robustes (c'est-à-dire être répétables dans de nombreuses conditions d'observations).
3. Aucun énoncé d'observation accepté comme valide ne doit contredire l'énoncé à caractère universel que l'on a déduit.

L'induction amplifiante peut être formalisée sous la forme du principe suivant⁶⁴ :

Principe de l'induction amplifiante

Si un grand nombre de A ont été observés dans des circonstances très variées, et si l'on observe que tous les A sans exceptions possède la propriété B, alors tous les A possèdent la propriété B.

Ce principe est-il valide ? Comme il a déjà été dit, l'incertitude qui pèse sur la validation des énoncés d'observation interdit toute réponse positive définitive. Une réponse positive basée sur le raisonnement suivant

1. le principe d'induction s'est révélé être juste dans le cas 1
2. le principe d'induction s'est révélé être juste dans le cas 2
3. ...
4. Donc le principe d'induction est valable dans tous les cas

n'est pas logiquement admissible, car ce raisonnement est circulaire: on valide le principe d'induction en faisant appel au principe d'induction (on parle alors de méta-induction) ! Il apparaît donc impossible de valider de manière certaine ce principe : ce problème est connu sous le nom de **problème de l'induction**. De plus, il faut noter que les notions de « grand nombre » et de « circonstances très variées » sont floues et subjectives. Elles doivent être comprises comme étant des conditions suffisantes pour emporter le consensus au sein de la communauté scientifique.

⁶⁴ On reprend ici la forme du principe d'induction formulée par A. Chalmers, « Qu'est-ce que la science », Coll. Biblio Essais, Livre de poche, p.27

Une première tentative pour résoudre le problème de l'induction consiste à affaiblir la portée de l'énoncé universel en lui donnant un caractère probabiliste :

Principe de l'induction amplifiante probabiliste

Si un grand nombre de A ont été observés dans des circonstances très variées, et si l'on observe que tous les A sans exceptions possède la propriété B, alors tous les A possèdent *probablement* la propriété B.

Un examen approfondi montre que cette solution n'est pas pleinement satisfaisante, car elle ne lève pas l'incertitude qui porte sur les énoncés d'observation, et elle nécessite d'être capable d'assigner un degré de probabilité à un énoncé, ce qui n'est pas faisable dans le cas le plus général. Une autre possibilité consiste à ne pas tenter d'assigner une probabilité à une loi ou une théorie scientifique, mais aux prédictions individuelles (énoncés singuliers) que l'on peut déduire de celles-ci. Ceci est possible, mais ne résout pas complètement le problème, car on peut déduire une infinité de prédictions individuelles à partir d'un énoncé universel, et que l'on ne peut pas réduire une loi à un nombre fini d'observations.

La théorie de l'induction est liée à celle de la vérification. Aussi le **paradoxe de la confirmation de Hempel** est-il également un problème pour la théorie de l'induction. Ce paradoxe repose sur le fait qu'un énoncé universel comme « tous les A sont B » est logiquement strictement équivalent à sa contraposée « Tous les non-B sont non-A ». Prouver la proposition initiale ou sa contraposée revient à prouver les deux. Ce qui conduit à des exemples paradoxaux, comme celui des corbeaux de Hempel. En partant des deux propositions :

1. Tous les corbeaux sont noirs
2. Tous les objets qui ne sont pas noirs ne sont pas des corbeaux

on observe que l'on valide l'hypothèse que tous les corbeaux sont noirs à chaque fois que l'on observe un objet qui n'est pas noir (et qui n'est pas un corbeau). Un ornithologue pourrait donc se contenter d'observer des fleurs et des poissons pour étudier les corbeaux ! Résultat qui est à l'évidence absurde.

Le problème de l'induction se pose également dans d'autres domaines, comme par exemple la théorie du langage. Ainsi, le philosophe Saul Kripke, sur la base des écrits de Ludwig Wittgenstein, défend la thèse selon laquelle l'usage des mots dans le langage commun pose le problème de l'induction. La question est de savoir comment utiliser un mot dans un nouveau contexte. Si l'on fait abstraction des mots dont le sens est explicitement et entièrement défini à l'aide d'autres mots (dans un dictionnaire, par exemple), c'est l'usage fait antérieurement d'un mot qui détermine son sens et nous guide. Pour Kripke, connaître le sens d'un mot revient à savoir l'utiliser selon une règle ; or il montre que les usages antérieurs ne permettent pas de déterminer de manière univoque cette règle d'usage. On retrouve donc bien le problème de l'induction définit auparavant. La solution apportée par Kripke à ce problème est de dire qu'il ne faut pas chercher à trouver une justification externe à l'induction. Selon lui, les raisonnements inductifs tout comme l'usage des mots font partie de ce que Wittgenstein appelle une « *forme de vie* », c'est-à-dire une ensemble cohérent de pratiques. En conséquence, c'est à l'intérieur de cet ensemble de pratiques que

l'induction à un sens, et il faut donc chercher des justifications internes. Cette position est attaquant, puisqu'elle relève du cohérentisme (voir au chapitre 1.4.4).

Le problème de l'induction reste irrésolu. C'est pourquoi des épistémologues ont tenté de refonder la science sans faire appel au principe de l'induction. La théorie du falsificationisme fondée par Karl Popper en est un exemple majeur.

5.6 Le falsificationisme : Popper

5.6.1 La théorie falsificationniste

Intéressons-nous maintenant au troisième niveau de réflexion évoqué au début de ce chapitre. Comment distinguer une science d'une discipline non-scientifique ? Comme nous venons de le voir, les énoncés théoriques de portée universelle (comme les lois) ne sont pas strictement vérifiables, même en ayant recours à l'expérience. Faut-il donc en conclure qu'il n'existe pas de solution, et que la distinction entre la science et le reste des connaissances humaines ne peut être que purement subjective ? Non, répond l'épistémologue Karl Popper (1902-1994) qui, en 1932, propose une nouvelle théorie : le **falsificationisme**.

Cette thèse repose sur le fait que si la confrontation à l'expérience ne peut jamais permettre de répondre par « oui » à la question de la valeur de vérité pour un énoncé, elle peut en revanche répondre « non ». Pour répondre par oui, il faudrait vérifier une infinité d'énoncés d'observation. Pour répondre par non, il suffit de trouver un unique contre-exemple. Il y a donc une dissymétrie fondamentale entre les deux réponses possibles. Donc, si on ne peut pas vérifier intégralement, on peut en revanche aisément prouver la fausseté d'un énoncé. Abandonnant l'idée défendue par les empiristes logiques que c'est la vérification ou la confirmation partielle qui doit être le critère de scientificité, Popper avance que

Critère de scientificité falsificationniste

une théorie empirique est scientifique si elle est **falsifiable**, c'est-à-dire si elle est susceptible d'être infirmée par l'expérience.

C'est la thèse centrale du falsificationisme. En conséquence, ne peuvent être inclus dans une théorie, pour qu'elle soit scientifique, que des *énoncés réfutables*. Ceci semble aisé dans le cas des énoncés d'observation. Pour les énoncés théoriques, le cas est un peu plus compliqué. Pour qu'un énoncé théorique soit réfutable (on dit encore falsifiable), il faut qu'on puisse lui associer, par déduction, au moins un énoncé d'observation falsifiable. Les énoncés métaphysiques, puisqu'ils ne sont par définition associables à aucun énoncé d'observation, sont exclus du champ de la science. Le falsificationisme ne rompt donc pas le lien avec l'expérience, il modifie seulement le critère de scientificité proposé par l'empirisme logique.

Du point de vue falsificationniste, une théorie scientifique est d'autant plus digne d'intérêt que le nombre d'énoncés d'observation qui permettent de la tester est grand. Ces énoncés d'observation falsifiables déduits de la théorie sont appelés ses **falsificateurs virtuels**. Donc, les théories scientifiques les plus intéressantes sont

celles qui sont hautement falsifiables, c'est-à-dire celles qui ont le plus haut degré de falsifiabilité. Une conséquence est que l'on doit s'attacher à ne considérer que les énoncés les plus précis et les plus clairs, car ce sont eux qui permettent les tests les plus sûrs.

Dans cette approche, on ne peut donc pas dire avec certitude ce qui est vrai. Mais on peut dire ce qui est faux : est faux tout énoncé falsifiable qui est falsifié, c'est-à-dire qui a été réfuté par l'expérience. Mais quel est le statut d'un énoncé falsifiable mais non falsifié (c'est-à-dire un énoncé pour lequel on n'a pas découvert de contre exemple jusqu'à ce jour) ? Il n'est pas vrai, puisqu'on ne pourra jamais conclure cela. On dit en revanche qu'il est **corroboré** par les faits acquis : ceux-ci ne le contredisent pas, ce qui intuitivement lui donne une certaine valeur de vérité. Mais attention : alors que vrai et faux sont des attributs absolus, corroboré n'a de sens que relativement à la base empirique utilisée pour effectuer le test de falsification. Une théorie n'est donc jamais corroborée dans l'absolu.

La thèse défendue par Popper propose une démarcation claire entre science et non-science. Sont notamment exclues du champ de la science les théories qui ne produisent que des énoncés trop vagues pour pouvoir être réfutés de manière indiscutable (Popper cite l'astrologie), ou celles qui expliquent tout (Popper cite la psychanalyse adlérienne). Pour pouvoir être scientifique, une théorie doit être falsifiable, ce qui implique qu'elle ne doit pas par construction être compatible avec tous les faits dans l'univers !

Le falsificationisme est-il une épistémologie descriptive ou normative ? Les avis sont partagés. Popper défendait un caractère descriptif. D'autres épistémologues, comme Imre Lakatos, soutiennent au contraire qu'il s'agit d'une épistémologie normative, car l'histoire des sciences ne permet pas de dégager des exemples « purs » de fonctionnement falsificationniste d'une science. Il arrive souvent que des énoncés théoriques soient conservés, bien qu'ils aient été falsifiés. En pratique, on ne dénie pas toute valeur à une théorie si l'on découvre un fait avec lequel elle est incompatible, car la corroboration par de nombreux autres faits semble indiquer qu'elle n'est pas une pure fiction.

Il existe donc des variantes plus élaborées, plus souples, de la thèse originale de Popper. Ces variantes « sophistiquées » prennent en compte la valeur de certaines corroborations, en distinguant les **corroborations peu significatives**, dont le contenu empirique ne se distingue que très peu d'autres corroborations antérieures, des **corroborations décisives**, qui amènent un contenu réellement novateur.

Avant de conclure, il convient de discuter les limites du falsificationisme. Un premier problème est que les tests auxquels sont soumis les énoncés d'observation pour tenter de les falsifier reposent sur la base empirique disponible. Le manque de complétude de celle-ci ainsi que l'incertitude qui pèse sur les observations interdisent donc de considérer une falsification comme définitive et absolument sûre. Les falsifications sont **faillibles**. En conséquence, considérer un énoncé comme falsifié est le fruit d'une décision, et possède une part de subjectivité. Popper donnait une grande importance à la décision consciente des scientifiques : « *notre acceptation des énoncés de base résulte d'une décision ou d'un accord, et à cet égard ces énoncés sont des conventions.* »⁶⁵ Ensuite, la falsification d'un énoncé est relative à une base

⁶⁵ cité par A. Chalmers, « Qu'est-ce que la science ? », Coll. Biblio essais, Livre de poche, p.110

empirique, et hérite donc du caractère incertain de celle-ci. Sur ce plan, la falsification souffre donc des mêmes problèmes que la confirmation de l'empirisme logique.

5.6.2 Le concept de vérisimilarité de Popper

Popper interprète la science comme la recherche de la vérité, et le progrès scientifique comme un mouvement de rapprochement vers la vérité (au sens de la vérité-correspondance). Mais, les théories étant imparfaites, elles ne peuvent au mieux être interprétées que comme des approximations de la vérité. Un apport important de Popper à la théorie de la vérité comme correspondance entre la science et la réalité est son travail sur **l'approximation vers la vérité**, qu'il nomme la **vérissimilarité**. Pour juger de la vérissimilarité d'une théorie scientifique, Popper propose de distinguer son **contenu de vérité**, qui désigne l'ensemble de ses conséquences vraies, et son **contenu de fausseté**, qui représente l'ensemble de ses conséquences fausses. Ces deux notions ne permettent pas de juger la qualité d'une théorie dans l'absolu : elles permettent seulement d'opérer un jugement relatif, par comparaison avec une autre théorie : « *Si l'on pose que les contenus de vérité et les contenus de fausseté de deux théories t_1 et t_2 sont comparables, on pourra dire que t_2 ressemble plus étroitement à la vérité ou correspond mieux aux faits que t_1 si et seulement si :*

1. *le contenu de vérité de t_2 est supérieur à celui de t_1 , sans qu'il en soit de même de son contenu de fausseté*
2. *le contenu de fausseté de t_1 est supérieur à celui de t_2 , sans qu'il en aille de même de son contenu de vérité.* »⁶⁶

Notons, avec Alan Chalmers que, dans une perspective réaliste (qui est celle de Popper), le concept de vérissimilarité souffre d'une faiblesse. Car, si les concepts de base d'une théorie sont faux, alors toute la théorie est fautive : son contenu de vérité est vide. L'exemple de Chalmers est la comparaison de la mécanique newtonienne et de la théorie de la relativité. Dans ses deux théories, les concepts de masse sont incommensurables : la masse newtonienne est une propriété d'un corps, alors qu'il s'agit d'une relation entre un objet physique et un système de référence chez Einstein. Si l'on considère que la théorie de la relativité est la plus proche de la vérité, alors la mécanique newtonienne est intégralement fautive, puisqu'elle repose sur un concept erroné. Cette faiblesse disparaît dans une perspective instrumentaliste, pour laquelle l'efficacité prédictive est le critère dominant (et il apparaît que la mécanique newtonienne est une bonne approximation de la physique relativiste pour des vitesses faibles).

5.7 Le holisme épistémologique : Duhem, Quine

Revenons sur le critère de scientificité défendu par les empiristes logiques : est scientifique toute théorie dont les énoncés théoriques sont vérifiables. Une théorie scientifique est vraie si tous ses énoncés sont vérifiés. La démarche classique consiste donc, suivant en cela le réductionnisme cartésien, à prendre les énoncés théoriques qui constituent la théorie et à les vérifier un à un, méthodiquement.

⁶⁶ K. Popper, « Conjectures et réfutations »

Cela est-il possible ? Non, répondent les partisans du **holisme épistémologique**, pour lesquels

il n'est pas possible de vérifier un énoncé de manière isolée, car tous les énoncés sont interconnectés.

Cette interconnexion a été décrite sous de nombreux aspects dans les chapitres et les sections précédents. Différents points de vue s'expriment au sein du holisme épistémologique.

Pour le physicien et épistémologue français Pierre Duhem (1861-1916), c'est toute la théorie qui est mise à l'épreuve à chaque fois qu'un énoncé est soumis à la vérification. Tout d'abord, les énoncés théoriques sont liés les uns aux autres. Ensuite, comme nous l'avons déjà vu, les instruments de mesure qui sont utilisés par l'expérimentateur ne sont eux-mêmes pas neutres, puisque ce sont des « théories matérialisées ». On peut noter que, de plus, les instruments font souvent appel à



Figure 51: P. Duhem

des théories différentes (chimie, mécanique, optique, ...). Les résultats d'observation sont donc interprétés dans un cadre théorique qui leur donne du sens. Tenter de valider un énoncé d'observation, c'est donc éprouver non seulement toute la théorie à laquelle il appartient, mais également tout un ensemble de disciplines scientifiques. En conséquence, pour Duhem, c'est la physique tout entière qui doit être considérée comme un tout indissociable (*La théorie physique, son objet, sa structure*, 1906): « *La science physique, c'est comme un système que l'on doit prendre tout entier ; c'est un organisme dont on ne peut faire fonctionner une partie sans que les parties les plus éloignées de celle-là entrent en jeu.* »

Un point de vue plus radical est soutenu par Willard V.O. Quine (1908-2000) : pour lui, les connexions ne se limitent pas à la seule physique (ou aux sciences empiriques), mais à l'ensemble de la connaissance humaine. Pour Quine, l'ensemble de nos connaissances forme un réseau dynamique et cohérent (*Les deux dogmes de l'empirisme*, 1951) : « *La totalité de ce qu'il est convenu d'appeler notre savoir ou nos croyances, des faits les plus anecdotiques de l'histoire et de la géographie aux lois les plus profondes de la physique atomique ou même des mathématiques pures ou de la logique, est une étoffe tissée par l'homme, et dont le contact avec l'expérience ne se fait qu'aux contours. Ou encore, pour changer d'image, l'ensemble de la science est comparable à un champ de forces, dont les frontières seraient l'expérience.* » La réfutation d'un énoncé montre que le système doit être modifié, mais sans indiquer quelles parties de ce système doivent être ajustées. Les réajustements seront effectués en raison des liens logiques qui existent entre les différents énoncés.

Le holisme épistémologique a des conséquences importantes sur la possibilité de vérifier les énoncés par recours à l'expérience :

Conséquences du holisme épistémologique

La théorie est sous-déterminée par l'expérience : en cas de conflit entre deux énoncés, la logique ne dit pas quelles parties du système global doivent être modifiées. Il y a donc a priori plusieurs modifications satisfaisantes possibles.

Il peut exister des théories distinctes mais empiriquement équivalentes, c'est-à-dire des théories qui possèdent la même base empirique sans être pour autant identiques. Cela est une conséquence de la sous-détermination.

Aucun énoncé n'est à l'abri de la réfutation et tout énoncé peut toujours être sauvé de la réfutation, puisque les ajustements du système des connaissances sont globales, non locales.

Le troisième point rend caduque la notion d'expérience cruciale développée par Bacon. De façon encore plus profonde, le holisme au sens de Quine rend toute théorie scientifique en partie subjective : puisque rien n'impose les ajustements de l'ensemble des énoncés, chaque modification est le fruit d'une décision arbitraire.

Mais revenons au processus par lequel les scientifiques font évoluer les théories. Quel est-il, puisqu'il n'existe pas d'expérience cruciale qui permette de trancher en faveur d'une hypothèse au détriment d'une autre ? Est-il le fruit du hasard, d'une marche aveugle ? Non, répondent les partisans du holisme épistémologique. Si les expériences cruciales telles que les imaginait Bacon n'existent pas, il existe en revanche des **expériences psychologiquement cruciales**, qui vont motiver certaines décisions au sein d'une communauté scientifique. Les critères qui vont faire pencher la balance en faveur d'une solution peuvent être divers : pragmatiques (loi du moindre effort), esthétiques ou autres.

Puisque aucune hypothèse ne peut être réfutée dans l'absolu, le holisme épistémologique est-il compatible avec le falsificationisme ? Oui, répond Popper, si l'on s'abstient, pour éviter de réfuter un énoncé, de faire appel à des **hypothèses ad hoc**, c'est-à-dire, au sens de Popper, des hypothèses qui ne sont pas testables indépendamment de l'effet à expliquer⁶⁷. Les deux épistémologies ne sont donc pas fondamentalement incompatibles.

5.8 Efficacité et production de technologies

Comme il a été dit au tout début de ce chapitre, la valeur accordée à une discipline scientifique peut être basée sur d'autres critères que sont degré de véracité. L'**efficacité prédictive** et la production d'outils technologiques performants (c'est-à-dire la **valeur utilitaire**) sont également des critères forts qui motivent de très nombreuses recherches. La mise au point de nouvelles technologies efficaces basées

⁶⁷ Un exemple d'hypothèse ad hoc est fourni par l'histoire de l'astronomie. Galilée, à l'aide de sa lunette astronomique, observa l'existence de cratères sur la Lune. Cette observation allait à l'encontre du dogme aristotélien (alors dominant) suivant lequel les objets célestes sont absolument parfaits et donc parfaitement sphériques. Pour rétablir l'accord avec le dogme de la sphéricité de la Lune, un des détracteurs de Galilée émit l'hypothèse que la Lune était recouverte d'une substance parfaitement invisible qui comblait les cratères, redonnant ainsi à la Lune une forme parfaitement sphérique. Cette hypothèse, invérifiable par Galilée, est une hypothèse ad hoc.

sur des énoncés théoriques peut être considérée comme une autre forme de corroboration plus complexe que celle évoquée dans ce qui précède. Ce qui pourrait être énoncé comme « *c'est vrai, puisque ça marche !* ». Nous allons illustrer cette efficacité technologique par quelques exemples, qui complètent la discussion sur les appareils de mesure (chambre à bulles, microscope à effet tunnel, lentille gravitationnelle) menée au paragraphe 4.2.4.

Mais notons que certaines avancées technologiques sont en avance sur la théorie : les ingénieurs conçoivent parfois des systèmes innovants en se basant sur leur intuition et une approche de type erreur/essai pour lesquels on ne dispose pas de d'explications théoriques complètes. Bien sur, l'intuition est guidée par les connaissances scientifiques disponibles. Et l'existence de technologies performantes au moins partiellement inexpliquées est une source de motivation et d'inspiration pour les théoriciens.

5.8.1 Relativité générale et localisation par système GPS⁶⁸

La localisation par satellite est aujourd'hui très courante, le système le plus connu étant le système GPS (*Global Positioning System*) dont le premier satellite fût lancé en juin 1977. Le système est simple : en effectuant une triangulation à partir de 3 des satellites GPS qui sont en orbites à 20 000 kilomètres d'altitude, on peut obtenir une localisation avec une précision de l'ordre du millimètre (puisque trois sphères n'ont que deux points communs, dont un seul représente une position possible, l'autre étant situé dans les profondeurs de la Terre ou dans l'espace). Pour mesurer la distance entre le récepteur GPS et un satellite, il a été choisi (pour des raisons de furtivité : n'oublions pas que le système GPS est à l'origine développé pour des applications militaires !) d'utiliser des signaux radios émis par les satellites à des instants connus. En mesurant le retard entre le temps d'émission et la réception, on obtient une estimation de la distance. Un point crucial est donc la synchronisation entre les horloges des satellites et celle du récepteur. Pour garantir la précision du système, chaque satellite GPS contient une horloge atomique. Mais cela ne suffit pas ! En effet, pour obtenir la meilleure précision possible, il faut tenir compte de certains effets prédits par la théorie de la relativité générale et les corriger. Ces effets sont les suivants :

- L'effet Sagnac, qui est dû à la rotation de la Terre. Deux horloges tournant l'une dans le sens de la Terre et l'autre dans le sens opposé, se décalent. Pour le système GPS, cela représente un retard de propagation des signaux de 207 nanosecondes après une rotation des satellites autour de la Terre (soit 62 mètres de trajet de la lumière).
- Le décalage gravitationnel de la lumière⁶⁹ vers le rouge ou le bleu. Cet effet représente 60 microsecondes, soit un décalage de position de 18 kilomètres, par jour.
- Les horloges atomiques embarquées dans les satellites sont aussi soumises au champ de gravité de la Terre et tournent rapidement autour de celle-ci,

⁶⁸ Voir « La localisation par satellite », T. Herring, Dossier Pour la Science No 38, 2003

⁶⁹ Cet effet ne sera observé expérimentalement qu'en 1960 (soit 45 ans après les travaux d'Einstein !) par Robert Pound et Glen Rebka, qui ont montré que la fréquence de photons émis depuis le haut de la tour Jefferson de l'Université d'Harvard augmente lors de la chute en raison de la variation du champ de gravité terrestre avec l'altitude.

induisant un effet Doppler relativiste. La somme de ces deux effets représente un décalage de 21 microsecondes (soit un décalage de position de 6,5 kilomètres) par jour.

Ainsi, comme on peut le voir, la physique théorique peut avoir des applications très concrètes dans la vie de tous les jours.

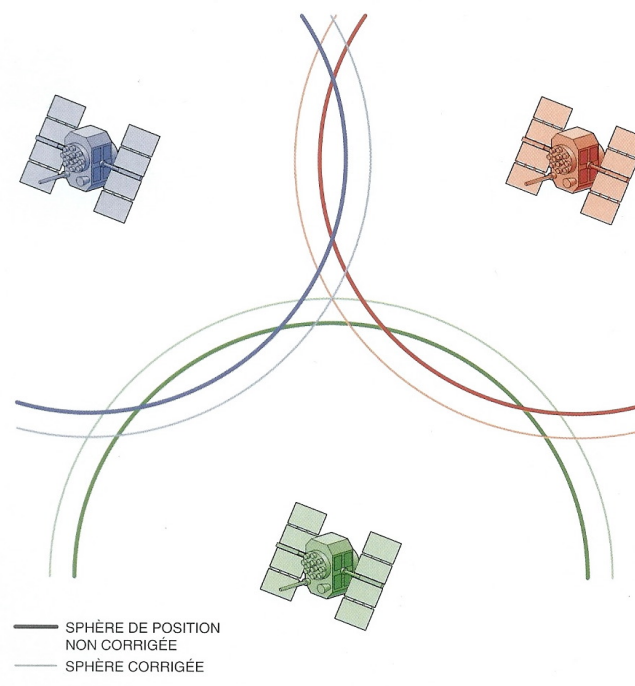


Figure 52: schéma du principe de calcul de la position par triangulation

5.8.2 Manipuler les molécules : pinces et ciseaux optiques⁷⁰

Comment manipuler des objets très petits (de la taille de quelques molécules ou d'une macromolécule) sans les détruire ? Comment sculpter la matière à très petite échelle, ou découper des segments au sein de molécules longues ? Ces questions sont aujourd'hui d'une très grande importance pour les biologistes, qui ont besoin de manipuler les cellules vivantes et leurs organites sans les détruire. La solution est fournie par l'optique : on peut construire des pinces et des ciseaux d'une grande précision à l'aide de rayons laser ! L'action d'un rayon laser dépend de l'absorption de la cible et de la radiance (mesurée en Watts par unité de surface), c'est-à-dire de l'énergie qui atteint la cible pendant une durée déterminée.

La fonction de ciseaux est obtenue au moyen de différents mécanismes : création d'ondes de choc, échauffement ou encore rupture des liaisons moléculaires. On utilise généralement des impulsions brèves de forte énergie.

La fonction de pince est obtenue de manière plus subtile au moyen d'une paire de rayons lasers symétriques. Lors de la réfraction d'un faisceau laser par une cible transparente, la quantité de mouvement de celui-ci est modifiée. La conservation

⁷⁰ Voir « Pinces et ciseaux optiques », M. Berns, Dossier Pour La Science No 53, 2006

globale de la quantité de mouvement du système (laser+cible) implique que la quantité de mouvement de la cible est modifiée, et donc que cette dernière est soumise à une force. Quand le point de focalisation des deux faisceaux est situé entre le centre de la cible et la lentille, la force résultante attire la cible vers la lentille. Par contre, si il est situé au-delà du centre de la cible, celle-ci sera repoussée. On peut également produire des déplacements latéraux. On utilise pour cela des lasers continus de faible intensité, pour obtenir des forces de l'ordre du piconewton.

La possibilité de manipuler des cellules au moyen de rayons laser a été découverte dans les années 1970 par Arthur Ashkin, des Laboratoires *AT&T Bell*, puis Steven Chu (prix Nobel de Physique en 1997) a montré qu'il est possible de manipuler des molécules au moyen de pinces optiques. Celles-ci ont aujourd'hui de nombreuses applications : elles permettent aujourd'hui de mesurer la position d'un enzyme avec une précision de 1 angström (soit 10^{-10} mètre, c'est-à-dire la taille d'un atome !) et de manipuler des molécules avec une grande précision. Utiliser conjointement avec les ciseaux optiques, elles permettent également d'effectuer des actes de chirurgie cellulaire. Notons que l'idée que la lumière puisse engendrer un effort mécanique n'est pas récente : la notion de pression de radiation (c'est-à-dire la pression exercée par les ondes électromagnétiques à la surface d'un corps éclairé) avait été anticipée par Johannes Kepler qui, à la fin du XVI^e siècle, avait émis l'hypothèse que la queue d'une comète est constituée de grains de poussière poussés par la lumière du Soleil. Il cherchait à expliquer pourquoi, lorsqu'une comète s'éloigne du Soleil, sa queue la précède.

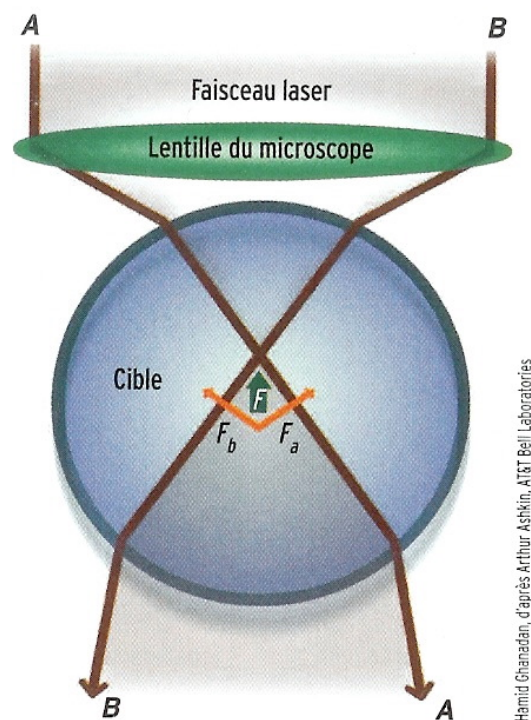


Figure 53: schéma du principe des pinces optiques

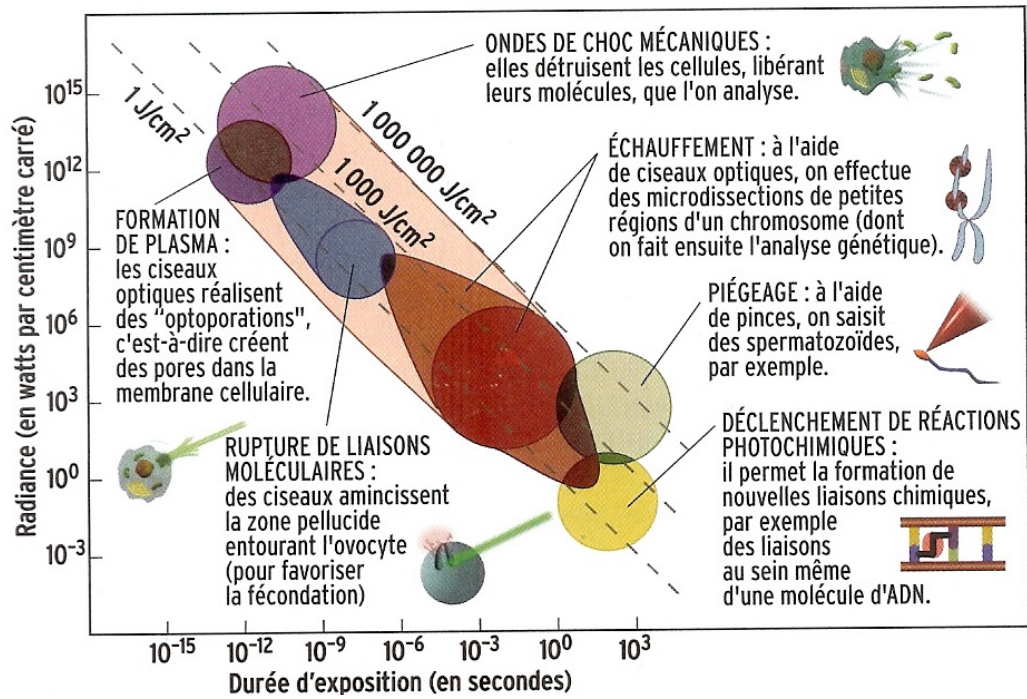


Figure 54: régimes de fonctionnement des outils optiques, et applications correspondantes

6 La science explique-t-elle quelque chose ?

Cette question peut sembler provocante. En effet, il est généralement admis que la science a une forte valeur explicative : on attend d'elle qu'elle nous renseigne sur le monde, et qu'elle nous dise pourquoi les phénomènes ont lieu.

Mais, à y regarder de plus près, nous allons voir que la valeur explicative des sciences empiriques n'est pas une question triviale, et qu'il est nécessaire d'articuler la réflexion autour de quatre notions : **décrire**, **expliquer**, **comprendre** et **interpréter**.

6.1 Décrire ou expliquer ?

6.1.1 Que veut dire expliquer ?

L'opposition décrire/expliquer semble au premier abord simple à concevoir.

Pour le sens commun, **décrire**, c'est **montrer**, c'est-à-dire rapporter les faits de la manière la plus fidèle possible au moyen d'une suite d'énoncés d'observation. Décrire consiste à indiquer à une tierce personne comment les phénomènes se déroulent.

Expliquer, c'est **dire pourquoi** les phénomènes se déroulent de la manière observée.

Illustrons cela dans le cas de la dynamique des gaz parfaits. Décrire, c'est énoncer les lois associées au phénomène observé. Dans le cas d'une réduction du volume à température constante, le phénomène observé (hausse de la pression) est décrit par la loi de Boyle-Mariotte, qui dit que pression et volume varient en proportions inverses. Mais la loi de Boyle-Mariotte ne dit pas pourquoi pression et volume varient de cette manière. Pour expliquer, il faut mener une analyse plus profonde. C'est ce que fournit le modèle de Maxwell en théorie cinétique des gaz : il explique le phénomène par l'augmentation de la fréquence des chocs entre les molécules du gaz sur les parois du contenant.

Mais cela est-il aussi simple ? Non. En effet, l'explication doit remporter la conviction de la personne à qui elle est fournie, et la valeur explicative d'une suite d'énoncés est donc en partie subjective, car de nature psychologique. La valeur explicative n'est donc pas absolue : elle dépend des croyances (au sens large) des individus. Une suite d'énoncés qui n'est pas reconnue comme étant une explication est assimilée à une description, peut être plus fine car faisant appel à des énoncés théoriques plus complexes que la description originale du phénomène. La valeur explicative est donc associée à l'appel du niveau de description le plus profond, c'est-à-dire celui qui est admis comme le niveau ultime par les individus. Le caractère relatif de la valeur explicative est donc évident, car ce niveau ultime peut varier suivant les individus, les époques et les disciplines scientifiques.

Par exemple, Newton et Descartes invoquaient Dieu comme cause première des mouvements naturels. Plus tard, le mathématicien et physicien Pierre-Simon Laplace (1749-1827), à la remarque Napoléon Bonaparte sur la place de Dieu dans son ouvrage *Exposition du système du monde*, « *M. Laplace, on me dit que vous avez*

écrit ce volumineux ouvrage sur le système de l'Univers sans faire une seule fois mention de son Créateur » répondit « Sire, je n'ai pas eu besoin de cette hypothèse », ce à quoi Napoléon répondit à son tour « Ah ! C'est une belle hypothèse ; elle explique beaucoup de choses ».

Il est convenu d'opérer une distinction entre **explication au sens fort** et **explication au sens faible**. On appelle explication au sens fort une explication qui fait intervenir des causes inobservables (dites encore métaphysiques), par opposition à l'explication au sens faible, qui ne fait intervenir que des causes observables. Cette distinction permet de rendre compte des différences d'opinion chez les scientifiques. Pour certains, il n'y a d'explication qu'au sens fort : l'explication doit faire intervenir un principe premier, métaphysique (Dieu, ...), et l'explication au sens faible n'est pas considérée comme une explication, mais comme une description (comment au lieu de pourquoi). Pour les autres, seule l'explication au sens faible est recevable, et c'est bien elle qui est aujourd'hui très majoritairement considérée comme l'explication scientifique par excellence. Cette distinction n'est pas triviale : Pierre Duhem a défendu l'idée que la notion même d'explication était étrangère au champ scientifique, car elle repose nécessairement (selon lui) sur des concepts métaphysiques qui échappent à l'expérimentation.

Pour résumer la position la plus communément admise de nos jours chez les scientifiques, on peut dire que

à l'opposition description/explication, il faut substituer une hiérarchie de niveaux de description.

6.1.2 Explication nomologico-déductive (N-D)

Il existe, dans les sciences empiriques, un type d'explication qui occupe une place particulière : c'est l'**explication** (ou la méthode) **nomologico-déductive** (en abrégé, **méthode N-D**), qui consiste à déduire les phénomènes observables des conditions particulières et des lois (**nomos** en grec). Le schéma de ce type d'explication proposé en 1948 par Carl Hempel et Paul Oppenheim est

Explication ou méthode N-D

Explanans :

Loi universelle, dite **loi de couverture** (exemple : Boyle-Mariotte)
Faits particuliers (exemple : on réduit le volume)

Explanandum : augmentation de la pression (ce qui est observé)

Ce type d'explication est dit **légaliste**, car il repose sur des lois, par opposition aux **explications non-légalistes**. En anticipant la discussion sur la notion de cause développée au chapitre 10, on peut remarquer que, dans l'explication N-D, on appelle souvent **cause** l'explanans et **effet** l'explanandum. On parle en conséquence d'**explication causale**. Les sciences empiriques qui se basent sur la méthode N-D sont dites **explicatives**. Elles sont également **prédictives**, car les lois permettent de prévoir ce qui sera observé avant que l'expérience n'ait lieu.

L'explication N-D est l'héritière d'une longue tradition scientifique qui remonte à Aristote, pour qui une connaissance scientifique doit être exprimée comme la conclusion d'un raisonnement. Elle est compatible avec la théorie aristotélicienne en plusieurs points : puisqu'elle est basée sur une loi, elle procède de connaissances universelles ; elle vérifie la contrainte de nécessaire généralité de la connaissance, puisque les conséquences de la loi de couverture sont nécessaires ; enfin, la connaissance scientifique est conçue comme la connaissance des causes. Toutefois, elle diffère de la théorie aristotélicienne en plusieurs points :

- Elle permet d'expliquer les faits singuliers, alors que pour Aristote les faits singuliers ne relèvent pas du champ de la connaissance scientifique
- Le rapport entre cause et explication scientifique n'est pas le même : alors que chez Aristote la théorie de la connaissance est fondée sur une théorie plus fondamentale de la causalité, la théorie N-D conçoit un lien plus faible, puisqu'il est seulement demandé que l'explication rende compte de la causalité.

L'explication N-D rencontre plusieurs problèmes.

- Un premier problème est le recours à une loi universelle, que l'on ne sait pas toujours distinguer d'une généralisation accidentelle (voir chapitre 3.1.2) dont la véracité n'est pas établie de manière universelle.
- Le deuxième problème vient de ce que le schéma d'explication N-D n'impose rien quant à la succession dans le temps des faits singuliers qui entre dans l'explanans (cause) et l'explanandum (effet) : c'est le **problème de la précedence temporelle**. Alors que le sens commun et la plupart des théories de la causalité imposent que la cause doit précéder l'effet, il est possible de construire des explications N-D rétroactives. Par exemple, connaissant les positions actuelles de la Terre, de la Lune et du Soleil, il est possible de retrouver les dates des éclipses qui ont déjà eu lieu, et donc de les expliquer en utilisant les lois de Kepler ou de Newton. Dans un tel schéma, on explique des événements passés par le présent ! Il est possible de construire des exemples plus compliqués en faisant intervenir le principe de moindre action (voir chapitres 8.1.3 et 12.6.1) qui fait intervenir l'état final pour « expliquer » l'évolution d'un système physique.
- Le troisième problème provient de l'identification qui est faite dans la méthode N-D entre explication et causalité. Il est possible de construire des exemples pour lesquels l'explication de l'effet par sa cause n'est pas satisfaisante. L'exemple de la taille de l'ombre d'une tour a été proposé par Bromberger en 1966. La taille de l'ombre est parfaitement explicable à partir de la position du Soleil dans le ciel et de la taille de la tour dans le cadre d'un schéma causal N-D. Mais on peut également, à l'inverse, déduire la taille de la tour à partir de la taille de l'ombre, connaissant la position du Soleil. Mais ce raisonnement n'est pas satisfaisant comme « explication » de la taille de la tour ! La théorie N-D est donc incomplète en ce sens, puisqu'elle ne distingue pas ces deux cas.
- Un quatrième problème vient de l'identification faite par Hempel et Oppenheim entre explication et prédiction. Selon le modèle N-D, le même modèle logique s'applique dans les deux cas. Mais les deux procédés sont différents : dans le cas de l'explication, le sujet à connaissance de l'explanans et de l'explanandum, qui sont deux phénomènes ayant une existence concrète.

C'est le lien entre eux qui fait l'objet de l'explication. Dans le cas de la prédiction, on discute de quelque chose qui n'existe pas, ou pas encore (il s'agit d'une proposition contra-factuelle), et il n'y a donc pas de connaissance.

6.1.3 Explication inductive-statistique (I-S)

L'explication N-D repose sur l'emploi d'une loi déterministe. Hempel a également travaillé sur des schémas explicatifs basés sur des lois statistiques, qu'il nomme **modèle inductif statistique (I-S)**.

Explication ou méthode I-S

Explanans :

Loi universelle statistique, probabilité $P(A|B)$
Faits particuliers (x est A)

Explanandum : probabilité que x soit B

Le modèle inductif statistique est d'un maniement beaucoup plus compliqué que le modèle N-D, surtout en ce qui concerne le lien que l'on peut établir entre causalité et explication. Tout d'abord, il faut retenir que l'explanandum n'est pas déductible de l'explanans au sens où il l'était avec le modèle N-D ; on peut seulement dire ici que la vérité des prémisses rend probable l'explanans à un certain degré, cette probabilité étant souvent notée $P(A|B)$. Ensuite, l'explication probabiliste (ou la causalité probabiliste) soulève de nombreux problèmes, parmi lesquels :

- Il existe des explications I-S qui conduisent à une explication et sa contradictoire à partir des mêmes prémisses. C'est pourquoi Hempel indiquait que les explications statistiques ne peuvent être considérées de manière isolée, et qu'elles doivent être examinées dans un cadre plus large qui permet de trancher. Il s'agit de **l'exigence de spécificité maximale** des prémisses par rapport à l'explanandum formulée par Hempel. Le problème non résolu ici est que la spécificité maximale peut demander une connaissance complète de l'état de l'univers, ce qui est impossible.
- La théorie de l'explication statistique est basée sur une **exigence de probabilité élevée** (la cause statistique étant associée avec une probabilité élevée de produire l'effet observé). Or le critère de probabilité élevé est subjectif, et de plus il existe des exemples où de telles probabilités ne sont pas pertinentes.
- Il existe des situations dans lesquelles prédiction et causalité sont découplées : l'explanans n'est pas assimilable à la cause, pas plus que l'explanandum à l'effet. Il faut parler donc d'**indices de causalité** plutôt que de lien causal au sens strict lorsqu'on tente de donner une explication.

6.1.4 Classification de Halbwachs

Il est possible de distinguer plusieurs types d'explication dans les sciences empiriques, en fonction des éléments invoqués. L'épistémologue contemporain

Francis Halbwachs⁷¹ propose ainsi de distinguer trois sortes d'explications dans les sciences empiriques :

Les 3 formes d'explication selon Halbwachs

- L'**explication homogène** (aussi appelée **explication formelle**), qui ne fait intervenir aucun élément extérieur au système dont on cherche à expliquer l'évolution. L'explication portera donc sur des variables internes de ce système et sur les relations, les lois, qui lient ces variables. Le plus souvent, de telles explications font appel à la notion de conservation d'une quantité (énergie, masse, ...) ou de propriétés de symétrie du système considéré.
- L'**explication causale** (encore appelée **explication hétérogène**), qui est basée sur l'interaction entre le système et le monde extérieur, ou précisément une sous-partie du monde extérieur représentée par un ou plusieurs objets. L'explication repose alors sur la notion d'échange entre les différents objets en interaction.
- L'**explication bathyène** (du grec *bathus*, « profond »), qui consiste à se référer à un autre niveau de description. Il s'agit ici de faire référence à un autre modèle du même système physique, qui le plus souvent repose sur une description à une échelle plus petite.

Illustrons maintenant cette classification.

Nous avons déjà rencontré une explication homogène, dans le cas du gaz parfait : à température constante, la variation de pression est expliquée par la constance du produit de la pression et du volume (loi de Boyle-Mariotte). Cette explication ne fait pas intervenir le monde extérieur : seules des quantités associées au gaz (pression, volume) sont invoquées.

Un exemple fameux, utilisé depuis des siècles par de nombreux philosophes et physiciens pour illustrer la notion de relation causale est celui du choc de deux boules de billard, l'une étant initialement au repos. Qu'est-ce qui cause le mouvement de la seconde ? La réponse classique est le choc avec la première boule. Lorsque l'on considère la seconde boule, l'explication de sa mise en mouvement lui est extérieure : c'est son interaction avec le monde extérieur (en l'occurrence la première boule) qui est évoquée pour expliquer le changement de ses paramètres internes (quantité de mouvement, énergie cinétique).

L'explication bathyène est parfaitement illustrée par l'exemple de la théorie cinétique des gaz. Les variations des quantités macroscopiques (pression, température, masse volumique) du gaz sont expliquées en faisant référence aux changements dans le comportement des molécules qui forment le gaz. L'explication réside ici dans le passage d'une description macroscopique (le gaz, milieu continu) à une description microscopique (le gaz, ensemble de molécules en interaction).

Ces types d'explication sont-ils fondamentalement opposés ? Non. Il faut retenir que le choix d'un type d'explication pour un phénomène observé n'est pas imposé par l'explication : il résulte d'un choix du scientifique. En modifiant par exemple sa

⁷¹ Article « causalité linéaire et causalité circulaire en physique », dans « Les théories de la causalité », PUF, 1971

définition du système dont on suit l'évolution, il est possible de changer de type d'explication. Reprenons l'exemple du choc des boules de billard. En ne considérant que l'une des boules, nous avons affaire à une explication causale. Mais que se passe-t-il si maintenant nous considérons le système formé par les deux boules ? La description du système ne fait plus intervenir d'éléments extérieurs à celui-ci, et les propriétés des deux boules (quantité de mouvement, énergie cinétique) peuvent être à tout moment décrites au moyen de lois de conservation (ces deux quantités sont des invariants). D'une explication causale nous sommes donc passés à une explication homogène !

Il est bien sûr possible de donner une explication bathyène pour le choc des boules de billard. Ce problème, qui est associé à celui du rebond élastique d'une sphère solide est un phénomène très complexe pour le physicien moderne.

En mécanique quantique, la condition de non-pénétration entre les corps rigides élastiques résulte du *principe de Pauli*, qui interdit aux atomes de se recouvrir. Mais le choc fait intervenir plusieurs forces, comme celle due à l'électrification et les forces de van der Waals.

La description du choc élastique selon le modèle de la mécanique des milieux continus fit une avancée considérable avec les travaux de Heinrich Rudolf Hertz (1857-1894) qui, en 1881 proposa une description de la déformation d'une sphère élastique appuyée sur un plan rigide avec une force arbitraire. Cette solution, connue sous le nom de *contact de Hertz*, montre que la surface de la sphère s'aplatit pour former un petit disque autour du point de contact initial. Le rayon de ce disque est proportionnel à la racine carrée du produit du rayon initial de la sphère et de l'écrasement (défini comme la variation de la distance entre le point de contact et le centre de la sphère). Les résultats de Hertz prédisent également d'autres propriétés du choc élastique. Ainsi, le choc élastique ne dissipe presque pas d'énergie (et les lois de conservation globales employées par Huygens et Leibniz pour l'énergie cinétique et la quantité de mouvement restent valables) si la vitesse relative au moment de l'impact est petite devant la vitesse du son dans le solide⁷², et la dissipation d'énergie est partiellement due à la propagation d'ondes élastiques au sein du solide. La durée du choc élastique est du même ordre de grandeur que le temps mis par une onde sonore pour traverser la sphère, multiplié par le quotient entre la vitesse d'impact et la vitesse des ondes sonores élastiques à la puissance 1/5. Par exemple, un choc élastique entre la Terre et la Lune durerait environ ... une journée⁷³.

6.1.5 Explications non-causales

Notons que toutes les explications scientifiques ne sont pas des explications causales. Voyons quelques exemples :

- En mathématiques, les explications tirées à partir de théorèmes ou d'axiomes sont complètement étrangères au schéma causal N-D
- Il existe des explications légalistes non-causales, comme celles basées sur la loi de Boyle-Mariotte. En effet, cette loi s'écrit $PV = \text{constante}$, et la pression

⁷² Rappelons que la vitesse du son dans l'acier est de l'ordre de 5000 mètres par seconde, contre 340 mètres par seconde dans l'air « aux conditions normales de température et de pression ».

⁷³ B. Audoly et Y. Pomeau, « Contact et impact de sphères », dans « La sphère sous toutes ses formes » Dossier Pour La Science No 41, 2003

et le volume sont mesurés au même instant. Il n'y a donc pas de notion de précedence temporelle lorsqu'on déduit la valeur de l'une de ces deux grandeurs à partir de la loi de Boyle-Mariotte et de la mesure de l'autre.

Comme il a été vu plus haut, le caractère explicatif d'une proposition revêt un caractère psychologique, puisqu'une explication doit être jugée satisfaisante pour être prise comme telle. Se pose donc le problème du niveau auquel le caractère satisfaisant est reconnu, un esprit infiniment critique n'étant jamais satisfait. On rencontre ici un problème analogue au **problème de la régression épistémique** déjà rencontré au chapitre 1.

Cette dimension psychologique de l'explication a conduit certains philosophes à élaborer des théories de l'explication très différentes de la théorie logique de Hempel et Oppenheim. Un exemple est **la théorie des épisodes explicatifs** de Bromberger. L'explication est analysée comme le résultat d'épisodes d'interactions localisés dans l'espace et le temps entre un *tuteur* et un *auditeur*. D'après Bromberger, le tuteur aura donné une explication à l'auditeur lors de cette interaction communicative si les conditions suivantes sont remplies :

3 critères de Bromberger

- Le tuteur est rationnel et connaît la réponse à la question de l'auditeur au moment de l'interaction
- le tuteur présente à l'auditeur les faits qu'il juge nécessaires pour que ce dernier acquiert une bonne connaissance de la réponse à sa question
- le tuteur instruit l'auditeur de la manière qu'il juge adéquate

Cette théorie, tout comme la théorie N-D, n'est pas exempte de problème. Un exemple soulevé par Bromberger lui-même, est celui des *ornières-p*. Une ornière-p est une situation dans laquelle toutes les explications que l'on peut fournir à l'auditeur sont en contradiction avec des objections décisives que celui pourrait formuler.

6.2 Expliquer ou comprendre ?

La philosophie des sciences distingue entre comprendre et expliquer. Comprendre est associé à un sentiment de familiarité, à la perception d'un lien intime avec le phénomène analysé.

Aussi expliquer est employé dans le cadre des sciences de la nature, alors que comprendre s'applique aux sciences humaines : on explique les phénomènes naturels, on comprend les faits humains. Les premiers sont associés à des causes, les seconds à des raisons qui traduisent le fait que les humains ont des intentions.

Expliquer est également souvent associé à l'explication N-D, alors que comprendre réfère souvent à des méthodes non-légalistes.

6.1 Interpréter, expliquer et comprendre

Venons-en maintenant à l'interprétation, qui est différente des notions décrites précédemment. Dire que l'on interprète quelque chose a deux implications. Tout d'abord, qu'il pré-existe quelque chose d'indépendant de l'interprétation. Ensuite, que plusieurs interprétations sont possibles : l'interprétation a donc un caractère subjectif. Toute explication est une interprétation. L'interprétation peut (cas des sciences humaines) ou non (cas des sciences de la nature) être accompagnée de la compréhension.

On dissocie les **sciences explicatives**, basées sur l'explication N-D, des **sciences interprétatives**, qui font appel à l'explication non-légaliste par des **raisons**. Cette distinction recouvre celle évoquée précédemment entre sciences de la nature et sciences humaines.

6.2 Sciences prédictives et sciences anticipatives

6.2.1 Prédiction, anticipation et scientificité

Les distinctions qui viennent d'être effectuées permettent également de différencier les sciences selon la manière qu'elles offrent pour prévoir les phénomènes. Pour une science empirique, cette capacité de prévision a une valeur très forte aux yeux des scientifiques, et elle est considérée comme un critère de scientificité. La capacité de prévoir permet la production d'énoncés d'observation réfutables, ce qui est essentiel pour la méthode expérimentale et nécessaire d'après les falsificationnistes. Elle confère aux sciences empiriques une grande valeur de vérité, car elle semble indiquer que les énoncés théoriques (théories, lois, ...) ne sont pas de pures fictions, c'est-à-dire que, d'une manière ou d'une autre, elles ont un lien avec le monde réel. Sur un plan plus pratique, c'est la possibilité de prévoir les phénomènes qui fait des sciences empiriques des outils pour agir sur le monde, et qui donc leur confère une certaine efficacité.

Comme nous l'avons déjà vu, l'explication N-D fait intervenir des lois universelles qui, grâce au calcul, permettent de prévoir avec exactitude une évolution pour un ensemble de conditions initiales fixées, et donc de formuler des énoncés d'observation. Les sciences basées sur de telles explications, c'est-à-dire les sciences explicatives, sont en conséquence appelées sciences prédictives, et l'on dit que la méthode N-D est une structure prédictive.

Cela veut-il dire pour autant que les sciences interprétatives n'ont aucune capacité à prévoir les phénomènes ? Non. Mais, l'interprétation compréhensive conduit souvent à des prévisions plus vagues, plus difficilement quantifiables que la méthode N-D. On parle alors d'anticipation plutôt que de prédiction, et de sciences anticipatives.

6.2.2 Préviation = validation ? La découverte de Neptune

Un exemple célèbre de préviation à partir de la mécanique Newtonienne est celui de la découverte de la planète Neptune (la huitième planète du système solaire) par l'astronome français Jean Joseph Urbain Le Verrier (1811-1877).



La planète Uranus (la septième du système solaire) a été découverte en 1781 par l'astronome hollandais William Herschel (1738-1822). Il est le premier à identifier Uranus comme une planète et non comme une étoile. Mais l'on s'aperçoit très tôt que la trajectoire d'Uranus ne correspond pas aux prédictions que l'on peut faire grâce aux lois de Newton et de Kepler. De plus, les anomalies sont plus grandes que les erreurs de précision des télescopes d'alors et que celles des calculs manuels. Se pose donc la question : faut-il réviser ces lois, les compléter par d'autres, ou faire l'hypothèse qu'elles sont valides et qu'il faut trouver une solution compatible avec celles-ci ?

C'est la dernière solution qui sera tout de suite adoptée. Dès 1835, des astronomes comme l'anglais George Biddel Airy (1801-1892) et le français François Arago (1786-1853) ou encore l'allemand Friedrich Bessel émettent l'hypothèse que ces irrégularités proviennent de l'existence d'une neuvième planète qui, par attraction gravitationnelle, perturberait l'orbite d'Uranus.

En 1845, Arago persuade Le Verrier de s'attaquer à ce problème, qui est révolutionnaire : au lieu de justifier a posteriori les observations, il s'agit de **prédire à l'avance** l'existence d'un objet. Après 11 mois de travail et 10 000 pages de calcul, Le Verrier présente trois mémoires à l'Académie des Sciences. Le dernier, décisif, date du 31 août 1846, et s'intitule *Sur la planète qui produit les anomalies observées dans le comportement d'Uranus. Détermination de sa masse, de son orbite et de sa position actuelle*.



figure 56. J.C. Adams

La prédiction théorique est réalisée. Il s'agit maintenant de la vérifier par l'observation. Le 18 septembre, Le Verrier écrit à Johann Galle, à l'observatoire de Berlin. Le soir même de la réception de la lettre, le 23 septembre, Galle pointe son télescope sur la position prédite par Le Verrier, et y découvre un objet jusque là inconnu : Neptune. Cette observation clôt le problème, et démontre la puissance prédictive des lois de Newton.

La découverte de Neptune révèle également que la postérité peut être cruelle. Un brillant astronome anglais de 22 ans, John Couch Adams, avait entrepris le même travail mathématique à Cambridge. Il avait prédit la position de Neptune en septembre 1845, soit presque un an avant Le Verrier (qui ignorait ses travaux). Il écrivit d'abord à un astronome de Greenwich, puis, ne recevant pas de réponse, au directeur de l'observatoire de Greenwich, George Airy. Ce dernier, peu intéressé, demanda des précisions si futiles que Adams abandonna.

6.2.3 Et lorsque cela ne marche pas ? Les trajectoires de Mercure et de la Lune

La découverte de Neptune est un argument en faveur de la validité des lois de Newton. Mais que penser lorsque la mécanique newtonienne ne parvient pas à expliquer un phénomène astronomique, au sens où les observations ne sont pas compatibles avec les énoncés théoriques et les énoncés d'observation déduits des lois de Newton ?

Un exemple similaire à celui de la découverte de Neptune est celui du déplacement du périhélie⁷⁴ de Mercure, la première planète du système solaire, par rapport à la théorie newtonienne. Pour expliquer ces irrégularités, Le Verrier et d'autres tentèrent d'utiliser la même « ruse » que dans le cas d'Uranus : faire l'hypothèse de l'existence d'une nouvelle planète (nommée Vulcain) qui perturberait l'orbite de Mercure. Les calculs furent menés, mais cette hypothétique planète ne fût jamais observée. Il fallait donc se rendre à l'évidence : ces anomalies échappaient à la mécanique newtonienne !

Un second exemple est la trajectoire de la Lune : celle-ci se trouvait parfois légèrement en avance ou en retard par rapport aux prédictions théoriques. De plus, les anomalies semblaient évoluer de manière hasardeuse et imprévisible.

Que déduire de ces deux cas où les calculs réalisés à partir des lois de Newton ne permettent pas d'établir une correspondance avec les faits observés ?

Dans le premier cas, celui de la trajectoire de Mercure, ce sont les phénomènes physiques mis en jeu qui échappent à la mécanique Newtonienne . Il a fallu attendre qu'Albert Einstein (1879-1955, prix Nobel de Physique 1921) révise la théorie de la gravitation dans le cadre de la théorie de la Relativité Générale en 1915 pour pouvoir expliquer la trajectoire de Mercure, après presque 50 années de mise en échec de la théorie newtonienne.

Le second cas est différent : la trajectoire de la Lune est bien gouvernée par la physique contenue dans les lois de Newton, mais c'est la prédiction (sur des temps de longueur arbitraire et avec une précision infinie) grâce au calcul mathématique qui est impossible ! Pour comprendre cela, il a fallu attendre 1890, année où Henri Poincaré (1854-1912) analyse la structure des équations et introduit les bases de ce qui est aujourd'hui appelé la **théorie du chaos**. La trajectoire de la Lune est chaotique, et n'est donc pas prévisible, au sens traditionnel du terme utilisé depuis Laplace (cela sera discuté plus en profondeur dans le chapitre 10).

Ces deux exemples montrent la complexité de l'interprétation de l'incompatibilité entre une théorie et des énoncés d'observation. Dans le cas de Mercure, c'est la théorie qui doit être étendue, pour rendre compte de nouveaux mécanismes. Dans le second cas, la théorie n'a pas besoin d'être étendue, mais c'est l'idée de prédiction qui doit être révisée !

⁷⁴ Conformément aux lois de Képler, Mercure suit une trajectoire elliptique. Le périhélie désigne la rotation des axes de l'ellipse dans le plan de celle-ci. Pour Mercure, elle est de 42 secondes d'angle par siècle.

7 La science nous apprend-elle quelque chose sur le monde ?

L'intitulé de ce chapitre peut sembler absurde au premier abord, s'agissant des sciences empiriques qui se basent sur des énoncés d'observation. Mais nous allons voir que le problème est plus complexe, car il s'agit de savoir si les sciences empiriques nous renseignent sur le monde tel qu'il est (la réalité extra-linguistique) et non tel que nous le percevons. Ce qui est posé, c'est la question de la relation entre les sciences empiriques et leur objet. La théorie est-elle une description exacte de la nature ? Possède-t-elle une part de subjectivité ? N'est-elle qu'une pure convention employée par commodité pour organiser nos perceptions ?

7.1 L'opposition réalisme/antiréalisme

7.1.1 Réalisme et idéalisme

La première opposition abordée ici est celle du réalisme et de l'antiréalisme.

Le **réalisme métaphysique** affirme l'existence d'une réalité indépendante de l'homme. Au contraire, selon l'**idéalisme**, le monde extérieur n'a pas d'existence autonome, et n'est qu'une représentation des sujets humains.

L'idéalisme a des racines très anciennes, par exemple chez Platon, pour qui ce sont les **Idées** qui recèlent la vérité du monde. Chez Platon, le monde des Idées est parfait, immuable et inaccessible aux sens, et le monde sensible est lui éphémère et dégradé. C'est ce qu'il illustre avec sa fameuse allégorie de la caverne. Dans sa conception extrême, défendue par le philosophe anglais George Berkeley (1685-1753), l'idéalisme conduit à la proposition que le monde matériel disparaîtrait si tous les hommes disparaissaient.



Figure 57: G. Berkeley

Pour les partisans du **réalisme scientifique**, la théorie scientifique parfaite vers laquelle tend la théorie actuelle décrit exactement la réalité extérieure à l'homme, donc nous renseigne sur le monde en soi (c'est le **réalisme essentialiste**). A l'inverse, l'**antiréalisme** (sa variante la plus répandue étant l'**instrumentalisme**) soutient que les théories ne sont que des instruments commodes pour organiser nos perceptions et agir sur le monde, et qu'elles ne nous apportent rien de plus. Le réalisme est caractérisé par la métaphore de la **théorie-reflet**, l'antiréalisme par celle de la **théorie-outil**. Un exemple célèbre de position antiréaliste est la préface du livre de Nicolas Copernic *Des Révolutions des orbés célestes* écrite par son contemporain Andreas Osiander : « *C'est le propre de l'astronome de colliger, par une observation diligente et habile, l'histoire des mouvements célestes. Puis d'en [rechercher] les causes, ou bien – puisque d'aucune manière il ne peut en assigner de vraies – d'imaginer et d'inventer des hypothèses*

quelconques, à l'aide desquels ces mouvements (aussi bien dans l'avenir que dans le passé) pourraient être exactement calculés conformément aux principes de la géométrie. Or, ces deux tâches, l'auteur les a remplies de façon excellente. Car en effet il n'est pas nécessaire que ces hypothèses soient vraies où même vraisemblables ; une seule chose suffit : qu'elles offrent des calculs conformes aux observations. »

Puisque la réalité pré-existe à la théorie qui tente de la décrire, selon le réalisme, le chercheur découvre des objets, au sens où un explorateur découvre un nouveau territoire. À l'inverse, pour l'antiréaliste, le chercheur invente ses objets.

Le réalisme scientifique implique le réalisme métaphysique. Mais l'antiréalisme scientifique n'implique pas forcément l'adhésion à l'idéalisme : le fait que la théorie scientifique ne reflète pas un monde extérieur autonome n'implique pas que ce dernier n'existe pas, mais signifie seulement qu'il est inaccessible à la science. L'antiréalisme scientifique est donc compatible avec le réalisme métaphysique et l'idéalisme.

Le réalisme scientifique pose une autre question : la théorie idéale, celle qui serait vraiment la théorie-reflet, existe-t-elle aujourd'hui ? Non, à l'évidence. C'est un idéal, une asymptote vers laquelle évoluent les sciences. Mais alors, quel est le statut des théories existantes ? Sont-elles fausses pour autant ? Non. On peut distinguer ici deux positions. Le **réalisme naïf**, qui considère que chaque nouvel énoncé vérifié est vrai, et cela pour toujours. Le progrès scientifique serait donc un processus strictement cumulatif. Cette thèse est à l'évidence en contradiction avec les leçons de l'histoire des sciences. Il faut donc lui préférer le **réalisme convergent**, qui stipule que c'est l'ensemble de la théorie qui évolue et cela de manière globale. C'est à ce niveau global qu'il faut attendre la ressemblance avec la réalité extérieure. Certains énoncés peuvent donc être réfutés au cours de l'évolution de la science.

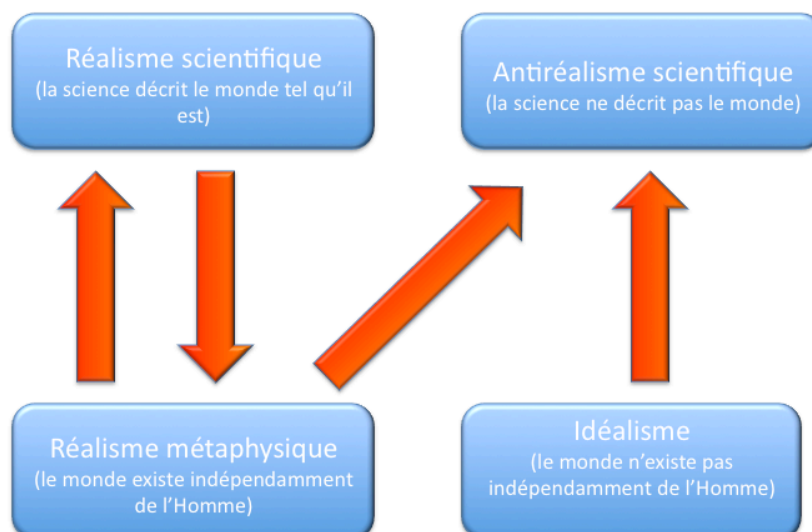


Figure 58 Schéma des relations entre réalisme et antiréalisme

7.1.2 Retour sur la question de la vérité

Ces deux positions, réalisme et antiréalisme, ont bien évidemment des répercussions sur la question de la valeur de vérité des énoncés et des théories. Le réalisme, puisqu'il suppose l'existence d'une réalité extérieure à la connaissance humaine, utilise le critère de vérité-correspondance. De manière opposée, l'antiréalisme rejette la vérité-correspondance. L'antiréalisme conduit-il pour autant à rejeter toute notion de vérité ? Non. Rejeter l'idée de la théorie-reflet n'est pas dire que le monde extérieur autonome, s'il existe, n'exerce aucune contrainte sur l'élaboration des théories scientifiques. C'est seulement dire que celles-ci ne reproduisent pas la structure intime du réel.

Affinons l'analyse en considérant le statut vrai/faux d'un énoncé. Pour un réaliste, dire qu'une proposition est vraie signifie tout d'abord que les mots (masse, force, énergie, atome, solide rigide, ...) qui composent les énoncés correspondent à des entités qui existent indépendamment dans la réalité extra-linguistique, et que le contenu de l'énoncé est vrai (la seconde loi de Newton est vraie : les phénomènes réels la présentent également). Comme la réalité est immuable, le caractère vrai ou faux d'un énoncé est absolu, et il est attaché à l'énoncé, indépendamment de l'activité humaine. On parle de **réalisme sémantique**. Pour l'**anti-réalisme sémantique**, les termes vrai et faux sont utilisables, mais ils n'ont pas de caractère absolu : ils sont relatifs aux connaissances actuelles, aux croyances, etc. Vrai et faux résultent de décisions humaines, qui peuvent évoluer dans le temps.

7.1.3 La théorie scientifique est-elle unique ?

Le réalisme scientifique impose de répondre par l'affirmative : puisqu'il existe une réalité extérieure autonome unique, et que la théorie est le reflet de celle-ci, il n'existe qu'une seule théorie vraie. Pour les antiréalistes, la réponse est négative. Puisqu'il n'y a pas de réalité autonome ou que celle-ci est inatteignable, il peut exister plusieurs théories concurrentes, qui présenteront des avantages différents, et qui pourront être employées selon les besoins. Ce point de vue était par exemple celui de Henri Poincaré, qui était un anti-réaliste scientifique (voir le chapitre 7.3). Il écrivait⁷⁵ : *« On ne doit pas se flatter d'éviter toute contradiction ; mais il faut en prendre son parti. Deux théories contradictoires peuvent, en effet, pourvu qu'on ne les mêle pas, et qu'on n'y cherche pas le fond des choses, être toutes deux d'utiles instruments de recherche. »*

7.1.4 Peut-on trancher le débat ?

Y-a-t-il un moyen décisif de départager réalisme scientifique et antiréalisme ? Non. Mais examinons les arguments du débat.

Tout d'abord, il faut noter qu'il n'existe pas, par définition, de moyen direct de prouver le réalisme, puisque la réalité extérieure autonome, si elle existe, n'est pas accessible directement. Elle n'est appréhendée qu'au travers des théories. C'est ce qu'exprime Albert Einstein (1879-1955, prix Nobel de Physique 1921) dans sa lettre

⁷⁵ Cité dans « La quête de l'unité », E Klein, M. Lachière-Rey, coll. Biblio essai, Albin Michel, 1996

du 19 juin 1935 à Erwin Schrödinger (1887-1961, prix Nobel de Physique 1933) :



Figure 59: A. Einstein

« *La physique décrit la réalité. Or nous ne savons pas ce qu'est la réalité, nous ne la connaissons qu'à travers la description qu'en donne la physique.* » Nous avons déjà évoqué les principaux points qui interdisent la démonstration directe :

- Les énoncés d'observation ne sont pas vérifiables au sens strict. Leur acceptation ou leur réfutation est une décision, un jugement, qui n'est que motivé et non imposé par l'expérience.
- Les énoncés théoriques de portée universelle ne sont pas vérifiables au sens strict. Ici aussi, les accepter ou les rejeter relève d'un choix.
- Pour les tenants du holisme épistémologique, la question de la valeur de vérité d'un énoncé pris isolément n'a pas de sens.

Puisqu'il n'existe pas de preuve directe en faveur du réalisme, peut-on trouver des preuves indirectes ? Oui, répondent les réalistes, en invoquant **l'argument du miracle de l'efficacité prédictive** : si une théorie est capable de prédire à l'avance les phénomènes, c'est qu'elle correspond à la réalité extérieure. Ce postulat est parfois appelé **principe de réalité d'Einstein**, car il a été formalisé dans les critiques d'Einstein sur la mécanique quantique (voir le chapitre 11.4). L'efficacité prédictive apparaîtrait alors comme un critère empirique de la vérité d'une théorie. Cet argument a d'ailleurs aujourd'hui beaucoup de poids chez de nombreux scientifiques. Non, répondent les antiréalistes, en faisant remarquer qu'il existe des théories empiriquement équivalentes, qui ont des efficacités prédictives très semblables, mais qui sont très différentes, voire contradictoires. C'est le cas des théories corpusculaire et ondulatoire de la lumière. L'argument de l'efficacité prédictive ne clôt donc pas le débat, car il n'impose pas le consensus. De plus, le contre-argument de l'existence de théories contradictoires ayant la même base empirique peut être contré en invoquant l'imprécision des lois : puisque les instruments de mesure sont d'une précision limitée, les énoncés théoriques qu'on en tire souffrent de cette imprécision, et c'est elle qui permettrait cette coexistence, en générant une marge floue sur l'exactitude des prédictions. La vraie théorie-reflet, elle, serait exacte et unique.

Le débat reste ouvert, chacun étant libre de se forger ses propres convictions.

7.2 Le phénoménisme

La science peut-elle aller au-delà des faits qu'elle décrit ? Non, répondent les partisans du **phénoménisme**, pour qui

la science doit se limiter à représenter les phénomènes sans aller au-delà de l'expérience.

La première (et lourde) conséquence du phénoménisme est que l'explication n'est pas du ressort de la science. Celle-ci doit se restreindre à la description des faits. Cette position a par exemple été défendue par le mécanicien et épistémologue Ernst Mach (1838-1916), qui soutenait une version dite sensualiste du phénoménisme. Dans cette



Figure 60: E. Mach

variante rigoureuse, la science doit se borner à ne considérer que les faits sensibles, les lois n'étant que des résumés commodes pour rendre compte d'un grand nombre d'expériences (*La mécanique, exposé historique et critique de son développement*, 1883): « Les notions abstraites qui constituent l'outillage de la physique, les concepts de masse, force, atome, qui n'ont pas d'autre rôle que de rappeler les expériences systématisées dans un but d'économie, sont dotées par la plupart des investigateurs de la nature d'une existence réelle, en dehors de la pensée. Bien plus, on en arrive à croire que ces masses et ces forces sont les choses essentielles à rechercher [...] L'homme qui ne connaîtrait l'univers que par le théâtre, et qui viendrait à découvrir les trucs et les machinations de la scène, pourrait penser que l'univers réel a aussi des ficelles et qu'il suffirait de les tirer pour acquérir la connaissance ultime de toutes choses. Nous ne devons pas considérer

comme bases de l'univers réel des moyens intellectuels auxiliaires dont nous nous servons pour la représentation du monde sur la scène de la pensée. » Ainsi « il n'existe pas de loi de la réfraction, mais rien que de multiples cas de ce phénomène. » Pour Mach, seules les perceptions directes ont une existence : « ce ne sont pas les choses (les objets, le corps), mais bien les couleurs, les tons, les pressions, les espaces, les durées, (ce que nous appelons d'habitude les sensations), qui sont les véritables éléments du monde. »

Tous les phénoménismes ne se résument pas au **sensualisme**. Une autre variante, appelée **physicalisme**, soutient que ce sont les objets de la vie quotidienne (par exemple une table) qui constituent les véritables éléments que le scientifique doit considérer.

Le phénoménisme « pur » est-il possible ? Non, répondent ses détracteurs, parmi lesquels le physicien allemand Ludwig Boltzmann (1844-1906), qui remarque que les équations utilisées par la physique font appel à des cas simplifiés, idéalisés, et qu'elles ne permettent pas de décrire les phénomènes réels dans toute leur complexité. Cette phase d'idéalisation dépassant la simple perception, le phénoménisme intégral n'est pas tenable.

La science peut-elle décrire le monde réel, conçu comme autonome et extérieur à l'observateur, selon les phénoménistes ? Oui et non. Non, en ce sens que le phénoménisme est incompatible avec le réalisme essentialiste : puisque le physicien doit se restreindre à décrire les phénomènes, l'idée d'une théorie qui reflète parfaitement le monde réel et ses mécanismes n'a pas de sens. Oui, en ce sens que les perceptions, elles, peuvent se rapporter à la réalité extra-linguistique. Il s'agirait alors d'un réalisme à propos des perceptions et non de théories. Et, bien sûr, le phénoménisme est parfaitement compatible



Figure 61: L. Boltzmann

avec l'antiréalisme.

7.3 Le conventionnalisme

Le conventionnalisme, dont on attribue la fondation à Henri Poincaré (1844-1912), est la thèse qui soutient que

les théories, les énoncés scientifiques sont des conventions passées entre les scientifiques par raison de commodité ou d'efficacité, et elles résultent de décisions arbitraires. Les lois, les théories ne sont pas obligatoires, et elles auraient pu être autres.

Pour le conventionnalisme radical (illustré par exemple par Pierre Duhem), les observations ne contraignent en rien le scientifique, et les théories sont purement arbitraires, ce qui le rend incompatible avec le réalisme scientifique. Pour le conventionnalisme modéré, défendu par Poincaré, si l'expérience n'impose rien, elle guide en revanche le scientifique. Une théorie n'est donc pas complètement arbitraire, elle ne peut pas dire n'importe quoi, car les énoncés d'observation permettent d'exclure les théories aberrantes. L'argument de Poincaré pour défendre le



Figure 62: H. Poincaré

conventionnalisme dans les sciences empiriques repose sur le fait qu'il est impossible de vérifier au sens strict les énoncés d'observation. Comme on l'a déjà vu, ceux-ci ne peuvent que guider le scientifique dans ses choix. Cette forme de conventionnalisme modéré est compatible avec le réalisme scientifique : une théorie, même partiellement subjective, peut retenir quelque chose du monde autonome extérieur.

7.4 Le réductionnisme

7.4.1 Isolationnisme et constructionnisme

La démarche scientifique moderne est basée sur le *réductionnisme de méthode* (voir le chapitre 5.5.1): confronté au monde très complexe qui nous environne, le scientifique isole du reste de l'univers les objets et les phénomènes sur lesquels il souhaite travailler. Sans cette nécessaire séparation, qui apparaît explicitement dans les définitions de la science vues au chapitre 2, la démarche scientifique, telle qu'elle est conçue de nos jours, n'est pas concevable, car une trop grande complexité rend l'analyse impossible. C'est également cette démarche qui est à la base de la distinction entre les différentes disciplines et sous-disciplines scientifiques. Il faut également noter que l'accroissement des connaissances scientifiques rend aujourd'hui impossible pour une seule personne de connaître l'ensemble des résultats relevant d'

une discipline ou même d'une sous-discipline dans laquelle elle serait spécialisée, ce qui rend nécessaire en pratique, à la fois pour la recherche et l'enseignement, la définition de régions limitées du savoir scientifique. Aussi chaque discipline (ou sous-discipline reconnue comme telle) aborde-t-elle son domaine d'étude d'une manière qui lui est propre ; elle construit ses paradigmes et ses programmes de recherche. Bien entendu, il existe des interactions et des échanges avec d'autres disciplines et sous-disciplines qui traitent de sujets proches ou qui utilisent des outils théoriques (les mathématiques ont ici un rôle privilégié, car toutes les sciences empiriques sont aujourd'hui mathématisées) ou expérimentaux similaires. Cette identification d'objets d'études propres correspond à la méthode réductionniste cartésienne (voir le chapitre 5.5.1) ; on parle d'**isolationnisme**⁷⁶ ou encore de **réductionnisme de méthode**. Cette démarche est pratiquée dans toutes les disciplines, y compris les sciences humaines et sociales. Elle est considérée par beaucoup d'épistémologues, par exemple Karl Popper, comme fondatrice de la science moderne. Rappelons toutefois qu'un écueil pour le réductionnisme de méthode est la description des **phénomènes émergents**.

A cet **isolationnisme global** (puisque appliqué à l'échelle des disciplines scientifiques), il faut ajouter un **isolationnisme local**, interne à chaque sous-discipline, qui correspond au fait que chaque théorie scientifique est basée sur des objets (atomes, ondes, lois, principes, ...) qui sont bien répertoriés et en nombre fini, et qui lui sont propres.

La méthode cartésienne comprend une seconde phase : la reconstruction. Une fois que le problème de départ a été scindé en plusieurs sous-problèmes plus simples, la solution globale est obtenue en combinant les solutions de chacun de ces sous-problèmes après qu'ils aient été résolus. On parle alors de **constructionnisme**.

La démarche scientifique est donc toujours écartelée entre une démarche simplificatrice et un accroissement continu de la complexité son objet d'étude, qui croît au fur et à mesure que les succès s'accumulent. C'est ce que note H. Poincaré dans *La science et l'hypothèse* : « Dans l'histoire du développement de la physique, on distingue deux tendances inverses : d'une part, on découvre à chaque instant des liens nouveaux entre des objets qui semblaient devoir rester à jamais séparés ; les faits épars cessent d'être étrangers les uns aux autres ; ils tendent à s'ordonner en une imposante synthèse. La science marche vers l'unité et la simplicité. D'autre part, l'observation nous révèle tous les jours des phénomènes nouveaux : il faut qu'ils attendent longtemps pour leur place et parfois, pour leur en faire une, on doit démolir un coin de l'édifice. Dans les phénomènes connus eux-mêmes, où nos sens grossiers nous montraient l'uniformité, nous apercevons des détails de jour en jour plus variés ; ce que nous croyions simple redevient complexe et la science paraît marcher vers la variété et la complication. »⁷⁷

Si le recours à l'isolationnisme est couramment accepté, tant au niveau local qu'au niveau global, le constructionnisme, quant à lui, soulève de nombreuses questions.

⁷⁶ Pour une discussion détaillée du réductionnisme, de l'isolationnisme et du constructionnisme et de l'évolution de la physique, voir « La quête de l'unité », E. Klein, M. Lachièze-Rey, collection « essais », Albin Michel, 1996

⁷⁷ cité dans « La quête de l'unité », E. Klein, M. Lachièze-Rey, collection « essais », Albin Michel, 1996

Au niveau global, le constructionnisme fait appel à la notion d'unité de la science. Dans sa **version radicale** ou **ontologique**, il s'agit d'arriver à une théorie scientifique unique qui couvrirait l'ensemble des champs disciplinaires actuels, allant des sciences de la nature aux sciences humaines et sociales. Cette idée d'une « théorie ultime » peut être vue comme une survivance du programme mécaniste, qui visait à expliquer l'ensemble de l'univers, y compris le fonctionnement des êtres vivants considérés comme des automates, sur la base de la mécanique. Ce programme était énoncé de manière très simple par Descartes : « *L'univers est une machine où il n'y a rien d'autre à considérer que les figures et les mouvements de ses parties* ». Une telle théorie n'est plus jugée possible ni envisagée de nos jours : la biologie n'est pas une chimie appliquée, la psychologie et la sociologie ne sont pas réductibles à la biochimie. Dans sa **version modérée**, l'unité de la science est une **unité méthodologique** ; c'est dans le partage de méthodes d'acquisition des données, dans les procédures de validation des hypothèses ou encore dans l'emploi de certains outils théoriques (comme les mathématiques : les statistiques et les probabilités sont employées aussi bien par les sciences humaines que les sciences de la nature) que se trouvent les points communs aux disciplines scientifiques et que peut être comprise la notion d'unité du champ scientifique.

Au niveau local, c'est-à-dire au sein d'une discipline ou d'une sous-discipline, la portée du constructionnisme est également sujette à débat. On peut à ce niveau faire la distinction entre unification horizontale et unification verticale. L'**unification horizontale** résulte de la fusion de branches ou de sous-disciplines jusqu'alors séparées au sein d'une même discipline pour donner naissance à une nouvelle théorie plus puissante. Un exemple est la fusion de la théorie de la lumière et de l'électromagnétisme grâce à la théorie de Maxwell au XIX^e siècle. L'**unification verticale** désigne l'émergence d'une nouvelle théorie dont l'impact déborde le champ couvert par les théories initialement fusionnées. Ceci est illustré par la mécanique quantique, qui unifie matière et rayonnement.

Une question longuement débattue est celle de l'unification de la physique : une unification locale totale de cette discipline est-elle possible ? Une première question qui se pose est celle de la définition de ce que l'on appelle la physique. Dans son sens le plus général, la physique est la science qui a pour objet l'ensemble du monde inanimé (bien qu'il existe une biophysique et une éconophysique !), et dont la chimie ou la mécanique ne sont que des sous-disciplines. Ce caractère flou des frontières de la discipline complique bien sûr la réflexion sur son unité. Mais retenons que, même dans un sens très restreint, la physique comprend les sciences de la matière.

L'histoire de la physique présente plusieurs exemples d'unifications : outre l'unification des théories de l'électricité et du magnétisme par Maxwell au XIX^e siècle, on peut citer l'unification de l'électromagnétisme et de la dynamique par Einstein au XX^e siècle. La physique quantique donnera un cadre descriptif unifié pour la matière et le rayonnement. Plus tard, la mécanique quantique et la relativité restreinte seront unifiées au sein de la théorie quantique des champs, qui permettra le développement de théories unifiées pour trois des quatre interactions fondamentales (interactions faible et forte, force électromagnétique). Est-il imaginable que la physique puisse être condensée en une seule et unique théorie fondamentale ? Tout au long de l'histoire, des penseurs ou des chercheurs ont cru pouvoir répondre positivement à cette question, principalement parce qu'ils pensaient que la théorie physique était complétée ou en passe de l'être. Ainsi Francis Bacon écrivait au XVI^e siècle : « *Il n'y a en réalité qu'une poignée de phénomènes particuliers des arts et des*

sciences. *La découverte de toutes les causes et de toutes les sciences ne serait qu'un travail de quelques années.* »⁷⁸ A plusieurs siècles de distance, après l'avènement de la physique newtonienne, de l'électromagnétisme et la naissance de la physique statistique, Lord Kelvin, à la fin du XIX^e siècle, pensait que la physique était complétée. Environ trente ans plus tard, après le développement de la théorie de la relativité et de la mécanique quantique, un des pères de cette dernière, le prix Nobel de Physique Max Born, écrivait à la fin des années 1920 que « *La physique serait achevée dans six mois.* » La réponse moderne est négative. La raison principale est que la description du monde par les différentes branches de la physique fait appel à des niveaux de description différents (allant de l'échelle cosmologique à la physique des particules, en passant par la mécanique des milieux continus, la chimie et la physique nucléaire) pour lesquels des théories et des modèles adéquats ont été bâtis. Une unification totale signifierait qu'il est possible de **déduire** toutes les lois aux différents niveaux de description employés au sein des différentes branches de la physique à partir du niveau de description le plus fondamental, celui de la physique des particules. Or ceci est impossible. C'est ce que notait le physicien A.J Legget en 1983, lors d'une conférence internationale consacrée au réductionnisme : « *Même dans le cas où selon un croyance largement répandue une théorie « de niveau supérieur » peut être réduite à une théorie « de niveau inférieur », par exemple la physique du solide à la physique atomique et à l'électromagnétisme, il s'agit d'une complète illusion. Par exemple, je défie quiconque dans cette salle de prouver rigoureusement la loi d'Ohm pour un échantillon réel, à partir de la théorie atomique et de l'électromagnétisme.* »⁷⁹ Outre les difficultés techniques, le problème rencontré ici est un **problème d'intelligibilité** (et donc d'utilité en pratique) d'une telle théorie « totale ». C'est ce que notait Dirac en 1929⁸⁰ : « *Les lois physiques fondamentales nécessaires pour la théorie mathématique de la plus grande partie de la physique et de la totalité de la chimie sont ainsi complètement connues, et la difficulté est seulement que l'application de ces lois conduit à des équations beaucoup trop compliquées pour qu'on puisse les résoudre.* » Les modèles et les théories ne sont pas de simples poupées russes, emboîtées les unes dans les autres et déductibles les unes des autres par des procédés mathématiques. Chaque niveau de description fait appel à des objets qui lui sont propres, et qui sont pertinents pour fournir un cadre explicatif pour les phénomènes qu'il étudie. Par exemple, la température, la pression et la masse volumique d'un gaz définies en mécanique des milieux continus sont plus utiles que les particules élémentaires pour décrire le comportement d'un gaz. Mais il existe des liens entre les différents niveaux de description de la nature : même si le chimiste n'utilise pas la physique des particules pour construire sa description des ions et des molécules, les résultats de la physique fondamentale lui offrent des concepts et des résultats qui peuvent le guider dans sa propre recherche. De même, certains concepts théoriques et outils mathématiques sont partagés par de nombreuses branches de la physique : théorie des groupes, symétries, statistiques et probabilités, théorie des bifurcations et des catastrophes ... Il existe donc bien une unité de la physique, mais plus floue et complexe qu'une simple hiérarchie de théories gigognes.

⁷⁸ cité dans « La quête de l'unité », E. Klein, M. Lachièze-Rey, coll. « essais », Albin Michel, 1996

⁷⁹ cité dans « La quête de l'unité », E. Klein, M. Lachièze-Rey, coll. « essais », Albin Michel, 1996

⁸⁰ cité dans « La quête de l'unité », E. Klein, M. Lachièze-Rey, coll. « essais », Albin Michel, 1996

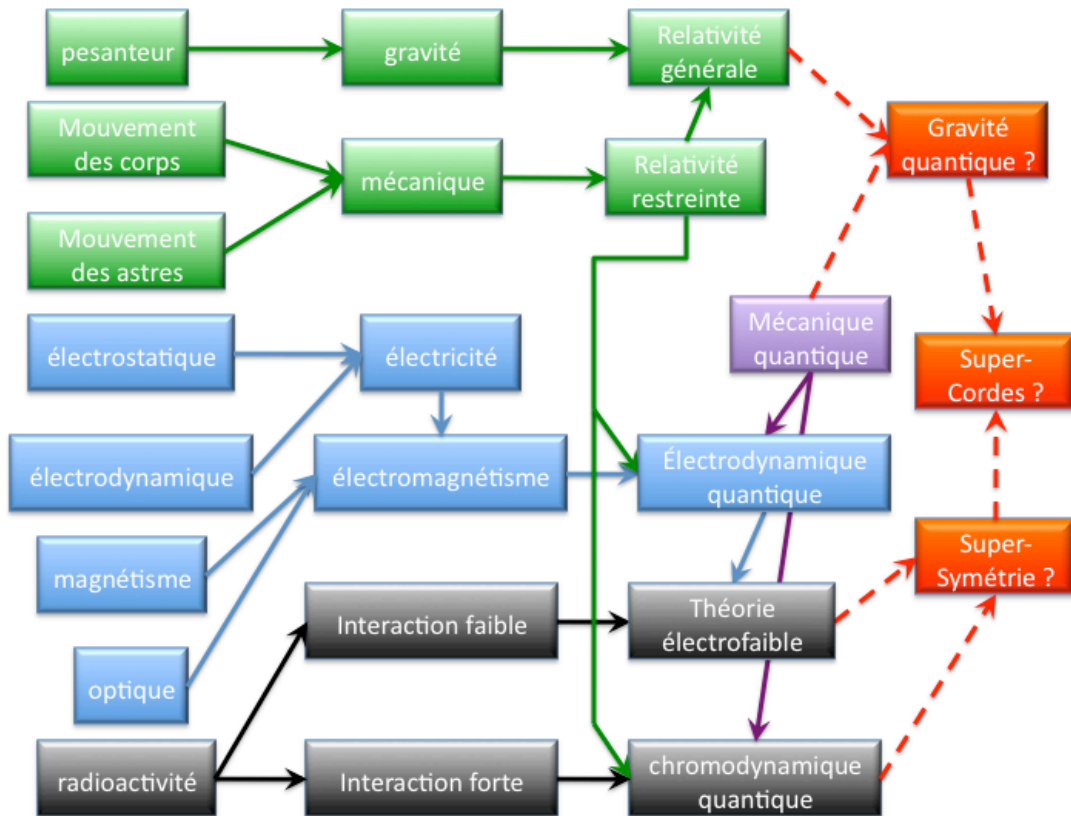


Figure 63 schéma de l'évolution de la physique par le biais d'unifications successives. Les théories en rouge sont encore en développement ou à développer, et les flèches en pointillé représentent les liens espérés (adapté d'après « L'univers des particules », M. Crozon, coll. Points Sciences, Seuil, 1999)

Notons que le réductionnisme est neutre vis-à-vis du réalisme et de l'antiréalisme : la démarche réductionniste est envisageable dans les deux cas.

7.4.2 Un exemple issu de la physique moderne : le modèle standard

La discussion précédente va maintenant être illustrée par exemple aujourd'hui célèbre de programme d'unification locale dans le domaine de la physique fondamentale : le **modèle standard**. Le modèle standard est le niveau le plus fondamental de description de la matière et des forces présentes dans l'univers qui soit très fortement corroboré par les données expérimentales.

Selon le modèle standard, il existe **quatre forces** ou **interactions fondamentales** : *la gravitation, l'interaction forte, l'interaction faible et l'interaction électromagnétique*. Chacune de ces interactions est portée par une particule spécifique (voir les tableaux ci-dessous). Il y a en tout 12 de ces particules ; elles forment la famille des **bosons**.

La matière est, quant à elle, constituée par une autre famille de particules : les **fermions**. Ils sont regroupés en deux sous-familles : les **leptons**, qui sont insensibles à l'interaction forte, et les **quarks**, qui y sont sensibles. Il y a en tout 6 leptons et 6 quarks, soit 12 particules élémentaires, auxquelles il faut ajouter les 12 **antiparticules**

associées pour la formation de l'**antimatière**⁸¹. Les fermions sont classés en trois familles, chacune comprenant 2 quarks et 2 leptons, dont une seule constitue la matière stable, les fermions des deux autres familles n'apparaissant pas dans des configurations stables.

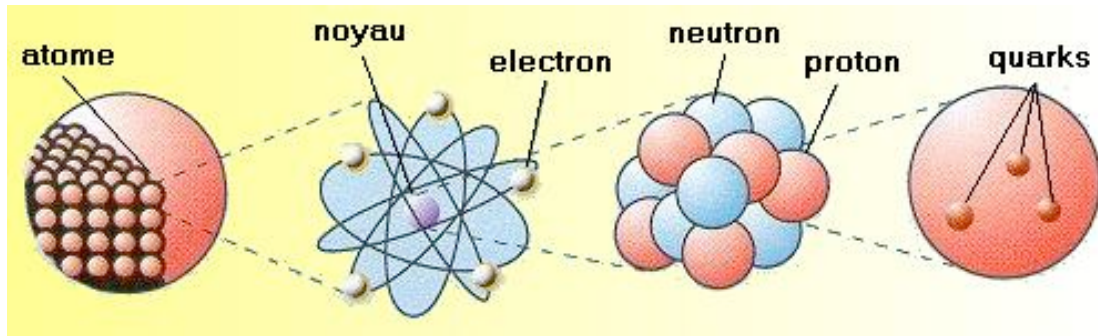


Figure 64 Illustration schématique des niveaux de description menant au Modèle Standard. On représente ici les particules comme des corpuscules, en négligeant leur caractère quantique.

Tableau 5 Caractéristiques des 4 interactions fondamentales connues. Le graviton n'a jamais été observé à ce jour, car il correspondrait à des niveaux d'énergie hors d'atteinte des moyens expérimentaux. La boson de Higgs, prédit par la théorie électrofaible et qui serait responsable de la masse, n'a pas été observé à ce jour et n'est pas figuré dans le tableau.

Interaction	Portée	Intensité	Temps caractéristique	Particule porteuse (boson)	Rôle
gravitation	infinie	10^{-39}		graviton ?	Structuration de l'Univers à grande échelle
électromagnétique	infinie	10^{-2}	10^{-16} s	photon	Lumière, cohésion des atomes, molécules, cristaux
forte	$\sim 10^{-15}$ m	1	10^{-23} s	8 gluons	Cohésion des protons et des noyaux atomiques
faible	10^{-17} m	10^{-5}	1 à 10^{-12} s	W^+ , W^- , Z^0	Fusion nucléaire, radioactivité β

⁸¹ L'antimatière a été prédite théoriquement par Paul Dirac en 1930 et découverte en 1932 par Carl Anderson au moyen d'une chambre à brouillard.

Tableau 6 Particules élémentaires de la famille des fermions. La masse est exprimée en électron-Volt (eV)⁸². La masse des neutrinos est théoriquement nulle, et très difficile à mesurer (on ne possède que des limites supérieures). Les quarks ne sont pas observables de manière isolée : on ne possède donc que des estimations indirectes.

		Leptons		Quarks	
Charge électrique		0	+1	- 1/3	+ 2/3
Famille 1 (constituants matière stable)	Nom	Neutrino électronique ν_e	Electron e	Down d	Up u
	Masse	< 10 eV	0,511 MeV	3-9 MeV	1-5 MeV
	Découverte		1897	1968	1968
Famille 2	Nom	Neutrino muonique ν_μ	Muon μ	Strange s	Charm c
	Masse	< 0,17 MeV	105,66 MeV	60-170 MeV	1,1-1,4 GeV
	Découverte	1962	1937 (confirmation 1945)	1968	1974
Famille 3	Nom	Neutrino tau- ique ν_τ	Tau τ	Bottom (beauty) b	Top (truth) t
	Masse	< 18,2 MeV	1777,05 MeV	4,1-4,4 GeV	~175 GeV
	Découverte	1976	1976	1977	1996
Sensibilité aux interactions		faible	Faible, électromagnétique	Faible, forte, électromagnétique	

⁸² 1 eV est défini comme l'énergie acquise par un électron soumis à une différence de potentiel de 1 Volt, soit $1,6 \cdot 10^{-19}$ joule. On fait le lien avec la masse en utilisant la fameuse formule d'Einstein $E=mc^2$, ou c désigne la vitesse de la lumière. On trouve $1 \text{ eV} = 1,6 \cdot 10^{-36} \text{ kg}$

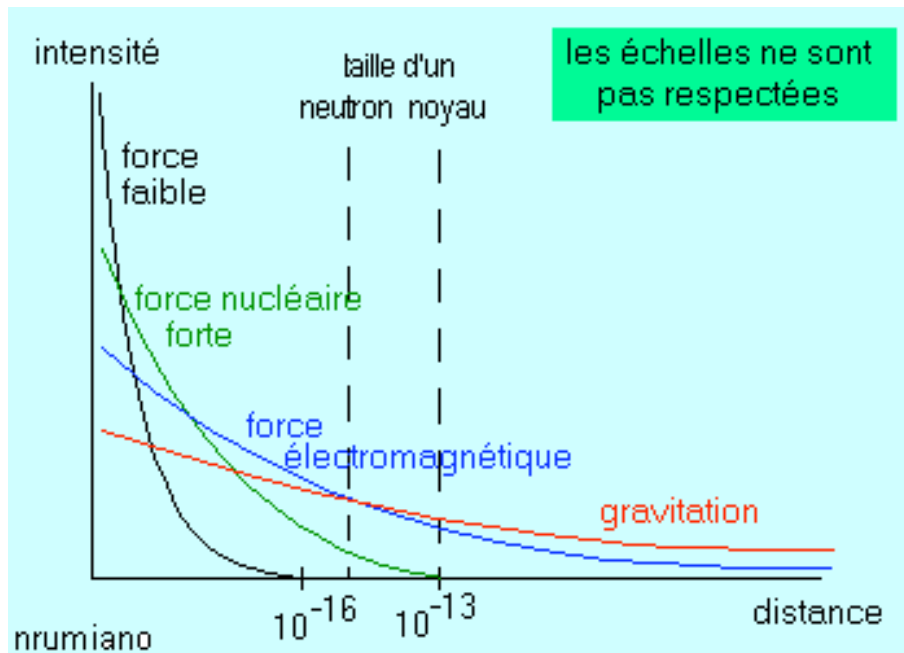


Figure 65 schéma de la variation d'intensité des interactions fondamentales en fonction de la distance

Le modèle standard a été défini dans une volonté d'unification de la théorie électrofaible et de la chromodynamique quantique dans le but d'obtenir un cadre unifié pour 3 interactions fondamentales : l'interaction électromagnétique, l'interaction faible et l'interaction forte.

La **théorie électrofaible**, développée durant les années 1950 et 1960, est une théorie quantique des champs qui unifie l'interaction électromagnétique et l'interaction faible. L'interaction électromagnétique était jusque là décrite de manière isolée par une autre théorie quantique des champs, l'**électrodynamique quantique**, développée durant les années 1930. Cette théorie elle-même issue de trois théories antérieures : l'**électromagnétisme**, la **relativité restreinte** et la **mécanique quantique**. Elle permet de décrire les interactions électromagnétiques entre deux objets par l'échange de photons. A une force est donc associée une particule porteuse (un boson : le photon), dont la vitesse de propagation ne peut dépasser la vitesse de la lumière dans le vide. Cette théorie, bien que limitée dans son champ d'application, est très précise et en très bon accord avec les observations, ce qui a conduit à rechercher son extension pour les autres forces fondamentales. Cette extension a été pleinement réalisée pour la force faible au moyen d'outils mathématiques très évolués (théorie des groupes, symétries, ...), conduisant à la théorie électrofaible. Ce développement a valu le prix Nobel de Physique en 1979 à Sheldon Glashow, Abdus Salam et Steven Weinberg. Les développements mathématiques conduisaient à la prédiction de l'existence de nouvelles particules : les bosons W^+ , W^- et Z^0 , qui ne seront découverts que plus tard au début des années 1980. Cette corroboration a posteriori est considérée comme un argument fort en faveur de la théorie électrofaible. Mais cette théorie prédit également l'existence d'un autre boson, appelé **boson de Higgs**, associé à un champ de force dont l'effet est de donner une masse non-nulle aux bosons autres que le photon. Cette particule de Higgs joue un rôle central, mais il n'existe aucune preuve directe de son existence à ce jour.

La **chromodynamique quantique** est une théorie de l'interaction forte, qui assure la cohésion du noyau atomique (qui sinon éclaterait sous l'action de la répulsion électromagnétique). Elle est basée sur l'existence des **quarks**, dont elle prédit le nombre et les propriétés : charge électrique, étrangeté (-1, 0, 1), nombre baryonique ($1/3$ ou $-1/3$) et couleur⁸³. Les quarks sont des particules étranges, qui vont toujours par groupe de 3. La cohésion du proton, qui est formé de quarks, est assurée par une force, appelée **force de couleur** (d'où le nom de **chromodynamique**), qui est portée par les **gluons** qui assurent les échanges de couleur entre les quarks. Les quarks ne sont pas observables directement, car ils ne restent jamais isolés et sont confinés au sein des particules lourdes qu'ils forment. Les gluons, quant à eux, ont été observés en 1978.

Le modèle standard est extrêmement puissant, car il permet de réduire la description des centaines de particules observées à 12 particules fondamentales qui interagissent via 3 forces fondamentales. Il est corroboré par de très nombreuses expériences. Toutefois, les théoriciens lui trouvent plusieurs faiblesses :

- Le degré d'unification de l'interaction forte avec les deux autres est moindre que celui atteint entre l'interaction électromagnétique et l'interaction faible. Certains y voient une juxtaposition formelle plutôt qu'une unification réelle. Des tentatives d'unification plus poussées, comme la *théorie de supersymétrie*, ont été avancées. Cette théorie prédit l'existence de nouvelles particules. L'unification entre les trois forces auraient lieu à des niveaux d'énergie de l'ordre de 10^{16} GeV, ce qui correspond à des échelles⁸⁴ de l'ordre de 10^{-30} à 10^{-32} m. Ce niveau d'énergie est très supérieur aux capacités des moyens expérimentaux existants ou en préparation, ce qui semble interdire tout espoir de corroboration directe.
- Le modèle standard nécessite l'introduction d'une vingtaine de constantes qui ne sont pas déterminées par la théorie.
- Il ne dit rien sur la différence entre les quarks et les leptons
- Il ne dit pas pourquoi il existe 3 familles de fermions, et pas 2, 4 ou un autre nombre.
- Certaines corroborations sont toujours manquantes : la masse de certains fermions, les neutinos, n'ont toujours pas été mesurées ; de même, le boson de Higgs n'a toujours pas été observé.
- La description par 12 fermions et 12 bosons semblent trop compliquée pour certains physiciens, qui cherchent une théorie faisant apparaître moins de constituants élémentaires.
- Le modèle standard ne permet pas de prendre en compte la quatrième interaction, à savoir l'interaction gravitationnelle. L'unification des 4 interactions est l'objet de théories plus récentes, comme la théorie des

⁸³ La couleur est une propriété inventée pour satisfaire le **principe de Pauli**, qui stipule que 2 fermions ne peuvent pas être dans des états quantiques identiques. Pour résoudre certains problèmes théoriques, il est apparu qu'il fallait donner une nouvelle propriété aux quarks au sein d'une triade pour éviter qu'ils soient identiques. Cette propriété a été appelée couleur, car la somme des couleurs au sein d'un groupe de 3 quarks s'annule, comme lors de la composition des couleurs en base RGB.

⁸⁴ Le lien entre l'énergie et une échelle caractéristique est donné par la relation de de Broglie, qui indique que la longueur d'onde de l'onde associée à une particule est inversement proportionnelle à la quantité de mouvement (produit de la masse par la vitesse) de celle-ci. Donc, plus la masse ou la vitesse d'une particule est grande, et plus petite est la longueur d'onde, qui est prise comme échelle caractéristique.

supercordes, qui inclut une formulation quantique de la gravitation. De telles théories implique une augmentation du nombre de dimensions (entre 10 et 23), et des échelles de l'ordre de l'échelle de Planck (10^{-35} m), soit un niveau d'énergie de 10^{19} GeV très supérieur aux limites imaginées des dispositifs expérimentaux.

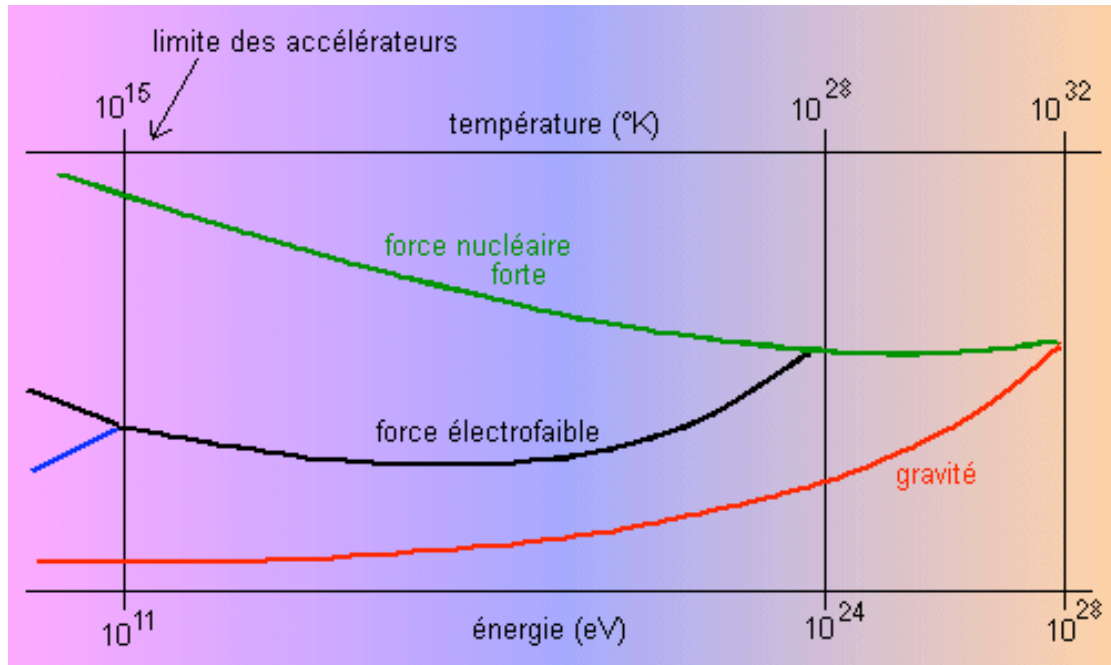


Figure 66 Représentation schématique de l'unification des interactions fondamentales en fonction de la température ou de l'énergie

8 Les mathématiques dans les sciences empiriques

Nous allons maintenant discuter de la place et du rôle des mathématiques dans les sciences empiriques « dures », à savoir la physique et la mécanique, qui sont les deux disciplines dans lesquelles la mathématisation est la plus avancée. Avant d'aller plus loin, il est important de noter que l'utilisation des mathématiques est « relativement récente », puisqu'il est classique d'attribuer cette démarche à Galilée qui écrivait en 1623 dans son texte intitulé *Il Saggiatore (L'Essayeur)* « *La philosophie est écrite dans cet immense livre qui se tient toujours ouvert devant nos yeux, je veux dire l'Univers, mais on ne peut le comprendre si on ne s'applique pas d'abord à en comprendre la langue et à connaître les caractères avec lesquels il est écrit. Il est écrit en langage mathématique ... et ses caractères sont des triangles, des cercles et autres figures géométriques ...* » C'est en effet lui qui, le premier, a appliqué les mathématiques à la mécanique et la mécanique à l'astronomie, formant un ensemble de savoirs cohérents. Auparavant, pour donner une description abrégée (voire simpliste) de l'emploi des mathématiques dans la description du monde, on peut dire que la possibilité d'appliquer les mathématiques au monde était défendue par les savants platoniciens et combattue par les aristotéliens.

8.1 Un exemple : la mécanique

8.1.1 Exemple 1 : le calcul quantitatif

Commençons par un exemple simple. On considère un corps pesant de masse m , dans le vide, initialement au repos. Quelle distance D aura-t-il parcouru au temps t ? La réponse s'obtient à l'aide de la théorie du calcul intégral en intégrant deux fois la seconde loi de Newton, ou une fois la loi sur la quantité de mouvement. La réponse (bien connue depuis Galilée) est :

$$D = \frac{1}{2}gt^2$$

8.1.2 Exemple 2 : la manipulation symbolique

Passons maintenant à un second problème, moins trivial: l'utilisation des mathématiques dans la formulation de la mécanique newtonienne. Partons de la seconde loi de Newton

$$\sum \vec{F} = m\vec{\gamma}$$

et exprimons la en remplaçant l'accélération par son expression en fonction de la vitesse :

$$\sum \vec{F} = m \frac{d\vec{v}}{dt}$$

Cette nouvelle formulation est équivalente à la première, mais elle apparaît sous la forme d'une équation différentielle, et cela parce que nous avons exprimé l'accélération comme étant la dérivée de la vitesse par rapport au temps.

La seconde loi de Newton est exprimée en fonction de l'accélération. Peut-on trouver une expression analogue pour la vitesse ? Oui, en intégrant une fois sa forme usuelle. Nous obtenons (en assumant que la masse reste constante)

$$m\vec{v}(t) = m\vec{v}(t_0) + \int_{t_0}^t \sum \vec{F}(\tau) d\tau$$

une loi qui donne la valeur de la quantité de mouvement en fonction de la valeur à l'instant initial t_0 . Elle est équivalente à la première, mais ne fait plus intervenir l'accélération. On peut donc exprimer les trois lois de Newton sans faire intervenir le concept d'accélération. Que peut-on déduire, à partir de cette loi ?

Tout d'abord, qu'à tout instant t tel que

$$\int_{t_0}^t \sum \vec{F}(\tau) d\tau = 0$$

la quantité de mouvement est égale à celle de l'instant initial. Ensuite que, pour un système tel que la résultante des forces est nulle à chaque instant depuis l'instant initial, la quantité de mouvement est constante au cours du temps. Si la masse est constante, on en déduit que pour un tel système la vitesse est obligatoirement constante et égale à sa valeur initiale. Si elle est initialement nulle, le corps restera au repos. Sinon, le corps observera un mouvement rectiligne uniforme. Nous venons donc de déduire la première loi de Newton. Chose importante, nous ne l'avons pas utilisée pendant la suite de nos opérations. Nous pouvons donc énoncer

$$\sum \vec{F} = \vec{0} \iff \text{Mouvement Rectiligne Uniforme ou repos}$$

Qu'avons-nous fait ? Nous avons utilisé le calcul différentiel pour trouver une nouvelle loi, équivalente à la seconde loi de Newton, mais qui ne fait pas intervenir l'accélération. De plus, nous en avons déduit une autre loi, qui se trouve être exactement la première loi de Newton. Ce qu'il faut voir ici, c'est que le calcul est purement formel : durant tout le temps du développement, nous n'avons fait que manipuler des symboles, sans en connaître la signification physique. Nous n'avons fait qu'appliquer des règles d'intégration données par les mathématiques. Ce qui est remarquable, c'est que l'outil mathématique nous a permis de trouver une nouvelle loi et d'en retrouver une autre. Encore plus remarquable, on peut, en fin d'opération, redonner un sens physique au symbole et en déduire quelque chose sur la dynamique, et cela sans avoir recours à la méthode expérimentale ! Celle-ci, et c'est heureux, confirme nos nouvelles lois. Nos nouvelles lois sont d'ailleurs, pour reprendre des

distinctions déjà rencontrées, *empiriquement équivalentes* aux deux premières lois de Newton.

Par contre, dans le premier exemple, nous n'avons pas cherché, par des manipulations formelles, à formuler de nouvelles lois. Nous nous sommes intéressés à un fait particulier : la distance parcourue par un objet identifié, concret. Ce que nous avons fait, c'est utiliser les mathématiques et les lois de la dynamique pour prédire à l'avance une valeur numérique, D . Ici, les mathématiques nous ont offert une capacité de prédiction.

8.1.3 Dédire les lois de la physique : symétries et principes variationnels

Le second exemple discuté plus haut n'est qu'illustratif, car la nouvelle formulation n'apporte rien par rapport à la formulation classique des lois de la mécanique newtonienne. Elle est même sans aucun doute moins pratique. Mais nous allons voir maintenant que les mathématiques, couplées à des grands principes généraux de la physique classique, permettent de **dédire** les lois de Newton et celles d'autres branches de la physique classique, comme les lois de Maxwell en électromagnétisme, mais aussi celles de la théorie de la relativité ou encore celle de la théorie quantique des champs. Le point important est que les lois sont déduites de l'analyse mathématique, et non pas obtenues comme le fruit d'observations et d'expérimentations. Nous allons nous intéresser, sans aborder les détails techniques, à l'obtention des lois de la mécanique classique.

Tout d'abord, il faut se donner des grands principes généraux dans l'établissement des lois d'une science empirique.

La première question est celle du choix des types de variables que l'on va utiliser pour décrire un système physique. Ainsi, en mécanique, on introduit la position dans l'espace-temps, la vitesse, la masse, l'énergie (potentielle, cinétique ou totale), la quantité de mouvement, ... Pour décrire la trajectoire d'un point matériel, il faut considérer 6 valeurs différentes, appelées **degrés de liberté** : les trois coordonnées de la position spatiale, et les trois composantes de la vitesse à un instant donné. Le problème engendre donc un espace de dimension 6, appelé **espace de configuration**. Notons qu'il est possible de trouver d'autres types de degrés de liberté que la vitesse et la position pour représenter le système. Pour définir ces grandeurs et les grandeurs dérivées, il faut se doter au préalable d'un référentiel, et se pose la question de leur variation si on change de référentiel. Les deux sortes de variables utilisées en physique sont :

- Les **quantités invariantes**, dont la valeur ne dépend pas du référentiel choisi. L'énergie totale (cinétique+potentielle), qui est un concept central en mécanique et en physique est une quantité invariante pour un système isolé (si l'on fait abstraction des phénomènes dissipatifs, comme dans le cadre de la mécanique Newtonienne). D'autres quantités invariantes en mécanique newtonienne sont : la masse, la distance et la durée. Notons que toutes ses quantités sont des **scalaires**.

- Les quantités **covariantes** (étymologiquement « qui varient avec »), qui sont des quantités dont on sait déduire le changement de valeur à partir de la formule de changement de référentiel. Deux quantités covariantes importantes en mécanique sont la quantité de mouvement et le moment cinétique (qui sont des quantités conservées pour un système isolé). Ce sont des quantités **vectérielles**.

Notons qu'ils existent des quantités qui ne sont ni invariantes, ni covariantes. Celles-ci seront généralement rejetées car elles dépendent intrinsèquement du référentiel choisi.

Venons-en maintenant aux symétries que doivent vérifier les lois physiques pour être acceptables. L'espace et le temps étant homogènes (c'est le **principe copernicien**), les lois de la physique ne doivent pas dépendre :

- De la position en espace, c'est-à-dire être invariantes par **translation en espace**
- De l'orientation du repère, et donc être invariantes par une **rotation du référentiel**
- De l'instant, et donc être invariantes par **translation en temps**

Venons en maintenant au second grand principe : le **principe de moindre action**, né des discussions entre Descartes et Fermat sur l'optique géométrique (voir le paragraphe 12.6.1 pour une discussion historique). Ce principe stipule que la solution physique, c'est-à-dire celle qui est observée, est celle qui minimise une quantité nommée **action**. Par exemple, il existe une infinité de trajectoires d'un point matériel soumis à un champ de gravité pour se rendre d'un point A à un point B. Mais une seule trajectoire sera suivie par le point (et toujours la même) : c'est celle qui minimise une quantité appelée **intégrale d'action**. En mécanique, l'intégrale d'action est définie comme étant **l'intégrale en temps le long d'une trajectoire de la différence entre l'énergie cinétique et l'énergie potentielle** (cette différence est appelée le **lagrangien**)⁸⁵. Le principe de moindre action fait partie de la famille des principes variationnels, car il est exprimé comme une intégrale, et la solution du problème minimise la valeur de cette intégrale (et donc toute modification, même légère, de la trajectoire solution conduira à une augmentation, donc une variation, de l'intégrale d'action – d'où l'adjectif variationnel).

Le résultat tout à fait remarquable de la mécanique rationnelle développée par Maupertuis, Euler puis Lagrange et Hamilton sur les bases posées par Descartes et Fermat est que les quantités invariantes et les lois qui leur sont associées se déduisent du couplage entre le principe copernicien et le principe de moindre action (voir le tableau ci-dessous). Cet édifice théorique est couronné par le **théorème de Noether**⁸⁶, qui établit la **correspondance entre relativité** (c'est-à-dire la non-observabilité de certaines entités absolues), **symétrie** (c'est-à-dire invariance par des transformations de symétrie – translation, rotation, ...) et **lois de conservation** (c'est-à-dire conservation et donc observabilité de certaines quantités). Ce résultat se généralise aux autres disciplines déjà citées dans ce chapitre.

⁸⁵ Le principe de moindre action s'exprime de la même manière, mais avec des lagrangiens différents, en électromagnétisme, en théorie de la relativité et en mécanique quantique.

⁸⁶ Emmy Noether (1882-1935), mathématicienne allemande

La formalisation de ce qui vient d'être dit est basée sur de nombreuses théories mathématiques : calcul différentiel, calcul intégral, théorie des groupes, ... On voit donc ici que les mathématiques permettent, à partir de deux principes très généraux et de la définition des lagrangiens appropriés, déduire les quantités conservées (qui seront celles à mesurer en priorité lors des expériences) et les équations associées.

Tableau 7: dérivation des lois de la mécanique classique

symétrie	Loi de conservation	Non-observable
Translation dans l'espace	Quantité de mouvement	Origine de l'espace
Rotation dans l'espace	Moment cinétique	Direction privilégiée
Translation dans le temps	Energie totale	Origine du temps

8.2 Les apports de la mathématisation

Au vu de l'exemple précédent, nous pouvons réfléchir sur l'apport de la mathématisation des sciences empiriques dures (mécanique, physique). Quels sont les apports des mathématiques ? On peut distinguer plusieurs fonctions :

- L'emploi de mathématiques permet la production d'énoncés d'observation à partir des lois et d'hypothèses particulières (conditions aux limites, conditions initiales), et donc participe de la démarche de vérification des théories mathématisées. C'est le calcul qui nous permet de faire des prédictions, puis de vérifier que celles-ci sont réalisées lors d'expériences adéquates. C'est ce que nous avons fait dans le second exemple au début de ce chapitre.
- Les mathématiques ont une fonction de systématisation : elles permettent d'unifier une théorie en fournissant un cadre qui permet de relier ses différents composants, de les combiner et de vérifier leur compatibilité. Dans le premier exemple, nous avons pu vérifier que les deux premières lois de Newton sont compatibles, et que la dynamique de Newton vérifie le critère de la vérité-cohérence.
- Les mathématiques permettent de condenser un ensemble (de taille quelconque) d'énoncés d'observation en un unique énoncé théorique. Ainsi, les lois de Newton ou la loi de Boyle-Mariotte représentent les résultats d'une infinité potentielle d'expériences. Et cela parce que les mathématiques opèrent symboliquement : les arguments d'une fonction sont des variables muettes. Ceci se prête bien à la mise en forme d'énoncés qui portent sur des concepts tels que force, masse, ...
- Les mathématiques sont un langage commun pour tous les scientifiques, indépendamment des facteurs psychologiques, politiques, culturels, ... Elles fonctionnent donc comme une sorte de langage scientifique universel.
- Enfin, les mathématiques sont un vecteur d'innovation, elles sont fécondes : la manipulation symbolique permet de dériver de nouvelles lois. Dans notre premier exemple, nous avons « découvert » la première loi de Newton en ne nous servant que de la seconde et du calcul différentiel.

Mais l'emploi des mathématiques est-il une garantie ? Toute nouvelle relation dérivée à partir du seul calcul symbolique doit-elle automatiquement être intégrée à la théorie ? Sera-t-elle automatiquement vérifiée ? Non. Et justement parce que ce qui fait la force du calcul symbolique en fait également la faiblesse : énoncer une loi sous la forme d'une relation mathématique, c'est faire abstraction du contenu physique de chacune des variables, et c'est donc oublier que les expériences qui ont permis de formuler la loi lui confèrent un domaine de validité, en dehors duquel d'autres mécanismes physiques peuvent entrer en jeu et conduire à des résultats et des lois très différents.

Prenons un exemple : la loi de Boyle-Mariotte pour les transformations isothermes, $PV = \text{constante}$. Pour le mathématicien, cette relation définit une hyperbole dans le plan (P, V) , qui à toute valeur non-nulle de la pression associe une valeur du volume. La première restriction imposée par le physicien sera de dire que la loi n'est valide que pour les valeurs positives (mais la fonction est mathématiquement définie pour les valeurs négatives également !). Ensuite, des expériences répétées montreront que la loi n'est plus valide pour certaines valeurs de la pression : on peut observer la condensation, la solidification (voir les figures ci-dessous). Il existe donc un **domaine de validité de la loi** qui ne peut être donné que par d'autres sources, c'est-à-dire en utilisant une base empirique quelconque. Il faut, dans le cas présent, y intégrer une thermodynamique beaucoup plus complexe.

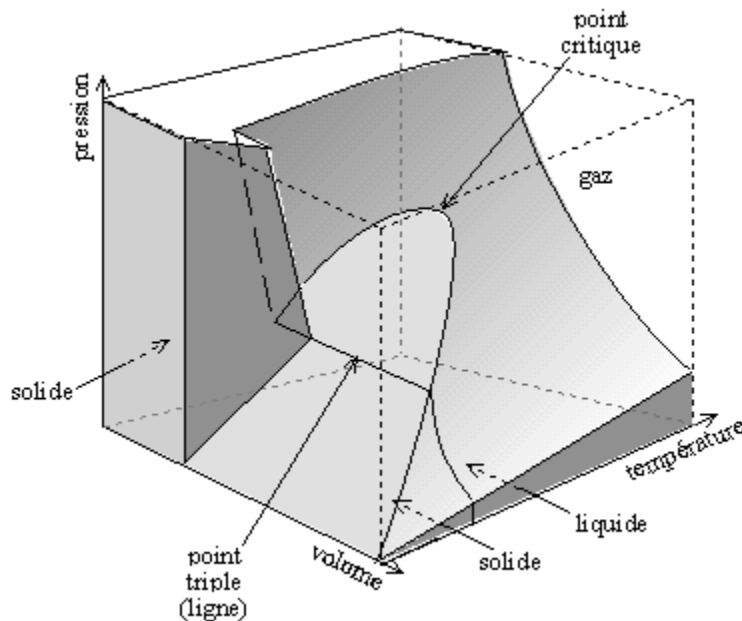


Figure 67: diagramme de changement de phase (P,V,T) d'un corps pur

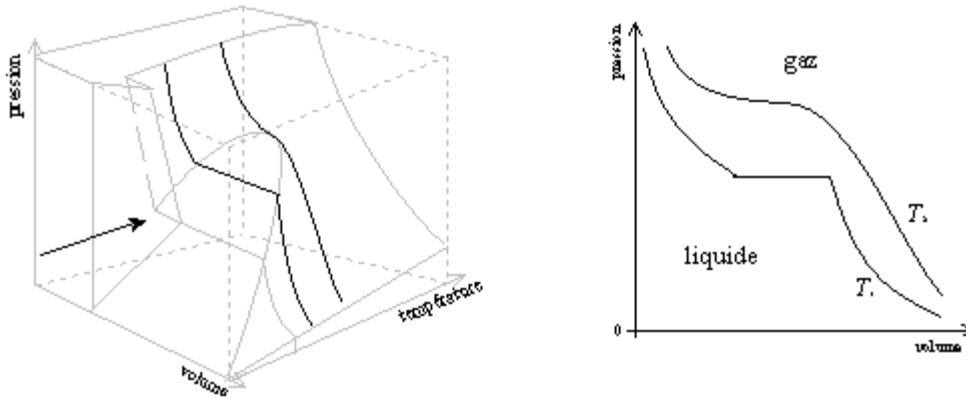


Figure 68: coupe dans le plan (P,V)

Le seul calcul symbolique ne suffit donc pas : il guide, il suggère, mais il n'a pas force de vérification. Le domaine de validité correspond aux hypothèses qui conditionnent la validité d'un théorème. De la même manière, derrière toute loi exprimée dans le cadre des sciences empiriques se cachent des hypothèses, qui définissent son domaine de validité. C'est ce que dit Claude Bernard, lorsqu'il s'attaque à ceux qu'il nomme les « systématiques », c'est-à-dire les partisans inconditionnels d'une dérivation formelle sans aucun rapport à l'expérience : *« Ces hommes partent d'une idée fondée plus ou moins sur l'observation et qu'ils considèrent comme une vérité absolue. Alors ils raisonnent logiquement et sans expérimenter, et arrivent, de conséquence en conséquence, à construire un système qui est logique, mais qui n'a aucune réalité scientifique ... Cette foi trop grande dans le raisonnement, qui conduit un physiologiste à une fausse simplification des choses, tient d'une part à l'ignorance de la science dont il parle, et d'autre part à l'absence du sentiment de complexité des phénomènes naturels. C'est pourquoi nous voyons quelquefois des mathématiciens purs, très grands esprits d'ailleurs, tomber dans les erreurs de ce genre ; ils simplifient trop et raisonnent sur les phénomènes tels qu'ils les font dans leur esprit, mais non tels qu'ils sont dans la nature. »*

La mathématisation est-elle importante, au sens de la valeur accordée à une théorie ? Oui. Pour certains, c'est même un critère majeur de scientificité.

8.3 Poser et résoudre une équation : histoire d'un long processus

Tout étudiant en science « dure » est rôdé au calcul, et a reçu un enseignement conséquent en mathématiques. Avec l'habitude, il nous semble naturel d'employer les mathématiques pour exprimer les énoncés scientifiques et résoudre des problèmes. Mais, outre l'évolution des concepts de la physique et de la mécanique, la formation des théories telles que nous les connaissons a nécessité l'évolution concomitante des mathématiques, d'ailleurs souvent aiguillonnée par les besoins de la mécanique et de la physique. Sans les outils mathématiques requis, la plupart des expressions que nous manipulons aujourd'hui seraient difficilement formulables. Revenons sur la loi de Boyle-Mariotte, dont l'expression semble mathématiquement triviale. Comme nous l'avons déjà dit,

$$P \times V = \text{Constante}$$

est un cas particulier de fonction hyperbolique, que l'on peut écrire sous la forme

$$y = f(x) = \frac{a}{x}$$

où a est une constante, x une variable muette qui est dans notre cas la pression ou le volume, et y la seconde variable, exprimée comme une fonction dépendante de la première. A première vue, pour tout bachelier, rien que de simple et d'ordinaire. Nous n'utilisons ici que des objets et des concepts mathématiques très simples : une inconnue, une fonction, une division (notée classiquement), un signe égal, une constante notée par une lettre. En réalité, il a fallu des millénaires pour conceptualiser ces objets et unifier leurs notations ! Ce n'est que très récemment, au vu de la longue histoire des mathématiques, qu'il est devenu possible d'écrire la loi de Boyle-Mariotte ou les lois de Newton !

Nous allons passer en revue quelques grandes avancées (en nous limitant aux notions les plus simples vues dans l'enseignement secondaire).

8.3.1 Une brève histoire du calcul

- **XVIII^e siècle avant notre ère** : les savants mésopotamiens savent résoudre des équations du second degré, ainsi que certaines équations du troisième et du quatrième degré. Ils savent également résoudre des systèmes de deux équations à deux inconnues. Plus précisément, ils connaissaient les solutions de certains problèmes, mais ne possédaient pas de méthodes générales.
- **XVI^e siècle avant notre ère** : les égyptiens connaissent une algèbre rudimentaire.
- **V^e siècle avant notre ère** : époque de Pythagore en Grèce. Découverte de l'irrationalité de la racine de 2 ; étude du pair et de l'impair ; théorie des nombres figurés (nombres entiers représentables sous forme géométrique).
- **III^e siècle avant notre ère** :
 - Archimède : Extension de la numération grecque aux très grands nombres ; démonstration que $3 + 10/71 < \pi < 3 + 10/70$; algorithme d'Archimède permettant de calculer π avec une précision donnée
 - Euclide : fondements de la géométrie classique ; étude des nombres pairs et impairs; étude des nombres premiers ; puissance entière des fractions
- **III^e siècle (de notre ère)** : Diophante d'Alexandrie publie son *Arithmetica* : pour la première fois dans l'histoire des mathématiques grecques, les fractions sont reconnues comme des nombres ; traitement systématique des équations du premier degré et de degré supérieur ; problèmes d'analyse indéterminée, qui préfigurent l'algèbre moderne.
- **IV^e-V^e siècles** : découverte du zéro et de la numération décimale de position d'origine indienne. Cette nouveauté permettra l'essor de l'idée généralisatrice du nombre et de l'algèbre
- **628** : le mathématicien et astronome indien Brahmagupta publie son *Brahmasphutasiddhanta* : apparition d'une parfaite maîtrise de la numération

décimale de position au moyen des neuf chiffres et du zéro ; méthodes de calcul très proches de méthodes actuelles ; introduction de règles algébriques fondamentales – nombres positifs et négatifs, théorisation du zéro comme « nombre nul » ; infini mathématique défini comme l'inverse du zéro

- **629** : le mathématicien indien Bhâskara publie son *Commentaire de l'Âryabhatîya* : utilisation parfaite du zéro, de la numération décimale de position, de la règle de trois ; calculs sur les fractions arithmétiques, notées sous forme moderne (mais sans la barre de fraction, qui ne sera introduite par les mathématiciens arabes que quelques siècles plus tard)
- **820-850** : époque du mathématicien et astronome musulman Al Khuwarizmi, dont les œuvres contribueront largement à la propagation des chiffres indiens et des méthodes de calcul et de l'algèbre indiennes en pays d'Islam et dans l'Europe chrétienne. Son nom est bien connu (inconsciemment) du grand public : latinisé en *algorismus*, puis francisé en *algorisme*, il donnera naissance à *algorithme*. Le titre de son ouvrage *Al jabr wa'l muqabala* (« transposition et réduction » : transposition d'un terme négatif d'un membre à l'autre d'une équation et réduction des termes semblables) donnera naissance au mot *algèbre*.
- **X^e siècle** :
 - Abu Kamil : continuation des travaux de Al Khuwarizmi ; développements algébriques qui serviront de base à Fibonacci en 1202
 - Al Karaji : travaux sur l'arithmétique des fractions ; sur la base des travaux de Diophante et d'Abu Kamil, travaux sur les équations de degré 2 et 2n ; maniement souple des nombres irrationnels, qui initie l'évolution qui conduira à l'élimination de l'algèbre géométrique d'origine grecque grâce au développement du symbolisme mathématique.
- **1202** : le mathématicien italien Leonard de Pise (surnommé Fibonacci) publie son *Liber Abaci* (« Traité de l'abaque ») où il introduit dans le monde occidental l'art du calcul au moyen des neuf chiffres d'origine indienne, après des voyages en Afrique du nord et au Proche-Orient.
- **XV^e siècle** :
 - Travaux du mathématicien persan Ghiyat ad din Ghamshid al Kashi : algèbre, formule du binôme, calcul des fractions décimales, puissances d'exposant entier, racines n-ièmes, nombres irrationnels, ...
 - **1484** : Nicolas Chuquet utilise les méthodes indiennes et le zéro, et introduit pour la première fois en Europe occidentale les nombres négatifs ; proposition d'une terminologie parfaitement élaborée pour désigner les puissances de 10 ; introduction d'une notation exponentielle un peu semblable à la forme moderne qui comprend des exposants négatifs et positifs
 - **1489** : le mathématicien allemand Johann Widmann d'Eger introduit les signes + et – en place des lettres p (initiale de piu – plus) et m (pour minus – moins)
- **XVI^e siècle** :
 - Le mathématicien allemand Rudolff introduit le symbole de la racine carrée $\sqrt{\quad}$, qui constitue une abréviation de l'initiale du nom latin du radical, R

- **1545-1560** : les mathématiciens italiens Gerolamo Cardano et Raffaele Bombelli introduisent les « nombres imaginaires », appelés aujourd'hui nombres complexes.
- **1557** : le mathématicien anglais Robert Recorde introduit le symbole de l'égalité =
- **1582** : le mathématicien belge Simon Stévin publie son *De Thiende*, qui est le premier ouvrage européen consacré à la théorie générale des fractions décimales (qui sont connues des arabes depuis le temps d'Al Uqlidisi en 952 et des européens depuis Emmanuel Bonfils de Tarascon en 1350) ; il est le premier à les substituer aux fractions ordinaires ; pas décisif vers la notation actuelle des nombres décimaux en symbolisant l'ordre des unités décimales : il écrit 679 (0) 5(1) 6(2) pour la forme moderne 679,56
- **1592** : le mathématicien italien Magini introduit la notation anglo-saxonne actuelle du point décimal : 679.56
- **1594** : le mathématicien écossais John Napier of Merchiston (dit Néper) découvre la correspondance entre les termes d'une progression géométrique et d'une progression arithmétique, qui sera à l'origine des logarithmes.
- **XVII^e siècle**
 - **1608** : Le mathématicien néerlandais Willebrord Snellius introduit la notation décimale à virgule : 679,56
 - **1614** : Néper élabore la théorie des logarithmes (il introduit ce terme) et publie la première table de logarithme (base e)
 - **1615** : le mathématicien anglais Henry Briggs introduit les logarithmes décimaux et publie une table pour les 31 000 premiers nombres entiers avec 14 décimales !
 - **1631** : le mathématicien anglais Thomas Hariot introduit les symboles de l'inégalité : < et >
 - **1632** : le mathématicien anglais William Oughtred introduit le symbole de la multiplication : x
 - **1637** : René Descartes invente la géométrie analytique. Il crée la notation algébrique moderne et introduit la notation exponentielle moderne pour les exposants positifs.
 - **1640** : Le mathématicien français Pierre de Fermat établit les bases de l'analyse combinatoire
 - **1654** : en collaboration avec Pierre de Fermat et Georges de Méré, Blaise Pascal établit les principes du calcul des probabilités
 - **1656** : le mathématicien anglais John Wallis étend la notation exponentielle de Descartes aux exposants négatifs ou fractionnaires
 - **1670-1684** : Isaac Newton et Gottfried Wilhelm Leibniz établissent les fondements du calcul infinitésimal moderne
 - **1672** : Leibniz invente le calcul différentiel et intégral
 - **1696** : publication du premier traité classique de calcul intégral par le marquis de l'Hospital : *Analyse des infiniment petits*.
- **XVIII^e siècle** :
 - **1748** : Leonhard Euler publie un livre traitant de l'étude générale des fonctions : fonctions exponentielles, logarithmiques, trigonométriques ; développements en série ou en produit infinis

- **1750** : première étude approfondie des systèmes d'équations par le mathématicien suisse Gabriel Cramer
- **1770** : le mathématicien français Alexandre Vandermonde invente la théorie des substitutions
- **1797** : le mathématicien français Louis Lagrange crée la théorie des fonctions analytiques, qui fonde l'analyse mathématique moderne.
- **XIX^e siècle** :
 - **1807-1822** : Joseph Fourier étudie les séries trigonométriques (séries de Fourier)
 - **1821-1893** : Augustin Cauchy démontre l'existence des solutions des équations différentielles et aux dérivées partielles. La théorie des équations différentielles sera ensuite développée par Lipschitz et Picard.
 - **1828** : le mathématicien irlandais William Rowan Hamilton donne la première définition rigoureuse des nombres complexes.
 - **1843-1844** : W.R. Hamilton et le mathématicien allemand Grassmann publient les premiers travaux faisant apparaître le calcul vectoriel
 - **1858** : l'avocat et mathématicien Arthur Cayley fonde le calcul matriciel

8.3.2 Une « invention » révolutionnaire : l'inconnue

Considérons l'énoncé suivant :

Trouver x tel que $f(x)=0$

C'est la forme classique que prend l'énoncé d'un problème en science : il consiste à chercher tous les éléments x qui satisfont au critère retenu (des valeurs dans le cas où l'on cherche des nombres). Ici, le symbole x représente toutes les solutions possibles, c'est-à-dire quelque chose que l'on ne connaît pas. Les manipulations formelles qui permettent la résolution du problème ont pour objectif d'arriver à écrire une ligne du type $x= \dots$, c'est-à-dire de déterminer x . La puissance de la plupart de nos méthodes de résolution vient de notre capacité à manipuler l'inconnue, c'est-à-dire à la faire intervenir dans des équations qui incluent des symboles associés à des choses connues (fonctions, opérations, paramètres). Cela semble simple et naturel, alors que c'est loin d'avoir toujours été le cas : l'« invention » de l'inconnue est récente dans l'histoire des mathématiques.

En effet, symboliser ce que l'on ne connaît pas, donc nommer l'inconnue, est contraire à la philosophie platonicienne, dont l'influence est très grande jusqu'à la Renaissance. En effet, Platon donne cinq facteurs de la connaissance (*Lettre VII, 342a-e*) :

La connaissance selon Platon

1. le nom
2. la définition
3. la représentation ou figure

4. la connaissance en elle-même
5. la chose en elle-même

Platon illustre cela par l'exemple du cercle. On peut distinguer le nom « cercle », la définition du cercle (l'ensemble des points situés à une même distance du centre), la figure que l'on dessine, la connaissance du cercle qui est une représentation intellectuelle et enfin la chose, c'est-à-dire l'essence du cercle en soi.

La théorie platonicienne assimile ce qui est connu à ce qui est nommé, défini et figuré. Dans cette perspective, il est impossible de connaître ce qui ne peut être tracé, représenté matériellement. Et inversement, on ne peut concevoir de figurer ce que l'on ne connaît pas. L'inconnue, telle que nous la manipulons, n'est donc pas désignable ou même imaginable. Il faudra attendre les mathématiciens de la Renaissance pour que lentement émerge l'inconnue.

Mais il est important de noter que la difficulté à conceptualiser l'inconnue n'est pas une spécificité de la philosophie platonicienne : elle est également présente dans des civilisations beaucoup plus anciennes, qui avaient pourtant développé des éléments de calcul. La raison majeure est que le calcul abstrait, tel que nous pratiquons, n'était pas connu : tout calcul se rapportait à des choses matérielles (des marchandises, des champs, du bétail, ...) et avait une fonction bien précise : calcul d'impôt, cadastre, ... Il reste encore des traces de cela dans notre langage : le mot calcul est dérivé du latin **calculus**, qui veut dire « caillou », mais aussi « boule, jeton, pion », et qui fait référence aux méthodes initiales de calcul par dénombrement de cailloux qui datent de l'époque préhistorique. De la même manière, notre inconnue moderne est désignée par les mots « masse, tas, monceau, quantité » par le scribe égyptien Ahmes en 1550 avant notre ère. Le mathématicien grec Diophante y fait référence en employant l'expression « un nombre d'unités indéterminé » et introduit un symbolisme algébrique élémentaire en désignant l'inconnue d'un problème par un signe (la première syllabe du mot grec **arithmos**, qui signifie « le nombre »). Le mathématicien indien Brahmagupta utilise lui l'expression « autant que ». Les mathématiciens arabes et occidentaux médiévaux utilisent le terme « chose » (**res** en latin), ainsi que les savants chinois. On le voit, dans tous les cas, le terme fait référence à quelque chose de matérialisable.

La première forme « moderne » (dans le concept au moins) apparaît en 1487 dans un ouvrage du mathématicien italien Lucas Pacioli qui décrit de nombreuses méthodes de résolution des équations algébriques et les règles de calcul indiennes : il emploie le terme **cosa** (« la chose ») et parfois le terme **quantita** pour une inconnue secondaire. Ce n'est qu'en 1591 que le mathématicien français François Viète introduit une expression littérale dans les expressions algébriques : les inconnues sont désignées par des voyelles et les constantes indéterminées par des consonnes. La notation moderne, où les données sont représentées par les premières lettres de l'alphabet et les inconnues par les dernières (x, y, z) est introduite en 1637 par René Descartes.

9 La question du progrès scientifique

« Nous sommes comme des nains juchés sur des épaules de géants. Nous voyons ainsi davantage et plus loin qu'eux, non parce que notre vue est plus aiguë ou notre taille plus haute, mais parce qu'ils nous portent en l'air et nous élèvent de toute leur hauteur gigantesque. » Cette phrase, attribuée⁸⁷ à Bernard de Chartes, un des maîtres de la scolastique au XII^e siècle, exprime la continuité de l'activité intellectuelle et donne une vision cumulative de l'évolution du savoir. Cette même image est reprise des siècles plus tard par Isaac Newton dans une lettre à Robert Hooke datant de 1676 : *« si j'ai pu voir aussi loin, c'est parce que j'étais juché sur les épaules de géants. »*

L'idée de « progrès de la science », ou plutôt celle qu'il existe un progrès scientifique continu depuis plusieurs siècles est profondément ancrée dans le grand public. Elle est généralement associée à l'évolution technologique, particulièrement frappante depuis la seconde moitié du XX^e siècle. Mais qu'est-ce que le progrès scientifique ? Que la science change, c'est une évidence : la physique et la mécanique d'aujourd'hui sont différentes de celles connues au temps de Newton ; les bases empiriques ont été considérablement accrues, tant en volume qu'en précision des données, les théories et les modèles ont considérablement évolué et de très nombreuses nouvelles disciplines sont apparues. Cette évolution des sciences empiriques (mais les sciences formelles ont elles aussi connu des développements spectaculaires) appelle de nombreuses questions. Le contenu de ce chapitre va mettre l'accent sur les trois questions suivantes :

- Y a-t-il progrès scientifique, au sens où les sciences modernes sont supérieures (et en quoi ?) à celles des siècles passés ? Quel est donc le sens exact du terme progrès lorsqu'il s'agit des sciences ?
- Comment les sciences évoluent-elles ? De manière continue, ou au contraire subissent-elles des grandes ruptures, des révolutions scientifiques ?
- Quels sont les facteurs qui déterminent cette évolution ? Le développement scientifique est-il influencé par des facteurs extérieurs à la science, comme des contraintes sociales ou sociologiques, économiques, religieuses, politiques ou philosophiques ?

Autant de questions qui sont nécessaires pour comprendre la science en tant que processus historique intégré à l'histoire humaine. Nous allons tout d'abord nous intéresser aux facteurs qui, selon les épistémologues, peuvent intervenir dans le processus d'évolution des sciences empiriques. Puis, dans un second temps, nous verrons comment les grands courants théoriques dans le domaine de l'épistémologie conçoivent la dynamique du changement des sciences (et donc établissent une hiérarchie entre les différents facteurs de changement imaginables).

⁸⁷ Cité par M. Rouche, « Histoire de l'enseignement et de l'éducation », 2 tomes, tempus

9.1 Contraintes sur l'acceptation des théories

9.1.1 Les types de contrainte

Puisque l'activité scientifique est une activité publique et humaine, elle s'effectue dans un contexte qui exerce des contraintes sur elle. En conséquence, il est légitime de se poser des questions sur la nature de ces contraintes et sur leur influence.

Quelles sont ces contraintes ? On peut tout d'abord remarquer qu'elles s'exercent à plusieurs niveaux : aux **contraintes individuelles** qui s'exercent sur un individu, s'ajoutent les **contraintes collectives** (qui ne jouent que pour un groupe d'individus) et les **contraintes universelles**, valables pour tout individu.

On peut classer les contraintes suivant plusieurs critères :

- Le type de l'origine des contraintes. On peut citer l'objet de l'étude (terme qui regroupe les données sensibles, les méthodes, le monde extérieur, ...), la personnalité de l'individu, la structuration et le fonctionnement d'une communauté scientifique (organisation en laboratoires, en équipes, modes de dissémination du savoir – articles, livres, colloques, congrès -, recrutement et promotion, notions et valeurs dominantes imposées : vérité, efficacité, ...)
- Le type de déterminisme associé aux contraintes. On distingue tout d'abord ici les **contraintes contingentes** et les **contraintes nécessaires**. La contrainte est nécessaire si elle est nécessairement imposée par la source. Si au contraire l'existence de la source n'implique pas obligatoirement celle de la contrainte, cette dernière est dite contingente. Dans un second temps, on distingue les **contraintes conditionnantes** et les **contraintes absolument déterminantes**. Les contraintes conditionnantes sont celles qui ne font que participer au résultat, sans l'imposer complètement : d'autres effets auraient pu être observés. A l'inverse, les contraintes absolument déterminantes imposent de manière stricte le résultat, qui n'aurait pas pu être différent.
- La valeur conférée aux contraintes. On classe ici les contraintes suivant un jugement sur leur valeur : la contrainte est-elle bonne ou mauvaise ? Légitime ou illégitime ? Cette appréciation subjective dépend bien entendu des croyances de celui qui l'émet. Toutefois, les communautés scientifiques actuelles tendent à s'accorder autour des jugements suivants :
 - Sont valorisées : les contraintes universelles, les contraintes nécessaires
 - Sont dévalorisées : les contraintes contingentes, les contraintes psychologiques et sociales.

On le voit, le concept de science retenu aujourd'hui valorise fortement la référence à un monde extérieur autonome (réalisme), qui est perçu comme une garantie d'objectivité, ou au moins comme un moyen de se prévenir contre le développement de théories qui ne soient que de pures fictions. Tout ce qui provient du caractère humain de l'activité scientifique ou encore des aléas historiques est jugé négativement, car généralement considéré comme subjectif et contingent.

9.1.2 Une évolution strictement objective ? L'épistémologie naïve

L'évolution des sciences ne se ferait-elle que sous la pression de contraintes objectives, c'est-à-dire nécessaires et universelles ? C'est la thèse connue sous le nom d'**épistémologie naïve**. Il s'agit là d'un cas d'école, d'une position extrême qui aujourd'hui n'a plus de défenseur. Il est toutefois intéressant, car il correspond à une certaine vision de la science assez répandue dans le grand public.

Le postulat de base de l'épistémologie naïve est que la science, grâce aux méthodes qui sont les siennes, possède un moyen infallible de retenir les meilleures théories.

Notons ici que l'adjectif meilleur renvoie à des critères différents suivant les épistémologues : à la vérité-correspondance pour un réaliste, à l'efficacité pour un antiréaliste. Pour l'épistémologie naïve, ne sont donc légitimes que les contraintes nécessaires et universelles. Les contraintes objectives entrent dans ces catégories, mais on peut aussi y ranger les contraintes subjectives universelles (celles qui sont liées à l'individu en tant qu'humain : limites physiologiques, ...)

Cette thèse est bien évidemment en contradiction avec les enseignements de l'histoire et de la sociologie des sciences, qui démontrent que de nombreuses autres contraintes ont agi et agissent encore sur le développement des disciplines scientifiques. En effet, pour l'épistémologie naïve, puisque seules les contraintes nécessaires et universelles influent sur les sciences, les contraintes sociales ne jouent pas et la sociologie des sciences ne peut que décrire le contexte au sein duquel se développe une discipline, sans permettre d'établir un lien entre les contraintes sociales et le contenu des énoncés scientifiques. Donc, pour l'épistémologie naïve, le contexte social est neutre.

9.1.3 Une évolution partiellement subjective ? Internalisme et externalisme

Venons en maintenant aux épistémologies qui considèrent que l'évolution des disciplines scientifiques n'est pas purement objective. L'**internalisme** et l'**externalisme** sont d'abord des méthodes de description appartenant à l'histoire des sciences : l'internalisme vise à décrire l'évolution des sciences en mettant l'accent sur les facteurs internes au champ scientifique, alors que l'externalisme consiste à faire porter l'analyse sur les facteurs externes au champ scientifique. Cette dichotomie est à rapprocher de celle déjà rencontrée au chapitre 1 lors de la discussion sur la théorie de la connaissance.

Mais, décrire l'évolution d'une discipline en choisissant de mettre en lumière les facteurs internes et les facteurs externes, c'est implicitement accorder plus d'importance à l'un ou l'autre type de contraintes. C'est donc également faire des hypothèses sur les contraintes les plus importantes qui influent sur l'évolution des sciences.

Pour les défenseurs de l'internalisme, ce sont les facteurs internes qui guident l'évolution des sciences. Les facteurs extra scientifiques (économiques, sociologiques, religieux, ...) n'ont alors pas d'importance ou n'ont qu'un rôle très mineur. Dans sa version la plus extrémiste, l'internalisme est donc très proche de l'épistémologie naïve. Les partisans de l'externalisme, quant à eux, donnent une importance réelle

aux facteurs externes. Dans sa version la plus extrême, l'externalisme fait de la science le simple produit des contraintes économiques, sociales, religieuses et politiques.

Entre les deux extrêmes que constituent l'internalisme pur et l'externalisme radical se situe la thèse aujourd'hui la plus répandue, dite **thèse faible**, selon laquelle c'est la somme des contraintes externes et internes qui fait avancer les sciences. Elle est dite faible, car elle soutient que les facteurs externes ne sont que des contraintes conditionnantes, et donc que le contenu des théories scientifiques n'est pas réductible aux facteurs sociaux. Cette thèse faible a été défendue, entre autres, par Alexandre Koyré, Thomas Kuhn et Gaston Bachelard.

9.1.4 La notion d'obstacle épistémologique selon Bachelard

Gaston Bachelard a beaucoup travaillé sur le problème du progrès de la connaissance.

Il a développé une philosophie originale, communément appelée « *la philosophie du non* ». Cette appellation fait référence à l'idée de Bachelard selon laquelle « *Deux hommes, s'ils veulent s'entendre vraiment, ont dû d'abord se contredire. La vérité est fille de la discussion, non pas fille de la sympathie.* » L'évolution scientifique se fait donc par rupture avec les théories admises antérieurement : « *On connaît contre une connaissance antérieure.* » Il rejoint ici le philosophe français Alain (de son vrai nom Emile Chartier, 1868-1951) qui écrit : « *Penser, c'est dire non.* » Partant de ce principe, Bachelard identifie les facteurs qui vont à l'encontre de l'évolution scientifique, c'est-à-dire ceux qui retardent ou bloquent le développement de nouvelles théories : ce sont les **obstacles épistémologiques**.

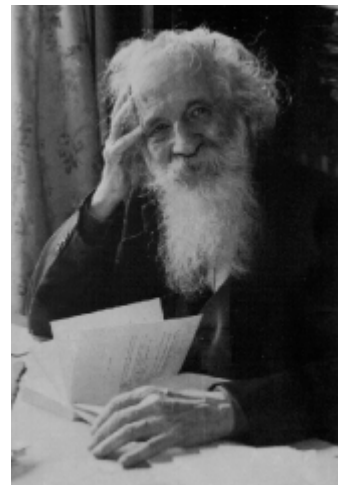


Figure 69: G. Bachelard

« Quand on cherche les conditions psychologiques des progrès de la science, on arrive bientôt à cette conviction que c'est en termes d'obstacles qu'il faut poser le problème de la connaissance scientifique. Et il ne s'agit pas de considérer des obstacles externes comme la complexité et la fugacité des phénomènes, ni d'incriminer la faiblesse des sens et de l'esprit humain : c'est dans l'acte même de connaître, intimement, qu'apparaissent, par une sorte de nécessité fonctionnelle, des lenteurs et des troubles. C'est là que nous montrerons des causes de stagnation et même de régression, c'est là que nous décèlerons des causes d'inertie que nous appellerons des obstacles épistémologiques. La connaissance du réel est une lumière qui projette toujours quelque part des ombres. Elle n'est jamais immédiate et pleine. Les révélations du réel sont toujours récurrentes. Le réel n'est jamais « ce qu'on pourrait croire », mais il est toujours ce qu'on aurait dû penser. La pensée empirique est claire après coup, quand l'appareil des raisons a été mis au point. En revenant sur un passé d'erreurs, on trouve la vérité en un véritable repentir intellectuel. En fait, on connaît contre une connaissance antérieure, en détruisant des connaissances mal faites, en surmontant ce qui, dans l'esprit même fait obstacle à la spiritualisation.

L'idée de partir de zéro pour fonder et accroître son bien ne peut venir que dans des cultures de simple juxtaposition où un fait connu est immédiatement une richesse. Mais devant le mystère du réel, l'âme ne peut se faire, par décret, ingénue. Il est alors impossible de faire d'un seul coup table rase des connaissances usuelles. Face au réel, ce qu'on croit savoir offusque ce qu'on devrait savoir. Quand il se présente à la culture scientifique, l'esprit n'est jamais jeune. Il est même très vieux, car il a l'âge de ses préjugés. Accéder à la science, c'est, spirituellement, rajeunir, c'est accepter une mutation brusque qui doit contredire un passé.

La science, dans son besoin d'achèvement comme dans son principe, s'oppose absolument à l'opinion. S'il lui arrive, sur un point particulier, de légitimer l'opinion, c'est pour d'autres raisons que celles qui fondent l'opinion ; de sorte que l'opinion a, en droit, toujours tort. L'opinion pense mal ; elle ne pense pas ; elle traduit des besoins en connaissances. En désignant les objets par leur utilité, elle s'interdit de les connaître. On ne peut rien fonder sur l'opinion : il faut d'abord la détruire. Elle est le premier obstacle à surmonter. Il ne suffirait pas, par exemple, de la rectifier en des points particuliers, en maintenant, comme une sorte de morale provisoire, une connaissance vulgaire provisoire. L'esprit scientifique nous interdit d'avoir une opinion sur des questions que nous ne comprenons pas, sur des questions que nous ne savons pas formuler clairement. Avant tout, il faut savoir poser des problèmes. Et quoi qu'on dise, dans la vie scientifique, les problèmes ne se posent pas d'eux-mêmes. C'est précisément ce sens du problème qui donne la marque du véritable esprit scientifique. Pour un esprit scientifique, toute connaissance est une réponse à une question. S'il n'y a pas eu de question, il ne peut y avoir connaissance scientifique. Rien ne va de soi. Rien n'est donné. Tout est construit.

Une connaissance acquise par un effort scientifique peut elle-même décliner. La question abstraite et franche s'use : la réponse concrète reste. Dès lors, l'activité spirituelle s'invertit et se bloque. Un obstacle épistémologique s'incruste sur la connaissance non questionnée. Des habitudes intellectuelles qui furent utiles et saines peuvent, à la longue, entraver la recherche. « Notre esprit, dit justement M. Bergson, a une irrésistible tendance à considérer comme plus claire l'idée qui lui sert le plus souvent ». L'idée gagne ainsi une clarté intrinsèque abusive. A l'usage, les idées se valorisent indûment. Une valeur en soi s'oppose à la circulation des valeurs. C'est un facteur d'inertie pour l'esprit. Parfois une idée dominante polarise un esprit dans sa totalité. Un épistémologue irrévérencieux disait, il y a quelques vingt ans, que les grands hommes sont utiles à la science dans la première moitié de leur vie, nuisibles dans la seconde moitié. L'instinct formatif est si persistant chez certains hommes de pensée qu'on ne doit pas s'alarmer de cette boutade. Mais enfin l'instinct formatif finit par céder devant l'instinct conservatif. Il vient un temps où l'esprit aime mieux ce qui confirme son savoir que ce qui le contredit, où il aime mieux les réponses que les questions. Alors l'instinct conservatif domine, la croissance spirituelle s'arrête. »

Pour Bachelard, la notion d'obstacle épistémologique recouvrent les contraintes non-techniques, c'est-à-dire les contraintes psychologiques et culturelles.

Plus particulièrement, il identifie :

- **L'expérience immédiate**, c'est-à-dire l'expérience menée sans cadre théorique digne de ce nom pour l'interpréter. Il fait référence ici aux

expériences purement ludiques réalisées pour le grand public ou les élèves, qui donne toute la place au spectaculaire sans donner d'information sur les éléments d'interprétation scientifique. Bachelard cite l'exemple suivant : « *En 1748, Franklin et ses amis tuèrent un dindon par la commotion électrique, le firent rôtir avec un tournebroche électrique, devant un feu allumé par la bouteille électrique⁸⁸ : ensuite ils burent à la santé de tous les électriciens célèbres d'Angleterre, de Hollande, de France et d'Allemagne, dans des verres électrisés, et au bruit d'une décharge d'une batterie électrique.* »

- **La connaissance générale**, qui consiste à généraliser de manière inadéquate un concept, menant à un usage inapproprié et au masquage des vrais éléments d'interprétation. Bachelard illustre cela en parlant de la manière dont la théorie des corps flottants est enseignée : « *L'équilibre des corps flottants fait l'objet d'une intuition familière qui est un tissu d'erreurs. D'une manière plus ou moins nette, on attribue une activité au corps qui flotte, mieux au corps qui nage. Si l'on essaie avec la main d'enfoncer un morceau de bois dans l'eau, il résiste. On n'attribue pas facilement la résistance à l'eau. Il est dès lors assez difficile de faire comprendre le principe d'Archimède dans son étonnante simplicité mathématique si l'on a pas d'abord critiqué et désorganisé le complexe des intuitions premières. En particulier sans cette psychanalyse des erreurs initiales, on ne fera jamais comprendre que le corps qui émerge et le corps complètement immergé obéissent à la même loi.* »
- **La connaissance pragmatique**, qui consiste à vouloir expliquer les phénomènes par leur utilité au sein d'un monde gouverné par un principe supérieur. Bachelard cite ici Voltaire, lorsqu'il raille Newton : « *Loin que les comètes soient dangereuses ... elles sont, selon (Newton), de nouveaux bienfaits du Créateur... (Newton) soupçonne que les vapeurs qui sortent d'elles sont attirées dans les orbites des planètes, et servent à renouveler l'humidité de ces globes terrestres qui diminue toujours. Il pense encore que la partie la plus élastique et la plus subtile de l'air que nous respirons nous vient des comètes... Il me semble que c'est deviner en sage, et que si c'est se tromper, c'est se tromper en grand homme.* »
- **L'obstacle substantialiste**, qui consiste à chercher une substance comme support de tout phénomène. Par exemple, au XVIII^e siècle, on a tenté d'expliquer les propriétés du fer aimanté par l'existence d'une sorte de colle, le « flegme », qu'il contiendrait. Et on a cru voir ce flegme lorsque le fer incandescent est trempé dans l'eau.
- **L'obstacle animiste**, qui consiste à attribuer aux objets inanimés des propriétés des êtres vivants : maladie, vieillesse, mort ou encore des intentions ou des sentiments. Ainsi, au XVIII^e siècle, la rouille était considérée comme une maladie du fer, qui pouvait être guérie ! Autre exemple : à cette époque, lorsqu'on fermait une mine, on y semait de la limaille de fer, pour féconder la terre et faire germer de nouveaux minerais !
- **L'obstacle psychanalytique**. Selon Bachelard, l'inconscient a une influence sur la réflexion, et suggère des métaphores, notamment à connotation sexuelle.
- Enfin, on peut citer **l'obstacle de l'autorité** : la soumission aux idées reçues pour la seule raison qu'elles sont enseignées ou défendues par des scientifiques connus est dangereuse. Ce que traduit Bachelard par : « *les*

⁸⁸ Il s'agit de la bouteille de Leyde, qui est l'ancêtre de la pile électrique.

grands hommes sont utiles à la science dans la première moitié de leur vie, nuisibles dans la seconde moitié. » Il rejoint en cela Cicéron, qui écrivait : « *A ceux qui veulent apprendre, c'est souvent une gêne que l'autorité de ceux qui leur donnent leur enseignement.* »

9.1.5 Un cas d'école : l'affaire Lyssenko

Un exemple célèbre d'influence externe sur l'évolution scientifique est fourni par « *l'affaire Lyssenko* », qui se déroula en Union soviétique. C'est le cas le plus spectaculaire de mystification politique qui ait affecté la communauté scientifique au XX^e siècle. Trofim Lyssenko commença sa carrière de biologiste en 1926 en Ukraine, où il travaillait sur la vernalisation du blé pour accroître les récoltes. Après avoir obtenu un soutien fort du pouvoir politique (Staline le soutient publiquement), il voit ses théories sur la culture du blé appliquées à grande échelle en 1935. Dès 1936, il attaque les théories « bourgeoises » et remet en cause le rôle des chromosomes et des gènes dans l'hérédité, et esquisse une nouvelle théorie qu'il juge plus conforme au matérialisme dialectique et au marxisme. L'intérêt du pouvoir politique est double : en soutenant Lyssenko, il tente de briser la résistance des paysans au collectivisme forcé et celle des intellectuels qui refusaient l'embrigadement idéologique. En 1937, la plupart des adversaires de Lyssenko sont arrêtés et fusillés. En 1938, le congrès mondial de génétique prévu à Moscou est annulé. Lyssenko reprend l'offensive dès la fin de la seconde guerre mondiale. La génétique est officiellement bannie d'Union soviétique en 1948 au profit d'une « science prolétarienne de l'hérédité », qui nie la réalité du gène et du chromosome. La génétique est présentée comme une science « *bourgeoise* » fondée par « *le moine autrichien Gregor Mendel* ». Puisque les nazis avaient exploité la théorie de l'hérédité, et que celle-ci avait également permis au biologiste américain Thomas Morgan d'introduire la notion de mutation en 1910, la génétique est présentée comme une pensée « *cléricale, nazie-mystique et américaine* ». La théorie de Lyssenko est fondée sur l'idée « *la concurrence n'a lieu qu'entre espèces : le loup mange le lièvre, les lièvres ne se mangent pas entre eux : le lièvre broute l'herbe. Le froment n'étouffe pas le froment* ». Il faut donc planter les cultures « par nids » et attendre que les pousses les moins fortes se sacrifient pour le bien de l'espèce. Ceci conduira à un désastre agricole à partir de 1949-1950. Mais la carrière de Lyssenko ne s'arrêtera pas là : après celui de Staline, il recevra le soutien de Khrouchtchev et le lyssenkisme restera la doctrine officielle en Union soviétique jusqu'en ... 1965 !

9.2 Conceptions du progrès scientifique

Nous venons de passer en revue différents problèmes associés à l'interprétation de l'évolution des sciences. Mais il reste une question : cette évolution est-elle un progrès ? Répondre à cette question, c'est tout d'abord définir ce que l'on entend par progrès. Le mot progrès inclut bien la notion de changement, d'évolution. Mais il fait plus que cela, car il introduit également l'idée d'une amélioration. Pour définir le progrès, il faut donc indiquer ce qui a été amélioré.

Parmi les définitions les plus classiques du progrès scientifique, on peut noter celles qui font référence à l'efficacité prédictive (c'est celle retenue par les antiréalistes) ou

encore le nombre de vérités disponibles (pour les réalistes convergents). Mais l'amélioration comporte une part de subjectivité : dire qu'il y a eu amélioration c'est émettre un jugement. Des facteurs psychologiques entrent donc en jeu. Par exemple, beaucoup de scientifiques attendent d'une nouvelle théorie qu'elle soit plus simple, plus élégante.

Nous allons maintenant nous intéresser aux deux grandes manières d'analyser le progrès scientifique.

9.2.1 La conception téléologique du progrès

Une première manière de penser le progrès est de voir celui-ci dirigé vers une fin dernière, prédéterminée et ayant une existence en soi. C'est la **conception téléologique**, ou encore **finaliste**. Le concept de téléologie implique que c'est la fin dernière qui conditionne les étapes intermédiaires et qui rend nécessaire le progrès en le déterminant. Le progrès est donc guidé par sa fin.

Pour les réalistes convergents, cette fin est la théorie ultime, celle qui reflète parfaitement la réalité extérieure autonome. Pour les antiréalistes, la fin est la théorie unique déterminée par l'objet d'étude et la méthode scientifique.

Quelle que soit l'école de pensée considérée (réaliste ou antiréaliste), le progrès est pensé comme cumulatif et continu. Il obéit à des nécessités internes dictées par la théorie ultime, qui ne sont pas des facteurs contingents et sont absolument déterminants. Le progrès n'aurait donc pas pu être autre que ce qu'il est.

La principale objection opposée à l'interprétation finaliste est bien sûr celle que le passé ne peut pas être déterminé par le futur et que cette vision résulte d'une mise en perspective par l'historien des sciences. Ceci lui est rendu possible car il connaît le présent et le passé, alors que les scientifiques ne connaissent pas le devenir de leur discipline.

9.2.2 La conception évolutionniste du progrès

A la conception téléologique s'oppose la **conception évolutionniste** du progrès scientifique. Il s'agit là de l'emploi du paradigme de la sélection naturelle proposée par Charles Darwin (1809-1882) en biologie. Selon cette thèse, ce sont les théories victorieuses (car plus adaptées) qui sont retenues.

Le progrès est donc aveugle, en ce sens qu'il n'est pas orienté d'avance. Il ne progresse pas vers un but unique. L'idée d'une concurrence entre différentes théories disponibles au même moment permet de concevoir le progrès comme non nécessaire et non continu. L'idée de sélection introduit un facteur contingent : elle se fait en présence de contraintes externes au champ scientifique. Les théories survivantes ne sont donc pas forcément les théories optimales en terme de vérité et d'efficacité, mais celles qui satisfont un ensemble plus globale de contraintes, dont l'efficacité ou la vérité font partie.

9.3 La dynamique des sciences : continuité ou révolutions ?

Deux points de vue s'opposent dans l'interprétation de l'évolution des sciences :

- Selon le **schéma évolutif continu (thèse continuiste)**, les sciences progressent en suivant un processus cumulatif continu : les nouvelles connaissances s'ajoutent aux anciennes, donnant naissance à un processus régulier d'évolution. Cette analyse a été défendue par de nombreux philosophes et épistémologues positivistes au début du XX^e siècle, dont Pierre Duhem.
- Selon le **schéma évolutif discontinu (thèse discontinuiste)**, le développement des disciplines scientifiques fait apparaître des ruptures radicales, des discontinuités profondes, qui se traduisent par l'abandon de certains modèles ou certaines théories, qui sont remplacées par de nouvelles jugées très différentes, voire contradictoires. Cette thèse est celle aujourd'hui la plus couramment acceptée, et a été développée, entre autres, par des philosophes et des épistémologues comme Alexandre Koyré (1892-1962), Gaston Bachelard (1884-1962) et Thomas Kuhn (1922-1996).

Il est important de retenir que, ce qui oppose les défenseurs de ces deux thèses, c'est l'interprétation qui est faite de l'histoire des disciplines scientifiques : la très grande majorité des intervenants sont tout à fait d'accord sur la reconnaissance des faits, des dates, et des contributions des savants des siècles passés.

9.3.1 Evolution discontinue et révolutions scientifiques : Kuhn

La thèse de l'évolution discontinue emploie le concept de **révolution scientifique** pour représenter les ruptures qu'elle identifie dans l'évolution d'une discipline scientifique. Ce terme a été popularisé par l'épistémologue Thomas Kuhn (1922-1996) dans son ouvrage *La structure des révolutions scientifiques* (1962).

Mais qu'est-ce, au juste, qu'une révolution scientifique ? L'intuition permet d'identifier des étapes marquantes dans l'évolution des sciences, comme par exemple en astronomie le passage du modèle géocentrique (le Soleil et les planètes tournent autour de la Terre) au modèle héliocentrique (les planètes tournent autour du Soleil).

Ce passage est souvent appelé la *révolution copernicienne*, en référence aux travaux de Nicolas Copernic (1473-1543) qui y participa.

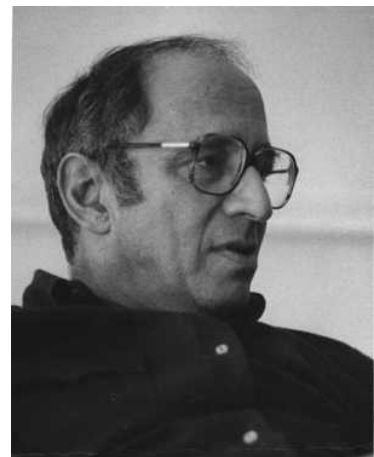


Figure 70 : T. Kuhn

Il est difficile de donner une définition plus précise, car le concept de révolution scientifique est en partie subjectif : il repose sur un jugement portant sur l'importance d'un changement. Et, comme tout jugement, il est conditionné par les croyances (au sens large) de celui qui l'émet. Il faut également retenir que l'identification du moment où intervient une révolution scientifique est également problématique. En

effet, faut-il dater la révolution du moment où une nouvelle idée, un nouveau modèle ou une nouvelle théorie est formulée pour la première fois, ou la faire remonter au moment de la prise de conscience de la communauté scientifique, c'est-à-dire au moment où cette nouveauté va être utilisée par la majorité des scientifiques qui vont reconnaître son caractère innovant ? Les deux solutions sont acceptables, et utilisées par les épistémologues. Dans le premier cas (critère de la première apparition), la révolution scientifique est considérée comme un élément objectif, la prise de conscience n'étant qu'un épiphénomène historique. Dans le second cas (critère de la prise de conscience), la révolution scientifique est avant tout un fait psychologique, donc subjectif. Mais reste encore à déterminer à qui appliquer ce critère de prise de conscience : à un seul individu ? A une communauté scientifique ? Aux historiens des sciences et aux épistémologues ? Ici aussi, le débat est ouvert.

Tout changement est-il une révolution ? Non, et c'est ici que se situe le point le plus délicat. En effet, la thèse continuiste admet aussi l'idée que les théories changent, mais sans parler de révolution. Le caractère révolutionnaire est le plus souvent attaché à la profondeur conceptuelle du changement, plutôt qu'à son « étendue ». Il peut donc y avoir des petites et des grandes révolutions : des bouleversements profonds peuvent intervenir sur des champs scientifiques très restreints, sans aucune répercussion sur les autres.

Maintenant que nous avons mieux précisé le concept de révolution scientifique, intéressons-nous à la dynamique d'évolution des sciences telle que la conçoivent les partisans de la thèse discontinuiste. Selon cette interprétation, on peut distinguer plusieurs phases dans l'évolution historique d'une science.

Tout d'abord, rappelons-nous qu'une théorie est sensée rendre compte des faits sur laquelle elle est basée. Donc, après l'élaboration d'un modèle ou d'une théorie, tous les faits connus (la base empirique) sont compatibles avec l'ensemble des énoncés théoriques et des énoncés d'observation qui s'en déduisent. Sans quoi, la théorie ou le modèle ne seraient pas retenus. Ce nouveau modèle permet aux scientifiques, pendant un certain temps, de progresser et de rendre compte de nouveaux faits, qui n'étaient pas pris en compte dans la phase d'élaboration. C'est la phase dite de **science normale**. Mais, au fur et à mesure que la base empirique croît (de nouvelles expériences sont menées), des contradictions apparaissent : certaines observations nouvelles échappent à la théorie (elle ne peut les expliquer), ou entrent en contradiction avec elle. On entre alors dans une phase de complexification de la théorie ou du modèle : pour rendre compte des nouveaux résultats expérimentaux, certaines parties anciennes sont révisées, et d'autres lui sont ajoutées. Ce processus de complexification permet de continuer à progresser, mais au prix d'efforts de plus en plus lourds : la théorie est de plus en plus compliquée, son emploi est de moins en moins aisé. De plus, il est de plus en plus difficile d'assurer son unité et sa cohérence : les différents ajouts peuvent parfois induire des conséquences contradictoires. La théorie n'est alors plus vraie, au sens de la vérité-cohérence. Cette phase est appelée la **conjecture**. Lorsqu'elle est devenue trop compliquée d'emploi, ou lorsqu'elle a atteint un degré de contradiction interne jugé intolérable, ou encore que l'on n'arrive plus à la modifier de manière satisfaisante pour rendre compte de nouveaux faits, la théorie entre dans une phase de crise, qui aboutira à une refonte très profonde de ses constituants : une **révolution scientifique**. On distingue plusieurs moments dans la révolution scientifique. Tout d'abord intervient la **coupure épistémologique**, qui correspond au passage par un point de non-retour : on ne cherche plus à revitaliser

l'ancienne théorie, dont il est acquis qu'elle ne peut plus être simplement modifiée ou étendue. Après la coupure épistémologique intervient la **rupture épistémologique**, qui est un processus irréversible d'invalidation et de réfutation de l'ancienne théorie. Pendant ces deux dernières phases, la science est dite **science révolutionnaire**, par opposition à la science normale.

On arrive donc à la définition d'un processus cyclique, au cours duquel des phases d'évolution continue sont séparées par des changements radicaux.

Mais, par analogie avec les révolutions politiques, ces changements doivent-ils être brusques pour que l'on puisse parler de révolution scientifiques ? Non, pas forcément, puisque ce concept repose surtout sur la refonte des bases théoriques. De plus, le caractère instantané de la révolution peut également dépendre du critère employé pour définir la révolution : le critère objectif s'accorde parfaitement avec un caractère soudain, alors que le critère subjectif conduit généralement à l'identification de durée plus longue (il faut le temps de « convaincre » une communauté scientifique).

Le scénario discontinuiste est illustré par le diagramme ci-dessous.

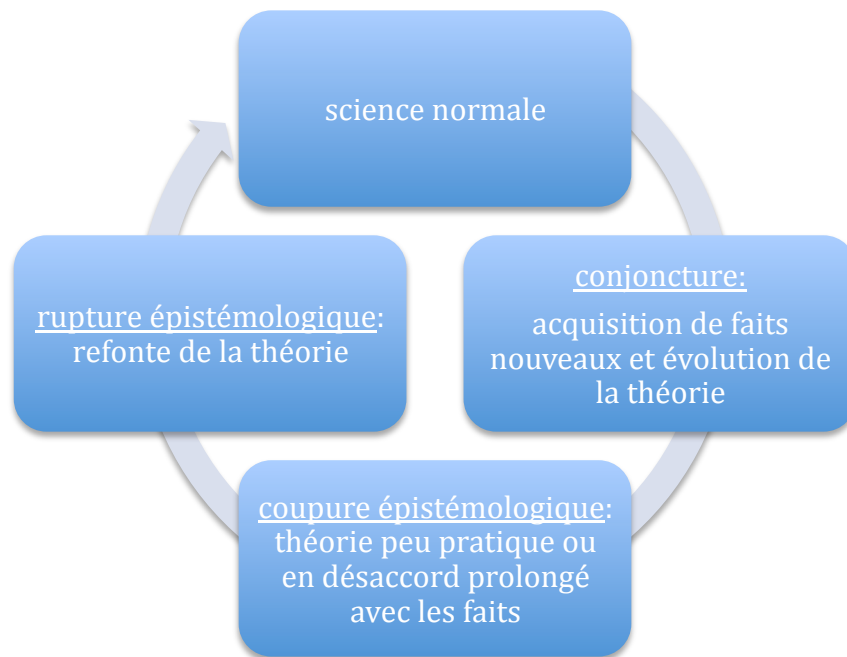


Figure 71 dynamique des sciences selon la thèse discontinuiste

L'analyse de Kuhn de l'évolution des sciences fait donc apparaître deux types de dynamiques scientifiques : l'évolution en période de science normale, et la révolution scientifique.

La période de science normale n'est pas une période de stagnation : elle correspond à une phase du processus durant laquelle le nombre et la gravité des anomalies contenues dans le paradigme sont suffisamment faibles pour que la communauté scientifique puisse résoudre les énigmes qui se présentent à elle sans trop de difficulté. La résolution de ces énigmes au moyen des lois universelles du paradigme conduit à un enrichissement continu et cumulatif de celui-ci. Le point important est que les

prémises du paradigme ne sont jamais remis en question durant cette phase, et que seules des hypothèses auxiliaires peuvent être remises en question.

Le passage de la période de science normale à la période de crise est pour Kuhn un changement majeur. Les fondements du paradigme sont remis en question, et s'ouvre alors pour les scientifiques une période de grande incertitude. La révolution est achevée lorsqu'un nouveau paradigme est apparu et a remplacé le paradigme en crise. Mais quels sont les critères qui permettent tout d'abord de décréter qu'un paradigme est en crise et que la période de science normale est finie, puis de choisir de le remplacer par un autre paradigme ? Pour ces deux questions, Kuhn admet qu'il n'existe pas de critère objectif. Il cite de nombreux critères : « *la précision de la prédiction, en particulier de la prédiction quantitative ; le rapport entre le nombre de sujets quotidiens et de sujets ésotériques ; et le nombre de problèmes différents résolus* » mais aussi « *la simplicité, l'étendue et la compatibilité avec d'autres spécialités* ». Dans sa théorie, les facteurs psychologiques et sociaux occupent une grande place. Plus précisément, Kuhn considère que des paradigmes concurrents sont **incommensurables**, au sens où leurs concepts sont si radicalement différents qu'il est impossible de « traduire » un paradigme en utilisant les concepts de son rival. Il ne peut donc pas y avoir de comparaison à proprement parler entre deux paradigmes, et le choix des scientifiques est assimilé par Kuhn à une « conversion religieuse » (il utilise lui-même cette analogie). L'analyse des révolutions scientifiques relève donc autant de la psychologie et de la sociologie que de l'analyse scientifique. Le changement de paradigme est le résultat d'une dynamique complexe, qui, selon les termes de Kuhn, provient « *d'une modification croissante de la distribution des persuasions professionnelles* »⁸⁹. Toujours selon lui, « *aucun argument ne saurait être rendu contraignant sur la plan de la logique ou même des probabilités* » et, définitive, il n'existe « *aucune autorité supérieure à l'assentiment du groupe intéressé* ». Le moyen par lequel les critères qui vont guider le choix entre deux paradigmes rivaux est imposé doit « *en dernière analyse, être psychologique ou sociologique. Autrement dit, il doit être une description du système de valeurs d'une idéologie, en même temps qu'une analyse des institutions à travers lesquelles ce système se renforce.* »

9.3.2 La thèse continuiste

Encore une fois, il est important de se rappeler que ce qui oppose les thèses continuiste et discontinuiste, c'est le jugement portant sur la radicalité du changement, et non la reconnaissance de son existence. Quelle continuité peut-on identifier dans les évolutions historiques des sciences ?

Les partisans de la thèse continuiste identifient deux types de continuité :

- La **continuité structurale** : la structure du formalisme mathématique employé reste la même. Cette continuité est en effet observée, et l'on remarque que dans certains cas les équations associées aux énoncés théoriques de l'ancienne théorie sont retrouvées comme des cas limites des équations de la nouvelle théorie.

⁸⁹ T. Kuhn « La structure des révolutions scientifiques », cité par A. Chalmers, « Qu'est-ce que la science ? », Coll. Biblio essais, Livre de poche

- La **continuité ontologique** : les entités physiques (masse, force, ...) restent les mêmes. Cette continuité est aujourd'hui difficile à défendre dans le cas général, au vu des développements de la mécanique quantique, dont les objets sont très différents de ceux de la mécanique classique. La continuité ontologique est également peu compatible avec la notion d'incommensurabilité des paradigmes.

9.4 Falsificationisme, progrès et évolution des théories scientifiques

Comme il a déjà été vu, la théorie falsificationniste donne une grande importance à la possibilité de falsifier les énoncés scientifiques. Pour un falsificationniste « naïf » (nous avons également vu que des variantes « assouplies » ou « sophistiquées » du falsificationnisme ont été proposées), il est impossible de prouver qu'une théorie scientifique est vraie mais il est possible de la réfuter. Quelles conséquences cette définition de la scientificité a-t-elle sur l'analyse de l'évolution des sciences ? C'est ce que nous allons voir maintenant.

Pour un falsificationniste « naïf », une nouvelle théorie scientifique sera supérieure à une théorie antérieure si elle est capable de résister à des tests qui avaient falsifié cette dernière, c'est-à-dire si elle permet de prendre en compte plus d'énoncés d'observation que la précédente. Le progrès scientifique est, dans cette perspective, conçu comme l'accroissement de la base empirique effectivement prise en compte par les théories scientifiques.

Cette analyse est affinée par les versions sophistiquées du falsificationnisme, selon lesquelles le degré de falsifiabilité d'un énoncé ou d'une théorie n'est pas suffisant pour pouvoir parler de progrès : il faut également que la nouvelle théorie soit plus falsifiable que celle qu'elle est destinée à remplacer. La bonne question n'est donc plus de savoir si une théorie est falsifiable (version « naïve »), mais de savoir si elle est plus falsifiable que les autres théories concurrentes (version « sophistiquée »). On passe d'une analyse statique qui porte sur le degré absolu de falsifiabilité d'une théorie à une version dynamique basée sur une mesure relative du degré de falsifiabilité. Une conséquence de ce changement est le passage d'une théorie normative de la définition de ce qu'est la science à une théorie du progrès scientifique. Le progrès scientifique consiste donc en la falsification d'hypothèses mais également, comme le souligne Popper en la corroboration d'autres hypothèses. Toutes les falsifications et corroborations ont-elles la même importance pour le progrès de la science ? Non. *Les vrais moteurs du progrès scientifique sont les falsifications de conjectures prudentes et les corroborations de conjectures audacieuses*, car elles nous apprennent beaucoup. Inversement, la corroboration d'une conjecture prudente et la falsification d'une conjecture audacieuse sont de peu d'utilité car elles apportent beaucoup moins d'information.

Il reste bien sûr à définir ce qu'est une conjecture audacieuse. Cette notion est définie relativement au **savoir acquis**, qui désigne la totalité des connaissances contenues dans l'ensemble des disciplines scientifiques à un instant donné. On dira qu'une hypothèse est audacieuse si ses affirmations apparaissent comme contradictoires par rapport au savoir acquis. Par exemple, l'hypothèse héliocentrique émise en 1543 par

Copernic était audacieuse, car elle se heurtait à l'hypothèse géocentrique admise jusque-là. Par opposition, une conjecture prudente est une conjecture qui ne remet pas en cause le savoir acquis. Un point important est le caractère historique du savoir acquis (il s'agit de l'ensemble des connaissances à un instant donné), qui fait que le caractère audacieux ou prudent d'une hypothèse peut varier au cours du temps. Ainsi, prouver de nos jours la validité du modèle héliocentrique n'aurait que très peu d'intérêt, alors qu'il s'agissait d'un fait d'une très grande importance au XVI^e siècle.

Un point important dans la théorie falsificationniste du progrès scientifique est la place accordée aux décisions des scientifiques dans le processus de falsification des énoncés. La dynamique du progrès scientifique possède donc une part non-négligeable de subjectivité. Celle-ci est d'autant plus grande que les théories sont complexes et ramifiées : on retrouve ici les arguments du holisme épistémologique. Ceci est illustré par l'exemple inventé par Imre Lakatos⁹⁰ : *« C'est l'histoire d'un cas imaginaire de comportement déviant d'une planète. Un physicien de l'ère pré-newtonienne prend comme point de départ la mécanique newtonienne et sa loi de la gravitation, N , des conditions initiales I , et calcule la trajectoire d'une petite planète nouvellement découverte, p . Mais la planète dévie de la trajectoire calculée. Notre physicien newtonien va-t-il considérer que cette déviation, interdite par la théorie de Newton, une fois établie, réfute la théorie N ? Non . Il suggère qu'il doit y avoir une planète p' jusqu'ici inconnue, qui perturbe la trajectoire de p . Il calcule la masse, l'orbite, etc., de cette planète hypothétique et demande ensuite à un astronome expérimentateur de tester son hypothèse. La planète p' est si petite que même les télescopes les plus puissants dont on dispose n'ont aucune chance de l'observer ; l'astronome expérimentateur rédige une demande de crédits de recherche dans le but d'en construire un plus gros encore. En trois ans, le nouveau télescope est prêt. Si l'on découvrait effectivement la nouvelle planète p' , ce fait devrait être célébré comme une nouvelle victoire de la science newtonienne. Mais tel n'est pas le cas. Notre savant va-t-il abandonner la théorie de Newton et son hypothèse d'une planète perturbatrice ? Non. Il suggère qu'un nuage de poussière cosmique nous cache la planète. Il calcule la position et les propriétés de ce nuage et demande des crédits de recherche pour envoyer un satellite tester ses calculs. Si les instruments du satellite (peut-être de nouveaux, fondés sur une théorie peu testée) avaient enregistré l'existence du nuage supposé, le résultat aurait été acclamé comme une éclatante victoire de la science newtonienne. Mais on ne trouve pas ce nuage. Notre savant va-t-il abandonner la théorie de Newton en même temps que l'idée d'une planète perturbatrice et d'un nuage qui la cache ? Non. Il suggère l'existence d'un champ magnétique dans cette région de l'univers qui perturbe les instruments du satellite. Un nouveau satellite est envoyé. Si on trouvait le champ magnétique, les newtoniens célébreraient une victoire sensationnelle. Mais ce n'est pas le cas. Est-ce considéré comme une réfutation de la science newtonienne ? Non. Soit on avance une autre hypothèse auxiliaire ingénieuse, soit ... on enterre toute cette histoire dans les volumes poussiéreux de périodiques et l'on en entend plus jamais parler. »*

Une faiblesse de la théorie falsificationniste est que l'histoire des sciences montre que des théories sont toujours utilisées malgré leur falsification. Par exemple, la mécanique newtonienne est aujourd'hui falsifiée par de nombreuses observations, et

⁹⁰ Cité par A. Chalmers, « Qu'est-ce-que la science ? », Coll. Biblio essais, Livre de poche, p.114

devrait donc systématiquement être remplacée par la mécanique relativiste. Ce qui n'est pas le cas.

9.5 L'évolution des programmes de recherche et des théories selon Lakatos

Comme il a été dit au chapitre 3, la réflexion de Imre Lakatos s'est principalement orientée autour de l'évolution scientifique au moyen de la théorie des programmes de recherche. C'est au travers de ces programmes, considérés comme des grilles de lecture, que Lakatos tente d'expliquer le développement des théories scientifiques. Nous allons maintenant voir quelle dynamique des sciences cette structuration en programme permet d'identifier. Avant d'aller plus loin, il est important de garder en mémoire que, pour Lakatos, la théorie des programmes de recherche était une réponse à l'incapacité des **théories inductivistes** (c'est-à-dire des théories qui, comme l'empirisme logique, font de l'expérimentation et de l'observation les moteurs du développement des sciences empiriques) et des théories falsificationnistes à rendre compte du développement des théories scientifiques tel qu'il est rapporté par l'histoire des sciences.

La question est double : comment évolue un programme de recherche, et pour quelle raison un programme de recherche prend-il le pas sur un programme concurrent ?

Commençons par la première question. Quels sont les moyens autorisés pour développer le programme de recherche, c'est-à-dire développer sa ceinture protectrice (rappelons que le noyau dur du programme est par définition infalsifiable et interchangeable) ? Tous, selon Lakatos, à l'exception de l'emploi d'hypothèses ad hoc. Le concept d'hypothèse ad hoc est identique à celui déjà rencontré au chapitre 5 lors de la présentation du falsificationisme. Une hypothèse ad hoc est une hypothèse qui n'est pas testable (*i.e.* falsifiable) indépendamment de la théorie dans laquelle on l'insère, c'est-à-dire que la théorie modifiée ne peut pas être testée d'une façon qui ne soit pas en même temps un test pour la théorie de départ. Les évolutions sont guidées par l'heuristique positive, qui indique des voies de recherche. Ces voies de recherche peuvent prendre la forme de nouvelles expériences, de la mise au point de nouveaux instruments de mesure ou encore de l'utilisation (voire du développement, comme le montre l'histoire des sciences !) de nouveaux outils mathématiques.

Venons-en à la seconde question, qui implicitement porte sur la comparaison de programmes de recherches concurrents. Ce problème est beaucoup plus compliqué que le précédent. Pour Lakatos, il convient de distinguer les **programmes de recherche progressistes** et les **programmes dégénérescents**. Un programme de recherche qui permet de nouvelles découvertes sera dit progressiste, alors qu'un programme devenu stérile sera dit dégénérescent. Le progrès scientifique est alors expliqué comme l'abandon des programmes dégénérescents au profit des programmes plus progressistes. Mais quand peut-on dire qu'un programme est dégénérescent, c'est-à-dire qu'il est devenu stérile et rempli de moins en moins bien le second critère définitoire de Lakatos (voir la définition d'un programme de recherche donnée au chapitre 3) ? Ici, le facteur temps intervient, et on rencontre une fois de plus un élément subjectif puisqu'il s'agit d'un jugement. Selon Lakatos lui-même, il est impossible de trouver une mesure objective de la stérilité d'un programme de

recherche. L'histoire des sciences fournit plusieurs exemples qui illustrent cette difficulté. En voici deux cités par A. Chalmers⁹¹.

Un premier exemple est donné par le programme de recherche copernicien. Il s'écoula effectivement beaucoup de temps avant que les prédictions de Copernic concernant la taille apparente de Vénus et la parallaxe des étoiles fixes ne soient vérifiées expérimentalement : plus de soixante-dix ans dans le premier cas, et plusieurs siècles dans le second.

Le développement des théories de l'électricité montre qu'un programme de recherche apparemment devenu stérile (donc dégénérescent) peut redevenir fécond (et donc progressiste). Un premier programme de recherche est la théorie (appelée parfois théorie de l'action à distance) qui considérait le courant électrique comme un flux de particules d'un certain type contenu dans les corps chargés électriquement et s'écoulant à travers les circuits électriques. Le second programme est la théorie du champ électrique de Faraday, selon laquelle les phénomènes électriques sont explicables en termes d'actions dans les milieux qui environnent les corps électrisés, plutôt que comme un flux de particules à travers les circuits électriques. Avant les succès de Faraday, la première théorie était majoritaire. Elle était féconde, car elle a permis de découvrir que l'on pouvait stocker l'électricité dans la bouteille de Leyde (l'ancêtre des piles actuelles) et elle a permis à Cavendish de développer sa théorie de l'attraction/répulsion entre les corps électriquement chargés de charges de signe opposé/identique en fonction de l'inverse du carré de la distance séparatrice. Puis les progrès de Faraday dans les années 1830 (sa découverte de l'induction électromagnétique, l'invention du moteur électrique, de la dynamo et du transformateur) firent de la théorie des champs la théorie dominante, car sa fécondité était supérieure à celle de sa rivale. Il faut noter que la définition du champ électrique donnée par Faraday était floue, et basée sur une analogie avec la mécanique newtonienne. Son succès fut encore renforcé près de soixante ans plus tard lorsque Heinrich Hertz prouva expérimentalement en 1888 l'existence des ondes radios prédites par la théorie du champ électromagnétique proposée par James Clerk Maxwell en 1864 mais restées inobservées jusqu'alors. Le programme de Maxwell était de donner à la théorie de Faraday une base solide en l'étayant par la mécanique de Newton, en la ramenant à des propriétés mécaniques d'un milieu mécanique particulier, l'éther. Cette théorie était alors une conjecture audacieuse (au sens discuté plus haut), puisqu'elle entraînait en contradiction avec le modèle dominant de la théorie de l'interaction à distance. C'est au cours de ces travaux que Maxwell introduisit un concept nouveau, le courant de déplacement, et rendit possible le développement de sa théorie du champ électromagnétique et de la théorie électromagnétique de la lumière. L'existence des ondes radio était contenue dans la théorie de Maxwell, mais Maxwell lui-même ne fit pas cette prédiction. Elle est l'œuvre de F.G. Fitzgerald, qui la réalisa en 1881, soit deux ans après la mort de Maxwell. Ce dernier ne connut donc jamais ni la prédiction, ni sa confirmation expérimentale.

Le premier programme n'était toutefois pas dégénérescent, puisqu'il conduisit à la découverte de l'électron. L'existence de l'électron avait été prédite de manière floue dans la première moitié du XIX^e siècle par un théoricien de l'action à distance, W. Weber. L'existence de l'électron fût ensuite prédite de manière beaucoup plus précise par H. Lorentz en 1892 et prouvée expérimentalement par J. Thomson durant la même

⁹¹ A. Chalmers, « Qu'est-ce-que la science ? », Coll. Biblio essais, Livre de poche.

décennie. La théorie de l'action à distance n'était donc pas devenue stérile, et son abandon complet aurait sans doute constitué une perte.

Il est intéressant de noter que la théorie moderne de l'électromagnétisme a hérité de ces deux théories : la théorie des champs et l'électron. Ceci prouve que les programmes de recherche ne sont pas complètement étanches. Ensuite, les prémisses de Maxwell sont aujourd'hui falsifiées : on sait que la théorie de Maxwell ne peut être déduite de celle de Newton. Néanmoins, la théorie qu'il a développée s'est révélée être éminemment féconde. Toutefois, il faut noter que les potentialités de la théorie de Maxwell n'étaient pas épuisées. Pendant que les physiciens maxwelliens « orthodoxes » Joseph Larmor et Oliver Lodge tentaient de développer des modèles pour l'éther, d'autres partisans de la théorie de Maxwell (H. Lorentz, H. Hertz) développaient un programme dérivé de celui de Maxwell dans lequel il n'était plus nécessaire de faire intervenir l'éther, donc sans tenter de faire le lien avec la mécanique des milieux continus. Ce programme dérivé mena jusqu'au développement de la théorie de la relativité restreinte par Albert Einstein et Henri Poincaré. Un autre constat est donc que la théorie du champ électromagnétique de Maxwell servit de base au développement de la théorie moderne des champs et à la remise en cause de la mécanique newtonienne, alors que Maxwell était un défenseur acharné de cette dernière. Ainsi, les résultats acquis peuvent parfois mener au développement d'une théorie « opposée » à celle initialement visée.

9.6 Le progrès scientifique est-il inéluctable ?

Les potentialités offertes par une théorie seront-elles forcément exploitées ? Une réponse positive à cette question conduit à dire que le progrès scientifique est inéluctable.

La réponse à cette question est bien sûr conditionnée par les prémisses philosophiques que l'on admet. Pour un externaliste radical, la réponse est négative, puisque la science est entièrement gouvernée par des facteurs sociaux et psychologiques. Il en est de même pour un individualiste radical, pour qui la science n'existe pas en dehors des croyances des individus. Mais notons que la réponse d'un internaliste « pur » n'est pas nécessairement positive.

Par contre, une certaine analyse objectiviste de la science, comme celle développée par Alan Chalmers, donne une réponse positive. La thèse de Chalmers est que toutes les opportunités objectives d'une théorie seront, un jour ou l'autre, saisies par quelqu'un. Il se base pour cela sur ce qu'il nomme une **hypothèse sociologique**, qui stipule qu'il existe une communauté scientifique maîtrisant les bases nécessaires (développant le programme de recherche, pour reprendre le schéma d'analyse de Lakatos). Si cette communauté est suffisamment nombreuse, alors toutes les opportunités objectives ouvertes seront exploitées, au bout d'un temps indéterminé. Plus le programme est fécond, plus il offre d'opportunités objectives et plus il est attractif, expliquant par cela les changements de paradigme ou de programme de recherche. Cette vision objectiviste a deux conséquences : tout d'abord, le progrès scientifique, assimilé aux révolutions scientifiques ou aux changements de programme de recherche, aura lieu inéluctablement. Ensuite, ce progrès est indépendant des individus. Le rôle accordé aux « génies de la science » est donc fortement diminué. La disparition d'un individu remarquable ne peut, au pire, que

retarder l'évolution de la théorie jusqu'à ce que ses découvertes soient faites par d'autres. Cette théorie ne dit pas à quelle vitesse l'évolution va se dérouler. Ce facteur est certainement impossible à évaluer, comme le montre l'histoire des sciences. Par exemple, on sait que les verres de lunettes ont été introduits dans l'Europe médiévale entre 1280 et 1285, et il a fallu attendre environ trois siècles pour que la lunette astronomique, qui est composée de deux de ces verres, soit mise au point. La technologie nécessaire était disponible dès le XIII^e siècle, et donc l'opportunité objective existait depuis ce moment. Pourquoi ce délai ?

10 Cause, causalité, déterminisme et hasard

Ce chapitre est dédié à l'examen de trois grandes notions rencontrées dans les sciences empiriques et étudiées par les épistémologues : la causalité, le déterminisme et le hasard. Ces trois notions occupent une place centrale dans la conception commune de la science. Elles semblent intuitivement relativement aisées à concevoir et à manipuler. Nous allons voir qu'un examen détaillé révèle que la réalité est tout à fait différente.

10.1 Cause et causalité : de Aristote au XXe siècle

La science « explique », elle dit donc « pourquoi » les phénomènes se produisent (voir le chapitre 6). Et, la plupart du temps, la réponse est fournie sous la forme d'un scénario, dans lequel les phénomènes s'enchaînent, les plus anciens étant présentés comme étant les causes des suivants (les effets). La notion de cause semble liée à celle d'explication, donc de science, et cela de manière indéfectible.

Avant d'entrer dans le cœur de la discussion, posons le problème de manière explicite. On considère un système physique, décrit par un ensemble de propriétés (masse, énergie, ...) qui lui sont propres, et qui évolue dans le temps. Un pré-requis est que ce système doit conserver son identité durant le changement : à tout moment, nous devons être capables de le distinguer du monde extérieur. La question de la cause est d'expliquer les évolutions du système.

Les sciences empiriques disposent de plusieurs modes explicatifs (voir le chapitre 6), qui tous visent à expliquer soit la **production** d'un phénomène ou d'une entité qui n'existait pas auparavant, soit la **détermination** des valeurs des paramètres qui décrivent un système lors de son évolution.

Les différentes théories philosophiques et épistémologiques (certaines faisant l'objet des sections ci-dessous) peuvent être rangées en deux grandes catégories :

- La première correspond aux théories selon lesquelles la relation de causalité existe dans une réalité extérieure autonome, indépendamment de l'observateur (on parle alors de *description sur le plan ontique*). Ici, on peut distinguer plusieurs conceptions :
 - Les théories selon lesquelles la causalité est bien réelle, mais n'est pas réductible. Deux variantes s'affrontent. Une variante selon laquelle la causalité est singulière par nature, et les relations causales sont complètement étrangères à l'existence de lois causales (mais les lois causales dépendent de la répétition régulière de ces relations singulières). Une seconde variante selon laquelle la causalité est une relation théorique universelle, qui est liée à l'existence de lois dans la nature.
 - Les théories selon lesquelles la causalité est réductible à une transmission (Salmon, Skyrms, ...). Un processus causal est identifié

par la transmission de quelque chose (information, énergie, ...) d'un système à l'autre, cette transmission étant identifiée comme étant la cause du changement d'état d'un des système (ce changement est l'effet).

Dans les deux cas, une question importante est de savoir si il est possible pour l'esprit humain de percevoir l'interaction causale (voir les discussions des travaux de Piaget et Whitehead dans ce chapitre).

- La seconde catégorie est celle des théories pour lesquelles l'explication causale est introduite par l'esprit humain pour ordonner et maîtriser sa perception du monde extérieur. La relation de causalité est dans ce cas une création de l'esprit humain, qui ne reflète pas strictement une réalité extralinguistique. Cet aspect psychologique a amené plusieurs philosophes à remettre en question la notion de causalité, ou à la considérer au mieux comme un mal nécessaire. Différentes variantes sont discutées dans ce chapitre à travers les présentations des travaux de Hume, Kant, Russel et Mach.

10.1.1 Les quatre causes selon Aristote

Aristote nommait **aïtion** (« responsable ») une entité qui détermine ce qu'une autre est, qui la fait changer ou la fait exister. Ce terme a ensuite été traduit en latin par **causa**, dont dérive notre mot cause. Ce concept est associé à la notion de production, et répond pour Aristote à la question « pourquoi » : « *nous ne croyons connaître rien avant d'en avoir saisi chaque fois le pourquoi* ». Dans la théorie d'Aristote, qui dominera la réflexion occidentale pendant près de deux mille ans, on distingue quatre sortes de cause :

Les quatre causes selon Aristote

1. La **cause matérielle** : la matière dont est faite une chose
2. La **cause formelle** : la structure ou le contenu formel d'une chose
3. La **cause efficiente** : ce qui est responsable de l'évolution observée
4. La **cause finale** : le but de l'existence d'une chose

Illustrons ces quatre causes en considérant l'exemple d'une maison : pourquoi est-elle capable de nous protéger des intempéries ?

La cause matérielle est la solidité des matériaux utilisés. La cause formelle est la manière dont ils ont été agencés (murs, toit). La cause efficiente, se sont les ouvriers qui ont bâti la maison. La cause finale, c'est le but que poursuivait les bâtisseurs de la maison (gagner de l'argent, par exemple).

Cet exemple met en jeu une cause efficiente basée sur des agents humains, donc douée d'intention, ce qui explique pour nous la possibilité d'envisager une cause finale. Qu'en est-il lorsque les agents sont inanimés (corps solides, ...) ? Pour Aristote, ce problème n'existe pas : dans sa philosophie, tout changement est orienté vers une fin, car le cosmos entier est un ordre, une harmonie. Toutes les parties de l'univers sont mutuellement complémentaires, et concourent au maintien de cette harmonie. La cause finale renvoie alors à ce principe de maintien de l'ordre universel.

La théorie aristotélicienne de la causalité a eu une importance énorme dans le développement des sciences, de la philosophie et de la théologie en Occident, puisqu'elle a dominé jusqu'au XVII^e siècle, soit pendant près de deux mille ans ! C'est à partir de cette époque que la notion de cause formelle est apparue moins attractive. Par exemple, Galilée et Képler ne se réfèrent à la notion de cause formelle que pour évoquer une certaine régularité mathématique dans les phénomènes observés. La notion de cause efficiente, intuitivement liée aux interactions mécaniques, sensibles, perceptibles est apparue à cette époque beaucoup plus attractive, de nombreux savants ont tenté de décrire l'ensemble des phénomènes à des interactions entre objets physiques. Ainsi, René Descartes propose d'expliquer le poids d'un corps comme résultant de l'impact sur la surface supérieure de ce corps de particules de l'éther ambiant. Ce souci d'associer à chaque phénomène une interaction matérielle fut à l'origine de nombreuses critiques du travail de Newton, lorsqu'il introduisit sa proposition d'interaction gravitationnelle à distance, qui ne faisait pas intervenir de contact matériel « physique » entre les corps, et qui donc s'éloignait de la conception « classique » de la cause efficiente. Ce type d'explication par la cause efficiente associée à des interactions matérielles est très répandu au XVIII^e siècle, durant lequel il fut appliqué dans de nombreux domaines émergents: électricité, magnétisme (théorie du flegme), chaleur (théorie du phlogistique). C'est au XIX^e siècle que les changements initiés en mécanique se répandent dans les autres domaines : la notion de cause efficiente par interaction matérielle est peu à peu abandonnée au profit de la mise en évidence de formes appropriées (les lois) et de la dérivation de leurs conséquences. Couplée aux développements des mathématiques et à la mathématisation de la physique et de la mécanique, l'explication s'est alors tournée vers l'élaboration d'équations différentielles.

La notion de cause dans les sciences empiriques a subi une révolution à partir du XX^e siècle, au cours duquel l'évolution des théories physiques a changé en profondeur les concepts de force, de champs, de matière, ... Dans sa version actuellement maniée par la plupart des scientifiques, l'explication causale est structurellement proche de l'explication aristotélicienne tout en étant radicalement différente dans son contenu : la cause finale est rejetée car anthropomorphe au profit d'une nécessité aveugle compatible avec la notion de déterminisme ; la cause formelle et la cause matérielle ont été abandonnées car elles correspondent plus à la définition acceptée d'une cause. Le développement de la physique statistique et de la mécanique quantique a également bouleversé la vision qu'ont les scientifiques de la causalité, en conduisant au développement des théories de la causalité statistique (ceci est discuté au chapitre 10.2).



Figure 72: D. Hume

10.1.2 La causalité selon Hume

Une remise en cause du concept aristotélicien de cause efficiente est due au philosophe anglais David Hume (1711-1776), dans ses deux ouvrages *A Treatise of Human Nature* (1738) et *An Inquiry Concerning Human Understanding* (1748).

Le premier argument de Hume est que le terme de cause est complexe, et que les définitions d'Aristote et les concepts qui y sont souvent rattachés (force, puissance, énergie, action, efficacité) le sont tout autant. Donc, pour Hume, il s'agit d'une notion vague, qui ne nous dit pas exactement ce qu'est une relation causale : il s'agit d'un cercle vicieux sémantique.

Le second argument de Hume est que la relation causale n'est pas immédiatement accessible aux sens. Il donne, pour illustrer cela, deux exemples.

Le premier concerne le choc de deux boules de billard, l'une étant initialement au repos. Qu'observons-nous ? Avant le choc, une boule est au repos, pendant que l'autre se déplace. Après le choc, les deux boules sont en mouvement (ou les rôles sont inversés entre les deux boules, selon les cas). L'explication classique consiste à dire que le mouvement de la première boule a causé celui de la seconde, et cette affirmation nous semble justifiée par notre observation. Faux, répond Hume ! A aucun moment nous n'avons observé la relation de causalité, c'est-à-dire la production du mouvement de la seconde boule. Nous avons simplement observé une suite de phénomènes, ordonnés dans le temps.

Le second exemple est celui du doigt plongé dans le feu. Au fur et à mesure que nous approchons notre doigt de la flamme, nous ressentons une sensation accrue de chaleur, qui, passée un certain seuil, se transforme en douleur. La sensation est-elle produite par la flamme ? Pour Hume, rien ne permet de l'affirmer : ce qui est perçu, c'est une série d'évènements et de sensations simultanés ou successifs.

La troisième critique de Hume porte sur le **principe de causalité**, qui, selon lui, ne peut être démontré. Avant d'aller plus loin, il est nécessaire de rappeler qu'il existe deux formulations de ce principe :

Expressions du principe de causalité

- **Forme 1** : Toute cause a nécessairement un effet, et tout effet a nécessairement une cause.
- **Forme 2** : Tout ce qui commence à exister doit nécessairement avoir une cause.

La première forme du principe n'apporte rien : elle n'est qu'une définition, et ne se prête donc pas à la démonstration. C'est à la seconde forme, qui porte sur la production des phénomènes, et à ses justifications que s'attaque Hume.

Tout d'abord, puisqu'il s'agit ici d'un principe général, il ne peut être démontré au moyen de l'induction : un seul contre-exemple suffirait pour l'infirmier, alors qu'une infinité d'observations positives ne feraient que le corroborer. De plus, pour Hume, l'observation est inefficace puisque la relation de causalité n'est pas observable.

Ce principe peut-il être établi au moyen de la déduction ? C'est ce qu'ont tenté de faire plusieurs philosophes, dont Hume réfute les arguments :

- *L'argument de Thomas Hobbes (1588-1679)* peut s'exprimer comme suit : les points de l'espace et les instants du temps étant indifférents aux objets, propriétés ou processus de changement qui les occupent, il faut nécessairement l'intervention d'une cause pour que les objets occupent la

place qu'ils ont et les phénomènes se déroulent comme ils le font. Sans cause, rien ne commencerait jamais à exister. Selon Hume, l'argument de Hobbes est vicié sur le plan logique. Tout d'abord, il établit une différence sans justification entre le temps et l'espace qui peuvent exister sans cause, et les objets et les phénomènes qui ont besoin d'une cause pour exister. Ensuite, la démonstration se base sur le principe de causalité !

- *L'argument de Samuel Clarke (1675-1729)* consiste à dire que ce qui arrive doit être causé par autre chose que lui-même, car pour causer il faut exister, et la cause est donc antérieure à l'effet. Sans quoi, pour causer sa propre existence, il faudrait causer avant d'exister, ce qui est absurde. La réfutation de Hume tient dans le constat que Clarke postule le principe de causalité, mais ne le démontre pas.
- *L'argument de John Locke (1632-1704)* : ce qui est produit sans cause est donc produit par rien, or rien ne peut causer car il n'existe pas. Ici aussi, remarque Hume, on postule que la cause est nécessaire pour exister. Une fois de plus, on postule sans démontrer.

Le concept de cause doit-il pour autant être complètement rejeté ? Non. Hume propose de rompre complètement avec la théorie d'Aristote en dissociant complètement les notions de production et de cause. Selon lui, il est légitime d'utiliser les termes de cause et d'effet pour distinguer deux phénomènes lorsque les conditions suivantes sont réunies :

Causalité selon Hume

1. la cause et l'effet doivent être contigus dans l'espace et le temps
2. la cause doit être antérieure à l'effet
3. le lien entre cause et effet doit être constant (observation répétable), ce qui est l'élément constitutif de la relation causale
4. la même cause a toujours le même effet et le même effet a toujours la même cause
5. si plusieurs causes différentes ont toutes le même effet, c'est qu'elles ont en commun une certaine qualité grâce à laquelle elles ont un effet commun
6. si un effet augmente ou diminue en fonction de l'augmentation ou de la diminution de la cause, alors il est un effet complexe, chacune de ses parties étant respectivement causée par une partie spécifique de la cause
7. une cause qui n'a pas toujours son effet spécifique n'est qu'une cause partielle

Comme on peut le voir, la notion de production de la théorie aristotélicienne est remplacée par celle de **corrélation**.

Mais, si le concept de relation causale est si difficile à définir et à vérifier, pourquoi y faisons-nous référence avec autant d'obstination ? Par habitude et par commodité, répond Hume. Le concept de relation causale nous permet d'anticiper les phénomènes futurs, et cela de manière vraisemblable mais sans garantie.

10.1.3 La causalité chez Kant

Une contribution très importante au développement du concept de cause est due au philosophe allemand Emmanuel Kant (1724-1804).

Avant d'aborder sa théorie de la causalité, rappelons les éléments essentiels de la philosophie kantienne, qui est très marquée par la mécanique newtonienne.



Figure 73: E. Kant

Tout d'abord, Kant définit un concept comme quelque chose qui est opposé à l'intuition sensible (*Critique de la raison pure*, 1781, 1787): « toutes les intuitions, en tant que sensibles, reposent sur des affections, et les concepts, par conséquent, sur des fonctions. Or, j'entends par fonction l'unité de l'acte qui range diverses représentations sous une représentation commune. Les concepts reposent donc sur la spontanéité de la pensée, comme les intuitions sensibles, sur la réceptivité des impressions. » Toujours selon Kant, les concepts ne peuvent être définis de manière isolée : ils forment un groupe. L'identité de chaque concept est déterminée par les liens logiques qu'il entretient avec les autres concepts.

Les concepts sont les éléments du jugement, dont les formes les plus générales forment les **catégories**. Les catégories sont au nombre de quatre : **quantité, qualité, relation et modalité**, chacune comprenant trois divisions. La causalité apparaît dans la catégorie de la relation, qui est divisée en **substance, causalité et action réciproque**. Ces trois catégories sont reliées chez Kant à l'expérience par ce qu'il appelle les **analogies de l'expérience**. La première est celle du **principe de permanence de la substance** : ce qui fait qu'une entité est ce qu'elle est ne varie pas lorsque cette entité évolue. La seconde analogie est le **principe de la succession dans le temps suivant la loi de la causalité** : tous les changements se produisent suivant la relation de la cause à l'effet. Alors que chez Hume la causalité est une commodité dictée par l'habitude, et que chez d'autres elle résulte directement de la perception, elle est, chez Kant, une catégorie a priori de la pensée. La causalité, chez Kant, est une catégorie nécessaire de l'esprit, qui ne peut percevoir directement le temps, qui n'est perçu qu'à travers les changements. La causalité est la catégorie de la pensée qui permet à l'esprit de déterminer le rapport objectif qui existe entre les phénomènes, et non la succession temporelle. La causalité est donc fondatrice de notre connaissance du monde : ce n'est que « *parce que nous soumettons la succession des phénomènes, et, par suite, tout changement, à la loi de la causalité qu'est possible l'expérience même, c'est-à-dire la connaissance empirique des phénomènes.* » Cette nécessité induit la nécessité de l'existence d'une cause pour tout changement dans le monde.

Le principe de causalité induit donc un ordre dans le temps entre les phénomènes. Mais il n'induit pas le cours du temps : en effet, la cause et l'effet peuvent être simultanés, c'est-à-dire exister au même temps, sans faire intervenir le **principe des actions retardées**, qui implique l'antériorité de la cause par rapport à l'effet.

Kant applique son analyse à la mécanique dans son ouvrage intitulé *Premiers principes métaphysiques de la science de la nature*. La mécanique appartient à la catégorie de la relation et comprend, selon Kant, trois théorèmes :

Les 3 théorèmes de la mécanique selon Kant

- « *La quantité de matière comparée à tout autre ne peut s'évaluer qu'au moyen de la quantité de mouvement à une vitesse donnée.* »
- « *dans toutes les modifications de la nature corporelle, la quantité de matière reste la même dans l'ensemble sans augmentation, ni diminution.* »
- « *tout changement dans la matière a une cause externe (chaque corps persévère dans son état de repos ou de mouvement, en conservant même direction et même vitesse quand une cause externe ne l'oblige pas à quitter cet état.)* »

La philosophie de Kant aura un impact certain chez les scientifiques, notamment chez le physicien allemand Helmholtz, qui développera sa recherche en se basant sur l'idée de conservation de l'énergie, inspirée par le principe de permanence de la substance de Kant, qui préfigure les lois de conservation.

10.1.4 Mach : la théorie de la relation fonctionnelle

La conception traditionnelle de la cause qui, de Aristote à Kant, repose sur un certain degré de réalisme métaphysique, va être remise en cause par l'antiréaliste Ernst Mach dans le dernier quart du XIX^e siècle. En effet, pour Mach, qui est partisan du phénoménisme sensualiste, les fondements de la théorie classique de la causalité ne sont pas acceptables. Aussi écrit-il (*L'analyse des sensations*, 1872) : « *La conception ancienne de la causalité, que nous avons héritée de la tradition, est plutôt rigide : pour une quantité définie de cause résulte une quantité définie d'effet. Telle est l'expression typique d'une conception du monde primitive et pharmaceutique, comme c'est encore le cas dans la théorie des quatre éléments. L'emploi du terme de cause suffit à nous le rendre évident. Dans la nature, les connexions sont rarement assez simples pour que l'on puisse fournir, dans un cas donné, une seule et même cause, un seul et même effet. C'est pourquoi je me suis efforcé depuis longtemps de remplacer le concept de cause par le concept de fonction mathématique, ce qui signifie : dépendance mutuelle des phénomènes, ou plus précisément dépendance mutuelle des caractères qui appartiennent aux phénomènes. Nous pouvons élargir ou restreindre ce concept à notre gré, selon ce qui est requis par les faits que nous examinons.* »

Il émet une critique plus radicale en 1871 dans son ouvrage *La mécanique* : « *il n'y a aucune cause ni aucun effet dans la nature ; la nature n'a rien d'autre qu'une existence individuelle ; la nature simplement est ... L'essence de la connexion de la cause et de l'effet n'existe que dans l'abstraction que nous effectuons afin de reproduire mentalement les faits.* » Donc, pour Mach, la causalité est de l'ordre de la nécessité psychologique et provient de notre expérience du mouvement volontaire.

Un point important de la philosophie de Mach de la notion de causalité est sa critique de la physique newtonienne qui, selon lui, n'est pas capable de rendre compte de la complexité de la nature. Selon lui, les formes des relations mathématiques (lois) de la physique newtonienne ne permettent pas de rendre compte de l'interdépendance qui existe entre les phénomènes. En effet, les énoncés de la mécanique classique traitent des interactions qui existent entre des objets identifiés, bien définis, et en nombre limité. En bref, la mécanique newtonienne considère des systèmes isolés. Or, selon

Mach (*L'analyse des sensations*) : « il est impossible de concevoir aucun système totalement isolé du reste du monde, puisque la détermination du temps, et par conséquent celle des vitesses, dépendent d'un paramètre supposé, qui est défini par la trajectoire qu'a parcourue un corps situé en dehors du système (une planète). La dépendance effective de la position d'un corps quelconque de l'univers – même si nous ignorons tous les processus qui affectent cette position, et dont elle dépend directement par ailleurs – nous garantit la cohésion même de l'univers... Tous les rapports de dépendance reconnus et clairement établis se laissent envisager comme des relations réciproques de simultanéité. »

10.1.5 La causalité perceptible : Whitehead et Piaget

L'affirmation de Hume selon laquelle la relation causale échappe à l'observation a été critiquée par plusieurs auteurs.



Figure 75: J. Piaget



Figure 74: A. Whitehead

Pour le mathématicien, logicien et philosophe anglais Alfred North Whitehead (1861-1947), le rapport causal n'est pas lié à l'habitude, et Hume confond deux types d'expérience différents. Le premier est du type « **presentational immediacy** », où le sujet est un pur spectateur : c'est le cas du choc des boules de billard. Le second est du type « **causal efficacy** », où le sujet est personnellement impliqué, comme agent du changement ou comme subissant une action : c'est le cas du doigt plongé dans le feu. Dans le premier cas, le recours à l'explication causale peut sembler être une convention, une mesure de commodité intellectuelle. Dans le second cas, le rapport de causalité apparaît comme un sentiment instantané de certitude, comme une conviction intime non-réfléchie. Cette distinction est d'origine psychologique, et recouvre également des différences physiologiques. L'analyse des exemples donnés par Whitehead révèle que les expériences faisant intervenir un rapport du type « presentational immediacy » sont essentiellement des expériences visuelles et auditives, alors que celles qui se rapportent à la « causal efficacy » impliquent un contact physique « direct », c'est-à-dire le toucher (perception haptique, ou tactilo-kinesthésique dans le langage des psychologues).

Une seconde critique de la thèse de Hume a été formulée par le psychologue suisse Jean Piaget (1896-1980), qui a beaucoup travaillé sur la genèse et le développement des concepts chez l'enfant (c'est ce qu'on appelle la psychologie génétique). Les observations de Piaget et de son groupe montrent que la construction chez l'enfant de la représentation du monde extérieur passe par la manipulation répétée des objets et la découverte de la relation causale. C'est la répétition des expériences tactiles qui permet à l'enfant de construire

la notion d'objet permanent, d'espace et de temps. Donc, selon Piaget, il existe bien chez l'homme une expérience authentique, intime, de la causalité.

Faut-il pour autant rejeter complètement les conclusions de Hume ? Non. Car, si il apparaît aujourd'hui qu'il existe bien chez l'homme une perception de la causalité, la critique de Hume tient toujours en ce sens que le principe de causalité n'a toujours pas été établi comme un principe nécessaire et universel. En conséquence,

le principe de causalité doit être conservé comme un postulat nécessaire à l'intelligence humaine pour ordonner les perceptions et construire un modèle efficace du monde extérieur qui lui permette d'agir sur les objets et les phénomènes, en permettant l'analyse, l'explication et l'anticipation.

10.1.6 La causalité selon Russell : les lignes causales

Une autre remise en cause de la théorie classique de la causalité est fournie par le logicien et philosophe Bertrand Russell (1872-1970, prix Nobel de littérature 1950), qui écrit dans un article en 1913 : « *la loi de causalité, je pense, [...], est une relique du passé qui ne survit, comme la monarchie, que parce qu'elle est supposée, de manière erronée, ne pas faire de mal.* » En 1948, dans son livre *La connaissance humaine*, il développe une théorie alternative alors qu'il s'intéresse aux conditions nécessaires pour pouvoir valider la méthode scientifique. Cette théorie est basée sur le concept de **ligne causale**. Revenons tout d'abord sur les conditions nécessaires à la validation. Elles sont au nombre de cinq, exprimées sous la forme de postulats :

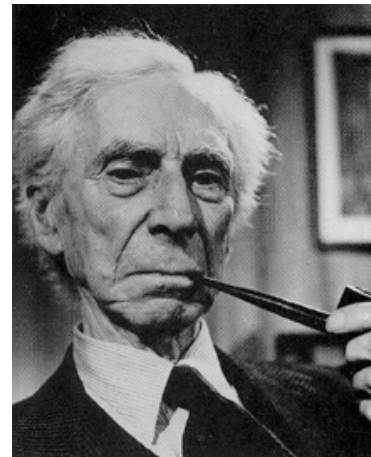


Figure 76: B. Russell

Les 5 postulats de Russell

1. **Postulat de quasi-permanence** : il existe une certaine permanence dans l'univers, car les choses ne changent que de manière continue
2. **Postulat des lignes causales séparables** : il existe souvent une persistance des objets et des processus sur des temps longs
3. **Postulat de la continuité spatio-temporelle** : il ne peut y avoir de relation causale qu'entre des objets contigus dans l'espace-temps
4. **Postulat structurel** : il est possible d'établir une inférence entre des événements complexes de structures semblables et un événement de même structure relié par des lignes causales à chaque événement.
5. **Postulat de l'analogie** : il est possible d'inférer l'existence d'un effet causal quand celui-ci n'est pas observable.

On le voit, la notion de ligne causale est centrale dans la théorie de Russell. Mais comment est définie une ligne causale ? « *J'appelle ligne causale une suite d'évènements si, certains étant donnés, il est possible d'inférer quelque chose sur les autres sans rien connaître du monde extérieur* ». La ligne causale est donc la trajectoire temporelle de quelque chose qui ne varie pas trop et qui reste isolé du monde qui l'environne. « *Le concept d'objet physique plus ou moins permanent, implique, selon le sens commun, la « substance », et quand la « substance » est rejetée nous devons trouver un autre moyen de définir l'identité d'un objet physique à différents instants. Je pense que cela doit être fait au moyen du concept de « ligne causale ».* »

La théorie de Russell a été critiquée récemment par Wesley Salmon. Une première critique est que la formulation de Russell porte sur la manière dont nous décrivons les phénomènes, et non sur les liens qui unissent les phénomènes eux-mêmes. La théorie de Russell ne répond donc pas directement aux objections de Mach. Ensuite, les lignes causales de Russell ne couvrent pas l'ensemble des relations causales possibles. Par exemple, on peut réaliser des inférences entre deux objets qui n'ont pas de lien causal entre eux (par exemple, deux nuages de même composition chimique ont la même signature spectrale, bien que n'étant pas liés par une relation causale). La seconde critique est que la théorie de Russell ne permet pas de distinguer entre pseudo-processus et processus causal. On nomme ici **pseudo-processus** (ou séquence irréaliste) des phénomènes qui n'ont pas de pouvoir d'action (par exemple, l'ombre d'un objet opaque).

10.1.7 Causalité et complexité : la cause INUS de Mackie

Les phénomènes observés sont souvent complexes, en ce sens qu'ils mettent en jeu de nombreux objets et phénomènes. Aristote distinguait, outre les quatre types fondamentaux de cause, les causes partielles et les causes totales, ainsi que les causes immédiates (proches) et médiates (éloignées). Ces distinctions sont parfois très difficiles à employer dans des cas concrets. Prenons le cas de la découverte de Neptune par Le Verrier. Quelles sont les causes de cette découverte ? Elles sont nombreuses : ses discussions avec Arago, la malchance de Adams, la diligence de Galle, l'existence des moyens expérimentaux et mathématiques nécessaires, la disponibilité des observations sur l'irrégularité de la trajectoire d'Uranus, la découverte d'Uranus, ... On voit s'établir une chaîne causale ramifiée, complexe, dans laquelle les facteurs s'entrecroisent. Il est donc presque impossible de classifier et d'ordonner chaque élément, voire même d'identifier les causes au sens aristotélicien. Pour palier ce problème et obtenir une notion de cause adaptée à la complexité, J.L. Mackie a publié en 1980 la nouvelle définition suivante pour la cause

Mackie : cause = INUS

La cause est une condition INUS (« Insufficient but Necessary part of an Unnecessary but Sufficient set of conditions »), c'est-à-dire un élément insuffisant mais nécessaire d'un ensemble non-nécessaire mais suffisant de conditions pour produire l'effet.

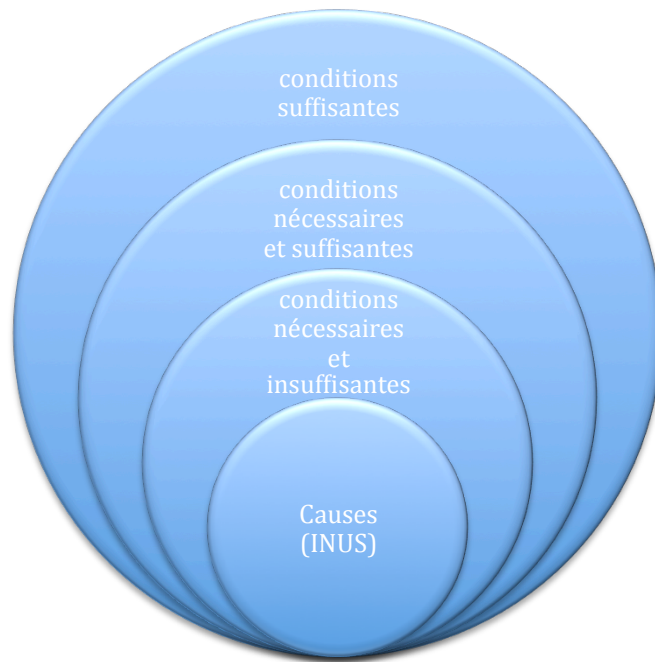


Figure 77 Illustration de la définition de la cause selon la théorie INUS

Cette définition est à rapprocher de celle donnée par J. Stuart Mill de la cause comme une « condition à la fois nécessaire et suffisante », qui résume la définition de Hume. La définition INUS est compatible avec celle de Hume, en ce sens qu'elle ne permet pas de distinguer entre cause (au sens de facteur responsable de la production) et corrélation. L'analyse de Mackie de la théorie de Hume est résumée dans le diagramme suivant :

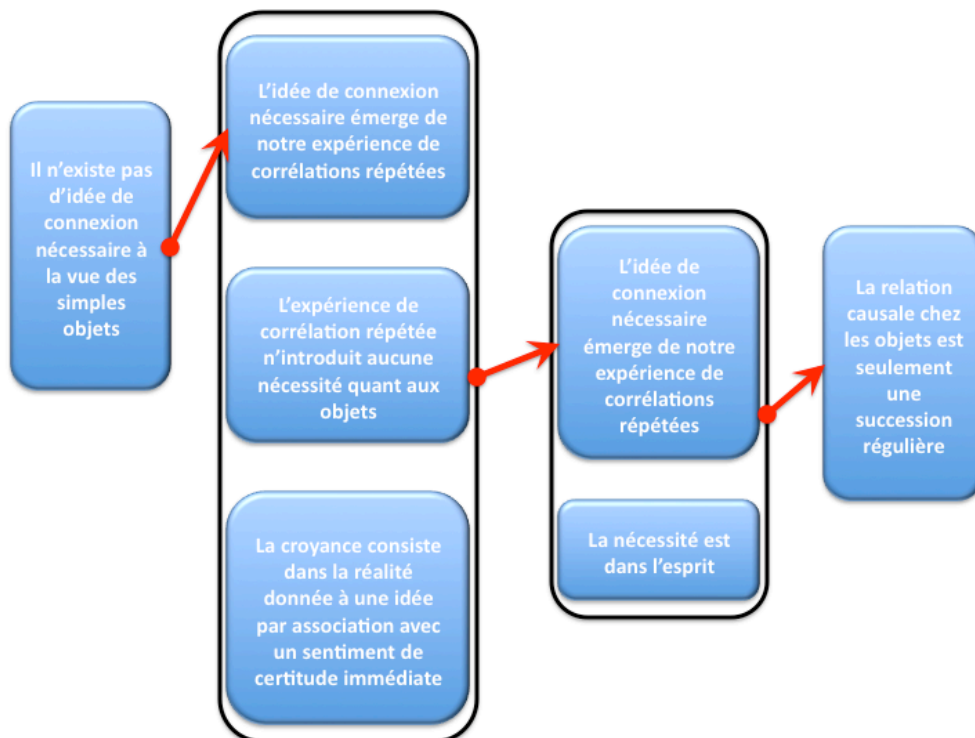


Figure 78: diagramme de Mackie pour la causalité selon Hume

où l'inférence conduisant à la proposition « l'idée de connexion nécessaire émerge de notre expérience de corrélations répétées » est analysée comme suit :

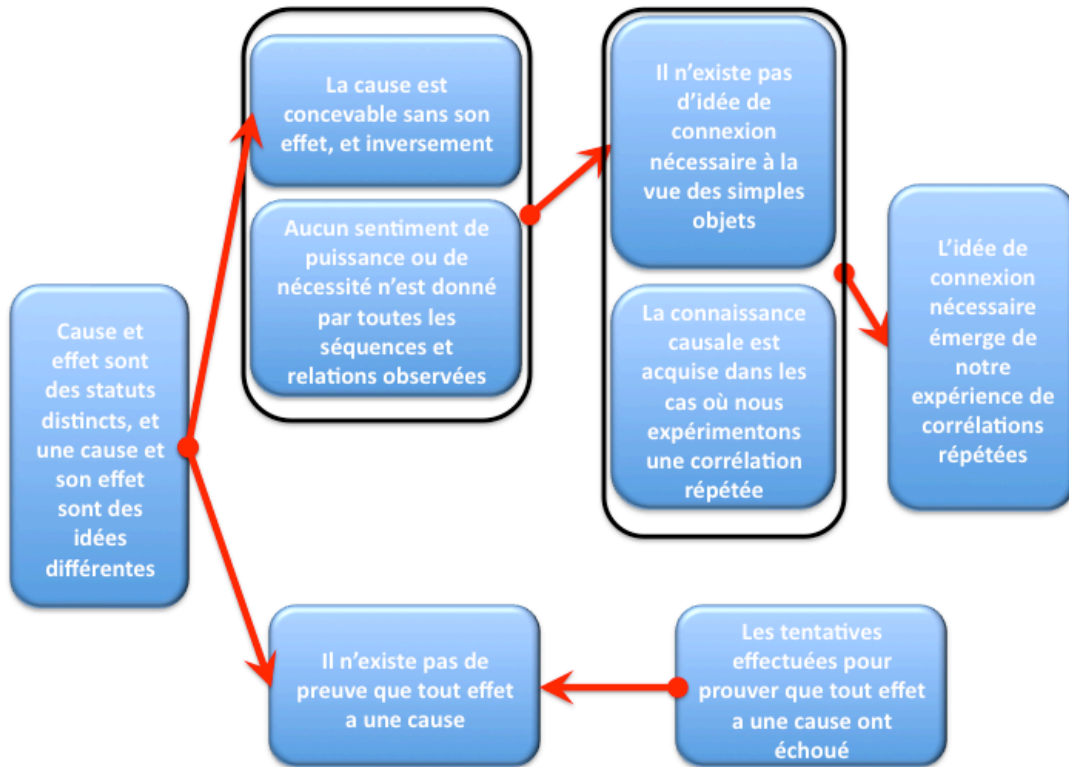


Figure 79: second diagramme de Mackie

Une propriété de la définition de Mackie est qu'elle permet l'établissement d'un lien causal rétroactif (voir figure ci-dessous). Considérons un événement A apparaissant à l'instant t_2 et un événement ultérieur B à l'instant $t_3 > t_2$. Il est possible que B ait été fixé à un temps antérieur $t_1 < t_2$ par une cause suffisante pour B, alors que A n'est pas encore fixé à cet instant t_1 . Il est alors possible d'envisager un lien causal rétroactif (mais pas d'agir sur le passé !) si B est une cause de A, en considérant A à t_1 .

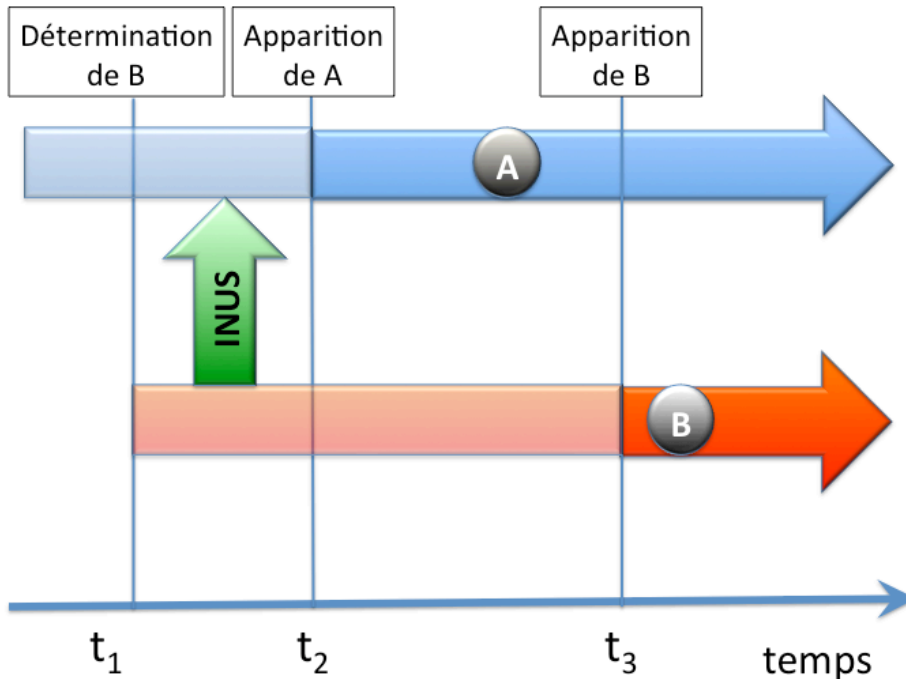


Figure 80 Illustration de lien causal rétroactif: A est déterminé par B avant l'observation de B

10.1.8 La causalité selon Salmon : la théorie MT

Sur la base de la théorie de Russell qu'il trouvait imparfaite, Wesley Salmon a proposé en 1984 une théorie de la causalité qui fait appel au **principe de transmission des marques** (« *mark transmission* » en anglais). L'objectif de Salmon est de bâtir une théorie qui soit compatible avec les assertions suivantes :

1. la causalité est une propriété objective du monde
2. la causalité est une propriété contingente du monde
3. une théorie de la causalité doit être compatible avec l'indéterminisme
4. la théorie devrait être indépendante du temps de manière à être consistante avec une théorie causale du temps

Dans la théorie de Salmon, la causalité est une caractéristique des processus continus plutôt qu'une relation entre événements. Cette théorie fait appel à deux éléments : la production et la propagation de l'influence causale, la propagation étant assurée par un processus causal. Chez Salmon, un processus est toute chose qui exhibe une constance structurelle au cours du temps. Pour distinguer entre processus causal et pseudo-processus, Salmon utilise le critère de transmission des marques : un processus est dit causal si il peut transmettre une modification locale de structure (une marque). Le **principe de transmission des marques** (principe MT) s'énonce comme suit :

Principe de transmission des marques (MT)

« MT : Soit P un processus qui, en l'absence d'interaction avec d'autres processus resterait uniforme par rapport à une caractéristique Q qu'il exhiberait continûment sur un intervalle spatio-temporel contenant les points A et B. Alors, une marque (qui consiste en la modification de Q en Q^*) qui a été introduite dans le processus P au moyen d'une seule interaction locale au point A, est transmise au point B si et seulement si P exhibe la modification Q^* en B ainsi qu'en tout les points situés entre A et B sans qu'aucune autre interaction ne soit nécessaire. »

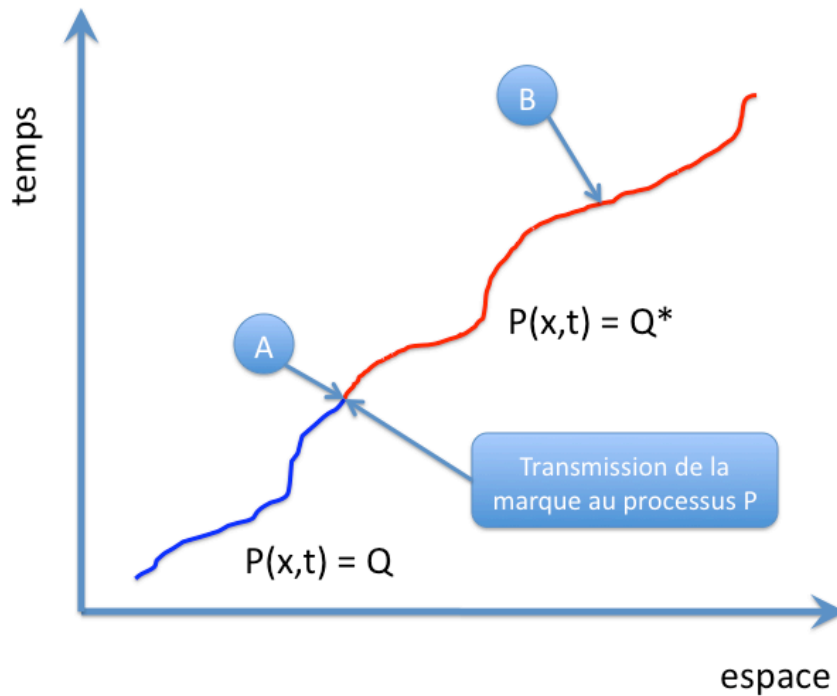


Figure 81 schéma de la transmission de marque

Le principe MT décrit les interactions causales, mais ne dit rien de la production du processus causal. La production causale est décrite par Salmon en terme de **fourche causale**, dont le rôle principal est celui qu'elle joue dans la production de l'ordre et de la structure des processus causaux. Les fourches causales sont des fourches statistiques, c'est-à-dire basée sur la notion de corrélations répétables, dont Salmon donne trois types :

Les fourches causales de Salmon

- La **fourche de conjonction**, dans laquelle deux processus émergent d'un type particulier de conditions initiales, et cela souvent de manière non-légaliste (c'est-à-dire que l'émergence n'est pas décrite par des lois scientifiques). Dans ce cas de figure, les deux processus sont statistiquement corrélés, ce qui peut être interprété en invoquant une cause commune.
- La **fourche d'interaction** : il s'agit de l'interaction entre deux processus qui conduit à une modification de ces deux processus, induisant une corrélation

entre eux qui ne peut pas être expliquée par une cause commune. De plus, l'interaction est gouvernée par des lois de conservation. Il s'agit ici d'une interaction causale.

- La **fourche parfaite**, définie comme la limite déterministe des fourches de conjonction et d'interaction.

Le concept d'interaction causale est défini plus précisément par Salmon comme suit :

Définition de l'interaction causale selon Salmon

« Soient P_1 et P_2 deux processus qui s'intersectent au point spatio-temporel S , qui appartient à l'histoire de ces deux processus. Soit Q une caractéristique que le processus P_1 aurait sur un intervalle contenant S si l'intersection avec P_2 n'existait pas ; Soit R une caractéristique que P_2 aurait sur un intervalle contenant S si l'intersection avec P_1 n'avait pas lieu. Alors, l'intersection de P_1 et P_2 constitue une interaction causale si P_1 a la caractéristique Q avant S et la caractéristique Q^* immédiatement après S et si P_2 a la caractéristique R avant S et la caractéristique R^* après S . »

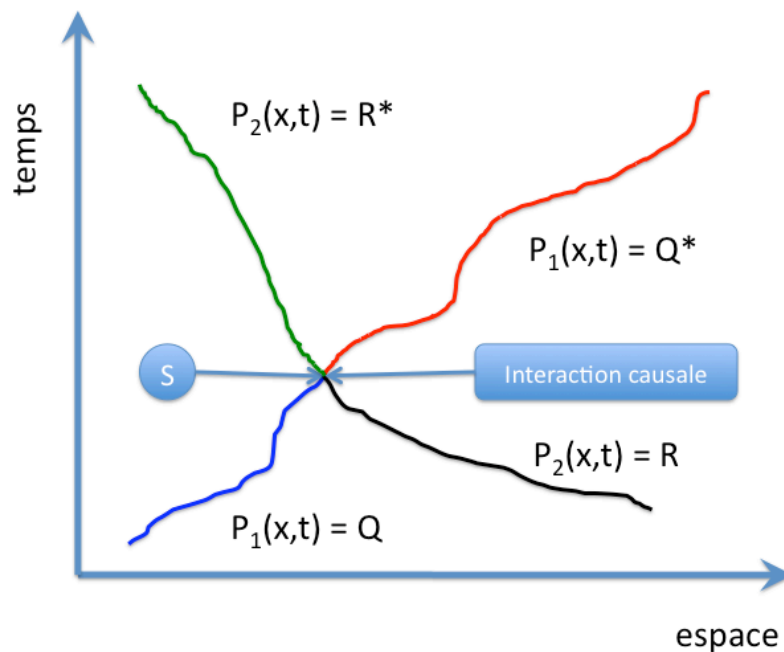


Figure 82 schéma de l'interaction causale de Salmon

La théorie de Salmon répond-elle aux critères évoqués plus haut ? Les analyses menées durant les années 1990 semblent montrer que non.

La première faille est que le principe MT exclut, dans sa formulation basée sur une certaine permanence, certains processus causaux rencontrés en physique atomique qui font intervenir des particules ayant une durée de vie très courte. Ensuite, cette définition fait appel à la notion de processus isolé, alors qu'il est reconnu qu'un tel processus ne peut exister. De plus, certains processus peuvent se propager seuls (exemple : radiation lumineuse, corps solide en mouvement), alors que d'autres ont

besoin d'interagir avec un milieu pour cela (exemple : les ondes acoustiques ont besoin d'un milieu matériel pour exister). Un autre point est que la notion de caractéristique est floue.

10.1.9 La théorie de la quantité conservée : Skyrms

Une autre théorie moderne de la causalité, appelée théorie de la quantité conservée, a été proposée en 1980 par B. Skyrms. Cette théorie est définie par les deux définitions suivantes :

Théorie de la quantité conservée

- Un *processus causal* est une trajectoire spatio-temporelle d'un objet qui possède une quantité conservée
- Une *interaction causale* est une intersection de trajectoires qui implique l'échange d'une quantité conservée.

Un processus est la trajectoire d'un objet, que celui-ci possède ou non une quantité conservée. On peut reconnaître ici une certaine similitude avec les théories de Helmholtz qui fondait son analyse sur la conservation de l'énergie.

Un objet, pour Skyrms, est toute entité trouvée dans les sciences empiriques (exemple : champs, particules, ondes). Les quantités conservées sont indiquées par la théorie : énergie-masse, quantité de mouvement, ...

10.2 Causalité et lois scientifiques

La causalité, nous l'avons vu, est intimement liée à la perception du temps. De plus, la notion de causalité semble également induire une certaine régularité dans les phénomènes : les mêmes causes conduisent aux mêmes effets. Plus précisément, c'est cette régularité dans le lien de production causale qui permet l'expérimentation : l'expérience vise à reproduire les effets en isolant les causes, par une simplification du monde extérieur dans le dispositif de laboratoire. Cette régularité est traduite par des lois. Se pose donc la question suivante : comment se traduit la relation de causalité dans les expressions mathématiques des lois ? C'est ce que nous allons voir maintenant.

Commençons par le cas le plus simple : celui des lois déterministes qui s'expriment comme une équation différentielle, comme par exemple la seconde loi de Newton :

$$\sum \vec{F} = m \frac{d\vec{v}}{dt}$$

Cette équation mathématique lie la variation temporelle de la vitesse à la résultante des forces. Si on considère l'évolution d'un corps solide durant un petit intervalle de temps dt (et pour simplifier on se place dans le cas où la résultante des forces et la masse sont constantes au cours du temps), on obtient la relation suivante :

$$d\vec{v} = \left(\sum \vec{F} \right) \frac{dt}{m}$$

Cette relation montre que la variation de vitesse est égale au produit de la résultante des forces par dt , le tout divisé par la masse. L'interprétation causale de ce résultat est immédiate : la variation de la vitesse est causée par la force, et est proportionnelle à la longueur de l'intervalle de temps. Si l'on supprime la cause (la force), l'effet disparaît : la vitesse reste constante (c'est le principe d'inertie).



Figure 83: J.L. Lagrange

Cet exemple montre que l'interprétation causale des lois différentielles est simple : l'effet est la variation de la quantité dérivée, la cause étant représentée par le second membre de l'équation. Notons toutefois que, bien que simple pour un observateur moderne, cette interprétation causale est le fruit d'une lente maturation. Initiée par Jean le Rond d'Alembert, pour qui le problème du changement du mouvement était « *le seul problème de la dynamique* », l'interprétation causale de la formulation différentielle est finalisée par Joseph-Louis Lagrange (1736-1813). On parle de **causalité lagrangienne** ou encore de **causalité différentielle**. Newton, pour sa part, refusait de considérer des causes physiques, et ne donnait qu'une signification mathématique au terme de force qui apparaît dans les équations.

Passons maintenant à l'autre forme de loi déterministe que nous avons rencontrée dans ce cours : les lois globales de conservation. Ces lois expriment la constance d'une quantité au cours du temps. L'exemple le plus simple est la loi de Boyle-Mariotte pour les transformations isothermes. Cette loi ne fait pas apparaître de différentielle par rapport au temps, et donc l'interprétation causale n'est pas aussi immédiate que dans le cas précédent. Ce type de loi permet une explication homogène, suivant la classification de Halbwachs : l'évolution des paramètres du système est régie par cette loi. Pour chaque valeur de la pression, on déduit une valeur du volume, et réciproquement. Il y a donc bien une prédiction, mais pas de chaîne causale au sens précédent. Cette loi décrit un état d'équilibre, pas un processus de variation. Elle permet de prédire l'état final que prendra le gaz, mais ne nous renseigne pas sur la manière dont les transformations s'opéreront dans le temps, et donc ne nous dit rien sur l'enchaînement cause-effet.

Ces deux types de loi sont-ils opposés de manière irréductible ? Non. Nous allons maintenant voir qu'il est possible (mais pas toujours facile !) de passer de l'une à l'autre, et donc de passer d'une explication homogène à une explication causale.

Partons de la loi de Boyle-Mariotte (cas à température constante), $PV = K$, où K est une constante, et dérivons la par rapport au temps. Nous obtenons la relation différentielle suivante :

$$\frac{d}{dt}(PV) = 0 = V \frac{dP}{dt} + P \frac{dV}{dt}$$

Plaçons-nous dans le cas où le gaz est comprimé ou raréfié dans un tube de section constante en bougeant un piston à la vitesse constante W . La variation du volume est donc proportionnelle à W (le facteur de proportion étant l'aire de la section du tube, notée S). Récrivons maintenant notre loi :

$$\frac{dP}{dt} = -\frac{SWK}{V^2}$$

Nous avons donc obtenu une loi différentielle pour l'évolution de la pression, dont la structure mathématique est similaire à celle de la seconde loi de Newton, et dont l'interprétation causale est réalisée de la même manière. Qu'avons-nous gagné en réalisant cette opération ? Nous avons maintenant à notre disposition une loi différentielle, qui nous indique à chaque instant comment évolue notre gaz. Mais il reste une trace de l'origine de notre loi différentielle : elle implique que notre gaz, lors de la compression, passe par une succession continue d'états d'équilibre. Ceci est-il vrai ? Oui, pour un observateur qui ne s'intéresse qu'à l'état du gaz entre deux instants séparés par des temps suffisamment longs. Le terme « suffisamment » est à comprendre de la manière suivante : l'état d'équilibre est atteint lorsque les mécanismes physiques qui propagent l'information depuis la surface du piston jusqu'au bout du tube ont eu le temps de rétablir un équilibre associé à la nouvelle position du piston (ce sont des ondes de pression : ondes acoustiques et ondes de choc). La validité de notre loi, qui n'est valable que pour les *transformations* dites *quasi-statiques*, est donc une question d'échelle d'observation.

Il est également possible de déduire des lois de conservation sous forme de relations intégrales, en utilisant des principes variationnels. C'est la base des mécaniques lagrangienne et hamiltonienne, dont les développements sont trop complexes pour être présentés ici. Il est même possible de déduire les lois de Newton à partir de concepts plus généraux de la physique : le principe de moindre action et les groupes de symétrie (groupe de Galilée). Cette possibilité a été abordée au chapitre 8.1.3.

Les exemples précédents posent la question suivante : peut-on ignorer la cause d'un phénomène tout en le décrivant par des lois ?

Cette question a été au cœur de la controverse qui a opposé Gottfried Leibniz (1646-1716) et Huygens à la fin du XVII^e siècle concernant l'analyse du choc des boules de billard. Cette question remonte en 1668, année où la Royal Society de Londres sollicite plusieurs savants pour résoudre le problème des lois de la transmission du mouvement lors des percussions. Huygens, dans sa lettre dans le *Journal des Sçavans* du 18 mars 1669, énonce sept lois, dont l'une préfigure la loi de conservation de l'énergie cinétique du système global. Mais le point central est qu'il suppose que les boules demeurent absolument rigides lors du choc et ne subissent aucune déformation. Cette hypothèse heurte la sensibilité de Leibniz, Mariotte et d'autres, pour qui seul le « *grand principe du ressort est la cause véritable du choc des corps.* »⁹² Mais ce qu'il est important de noter, c'est que Leibniz est parfaitement d'accord sur la loi proposée par Huygens concernant la conservation de l'énergie cinétique globale. C'est l'interprétation des mécanismes physiques, donc de la cause du mouvement de la seconde boule, qui est remise en cause. La controverse entre les deux savants ira en s'amplifiant jusqu'à la mort de Huygens en 1695, et cette question agitera les savants pendant encore plusieurs décennies. Elle montre que l'établissement des lois de la dynamique ne nécessite nullement d'être capable d'expliquer les phénomènes sous-jacents, c'est-à-dire d'en donner les causes.

⁹² Leibniz, 1692

L'exemple des boules de billard est instructif sur un autre point. Que peut-on, à partir des lois de Newton, déduire de la physique du choc ? Regardons la seconde loi de Newton, qui s'applique à tout instant à la seconde boule. Avant le choc, elle est au repos et son accélération est nulle. Après le choc, elle est en mouvement rectiligne uniforme, et son accélération est nulle. C'est pendant le choc que son accélération est non-nulle et qu'elle acquiert sa vitesse. La seconde loi fait apparaître la dérivée de la vitesse : cette quantité est donc continue, et même dérivable. En conséquence, le choc n'est pas instantané : il dure un temps très court, pendant lequel la résultante des forces sur la seconde boule croît avant de redevenir nulle (l'accélération n'est donc pas infinie). Cette durée dans le temps traduit l'existence de mécanismes à l'échelle atomique qui sont responsables de la transmission des efforts, mais ne les décrit pas (et les savants qui ont formulé les lois de la dynamique étaient bien loin de connaître ces mécanismes à l'échelle atomique). L'explication bathyène n'est donc pas nécessaire. Rappelons que, comme on l'a vu au chapitre 6, l'explication bathyène pour les chocs élastiques de sphères solides ne sera fournie qu'environ deux siècles plus tard, en 1881, par Heinrich Rudolf Hertz, dans le cadre de la mécanique des milieux continus.

Venons-en maintenant à la causalité probabiliste et aux lois statistiques, qui sont la base de la physique statistique et de la mécanique quantique. La causalité probabiliste a déjà été évoquée lors de l'analyse de l'explication statistique (chapitre 6.1). Une analyse de la causalité probabiliste est due à Dov Hugh Mellor. Pour ce philosophe, l'énoncé « A parce que B » a un sens causal probabiliste si et seulement si les conditions suivantes sont remplies :

Causalité probabiliste selon Mellor

1. A et B sont des faits.
2. **Exigence de grande probabilité** : la probabilité de A suivant B est plus grande que la probabilité de A suivant non-B.
3. **Augmentation de probabilité** : plus la probabilité que B soit réalisé augmente, plus celle que A le soit également augmente aussi.

Comme toutes les autres théories de la causalité, celle de Mellor n'est pas parfaite et peut être mise en défaut par des contre-exemples soigneusement construits. Un contre-exemple paradoxal pour la troisième condition a été proposé par Dorothy Edgington en 1997⁹³ : « *Nous sommes à la chasse au cerf à deux, nous n'avons qu'un fusil pour nous deux, et il ne nous reste plus de munitions que pour un seul coup. Un cerf apparaît. Tu es le tireur expert, moi le débutant inutile. Si je ne tire pas, tu le feras. Je tire, ce qui baisse la probabilité que le cerf soit tué. Par chance, je le tue. Mon tir a causé sa mort alors que, dans les circonstances, il aurait eu une plus grande chance d'être tué si je n'avais pas tiré.* » On arrive ici à la conclusion contre-intuitive que la cause (mon tir) est associée à une réduction de la probabilité de l'effet (la mort du cerf). La théorie de Mellor peut être améliorée pour éliminer cet exemple, mais d'autres contre-exemples plus sophistiqués peuvent à leur tour être proposés. En

⁹³ Cet exemple est traduit de l'anglais et cité dans le livre « la philosophie des sciences au XX^e siècle », A ; Barberousse, M. Kistler, P. Ludwig, Flammarion

conclusion, il faut retenir que, tout comme la causalité déterministe, la causalité probabiliste est un problème ouvert.

10.3 Le déterminisme

La notion de déterminisme physique s'est développée parallèlement à celle de causalité, avec laquelle elle est souvent confondue. Le concept de déterminisme n'est pas équivalent à celui de causalité, car il stipule quelque chose de complémentaire. Alors que le concept de causalité introduit une certaine régularité dans les phénomènes (enchaînement causal associé à un ordre d'intervention dans le temps), le déterminisme classique ajoute l'idée que le futur est entièrement déterminé par le passé.

L'idée du déterminisme a germé au fur et à mesure que se développait la science classique basée sur des lois mathématisées (XVII^e-XVIII^e siècles). Le déterminisme exprime l'idée que les faits sont gouvernés par une nécessité aveugle, ce qui est en rupture totale avec la vision héritée des grecs anciens. En effet, si détermination il y avait chez les philosophes de l'Antiquité comme Aristote, elle était basée sur une conception finaliste de l'univers (ce qui est traduit par l'idée de cause finale vue précédemment) : chaque être (humain : les grecs n'imaginaient pas que les choses aient un destin) a une destinée (*nemesis*) fixée par le dieu du destin (la *Moirai*) auquel il ne peut échapper. Ce destin fixe la durée de sa vie et les principaux événements qui s'y dérouleront (Œdipe tuera son père avant d'épouser sa mère) ; les dieux peuvent intervenir dans le cours du déroulement de cette destinée. Son caractère inéluctable en fait un **fatalisme**. « *Pour l'esprit grec, la prédétermination a une teinte religieuse et s'exprime plutôt par le concept de destin que par celui de cause. L'origine du fatalisme est anthropomorphique et explicable seulement par la projection naïve des valeurs humaines et des formes d'action humaines dans le cours de la nature.* » (H. Reichenbach, *L'avènement de la philosophie scientifique*).

Avant d'aller plus loin, notons que le terme **déterminisme** est d'usage assez récent. Il est utilisé au XIX^e siècle par certains philosophes allemands. L'idée, nous venons de le voir, est plus ancienne. On la retrouve chez Leibniz sous les termes de **détermination** et de **raison déterminante**. La doctrine de la nécessité causale chez Leibniz repose sur le **principe de raison suffisante**, qui fait que les choses sont telles qu'elles sont et non autrement. La raison suffisante apparaît également chez Pierre-Simon Laplace (voir ci-dessous), avec la signification du déterminisme.

10.3.1 Le déterminisme selon Laplace

En reprenant la causalité différentielle de Lagrange, on peut dire que le déterminisme y ajoute la notion de condition initiale. Cette conception du déterminisme est illustrée dans un texte célèbre de Pierre-Simon Laplace (1749-1827), qui fonde le **déterminisme laplacien**, qui fut le modèle conceptuel dominant jusqu'à l'avènement de la mécanique quantique et de la théorie du chaos. Elle sera notamment



Figure 84: P.S. Laplace

à l'origine des positions « déterministes » de Claude Bernard et Henri Poincaré. Le texte de Laplace se trouve au début de son *Essai philosophique sur les probabilités*, écrit en préface à la seconde édition de sa *Théorie analytique des probabilités* qui paraît en 1814 : « Tous les évènements, ceux mêmes qui par leur petitesse semblent ne pas tenir aux grandes lois de la nature, en sont une suite aussi nécessaire que les révolutions du Soleil. Dans l'ignorance des liens qui les unissent au système entier de l'univers, on les a fait dépendre des causes finales ou du hasard suivant qu'ils arrivaient et se succédaient avec régularité ou sans ordre apparent ;

mais ces causes imaginaires ont été successivement reculées avec les bornes de nos connaissances, et disparaissent entièrement devant la saine philosophie,

qui ne voit en elles que l'expression de l'ignorance où nous sommes des véritables causes.

Les évènements actuels ont avec les précédents une liaison fondée sur le principe évident, qu'une chose ne peut pas commencer d'être sans une cause qui les produise. Cet axiome, connu sous le nom de principe de raison suffisante, s'étend même aux actions que l'on juge indifférentes. La volonté la plus libre ne peut sans un motif déterminant leur donner naissance ; car si, toutes les circonstances de deux positions étant exactement semblables, elle agissait dans l'une et s'abstenait d'agir dans l'autre, son choix serait un effet sans cause ; elle serait alors, dit Leibniz, le hasard aveugle des épicuriens. L'opinion contraire est une illusion de l'esprit, qui, perdant de vue les raisons fugitives du choix de la volonté dans les choses indifférentes, se persuade qu'elle s'est déterminée d'elle-même et sans motifs.

Nous devons donc envisager l'état présent de l'univers comme l'effet de son état antérieur et comme la cause de celui qui va suivre. Une intelligence qui, pour un instant donné, connaîtrait toutes les forces dont la nature est animée et la situation respective des êtres qui la composent, si d'ailleurs elle était assez vaste pour soumettre ces données à l'analyse, embrasserait dans la même formule les mouvements des plus grands corps de l'univers et ceux du plus léger atome ; rien ne serait incertain pour elle, et l'avenir, comme le passé, serait présent à ses yeux. L'esprit humain offre, dans la perfection qu'il a su donner à l'astronomie, une faible esquisse de cette intelligence. Ses découvertes en mécanique et en géométrie, jointes à celle de la pesanteur universelle, l'ont mis à portée de comprendre dans les mêmes expressions analytiques les états passés et futurs du système du monde. En appliquant la même méthode à quelques autres objets de ses connaissances, il est parvenu à ramener à des lois générales les phénomènes observés et à prévoir ceux que les circonstances données doivent faire éclore. Tous ses efforts dans la recherche de la vérité tendent à le rapprocher sans cesse de l'intelligence que nous venons de concevoir, mais dont il restera toujours infiniment éloigné. »⁹⁴

⁹⁴ Il a été récemment évalué une limite au calcul dans l'univers depuis sa création. Celle-ci est de l'ordre de 2^{130} bits, ce qui signifie que tout ce qui nécessite un volume d'information plus grand n'aurait pas pu être calculé.

Notons qu'une position très similaire avait été développée près de soixante ans auparavant par Jean le Rond d'Alembert vers 1750 pour répondre à la question mise au concours par l'Académie de Berlin de savoir si « *les lois du mouvement des corps sont de nature nécessaire ou contingente.* » Dans l'Encyclopédie, d'Alembert développera également l'idée que « *l'Univers, pour qui saurait l'embrasser d'un seul point de vue, ne serait, s'il est permis de le dire, qu'un fait unique et une grande vérité* ».

Avant de discuter d'autres conceptions du déterminisme, revenons sur les caractéristiques du déterminisme laplacien. Il est :

- **Causal et légaliste** : la régularité des phénomènes provient, dans la conception de Laplace, de l'existence de lois. Tous les phénomènes font apparaître des lois. Par opposition, on peut imaginer un univers qui ne soit pas soumis à la causalité, ou encore qui soit causal mais pas légaliste.
- **Différentiel** : l'Intelligence de Laplace est à même de prédire le futur en intégrant les lois associés aux phénomènes, qui sont exprimées sous forme d'équations différentielles. On postule donc que toutes les lois sont différentielles (**postulat du caractère différentiel**). Cette conception est donc très influencée par la mécanique de Newton et ses succès. Par opposition, on peut imaginer un univers causal, légaliste, mais dont les lois ne prennent pas la forme d'équations différentielles, et qui donc ne se prêtent pas à la prédiction conçue comme la détermination de l'état de l'univers à l'instant $t+dt$ à partir de sa connaissance à l'instant t .
- **Universel** : c'est l'ensemble de l'Univers qui est le champ de la prévision pour Laplace. Pour que cela soit possible, cela implique que l'univers est isolé : puisque seul l'univers est connu, pour que l'efficacité prédictive existe, il est nécessaire qu'aucun élément extérieur n'interagisse. Par opposition, on pourrait faire l'hypothèse d'un univers qui serait composé de parties isolées les unes des autres, et qui serait soumis à un déterminisme différent dans chacune de ses parties (ou même pour lequel on trouverait des lois différentes dans chaque partie). Le caractère universel implique que les lois naturelles soient valables en tout point de l'espace et en tout temps (**postulat de la relativité du temps et de l'espace**). En conséquence, l'espace et le temps, ou plus précisément la position dans l'espace et le temps, ne doivent pas apparaître directement dans l'expression mathématique des lois (c'est le **principe de Maxwell**).
- **Temporel** : c'est la connaissance des lois et de l'état à un instant donné qui sont la base de la connaissance de l'Intelligence imaginée par Laplace. Celle-ci est donc une extrapolation de l'intelligence humaine : son mode de connaissance de l'univers est quantitativement différent, mais non qualitativement. Ceci est traduit par le caractère légaliste et différentiel de la causalité laplacienne. Ici aussi, on peut imaginer des solutions différentes, comme par exemple une Intelligence qui aurait une perception inversée par rapport à la nôtre entre espace et temps. Pour cette intelligence, c'est la donnée de l'état en un point qui serait nécessaire pour la prédiction, et les lois qu'elle pourrait écrire prendraient la forme d'équations différentielles en espace. On peut également imaginer une intelligence « hypercausale », qui aurait une connaissance naturelle de l'espace et du temps. Pour cette dernière, l'univers

ne serait qu'un objet réalisé en une seule fois, et les notions de loi, de causalité, et de déterminisme n'auraient d'ailleurs plus de sens.

- **Incompatible avec le hasard** : la capacité de prédiction de l'Intelligence est infinie, tant dans la durée que dans la précision. Donc, aucun phénomène n'échappe aux lois et à leur intégration temporelle. Pour que cela soit possible, il est nécessaire que la condition initiale soit connue de manière exhaustive et exacte, et on fait donc l'hypothèse qu'une telle connaissance est possible (**postulat de certitude**).

10.3.2 L'analyse de Kojève

L'idée de déterminisme en physique et en mécanique a été minutieusement analysée par le philosophe français Alexandre Kojève (1902-1968) dans son livre posthume *L'idée du déterminisme* (1990). Partant du déterminisme de Laplace, Kojève identifie plusieurs structures causales possibles pour l'univers :

1. La **structure causale exacte**, qui correspond à celle du déterminisme laplacien, qui est un **déterminisme causal exact** : chaque phénomène est entièrement et strictement déterminé par le passé et peut être décrit par un ensemble de lois. Laplace indique alors qu'il est possible de prédire, avec une précision théoriquement a priori infinie, chaque événement futur, ce que Kojève démontre comme étant impossible. En effet, considérons les deux variantes possibles de la conception laplacienne du déterminisme : le déterminisme à l'échelle de l'univers (déterminisme universel), ou alors sa restriction à un sous-espace fermé (déterminisme local).

Commençons par le cas du **déterminisme universel**, et interrogeons-nous sur ce qu'il implique pour un scientifique. Tout d'abord, la connaissance exacte et complète d'un univers infini est par définition impossible, puisqu'il est infini. Dans le cas d'un univers fini, la connaissance exacte et complète est également impossible en pratique : il faut un support matériel pour les informations, qui devrait donc être extérieur à l'univers pour représenter complètement celui-ci. Enfin, dans les deux cas, la connaissance exacte attendue par Laplace (vitesse et position de toutes les particules) est aujourd'hui reconnue comme impossible en vertu du principe d'indétermination de Heisenberg (voir le chapitre 11). La connaissance exacte et complète est donc impossible en pratique. Mais ce n'est pas le seul problème. Le déterminisme tel qu'il est imaginé par Laplace repose sur l'idée que la prédiction des événements futurs est possible, et cela avec une précision infinie, ce qui implique une connaissance infiniment précise de la condition initiale (l'univers à l'instant t). Ce besoin d'une précision infinie implique que toutes les interactions, aussi faibles soient-elles, entre les parties de l'univers, doivent être prises en compte. Deux états de l'univers ne sont égaux au sens qu'ils mèneront au même futur (c'est-à-dire qu'ils induiront les mêmes effets) que s'ils sont strictement identiques en tout point. La prédiction parfaite n'est donc possible que dans le cas d'un univers cyclique, soumis à un Eternel Retour.

Intéressons-nous maintenant au cas d'un univers ayant une **structure causale exacte locale**. Les objections précédentes liées au caractère infini de l'Univers

tombent, mais sont remplacées par d'autres. Tout d'abord, si chaque partie de l'univers est isolée des autres, un scientifique ne peut connaître que la partie dans laquelle il se trouve, les autres n'existent pas pour lui. Il se trouve confronté au cas d'un déterminisme causal exact universel pour sa partie d'univers, qui doit, ainsi que nous l'avons vu, être strictement cyclique pour permettre le déterminisme laplacien. Dans le cas où les portions d'univers interagissent, quelles sont les conditions nécessaires pour permettre une prédiction parfaite ? Un univers parfaitement cyclique remplit ces conditions. Sinon, il est nécessaire de connaître les actions exercées par les autres portions d'univers pendant la durée sur laquelle on veut effectuer la prédiction (disons entre les temps $t1$ et $t2$), l'état de notre portion d'univers à l'instant initial (ici $t1$) étant par ailleurs parfaitement connu. Une première solution est celle où les sous-univers sont tous identiques, et l'influence des autres univers peut être déduite de la connaissance de notre propre portion d'univers par symétrie. Si cette symétrie n'existe pas, on voit qu'il est nécessaire de connaître le futur (les actions que vont exercer les autres portions d'Univers pendant le temps sur lequel on veut réaliser la prédiction). La prédiction stricte est donc impossible : le déterminisme causal exact local ne permet pas la prédiction du futur, mais l'explication du passé.

2. La **structure causale approchée**, qui repose sur une relaxation de la contrainte d'exactitude qui apparaît dans le déterminisme laplacien (déterminisme causal exact). Alors que ce dernier repose sur le principe « *mêmes causes, mêmes effets* », la structure causale approchée est basée sur le principe « *causes semblables, effets semblables* ». La notion d'identité est remplacée par celle de similitude, qui permet d'intégrer les problèmes de précision des mesures et de négligence des interactions très faibles. C'est cette version du déterminisme, le déterminisme causal approché, qui est couramment utilisé. Notons que les relations entre structure causale exacte et structure causale approchée sont complexes, car elles sont relativement indépendantes. La causalité approchée n'implique pas la causalité exacte, et la causalité exacte n'implique la causalité approchée que si on postule l'existence d'une certaine continuité dans les phénomènes. Notons également que ces deux structures causales sont compatibles : un univers peut présenter soit l'une, soit l'autre ou encore les deux à la fois. La structure causale approchée permet de considérer des sous-systèmes comme étant isolés, et donc de prédire leur évolution à partir de données partielles de précision finie. Mais de telles **prédictions** ne peuvent être qu'**approximatives**.
3. Le **déterminisme statistique**. Alors que les déterminismes causaux reposent sur la description de chaque particule de la portion d'univers considérée, le déterminisme statistique repose sur la description et la l'évolution des propriétés statistiques d'un ensemble d'objets qui demeurent indifférenciés dans cette description. Une telle description de l'univers repose sur des lois statistiques, qui n'ont pas de sens pour un objet unique. Il est à noter que la description statistique peut être réalisée à partir d'objets identiques ou semblables, et que donc elle s'accommode de mesures de précision finie. La prédiction est elle aussi statistique : elle consistera par exemple à dire qu'au terme de l'évolution considérée d'un certain système physique, tel pourcentage de ses composants aura acquis la propriété X pendant que le reste aura acquis la propriété Y. Mais la prédiction ne dit pas ce que sera la

propriété d'un composant donné : on peut au mieux dégager une probabilité que le composant soit de type X ou Y.

Ce type de description statistique est généralement employé pour effectuer un changement d'échelle dans la description, comme dans le cas de la théorie cinétique des gaz : les propriétés macroscopiques du gaz (pression, masse volumique, température) sont définies à partir des propriétés statistiques de l'ensemble formé par les molécules qui le composent.

Nous avons vu que le déterminisme causal repose sur un certain nombre de postulats. Quels sont les postulats qui sous-tendent le déterminisme statistique ? C'est ce que nous allons voir maintenant.

Tout d'abord, pour que des lois statistiques puissent exister, il est nécessaire que l'univers possède une structure statistique, c'est-à-dire que les phénomènes qui s'y déroulent puissent être soumis au calcul des probabilités. La théorie mathématique des probabilités est « vraie » en soi, et n'entretient aucun rapport avec le monde sensible. Pour que ces résultats puissent être interprétés de la manière dont nous le faisons quotidiennement sous forme de prédictions statistiques, deux conditions doivent être réalisées :

1. Quelque chose dans le monde réel doit correspondre à la notion de probabilité
2. Les probabilités calculées au moyen des observations expérimentales doivent coïncider avec celles de la théorie mathématique.

Que cela implique-t-il ? Tout d'abord, rappelons que les statistiques sont établies en considérant des ensembles de réalisations souvent appelés séries. Les séries expérimentales sont par construction finies, et portent sur des observations. Les séries qui apparaissent dans la théorie sont le plus souvent infinies, et portent sur des réalisations identiques au vu des critères sélectionnés. Une notion très importante est celle de **convergence** : le pourcentage de réalisations possédant un caractère donné doit converger vers une valeur donnée lorsque le nombre de réalisations de la série augmente, pour tendre vers sa valeur asymptotique. Une condition est donc que les séries expérimentales soient convergentes. Notons que l'on peut imaginer des univers dans lesquels les séries ne sont pas convergentes au sens donné plus haut. Une seconde condition théorique est la vérification du **principe du système du jeu exclu** : la valeur des statistiques ne dépend pas du choix des éléments que l'on peut faire dans la série.

Quels sont les liens entre le déterminisme causal et le déterminisme statistique ?

Commençons par le déterminisme causal exact. Pour assurer la compatibilité, il faut restreindre un peu le déterminisme statistique. En effet, dans sa version « complète », celui-ci n'interdit pas que deux causes identiques produisent des effets différents. La structure causale exacte implique que les probabilités, dans certains cas, ne peuvent être que 0 ou 1, suivant que les causes sont identiques ou non. Pour assurer la compatibilité, et obtenir des probabilités comprises entre 0 et 1, il est donc nécessaire de considérer que l'analyse statistique ne peut s'appliquer qu'à des séries composées d'éléments différents sur le plan causal.

La compatibilité avec le déterminisme causal approché est naturelle et ne pose pas de question supplémentaire.

Au vu de ces discussions, il apparaît que le déterminisme statistique est compatible avec le déterminisme causal (exact ou approché). L'univers peut donc être décrit au moyen de ces deux déterminismes. Mais il faut noter qu'ils ne s'impliquent pas mutuellement : on admet généralement que l'existence d'une structure causale n'est pas une condition suffisante pour l'existence d'une structure statistique, et que l'existence de la seconde n'est pas une condition nécessaire de celle de la première. Mais cela n'implique bien évidemment pas que ces deux déterminismes soient complètement indépendants.

10.3.3 Légalisme, déterminisme, incomplétude et indéterminisme

Le déterminisme de Laplace repose sur l'hypothèse que tout système qui peut être décrit par un système de lois différentielles et des conditions initiales est prévisible, au sens où ses états futurs peuvent être prédits.

La question de savoir si un système physique évolue de manière déterministe peut être posée de manière générale. Mais il est impossible de dire dans l'absolu si un système est intrinsèquement déterministe ou non. *L'observation permet seulement de dire que, sur une certaine période, son évolution est décrite de manière satisfaisante par une théorie déterministe.*

Depuis les travaux de Henri Poincaré sur les trajectoires de certains astres du système solaire (notamment la lune), on sait que l'évolution de tous les systèmes décrits par des lois différentielles déterministes ne peut pas être prévue par le calcul mathématique avec une précision arbitrairement grande. Les solutions de certains systèmes d'équations différentielles sont si sensibles que, quel que soit le degré de précision (fini) avec lequel la condition initiale est connue, deux solutions initialement « équivalentes » finiront par diverger. Ce sont les systèmes dits **chaotiques**. Un exemple simple, basé sur la mécanique newtonienne, est le problème des trois corps. La causalité différentielle n'est donc pas une garantie du déterminisme exact. Par contre, les systèmes chaotiques peuvent avoir une certaine régularité statistique. Le déterminisme statistique peut donc s'appliquer, au moins pour certains d'entre eux (exemple : les gaz parfaits).

Les systèmes qui n'évoluent pas de manière déterministe sont appelés **systèmes indéterministes**. Il est en de même pour les théories physiques. Une **théorie indéterministe** est une théorie qui ne permet pas d'arriver à une description déterministe satisfaisante d'un système physique. Il faut alors distinguer deux cas fondamentalement différents. Pour cela, considérons un système physique qui évolue selon une trajectoire A ou une trajectoire B, et une théorie qui ne permet pas de dire quelle trajectoire est suivie par le système.

Dans le premier cas de figure, le système évolue réellement suivant A ou B, mais pas les deux à la fois, et la théorie n'est pas assez précise ou complète pour distinguer entre les deux options. Mais elle pourrait être améliorée, ou les données rendues plus

précises, pour donner une réponse satisfaisante. Le problème rencontré ici est un **problème d'incomplétude** : c'est le manque de précision des données et/ou des lois qui est en cause.

Dans le second cas de figure, le système n'évolue pas selon A ou B, mais selon les deux à la fois. La théorie indéterministe reflète alors un indéterminisme plus fondamental, et nous sommes conduits à conclure que le monde réel est indéterministe. C'est par exemple le cas de la mécanique quantique, qui admet l'existence d'états superposés pour un système physique.

10.4 Le hasard

Venons-en maintenant au troisième concept discuté dans ce chapitre : le hasard. Selon la tradition aristotélicienne, le hasard n'entre pas dans le champ de la science, car cette dernière ne traite de faits réguliers universels. Cette conception l'emportera jusqu'au XVII^e siècle. Il est également exclu du champ de la physique par le déterminisme laplacien. Il retrouve toute sa place depuis que celui-ci s'est retrouvé mis en défaut par la mécanique quantique et la théorie du chaos, qui ont mis à mal l'idée d'une prédictibilité telle que Laplace l'imaginait. La théorie du hasard doit beaucoup au développement des théories mathématiques des statistiques et des probabilités, tout d'abord avec les travaux de Fermat et Pascal sur les jeux de hasard, puis au XIX^e siècle avec les travaux sur les erreurs de mesure⁹⁵, les statistiques sociales et la mécanique statistique (théorie cinétique des gaz, thermodynamique), et au XX^e siècle avec la naissance de la mécanique quantique et l'axiomatisation vers 1930 de la théorie des probabilités par des mathématiciens tels que A.N. Kolmogorov.

Le concept de hasard revêt des significations différentes, qui vont être discutées plus bas. Un point commun est que le hasard apparaît toujours comme une cause à travers ses effets : c'est à travers l'observation de phénomènes « inexplicables », ou plutôt « imprévisibles », qu'il se manifeste. La question centrale est de savoir si le hasard, qui apparaît comme un défaut de la capacité prédictive des théories causales légalistes, traduit l'existence d'un indéterminisme réel, ontologique, ou s'il n'est que le reflet de l'imperfection des théories scientifiques et des méthodes mathématiques.

La description de l'évolution d'un système physique par des lois probabilistes soulève la question de l'interprétation de ces lois. Associer une probabilité aux divers états que peut prendre un système physique signifie-t-il que ce système va passer par tous ces états, si l'on attend suffisamment longtemps ? Répondre oui à cette question revient à faire appel à ce que l'on appelle la **théorie du hasard objectif** : les probabilités mesurent bien la possibilité de réalisation d'évènements réels.

⁹⁵ On peut notamment citer les travaux de certains mathématiciens comme Legendre, Laplace et Gauss concernant l'exploitation des données géographiques recueillies par Delambre et Méchain pendant l'expédition de la méridienne (1792-1799) qui avait pour but d'établir la longueur du méridien entre Dunkerque et Barcelone afin de définir un mètre étalon. Ces travaux conduisirent à la naissance de la théorie des erreurs, entre 1805 et 1811.

Ceci est illustré par la thermodynamique statistique, née au XIX^e siècle, et la notion d'irréversibilité. Le second principe de la thermodynamique indique que l'entropie d'un système physique isolé ne peut pas diminuer au cours du temps, ce qui détermine la flèche du temps. Ce principe, comme nous allons le voir, est basé sur un résultat « moyen » (au sens statistique). L'entropie peut être comprise comme une mesure du désordre (au sens par exemple où les atomes dans un cristal sont ordonnés suivant un motif régulier dans l'espace, et détruire le cristal revient à augmenter l'entropie, car les atomes sont alors répartis de manière moins ordonnée). Le sens commun nous indique de très nombreux exemples de transformations dites irréversibles : placé dans un verre d'eau chaude, un glaçon fond ; on n'observe jamais que le glaçon devienne plus froid et l'eau plus chaude. Ceci est dû au fait qu'il existe beaucoup plus de manières de répartir les molécules du liquide et du glaçon dans des états qui correspondent à ce que l'on observe (le glaçon fond et l'eau est plus froide) que dans l'état « improbable » jamais observé. Mais cela ne veut pas dire que l'état improbable est impossible. La théorie du hasard objectif nous dit que cet état est possible, mais qu'il a très peu de chance de se réaliser. Si on admet qu'un système physique passe par tous les états admissibles et que l'on attend très longtemps, on observera au moins une fois l'état improbable (on calcule, pour l'exemple du verre d'eau, qu'il faudrait attendre plusieurs milliards d'années pour que la probabilité que cet état se soit réalisé soit proche de 1). L'irréversibilité est donc à comprendre en un sens statistique : elle indique que les systèmes observés sont dans des états de forte probabilité, mais elle n'exclut pas dans l'absolu l'occurrence d'états improbables.

10.4.1 Le hasard selon Aristote

Afin de cerner le concept de hasard, Aristote établit plusieurs distinctions concernant les causes, qui toutes font intervenir le concept de hasard.

La première distinction porte sur la **cause essentielle** et la **cause accidentelle**. La cause essentielle (ou cause par soi) est celle dont l'effet est associé à son essence, c'est-à-dire à ce qui la distingue de toute autre chose et qui fait qu'elle est ce qu'elle est, et qui régit son mode d'action. Par opposition, la cause accidentelle ne produit pas d'effet correspondant à son essence. Aristote illustre cette distinction par les exemples suivants. L'art et le savoir-faire de l'architecte sont la cause par soi de la maison qu'il construit. La chute d'une pierre est la cause accidentelle de la blessure de celui qui la reçoit.

La seconde distinction porte sur trois catégories d'événements : ceux qui se produisent toujours d'une manière identique et qui sont donc nécessaires et constants, ceux qui se produisent fréquemment ou la plupart du temps, et enfin ceux qui se produisent exceptionnellement, c'est-à-dire par accident.

La troisième et dernière distinction porte sur les événements qui se produisent en vue d'une fin, et ceux qui se produisent sans finalité. Pour mieux comprendre cette dernière distinction, il faut se rappeler que pour Aristote, tout dans l'univers, y compris les objets inanimés, agit en vue d'une fin. Par exemple, les corps pesants chutent pour rejoindre la place qui est la leur dans l'univers. Les événements nécessaires ou fréquents se produisent en vue d'une fin, alors que les événements exceptionnels n'ont pas de finalité.

Sur la base de ces éléments, Aristote distingue le hasard (mot grec **automaton**, qui signifie « qui est sa propre cause »⁹⁶) et la fortune (mot grec **tukè**) :

« [...] la fortune est une cause par accident, survenant dans les choses qui, étant en vue de quelque fin, relèvent en outre du choix. [...] la fortune et le hasard sont des causes par accident, pour les choses susceptibles de ne se produire ni absolument, ni fréquemment, et en outre susceptibles d'être produites en vue d'une fin. [...] dans le domaine des choses qui ont lieu absolument en vue de quelque fin, quand les choses ont lieu sans avoir en vue le résultat et en ayant leur cause finale en dehors de lui, alors nous parlons d'effet de hasard ; et d'effets de fortune, pour tous ceux des effets de hasard qui, appartenant au genre des choses susceptibles d'être choisies, atteignent les êtres capables de choix. »

Selon ces définitions, un événement hasardeux est un événement qui n'a pas de cause essentielle, mais un nombre indéfini de causes accidentelles. Le hasard (au sens large du terme) est donc l'ensemble de ces causes accidentelles d'un événement exceptionnel qui n'est pas la fin visée des actes qui ont conduit à son existence. L'événement accidentel (l'effet dû au hasard) se substitue ou s'ajoute à la fin visée de l'action.

Un point important est que le concept de hasard aristotélicien est étranger à celui de légalité et de prévisibilité. Revenons sur certains aspects de ces définitions, qui peuvent être vus comme des failles.

Tout d'abord, la distinction entre événement exceptionnel et événement fréquent peut prêter à confusion. Il existe des phénomènes très rares (passage d'une comète avec une période très longue) qui sont parfaitement réguliers. Ce critère est donc à manipuler avec précaution.

Ensuite, la distinction entre cause accidentelle et cause essentielle se base sur l'idée que chaque être possède des caractéristiques qui lui sont propres et qui le définissent individuellement, et qui, si elles étaient modifiées, feraient de lui autre chose. Or, en pratique, notre connaissance du monde est imparfaite et incomplète, et évolue avec le temps. Il est donc a priori impossible de garantir que nous connaissons l'essence de tel ou tel être, ce qui rend très difficile de trancher la question de savoir si une cause est accidentelle ou essentielle.

Aristote développe une vision finaliste de l'univers : toutes les actions sont orientées vers une fin. Dans le cas des actions humaines, l'existence de nos intentions rend cette analyse possible. Dans le cas des phénomènes relatifs aux objets inanimés, cela nécessite de faire l'hypothèse de l'existence d'un principe ultime (Dieu, principe de conservation, ...) qui est étrangère à la science moderne.

Enfin, l'analyse d'Aristote repose sur la possibilité d'isoler et d'identifier les différentes actions/causes responsables de la production d'un effet. Ceci pose le problème de l'individuation : ce travail de décomposition, de séparation est-il possible ? Non, le plus souvent. Il existe tout d'abord un problème pratique : le très grand nombre d'interactions auxquelles est soumis un système physique interdit le plus souvent de dresser une liste exhaustive de celles-ci. Ensuite, apparaît un problème théorique : si la distinction entre deux objets macroscopiques est réalisable,

⁹⁶ Et donc, étant sa propre cause, un tel phénomène n'est pas prévisible, puisqu'il n'est pas corrélé à ce qu'il est possible de connaître avant qu'il n'apparaisse. Un tel événement est donc étranger à toute chaîne causale.

il est en revanche parfois impossible de faire la différence entre des objets microscopiques : comment faire la différence entre deux électrons ayant la même vitesse ?

10.4.2 Le hasard selon Cournot

Augustin Cournot a proposé une définition plus satisfaisante pour le scientifique moderne du hasard, en se basant sur la notion de **séries causales indépendantes**. Pour Cournot, l'idée de hasard *« est l'idée de rencontre entre des faits rationnellement indépendants les uns des autres, rencontre qui n'est elle-même qu'un pur fait, auquel on ne peut assigner de loi ni de raison.[...]Les évènements amenés par la combinaison ou la rencontre d'autres évènements qui appartiennent à des séries indépendantes les unes des autres, sont ce qu'on nomme des évènements fortuits, ou des résultats du hasard. »* Il est intéressant de noter que le concept de hasard est ici associé aux effets plutôt qu'aux causes, comme chez Aristote. Le point clé de la définition de Cournot est la notion de séries causales indépendantes. Celles-ci *« se développent parallèlement ou consécutivement, sans avoir les unes sur les autres la moindre influence, ou (ce qui reviendrait au même pour nous) sans exercer les unes sur les autres une influence qui puisse se manifester par des effets appréciables. Personne ne pensera sérieusement qu'en frappant la terre du pied il dérange le navigateur qui voyage aux antipodes, ou qu'il ébranle le système des satellites de Jupiter ; mais, en tout cas, le dérangement serait d'un tel ordre de petitesse, qu'il ne pourrait se manifester par aucun effet sensible pour nous, et que nous sommes parfaitement autorisés à n'en pas tenir compte. Il n'est pas impossible qu'un évènement arrivé à la Chine ou au Japon ait une influence sur les faits qui peuvent se passer à Paris où à Londres ; mais, en général, il est bien certain que la manière dont un bourgeois de Paris arrange sa journée n'est nullement influencée par ce qui se passe actuellement dans telle ville de Chine où jamais les européens n'ont pénétré. Il y a là comme deux petits mondes, dans chacun desquels on peut observer un enchaînement de causes et d'effets qui se développent simultanément, sans avoir entre eux de connexion, et sans exercer les uns sur les autres d'influence appréciable. »* Cette notion de séparabilité de chaînes causales est centrale, et anticipe la critique que fera Gaston Bachelard du déterminisme laplacien universel (Bachelard rejette l'idée d'un déterminisme unique à l'échelle de l'Univers et plaide pour des déterminismes multiples, localisés, sélectifs et pas rigoureusement reliés entre eux).

On peut retenir que, pour Cournot, la notion de hasard est fortement liée à la perception que nous avons des phénomènes : c'est l'extrême faiblesse des liens entre les chaînes causales qui les rend imperceptibles et nous fait dire que ces chaînes sont indépendantes. Il ne s'agit donc pas ici d'un hasard ontologique.

11 Brève incursion dans le monde quantique

Ce chapitre a pour but de présenter, très rapidement, les modifications dans la conception du monde introduites par le développement de la mécanique quantique, ainsi que les conséquences épistémologiques de ces modifications. Il n'est absolument pas question ici de discuter les aspects « techniques » et mathématiques de la mécanique quantique, ni de faire un cours de mécanique quantique, dont de nombreux aspects ne peuvent pas être traités indépendamment du formalisme mathématique sous-jacent.

11.1 Exemple introductif : l'expérience des fentes d'Young

La différence entre la mécanique classique et la mécanique quantique est parfaitement illustrée par l'expérience des fentes de Young. Cette expérience a été réalisée pour la première fois par le physicien (mais aussi égyptologue et médecin !) anglais Thomas Young (1773-1829) en 1803, qui voulait démontrer le caractère ondulatoire de la lumière.

Le principe consiste à éclairer un panneau opaque muni d'une ou plusieurs fentes, et d'observer le signal lumineux projeté sur un écran situé derrière le panneau.

Dans le cas d'une fente unique, l'intensité lumineuse décroît régulièrement lorsqu'on s'éloigne de la position de la fente.

Dans le cas de deux fentes parallèles, on observe que l'intensité lumineuse varie de manière beaucoup plus compliquée : on voit apparaître des franges d'interférences. Ces franges, qui proviennent de l'interaction entre les faisceaux lumineux issus des deux fentes, sont parfaitement prévisibles à partir de la théorie ondulatoire de la lumière (de telles interférences sont observées avec tous les types d'onde : ondes sonores, ondes capillaires à la surface de l'eau, ...) Une remarque importante est donc que l'intensité lumineuse mesurée sur l'écran n'est pas la somme de celles mesurées lorsqu'on fait deux expériences avec des fentes uniques. Dans ce dernier cas, il n'y a pas d'interférences.

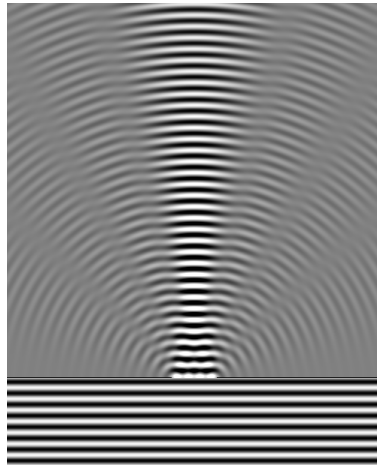


Figure 85 expérience de diffraction d'une onde par une fente unique

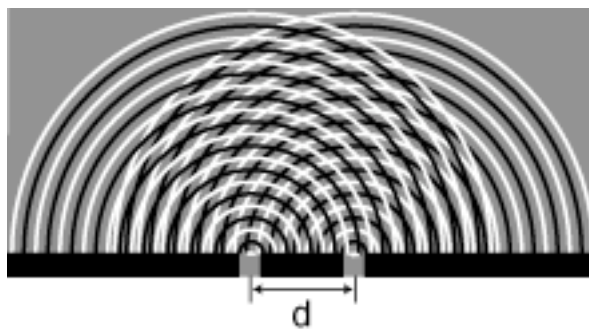


Figure 86 expérience de diffraction d'une onde par une fente double (on observe des franges d'interférence)

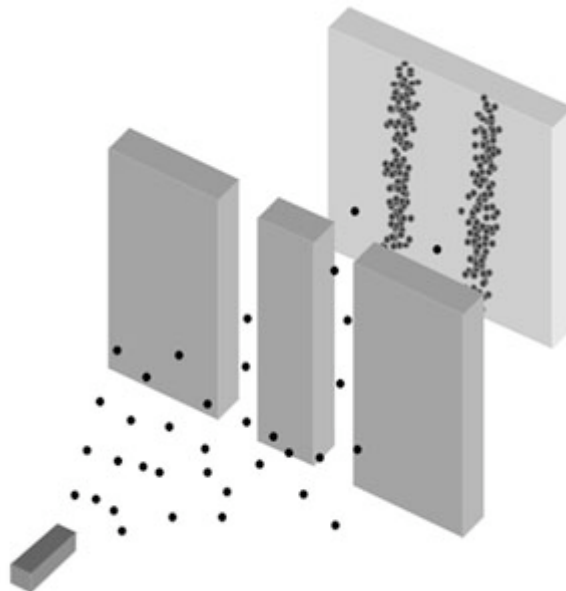


Figure 87 expérience de la fente double avec des corpuscules (pas d'interférence)

Une conclusion de cette expérience porte sur la **dualité onde-corpuscule** : si un seul photon est émis, on n'observe qu'un point lumineux sur l'écran, ce qui est compatible avec la théorie corpusculaire de la lumière. Si on émet un grand nombre de photons,

on observe les franges d'interférences, qui sont compatibles avec une théorie ondulatoire de la lumière. Notons que les franges apparaissent même si les photons sont émis successivement, un à un, et non tous en même temps. Les photons ne sont donc ni des corpuscules classiques (des billes de verre ou des balles de tennis ne produisent pas d'interférences), ni des ondes pures (un unique photon produit une seule petite tâche lumineuse sur l'écran, alors qu'une onde pure éclairerait tout l'écran): on ne peut rien conclure d'autre que le fait que ces deux modèles sont efficaces et pratiques pour décrire certains phénomènes observés. Ce **principe de complémentarité**, énoncé par Niels Bohr (1885-1962, prix Nobel de physique 1922) en 1927, est un des piliers de la mécanique quantique.

Cette expérience classique peut être répétée avec des électrons dans des conditions similaires. Les phénomènes de diffraction et d'interférences sont également observés avec des objets plus gros. L'expérience de diffraction a été réalisée en 1930-1933 avec des atomes d'hélium et des molécules d'hydrogène, et en 1976 et 1979 avec des atomes de néon (qui contiennent dix électrons, dix protons et dix neutrons). Mais pour ces objets massifs on ne peut plus se servir de fentes, car elles seraient trop fines et trop proches les unes des autres pour pouvoir les réaliser en pratique. La diffraction est obtenue en projetant les objets sur la surface d'un cristal.

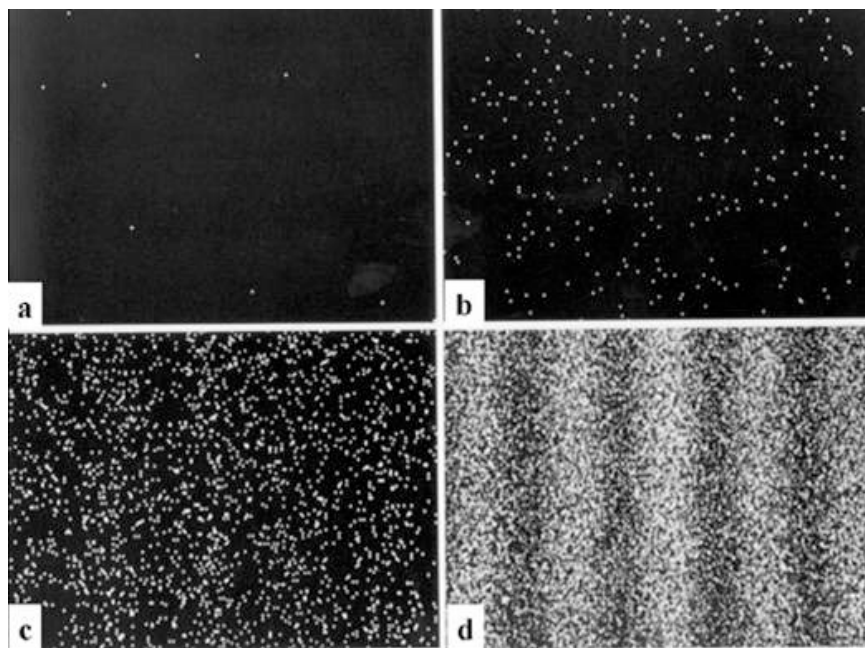


Figure 88 expérience de diffraction à 2 fentes. Les quantons (photons ou électrons) sont émis 1 par 1. Au début de l'expérience, on ne voit pas de frange, mais seulement quelques points associés chacun à un quanton, comme pour un corpuscule classique (a). A la fin de l'expérience, les franges sont apparues, comme pour une onde (d)

L'expérience est encore plus étrange si l'on souhaite maintenant savoir par quelle fente est passée chaque photon. Dans l'expérience à deux fentes décrite plus haut, on ne savait pas par quelle fente chaque photon est passé. Equipons maintenant le panneau d'un dispositif qui permet de savoir par quelle fente chaque photon émis passe (cela est réalisé en pratique avec une source lumineuse secondaire et un détecteur de photon). Cela revient à mesurer la position de chaque photon lorsqu'il passe par une fente. Les expériences avec une fente unique donnent des résultats

identiques avec et sans mesure de la position des photons au niveau de la fente, ce qui est intuitivement compréhensible. En revanche, une différence frappante est observée dans le cas d'une fente double : les interférences ont disparu ! Le résultat observé est celui que l'on obtiendrait en additionnant les résultats de deux expériences menées avec des fentes uniques : on retrouve un comportement classique pour des corpuscules, c'est-à-dire sans aucun effet quantique ! Cette modification radicale du résultat est un effet qui n'est explicable que dans le cadre de la mécanique quantique, ce que nous allons voir plus loin.

11.2 Paquet d'ondes et d'états superposés

Comment décrire les objets qui se conduisent de « manière quantique » ? La solution fut apportée par Louis de Broglie (1892-1987, prix Nobel de physique 1929) en 1923, qui propose de les décrire comme des paquets d'ondes⁹⁷. Ces paquets d'ondes évoluent en suivant l'équation trouvée par Erwin Schrödinger (1887-1961, prix Nobel de physique 1933) en 1925 et dont l'interprétation correcte fut trouvée par Max Born (1882-1970, prix Nobel de physique 1954) en 1926. Le point conceptuellement dur et fondamental de la mécanique quantique réside dans l'interprétation de ce formalisme (c'est cette interprétation qui vaudra son prix Nobel à Born). Ce que représente ce paquet d'ondes est la *densité de probabilité* de présence du quanton⁹⁸. Le point fondamental est que la mécanique quantique est intrinsèquement probabiliste, et est en cela radicalement différente de la physique classique, et cela parce que les phénomènes qu'elle étudie ne peuvent pas être décrits par une théorie déterministe. C'est ce point qui rebutera toujours Albert Einstein qui, dans sa fameuse lettre de 1926 à Max Born, écrit « *La mécanique quantique force le respect. Mais une voix intérieure me dit que ce n'est pas encore le nec plus ultra. La théorie nous apporte beaucoup de choses, mais elle nous approche à peine du secret du Vieux (Dieu). De toute façon je suis convaincu que lui, au moins, ne joue pas aux dés !* »

Revenons à l'expérience des deux fentes. Que nous dit le formalisme quantique ? Chaque quanton est décrit par une onde ou un paquet d'ondes de probabilité, qui occupe tout l'espace. Cette description ne nous indique pas où se trouve la quanton à un instant donné : la notion de trajectoire de la physique classique n'a plus de sens ici. Le point important est qu'il ne s'agit pas ici d'un problème d'incomplétude de la théorie, mais bien d'un indéterminisme plus fondamental. Il est impossible d'associer une position et une vitesse à un quanton (au sens classique), parce que celui-ci n'est pas localisé. Tant qu'il n'est pas l'objet d'une mesure, il occupe tout l'espace, avec une probabilité décrite par sa fonction d'onde associée. Donc, la question de savoir par quelle fente est passée un quanton n'a pas de réponse classique, et n'a pas de sens dans le cadre quantique : si il n'y a pas de défaut dans le montage expérimental, sa probabilité d'être passé par chacune des deux fentes est égale à 1/2. Chaque quanton est représenté par la superposition de l'état « passé par la première fente » et de l'état « passé par la seconde fente ». Ce **principe de superposition des états** est un des fondements de la mécanique quantique, et est basé sur l'emploi d'outils mathématiques adéquats (des vecteurs, dont les composantes peuvent être des

⁹⁷ Un paquet d'ondes résulte de la combinaison de plusieurs ondes, et il est localisé dans l'espace (alors qu'une onde simple emplit tout l'espace) au sens où il peut n'occuper qu'une portion limitée de l'espace. Mais il ne se réduit pas à un point matériel.

⁹⁸ Le terme quanton a été introduit par le physicien M. Bunge en 1967 pour désigner un objet au comportement quantique.

nombre complexes). Il se traduit par le fait que l'on peut *directement sommer les amplitudes de probabilité associées à chaque état possible*, qui sont des nombres complexes.

La possibilité de qu'il existe des interférences peut être illustrée de manière simple. Soit a_1 l'amplitude de probabilité qu'un quanton soit reçu sur l'écran en un point donné, lorsque seule la fente 1 est ouverte. La probabilité P_1 associée est définie comme le carré du module de a_1 , soit :

$$P_1 = a_1 a_1^* = |a_1|^2$$

De manière symétrique, lorsque seule la fente 2 est ouverte, la probabilité que le quanton soit reçu au même point sur l'écran est :

$$P_2 = a_2 a_2^* = |a_2|^2$$

Dans le cas où les deux fentes sont ouvertes, l'amplitude de probabilité est obtenue par sommation des deux précédentes : $a_{12} = a_1 + a_2$, et donc la probabilité de réception du quanton dans ce cas est

$$P_{12} = (a_1 + a_2)(a_1 + a_2)^* = |a_1 + a_2|^2 \neq P_1 + P_2$$

qui n'est pas égale à la somme des probabilités des deux expériences à fente unique, $P_1 + P_2$.

La superposition des états peut donner des résultats très perturbants lorsqu'elle est appliquée à des objets macroscopiques. L'exemple le plus célèbre est celui du chat de Schrödinger (1935). Considérons le dispositif suivant : un chat (initialement vivant) est enfermé dans une boîte. Dans cette boîte se trouve une fiole de gaz toxique, et un marteau qui peut briser la fiole. Le marteau est actionné par un dispositif automatique, relié à un appareil de mesure qui repère l'état d'un quanton (par exemple le passage par la fente 1 ou la fente 2 dans le cas évoqués précédemment). Si le quanton passe par la fente 1, le marteau est actionné et le gaz mortel se répand dans la boîte, et rien ne se passe dans l'autre cas. Dans quel état est le chat ? Ni mort, ni vivant, tant que la mesure n'est pas faite. Le système complet est décrit par la superposition des deux états (*fente 1 – marteau baissé – fiole brisée – chat mort*) et (*fente 2 – marteau relevé – fiole intacte – chat vivant*). On mesure ici combien la description quantique peut paraître étrange.

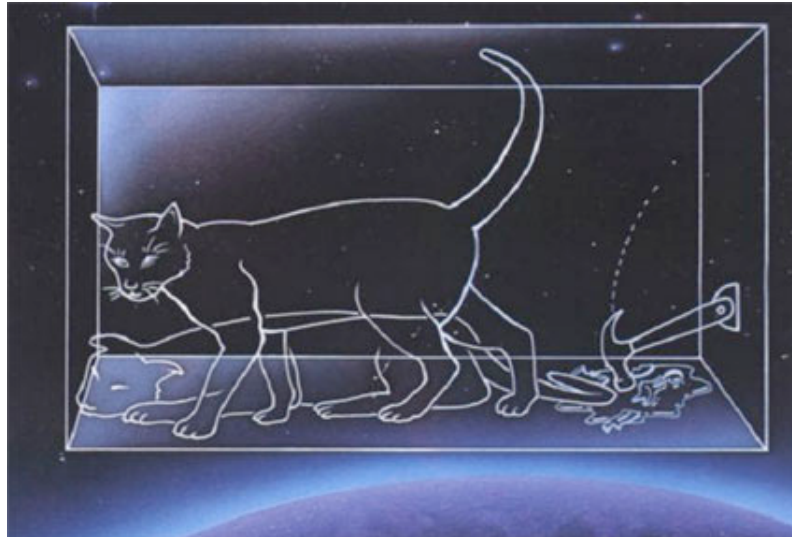


Figure 89 Etat quantique du chat de Schrödinger, qui résulte de la superposition de l'état "chat mort" et de l'état "chat vivant"

11.3 Mesure et réduction du paquet d'ondes

L'expérience menée en laboratoire montre que, lorsqu'on mesure la position de chaque quanton au niveau des fentes dans l'expérience de Young, les franges d'interférence disparaissent et le résultat obtenu est compatible avec une théorie corpusculaire. Le caractère « étrange » associé au comportement quantique a disparu.

Comment interpréter et comprendre cela ? Ce point particulièrement difficile a été l'objet de très nombreux débats et fait l'objet de ce que l'on appelle la **théorie de la mesure**. Sur le plan mathématique, effectuer une mesure de la position, cela revient à remplacer une répartition spatiale de probabilité de présence par la notion classique de position (le quanton est à une position unique lors de la mesure). On parle de **réduction du paquet d'ondes**. Il faut donc concevoir que, tant qu'il ne fait pas l'objet d'une mesure, un quanton n'a pas de position au sens classique et est correctement décrit au sens probabiliste par un paquet d'ondes et l'équation de Schrödinger, et que l'action de mesure rend cette description instantanément caduque et donne un caractère classique au quanton au moment de la mesure. Après la mesure, le quanton retrouve son état quantique, et est doit de nouveau être décrit avec un formalisme probabiliste.

La mesure apparaît donc comme un événement très particulier, qui perturbe de manière notable le comportement des quantons. Tout d'abord, il faut noter que la mesure est toujours le fruit d'une interaction, et qu'elle n'est donc jamais neutre. Si, sur les objets macroscopiques, cette interaction peut souvent être négligée (au sens où elle perturbe très peu l'état de l'objet que l'on mesure), sur les objets à comportement quantique, qui sont de petite taille et de faible énergie, cette interaction n'est pas négligeable. Or, il existe une intensité minimum aux interactions dans l'univers, mesurée par la constante introduite par Max Planck (1858-1947, prix Nobel de Physique 1918) en 1900 lors de son étude du rayonnement des corps noirs, $h = 6,622 \cdot 10^{-34}$ Joule par seconde. Il existe donc une perturbation minimale associée à toute mesure. Mesurer la position introduira donc une modification de la vitesse, et réciproquement. Une conséquence très forte est le **principe d'incertitude** – ou encore

principe d'indétermination – de Heisenberg. Ce principe ne porte pas, comme on le dit souvent, sur une imprécision de la mesure qui serait due à une imperfection de la technique employée. Le principe de Heisenberg énonce un fait beaucoup plus fondamental : il n'est pas pertinent d'associer à un quanton les deux propriétés classiques que sont la position et la vitesse à un instant donné, *car elles ne sont pas définies au même instant*. Il s'agit ici d'une propriété intrinsèque des quantons. De manière fondamentale, il faut retenir que dans le cadre quantique, il ne peut pas y avoir de connaissance du monde (au sens classique non probabiliste) sans perturber celui-ci par une action de mesure. Une connaissance du monde en soi n'a en conséquence pas de sens. Mesurer la position perturbe la vitesse du quanton, et on ne peut donc plus avoir accès à celle-ci. Inversement, mesurer sa vitesse modifie sa position, ce qui interdit de connaître cette dernière. Comme on identifie mesure et définition (au sens de connaissance), la connaissance, ou définition, de l'une est donc exclusive de la connaissance (ou définition) de l'autre.

Il existe un autre résultat dû à Heisenberg. Si l'on prépare un grand nombre de quantons dans un même état, et que l'on mesure la position de la moitié d'entre eux et l'impulsion de l'autre moitié, le produit de la dispersion des mesures de ces deux quantités sera toujours supérieur à une valeur constante fondamentale. Ceci n'est pas dû à une imperfection des appareils de mesure, mais au fait que chaque quanton est associé à une onde de probabilité, et que donc la mesure ne peut pas donner à chaque fois la même valeur : il existe une dispersion nécessaire des résultats.

La réduction du paquet d'onde est un phénomène « typiquement quantique », qui lorsqu'on tente de l'interpréter au moyen de la physique classique, conduit à des paradoxes. Un exemple est le paradoxe introduit en 1959 par Louis de Broglie. Considérons un électron enfermé dans une boîte. Sa fonction d'onde remplit toute la boîte, indiquant que l'électron peut occuper toutes les positions possibles dans cette boîte. Séparons maintenant la boîte en deux boîtes de volumes égaux. La fonction d'onde de l'électron emplit maintenant les deux boîtes, ce qui est compatible avec l'intuition que l'électron peut être dans chacune des boîtes. Laissons une boîte à Paris, et emmenons l'autre à bord d'un satellite en orbite. Ouvrons maintenant la boîte restée à Paris, et effectuons la mesure sur la présence de l'électron. Il y aura alors réduction du paquet d'onde dans la boîte à Paris, mais aussi, instantanément, dans la boîte embarquée sur le satellite (car l'électron est nécessairement dans l'une des deux boîtes). Prenons le cas où l'électron est dans la boîte restée sur Terre. L'interprétation physique classique nous indique qu'il y a toujours été, et que donc sa probabilité d'être dans la boîte sur le satellite a toujours été nulle. Ce n'est pas ce que dit la mécanique quantique, pour laquelle la fonction d'onde de l'électron n'est pas nulle, jusqu'à ce que la mesure réduise le paquet d'onde de manière instantanée. Ce caractère instantané est très déroutant, car il semble incompatible avec un des résultats fondamentaux de la théorie de la relativité, qui est qu'aucun signal ne peut se déplacer avec une vitesse supérieure à celle de la lumière dans le vide ! Cette complication a conduit à de nombreuses interprétations de la mécanique quantique, qui seront évoquées au chapitre 11.5.

Pour en revenir au chat de Schrödinger, la mécanique quantique que le chat ne sortira de son état superposé mort-vivant que lorsque la mesure sera réalisée !

11.4 Intrication quantique et non-séparabilité

La réduction du paquet d'ondes est un premier exemple de l'étrangeté du monde quantique, lorsqu'on tente de l'interpréter à l'aide des concepts de la physique classique.

Un second exemple, plus frappant encore, est celui de l'**intrication quantique** (on parle également d'**entrelacement** ou d'**enchevêtrement quantique**) qui viole le principe classique de séparabilité. Ce phénomène a été découvert par Schrödinger au début des années 1930. Le **principe de séparabilité**, déjà évoqué lors de la discussion sur les théories de la causalité, consiste à dire que les objets physiques sont **locaux** en espace et identifiables séparément les uns des autres. L'interaction entre deux objets ne peut se faire que par la transmission de quelque chose qui se propage à une vitesse au plus égale à celle de la lumière.

Or la mécanique quantique nous apprend que lorsque deux quantons interagissent, ils forment une **paire intriquée**. Après cette interaction, ils ne peuvent plus jamais être considérés isolément, quelle que soit la distance qui les sépare et cela aussi loin que l'on se place dans le futur. Sur le plan mathématique, chaque quanton perd son individualité et la paire est alors décrite par une fonction d'état propre (défini comme le *produit* des vecteurs d'état de chaque quanton – rappelons que la superposition des états et associée à la *somme* des vecteurs d'état). Une conséquence de cette intrication des quantons est qu'ils ne sont plus indépendants, faisant partie de la même paire intriquée. Donc, effectuer des mesures sur un des quantons donnera automatiquement des renseignements sur le second, sans mesure directe sur ce dernier, et cela instantanément et quelle que soit la distance qui sépare les deux quantons. Ce qui est incompatible avec le principe de séparabilité. Plus dérangent encore : c'est la mesure sur le premier quanton qui va induire la réduction du paquet d'ondes de la paire, et fixer l'état du second quanton ! En conséquence, on parle de **non-séparabilité** ou de **non-localité** de la mécanique quantique.

Cet aspect de la mécanique quantique a dérangé beaucoup de physiciens, poussant certains à refuser la théorie en l'état. Ce fut le cas d'Albert Einstein. En 1935, il publie avec deux collègues, Boris Podolsky et Nathan Rosen, un article de quatre pages qui restera célèbre et qui engendrera de très nombreuses recherches. Les auteurs y développent une critique de la non-séparabilité, connue sous le nom de **paradoxe EPR**, du nom des trois auteurs. Dans cet article, ils posent les prémisses suivantes :

3 prémisses du paradoxe EPR

1. Les prédictions de la physique quantique sont justes
2. **Principe de séparabilité** ou de **localité d'Einstein** : aucune influence ne peut se propager plus vite que la lumière
3. **Critère de réalité** : si, en ne perturbant aucunement un système, on peut prédire avec certitude (c'est-à-dire avec une probabilité égale à 1) la valeur d'une quantité physique, alors il existe un élément de réalité physique correspondant à cette quantité.

On peut reconnaître dans le critère de réalité une conception réaliste de la physique. Le principe de séparabilité est issu de la théorie de la relativité, qui peut être vue ici comme le point culminant du développement de la physique classique. En considérant une paire de quantons intriqués, ces auteurs ont montré l'incompatibilité de la mécanique quantique avec les conclusions de la physique classique, et en déduisaient

que la théorie de la mécanique quantique était incomplète, au sens où de nouvelles variables devaient être ajoutées pour retrouver une théorie compatibles avec les faits observés et les principes classiques de localité et de séparabilité. Cela est-il possible ? Il a fallu attendre presque 50 ans, avec les expériences conduites par le physicien français Alain Aspect au début des années 1980, pour avoir la certitude que les mécanismes quantiques ne peuvent pas être représentés par une théorie locale, quelle qu'elle soit. Cette conclusion s'appuie sur les inégalités trouvées par le physicien irlandais John Bell en 1964. Ces inégalités, dont le caractère technique est trop compliqué pour pouvoir être développé ici, ne peuvent être violées que dans le cas où le monde réel possède bien le caractère non-local décrit par le formalisme quantique. Les expériences d'Aspect ont été les premières à montrer cette violation avec assez de précision pour pouvoir conclure positivement sur le caractère non-locale des objets quantiques.

Un aspect « inconfortable » de la non-séparabilité quantique est qu'elle n'est pas compatible avec la théorie classique de la causalité, puisque cette dernière est basée sur la séparabilité (la cause et effet doivent pouvoir être définis isolément) et sur la succession dans le temps (les interactions instantanées posent donc problème). Notons également qu'aujourd'hui la théorie quantique n'est toujours pas compatible avec la théorie de la relativité générale, au sens où il n'existe toujours pas de théorie qui unifie les deux, et cela malgré de très nombreux efforts.

11.5 Conséquences épistémologiques

Les conséquences épistémologiques de la mécanique quantique sont immenses. L'incompatibilité intrinsèque de la description du monde donnée par la mécanique quantique avec notre perception intuitive (qui est en correspondance avec la physique classique) paraît en effet fondamentale et irréductible, ce qui a conduit les physiciens et les philosophes à développer de nombreuses interprétations de la théorie quantique.

Un premier point est que les phénomènes quantiques ne sont principalement observables qu'à très petite échelle, celle de la physique atomique. A l'échelle macroscopique, qui est celle de la physique classique, les phénomènes recouvrent un comportement classique. Certaines manifestations des propriétés quantiques à l'échelle macroscopique sont toutefois possibles, comme le montre l'existence des supra-conducteurs et des superfluides. La continuité entre la description quantique et la physique classique est assurée par le **principe de correspondance**, énoncé par Bohr en 1916 et corrigé en 1927 par Ehrenfest, qui dit que lorsque le nombre de quanta impliqués dépasse un certain seuil, les effets quantiques s'estompent. La physique classique peut donc être vue comme un cas limite de la physique quantique. La **théorie de la décohérence**, développée durant les années 1990, indique que ce sont les interactions avec le monde extérieur qui donnent leur caractère déterministe aux objets macroscopiques. Cette théorie a été corroborée expérimentalement, et résout partiellement le problème du chat de Schrödinger : étant un objet macroscopique, le chat est soumis au phénomène de décohérence, et obéit donc à un comportement classique. Il est donc mort ou vivant, et ne reste pas dans un état superposé.

La théorie de la décohérence et celle de la réduction du paquet d'ondes ne sont pas sans poser de problème. En effet, un quanton n'est jamais complètement isolé avant la

mesure. La réduction du paquet d'ondes lors d'une opération de mesure semble indiquer que la mesure est un type d'interaction particulier qui lui confère des propriétés spéciales. De même, puisque deux quanta qui ont interagis restent intriqués, si l'on remonte jusqu'au big-bang, il est pensable que toutes les particules présentes dans l'Univers sont intriquées. Ces difficultés d'interprétation des résultats de la mécanique quantique ont conduit au développement de plusieurs écoles de pensée.

L'école dominante est sans aucun doute l'**école de Copenhague** (représentée entre autres par Heisenberg, Bohr et Pauli), pour laquelle la théorie est un outil prédictif complet et un outil explicatif incomplet, et rien d'autre. Son formalisme ne traite pas du monde tel qu'il existe, mais de la connaissance que nous en avons. Chercher une signification plus profonde n'a pas de sens. Cette école est désignée comme école positiviste, école empiriste ou encore école opérationnaliste.

Les physiciens (parmi lesquels des chercheurs prestigieux) et philosophes qui pensent que la théorie quantique est plus qu'un outil formel efficace peuvent être regroupés de plusieurs manières. On peut par exemple distinguer les **matérialistes** (qui pensent qu'il n'existe rien d'autre dans le monde réel que les quantités qui apparaissent dans les théories physiques) et les **spiritualistes**, pour lesquels l'esprit humain (ou quelque chose d'autre de nature immatérielle comme la conscience) doit être pris en compte pour expliquer pourquoi il y a réduction du paquet d'ondes lorsqu'on fait une mesure et pas lors d'autres interactions. C'est par exemple la position du prix Nobel de Physique Eugène Wigner. Les différentes tentatives d'interprétation de l'intrication ou de la réduction du paquet d'ondes dans un cadre classique ont également poussé certains physiciens à remettre en cause les notions d'espace et de temps (le voyage dans le temps ayant été employé par certains pour expliquer ce qui ressemble à des interactions à distance instantanées), voire à proposer des théories basées sur l'existence d'univers multiples (un univers existe pour chaque état possible). On voit ici que ces interprétations sont du domaine métaphysique, et échappent au domaine scientifique tel que nous l'avons défini. Mais il faut retenir que, si les équations de la mécanique quantique sont aujourd'hui communément acceptées comme un outil efficace de prédiction, la question de leur sens et de leur lien avec une réalité extralinguistique reste un problème ouvert.

11.6 Vers de nouvelles applications : cryptographie quantique, informatique quantique et téléportation quantique

La physique quantique et ses résultats ouvrent des perspectives pratiques et technologiques tout à fait nouvelles. Nous allons très rapidement voir trois d'entre elles. Rappelons également le cas du microscope à effet tunnel, déjà vu au chapitre 4.2.7.

11.6.1 Ordinateur quantique

Les ordinateurs actuels sont basés sur le codage de l'information par des **bits**, qui peuvent prendre la valeur 0 ou 1. Au plan technologique, ce codage binaire est réalisé au moyen de commutateurs qui ont deux états stables.

A la suite des travaux initiés par Richard Feynman (1918-1988, prix Nobel de physique 1965) et d'autres physiciens comme David Deutsch, Paul Benioff et Charles Bennet dans les années 1970-1980, s'est posée la question de construire de nouveaux ordinateurs et de nouveaux algorithmes basés sur des bits quantiques (*quantum bits* - **Qubits**). Les Qubits ne sont pas binaires : ils peuvent prendre toutes les valeurs comprises entre 0 et 1. Un Qubit n'est pas matérialisé par un commutateur déterministe, mais par quanton dont les états superposés permettent de balayer les valeurs autres que 0 et 1. Mais il faut noter que le Qubit est de nature probabiliste : alors qu'un commutateur classique est dans un état associé à 0 ou 1, un Qubit n'est associé qu'à une probabilité de donner une valeur si une mesure est effectuée. Cette distinction fondamentale donne la possibilité de développer de nouveaux algorithmes aux performances parfois très supérieures à ceux basés sur les bits déterministes. La raison en est qu'en travaillant sur un Qubit, on peut effectuer des opérations sur toutes les valeurs qu'il peut prendre à la fois, ce qui revient à faire de très nombreux calculs déterministes classiques en parallèle. La superposition des états et l'intrication font que la puissance de calcul d'un ordinateur quantique double à chaque fois qu'on ajoute un Qubit. Un ordinateur à N Qubits a donc une puissance de l'ordre de 2^N . Pour N=300, on obtient une puissance supérieure au nombre de particules estimé pour l'ensemble de l'univers observable !

De tels algorithmes ont été proposés pour la factorisation des grands nombres. Ils sont efficaces, au sens où ils peuvent résoudre ce problème en un temps polynomial et non exponentiel comme les algorithmes classiques. Le premier algorithme de ce type a été proposé en 1994 par Peter Shor. Concernant la recherche au sein des grandes banques de données, l'algorithme par Lov Grover en 1996 est plus efficace que les algorithmes classiques.

La construction d'ordinateurs quantiques est encore loin d'être un problème résolu. Le premier ordinateur quantique a été construit par IBM en 1998 ; il avait 2 Qubits. Cette même équipe réalisa la première utilisation de l'algorithme de Grover sur un ordinateur avec 3 Qubits en 1999, puis en 2000 sur un ordinateur avec 5 Qubits. Le nombre 15 est factorisé au moyen de l'algorithme de Shor le 19 décembre 2001 chez IBM sur une machine à 7 Qubits basés sur des molécules de chloroforme (la durée de vie de la machine était de quelques minutes seulement !). Depuis, des ordinateurs quantiques comprenant plus de 10 Qubits ont été construits en utilisant différentes technologies (chimie, optique, circuits supraconducteurs ...) Les deux problèmes majeurs pour la construction d'ordinateurs quantiques est qu'il faut éviter la décohérence des Qubits (donc il faut les isoler du monde extérieur) et éviter la destruction de l'information associée à la réduction du paquet d'ondes lorsque la valeur d'un Qubit est mesurée.

Un ordinateur quantique serait-il supérieur à un ordinateur classique pour tous les types de problème ? Il semble aujourd'hui que non. Les problèmes informatiques peuvent être classés suivant leur **complexité**, qui mesure le nombre d'opérations nécessaires pour les résoudre. Plus précisément, la complexité mesure la manière dont

le nombre d'opérations varie en fonction du volume de données. Les travaux théoriques permettent de distinguer les classes de problèmes ci-dessous⁹⁹ :

- Les **problèmes de classe P** (*temps polynomial*), qui sont ceux qui peuvent être résolus en temps polynomial par un ordinateur classique (c'est-à-dire que si n est le volume de données, le nombre d'opérations varie comme n^q , où q est une puissance qui dépend du problème). Ce sont les problèmes les plus faciles à résoudre. Un exemple classique est celui de la **connexité d'un graphe**, illustré par le **problème de la carte routière** (complexité n^2): étant donné une carte routière, est-il possible d'atteindre chaque ville à partir de tout autre ville ?
- Les **problèmes de classe NP** (*temps polynomial non-déterministe*), qui sont les problèmes pour lesquels une solution proposée peut être testée en un temps polynomial (ce qui ne veut pas dire que l'on peut trouver une solution en un temps polynomial !). Tous les problèmes P sont bien sûr des problèmes NP, mais on ne sait pas aujourd'hui si tout problème NP est obligatoirement P (ce qui se note formellement comme $P = NP$?). De très nombreux problèmes de grand intérêt pratique sont de classe NP. On peut citer le problème de la factorisation d'un nombre entier comme un produit de nombres premiers : les meilleurs algorithmes classiques connus sont de complexité exponentielle (et donc supérieure à une complexité polynomiale).
- Les **problèmes NP-complets**. Ces problèmes de classe NP ont une propriété particulière : si on trouvait un **algorithme efficace** (c'est-à-dire de classe P) pour résoudre l'un d'eux, on démontre que cet algorithme serait adaptable et efficace pour tous les autres problèmes NP-complets. On ne connaît aujourd'hui d'algorithme P pour aucun des problèmes NP-complets. Ils sont en ce sens les plus difficiles des problèmes NP. Si on trouvait un algorithme P pour un problème NP-complet, on aurait démontré un des 7 problèmes identifiés comme les plus difficiles des mathématiques en 2000 (année mondiale des mathématiques). Des exemples célèbres de problèmes NP-complets sont le **problème du 3-coloriage** (étant donné une carte, est-il possible de la colorier avec 3 couleurs de telle sorte que 2 pays voisins ne soient jamais de la même couleur ?) et le **problème du voyageur de commerce** (tracer l'itinéraire de longueur minimale qui visite une seule fois toutes les villes d'une région).
- Les **problèmes de la classe PSPACE**, qui sont les problèmes qu'un ordinateur classique peut résoudre avec une quantité polynomiale de mémoire en un temps quelconque (polynomial ou exponentiel). C'est la classe la plus générale, qui contient toutes les autres. Des exemples de problème PSPACE mais pas NP sont les jeux d'échecs ou de go sur des grilles de dimension $n \times n$.
- Les **problèmes BQP** (*bounded-error quantum polynomial time*), qui sont les problèmes que l'on peut résoudre en un temps polynomial avec un ordinateur quantique avec une erreur aussi petite qu'on le veut. Cette classe contient bien sûr tous les problèmes P, mais aussi d'autres problèmes, comme la **factorisation des entiers** (l'algorithme de Shor est de complexité n^2). On n'a pas pu montrer jusqu'à aujourd'hui qu'un ordinateur quantique puisse résoudre de manière efficace (donc en temps polynomial) un problème de la

⁹⁹ Voir l'encyclopédie en ligne des classes de complexité de S. Aaronson : <http://www.complexityzoo.com>

classe NP-complet. Les algorithmes quantiques peuvent dans certains cas faire mieux pour ces problèmes, mais en un temps qui reste exponentiel¹⁰⁰.

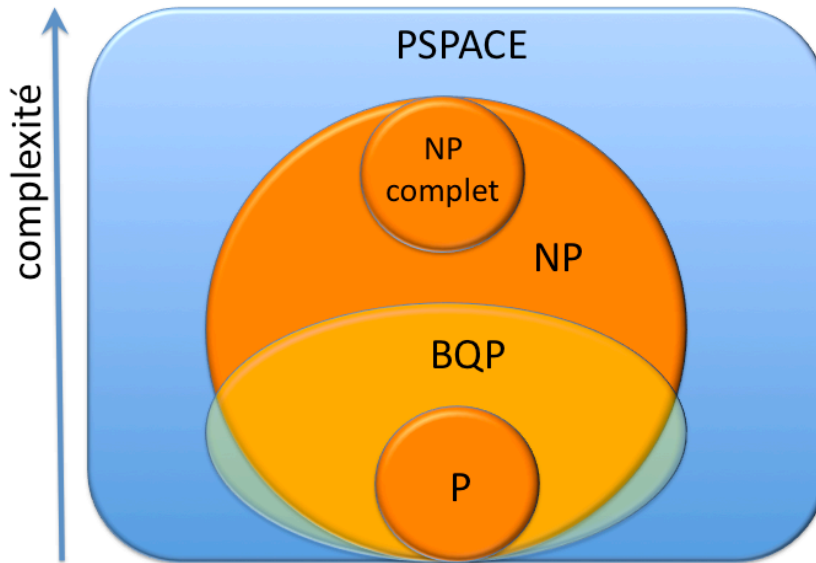


Figure 90 Classification des problèmes informatiques suivant leur complexité (adapté de "les limites du calcul quantique", S. Aaronson, Pour la Science, mai 2008)

11.6.2 Cryptographie quantique

La cryptographie (du grec *kryptos*, qui veut dire caché) est la discipline qui traite du codage et du décodage des messages, le but étant que seul un interlocuteur autorisé puisse avoir accès à l'information contenue dans le message en un temps suffisamment court pour que cette information soit utile. Le schéma classique d'un message crypté est le suivant. Un émetteur (traditionnellement appelé Alice) souhaite envoyer un message sécurisé à son interlocuteur, Bob. Pour cela, Alice et Bob partagent une clé secrète, qui sert à crypter (chez Alice) et décrypter (chez Bob) le message. Alice crypte le message avant de l'envoyer. L'information qui circule est donc sécurisée, puisque théoriquement inutilisable sans la clé, qui elle ne fait pas partie du message. Ceci est illustré sur la figure ci-dessous.

¹⁰⁰ Mais on a pu monter que des ordinateurs encore plus étranges que les ordinateurs quantiques, c'est-à-dire des ordinateurs capables de remonter dans le temps, pourraient résoudre tous les problèmes de PSPACE (y compris les problèmes NP-complets) en un temps polynomial. Le seul problème est bien sûr que de tels ordinateurs n'existent pas ! Il est même peut être possible de conjecturer que prouver $P=NP$ serait équivalent à la possibilité de remonter dans le temps !

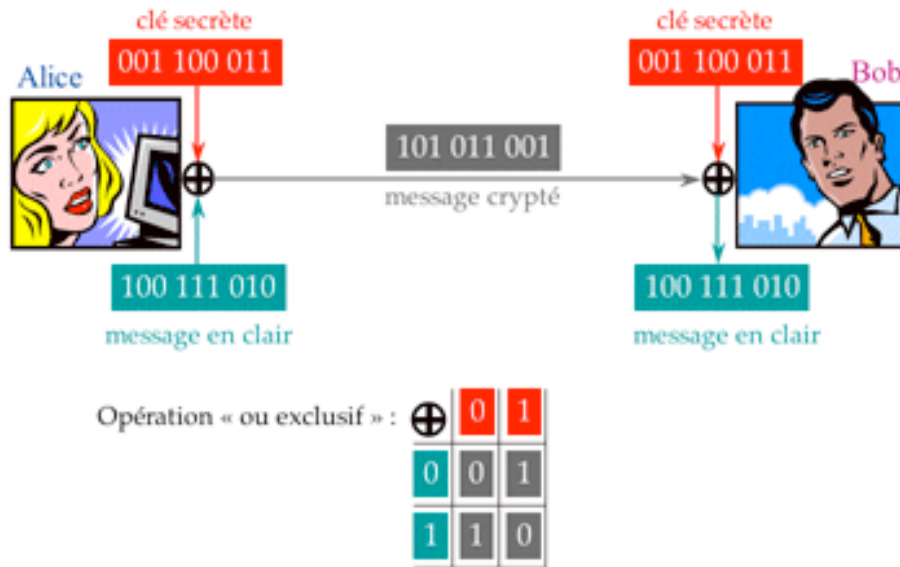


Figure 91 Principe d'une communication cryptée : Alice code son message pour Bob au moyen d'une clé secrète, dont Bob dispose et qu'il utilisera pour décrypter le message.

La cryptographie quantique, développée depuis le milieu des années 1980, consiste à transmettre l'information cryptée à l'aide de quantons (des photons polarisés, en pratique), et à se servir de leurs propriétés pour détecter si le message a été intercepté durant la communication par un tiers non-autorisé (appelé Eve, selon la tradition). Plus précisément, il s'agit de communiquer la clé secrète de manière confidentielle, puisque c'est cette clé qui garantit ensuite une communication sécurisée.

Pour cela, la clé va être codée sous forme binaire, c'est-à-dire une suite de 0 et de 1. Lors du message, les valeurs 0 et 1 vont être associées à différentes polarisations possibles des photons, ce choix étant effectué aléatoirement entre deux modes possibles pour chaque photon. Bob va réceptionner les photons et mesurer leur polarisation à l'aide d'un filtre dont la polarisation est elle-même aléatoire. Comme le processus est aléatoire, la moitié de l'information est perdue. Ensuite, pour donner à Bob le moyen de retrouver son choix de codage par polarisation, Alice va envoyer un message « en clair ». Il ne s'agit pour le moment que d'un moyen très compliqué d'envoyer une clé. Où se trouve l'intérêt de la cryptographie quantique ? Dans le fait que si Eve intercepte le message elle va nécessairement modifier aléatoirement la polarisation des photons. Donc, Bob recevra alors un message modifié, qui n'aura pas de signification, ce dont il s'apercevra lorsque Alice lui communiquera le bout de message en clair. L'intrusion est alors détectée, et il suffit à Bob de prévenir Alice pour que celle-ci n'envoie pas le message confidentiel.

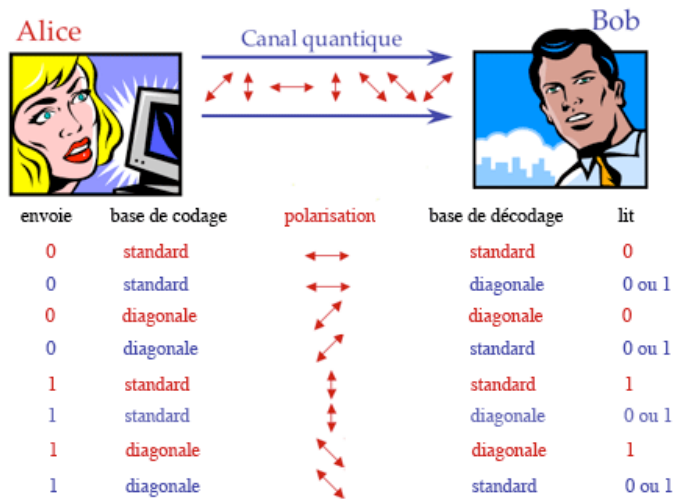


Figure 92 Schéma de base de la cryptographie quantique basée sur la polarisation des photons

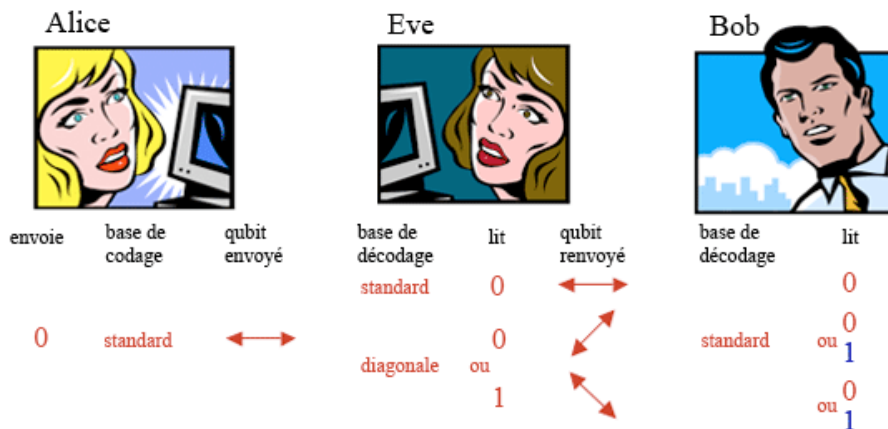


Figure 93 Schéma de principe d'une communication employant la cryptographie quantique : si Eve intercepte le message envoyé par Alice à Bob, la polarisation des photons est nécessairement modifiée, ce que Bob pourra détecter.

11.6.3 Téléportation et fax quantique

La téléportation quantique est basée sur la possibilité de réduire le paquet d'ondes d'une paire de quanton intriquée en n'agissant que sur un seul des deux quantons. Supposons que Alice souhaite « faxer » une information à Bob, et que cette information soit stockée à l'aide de photons. Pour cela, elle doit partager des paires de photons intriqués avec Bob. Pour partager son information avec Bob, elle doit intriquer chacun de ces photons porteurs d'information avec un photon d'une paire intriquée qu'elle partage avec Bob. Ensuite, il lui suffit d'effectuer une mesure sur la

nouvelle paire qu'elle vient de créer pour réduire son paquet d'ondes, et donc figer l'état des deux photons. Or l'un de ces deux photons était et est resté intriqué avec un photon que Bob a reçu. En effectuant sa mesure, Alice va donc également réduire le paquet d'ondes du photon de Bob, dont l'état va être déterminé par celui des photons qui contenait l'information d'Alice. En conséquence, Bob a accès aux informations que veut lui transmettre Alice à distance, le transfert se faisant de manière instantanée.

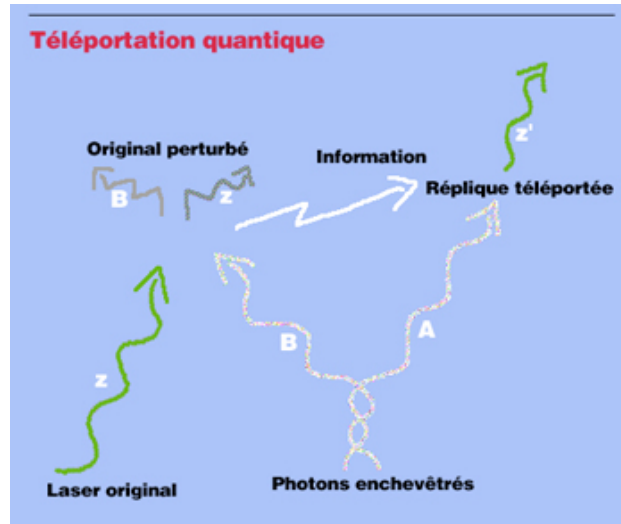


Figure 94 Schéma explicatif de la téléportation et du fonctionnement du fax quantique

Les premières expériences de téléportation quantique sur une paire de photons ont été réalisées en 1997 par Anton Zeilinger et son équipe. *Les termes de téléportation et de fax quantique sont à manipuler avec précaution.* Il faut d'abord noter que, contrairement à un fax classique, le fax quantique détruit l'information qu'il transmet (les photons d'Alice sont perturbés). Ensuite, la téléportation évoquée ici n'est pas celle des films de science fiction ! On ne téléporte ici ni matière ni énergie, mais seulement de l'information.

12 Histoire de la science du mouvement

12.1 La physique d'Aristote

La partie de l'œuvre d'Aristote de Stagire (384-322 av. JC) consacrée à ce que nous appelons aujourd'hui la mécanique et la physique sont indissociables du reste de sa pensée. En effet, la « physique aristotélicienne », bien qu'elle traite des phénomènes naturels observables par l'homme, est très différente d'une science dans la conception moderne de ce terme, car elle repose sur des principes métaphysiques. Nous allons passer en revue les principaux éléments de la théorie d'Aristote, qui a dominé la pensée occidentale pendant près de deux millénaires.

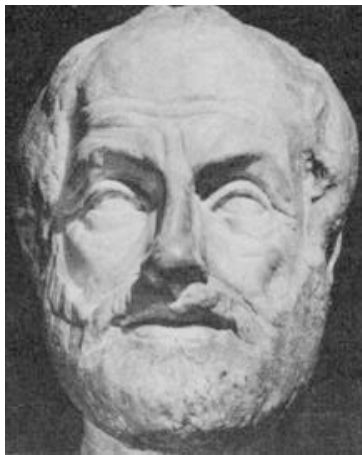


Figure 95: Aristote

La physique d'Aristote est une théorie complète et cohérente, basée sur les concepts d'ordre et de hiérarchie. Le critère de validité sur laquelle elle repose est un critère de logique formelle (vérité-cohérence) et sur l'adéquation avec des principes métaphysiques issus de la philosophie platonicienne.

Les conceptions physiques d'Aristote sont basées sur une distinction fondamentale : celle qui existe entre le monde sublunaire, qui contient la Terre, et le Ciel. Ces deux mondes sont de nature différente, et les lois qui les régissent ne sont pas les mêmes. Cette conception est donc en opposition totale avec la conception moderne selon laquelle l'espace et le temps sont homogènes¹⁰¹.

Le Ciel, qui contient les corps célestes, est le domaine de la perfection et de l'immutabilité. En conséquence, pour Aristote, les mouvements qui s'y déroulent ne peuvent être que parfaits et éternels. Le cercle étant l'incarnation de la perfection et de l'éternité (il n'a ni commencement ni fin), les mouvements dans le ciel sont des mouvements circulaires uniformes. Nous observons ici une transposition de l'idéalisme de Platon. On trouve dans le Ciel un élément particulier (le cinquième élément, les quatre autres formant le monde sublunaire) : l'éther, qui est incorruptible et immuable.

Le monde sublunaire est à l'opposé du ciel : il est corrompu et éphémère, car les objets qui s'y trouvent se forment, changent et se détruisent. Pour Aristote, qui rejette la théorie atomiste, les corps qui s'y trouvent sont composés, en accord avec la proposition d'Empédocle d'Agrigente (480 ?-430 ? av. JC), de quatre éléments (**éléments d'Empédocle**) : terre, eau, air et feu. Les corps sont plus ou moins lourds

¹⁰¹ Cette hypothèse d'homogénéité de l'espace et du temps, qui traduit le fait que les lois de la physique sont les mêmes en tout point de l'Univers et à chaque instant, est parfois appelée *principe copernicien*.

en fonction de leur composition relative. Notons ici que la distinction entre masse et poids n'existe pas chez Aristote. Un point très important de la théorie d'Aristote est l'idée que chaque élément possède un lieu naturel de repos dépendant de sa densité. L'ordre de stratification est : terre-eau-air-feu. Le lieu de repos est donc fonction de la composition du corps. Cette partie de l'œuvre d'Aristote aura une influence très profonde sur le développement de la chimie. Dans le monde sublunaire, Aristote distingue deux types de mouvements: les mouvements naturels et les mouvements violents ou forcés.

Le **mouvement naturel** est celui d'un corps vers son lieu de repos naturel (chute d'un corps pesant sur le sol, par exemple). Le **mouvement violent** est celui imposé à un objet et qui ne correspond pas à un mouvement naturel.

Venons-en maintenant aux principes de la mécanique d'Aristote, c'est-à-dire à son explication des mouvements dans le monde sublunaire. Pour lui, deux facteurs s'opposent : une action motrice et une action de résistance au mouvement. De manière abusive, puisque la notion de force lui était complètement inconnue, on parlerait aujourd'hui de force motrice F et de force de résistance R .

Le premier principe de la mécanique aristotélicienne est qu'il y a mouvement si et seulement si l'action motrice est supérieure à l'action de résistance. En employant un formalisme mathématique moderne complètement inconnu à cette époque, ce principe peut s'écrire comme suit :

Premier principe de la mécanique d'Aristote

$$F > R \Rightarrow \text{mouvement}$$

Le second principe aristotélicien peut s'énoncer de la manière suivante : il n'y a mouvement, c'est-à-dire vitesse, que si il y a action (force), soit :

Loi fondamentale de la mécanique d'Aristote

$$\text{Mouvement} \Leftrightarrow \text{Action}$$

Plus précisément, pour Aristote, le mouvement violent ne peut perdurer que si l'action motrice continue de jouer. Lorsque l'action cesse, le mouvement violent cesse pour laisser place à un mouvement naturel. Nous savons aujourd'hui que cela est faux, car ce principe est en opposition avec la première loi de Newton (principe d'inertie). Mais analysons la cohérence interne de cette mécanique.

Pour cela, considérons deux boules pesantes identiques, que l'on fait chuter au même moment et de la même hauteur, l'une dans l'air, l'autre dans l'eau. On observe que celle qui chute dans l'air atteint la première le sol. En répétant l'expérience avec des liquides de viscosités différentes, on arrive à la conclusion que la vitesse de chute V est d'autant plus faible que la force de résistance due à la viscosité est grande, et qu'il existe un rapport de proportionnalité inverse. En faisant varier le poids des boules en modifiant leur composition, on observe qu'il existe un rapport de proportionnalité entre le poids (force motrice) et la vitesse. Une conclusion logique de la mécanique d'Aristote est donc la loi (fausse !) suivante : la vitesse est proportionnelle à l'action et inversement proportionnelle à la résistance :

Cette loi est bien évidemment incompatible avec la seconde loi de Newton, qui lie la résultante des forces (et non pas leur rapport, car il n'y a pas de distinction de nature entre les forces dans la mécanique de Newton) avec l'accélération et non pas la vitesse.

Comment mettre en échec la mécanique aristotélicienne ?

En considérant par exemple la chute dans l'air de deux boules de même taille, mais de composition différente, l'une pesant dix fois plus lourd que l'autre¹⁰². La mécanique d'Aristote prévoit que la boule la plus lourde aura une vitesse dix fois plus grande, et donc que son temps de trajet sera dix fois plus court que celui de la boule légère. Or l'expérience révèle que les temps de trajet sont presque identiques !

Une autre difficulté de la mécanique d'Aristote est le problème du jet : pourquoi un objet, une fois qu'il a quitté la main de son lanceur, continue-t-il à se mouvoir ? Le problème est réel puisque, selon son principe fondamental, le mouvement doit être entretenu pour continuer à exister. Pour lever le paradoxe, Aristote invoque l'impossibilité du vide (doctrine qui 2500 ans plus tard nous fait encore dire, à la suite des scolastiques, que « la nature a horreur du vide »), qui n'est pas compatible avec l'existence d'un ordre cosmique et de place attribuée à chaque chose : dans le vide il n'y a pas de lieu naturel et il n'y a pas de lieu du tout – on ne peut placer quelque chose dans rien. Puisque le vide ne peut exister, l'air écarté sur le passage du projectile se dirige vers le sillage (sans quoi du vide existerait), formant des tourbillons qui font pousser le projectile, prolongeant ainsi l'action de la force motrice. Cette explication est, pour un mécanicien moderne, doublement fautive : elle attribue la continuation du mouvement à une action du milieu extérieur (et non pas à l'inertie du corps, qui lui est propre), et donne une action motrice au sillage (alors que l'aérodynamique moderne montre que cette force est une force de résistance : la traînée).

En conclusion, retenons les faiblesses suivantes de la mécanique d'Aristote :

1. Elle ne définit pas clairement la notion de vitesse
2. La notion d'accélération lui est complètement étrangère
3. Elle ne connaît pas la notion d'inertie (et lui est même contraire par certains aspects)

Gardons également en mémoire un point important de la physique d'Aristote, qui dominera la réflexion sur le mouvement jusqu'au XVIIe siècle : le statut particulier du mouvement circulaire, qui est l'image de la perfection (*Physique*, livre VIII) :

« Il est évident que le transport circulaire est le premier des transports. En effet tout transport, comme nous l'avons dit précédemment, est, ou circulaire, ou rectiligne ou mixte ; ceux-là sont nécessairement antérieurs à celui-ci, puisqu'il en est composé ; et le circulaire est antérieur au rectiligne, car il est plus simple et plus parfait. En effet il

¹⁰² Il s'agit d'un exemple développé par Galilée, dont le texte est repris plus bas

n'y a pas de transport sur une droite infinie, car un tel infini n'existe pas¹⁰³ ; et, s'il existait, rien ne serait ainsi mû, car l'impossible ne se produit pas et parcourir l'infini est impossible. Maintenant, le mouvement sur une droite finie, quand il est rebroussé, est imparfait et destructible. Or le parfait est antérieur à l'imparfait selon la nature, selon la notion, selon le temps ; et l'indestructible, au destructible. En outre, un mouvement qui peut être éternel est antérieur à celui qui ne le peut ; or le mouvement circulaire peut être éternel, tandis qu'aucun des autres, ni le transport, ni d'ailleurs aucun autre, ne le peut ; car un arrêt doit se produire et, s'il y a arrêt, le mouvement est détruit.

Il nous a paru d'autre part raisonnable que le mouvement circulaire fut un continu, et que le rectiligne ne le fut pas. Pour le rectiligne, en effet, le commencement, la fin et le milieu sont déterminés, et il a tout cela en soi-même, de sorte qu'il y a pour la chose mue un point de départ et un point d'arrivée (aux limites, en effet, il y a toujours repos, à la limite initiale comme à la terminale). Pour le circulaire tout cela est, au contraire, indéterminé ; car pourquoi, entre les points qui sont sur la ligne, celui-ci plutôt que celui-là serait une limite ? Chaque point en effet est au même titre commencement, milieu et fin ; et, par suite, une chose qui se meut en cercle est toujours au commencement comme à la fin, et elle n'y est jamais [...] Autre preuve qui se tire de la réciproque suivante : c'est parce que le transport circulaire est la mesure des mouvements, qu'il doit être le premier (car partout, c'est le premier qui est mesure) ; et c'est aussi parce qu'il est le premier qu'il est la mesure des autres. En outre, seul le transport circulaire peut aussi être uniforme. En effet les choses mues sur une droite ne sont pas transportées uniformément du commencement vers la fin ; car plus elles s'éloignent de l'état où elles sont en repos plus rapide est le transport ; et pour le seul transport circulaire, le commencement et la fin ne sont par nature pas en lui, mais hors de lui ... »

12.2 La physique médiévale

La physique médiévale se construit sur les bases de la physique d'Aristote, en accord avec ou en réaction contre ses énoncés. Le principal problème posé aux penseurs est celui du jet (persistance du mouvement), car l'explication ingénieuse d'Aristote ne semble pas satisfaisante.

12.2.1 Jean Philopon d'Alexandrie : l'*energeia*

Ainsi, dès le VI^e siècle, le commentateur chrétien d'Aristote Jean Philopon d'Alexandrie (490-566) remet cette explication en cause en avançant que l'action de l'air ne peut être qu'une action de résistance et non une action motrice. Le lanceur communique donc au projectile une « certaine puissance motrice immatérielle » nommée **energeia** qui permet au mouvement de se prolonger. Cette **energeia** va diminuer au cours du temps à cause de la résistance de l'air, et, lorsqu'elle aura complètement disparu, le mouvement s'arrêtera. Cette position sera discutée quelques siècles plus tard par Saint Thomas d'Aquin (1225-1274) et Roger Bacon (1219-1292).

¹⁰³ Il faut se rappeler ici qu'il s'agit des mouvements dans le monde sublunaire, qui est fini puisque entouré par le ciel

12.2.2 La naissance de la cinématique : le Merton College d'Oxford

La cinématique naît au XIV^e siècle durant lequel on voit apparaître les premières définitions rigoureuses des grandeurs cinématiques et les premières descriptions des divers mouvements. Notons que près de 2000 ans se sont écoulés depuis Aristote ! Ces travaux sont l'œuvre des scolastiques anglais du Merton College d'Oxford comme Thomas Bradwardine (1290-1349) et William de Heytesbury (1330-1371).

Un apport fondamental du Merton College est le passage d'une description purement qualitative (chez Aristote, la vitesse est grande ou petite) à une description quantitative : on traite maintenant des grandeurs physiques en parlant de leur intensité que l'on nommait leur degré (de vitesse, de chaleur, ...). C'est Bradwardine qui donne la définition de la vitesse V comme étant le rapport entre la distance parcourue d et le temps écoulé t

Définition de la vitesse selon Bradwardine

$$V=d/t$$

Il introduira également une représentation graphique permettant de représenter l'évolution des degrés. Le concept d'accélération est introduit par Heytesbury, qu'il nomme la « vitesse de la vitesse ».

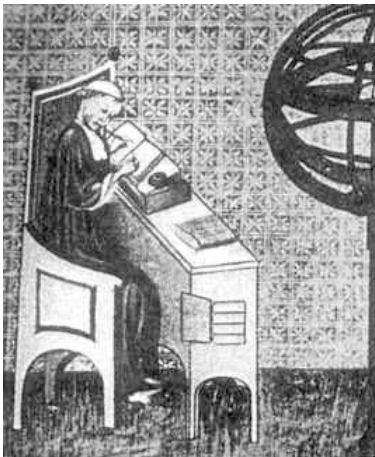


Figure 96: Nicole d'Oresme

C'est le français Nicole d'Oresme (1323-1382), un élève de Buridan, qui, faisant usage des travaux du Merton College, distingue en 1350 trois sortes de mouvement : le mouvement uniforme (motus uniformis), le mouvement uniformément varié (motus uniformiter difformis) et le mouvement difformément varié (motus difformiter difformis). Les deux premiers sont des mouvements de base de la cinématique moderne. C'est également à Oresme que l'on doit la démonstration (géométrique) de la règle de Merton (encore appelée règle de la vitesse moyenne), qui dit que dans le cas d'un mouvement rectiligne uniformément accéléré de vitesse initiale V_i et de vitesse finale V_f , la distance parcourue d est égale à celle parcourue pendant le même temps à la vitesse $V = (V_i + V_f)/2$. Oresme proposera également de décomposer les mouvements quelconques en une succession de mouvements uniformes durant de petits intervalles, mais il faudra attendre l'apparition du calcul différentiel et intégral pour que cette idée devienne féconde.

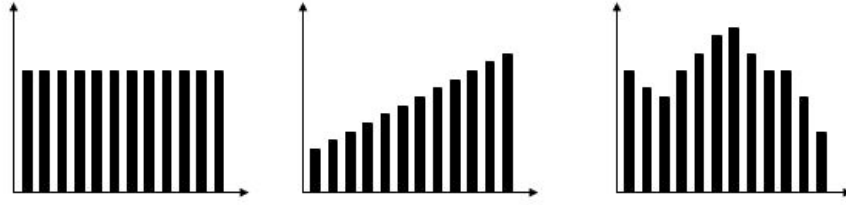


Figure 97: mode de représentation des mouvements selon l'école du Merton College. Gauche: mouvement uniforme, centre: mouvement uniformément accéléré, droite: mouvement difformément accéléré

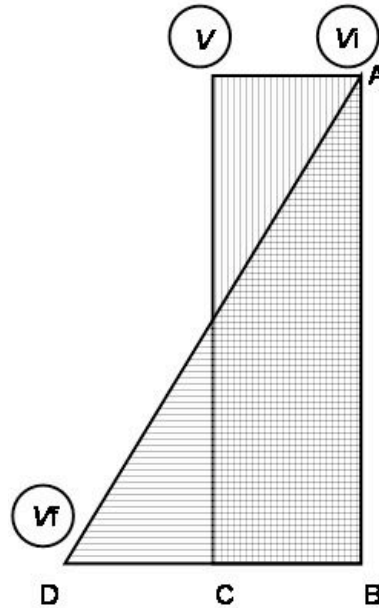


Figure 98: principe de la démonstration de la règle de Merton par Nicole d'Oresme: l'aire du triangle ADB est égale à celle du rectangle de côtés AB et BC

Si ces idées semblent modernes, il faut toutefois se rappeler que, pour leurs auteurs, elles étaient théoriques et ne reposaient sur aucune expérimentation. La réalité de ces mouvements et le fait qu'ils puissent correspondre aux phénomènes observés, comme la chute des corps, étaient des pensées étrangères pour les scolastiques de cette période.

12.2.3 La dynamique de l'*impetus* : Buridan

Une contribution majeure à la physique médiévale est celle de Jean Buridan (1300-1358), qui fût deux fois recteur de l'Université de Paris. Il reprend l'idée de Philopon et émet l'hypothèse que lors de son lancement un corps acquiert une sorte de force intérieure, l'*impetus*, qui s'épuise peu à peu sous l'effet de la résistance au mouvement. Il préfigure donc le principe d'inertie de la mécanique de Newton (*Questions sur les huit livres de la physique d'Aristote*) : « On cherche à savoir si un projectile, après avoir quitté la main du lanceur, est mû par l'air ou par quoi il est mû [...] Je juge cette question très difficile parce que Aristote, me semble-t-il, ne l'a pas bien résolue. Parce qu'il [...] soutient en un endroit que le projectile quitte rapidement le lieu où il était et que la nature ne permettant pas le vide envoie

rapidement de l'air derrière lui pour remplir le vide. L'air ainsi déplacé vient en contact avec le projectile et le pousse en avant. Cela se répète continuellement sur une certaine distance [...] Mais il me semble que plusieurs expériences montrent que cette méthode de procéder est sans valeur [...] » Plus loin, il considère l'exemple d'une « lance dont l'arrière aurait une forme conique aussi effilée que la pointe ; une telle lance, une fois partie, se déplacerait tout aussi rapidement que si l'arrière n'avait pas la forme conique. Mais il est certain que l'air qui suit la lance ne pourrait pas pousser de la manière indiquée sur une extrémité effilée puisque l'air serait aisément séparé par la pointe alors qu'elle pourrait pousser sur l'extrémité émoussée d'une lance et la déplacer ainsi vers l'avant.

Nous pouvons et nous devons donc dire que dans la pierre ou dans un autre projectile est imprimé quelque chose qui est la force motrice de ce projectile. Cela est évidemment mieux que de retomber dans l'affirmation que l'air doit continuer de mouvoir le projectile, car l'air paraît résister [...], le lanceur imprime un certain impetus ou force motrice dans le corps en mouvement, lequel impetus agit dans la direction vers laquelle le lanceur a mû le corps en mouvement, vers le haut ou vers le bas, latéralement ou circulairement. Et il imprimera dans ce corps un impetus dont le montant est le même que celui dont le moteur meut ce corps en mouvement plus rapidement. C'est par cet impetus que la pierre est mue après que le lanceur a cessé de la mouvoir. Mais cet impetus est continuellement diminué par l'air qui résiste et par la gravité de la pierre qui l'incline dans une direction contraire à celle dans laquelle l'impetus était naturellement disposé à la mouvoir. Ainsi le mouvement de la pierre devient plus lent jusqu'à ce que l'impetus soit tellement diminué ou corrompu que la gravité de la pierre l'emporte et fait descendre celle-ci en son lieu naturel ».

Buridan appliquera également cette théorie de l'impetus à la mécanique céleste pour expliquer le mouvement des astres (Dieu a imprimé son impetus à chaque astre lors de la création de l'univers) et la chute des corps (l'accélération est due à l'acquisition d'impetus sous l'effet de la gravitation¹⁰⁴).

La théorie du mouvement des projectiles de Buridan sera ensuite affinée par le scolastique parisien Albert de Saxe (1316-1390), qui distingue trois phases dans le mouvement :

1. une première phase durant laquelle l'impetus domine, induisant une trajectoire rectiligne
2. une deuxième phase, de transition, où l'impetus décline, donnant naissance à une trajectoire en forme d'arc de cercle
3. une troisième phase pendant laquelle la gravité domine et la trajectoire est verticale.

A la fin du XIV^e siècle, la dynamique de l'impetus a remplacé celle d'Aristote. Elle permettra les développements des savants italiens du XVI^e siècle. Elle jouera aussi un grand rôle dans la révolution copernicienne et le développement du modèle héliocentrique, puisqu'elle permet d'envisager le mouvement de la terre.

¹⁰⁴ Mais attention, il n'existe pas chez Buridan de théorie de la gravitation qui ressemble à celle de Newton !

12.3 La géométrisation du mouvement

12.3.1 Galilée : la chute des graves

L'œuvre de Galileo Galilei, dit Galilée (1564-1642) est immense, et couvre de très nombreux domaines, dont l'astronomie et la mécanique. On ne présentera ici que certains aspects des travaux du savant pisan, et principalement ceux associés à ses contributions à la formation de la mécanique moderne.

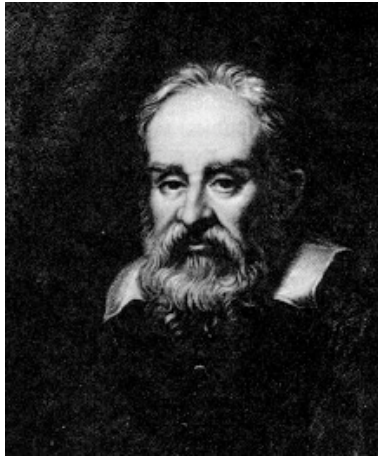


Figure 99: Galilée

Toutefois, notons quelques unes de ses autres contributions en mécanique :

- **1583** : isochronisme des petites oscillations du pendule. La légende veut, qu'après avoir observé les mouvements d'un lustre dans la cathédrale de Pise, Galilée ait eu l'idée qui le mena à montrer (par l'expérience) que la période d'un pendule ne dépend que de la longueur du fil (et pas de la masse du mobile)
- **1586** : mise au point d'une balance hydrostatique servant à la joaillerie. L'influence d'Archimède est grande dans cette invention.
- **1586-1587** : théorèmes sur le centre de gravité des corps, également très influencé par Archimède.
- **1593** : invention du thermoscope, sorte de thermomètre à gaz, malheureusement sensible aux variations de pression. Notons que les mesures de températures n'existent pas encore, et que de véritables échelles de température ne seront mises au point qu'au début du XVIIIe siècle.
- **1597** : invention du compas géométrique militaire, sorte de règle à calcul, que Galilée commercialisa à un prix très élevé. Il obtint gain de cause devant les tribunaux face à Baldassarre Capra, qui revendiquait la paternité de cette invention.
- **1612** : querelle des corps flottant avec le philosophe Ludovico delle Colombe, dit Colombo. Ce dernier défendait l'idée que la flottaison des corps est due à leur forme et non à leur densité. A l'inverse, Galilée entreprit de démontrer la justesse des conclusions d'Archimède, en se servant de la méthode expérimentale. Il faut noter qu'Archimède ne fournit aucune preuve de ce type. Galilée butta toutefois sur les effets de la tension superficielle¹⁰⁵.

Venons-en maintenant aux contributions à la science du mouvement. Dès 1590, dans son *De Motu* (Du Mouvement), Galilée, encore sous influence aristotélicienne, pose

¹⁰⁵ En effet, sous l'effet de la tension superficielle, des petits objets (aiguilles de métal, grains de poussière) peuvent flotter sur l'eau bien que leur masse volumique soit plus grande que celle de l'eau ! Ces effets moléculaires ne seront compris que beaucoup plus tard.

les bases d'une contestation de la mécanique d'Aristote et montre le besoin d'une mécanique nouvelle basée sur des expériences quantitatives et les mathématiques.

La réfutation des arguments d'Aristote concernant l'organisation du monde est le thème central de son livre *Dialogo sopra i due massimi sistemi del mondo – Tolemaico e Copernicano* (Dialogue sur les deux principaux systèmes du monde – le ptoléméen et le copernicien) paru en 1632. Galilée y défend l'homogénéité du monde (et donc prône l'abandon de l'opposition ciel/monde sublunaire). Il y invoque également le principe de relativité du mouvement, et se rapproche du principe d'inertie.

Sa grande œuvre mécanique est son ouvrage commencé en 1633 et paru en 1638 *Discorsi e dimonstrazioni matematiche intorno a due nuove scienze, attinenti alla meccanica e i movimenti locali* (Discours et démonstrations mathématiques concernant deux nouvelles sciences se rapportant à la mécanique et au mouvement local). Ces deux nouvelles sciences, en langage moderne, sont la résistance des matériaux et la cinématique.

La contribution de Galilée à la cinématique porte sur deux points : la chute libre des corps pesant (les graves) et la balistique (le mouvement des projectiles).

Concernant le premier point, Galilée remet en cause le principe de la mécanique d'Aristote qui dit que les corps lourds tombent plus vite que les corps légers. Il n'est pas le premier à réfuter ce principe aristotélicien, puisque Simon Stévin de Bruges avait en 1605 entamé une série d'expériences pour l'infirmer. Mais le génie de Galilée tient dans ce qu'il comprit que la chute des corps était un phénomène clé qui permettrait la compréhension des autres mouvements et aurait ainsi des répercussions en astronomie.

Il infirme le principe d'Aristote par l'absurde dans un passage de son dialogue entre trois personnages : Simplicio, qui défend les arguments traditionnels, Sagredo, l'honnête homme cultivé ouvert aux nouvelles idées, et enfin Salviati, qui représente Galilée.

« **SALVIATI** : *Je doute fort qu'Aristote ait jamais vérifié par expérience que deux pierres, l'une étant dix fois plus lourde que l'autre, lâchées au même instant d'une hauteur de cent brasses, tomberaient à des vitesses si inégales qu'au moment où l'une toucherait le sol, l'autre n'aurait encore parcouru que dix brasses.*

SIMPLICIO : *Ses propres paroles nous montrent pourtant qu'il l'a expérimenté, puisqu'il dit : Nous voyons que le plus pesant ... Ce nous voyons fait allusion à une expérience.*

SAGREDO : *Mais moi, signor Simplicio, je vous assure que j'ai l'épreuve qu'un boulet de canon de cent, de deux cents livres, ou plus encore, n'aura pas pris l'avance d'une palme, à son arrivée au terrain, sur une balle de mousquet d'une demi-livre, la hauteur de chute fût-elle de deux cents brasses.*

SALVIATI : *Mais, sans autre expérience, on prouvera clairement, par une démonstration courte et concluante, qu'il n'est pas vrai qu'un mobile lourd se meuve*

plus vite qu'un mobile moins lourd, pourvu que tous deux soient de la même matière, comme ceux, en somme, dont parle Aristote. Aussi vous demanderai-je, signor Simplicio, si vous admettez qu'à tout corps grave, en chute libre, corresponde une vitesse déterminée, telle qu'on ne puisse l'augmenter ou la réduire sinon par l'effet de quelque violence ou en lui opposant quelque résistance.

SIMPLICIO : *Il n'est pas douteux qu'un même mobile dans un même milieu ait, par nature, une vitesse réglée et déterminée qui ne pourra être accrue sinon par un impetus nouveau, ni diminuée sinon par un retard dû à un obstacle quelconque.*

SALVIATI : *Si donc nous avons deux mobiles dont les vitesses naturelles seraient inégales et si nous les joignons ensemble, il est manifeste que, dans une certaine mesure, le plus rapide serait retardé par le plus lent et le plus lent accéléré par le plus rapide. N'êtes-vous pas de mon avis ?*

SIMPLICIO : *Je pense que les choses se passeraient ainsi, très certainement.*

SALVIATI : *Mais s'il en est ainsi, et s'il est vrai, d'autre part, qu'une grosse pierre se meut à une vitesse disons de huit degrés, et une petite pierre à une vitesse de quatre degrés, leur composé devra se mouvoir à une vitesse inférieure à huit degrés, mais les deux pierres conjointes en font une plus grosse que la première qui se mouvait à huit degrés de vitesse ; donc, ce composé (qui pourtant est plus grand que la première pierre toute seule) se mouvra plus lentement que cette première pierre, pourtant moindre, ce qui est contraire à votre hypothèse. Vous voyez donc comment, si je suppose que le mobile le plus lourd se meut plus vite que le moins lourd, j'en arrive à conclure que c'est le plus lourd qui se meut le moins vite.*

SIMPLICIO : *Oh ! Voilà qui passe tout à fait ma conception. Votre raisonnement est en vérité fort bien conduit ; toutefois, j'ai de la peine à croire qu'une larme de plomb puisse se mouvoir aussi vite qu'un boulet d'artillerie.*

SALVIATI : *Vous auriez dû dire : un grain de sable aussi vite qu'une meule de moulin. Je ne voudrais pas, signor Simplicio, qu'à l'exemple de tant d'autres, détournant notre propos de son objet principal, vous attachiez à telle chose que j'ai dite et qui s'écartait de la vérité de l'épaisseur d'un cheveu, pour dissimuler, sous ce cheveu, l'erreur, aussi grosse qu'une amarre, qu'un autre a commise, Aristote dit : « Une boule de fer de cent livres, tombant d'une hauteur de cent brasses, arrive au sol avant qu'une boule d'une livre soit descendue d'une seule brasse ». Je dis, moi, qu'elles arrivent en même temps ; vous faites l'expérience et vous constatez que la plus grosse devance de deux doigts la plus petite, c'est-à-dire qu'au moment où la plus grosse boule touche la terre l'autre en est encore éloignée de deux doigts ; et vous voudriez maintenant, derrière ces deux doigts, cacher les quatre vingt dix neuf brasses d'Aristote, et, relevant mon erreur minime, passer sous silence une erreur énorme. »*

Cette très légère différence est due, nous le savons aujourd'hui, à la résistance de l'air. Mais un point méthodologique important est que Galilée introduit ici le **concept d'expérience idéalisée**, c'est-à-dire d'expérience dans laquelle seuls les aspects pertinents sont retenus. Ici, le frottement de l'air sur les mobiles est négligé, car il ne

participe pas du phénomène étudié. L'analyse de Galilée porte donc sur la **chute libre**, cas idéal réalisé lorsque les mobiles se déplacent dans le vide.

Galilée part du principe que la chute est un mouvement accéléré (et donc n'a pas de vitesse constante a priori, l'existence d'une vitesse de chute terminale étant associée à la résistance de l'air), et il va considérer possible et définir le **mouvement rectiligne uniformément accéléré**. Ainsi, dès 1604, il énonce les deux affirmations suivantes :

Enoncés de Galilée en 1604

- « *Je dis qu'un mouvement est uniformément accéléré quand, partant du repos, il reçoit en des temps égaux, des moments [des accroissements] égaux de vitesse* ». Il s'agit ici de la définition rigoureuse du mouvement rectiligne uniformément accéléré.
- « *Les corps en chute libre ont un mouvement rectiligne uniformément accéléré* »

Le second énoncé a pour conséquence directe (un calcul algébrique trivial le montre¹⁰⁶) que la distance parcourue lors d'un mouvement rectiligne uniformément accéléré varie comme le carré du temps écoulé depuis la mise en mouvement.

Comment vérifier cela ? Par l'expérience. Mais où trouver un endroit qui permette un temps de parcours suffisamment long et des mesures assez précises ? La légende, sans doute fautive, veut que Galilée ait d'abord vérifié sa théorie en lâchant des objets du haut de la tour de Pise. De manière plus subtile, il va mettre au point un dispositif expérimental qui lui permet de faire varier à volonté la valeur de l'accélération. Il fait donner la description de son montage expérimental par Salviati :

« Dans une règle ou une volige de bois d'une longueur d'environ douze brasses, d'une épaisseur d'une demi-brasse d'un côté et de trois doigts de large nous avons taillé une rainure un peu plus large qu'un doigt et très droite, et pour la rendre bien lisse, nous y avons collé un parchemin poli et lustré au possible. Après avoir donné à notre planche une certaine inclinaison en élevant une de ses extrémités d'une brasse ou deux au-dessus de l'horizontale, nous avons fait descendre dans la rainure une bille de bronze très dure, bien ronde et polie, en notant, suivant un mode que je décrirai ci-après, le temps qu'elle utilisait pour un voyage complet. Nous avons répété ce même acte beaucoup de fois pour bien nous assurer du temps nécessaire, de sorte que la marge d'erreur ne dépasse pas un dixième de battement de pouls. Après avoir fait descendre la même bille seulement sur un quart de la longueur de ce canal et mesuré le temps de sa descente, nous l'avons trouvé exactement égal à la moitié du temps trouvé antérieurement. Ensuite, nous avons refait l'expérience avec d'autres subdivisions et comparé le temps pour la longueur tout entière avec le temps pour la moitié, avec celui pour les deux tiers ou les trois quarts, bref pour n'importe quelle autre fraction, et dans toutes ces expériences, nous avons trouvé que les espaces traversés avaient le même rapport entre eux que le carré des temps, et cela pour toutes les inclinaisons du plan, c'est-à-dire de la rainure dans laquelle nous faisons descendre la bille ... »

¹⁰⁶ Le calcul différentiel moderne permet de faire cela immédiatement. Mais, à l'époque de Galilée, il n'était pas encore inventé ! Le développement analytique est donc un peu plus complexe.

[...] Pour mesurer le temps, nous prenions un grand seau rempli d'eau que nous attachions assez haut ; par un orifice étroit pratiqué dans son fond s'échappait un mince filet d'eau que l'on recueillait dans un petit récipient, tout le temps que la bille descendait dans la rainure. Les quantités ainsi recueillies étaient à chaque fois pesées à l'aide d'une balance très sensible, et les différences et proportions entre les poids nous donnaient les différences et proportions entre les temps ; la précision de l'expérience était telle que, comme je l'ai dit, aucune discordance significative n'apparut jamais entre les expériences, maintes et maintes fois répétées. »

La modernité de cette description est frappante. En effet, on y retrouve tous les concepts rencontrés lors de la discussion de la mesure : répétabilité, fiabilité, ...

Ce qui fait que Galilée est considéré comme l'un des pères fondateurs de la méthode expérimentale. Dès 1608, il mesure la vitesse finale en prolongeant le plan incliné par un plan horizontal, sur lequel il fait l'hypothèse que le mobile suit un **mouvement rectiligne uniforme**. Cette hypothèse lui fait anticiper le **principe d'inertie** :

« Un degré de vitesse quelconque, une fois communiqué à un mobile, s'imprime en lui de façon indélébile du seul fait de sa nature, et pourvu que soient supprimées les causes extérieures d'accélération et de ralentissement, ce qui n'a lieu que sur un plan horizontal ; sur un plan descendant, en effet, il existe déjà une cause d'accélération, et sur un plan ascendant une cause de ralentissement ; d'où il suit que le mouvement sur un plan horizontal est éternel ; car s'il est uniforme, il ne s'affaiblit ni ne diminue, et encore moins ne cesse. »

Ces expériences sur le plan incliné permettent également d'énoncer la loi suivante :

Tous les corps en chute libre ont la même accélération quelle que soit leur masse.

Après la chute libre, Galilée s'attaque au problème du mouvement des projectiles, et apportera une pierre fondatrice de la balistique. Il est en effet le premier à comprendre que le mouvement des projectiles est un mouvement composé : la composante verticale correspond à un mouvement de chute, tandis que le mouvement horizontal est un mouvement rectiligne uniforme (tout ceci en l'absence des effets de résistance de l'air). Pour vérifier cela, il imagine une nouvelle expérience :

« J'imagine qu'un mobile a été lancé sur un plan horizontal d'où l'on a écarté tout obstacle ; [...] le mobile que j'imagine doué d'une certaine gravité, parvenu à l'extrémité du plan et continuant sa course, ajoutera à son précédent mouvement uniforme et indélébile la tendance vers le bas que lui confère sa gravité : le résultat sera un mouvement composé d'un mouvement horizontal uniforme et d'un mouvement naturellement accéléré vers le bas que j'appelle projection. Nous démontrerons maintenant quelques unes de ses propriétés dont voici la première :

Théorème I – Proposition I : Un projectile qu'entraîne un mouvement composé d'un mouvement horizontal uniforme et d'un mouvement naturellement accéléré vers le bas, décrit au cours de son déplacement une trajectoire semi-parabolique. »

Une conséquence importante de cette décomposition que Galilée mettra en avant est que la trajectoire des projectiles est parabolique, et n'est pas un raccord de segments de droite et d'arcs de cercle.

12.3.2 Descartes : les lois du mouvement

René Descartes (1596-1650) est un mathématicien et un philosophe, dont l'œuvre en mécanique est moins développée que celle de Galilée. Sa physique, tout comme sa philosophie, est très emprunte de considérations religieuses.



Figure 100: R. Descartes

Avant de discuter de sa contribution au développement de la mécanique, revenons sur les points essentiels de sa physique. Tout d'abord, l'espace de Descartes n'est pas vide : il est rempli de tourbillons, qui sont responsables des mouvements des corps célestes. Descartes est un partisan de la théorie des éléments (qu'il « décrit à sa mode ») :

« Je conçois le premier, qu'on peut nommer l'Élément du Feu, comme une liqueur, la plus subtile et la plus pénétrante qui soit au Monde. Et ensuite de ce qui a été dit ci-dessus touchant à la

nature des corps liquides, je m'imagine que ses parties sont beaucoup plus petites, et se remuent beaucoup plus vite qu'aucune de celles des autres corps [...]

Pour le second, qu'on peut prendre pour l'Élément de l'Air, je le conçois bien aussi comme une liqueur très subtile, en le comparant avec le troisième ; mais, pour le comparer avec le premier, il est besoin d'attribuer quelque grosseur, et quelque figure, à chacune de ses parties et de les imaginer à peu près toutes rondes et jointes ensemble, ainsi que les grains de sable, et de poussière [...]

Après ces deux éléments, je n'en reçois plus qu'un troisième ; à savoir celui de la Terre, duquel je juge que les parties sont d'autant plus grosses et se remuent d'autant moins vite, à comparaison de celles du second, que font celles-ci à comparaison de celles du premier. Et même je crois que c'est assez de le concevoir comme une ou plusieurs grosses masses, dont les parties n'ont que fort peu ou point du tout de mouvement, qui leur fasse changer de situation à l'égard l'une de l'autre. »

La mécanique de Descartes est basée sur l'idée que l'Univers a été créé par Dieu puis livré à lui-même. Puisque le mouvement est continu depuis la création, Descartes développe l'idée de **conservation**, et préfigure par cela les lois de conservation.

Ainsi, dans son *Monde ou Traité de la Lumière*, qu'il renonce à publier après la condamnation de Galilée en 1633 et qui ne paraîtra qu'à titre posthume en 1664, Descartes énonce trois règles de la nature, qu'il retravaille sous la forme de lois de la nature dans ses *Principia Philosophiae* de 1644.

Première règle de la physique de Descartes

« La première est : que chaque partie de la matière, en particulier, continue toujours d'être en même état, pendant que la rencontre des autres ne la contraint point de changer. C'est-à-dire : si elle a quelque grosseur, elle ne deviendra jamais plus petite, sinon que les autres la divisent ; si elle est ronde ou carrée, elle ne changera jamais cette figure sans que les autres l'y contraignent ; si elle est arrêtée en quelque lieu, elle n'en partira jamais que les autres ne l'en chassent ; et si elle a une fois commencé à se mouvoir, elle continuera toujours avec une égale force jusqu'à ce que les autres l'arrêtent ou la retardent. »

Cette règle correspond à la première loi donnée dans la seconde partie du même ouvrage : *« La première loi de la nature : que chaque chose demeure en l'état qu'elle est, pendant que rien ne la change ... »*. Cette loi est également un principe de causalité : tout changement a une cause, rien n'est produit sans cause. Notons également que le terme de force est flou chez Descartes. Une autre implication importante est la rupture avant la théorie scolastique qui oppose mouvement naturel et mouvement violent. Cela se traduit par une conception du repos plus moderne : *« Et moi je conçois que le repos est aussi bien une qualité qui doit être attribuée à la matière pendant qu'elle demeure en une place, comme le mouvement en est une qui lui est attribuée, pendant qu'elle en change. »*

La seconde règle de Descartes traite de la mise en mouvement :

Deuxième règle de la physique de Descartes

« Je suppose pour seconde règle : que quand un corps en pousse un autre, il ne saurait lui donner aucun mouvement, qu'il n'en perde en même temps autant du sien ; ni lui en ôter, que le sien ne s'augmente d'autant. Cette règle jointe avec la précédente, se rapporte fort bien à toutes les expériences, dans lesquelles nous voyons qu'un corps commence ou cesse de se mouvoir, parce qu'il est poussé ou arrêté par quelque autre. »

La troisième règle est la suivante :

Troisième règle de la physique de Descartes

« J'ajouterai pour la troisième : que lorsqu'un corps se meut, encore que son mouvement se fasse le plus souvent en ligne courbe et qu'il ne s'en puisse jamais faire aucun, qui ne soit en quelque façon circulaire, ainsi qu'il a été dit ci-dessus, toutefois chacune de ces parties en particulier tend toujours à continuer le sien en ligne droite. Et ainsi leur action, c'est-à-dire l'inclination qu'elles ont à se mouvoir, est différente de leur mouvement. »

La combinaison de cette nouvelle règle avec la précédente permet à Descartes d'énoncer plus tard sa seconde loi de la physique : *« La deuxième loi de la nature : que tout corps qui se meut, tend à continuer son mouvement en ligne droite »*. Il s'agit du principe d'inertie. Notons que Descartes n'associe pas directement ici la notion d'inertie et celle de mouvement rectiligne uniforme. Il réalise cependant cette association dans une lettre de février 1643 à Huygens : *« Sur quoi je considère que la nature du mouvement est telle que, lorsqu'un corps a commencé à se mouvoir, cela*

suffit pour faire qu'il continue toujours après avec une même vitesse et en même ligne droite, jusqu'à ce qu'il soit arrêté ou détourné par quelque autre cause. »

Outre la notion de loi de conservation et le principe d'inertie, Descartes fait mention d'un troisième élément important : la notion d'inclination ou de tendance à se mouvoir, qui lui permet de distinguer le mouvement rectiligne uniforme des autres.

Une conséquence est la rupture avec l'idée aristotélicienne, reprise par les scolastiques, que le mouvement circulaire est naturel : *« De même, quand on fait tourner une pierre dans une fronde, non seulement elle va tout droit aussitôt qu'elle en est sortie, mais de plus, pendant tout le temps qu'elle y est, elle presse le milieu de la fronde, et fait tendre la corde ; montrant évidemment par là qu'elle a toujours inclination d'aller en droite ligne et qu'elle ne va en rond que par contrainte. »* Mais attention : la position de Descartes ne résulte pas d'une analyse mathématique des équations de la dynamique, qui ne sont pas encore posées. Ici aussi, c'est la référence à Dieu et à l'idée que le mouvement rectiligne uniforme est le plus simple à décrire qui sont à la base de son argumentation.

Les considérations de Descartes sont générales, et portent sur tous les mouvements. Descartes n'a pas, comme Galilée, produit d'étude détaillée de la chute des graves. Il donne toutefois une explication pour la pesanteur, qui n'est pour lui ni une qualité de la matière ni le résultat d'une attraction, mais une force exercée par le ciel : *« Je désire maintenant que vous considériez quelle est la pesanteur de cette Terre, c'est-à-dire la force qui unit toutes ses parties et qui fait qu'elles tendent toutes vers son centre, chacune plus ou moins, selon qu'elles sont plus ou moins grosses et solides ; laquelle n'est autre et ne consiste qu'en ce que les particules du petit Ciel qui l'environne, tournant beaucoup plus vite que les siennes autour de son centre, tendent aussi avec plus de force à s'en éloigner et par conséquent les y repoussent. »*



Figure 101: C. Huygens

12.3.3 Huygens : l'organisation déductive des sciences du mouvement et les chocs

L'apport du savant hollandais Christiaan Huygens (1629-1695) à la science du mouvement va consister à rapprocher et combiner le mouvement des graves tel qu'il a été géométrisé par Galilée et la science du mouvement telle qu'elle a été systématisée par Descartes. Tout comme Galilée, et à l'inverse de Descartes, Huygens, appelé par Colbert pour diriger l'Académie Royale des Sciences à sa création en 1666¹⁰⁷ sur demande de Louis XIV, est un théoricien doublé d'un ingénieur. On lui doit, par exemple, l'invention de l'horloge à pendule en 1673 et de la montre à ressort spiral réglant en 1675. Il aura également une forte activité en mathématiques, en astronomie et en optique. Nous

¹⁰⁷ Il quittera Paris en 1683 à la mort de Colbert, et n'y reviendra plus après la révocation de l'édit de Nantes en 1685.

allons porter ici notre attention sur trois thèmes particuliers : la chute des graves, le mouvement circulaire et les chocs élastiques.

Le travail de Huygens se base sur ce qu'il nomme trois « hypothèses », énoncées en 1673 dans son *Horologium oscillatorium*:

Les 3 hypothèses de Huygens

1. *Si la gravité n'existait pas et qu'aucune résistance d'air ne s'opposait au mouvement des corps, chacun d'eux continuerait son mouvement avec une vitesse uniforme en suivant une ligne droite.*
2. *Mais maintenant il arrive par l'action de la gravité, de quelque cause qu'elle provienne, que les corps se meuvent d'un mouvement composé de leur mouvement uniforme dans une direction quelconque et de celui de haut en bas qui est dû à la gravité.*
3. *On peut considérer ces deux mouvements séparément et l'un n'est pas empêché par l'autre.*

On reconnaît ici le principe d'inertie. Rappelons que ce principe n'avait pas été complètement énoncé par Galilée, qui n'arrivait pas à concevoir un corps sans gravité et pour qui le mouvement circulaire gardait un statut à part. Il avait été en revanche énoncé par Descartes, qui avait clairement émis le principe de conservation du mouvement rectiligne uniforme.

Concernant le mouvement des projectiles, la situation est différente : ici, Galilée avait clairement dégagé le principe de la composition des mouvements et identifié la nature parabolique de la trajectoire. L'apport de Huygens réside dans sa compréhension de l'importance de ces résultats pour analyser la chute des graves, qui va le conduire à « déduire la cause et les lois de l'accélération des corps pesants ». Il s'agit ici d'un traitement mathématique systématisé et unifiant, fondé sur la géométrie euclidienne et l'algèbre. Huygens donne ses résultats sous la forme de propositions, dont voici les trois premières:

Les 3 premières propositions de Huygens sur la chute libre

- Proposition I : *Dans des temps égaux les accroissements de la vitesse d'un corps tombant sont toujours égaux et les espaces parcourus durant des temps égaux depuis le commencement de la chute forment une série dont les différences successives sont constantes.*
- Proposition II : *L'espace parcouru pendant un certain temps par un corps qui commence sa chute en partant du repos est la moitié de l'espace que ce corps pourrait parcourir d'un mouvement uniforme avec la vitesse acquise par la chute au bout du temps considéré.¹⁰⁸*
- Proposition III : *Deux espaces parcourus par un corps tombant dans des temps quelconques, dont chacun est pris depuis le commencement de la chute, sont entre eux comme les carrés de ces temps, ou bien comme les carrés des vitesses acquises.*

La seconde contribution principale de Huygens réside dans son analyse du mouvement circulaire, qui, depuis Aristote jusqu'au début du XVII^e siècle, a

¹⁰⁸ Il s'agit ici du théorème du degré moyen, déjà explicitement donné par Galilée

conservé un statut particulier. Cette dominance du mouvement circulaire est par exemple encore présente chez Nicolas Copernic (1474-1543). Des brèches ont été ouvertes par Johannes Kepler (1571-1630) et Galilée, mais c'est Descartes qui va ouvrir la voie à la remise en cause de l'idée selon laquelle le mouvement circulaire est « naturel ». C'est sur la base cartésienne que va se développer l'analyse de Huygens, dont le traité sur la force centrifuge, *De vi Centrifuga*, rédigé en 1659, ne paraîtra à titre posthume qu'en 1703. Les propositions II et III de Huygens fournissent une évaluation quantitative de la force centrifuge en fonction de la vitesse de rotation et du rayon du cercle décrit :

Propositions de Huygens sur le mouvement circulaire

- Proposition II : *Lorsque des mobiles égaux tournent dans les mêmes ou d'égales circonférences ou roues avec des vitesses différentes mais l'un et l'autre d'un mouvement uniforme, la force centrifuge du plus rapide sera à celle du plus lent dans un rapport égal à celui du carré des vitesses.*
- Proposition III : *Lorsque deux mobiles égaux se meuvent avec la même vitesse suivant des circonférences inégales, leurs forces centrifuges seront inversement proportionnelles aux diamètres, de sorte que dans le cas de la plus petite circonférence la force nommée est la plus grande.*

On le voit, le mouvement circulaire a perdu son statut naturel hérité de la physique qualitative d'Aristote. Il s'agit maintenant d'un mouvement contraint.

Venons-en maintenant aux travaux de Huygens sur ce que nous appelons aujourd'hui les chocs élastiques. En 1668, la Royal Society of London prend contact avec plusieurs savants renommés pour trouver « les lois de la communication du mouvement » lors du choc entre deux corps solides. Le 18 mars 1669, Huygens énonce sept règles du mouvement dans le *Journal des Sçavans*. La sixième règle est énoncée comme suit :

6^e règle de Huygens concernant les chocs

- 6. La somme des produits faits de la grandeur de chaque corps dur, multipliée par le carré de sa vitesse, est toujours la même avant et après leur rencontre.

On reconnaît ici la loi de conservation de l'énergie cinétique. Mais il faut noter que cette quantité n'est pas nommée par Huygens, et cela pour la bonne raison qu'elle n'est pas encore identifiée à cette époque ! Leibniz, en 1692, verra dans cette quantité conservée une mesure de la « force absolue », qu'il appelle également « force vive »¹⁰⁹.

¹⁰⁹ Notons ici que le terme « force vive » fut l'objet de nombreuses polémiques, tant qu'il n'a pas été défini clairement. On peut notamment retenir la querelle entre les cartésiens et les leibniziens, qui trouve son origine dans l'article de Leibniz publié en 1686 dans le journal *Acta Eruditorum* et intitulé « Brève démonstration d'une erreur mémorable de Descartes et d'autres sur une loi naturelle selon laquelle ils soutiennent que Dieu conserve toujours une même quantité de mouvement et qu'ils ont le tort d'appliquer à toute la mécanique ». L'opposition vient de ce que Descartes soutient que c'est la quantité de mouvement qui est conservée lors d'un choc élastique, alors que Leibniz défend que c'est l'énergie cinétique qui est invariante, donnant naissance à deux définitions de la force vive. Il est remarquable que dans le cas des chocs élastiques les deux quantités sont invariantes !

Rappelons que la controverse très vive qui a opposé Huygens et Leibniz ne portait pas sur cette loi de conservation, mais sur l'explication donnée pour la transmission du mouvement. Alors que Leibniz défendait une analyse du choc basée sur un modèle déduit du ressort, Huygens ne voulait concevoir que des corps infiniment rigides. Ici, c'est le caractère flou de la notion de dureté d'un corps solide qui est à l'origine de la controverse. Cette controverse marquera les travaux bien au-delà de la mort de Huygens en 1695, jusqu'à ce que les développements en mécanique des milieux continus précisent la description des propriétés mécaniques des solides.

12.4 La science newtonienne

La contribution d'Isaac Newton (1643-1727) va consister à organiser la dynamique en la réduisant sous la forme de l'énoncé de trois lois, qui constituent la base de l'édifice de la mécanique classique. De plus, il va donner la solution du problème de la chute des graves en trouvant la loi de la gravitation universelle. Son œuvre s'étend bien au-delà de la mécanique : il apporte également des contributions majeures en mathématiques et en optique.



Figure 102: I. Newton

Ses travaux en mécanique sont regroupés dans son ouvrage *Philosophiae Naturalis Principia Mathematica* (Principes mathématiques de la philosophie naturelle) paru en 1687 (mais écrit plus tôt : depuis sa controverse sur la paternité de la loi de la gravitation universelle avec Robert Hooke (1365-1703) en 1671, Newton avait décidé de ne plus rien publier ! C'est son ami le grand astronome Edmund Halley (1656-1742) qui le convaincra de divulguer ses travaux). Ce livre pose les bases de la mécanique classique, et cela dans la plupart de ses branches.

La première partie de l'ouvrage porte sur la définition de concepts généraux : Newton y donne des définitions claires des concepts fondamentaux de la mécanique. Après les concepts de base :

Définitions de Newton

- définition de la masse : « *La quantité de matière se mesure par la densité et le volume pris ensemble* », « *je désigne la quantité de matière par les mots de corps ou de masse* »
- définition de la quantité de mouvement : « *La quantité de mouvement est le produit de la masse par la vitesse* »

apparaissent les définitions de plusieurs forces :

définitions des forces chez Newton

- « *La force interne de la matière (vis insita) est le pouvoir de résistance, par lequel chaque corps persévère, autant qu'il est en lui de le faire, dans son état actuel de repos ou de mouvement uniforme en ligne droite* ». Cette force

interne, également appelée par Newton « force d'inertie », ne doit pas être confondue avec une force telle que nous la concevons aujourd'hui, ni avec la « force vive » de Leibniz. Son rôle ne s'éclaire que par rapport à l'énoncé de la loi d'inertie, qui sera donné plus bas.

- « *La force imprimée est une action exercée sur le corps, qui a pour effet de changer son état de repos ou de mouvement rectiligne uniforme. Cette force consiste dans l'action seule, et elle ne persiste pas dans le corps dès que l'action vient à cesser. En effet, le corps persévère dans son nouvel état par la seule force d'inertie. En outre, la force imprimée a des origines diverses comme le choc, la pression, la force centripète.* »
- « *La force centripète (vis centripeta) est celle par laquelle des corps sont tirés, poussés ou tendent de quelque façon que ce soit vers quelque point, comme vers un centre* »

On peut noter l'absence d'une définition exacte de la vitesse, bien que ce concept soit employé par Newton. Par ailleurs, alors que Descartes et Huygens orientent leurs analyses sur la force centrifuge (qui est la conséquence d'un mouvement de rotation), Newton axe la sienne sur la force centripète, qui cause le mouvement.

D'autres contributions majeures de Newton sont ses définitions de l'espace absolu et du temps absolu :

Temps, espace, lieu et mouvement chez Newton

- « *I. Le temps absolu, vrai et mathématique, sans relation à rien d'extérieur, coule uniformément, et s'appelle durée. Le temps relatif, apparent et vulgaire, est cette mesure sensible et externe d'une partie de durée quelconque (égale ou inégale) prise du mouvement : telles sont les mesures d'heures, de jours, de mois, etc. dont on se sert ordinairement à la place du temps vrai.* »
- « *II. L'espace absolu, sans relation aux choses externes, demeure toujours similaire et immobile. L'espace relatif est cette mesure ou dimension mobile de l'espace absolu, laquelle tombe sous nos sens par sa relation aux corps, et que le vulgaire confond avec l'espace immobile. C'est ainsi, par exemple, qu'un espace, pris au-dedans de la terre ou dans le ciel, est déterminé par la situation qu'il a à l'égard de la terre. L'espace absolu et l'espace relatif sont les mêmes d'espèce et de grandeur ; mais ils ne le sont pas toujours de nombre ; car, par exemple, lorsque la terre change de place dans l'espace, l'espace qui contient notre air demeure le même par rapport à la terre, quoique l'air occupe nécessairement les différentes parties de l'espace dans lesquelles il passe, et qu'il change réellement sans cesse* ».
- « *III. Le lieu est la partie de l'espace occupée par un corps et par rapport à l'espace, il est relatif ou absolu [...]* »
- « *IV. Le mouvement absolu est la translation des corps d'un lieu absolu dans un autre lieu absolu, et le mouvement relatif est la translation d'un lieu relatif dans un autre lieu relatif [...]* »

L'espace de Newton est homogène et infini : un corps peut donc y observer un mouvement rectiligne uniforme sur des temps quelconques. De plus, il est vide (au contraire de l'espace de Descartes, qui est rempli de tourbillons).

Après les définitions, l'ouvrage de Newton est composé de trois livres. Le premier, appelé « Axiomes et lois du mouvement », présente les trois fameuses lois de la dynamique. Leurs formulations originales sont :

Les 3 lois de la dynamique de Newton

- « *Première loi : Tout corps persévère dans son état de repos ou de mouvement rectiligne uniforme à moins que des forces imprimées ne le contraignent à changer son état.* »
- « *Deuxième loi : Le changement du mouvement est proportionnel à la force motrice imprimée et se fait suivant la droite par laquelle cette force est imprimée.* »
- « *Troisième loi : L'action est toujours égale et opposée à la réaction : c'est-à-dire que les actions mutuelles de deux corps sont toujours égales et dirigées en sens contraire.* »

Ces trois lois appellent quelques remarques. Tout d'abord, elles ne sont pas écrites sous leur forme différentielle moderne. Newton n'a pas recours au calcul différentiel dans ces chapitres, et l'usage de ces outils mathématiques ne sera généralisé que plus tard par d'autres savants. Par exemple, dans l'expression de la seconde loi, il ne parle pas de temps ni d'incrément de temps. La forme originale de cette loi est donc plus proche de la description de la transmission du mouvement par un choc que de sa forme moderne. Ensuite, la troisième loi est formulée chez Newton assez tardivement, puisqu'elle n'apparaît pas dans les textes préliminaires datés de 1685.

Enfin, il faut noter que Newton, pour traiter du cas où plusieurs forces sont imprimées à un corps, édicte une loi de composition des mouvements, et non une loi de composition des forces : « *sous l'action de forces conjointes, un corps décrit la diagonale du parallélogramme, dans le même temps qu'il mettrait à décrire ses côtés, si elles agissaient séparément* ».

Passons maintenant à l'étude des chocs élastiques. Newton énonce clairement la conservation de la quantité de mouvement : « *La quantité de mouvement que l'on obtient en prenant la somme des mouvements qui se font vers le même côté, et la différence de ceux qui se font vers des côtés opposés, n'est pas changée par l'action des corps entre eux.* » Il édicte ensuite l'invariance du mouvement du centre de masse : « *Le centre commun de gravité de deux ou plusieurs corps, ne change pas son état de mouvement ou de repos, par les actions des corps entre eux ; et, en conséquence, le centre commun de gravité de corps agissant tous les uns sur les autres (en l'absence d'actions et d'empêchements extérieurs), est ou bien en repos, ou bien en mouvement rectiligne uniforme.* »

Une autre contribution très importante est l'introduction des repères inertiels : « *Les mouvements des corps inclus dans un espace donné sont les mêmes, entre eux, que cet espace soit au repos, ou qu'il se meuve uniformément en ligne droite sans mouvement circulaire.* »

Le deuxième livre des *Principia* traite du mouvement des corps dans un milieu résistant, et introduit les bases de la mécanique des fluides.

Le troisième livre, intitulé « du système du monde », présente la cosmologie de Newton. C'est ici qu'il introduit la loi de la gravitation universelle, au sein d'une théorie des forces centrales. Après avoir étudié théoriquement les mouvements de corps s'attirant mutuellement, Newton émet l'hypothèse que, puisque le mouvement des planètes correspond à ce type de solution, celles-ci s'attirent mutuellement, et que le moteur des mouvements célestes est une force de gravitation universelle :

Loi de la gravitation universelle de Newton

« Proposition VII. Théorème VII. La gravité appartient à tous les corps, et elle est proportionnelle à la quantité de matière que chaque corps contient.

On a prouvé ci-dessus que toutes les planètes gravitent mutuellement les unes vers les autres : que la gravité vers une planète quelconque, considérée à part, est réciproquement comme le carré de la distance au centre de cette planète : et que par conséquent la gravité dans toutes les planètes est proportionnelle à leur quantité de matière. Mais comme toutes les parties d'une planète quelconque A, pèsent sur une autre planète quelconque B, que la gravité d'une partie quelconque est à la gravité du tout, comme la matière de la partie est à la matière totale, et que, par la troisième loi du mouvement, l'action et la réaction sont toujours égales ; la planète B gravitera à son tour vers toutes les parties de la planète A, et sa gravité vers une partie quelconque sera à sa gravité vers toute la planète, comme la matière de cette partie à la matière totale.

Cor. 1. La gravité vers toute une planète, est donc composée de la gravité vers toutes ses parties. Nous en avons des exemples dans les attractions magnétiques et électriques. Car l'attraction vers le tout est composée des attractions vers chacune des parties. On verra qu'il en est de même dans la gravité, en supposant que plusieurs petites planètes s'unissent en un globe, et forment une grosse planète. Car on conçoit aisément par là que la force totale doit naître de la force des parties composantes. Si quelqu'un objecte que selon cette loi tous les corps d'ici bas devraient graviter les uns vers les autres, et que cependant cette gravité mutuelle n'est pas sensible : je répondrai, que cette gravité mutuelle des corps étant à leur gravité vers la terre, comme la masse de ces corps à la masse de la terre, elle n'est pas à beaucoup près assez forte pour pouvoir être aperçue.

Cor. 2. La gravité vers chaque particule égale d'un corps, est réciproquement comme le carré des distances des lieux de ces particules. »

Outre le mouvement des planètes, Newton utilise sa nouvelle loi de la gravitation universelle pour étudier le mouvement des comètes, pour analyser les marées et l'aplatissement de la terre aux pôles. Ce dernier phénomène sera vérifié par l'expédition en Laponie de 1736-1737 dirigée par Pierre-Louis Moreau de Maupertuis et celle au Pérou de 1735-1744 dirigée par Charles-Marie de la Condamine.

La forme moderne de la loi de la gravitation universelle est

$$F = G \frac{m_1 m_2}{r^2}$$

où G est une constante universelle¹¹⁰, la distance des centres de masse des deux corps de masses respectives m_1 et m_2 . La première mesure de la valeur de la constante de gravitation universelle ne sera réalisée que plus d'un siècle plus tard, en 1798, par Henry Cavendish (1731-1810), qui s'exclama « *j'ai pesé la terre !* »¹¹¹.

Si la loi de la gravitation universelle ouvre de nombreuses possibilités pour comprendre des phénomènes, son emploi peut se heurter à des difficultés mathématiques. Outre le problème des solutions chaotiques analysé par Poincaré en 1890, il est démontré depuis 1912 que le **problème des trois corps** n'admet pas de solution analytique générale.

12.5 L'algorithmisation de la science du mouvement

12.5.1 Leibniz : le calcul différentiel et intégral



Figure 103: G.W. Leibniz

La mécanique classique, telle qu'elle a été formalisée par Newton, ne fait pas usage du calcul différentiel et intégral, et cela bien que Newton ait par ailleurs développé les bases de ces méthodes.

C'est l'œuvre du mathématicien et philosophe Gottfried Wilhelm Leibniz (1647-1716) qui enrichira la mécanique classique. Dès 1676, Leibniz a développé les bases de sa théorie du calcul. Pourtant, il ne les publiera qu'en 1684 dans un article de six pages dans la revue *Acta Eruditorum* éditée à Leipzig. C'est dans cet article fondateur qu'il introduit les « *differentia* » (petits éléments infinitésimaux, d'où dérive le nom de calcul différentiel) et la

notation moderne dx pour ceux-ci. Il y introduit aussi les règles de manipulation de ces *differentia* pour les puissances et les racines. On y trouve, entre autres, la relation

$$dx^a = a x^{a-1} dx$$

Au-delà de ces règles de calcul, Leibniz indique la marche à suivre pour dériver l'équation différentielle associée à une équation donnée : « *Il en résulte qu'on peut écrire l'équation différentielle de toute équation donnée, en remplaçant simplement chaque membre (c'est-à-dire chaque partie qui pour former l'équation est seulement ajoutée ou retranchée) par sa quantité différentielle. Pour chacune des autres quantités (qui ne sont pas elles-mêmes des membres mais qui contribuent à former l'un d'entre eux), on fait intervenir sa quantité différentielle pour obtenir la quantité*

¹¹⁰ $G=6,67423.10^{-11} \text{ kg}^{-1} \cdot \text{m}^3 \cdot \text{s}^{-2}$

¹¹¹ La masse de la Terre est $5,9736.10^{24} \text{ kg}$

différentielle du membre lui-même, non pas par une simple substitution, mais en suivant l'algorithme que j'ai donné ci-dessus. »

Cet article sera suivi d'un second en 1686, dans lequel Leibniz introduit le symbole mathématique moderne pour désigner l'intégrale, et dans lequel il définit l'une par rapport à l'autre la dérivation et l'intégration (mais notons que le terme intégral ne sera introduit qu'en mai 1690 par Jacques Bernoulli).

Les articles de Leibniz sont d'une lecture peu commode pour ses contemporains, et la diffusion de ses travaux sera lente. Le premier à les reprendre est Jacques Bernoulli, en 1690, suivi de son frère Jean. Les deux frères vont entreprendre un développement systématique de cette théorie, qui sera poursuivi entre autres par Leonhard Euler (1707-1783) et Joseph-Louis Lagrange (1736-1813). La théorie leibnizienne arrive à Paris durant l'hiver 1691-1692, à l'occasion de cours particuliers payants donnés par Jean Bernoulli au marquis Guillaume de l'Hospital (1661-1704), qui publiera sur cette base en 1696 son fameux manuel intitulé *Analyse des infiniment petits pour l'intelligence des lignes courbes*. Mais la diffusion vers un public large commence dès 1693 lorsque l'Hospital entre à l'Académie Royale des Sciences de Paris. Il y rencontrera Pierre Varignon, qui sera celui qui opérera la synthèse de la mécanique newtonienne et du calcul leibnizien.



Figure 104: G. de l'Hospital

12.5.2 Varignon : l'algorithme de la cinématique

Pierre Varignon (1654-1722) est celui qui va introduire et systématiser l'emploi de la théorie naissante du calcul intégral et différentiel pour l'étude du mouvement, donnant naissance à ce qui est appelé aujourd'hui l'algorithme de la cinématique.



Figure 105: P. Varignon

Le travail de Varignon est réalisé en deux temps. Il introduit tout d'abord le concept de « *vitesse dans chaque instant* » (mémoire lu à la séance de l'Académie Royale des Sciences le samedi 5 juillet 1698), puis celui de « *force accélératrice dans chaque instant* » (mémoire lu à la séance de l'Académie Royale des Sciences le samedi 6 septembre 1698). Pour cela, il utilise les *differentia* de Leibniz :

« *Cela posé, les instants seront = dz ; l'espace parcouru dans chaque instant, sera = dx ; et la vitesse avec laquelle dx aura été parcourue sera = y.*

[...] *De sorte que cette vitesse (y), dans chaque instant pouvant être regardée comme uniforme, à cause que $y \pm dy = y$, la notion seule des vitesses uniformes donnera $y = dx/dz$ pour la règle de tous les mouvements variés comme on voudra, c'est-à-dire quelque rapport d'espace, de temps ou de vitesse, qu'on suppose ; la vitesse de chaque instant étant toujours et partout égale au quotient de l'espace parcouru dans chaque instant divisé par cette même différentielle de temps. » (Mémoire du 5 juillet 1698)*

Sur cette base, Varignon propose, toujours dans son premier mémoire, la « Règle générale » suivante :

Règles de Varignon (5 juillet 1698)

« Règle générale

<i>Des vitesses</i>	<i>ou</i>	<i>Des temps</i>	<i>ou</i>	<i>Des espaces</i>	<i>»</i>
$y = \frac{dx}{dz}$		$dz = \frac{dx}{y}$		$dx = ydz$	

Ces règles correspondent aux définitions modernes. Un point important est que Varignon considère ces règles comme pouvant s'appliquer à tout mouvement. C'est la force de cette analyse basée sur le calcul différentiel : tout mouvement peut être analysé **localement**, indépendamment de sa complexité globale.

Après avoir posé les définitions et les règles de calcul relatives à la vitesse, Varignon s'attaque au problème des forces centrales, et est conduit à définir l'accélération (mémoire de l'Académie Royale des Sciences, 30 janvier 1700) :

« De là on aura dx pour l'espace parcouru comme d'une vitesse uniforme v , à chaque instant ; dv pour l'accroissement de vitesse qui s'y fait ; ddx pour ce qui se parcourt d'espace en vertu de cet accroissement de vitesse ; et dt pour cet instant.

A ce compte, la vitesse ne consistant que dans un rapport d'espace parcouru d'un mouvement uniforme, au temps employé à le parcourir, l'on aura déjà $v = dx/dt$ pour une première Règle, laquelle donnera $dv = ddx/dt$ en faisant dt constante. »

Varignon dispose donc de deux relations différentielles : la première qui définit la vitesse, et la seconde qui mesure l'accroissement de vitesse. Il synthétise les deux dans le même mémoire :

« De plus les espaces parcourus par un corps mû d'une force constante et continuellement appliquée, telle qu'on conçoit d'ordinaire la pesanteur, étant en raison composée de cette force et des carrés des temps employés à les parcourir ; l'on aura aussi $ddx = y dt^2$ ou $y = ddx/dt^2 = dv/dt$. Ce qui fait encore une règle $y = dv/dt$, qui avec la précédente $v = dx/dt$, satisfait à tout ce qu'on se propose ici de résoudre. »

On retrouve ici les définitions différentielles modernes de la vitesse et de l'accélération. Elles donnent naissance à la cinématique moderne : tout problème peut maintenant être exprimé comme un problème différentiel, dont la résolution passera par une intégration. Cet algorithme de la cinématique est extrêmement puissant. Il permettra notamment à Varignon de s'attaquer avec beaucoup plus de succès que ses illustres prédécesseurs aux problèmes des forces centrales : le problème direct (trouver la trajectoire d'un mobile soumis à une force centrale) et le problème inverse (connaissant la trajectoire, déterminer la force centrale). Un second champ d'investigation pour Varignon sera la théorie du mouvement des projectiles. La nature parabolique du mouvement balistique dans le cas où la résistance de l'air est nulle est connue depuis Galilée. Varignon va généraliser en 1707 l'analyse en considérant

différentes lois pour la résistance de l'air : résistance proportionnelle à la vitesse ou à une puissance de la vitesse. Il considère des équations de la forme

$$\frac{d^2x}{dt^2} = \frac{du}{dt} = -kF(u) + f(t)$$

où $F(u)$ est la résistance de l'air (fonction de la vitesse u), k un coefficient balistique et $f(t)$ l'accélération du mouvement primitif.

La théorie balistique sera également développée par Jean Bernoulli (1657-1748), qui introduit les coordonnées intrinsèques pour décrire le mouvement.

12.6 Réorganisations de la mécanique



Figure 106: J.L. Lagrange

Les travaux de Varignon, en permettant l'emploi du calcul différentiel et intégral pour décrire le mouvement, marquent un tournant décisif : ils ouvrent la voie à de nombreux développements impossibles en dehors du cadre de ces outils mathématiques. Ainsi, de nombreuses réorganisations de la mécanique verront le jour au XVIII^e siècle, qui seront couronnées par la mécanique analytique de Joseph-Louis Lagrange (1736-1813), qui publie en 1788 son livre *Mécanique Analytique*. L'œuvre de Lagrange forme l'aboutissement des travaux antérieurs, et représente la forme la plus achevée de la mécanique analytique. Comme le note Ernst Mach, Lagrange « *s'applique à faire, une fois pour toutes, toutes les démonstrations nécessaires et à condenser le plus possible de choses dans une seule formule.* » Les travaux de Lagrange sont mathématiquement complexes, et ne seront pas présentés dans le cadre du cours.

Nous allons maintenant discuter de la genèse et du développement de quelques principes fondamentaux de la mécanique classique.

12.6.1 Le principe de moindre action



Figure 107: P. Fermat

Le principe de moindre action est un principe de la physique classique. Il trouve son origine dans la controverse entre René Descartes et Pierre Fermat (1601-1665) à propos des lois de réfraction de la lumière.

Fermat utilise la méthode d'*adégalisation*, qu'il a développée vers 1629 pour la recherche des extrema. Pour une fonction f dépendant d'une variable x , et considérant une quantité très petite y (mais attention, le calcul infinitésimal n'existe pas encore !), Fermat repère les positions des extrema comme les valeurs de x pour lesquelles $f(x+y) \approx f(x)$, puis en posant finalement $y=0$. On reconnaît ici les prémisses de la définition de la dérivée

d'une fonction par passage à la limite à partir d'une formule d'accroissement fini. Mais la notion de fonction n'est pas formalisée chez Fermat.

Au moyen de cet outil, Fermat affirme en 1657 que la lumière suit toujours une trajectoire qui minimise le temps de déplacement, et faisant l'hypothèse que la vitesse de la lumière est plus faible dans les milieux denses que dans les milieux moins denses, retrouve la loi de réfraction de Descartes de 1637. Si le résultat est le même, il conteste en revanche l'analyse de Descartes, qui suppose que la vitesse de la lumière est une fonction croissante de la densité du milieu et qui fait appel à une analogie avec le mouvement d'une balle. La controverse durera longtemps, puisqu'en 1682 Leibniz prend position en proposant l'hypothèse selon laquelle la lumière suit le chemin le plus aisé (mais sans reprendre la notion de Fermat de temps ou de trajet le plus court).



Figure 108 : P.L. Maupertuis

Le principe de minimisation de Fermat est repris en 1697 par Jean Bernoulli pour étudier la trajectoire qu'un point matériel pesant doit suivre pour aller d'un point à un autre sans vitesse initiale (problème de la courbe brachystochrome).

Le principe de moindre action sera finalement énoncé par Pierre-Louis Moreau de Maupertuis (1698-1759) dans son mémoire à l'Académie Royale des Sciences du 15 avril 1744 :

Principe de moindre action de Maupertuis (1744)

« En méditant profondément sur cette matière, j'ai pensé que la lumière, lorsqu'elle passe d'un milieu dans un autre, abandonnant déjà le chemin le plus court, qui est celui de la ligne droite, pouvait bien aussi ne pas suivre celui du temps le plus prompt. En effet, quelle préférence devrait-il y avoir ici du temps sur l'espace ? La lumière ne pouvant plus aller tout à la fois par le chemin le plus court, et par celui du temps le plus prompt, pourquoi irait-elle plutôt par l'un de ces chemins que par l'autre ? Aussi ne suit-elle aucun des deux ; elle prend une route qui a un avantage plus réel : le chemin qu'elle tient est celui par lequel la quantité d'action est la moindre.

Il faut maintenant expliquer ce que j'entends par la quantité d'action. Lorsqu'un corps est porté d'un point à un autre, il faut pour cela une certaine action : cette action dépend de la vitesse qu'a le corps, et de l'espace qu'il parcourt ; mais elle n'est ni la vitesse ni l'espace pris séparément. La quantité d'action est d'autant plus grande, que la vitesse du corps est grande, et que le chemin qu'il parcourt est plus long ; elle est proportionnelle à la somme des espaces multipliés chacun par la vitesse avec laquelle le corps les parcourt. »

Ce même principe sera exprimé par Lagrange en 1788 sous la forme suivante :

Principe de moindre action de Lagrange (1788)

« De là résulte donc ce théorème général, que, dans le mouvement d'un système quelconque de corps animés par des forces mutuelles d'attraction, ou tendantes à des autres fixes, et proportionnelles à des fonctions quelconques des distances, les courbes décrites par les différents corps, et leurs vitesses, sont nécessairement telles, que la somme des produits de chaque masse $[m]$ par l'intégrale de la vitesse $[u]$ multipliée par l'élément de la courbe $[ds]$ est un maximum ou un minimum $[m \int u ds]$ pourvu que l'on regarde les premiers et les derniers points de chaque courbe comme donnés, en sorte que les variations des coordonnées répondant à ces points soient nulles. »

12.6.2 Principe d'inertie et relativité du mouvement : des prémisses au principe de Mach

Nous allons maintenant étudier un des problèmes centraux de la mécanique : la définition du principe d'inertie et son lien avec la notion de relativité du mouvement. Le principe d'inertie est un des piliers de la mécanique, dont le manque a été ressenti pendant des siècles pour établir le système héliocentrique. En effet, sans principe d'inertie, il n'est pas possible d'accepter le mouvement diurne de la Terre.

Comme on l'a vu, l'idée d'un tel principe est présente chez Galilée, puisqu'il évoque la possibilité du mouvement sans action motrice continue. Pour illustrer cette idée, Galilée fait appel à des expériences virtuelles (il ne les a probablement jamais réalisées) aujourd'hui fameuses :

Galilée, Dialogue (1632)

« **SAGREDO** : S'il est vrai que l'impulsion par laquelle le navire avance reste imprimée de manière indélébile dans la pierre après qu'elle se soit séparée du mât, et s'il est également vrai que ce mouvement n'empêche ni ne ralentit le mouvement vers le bas qui est naturel à la pierre, il s'ensuit nécessairement un effet de nature étonnante. Admettons que le navire soit immobile et que le temps de la chute d'une pierre depuis le haut du mât soit équivalent à deux battements du pouls. Si l'on fait maintenant avancer le navire et qu'on laisse tomber la même pierre du même endroit, cette pierre, comme nous l'avons exposé, mettra également le temps de deux battements du pouls pour arriver en bas, et, pendant ce temps, le navire aura parcouru, par exemple, vingt brasses ; de sorte que le mouvement de la pierre aura été une ligne transversale bien plus longue que la première qui était droite et perpendiculaire, de même dimension que le mât. Cependant, la pierre l'aura franchie dans le même temps.

Et si maintenant on augmente encore la vitesse du navire, la pierre en tombant décrira une transversale encore plus longue, mais elle la franchira toujours dans le même temps de deux battements du pouls.

Supposons de même qu'on ait disposé une couleuvrine au sommet d'une tour et que l'on tire avec elle des tirs « pointe en blanc », c'est-à-dire parallèlement à l'horizon ; quelle que soit la charge [de poudre] de la couleuvrine, et donc que le boulet tombe à mille, quatre mille, six mille ou dix mille brasses, tous ces tirs se feront en des temps égaux entre eux, chacun étant égal au temps qu'il faudrait au boulet pour aller de la gueule de la couleuvrine jusqu'à terre lorsqu'on le laisse simplement tomber à la perpendiculaire, sans aucune impulsion. Et c'est une chose bien étonnante que, dans le même laps de temps très court qui est nécessaire à la chute verticale jusqu'à terre, le même boulet, expulsé par le feu, puisse franchir des distances de quatre cents, mille, quatre mille ou dix mille brasses et que ce boulet, dans tous les tirs de « pointe en blanc », reste toujours le même temps en l'air. »

Mais notons qu'il s'agit ici d'une expérience par la pensée : Galilée n'a jamais réalisé l'expérience consistant à lâcher un corps pesant du haut du mât d'un navire. Cette expérience sera réalisée à bord d'une galère en 1641 par le savant et philosophe Pierre Gassendi (1592-1655).

Galilée, Dialogue (1632)

« SALVIATI : Enfermez-vous avec un ami dans la plus vaste cabine d'un grand navire, et faites en sorte que s'y trouvent également des mouches, des papillons et d'autres petits animaux volants, qu'y soit déposé un grand récipient rempli d'eau dans lequel on aura mis des petits poissons ; suspendez également à bonne hauteur un petit sceau et disposez-le de manière à ce que l'eau se déverse goutte à goutte dans un autre récipient à col étroit que vous aurez disposé en dessous ; puis, alors que le navire est à l'arrêt, observez attentivement comment ces petits animaux volent avec des vitesses égales quel que soit l'endroit de la cabine vers lequel ils se dirigent ; vous pourrez voir les poissons nager indifféremment dans toutes les directions ; les gouttes d'eau tomberont toutes dans le récipient posé par terre ; si vous lancez un objet à votre ami, vous ne devrez pas fournir un effort plus important selon que vous le lancerez dans telle ou telle direction, à condition que les distances soient égales ; et si vous sautez à pieds joints, comme on dit, vous franchirez des espaces semblables dans toutes les directions.

Une fois que vous aurez observé attentivement tout cela – il ne fait aucun doute que si le navire est à l'arrêt les choses doivent se passer ainsi – faites se déplacer le navire à une vitesse aussi grande que vous voudrez ; pourvu que le mouvement soit uniforme et ne fluctue pas de-ci de-là, vous n'apercevrez aucun changement dans les effets nommés, et aucun d'entre eux ne vous permettra de savoir si le navire avance ou s'il est arrêté : si vous sautez, vous franchirez sur le plancher les mêmes distances qu'auparavant et, si le navire se déplace, vous n'en ferez pas pour autant des sauts plus grands vers la poupe que vers la proue, bien que, pendant que vous êtes en l'air, le plancher qui est en dessous ait glissé dans la direction opposée à celle de votre saut ; si vous jetez quelque objet à votre ami, il ne vous faudra pas le lancer avec plus de force pour qu'il lui parvienne, que votre ami se trouve vers la poupe et vous vers la proue, ou que ce soit le contraire ; les gouttes d'eau tomberont comme auparavant dans le récipient qu'on aura mis en dessous, sans qu'une seule goutte ne tombe du

côté de la poupe, bien que, pendant le temps où la goutte est en l'air, le navire ait parcouru plus d'un empan ; les poissons dans leur eau nageront sans plus d'effort vers l'une ou l'autre partie du récipient dans lequel on les aura mis et ils se dirigeront avec autant d'aisance vers la nourriture quel que soit l'endroit du bord du bocal où elle aura été placée ; enfin, les papillons et les mouches continueront à voler indifféremment dans toutes les directions. »

Ces deux exemples illustrent non seulement la composition des vitesses, mais également l'idée que les lois de la physique sont les mêmes que le navire soit au repos ou en mouvement rectiligne uniforme.

Mais le principe d'inertie lui-même n'est pas énoncé par Galilée : ce qu'Albert Einstein nomme Principe d'Inertie de Galilée :

Principe d'Inertie de Galilée selon Einstein

« Un point matériel livré à lui-même et suffisamment éloigné de tous les autres points effectue un mouvement rectiligne uniforme »

ne peut pas vraiment être attribué à Galilée qui écrit :

Principe d'inertie de Galilée

« Un degré de vitesse quelconque, une fois communiqué à un mobile, s'imprime en lui de façon indélébile du seul fait de sa nature, et pourvu que soit supprimées les causes extérieures d'accélération et de ralentissement, ce qui n'a lieu que sur un plan horizontal ; sur un plan descendant, en effet, il existe déjà une cause d'accélération, et sur un plan ascendant une cause de ralentissement ; d'où il suit que le mouvement sur un plan horizontal est éternel ; car s'il est uniforme, il ne s'affaiblit ni ne diminue, et encore moins ne cesse. »

La paternité historique est en réalité partagée par le physicien hollandais Isaac Beeckman (1588-1637) qui écrit en 1613 :

Principe d'inertie de Beeckman (1613)

« Ce qui est une fois mis mouvement, demeure en mouvement éternellement. »

et par son ami René Descartes (1596-1650) qui exprime le principe de conservation de la quantité de mouvement dans son *Traité du monde et de la lumière* écrit entre 1629 et 1633 :

Principe d'inertie de Descartes (1629-1633)

« La première est : que chaque partie de la matière, en particulier, continue toujours d'être en même état, pendant que la rencontre des autres ne la contraint point de changer. C'est-à-dire : si elle a quelque grosseur, elle ne deviendra jamais plus petite, sinon que les autres la divisent ; si elle est ronde ou carrée, elle ne changera jamais cette figure sans que les autres l'y contraignent ; si elle est arrêtée en quelque lieu, elle n'en partira jamais que les autres ne l'en chassent ; et si elle a une fois commencé à se mouvoir, elle continuera toujours avec une égale force jusqu'à ce que les autres l'arrêtent ou la retardent. »

On l'a vu, il est ensuite très clairement exprimé par Huygens en 1673 :

Principe d'inertie de Huygens (1673)

Proposition I : Dans des temps égaux les accroissements de la vitesse d'un corps tombant sont toujours égaux et les espaces parcourus durant des temps égaux depuis le commencement de la chute forment une série dont les différences successives sont constantes.

et il faut attendre Newton pour le voir mis sous sa forme définitive (première loi du mouvement dans les Principia).

Principe d'inertie de Newton (1687)

« Première loi : Tout corps persévère dans son état de repos ou de mouvement rectiligne uniforme à moins que des forces imprimées ne le contraignent à changer son état. »

Outre le principe d'inertie, Galilée a également anticipé le principe de la relativité du mouvement. Il prendra plusieurs exemples, à partir desquels il conceptualise le fait que les lois du mouvement ne sont pas modifiées dans ce cas.

Plus tard, Leibniz énonce un principe d'« équivalence des hypothèses » : dans tout système formé de corps en interaction, l'hypothèse que l'un des corps est au repos est équivalente à celles qui disent que c'est l'un des autres corps qui est au repos et les autres en mouvement. Cette hypothèse a des implications cosmologiques fortes, puisqu'elle remet en cause les raisonnements de Ptolémée et de Copernic. Pour Leibniz, ce principe de relativité est un principe philosophique qui découle de l'idée que l'espace n'est rien d'autre qu'une abstraction conceptuelle des relations géométriques qui existent entre les corps.

Ici aussi, il faut attendre Newton pour avoir une expression satisfaisante du principe de la relativité (ce qui fait que ce principe doit être appelé principe de la relativité classique plutôt que principe de la relativité galiléenne) :

Principe de relativité de Newton

« Les mouvements relatifs des corps enfermés dans un espace quelconque sont les mêmes que cet espace soit immobile ou qu'il se meuve uniformément le long d'une ligne droite, sans rotation »

La forme moderne de ce principe est exprimée comme suit :

Principe de la relativité classique, ou galiléenne

« Les lois de la physique sont les mêmes dans tous les référentiels en mouvement rectiligne uniforme les uns par rapport aux autres. »

De tels référentiels sont appelés référentiels galiléens ou encore référentiels inertiels. Ce concept ne sera théorisé qu'en 1885 par Ludwig Lange. Les transformations qui permettent de passer d'un référentiel inertiel à un autre sont les transformations galiléennes, et ont pour conséquence la loi de composition des vitesses.

Mais revenons à la mécanique de Newton. Ses trois lois sont parfaitement compatibles avec le principe de relativité, et Newton donne même la première forme « moderne » du principe de relativité classique. Notamment, les forces telles qu'il les introduit dans sa seconde loi du mouvement ne sont mesurées et définies que par l'accélération résultante d'un corps, et sont donc indépendantes de la vitesse du système dans lequel la mesure est effectuée. Mais Newton introduit également dans les Principia la notion d'espace et de temps absolus, et donc d'un référentiel absolu dans lequel tout corps pourrait être considéré comme vraiment en mouvement ou vraiment au repos, ce qui est a priori incompatible avec le principe de relativité qui implique l'égalité de tous les repères inertiels.

Cette question de l'existence d'un repère absolu dans lequel on pourrait distinguer tous les corps en mouvement rectiligne uniforme (et donc les repères inertiels) a été longuement débattue au XIXe siècle. En effet, la théorie de Newton ne répond pas à une question fondamentale : dans quel repère peut-on dire qu'un corps est au repos ou en mouvement rectiligne uniforme, c'est-à-dire dans quel référentiel la résultante des forces est-elle nulle ? En 1870, le mathématicien Carl Neumann propose que la loi d'inertie fasse implicitement référence à un corps situé quelque part dans l'univers, le **corps Alpha**, par rapport à qui le mouvement de toute particule entièrement libre est rectiligne, et auquel il est possible d'associer une échelle de temps telle que le mouvement soit uniforme. Un peu plus tard, en 1885, Ernst Mach pose que les lois de Newton font implicitement appel aux étoiles fixes comme référentiel d'espace et à la période de rotation de la Terre comme échelle de temps. Pour Mach, qui est un phénoméniste, la notion de repère absolu est une abstraction qui résulte de la pratique de la mesure de coordonnées par rapport aux étoiles fixes. Un point commun à ces réponses est qu'elles font toutes références à des objets matériels pour définir l'origine du référentiel. On retrouve ici la trace d'une idée très ancienne, présente par exemple chez Aristote et Leibniz, qui est que le vide, n'étant « rien », ne peut servir de référence.

La définition de Neumann a inspiré celle de Ludwig Lange, qui définit les systèmes inertiels comme suit :

Définition du repère inertiel de Lange (1885)

« Un système inertiel est un système de coordonnées par rapport auquel trois particules libres, issues d'une même origine et se déplaçant suivant des trajectoires non-coplanaires, se déplacent en ligne droite et parcourent des distances mutuellement proportionnelles. La loi d'inertie dit alors que toute quatrième particule, si elle est libre, se déplacera uniformément par rapport à tout référentiel inertiel. »

Une autre définition sera donnée à la même époque (1884) mais de manière indépendante par le physicien anglais James Thomson, qui définit

Définition du repère inertiel de Thomson (1884)

un repère inertiel comme un repère dans lequel la seconde loi de Newton est vérifiée, de telle manière que toute accélération correspond à l'imposition d'une force.



Figure 109: J. Thomson

Comme la définition de Lange, celle de Thomson implique que toutes les forces doivent appartenir à une paire action-réaction (donc correspondre à une interaction matérielle réelle), sans quoi les repères en rotation entreraient dans la catégorie des repères inertiels. Ce dernier point a été ajouté par R. Muirhead en 1887. La définition de Thomson possède deux avantages par rapport à celle de Lange. Tout d'abord, en faisant appel à la seconde loi de Newton plutôt qu'à la première, elle montre que l'on peut définir des repères inertiels sans savoir si des particules libres existent réellement dans l'univers. Ensuite, elle éclaire le lien fort qui existe entre les lois newtoniennes du mouvement et les

repères inertiels : ces lois assument l'existence d'au moins un repère inertiel dans l'univers. Une conséquence est que la question « par rapport à quel référentiel les lois du mouvement sont-elles valides ? » est mal posée, car les lois du mouvement définissent une classe de référentiels, ainsi qu'une procédure pour les construire.

Venons-en maintenant à ce qu'Einstein a appelé en 1918 le **principe de Mach**, en référence aux développements de Mach de 1893 concernant la relativité en mécanique, et plus précisément la relativité de la notion d'inertie. Tout d'abord, Mach s'oppose au concept de repère absolu de Newton :

« Si, dans un système spatial matériel, il y a des masses de vitesses différentes qui peuvent entrer en interactions mutuelles, ces masses sont sujettes à des forces à nos yeux. Nous ne pouvons décider combien ces forces sont grandes que lorsque nous connaissons les vitesses que vont atteindre les masses. Les masses au repos sont également des forces si toutes les masses ne sont pas au repos ... Toutes les masses et toutes les vitesses, et par conséquent toutes les forces, sont relatives. Il n'y a pas de décision concernant leur caractère relatif ou absolu que nous puissions prendre, ou que nous soyons forcé de prendre, ou dont nous pourrions tirer un quelconque avantage, intellectuel ou autre. »

Dans la mécanique de Newton, le produit de la masse inertielle d'un corps par son accélération est égal à la résultante des forces exercées sur lui. Pour une force donnée, l'accélération est donc inversement proportionnelle à la masse inertielle, et cette dernière mesure donc l'inertie du corps définie comme sa capacité à résister à une force.

Cette analyse simple conduit à un résultat complexe lorsque l'on considère l'expérience virtuelle suivante. Supposons que toute la matière de l'Univers disparaisse, à l'exception d'un corps, qui est donc en conséquence vraiment isolé. Aucune force ne s'exerce sur lui, le produit de sa masse par son accélération doit être nul. Comme toute matière dans l'Univers a disparu, il n'existe plus de système matériel par rapport auquel il serait possible de définir la vitesse ou l'accélération du

corps. C'est donc la masse qui, mathématiquement, doit être prise comme étant égale à zéro. En effet, Mach définit la masse au moyen de l'accélération et du principe de l'action et de la réaction (troisième loi de Newton) :

« Les corps de masses égales sont les corps qui, au cours d'interactions (gravitationnelles) mutuelles, produisent les uns sur les autres des accélérations de sens opposés et de même amplitude. »

La conclusion de Mach est la suivante :

« Le comportement des objets terrestres par rapport à la Terre est réductible au comportement de la Terre par rapport aux objets célestes lointains. Si nous disions que nous en savons plus sur les corps en mouvement que leur comportement observé par rapport aux corps célestes, nous nous rendrions coupables de falsification. En conséquence, quand nous disons qu'un corps observe un mouvement rectiligne uniforme, notre énoncé n'est rien de plus qu'une référence abrégée à l'ensemble de l'Univers.

[...] Les considérations précédentes montrent qu'il n'est pas nécessaire de référer la loi d'inertie à un repère absolu. Au contraire, on voit que les masses qui, dans le langage commun exercent des forces les unes sur les autres aussi bien que celles qui n'en exercent aucune, sont dans le même rapport à l'accélération. Nous pouvons regarder toutes les masses comme liées les unes aux autres. Que l'accélération joue un rôle prédominant dans les relations entre les masses doit être accepté comme un fait provenant de l'expérience ; cela ne doit pas pour autant exclure les tentatives pour élucider ce fait en le comparant à d'autres, incluant la découverte de nouveaux points de vue. »

Ceci peut se résumer comme suit :

Principe de Mach

La masse et l'inertie d'un corps ne sont pas des propriétés propres de ce corps, mais sont déterminées par ses relations avec le reste de l'univers.



Figure 110: J. Foucault

Le principe de Mach pose non seulement la relativité de la masse inertielle et de l'inertie, mais il implique également un **principe de non-séparabilité** : c'est l'interaction avec l'ensemble du reste de l'Univers qui définit les propriétés d'un corps.

Cela est illustré par une expérience simple, comme celle du **pendule de Foucault**. Lors de cette expérience réalisée en 1851, Jean Bernard Léon Foucault (1819-1868) observe que le plan des oscillations du pendule n'est pas fixe, et en déduit que la Terre tourne sur elle-même. Mais le plan d'oscillation est en réalité fixe, et c'est la rotation de la Terre qui donne lieu à une rotation apparente du plan d'oscillation. La question se pose donc : dans quel système de référence le plan d'oscillation du pendule est-il fixe ? Ou, de manière équivalente, autour de quoi la Terre tourne-t-elle ? Dans un référentiel lié au soleil, la rotation apparente existe également,

mais est beaucoup plus faible. Il en va de même pour le centre de notre galaxie, ou celui de notre amas galactique. Le plan d'oscillation du pendule serait fixe en considérant les objets galactiques les plus lointains. Le pendule de Foucault est donc sensible à l'univers tout entier. Mais la nature de ce lien reste à élucider, et notamment celui de la vitesse de transfert des informations liées aux interactions !

Il faut noter que, bien qu'Einstein ait indiqué que les travaux de Mach avaient eu une influence certaine sur lui, le principe de Mach n'est pas compatible avec la théorie de la relativité générale. La mécanique newtonienne et la relativité restreinte posent toutes les deux le principe de la relativité du mouvement tout en conservant une notion d'accélération absolue (l'accélération est la même pour tous les référentiels galiléens liés). L'étape suivante, qui mènera au développement de la théorie de la relativité générale, est de considérer des accélérations (donc des forces relatives), conduisant à une généralisation de la seconde loi de Newton qui est invariante par changement de repère non-Galiléen. Après le principe de relativité du mouvement, on introduit donc celui de la relativité des accélérations.

12.6.3 La relativité du mouvement et le vide quantique

Comme nous l'avons vu plus haut, dans le cadre de la mécanique classique, si le mouvement dans le vide est considéré comme possible (le vide est supposé n'offrir aucune résistance au mouvement et permettre une propagation parfaite des ondes électromagnétiques), le mouvement n'est en revanche pas définissable *par rapport* au vide. Le référentiel est toujours associé à un corps matériel.

Les résultats de la mécanique quantique concernant le vide bouleversent complètement la question de la relativité du mouvement et permettent d'utiliser le vide comme référentiel absolu pour définir les mouvements inertiels.

L'aspect révolutionnaire de la description du vide par la mécanique quantique est que le vide n'est plus « rien », et est le lieu d'interactions complexes qui lui donnent une énergie non-nulle. Analysons tout d'abord ce qu'implique l'existence d'un vide qui n'oppose aucune résistance au mouvement. Tout d'abord, aucune matière ne doit être présente, sans quoi des collisions auraient lieu et le mouvement du corps se déplaçant dans le vide serait modifié. Mais, de plus, aucun rayonnement électromagnétique ne doit être présent, faute de quoi le corps en mouvement subirait une pression électromagnétique¹¹² qui le freinerait ou l'accélérerait. Pour obtenir un vide dynamiquement « inerte », il faut donc imaginer une enceinte délimitant un volume ne contenant aucune matière et aucun rayonnement. Or on sait qu'à toute température non-nulle, l'espace est rempli par un rayonnement électromagnétique d'origine thermique (c'est le *rayonnement du corps noir*). Il faudrait donc tendre vers le zéro absolu pour voir s'annuler la pression électromagnétique. Mais les résultats de Max Planck (1912) puis de Nernst (1916) montrent que, même à température nulle, des fluctuations (dites « fluctuations de point zéro ») persistent : le vide est traversé en permanence par des fluctuations qui se propagent à la vitesse de la lumière. Elles sont associées à des particules fugaces (appelées « particules virtuelles ») dont le temps de

¹¹² Cette pression est due à l'impact des photons associés au rayonnement électromagnétique sur la surface du corps en mouvement. Elle est donc comparable à celle exercée par les molécules d'un fluide sur la surface d'un corps immergé.

vie est très court Le vide possède donc une énergie non nulle ! Le vide quantique est donc l'état qui minimise l'énergie, sans l'annuler.

Cette nouvelle conception du vide a de nombreuses conséquences. Voyons celles qui concernent le mouvement d'un corps matériel dans le vide¹¹³. L'existence des fluctuations de point zéro génère automatiquement l'existence d'une pression sur ce corps. Un effet spectaculaire, observé expérimentalement, est l'effet Casimir.

Cet effet Casimir, prédit par le physicien hollandais Hendrick Casimir en 1948, n'a été observé expérimentalement qu'en 1958 par M.J. Spaarnay. Il consiste en l'apparition d'une force qui tend à rapprocher deux corps lorsque ceux-ci sont suffisamment proches. En rapprochant deux plaques métalliques dans le vide, lorsque la distance séparatrice est très faible, on rompt la symétrie entre les faces externes et internes des plaques : la pression exercée par les particules virtuelles est plus forte sur les faces externes, conduisant à la création d'une force (la *force de Casimir*) qui tend à rapprocher les deux plaques¹¹⁴. On peut donc générer du mouvement à partir du vide !

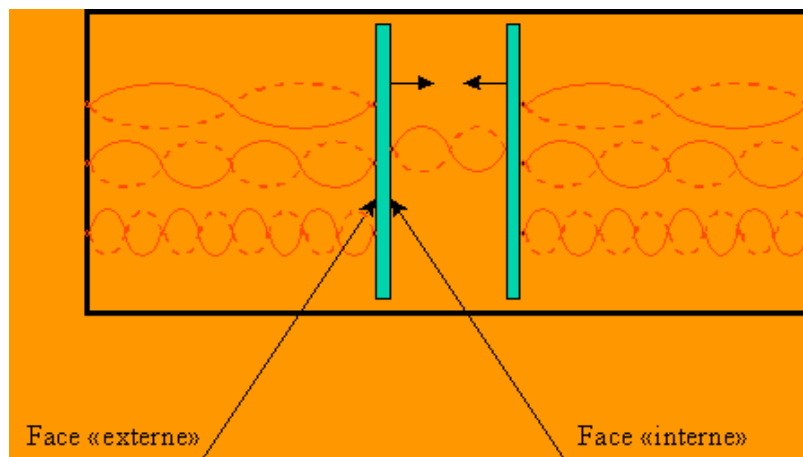


Figure 111: illustration de l'effet Casimir

La question de la définition du mouvement absolu est directement liée à la généralisation de la force de Casimir sur les corps qui se déplacent dans le vide au zéro absolu, donc en interaction avec les seules fluctuations de point zéro. La théorie quantique prédit¹¹⁵ que ces fluctuations exercent une force nulle sur tous les corps en mouvement uniforme ou uniformément accéléré, et une force dissipative pour tous les autres mouvements. Le vide exerce donc une résistance qui tend à ramener tous les corps vers un mouvement uniformément accéléré. Une conclusion très intéressante est donc que le vide, et plus précisément les fluctuations de point zéro, permettent naturellement de distinguer les mouvements uniformément accélérés. Ils forment donc

¹¹³ Un problème connexe est celui de la théorie quantique de la gravitation, car les fluctuations de point zéro sont sensibles à l'existence d'un champ gravitationnel et génèrent elles-mêmes des fluctuations du champ de gravité.

¹¹⁴ L'intensité de la force d'attraction engendrée est de l'ordre de $0,1 \mu\text{N}$ pour une surface des plaques égale à 1 cm^2 et une distance de séparation égale à $1 \mu\text{m}$.

¹¹⁵ Alors que l'effet Casimir a déjà été mis en évidence expérimentalement sur deux miroirs, la force « généralisée » ne l'a toujours pas été, du fait de sa petitesse et des problèmes technologiques posés.

un référentiel qui permet de distinguer les mouvements non-uniformément accélérés. Il ne s'agit malheureusement pas là d'un référentiel absolu, tel que défini par Newton, car ils ne permettent pas d'isoler les seuls mouvements uniformes (les référentiels Galiléens). Mais une partie du chemin a été parcourue !