

INTRO UNE INTRODUCTION À DUCTION TION

LE TEMPS : MESURABLE, RÉVERSIBLE, INSAISSABLE ?

Mathias Fink,
Michel Le Bellac
et Michèle Leduc

Hommage à Roger Maynard

edp sciences

Le temps : mesurable, réversible, insaisissable ?

**Mathias Fink
Michel Le Bellac
Michèle Leduc**

Hommage à Roger Maynard



17, avenue du Hoggar
Parc d'activités de Courtabœuf, BP 112
91944 Les Ulis Cedex A, France

Table des matières

Avant-Propos	vii
1 Le temps newtonien	1
1.1 Le temps subjectif	1
1.2 L'espace-temps newtonien	5
1.3 Causalité et simultanéité	7
2 Les différentes échelles de temps	11
2.1 Du temps de Planck à la milliseconde	11
2.2 De la milliseconde à mille années	15
2.3 De mille années à 10 milliards d'années	16
3 La mesure du temps	19
3.1 Repérer le temps : depuis la préhistoire	19
3.2 Mesurer les durées avec un phénomène continu : premières horloges	22
3.3 Mesurer le temps avec un phénomène périodique	24
3.4 Les horloges à atomes froids	36
3.5 Exactitude, stabilité, performances d'une horloge	43
3.6 Définition de l'unité du temps : la seconde	46
4 Le temps de la relativité restreinte	51
4.1 Vitesse limite	51
4.2 Temps propre	53
4.3 Espace-temps de Minkowski	58
4.4 Synchronisation des horloges et transformation du temps	61
4.5 Le paradoxe des jumeaux	67

5	Le temps de la relativité générale	73
5.1	Décalage vers le rouge gravitationnel	73
5.2	Application au GPS	77
5.3	Le temps en gravité forte	82
5.4	Le temps cosmologique	86
6	Le standard de temps universel	91
6.1	Connaître le temps sur la Terre	91
6.2	Les transferts temps/fréquence (T/F)	93
6.3	Une horloge sur la plateforme spatiale internationale (ISS)	103
6.4	Le transfert temps/fréquence par fibre optique	109
6.5	Le temps universel	113
6.6	Quelques applications fondamentales de la mesure du temps ...	113
6.7	Conclusion	116
7	La flèche du temps	117
7.1	Le second principe et l'entropie	119
7.2	L'entropie de Boltzmann	125
7.3	Énergie et entropie	131
7.4	Réversibilité microscopique et irréversibilité macroscopique ...	133
7.5	Entropie et information	141
8	Renversement du temps et ondes	145
8.1	Acoustique et renversement du temps	147
8.2	Principe d'Huygens et miroir à retournement temporel (MRT) ..	149
8.3	Causalité et limite de diffraction	151
8.4	Complexité, codage spatio-temporel et degrés de liberté	153
8.5	Communication sous-marine et objet tactiles	159
8.6	Super-résolution et résonateurs sub-longueur d'onde	164
8.7	Imagerie ultrarapide et retournement temporel	169
8.8	Thérapie	171
8.9	Sonar et radar à retournement temporel	172
8.10	Conclusion	174
	Bibliographie	175
	Index	177

Avant-propos

Ce livre est issu d'une réflexion sur le temps menée par Roger Maynard à partir de ses travaux de physicien sur les ondes, le désordre et la complexité. En nous inspirant des nombreuses conférences grand public qu'il avait données sur ce sujet, nous avons le projet d'écrire avec lui un livre sur le temps, en additionnant nos compétences sur divers aspects de la question. Malheureusement la disparition prématurée de Roger en juin 2015 nous a conduits à reprendre sans lui ce projet qui lui tenait à cœur et que nous avons écrit en hommage à sa mémoire. Roger aimait commencer ses conférences sur le temps par une citation de Saint-Augustin :

« Si personne ne me le demande, je sais ce qu'est le temps. Si je dois l'expliquer à quelqu'un qui me le demande, je ne le sais plus. »

Depuis l'Antiquité, nombreux sont les philosophes et les scientifiques qui ont réfléchi sur la nature du temps. Le présent ouvrage est écrit par des physiciens et repose sur un ensemble de connaissances théoriques et expérimentales dont nous disposons aujourd'hui. Faute de compétences sur le sujet, les aspects psychologiques de la perception du temps et le temps en biologie ne seront pas traités dans ce livre.

Nous rappelons brièvement quelques aspects historiques de la vision du temps par les grands scientifiques tels que Newton, Boltzmann, Einstein et d'autres. Nous exposons dans un premier temps la théorie du temps absolu conçue par Newton, qui sous-tend encore aujourd'hui notre perception intuitive. Nous abordons ensuite les bouleversements apportés par la relativité restreinte et la relativité générale, que l'on doit prendre en compte pour la navigation dans l'espace, le GPS et la définition d'un temps universel sur Terre. Nous discutons les progrès récents obtenus à l'aide des horloges atomiques les plus avancées, ainsi que les méthodes à base de satellites et de fibres optiques pour transférer le temps et synchroniser les horloges sur l'ensemble de la surface de la Terre.

Un problème fascinait Roger Maynard : celui du sens de l'écoulement du temps. Pourquoi l'évolution spontanée se déroule-t-elle toujours dans le même sens ? Pourquoi vieillissons-nous inexorablement ? La réponse fait intervenir le

concept d'entropie, dont la croissance constante fixe la direction de la flèche du temps. Nous examinons la réponse de nature statistique donnée par Boltzmann, qui doit cependant être complétée par des considérations d'ordre cosmologique.

Enfin un autre problème passionnait Roger Maynard : est-il possible de fabriquer un dispositif permettant de renverser localement la flèche du temps ? Est-il possible, pour un système de particules évoluant de l'ordre vers le désordre, d'inverser sur commande la dynamique de ces particules pour revenir à l'ordre initial ? Cela mène au concept de miroir à retournement temporel, qui est beaucoup plus facile à mettre en œuvre pour les ondes que pour les particules et qui sera développé dans le dernier chapitre.

Ce livre, abondamment illustré, se veut accessible à un public non spécialiste mais curieux des développements de la science contemporaine. Quelques passages plus complexes sont isolés dans des encadrés qui peuvent être omis en première lecture. Nous souhaitons que cet ouvrage apporte au lecteur une vision du temps qui intègre l'essentiel de ce que savent aujourd'hui les physiciens, tout en soulignant les interrogations qui demeurent sur sa nature.

Remerciements. Mathias Fink remercie les membres de l'Institut Langevin qui ont participé à la recherche sur les miroirs à retournement temporel. Michel Le Bellac remercie Olivier Darrigol, Jean-Marc Lévy-Leblond et Kirone Mallick pour leurs critiques constructives et leurs commentaires, ainsi que Agnès Mougeot, Vincent Le Bellac et André Sénik pour les références philosophiques. Michèle Leduc remercie tout particulièrement Noël Dimarcq du laboratoire SYRTE pour ses conseils et pour plusieurs figures illustrant les chapitres sur la mesure et le transfert du temps. Elle remercie également Arnaud Landragin, Sébastien Bize, Christophe Salomon et Saïda Guélati.

Mathias Fink, Michel Le Bellac et Michèle Leduc



Roger Maynard (1938–2015).

1

Le temps newtonien

Qu'est-ce que le temps ? Cette question a préoccupé les plus grands penseurs, depuis les philosophes de l'Antiquité jusqu'aux scientifiques et philosophes contemporains. En dépit de discussions remontant à plus de deux mille ans, aucun consensus n'émerge aujourd'hui sur la nature du temps. Nous n'aborderons pas les aspects ontologiques, par exemple : le temps est-il un concept permettant de structurer notre pensée, ou possède-t-il une « réalité objective » ? De même nous laisserons de côté les aspects psychiques de la perception du temps et nous nous limiterons plus modestement à ce que l'on peut appeler « le temps des physiciens », qui remonte en fait à Galilée, à Newton et à Huygens. Le § 1.1 introduit très brièvement les deux conceptions principales d'un tel temps. Dans la première, le temps s'écoule du passé vers le futur et l'instant présent joue un rôle privilégié. Dans la seconde, souvent appelée celle de « l'univers-bloc », le temps est un simple balisage qui permet d'ordonner les événements et l'instant présent ne joue aucun rôle particulier. Les § 1.2 et 1.3 introduisent le temps absolu, ou universel, tel qu'il était conçu par exemple par Newton avant la révolution einsteinienne. Ce temps absolu, qui s'écoule quoi qu'il arrive, nous permettra de définir l'espace-temps newtonien comme étape préliminaire à l'espace-temps relativiste du chapitre 4, ainsi que les notions de simultanéité et de causalité.

1.1 Le temps subjectif

Envisageons une situation familière que tout un chacun peut imaginer. Confortablement assis dans un avion reliant Paris à New York, nous entendons l'annonce du pilote : « *Il est 15h30, heure de Paris, nous allons décoller dans quelques instants, la durée estimée du vol est de huit heures et nous atteindrons*

New York à 17 h 30, heure locale ». Cette annonce semble aller de soi, mais elle contient plusieurs hypothèses implicites sur la nature du temps. Tout d'abord, nous avons conscience de vivre l'instant présent, aujourd'hui le 1^{er} juillet 2015 à 15 h 30 : cet instant présent correspond à ce qui est en train de se passer et il joue pour nous un rôle particulier. Tout en pensant au passé, à l'embouteillage sur l'autoroute qui a failli nous faire rater l'avion, un passé qui a cessé d'exister puisque nous sommes bien installés dans notre fauteuil, nous envisageons un futur qui n'existe pas encore, notre arrivée à New-York, et nous nous préparons à occuper les huit heures de vol pendant lesquelles le temps va s'écouler dans une direction déterminée, du passé vers le futur. Nous avons la sensation d'un écoulement du temps : le futur est ouvert jusqu'à ce qu'il devienne le présent, tandis que le passé est fixé une fois pour toutes et ne peut être modifié, ce que nous dit Lady Macbeth :

« *ce qui est fait est fait... ce qui est fait ne peut être défait.* »

Enfin cela ne nous pose pas de problème que les horloges de Paris et de New-York soient décalées de 6 heures, c'est un effet de fuseau horaire, et nous ne mettons pas en doute que ces horloges puissent être synchronisées sur une horloge de référence, par exemple celle du méridien de Greenwich : l'effet de fuseau horaire est parfaitement banal.

Cependant ce concept intuitif du temps, qui est celui de la vie courante, est éminemment subjectif : nous faisons jouer un rôle privilégié au lieu où nous nous trouvons, Paris, et à l'instant présent, le 1^{er} juillet 2015 à 15 h 30, et huit heures vont s'écouler entre le décollage et l'atterrissage. Une première difficulté de ce point de vue est que « l'écoulement du temps » est une notion qui ne repose pas sur des bases très solides. On utilise souvent la métaphore du fleuve, mais dans quel lit s'écoule le temps, ou bien quelle est la vitesse de cet écoulement ? Une deuxième difficulté tient à l'incompatibilité de la notion d'instant présent avec la relativité einsteinienne. Nous avons supposé implicitement l'existence d'une horloge maîtresse à laquelle tous les observateurs, qu'ils soient situés sur Terre ou dans l'espace, peuvent se référer. On pourrait donc considérer que l'instant présent n'est pas subjectif puisqu'il pourrait être choisi identique pour tous les observateurs. Cependant nous verrons au chapitre 4 qu'il devient problématique de synchroniser deux horloges éloignées, ou de décider de la simultanéité de deux événements se passant en des endroits différents. En effet, en raison de la vitesse finie de propagation des signaux, synchronisation des horloges et définition de la simultanéité doivent être réexaminées. En relativité einsteinienne, chaque observateur dispose d'un temps qui lui est propre, qui est d'ailleurs appelé son *temps propre*, et l'instant présent est valable seulement pour un observateur particulier. Cependant les temps propres des divers

observateurs peuvent être reliés : l'instant présent n'est pas le même pour deux observateurs différents, mais il existe une relation précise entre leurs instants présents, qui dépend en relativité restreinte des trajectoires des deux observateurs et de leur vitesse sur ces trajectoires. En conséquence, faut-il nécessairement rejeter les notions intuitives du temps et de son écoulement ? Nous examinerons cette question à la fin du chapitre 5.

Laissant de côté pour l'instant les modifications introduites par la relativité, mentionnons que les notions intuitives d'instant présent et d'écoulement du temps ont été débattues depuis l'Antiquité. Vers 500 av. JC, Héraclite défendait la conception intuitive du temps :

« Tout s'écoule et rien ne dure ; toute chose cède la place et rien ne reste fixe. »

mais Parménide d'Élée proposait une vision opposée quelques décennies plus tard :

« Il reste alors un seul mot qui exprime le véritable chemin : Est. Et sur ce chemin on trouve bien des signaux montrant que Ce Qui Est n'a pas de commencement et ne sera jamais détruit. »

Parménide a été relayé depuis par nombre de scientifiques et de philosophes qui doutent de ou rejettent la conception d'écoulement du temps. Parmi eux on trouve Einstein, qui écrit à la mort de son ami Michele Besso :

« Michele a quitté ce monde étrange juste avant moi, mais cela n'a aucune importance. Pour nous, physiciens de conviction, la distinction entre le passé, le présent et le futur est une illusion, bien que tenace. »

En effet, aux yeux de ces scientifiques et philosophes, le présent est une notion subjective et « maintenant » est lié à un point de vue particulier, tout comme « ici » est lié à une position particulière. Dans ce point de vue, le temps ne joue pas un rôle différent de celui de l'espace. Si nous dessinons la carte d'un pays, peu importe le lieu où nous sommes situés sur cette carte. De même nous pouvons considérer le temps comme un simple balisage, qui range une série d'événements dans un ordre déterminé, l'instant présent n'ayant pas de signification particulière. Nous avons depuis longtemps abandonné l'idée que notre position dans l'Univers jouait un rôle privilégié, et de même il n'y a aucune raison pour que notre instant présent ait une signification objective : les notions de « ici » et « maintenant » reflètent un point de vue subjectif. Rejeter ce point de vue subjectif mène à la notion d'*univers bloc*, illustrée par exemple sur la figure 4.6b et par les commentaires associés, où la réalité est vue comme une entité unique dont le temps est un ingrédient, plutôt que comme une entité qui évolue au

cours du temps. Il n'existe pas plus de notion objective d'instant présent qu'il n'existe d'écoulement du temps. Le philosophe des sciences Huw Price résume ce point de vue dans un jeu de mots intraduisible : de même qu'une carte est un point vue de « *nowhere* », de nulle part, notre position sur la carte n'a aucune signification objective, de même nous devons imaginer le temps de « *nowhen* », sans référence à notre instant présent. Price ajoute qu'il est parfaitement abusif de déduire de ce point de vue que le temps n'existerait pas, que le temps serait une illusion. Nous pouvons décrire une série d'événements passés en associant à chaque événement son propre instant présent, l'instant où cet événement s'est produit, mais tous ces instants présents ont le même caractère, on ne peut pas en privilégier un plutôt qu'un autre. Notre série d'événements passés est tout à fait réelle mais elle peut être décrite en omettant toute idée d'instant présent privilégié.

Un autre problème que nous examinerons au chapitre 7 est celui de l'orientation du temps, aussi appelé problème de la flèche du temps. Nous pouvons orienter un axe d'espace, par exemple un axe vertical, mais nous savons pertinemment que la direction est arbitraire : l'orientation peut être choisie du bas vers le haut, ou l'inverse. Mais *a priori* il n'en est pas de même pour le temps : nous observons apparemment une orientation privilégiée, du passé vers le futur, et il existe donc une asymétrie de l'orientation du temps. Mais est-ce le temps lui-même qui est intrinsèquement asymétrique, observons-nous une asymétrie *du* temps, ou bien déduisons-nous son orientation de l'observation de processus qui sont asymétriques *dans* le temps ? Une analogie permettra de mieux comprendre cette distinction : imaginons une progression dans un long couloir où sont disposées des tables. Dans le premier cas, l'analogie est celle d'un couloir qui se rétrécit au fur et à mesure de notre progression, le couloir lui-même est asymétrique. Dans le second cas, le couloir est de largeur constante, mais les tables sont de plus en plus étroites au fur et à mesure de notre progression, c'est la disposition dans le couloir qui est asymétrique et non le couloir lui-même. C'est cette seconde option, celle de processus asymétriques dans le temps, que nous développerons au chapitre 7. En effet, la plupart des processus de la vie courante évoluent dans un seul sens. Quand nous plongeons un fer chauffé au rouge dans un récipient d'eau froide, nous nous attendons à ce que les températures de l'eau et du fer s'égalisent, et nous serions très surpris de voir le fer devenir de plus en plus chaud et l'eau de plus en plus froide, bien que cela soit parfaitement autorisé par la conservation de l'énergie. La plupart des processus de la vie courante nous conduisent à accepter l'idée que le temps ne s'écoule que dans un seul sens, du passé vers le futur : c'est la *flèche du temps*. Nous verrons que cette flèche du temps peut être mise en correspondance avec la croissance d'une quantité physique, l'entropie, introduite au XIX^e siècle pour l'étude des

machines à vapeur. L'évolution spontanée de l'entropie se fait toujours dans le sens de l'entropie croissante, ce qui permet de fixer le sens d'écoulement du temps.

1.2 L'espace-temps newtonien



FIGURE 1.1. Isaac Newton (1643-1727).

Jusqu'à la fin de ce chapitre, nous nous plaçons dans le cadre habituel d'un instant présent privilégié et d'écoulement du temps. Ces notions intuitives du temps sont implicites chez la plupart des physiciens depuis Galilée et Newton, et nous allons appeler ce temps le *temps newtonien*. Bien que notre sujet principal soit le temps, il est nécessaire d'introduire parallèlement l'espace car, en relativité, l'espace et le temps se combinent pour former une entité unique, *l'espace-temps*, un concept que nous allons introduire d'abord dans le cas newtonien, le plus familier et le plus intuitif. Les notions newtoniennes d'espace et de temps sont celles dont nous nous servons quotidiennement sans trop y réfléchir. En ce qui concerne le temps, nous supposons l'existence d'une horloge maîtresse universelle qui donne l'heure en tout point du globe et même de l'espace, ce qui n'est pas trop éloigné de la réalité présente. En effet, les physiciens ont défini un temps atomique international grâce à un réseau d'horloges atomiques qui sont périodiquement synchronisées : voir le chapitre 6. Bien sûr l'origine des temps n'est pas identique en tout point du globe : on peut choisir le temps de Paris, celui de New York, etc. mais cela ne pose aucun problème si l'on dispose d'un temps universel. Nous avons déjà noté que le concept de fuseau horaire est tout à fait banal et ne pose pas de problèmes de principe. Les distances sont

mesurées à l'aide de règles rigides étalonnées sur une règle maîtresse, le mètre étalon conservé au Pavillon de Breteuil à Sèvres. Cet énoncé, familier aux élèves de terminale scientifique des années 1960, est aujourd'hui obsolète et il sera corrigé au chapitre 4. L'espace-temps possède quatre dimensions, une de temps et trois d'espace. Comme il n'est pas possible de dessiner dans quatre dimensions, on a souvent recours pour les figures à des espaces-temps à trois dimensions avec deux dimensions d'espace et une de temps, comme sur la figure 1.3, voire à une seule dimension d'espace, comme sur la figure 1.2.

La notion de base de l'espace-temps est celle d'événement. Un événement est défini par sa position dans l'espace et par l'instant où il se produit, sa date t . Par exemple un TGV et un TER se croisent sur la ligne Bordeaux-Paris : l'événement « croisement des deux trains » est défini par sa position, 300 km de Paris, et sa date, $t = 13\text{ h }00$ le 1^{er} juillet 2015. Naturellement le croisement des deux trains n'est pas instantané, et il nous faut être un peu plus précis, en définissant par exemple l'instant du croisement par le moment où les deux têtes de train coïncident. Cette définition comporte encore une certaine imprécision : comme toujours en physique, temps et distances ne peuvent être mesurés qu'avec une certaine marge d'erreur. Si l'on fait abstraction de cette imprécision, un événement est défini par sa coordonnée de temps, ou date, t et sa position dans l'espace, par exemple son abscisse x si les événements ont lieu sur une droite comme sur la figure 1.2.

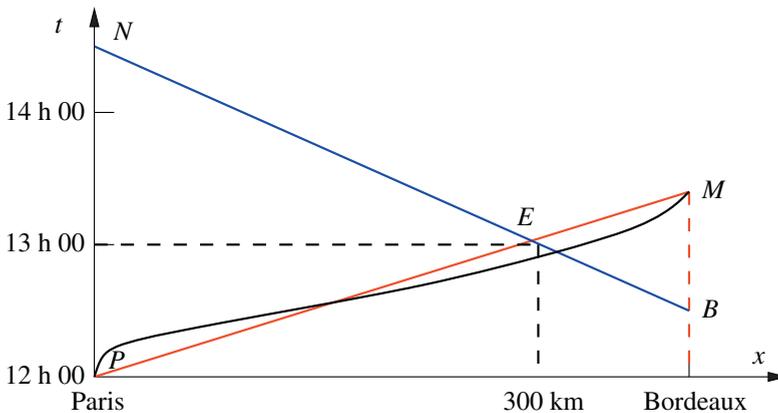


FIGURE 1.2. Diagramme d'espace-temps : lignes d'univers du TGV Paris-Bordeaux (rouge) et du TER Bordeaux-Paris (bleu). Les deux trains se croisent à 13 h 00, à 300 km de Paris. L'événement E « croisement des deux trains » a pour coordonnée de temps $t = 13\text{ h }00$ et pour coordonnée d'espace $x = 300\text{ km}$. Comme le TGV va plus vite que le TER, la pente de sa ligne d'univers est plus faible en valeur absolue que celle du TER. Une ligne d'univers plus réaliste, tenant compte de l'accélération au départ de Paris et du freinage à l'arrivée à Bordeaux, est tracée en noir pour le TGV.

Un concept fondamental pour la suite est celui de ligne d'univers. Reprenons l'exemple ci-dessus du TGV Paris-Bordeaux et portons sur un graphique sa position en fonction du temps. L'usage en relativité est de porter le temps sur l'axe vertical (ordonnées) et la position sur l'axe horizontal (abscisses). La courbe donnant la position du TGV en fonction du temps est appelée sa *ligne d'univers*. Le type de graphique illustré sur la figure 1.2, appelé *diagramme d'espace-temps*, était familier aux agents de la SNCF avant que les ordinateurs ne relèguent ces graphiques tracés à la main au rayon des souvenirs. Supposons que le TGV Paris-Bordeaux quitte Paris à 12 h 00 avec une vitesse de 300 km/h et qu'un TER Bordeaux-Paris quitte Bordeaux à 12 h 30 avec une vitesse de 200 km/h. La détermination graphique des coordonnées du croisement des deux trains est alors un exercice d'école primaire : la distance Paris-Bordeaux étant de 400 km, et en admettant que les deux trains roulent toujours à vitesse constante, à quelle heure vont-ils se croiser, autrement dit quelle sera la date de l'événement « croisement des deux trains » ? En raison de la vitesse constante, la courbe donnant la distance parcourue en fonction du temps est dans les deux cas un segment de droite : les segments de droite PM et BN sont les lignes d'univers respectives du TGV Paris-Bordeaux et du TER Bordeaux-Paris. Comme nous avons porté le temps t suivant l'axe vertical et la distance x suivant l'axe horizontal, la pente des lignes d'univers est inversement proportionnelle à la vitesse : la ligne d'univers du TGV (en rouge), plus rapide, a une pente moins forte que celle du TER (en bleu), plus lent. L'événement E « croisement des deux trains » a pour coordonnée de temps $t = 13$ h 00 et pour coordonnée d'espace $x = 300$ km, sa distance de Paris. En fait, la vitesse des trains n'est pas constante et les lignes d'univers sont des courbes, et non des droites : voir la courbe plus réaliste tracée en noir sur la figure 1.2. Le concept de ligne d'univers trouve une place naturelle dans le cadre de l'univers-bloc : l'histoire du TGV entre Paris et Bordeaux est résumée par la donnée de sa ligne d'univers entre ces deux villes.

1.3 Causalité et simultanéité

Un ensemble de règles rigides permet de mesurer la position d'un événement dans l'espace, mais comment les physiciens mesurent-ils le temps ? Ce point sera traité en détail au chapitre 3, et nous nous contentons ici de quelques indications sommaires. Le fonctionnement d'une horloge repose toujours sur un phénomène périodique, par exemple l'oscillation périodique du balancier de l'horloge de nos grands-mères. Aujourd'hui, les montres à quartz sont fondées sur les vibrations périodiques d'un cristal, et les horloges les plus précises, celles dont nous allons faire un usage intensif par la suite, sont les *horloges atomiques*,

où la vibration périodique est celle d'un rayonnement électromagnétique défini par l'écart entre deux niveaux d'énergie spécifiques d'un atome ou bien d'un *ion*, un atome ayant perdu un ou plusieurs électrons. De fait, la définition de la seconde repose aujourd'hui sur la vibration d'un rayonnement électromagnétique particulier de l'atome de césium : par définition, cette vibration effectue 9 192 631 770 cycles dans une seconde, environ dix milliards, ou 10^{10} cycles par seconde. Autrement dit, la *fréquence* de la vibration est de 9 192 631 770 Hz. Le *hertz* (Hz) est l'unité de fréquence, familière en radio : par exemple à Nice, France Culture émet sur une fréquence de 101,9 MHz (mégahertz), environ 10^8 Hz. La précision relative des horloges à césium est de l'ordre de 10^{-13} , ce qui correspond à une dérive d'une seconde tous les 300 000 ans. C'est une variante de ces horloges, l'horloge à rubidium, qui est embarquée dans les satellites du GPS. En refroidissant les atomes, on arrive à une précision relative de 10^{-15} , soit une dérive d'une seconde tous les 30 millions d'années. Il existe des horloges encore plus précises, les horloges atomiques optiques, dont la précision relative peut approcher 10^{-18} . Le lecteur trouvera un exposé détaillé sur ces horloges au § 3.4.

Dans l'espace-temps de Newton, définir la *simultanéité* de deux événements ne pose aucun problème, étant donné que leur date est donnée par un temps absolu. Deux événements éloignés seront considérés comme simultanés si les horloges associées, synchronisées sur une horloge maîtresse universelle, indiquent la même heure. On définit de même le passé et le futur d'un événement : tous les événements dont la date est antérieure se situent dans le passé, tous ceux dont la date est postérieure dans le futur. Cela permet de définir une notion de *causalité* : un événement ne peut avoir été influencé que par des événements situés dans son passé, et il ne peut influencer que des événements situés dans son futur. *En physique newtonienne, la structure causale se confond avec la structure chronologique*, ce qui ne sera plus le cas en relativité. De plus, tout comme le temps, l'espace newtonien est absolu, on peut donc définir une notion de repos absolu : un objet est au repos s'il est immobile dans cet espace. Cette notion de repos absolu est battue en brèche par le principe de relativité que nous verrons au chapitre 4, mais que l'on peut déjà introduire en physique newtonienne. Le schéma tracé sur la figure 1.3 résume la conception newtonienne de l'espace-temps.

On peut se poser la question de la fiabilité des horloges lorsqu'elles sont soumises à des influences brutales. Que se passe-t-il par exemple quand on les met dans une fusée qui accélère ? Il est évident que le fonctionnement d'une horloge à balancier serait profondément perturbé si elle était embarquée dans une fusée ou une station spatiale : en impesanteur, le balancier n'oscillerait plus ! Pour estimer la fiabilité d'une horloge atomique, il faut comparer l'accélération

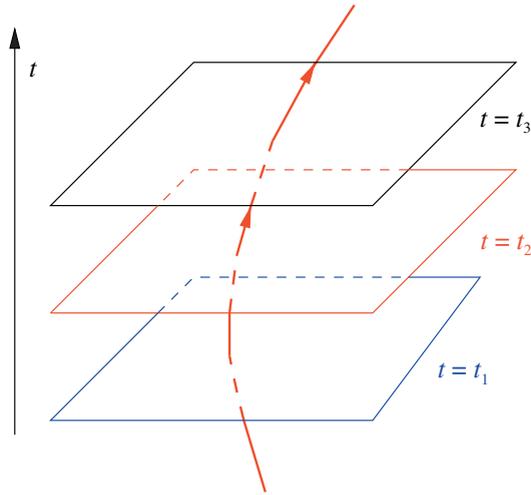


FIGURE 1.3. Espace-temps à trois dimensions de Newton, deux dimensions d'espace horizontales et une de temps verticale. Les surfaces de simultanéité contiennent tous les événements simultanés, ayant la même date. Ce sont des plans, le plan $t = t_1$ par exemple. Tous les événements situés dans un tel plan sont simultanés. On a aussi représenté en rouge la ligne d'univers d'un observateur.

des électrons sur leur orbite dans un atome, à l'origine de la structure des niveaux d'énergie, avec celle d'une fusée : on trouve que l'accélération des électrons sur leur orbite l'emporte par un facteur de l'ordre de 10^{22} , même si l'accélération de la fusée vaut 10 fois celle de la pesanteur. Les horloges atomiques sont donc insensibles aux accélérations usuelles et fonctionneront sans problème, même si leur parcours est un peu chahuté. Nous insistons sur l'idée que c'est le *mécanisme* des horloges qui n'est pas affecté par les accélérations, car nous verrons bientôt l'influence de la vitesse ou de la gravité sur les *indications* des horloges telles qu'elles peuvent être lues par différents observateurs.

2

Les différentes échelles de temps

Les durées considérées par les physiciens s'étalent sur une plage de 60 ordres de grandeur, depuis le temps de Planck (environ 10^{-43} s) jusqu'à l'âge de l'Univers ($4,36 \times 10^{17}$ s) : la figure 2.1 donne quelques repères dans ces différentes échelles de temps. Rappelons en deux lignes la notation exponentielle dont nous nous sommes servis. Pour les très grands nombres, 10^N représente 1 suivi de N zéros, par exemple un million = $1\,000\,000 = 10^6$. Pour les très petits nombres, 10^{-N} représente 1 précédé de $(N - 1)$ zéros après la virgule, par exemple un milliardième = $0,000\,001 = 10^{-6}$.

2.1 Du temps de Planck à la milliseconde

La durée la plus courte que peuvent envisager les physiciens est le *temps de Planck* t_{Planck} , formé par les trois constantes fondamentales de la physique, la *constante de Planck* h qui fixe l'échelle des phénomènes quantiques, la *vitesse de la lumière* c fondamentale en relativité, et la *constante de gravitation* G (le facteur 2π est une convention dans la définition) :

$$t_{\text{Planck}} = \sqrt{\frac{Gh}{2\pi c^5}} = 5,4 \times 10^{-43} \text{ s.} \quad (2.1)$$

C'est l'échelle de temps où les phénomènes quantiques ne peuvent plus être négligés en relativité générale, la théorie relativiste de la gravitation (chapitre 5). Ce temps joue un rôle crucial dans l'histoire du Big Bang : étant donné qu'il n'existe pas pour l'instant de théorie quantique de la gravitation, tout ce qui peut éventuellement se passer à des temps plus petits que t_{Planck} nous est aujourd'hui inaccessible. Un autre de point de repère important dans l'histoire du Big Bang est

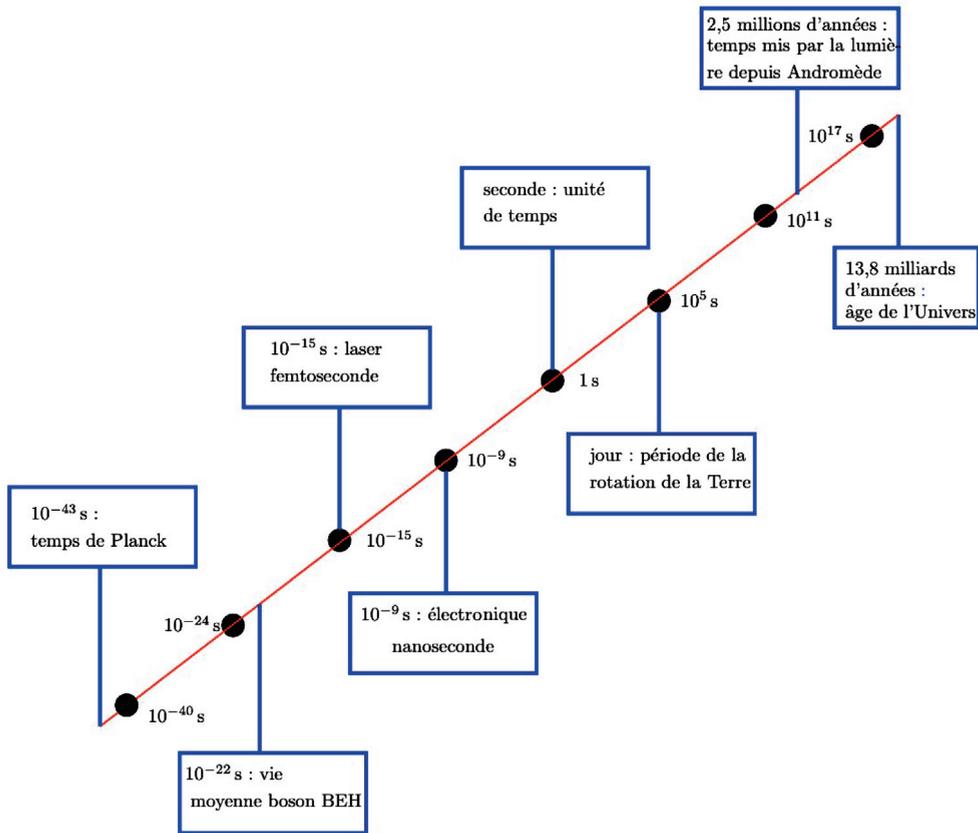


FIGURE 2.1. Les différentes échelles de temps dans le monde de la physique.

la fin de l'inflation, environ 10^{-35} s. Si l'on passe à la physique des particules élémentaires, la durée de vie du boson BEH (Brout-Englert-Higgs) mis en évidence au collisionneur LHC en 2012 est de 10^{-22} s (figure 2.2). En physique atomique, les durées les plus courtes rencontrées aujourd'hui s'expriment en *attosecondes* : une attoseconde vaut 10^{-18} s (figure 2.3). Les impulsions les plus brèves obtenues avec un laser durent quelques dizaines d'attosecondes. Un changement de conformation d'une molécule prend quelques dizaines de *femtosecondes*, où une femtoseconde vaut 10^{-15} s (figure 2.4). Les lasers femtosecondes, qui donnent des intensités très fortes au pic de leurs impulsions très courtes, sont largement utilisés en optique non linéaire pour produire de nouvelles longueurs d'onde. L'échelle de temps caractéristique de l'électronique moderne est le milliardième de seconde, ou *nanoseconde*, 10^{-9} s. En une nanoseconde, la lumière parcourt

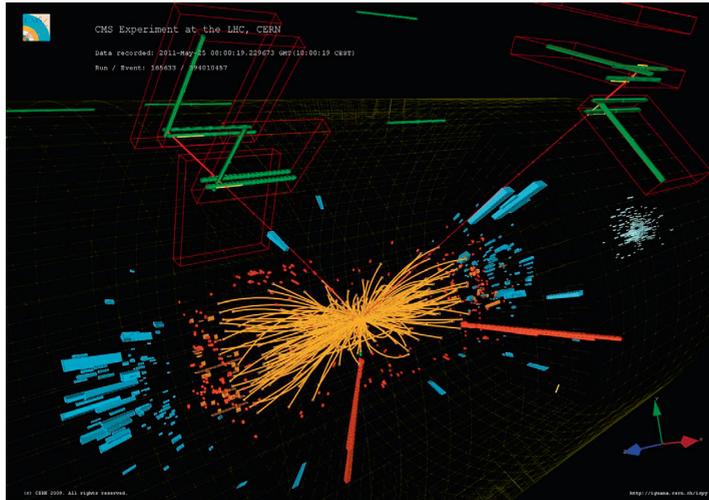


FIGURE 2.2. Un événement correspondant vraisemblablement à un boson BEH (plus connu sous le nom de boson de Higgs) observé dans le détecteur CMS du LHC au CERN. Cette désintégration se fait en deux bosons Z^0 , dont l'un donne deux muons (traces vertes) et l'autre deux électrons (traces rouges). La durée de vie du boson BEH est de 10^{-22} s. Reproduit avec l'autorisation du CERN.

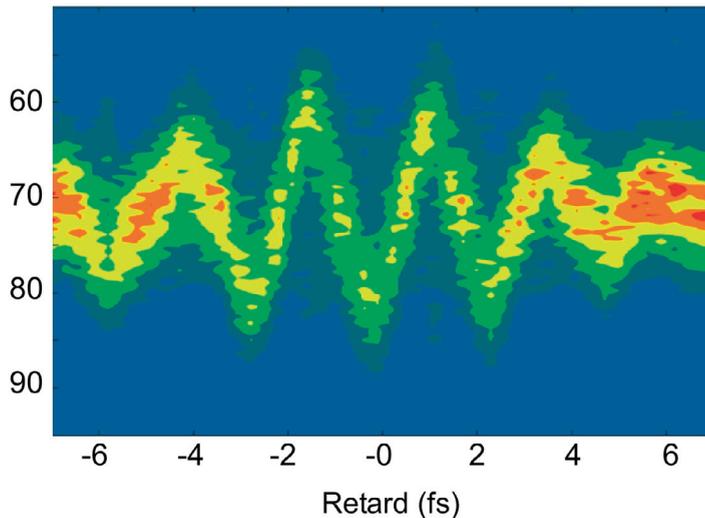


FIGURE 2.3. Une impulsion laser de quelques dizaines d'attosecondes (10^{-18} s) est focalisée sur un gaz, où elle ionise les électrons. La figure montre la distribution énergétique des électrons créés par photo-ionisation sous l'effet des impulsions attosecondes et modifiée par un champ infrarouge, en fonction du retard relatif par rapport à celui-ci. La structure femtoseconde de la dynamique des électrons est clairement résolue.

30 cm, et le temps d'accès à la mémoire d'un ordinateur est de quelques dizaines de nanosecondes (figure 2.5).

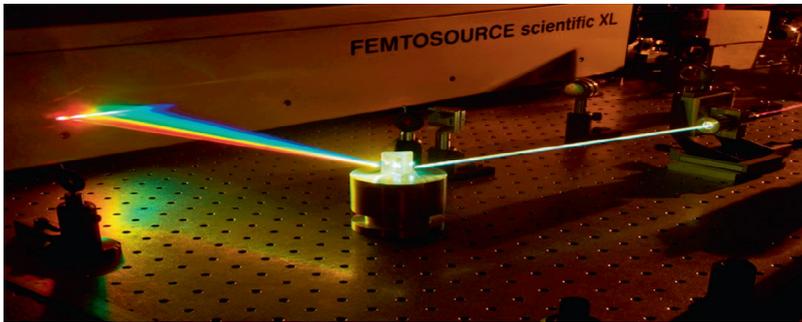


FIGURE 2.4. Le faisceau d'un laser femtoseconde (10^{-15} s) est focalisé dans une lame de saphir (à droite). L'interaction non linéaire dans le saphir produit une lumière blanche, ce qui donne un « laser blanc ». Cette lumière blanche est ensuite décomposée spectralement par un prisme.

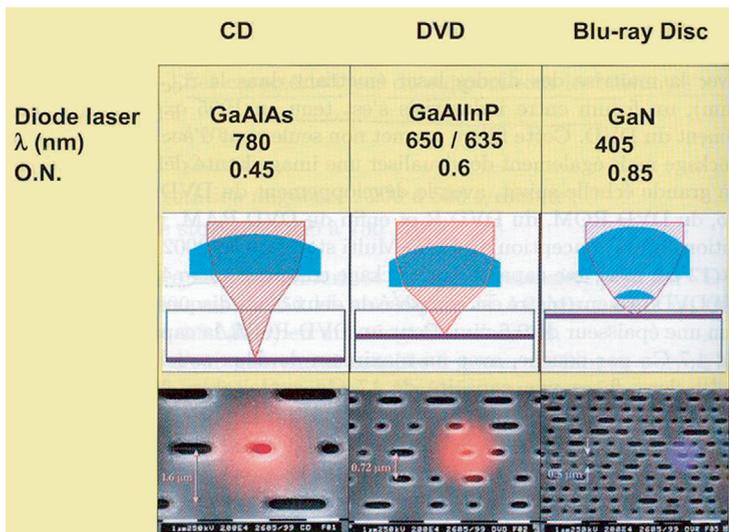


FIGURE 2.5. Schéma de lecture d'un CD ou d'un DVD avec un laser de longueur d'onde λ . Dans le cas d'un DVD Blu Ray (à droite) la distance entre les puits est de $0,5 \mu\text{m}$ et le temps entre la lecture de deux puits successifs est de quelques nanosecondes (10^{-9} s).

2.2 De la milliseconde à mille années

Avec le millième de seconde, ou *milliseconde* (10^{-3} s), nous entrons dans les durées à notre échelle. La durée de transmission d'un message nerveux est de quelques millisecondes. La seconde est l'unité de base pour la mesure du temps ; elle était définie il y a encore 50 ans comme la fraction $1/86\,400$ du jour solaire moyen. Elle est définie aujourd'hui à partir d'horloges atomiques au césium (figure 2.6). Une heure vaut 3 600 secondes, et une cellule se divise en quelques dizaines d'heures. Le jour solaire vaut 86 400 s, et l'activité de tous les organismes vivants suit un cycle d'environ un jour, ou rythme circadien. Passons maintenant dans le domaine de l'astronomie : la durée d'une révolution complète de la Terre sur son orbite autour du Soleil est d'un an, soit $3,16 \times 10^7$ s (figure 2.7). La lumière met quatre ans pour nous parvenir de l'étoile la plus proche. La plupart des exemples ci-dessus ont été pris dans le monde de la physique et de l'astronomie. Il est intéressant aussi de mentionner les échelles de temps pour la vie sur la Terre qui est apparue il y a 3,5 millions d'années.

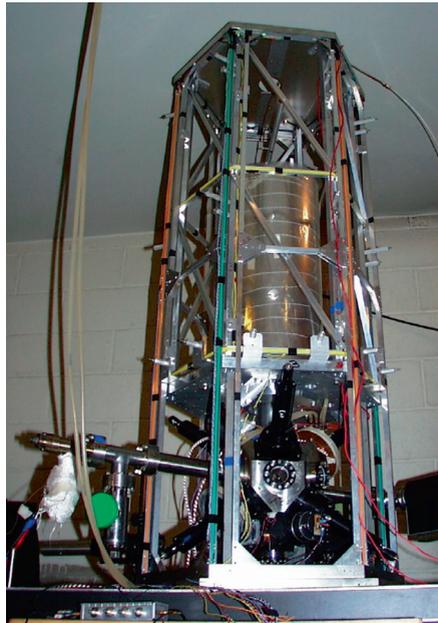


FIGURE 2.6. L'horloge à fontaine d'atomes froids de césium à l'Observatoire de Paris utilisée pour la définition de la seconde. La définition actuelle de la seconde est la durée de 9 192 631 770 oscillations de l'atome de césium 133 entre les deux niveaux hyperfins de son état fondamental.

Certains insectes ne vivent qu'un jour, les baleines jusqu'à 200 ans, et la durée de vie d'un séquoia peut dépasser mille ans.

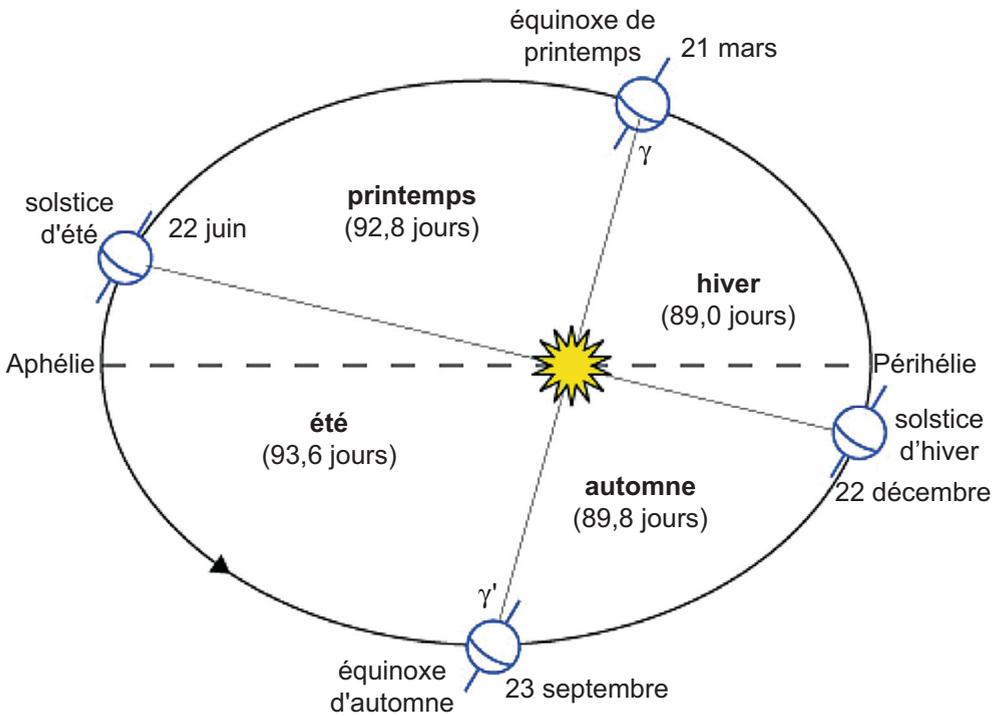


FIGURE 2.7. L'orbite terrestre.

2.3 De mille années à 10 milliards d'années

Il faut 100 000 ans à la lumière pour traverser notre galaxie. L'homme moderne (*Homo sapiens*) est apparu en Afrique il y a environ 200 000 ans, et les premiers représentants du genre *Homo* il y a environ deux millions d'années. Il faut 2,5 millions ($2,5 \times 10^6$) d'années à la lumière pour nous parvenir depuis la galaxie la plus proche, Andromède (figure 2.8). L'âge du système solaire est de 4,5 milliards d'années. Enfin l'âge de l'Univers est estimé aux environs de 13,8 milliards d'années et la mesure du fond diffus cosmologique par le satellite Planck (figure 2.9) nous donne une image de l'Univers 380 000 années après le Big Bang.



FIGURE 2.8. La nébuleuse (galaxie) d'Andromède. La lumière émise par cette galaxie met 2,5 millions d'années à nous parvenir. Photographie par R. Gendler.

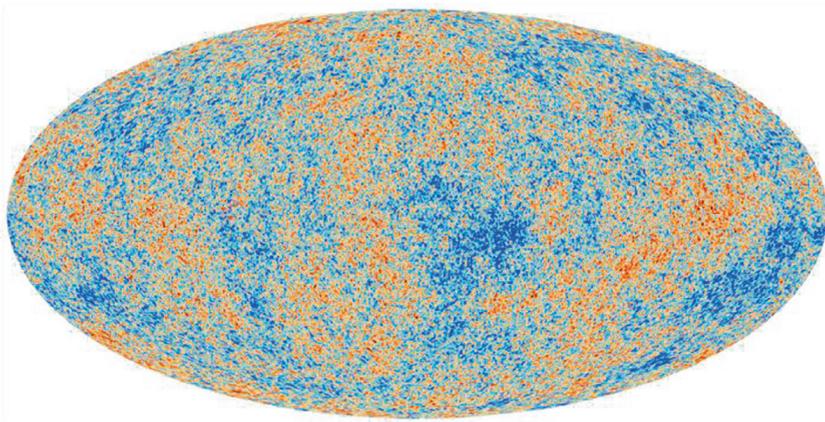


FIGURE 2.9. Le fond diffus cosmologique vu par le satellite Planck. C'est une image de l'Univers 380 000 ans après le Big Bang, soit environ 13,8 milliards d'années. La figure montre les fluctuations en température, de l'ordre de 10^{-5} en valeur relative, autour d'une température moyenne de 2,7 K.

3

La mesure du temps

Dans ce chapitre, nous décrivons les principales méthodes de mesure du temps, depuis les cadrans solaires de l'Antiquité jusqu'aux horloges atomiques les plus récentes. En fait, la mesure du temps recouvre deux notions qui ne sont pas indépendantes, mais néanmoins distinctes : (1) la mesure des durées et (2) la détermination des dates. La mesure de la durée d'un processus s'effectue à l'aide d'une horloge située au voisinage de l'endroit où il se déroule, c'est donc une mesure locale. Pour déterminer la date d'un événement, par exemple celle de la fin d'un processus, on peut évidemment décider d'un instant zéro pour l'horloge et mesurer les dates à partir de cet instant zéro. Mais la date d'un événement devrait être accessible à toute personne sur Terre ou même dans l'espace, ce qui pose le problème du transfert du temps qui sera examiné au chapitre 6. Enfin il est souhaitable que la date soit en accord avec les phénomènes astronomiques courants : les horloges atomiques utilisées pour mesurer les durées sont infiniment plus stables que la rotation de la Terre, et cela nécessite l'ajustement périodique du temps astronomique sur le temps atomique, faute de quoi on pourrait voir un jour le Soleil se lever à minuit !

3.1 Repérer le temps : depuis la préhistoire

Il a toujours été nécessaire pour l'humanité de construire des calendriers et de repérer le temps, au-delà de l'alternance du jour et de la nuit marqués par l'apparition et le coucher du Soleil. La plupart des civilisations depuis les plus anciennes et sur toutes les parties du globe terrestre ont développé des méthodes pour repérer les dates marquant périodiquement d'importantes fêtes religieuses ou simplement la planification des travaux agricoles. La mesure du temps a

toujours été reliée à l'astronomie et son repérage faisait appel à l'observation du mouvement des astres les plus proches, le Soleil et la Lune, si manifestement extérieurs au contrôle des hommes et par là-même sources de mystère et d'adoration. Encore à l'époque actuelle, les travaux des chercheurs s'efforçant d'améliorer la précision des horloges jusqu'à des limites extrêmes sont largement motivés par les applications aux grands instruments de l'astrophysique. Nous verrons ultérieurement comment aujourd'hui, à l'inverse de l'histoire ancienne, c'est la mesure précise du temps qui permet d'approfondir la connaissance du système solaire et de l'ensemble de l'Univers.

Les plus anciens vestiges d'horloges préhistoriques sont sans doute les impressionnantes et mystérieuses structures de grès de Stonehenge dans le Sud de l'Angleterre, qui remontent à plus de 4 000 ans (figure 3.1). Alors que ces structures ont longtemps été considérées comme de simples monuments votifs, les archéologues pensent aujourd'hui qu'ils étaient plutôt de gigantesques observatoires destinés à observer le mouvement des corps célestes et à établir des calendriers.



FIGURE 3.1. Les structures de pierres néolithiques de Stonehenge en Angleterre, un des premiers observatoires des astres utilisés pour repérer le temps et établir des calendriers.

De nombreuses civilisations disparues ont laissé des vestiges de calendriers élaborés, particulièrement les Mayas, les Incas et les Aztèques en Amérique centrale et du Sud. Remarquable est Cuzco, la capitale des Incas, où l'on a découvert que le plan de la ville lui-même était un gigantesque calendrier : à travers la ville, des passages dégagés offraient une vision du Soleil à son lever et à son coucher

lors d'occasions importantes. Les Aztèques aussi avaient plusieurs calendriers d'une grande complexité, comme le montrent les ruines de leur capitale Tenochtitlan ; ils observaient en particulier le passage des Pléiades à certaines périodes et à certains points du ciel, ce qui marquait le bon vouloir des dieux à leur égard.

Ce sont les Égyptiens qui eurent les premiers l'idée de diviser la journée en unités plus petites. On a retrouvé des obélisques isolés datant de 3 500 ans avant Jésus-Christ, qui par leur ombre portée servaient à l'évidence de repères pour l'écoulement du temps au cours de la journée. La lumière a longtemps été utilisée pour la mesure du temps et c'est encore la lumière qui sert aujourd'hui dans les horloges les plus précises qui utilisent des lasers. Dans l'Antiquité, c'était d'ailleurs plutôt l'ombre que la lumière qui servait de repère (figures 3.2 et 3.3).

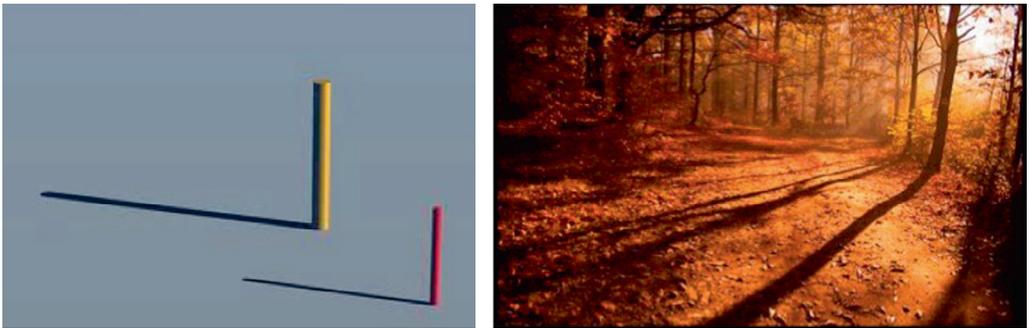


FIGURE 3.2. Dès l'Antiquité, l'humanité avait utilisé l'ombre pour mesurer le temps au cours d'une journée.



FIGURE 3.3. Le repérage de l'heure par l'ombre portée est effectué grâce à des cadrans solaires, qui décorent encore aujourd'hui jardins et monuments publics.

On avait très rapidement compris que la lumière se propageait en ligne droite, mais les savants grecs comme Ptolémée et Euclide pensaient que la lumière allait de l'œil vers l'objet source et non le contraire, ce qui n'a pas empêché l'élaboration des cadrans solaires. Cependant il faudra attendre le développement de la science arabo-islamique aux X^e et XI^e siècles pour que la notion de rayon lumineux soit véritablement conceptualisée.

Les anciens cadrans solaires ont longtemps permis aux hommes de repérer l'heure pendant la journée. Pourtant cela n'a pas suffi pour établir des calendriers. En effet, les cycles des jours, des mois et des années ne se recourent pas parfaitement : il faut 365 jours 1/4 à la Terre pour faire son tour complet autour du Soleil, alors que la Lune accomplit 13 cycles de rotation autour de la Terre en 364 jours. Cela a longtemps donné du fil à retordre aux mathématiciens, astronomes et autres concepteurs de calendriers.

D'autres techniques de mesure du temps ont été mises en œuvre, qui reposent sur des principes autres qu'astronomiques et ont été progressivement de plus en plus élaborées. De fait, nombre d'activités humaines requièrent une connaissance de la date et de l'heure. Or nous verrons qu'il est plus facile de mesurer des durées que de repérer des dates. En général les mesures consistent à évaluer des délais à partir de repères fixés dans le temps. La suite de ce chapitre est consacrée à la description des différentes méthodes qui ont servi au fil des siècles à mesurer des durées et à construire des montres et des horloges. Nous détaillerons les principes qui ont permis des améliorations de plus en plus spectaculaires de la précision de ces instruments, évoquant au passage certaines de leurs applications qui ont engendré le besoin d'améliorer encore cette précision. Les méthodes utilisées aujourd'hui pour transférer le temps d'un point à un autre du globe et synchroniser les horloges entre elles afin de définir un temps universel sont décrites au chapitre 6.

3.2 Mesurer les durées avec un phénomène continu : premières horloges

Se servir de l'écoulement d'un matériau est une méthode possible pour mesurer une durée, et cette méthode a été utilisée depuis l'Antiquité pour la vie courante. Le matériau peut être un fluide comme de l'eau ou un matériau granulaire comme du sable. La vitesse d'écoulement d'un fluide est liée à son énergie potentielle dans le récipient, qui décroît quand le fluide s'écoule. Elle dépend de la hauteur du fluide dans le récipient, qui dépend elle-même de la forme du récipient. L'écoulement n'est donc pas régulier et la mesure ne repose pas en général sur un phénomène strictement linéaire en fonction du temps. Il est

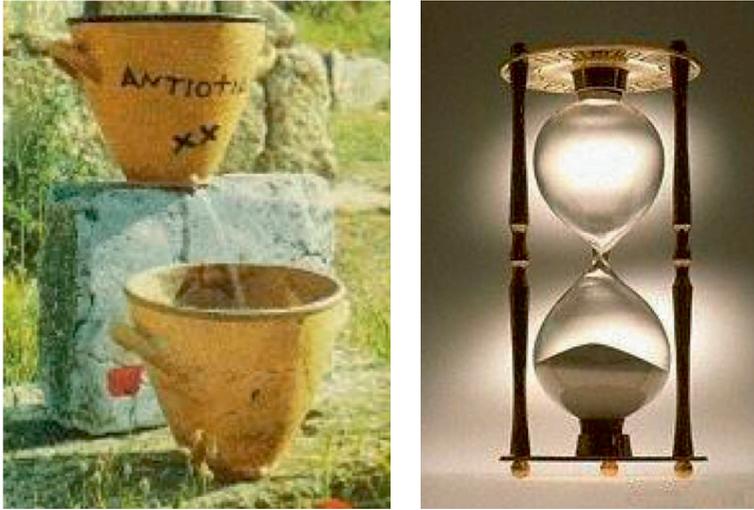


FIGURE 3.4. Mesure d'une durée par écoulement. À gauche, une clepsydra à eau grecque. À droite, un sablier contemporain.

possible de calculer des formes de vases où la section varie de telle sorte que le débit soit constant en appliquant la loi de Bernoulli. Les peuples de l'Antiquité avaient conscience du problème et certaines de leurs clepsydras avaient des formes étudiées pour que l'écoulement soit le plus régulier possible (figure 3.4). Les plus anciennes clepsydras retrouvées à Karnak datent du temps des Égyptiens ; leur précision était de l'ordre de 5 à 10 minutes, et elle fut améliorée par les Grecs. Ce sont les Perses et plus tard les Chinois qui poussèrent le plus loin le principe de la clepsydra.

En effet, vers le X^e siècle, les Chinois commencèrent à construire des horloges à eau, toujours fondées sur la durée d'un écoulement, mais en y ajoutant des systèmes mécaniques complexes pour remonter l'eau à base de roues, d'engrenages et de petits récipients qui se remplissaient et se vidaient à tour de rôle. La figure 3.5 montre une horloge à eau chinoise du XIV^e siècle de 10 m de hauteur. De telles horloges se fabriquaient aussi en Inde, incluant parfois des automates. Elles étaient d'authentiques symboles de la grandeur des États.

Des variantes des horloges à eau chinoises ont aussi été construites en Europe, en particulier en Allemagne dès le XIII^e siècle, et elles connurent une grande popularité. Mais elles étaient très peu précises et de plus elles avaient l'inconvénient de geler l'hiver ! L'eau fut bientôt remplacée par du sable, mais le poids limitait la mesure à de courts intervalles de temps. C'est malgré tout avec de telles horloges que les marins mesurèrent pendant longtemps la vitesse

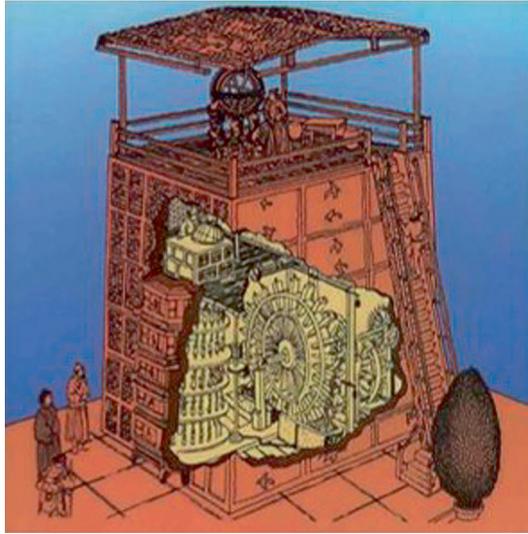


FIGURE 3.5. Une horloge à eau chinoise très sophistiquée du XIV^e siècle.

de leur bateau, laissant flotter librement une corde pourvue de nœuds réguliers dont ils mesuraient le défilement pendant un intervalle de temps repéré par l'horloge à bord. On mesure encore aujourd'hui la vitesse d'un navire en nœuds (mais pas avec une corde!) : rappelons qu'un nœud est une *vitesse* correspondant à un mille nautique (1 852 m) par heure.

Sur un principe voisin, les horloges mécaniques ont succédé aux horloges à écoulement à partir du XIV^e siècle, à base de poids attachés à une corde enroulée autour d'un cylindre. Les poids furent ensuite remplacés par des ressorts, sans améliorer pour autant la médiocre reproductibilité des mesures liée à la friction variable de la corde sur le cylindre.

3.3 Mesurer le temps avec un phénomène périodique

Pour mesurer une durée avec une fiabilité meilleure que celle d'un écoulement, on peut se fonder sur une méthode alternative exploitant un phénomène périodique. L'avantage d'une telle méthode est que la période de l'oscillation est une propriété du système lui-même et ne dépend pas en première approximation de facteurs extérieurs. Un *oscillateur* peut être considéré comme une règle graduée destinée à mesurer le temps. La figure 3.6 montre l'analogie avec la mesure d'une longueur à l'aide d'une règle dont on compte les graduations. Pour mesurer une durée, il suffit de compter le nombre de périodes des oscillations entre

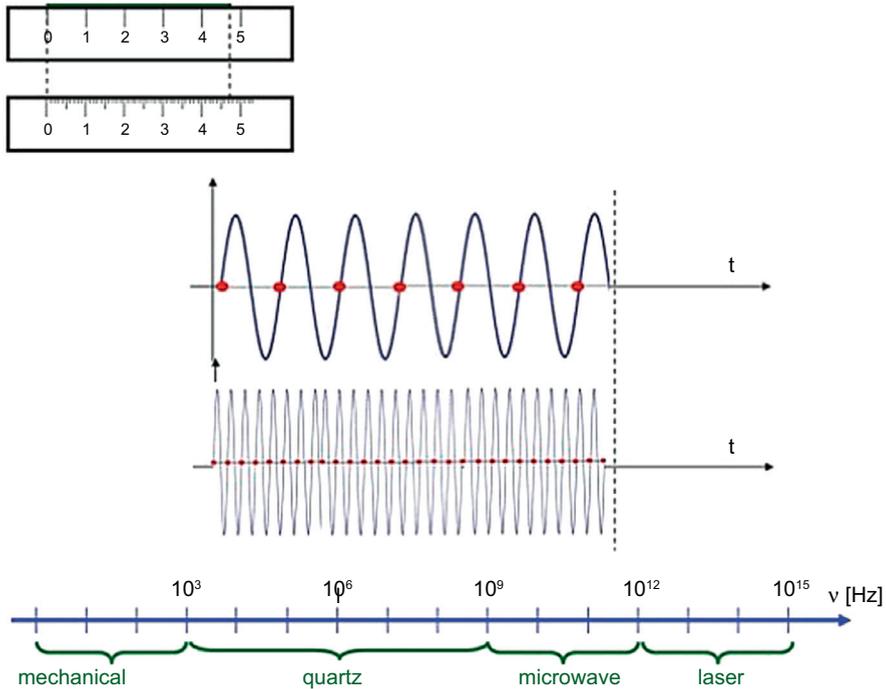


FIGURE 3.6. Principe de la mesure d'une durée avec un oscillateur et analogie avec la mesure d'une longueur avec une règle graduée (en haut de la figure) : plus les graduations de la règle sont serrées, plus grande est la précision sur la mesure de la longueur. Il en est de même pour la mesure du temps avec un oscillateur : plus la fréquence du phénomène périodique est élevée, plus la mesure de la durée est précise. La partie inférieure de la figure indique l'ordre de grandeur des fréquences pour divers types d'oscillateurs servant à la mesure du temps, depuis les oscillateurs mécaniques jusqu'aux lasers.

un instant de départ et un instant d'arrivée. Avec une règle, plus les graduations sont fines, plus précise est la mesure. De même, pour un intervalle de temps donné, plus la fréquence oscillatoire est grande, plus grand est le nombre de périodes décomptées et plus la mesure d'une durée est précise, comme le montre la figure 3.6.

Les oscillateurs utilisés pour les horloges sont très variés. Ils peuvent être mécaniques avec une fréquence se situant dans une gamme allant du hertz (Hz) au kilohertz (kHz), ou électriques, essentiellement à quartz dans le domaine du kHz au gigahertz ($1 \text{ GHz} = 10^9 \text{ Hz}$). À des fréquences plus élevées, on trouve les horloges atomiques dans le domaine des micro-ondes, et enfin les lasers dans une gamme de fréquence encore plus élevée. Tous les instruments fondés sur le principe du décompte d'oscillations ont une précision croissante lorsque la gamme de fréquences utilisée augmente.

3.3.1 Les oscillateurs mécaniques

Les *oscillateurs mécaniques* sont de nature extrêmement variée. Les oscillateurs à pendule ont permis de construire des horloges qui ont servi pendant des siècles. Ils ont de nombreux défauts que les horlogers ont cherché à compenser de mieux en mieux, mais qui restent intrinsèquement difficiles à corriger. En première approximation la fréquence d'oscillation ν d'un pendule est donnée par :

$$\nu = \frac{1}{2\pi} \sqrt{\frac{g}{l}} \left(1 - \frac{1}{16} \theta_0^2 \right), \quad (3.1)$$

où l est la longueur du pendule, g l'accélération de la pesanteur ($g \simeq 9,81 \text{ m}\cdot\text{s}^{-2}$ à Paris) et θ_0 l'amplitude de l'oscillation. La longueur varie en fonction de l'environnement, avec la température sous l'effet de la dilatation ou de l'humidité de l'air. On voit aussi sur la formule (3.1) que la fréquence dépend légèrement de l'amplitude des oscillations, que l'on rend aussi faibles que possible pour les pendules commerciaux afin de bénéficier de l'isochronisme des petites oscillations.

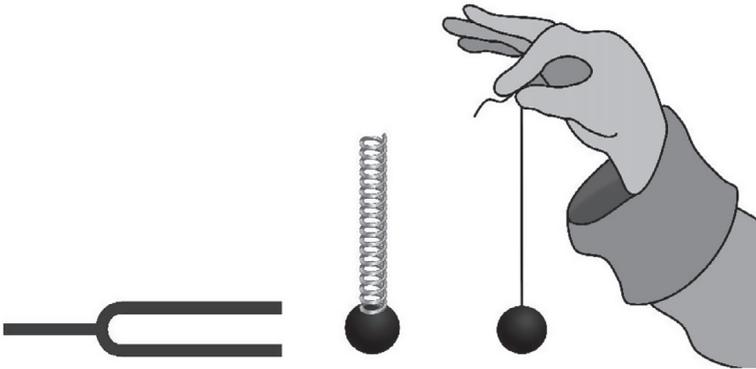


FIGURE 3.7. Toutes sortes d'oscillateurs mécaniques peuvent servir à mesurer le temps : vibrations du diapason, extensions-contractions du ressort, oscillations du pendule.

Enfin l'on remarque que la fréquence dépend de l'accélération de la pesanteur g , qui n'est pas une constante universelle. En premier lieu, elle dépend de l'altitude z au-dessus du sol. En effet, g est inversement proportionnel au carré de la distance R au centre de la Terre. Pour une variation de l'altitude z petite devant R , la variation relative de l'accélération de la gravité vaut $-2z/R$, soit $-3,086 \times 10^{-7}$ par mètre à faible altitude. De plus, les variations de densité du sous-sol entraînent des variations locales de la gravité, si bien qu'à altitude constante les pendules n'indiquent pas la même heure partout sur la Terre.

Enfin les forces de marée, dues à l'influence combinée de la Lune et du Soleil, agissent sur g et la correction correspondante varie au cours de la journée ; elle est de l'ordre de 2×10^{-7} à la latitude de 45° . En pratique, ces variations de g sont négligeables pour les pendules, car la précision des meilleures pendules ne dépasse pas 1 seconde sur un jour, soit une variation relative de 10^{-5} , très grande par rapport à des variations relatives de g l'ordre de 10^{-7} . L'effet principal est la dépendance de l'accélération de la pesanteur en fonction de la latitude : g augmente de $0,0052 \text{ ms}^{-2}$, soit environ 10^{-3} en valeur relative, entre l'équateur et le pôle, et cette variation n'est pas négligeable. Nous y reviendrons dans l'encadré 5.2. Pour clore cette discussion des erreurs affectant les horloges à pendule, on doit mentionner l'influence possible des vibrations mécaniques et celle des champs électromagnétiques : une bonne horloge à pendule doit comporter un socle très stable et être blindée contre les fluctuations des champs extérieurs.

On attribue à Galilée la paternité de l'idée de l'horloge à pendule, après qu'il eut constaté que la fréquence d'oscillation d'un pendule ne dépendait pas du poids qui pend à l'extrémité de la tige de l'oscillateur. Galilée est mort en 1642 avant d'avoir eu le temps de construire son horloge, et c'est à Christian Huygens que revient la réalisation de la première horloge à pendule en 1656 : sa stabilité était de 10 secondes par jour, un progrès spectaculaire par rapport aux horloges mécaniques précédentes. Pour l'oscillateur interne de l'instrument, une alternative au pendule est un ressort sous la forme d'une spirale métallique vibrante. C'est sur une idée de Robert Hooke, son collègue initiateur avec lui de la théorie ondulatoire de la lumière, que Christian Huygens construisit la première horloge au mouvement régulé par un ressort. Notons que c'est sur ce principe qu'ont fonctionné pendant deux siècles les montres-bracelets, celles que nous devons remonter manuellement il n'y a pas si longtemps.

Le problème de la stabilité des horloges a été capital pour la navigation pendant des siècles. Il est évidemment essentiel pour un navire de connaître sa position le plus précisément possible. Or, si la mesure de la latitude peut être faite aisément par l'observation des astres (sauf si la mer est trop agitée !), celle de la longitude requiert un repérage du temps à bord : le passage du Soleil au zénith définit midi au point où se trouve le navire, et la comparaison avec l'heure du méridien de Greenwich par exemple permet d'en déduire la longitude. Mais cette comparaison exige que l'on dispose à bord d'un chronomètre synchronisé sur l'heure de Greenwich et qui conserve cette heure. Un bon chronomètre est aussi indispensable pour la navigation à l'estime où l'on suit en permanence la route du navire en mesurant à chaque instant sa vitesse et son cap. Depuis les grandes explorations comme celles de Christophe Colomb, le problème de la longitude a préoccupé les marines de tous les pays. Le gouvernement britannique lança en 1714 un concours doté d'un prix de 20 000 livres pour la

fabrication du meilleur chronomètre de marine. Le prix fut finalement remporté par John Harrison, horloger, ingénieur et astronome, au bout de quarante ans d'efforts face à une concurrence féroce. Il a construit pas moins de quatre prototypes qu'il a sans cesse améliorés. Le chronomètre de Harrison ne fonctionne pas sur le principe du pendule, trop sensible aux mouvements du bateau. Il utilise des ressorts qui s'enroulent et se déroulent. Sa précision est d'un tiers de seconde par jour. Aujourd'hui le positionnement des bateaux est obtenu grâce au GPS, dont nous verrons le principe aux chapitres 5 et 6.

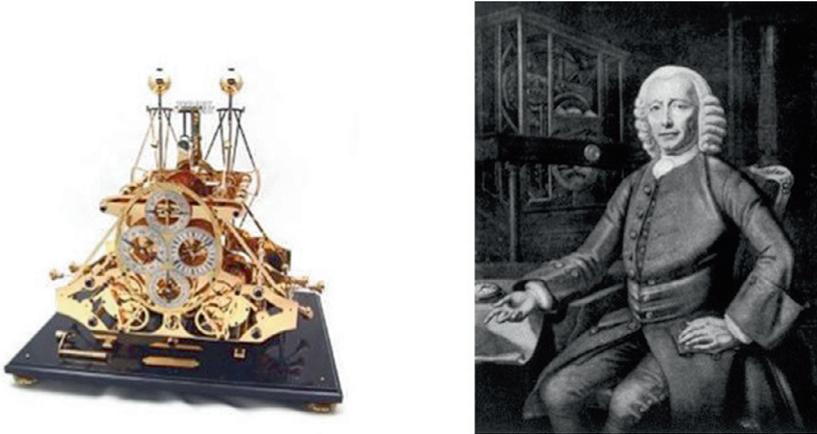


FIGURE 3.8. L'horloger anglais John Harrison (1693-1776) a perfectionné sans cesse la précision de la mesure du temps. Il est l'inventeur du chronomètre de marine et fut récompensé par le roi George III.

Les horloges exploitant un phénomène périodique comportent trois éléments :

1. un résonateur décrit plus en détail ci-dessous (par exemple un pendule ou un ressort) ;
2. une source d'énergie qui entretient le mouvement (par exemple un poids que l'on remonte, une manivelle que l'on tourne, etc.) ;
3. un mécanisme capable de compter les périodes de l'oscillation.

Une horloge idéale serait celle où l'oscillation une fois déclenchée se prolongerait indéfiniment, ce qui est impossible quel que soit le mécanisme en jeu, car tout mécanisme comporte toujours des pertes. Une caractéristique importante de l'horloge est le *facteur de qualité* Q de son oscillateur, c'est-à-dire très grossièrement le nombre d'oscillations qui se produisent une fois lancées avant d'être complètement amorties en l'absence de toute excitation extérieure. Les horloges mécaniques de bonne qualité ont des facteurs Q de quelques milliers.

Des progrès considérables ont été réalisés avec l'arrivée des oscillateurs électriques, puis atomiques, pour lesquels les valeurs de Q sont beaucoup plus élevées.

3.3.2 Les oscillateurs électriques

Après la découverte de l'effet piézoélectrique par Pierre Curie et son frère Jacques en 1880 et d'un mécanisme capable de compter les périodes de l'oscillation, les oscillateurs à quartz remplacèrent les oscillateurs mécaniques dans la plupart des horloges et des montres. Le mécanisme de l'effet piézoélectrique est le suivant : lorsque l'on applique une déformation mécanique à un cristal de quartz, on génère des charges électriques positives et négatives sur les bords du cristal, ce qui induit un champ électrique. Ce mécanisme est réversible car, inversement, lorsqu'on applique un champ électrique on produit une déformation du cristal.

Le principe des *horloges à quartz* consiste à exciter un diapason en quartz avec une tension électrique. Celui-ci vibre alors à une fréquence caractéristique qui dépend des dimensions du cristal. En quelque sorte, on peut considérer que l'horloge à quartz est un autre type d'horloge mécanique. Inversement, quand le cristal vibre, il induit un courant électrique qui oscille à la fréquence de vibration caractéristique du cristal et sert de signal de référence pour l'horloge. Les deux phénomènes se produisent simultanément dans l'horloge. Le système fonctionne sur le mode rétro-action (*feedback*) ; il auto-régule lui-même la stabilité de la fréquence produite, constamment égale ou très proche de la fréquence de résonance du cristal. Un circuit intégré entretient les vibrations du cristal. Les vibrations électriques générées sont mises en forme électroniquement et alimentent un moteur synchrone qui fait tourner les aiguilles de l'horloge (figure 3.9). La friction interne du cristal est si faible que son facteur de qualité Q est très élevé, allant de 100 000 à 2 millions. La fréquence de résonance dépend des dimensions des cristaux et de la façon dont ils sont taillés. Plus leur dimension est petite, plus la fréquence est élevée, typiquement quelques millions de cycles par seconde pour 1 mm d'épaisseur. Ce système est utilisé non seulement pour les horloges mais aussi dans la plupart des montres-bracelets avec de tous petits cristaux et une extrême miniaturisation de la circuiterie électronique (figure 3.10).

Plusieurs facteurs contribuent à la dérive de la stabilité des horloges à quartz. Le facteur principal est la dépendance en température T de la fréquence ν , donnée par :

$$\nu = \nu_0 [1 - 4 \times 10^{-8} (T - T_0)^2] . \quad (3.2)$$

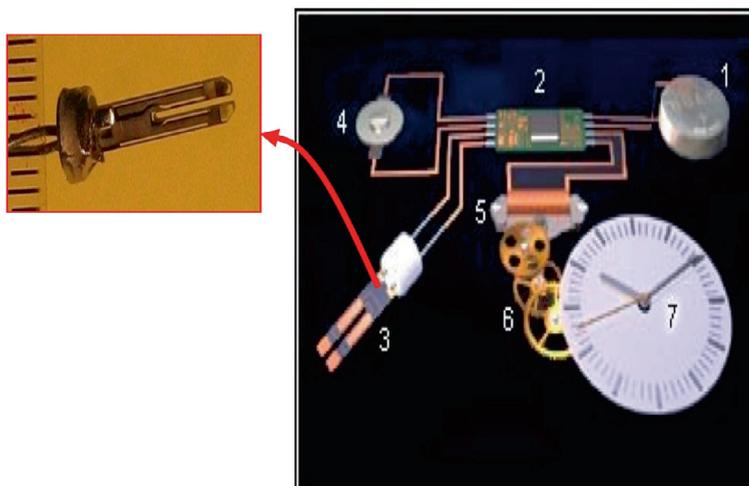


FIGURE 3.9. Principe de l'horloge à quartz, fondée sur l'effet piézoélectrique. Un diapason en quartz (en zoom sur la figure de gauche, en 3 sur la figure de droite) vibre sous l'effet d'une excitation électrique. En 1 la pile d'alimentation, en 2 et 4 la circuiterie électronique qui met en forme et contrôle la fréquence du signal électrique émis, en 5 le moteur qui alimente la mécanique en 6, laquelle fait tourner les aiguilles en 7.

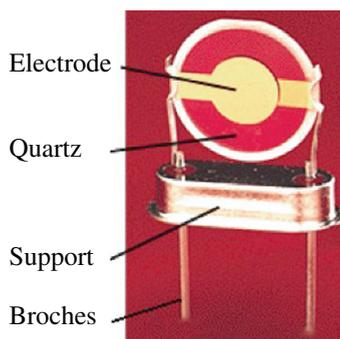


FIGURE 3.10. Principe d'un oscillateur à quartz de type commercial (figure de droite). L'excitation électrique est apportée par les électrodes de part et d'autre du disque cristallin. La tension électrique d'oscillation apparaît sur les broches. Ces oscillateurs sont utilisés dans les horloges, dans les montres, mais aussi dans beaucoup d'autres instruments de la vie courante.

Un écart de 10 degrés en température se traduit par une dérive de la mesure du temps de 2 minutes sur un an. Des instabilités s'introduisent également avec le vieillissement du matériau, typiquement sa contamination progressive par des impuretés, ou encore son usure liée aux vibrations mécaniques. Des horloges très perfectionnées à base d'oscillateurs à quartz ont été construites avec des

montages à ponts doubles ou triples placés dans des enceintes étanches. Elles atteignent une précision record de 1 milliseconde par mois. Les quartz servent encore dans presque toutes les montres-bracelets dont le marché mondial se chiffre en dizaines de milliards de dollars. Ils sont aussi un élément important des horloges atomiques plus précises décrites ci-dessous.

3.3.3 Les oscillateurs atomiques

Un progrès décisif pour les horloges a été effectué quand on a eu l'idée d'utiliser la fréquence d'une *transition atomique* comme oscillateur, cette fréquence étant définie beaucoup plus précisément que celle de la vibration d'un cristal de quartz. Les progrès des horloges ont ainsi suivi le développement de la mécanique quantique tout au long du XX^e siècle. Rappelons par exemple les avancées autour des années 1910 du physicien britannique Ernest Rutherford et du danois Niels Bohr, qui avaient collaboré pour comprendre la nature des atomes. Rutherford avait montré, en bombardant des atomes par des particules alpha issues de matériaux radioactifs, que les atomes étaient constitués d'un cœur central, ou noyau, autour duquel gravitaient les électrons comme des planètes autour du Soleil. Mais il restait à comprendre pourquoi les électrons ne perdaient pas leur énergie progressivement sur leurs orbites jusqu'à tomber sur le noyau. Bohr émit l'hypothèse révolutionnaire que les électrons se mouvaient dans l'atome autour du noyau sur des orbites d'énergie bien définie, qu'ils pouvaient quitter en absorbant ou émettant du rayonnement électromagnétique de fréquence proportionnelle à la différence des énergies entre les orbites initiale et finale. Dans la conception d'Einstein, cette relation correspond à l'émission ou à l'absorption de quanta lumineux ou *photons* (suivant le terme introduit par Gilbert Lewis en 1926), d'énergie proportionnelle à la fréquence de la lumière. Les atomes possèdent des niveaux discrets dont les énergies sont quantifiées. Ainsi un atome peut faire des « sauts » d'un *niveau d'énergie* E_1 à un niveau d'énergie E_2 en émettant ou en absorbant un photon dont la fréquence ν et l'énergie $h\nu$ sont des caractéristiques de l'atome (figure 3.11). Plus précisément, la loi de conservation de l'énergie s'écrit sous la forme de la *relation de Planck-Einstein* :

$$E_2 - E_1 = h\nu, \quad (3.3)$$

où h est la constante de Planck, une constante fondamentale de la physique ($h = 6,63 \times 10^{-34}$ J.s).

Le progrès apporté par l'utilisation d'une fréquence atomique comme référence pour fabriquer une horloge est devenu évident dès le milieu du XX^e siècle. En effet, une telle fréquence présente des avantages considérables : elle est bien définie, elle est la même partout sur la Terre, ou plus exactement sur l'ensemble

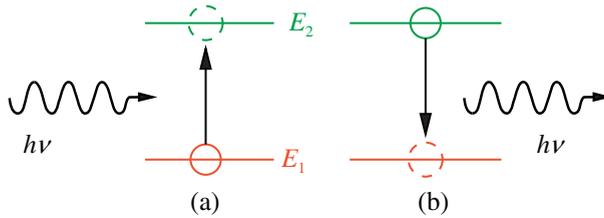


FIGURE 3.11. Émission et absorption d'un photon entre deux niveaux d'énergie d'un atome. (a) L'atome, initialement dans le niveau E_1 , absorbe un photon d'énergie $h\nu$ et se retrouve dans le niveau E_2 : $h\nu = E_2 - E_1$. (b) L'atome, initialement dans le niveau E_2 , émet un photon d'énergie $h\nu$ et se retrouve dans le niveau E_1 .

du géoïde (encadré 5.2), elle ne dépend ni de la température ni de l'environnement en l'absence de champs magnétiques ou électriques. Elle peut servir de résonateur pour l'horloge, avec une *fréquence de résonance* qui est une propriété de l'atome lui-même. Le facteur de qualité Q est de l'ordre de 100 millions, 100 fois plus élevé que celui des quartz. Le principe de *l'horloge atomique* consiste à utiliser la fréquence de la transition atomique comme référence sur laquelle est asservie électroniquement la fréquence d'un oscillateur, par exemple à quartz. Ce principe est illustré sur la figure 3.12. Dans ce dispositif, l'oscillateur à quartz a une double fonction : il sert d'une part à induire la transition atomique entre les niveaux E_1 et E_2 car il génère une fréquence très proche de la fréquence atomique, qui d'autre part est transmise pour la lecture de l'horloge. En résumé, grâce à un mécanisme d'asservissement, l'oscillation du quartz est maintenue à la fréquence de résonance de la transition atomique, ce qui est illustré sur la figure 3.12.

3.3.4 L'horloge à jet de césium

C'est en 1949 que le National Bureau of Standards (NBS) annonça qu'il avait réussi à faire fonctionner une horloge fondée sur une transition atomique. En fait, il ne s'agissait pas vraiment d'un atome mais plutôt d'une molécule, l'ammoniac, qui contient un atome d'azote et trois atomes d'hydrogène. La fréquence naturelle utilisée était de 23 970 MHz (mégahertz) située dans le domaine micro-onde et correspondant à une transition de type *sis-trans*, c'est-à-dire entre deux configurations géométriques de cette molécule. Cette horloge très innovante s'est pourtant révélée assez peu fiable pour plusieurs raisons. La première est que la vibration résultant des propriétés de plusieurs atomes était moins stable intrinsèquement que celle d'un atome unique. En outre les collisions des molécules sur les parois des récipients modifiaient considérablement la fréquence de

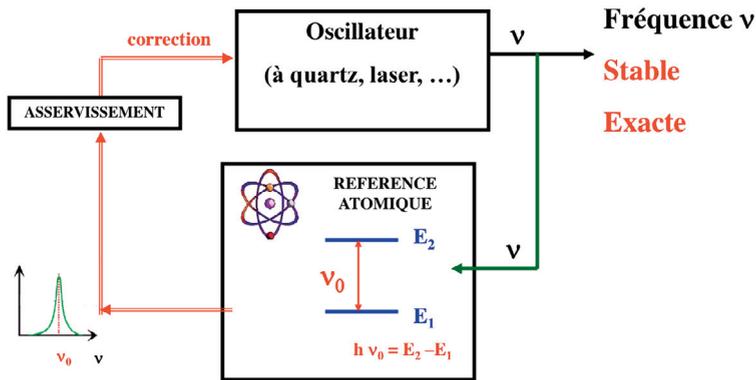


FIGURE 3.12. Principe d'une horloge atomique. On induit à l'aide d'un oscillateur à quartz la transition entre deux niveaux d'énergie d'un atome. La fréquence atomique ainsi générée sert de référence pour asservir en retour la fréquence de l'oscillateur à quartz. Elle est beaucoup plus stable que celle du quartz.

résonance d'une façon non contrôlée. Enfin la fréquence était décalée par l'effet Doppler, résultant du passage des atomes en mouvement désordonné dans le champ d'interrogation micro-onde : l'effet Doppler est la modification de la fréquence lorsque l'objet émetteur ou l'observateur sont en mouvement. Cet effet est familier pour les ondes sonores : la fréquence de la sirène d'une ambulance est perçue comme plus aiguë lorsque l'ambulance s'approche, et plus basse quand elle s'éloigne.

La plupart de ces défauts ont été largement atténués lorsque l'ammoniac a été remplacé par un atome. La première horloge à césium a été construite par Louis Essen et Jack Parry en 1955 au National Physical Laboratory (NPL) en Grande-Bretagne. Le césium a été privilégié d'abord pour des raisons pratiques, car il est facilement manipulable. Il se présente sous une forme métallique à la température ordinaire et fond à la température accessible de 28° C. C'est un élément alcalin, c'est-à-dire qu'il possède un électron périphérique gravitant autour d'un cœur constitué du noyau et de son cortège d'électrons formant des couches fermées. Cet électron sur une couche externe de l'atome peut osciller entre deux niveaux E_1 et E_2 , qui sont par choix pour l'horloge les deux niveaux hyperfins de l'état fondamental de l'atome. La dégénérescence de cet état fondamental en deux sous-niveaux, appelée dégénérescence hyperfine, est liée à l'interaction magnétique de l'électron périphérique avec le moment magnétique (aimant) porté par le noyau. Ces deux niveaux sont séparés comme pour l'ammoniac par une énergie située dans le domaine micro-onde ; elle est très bien définie et égale à 9 192 631 770 Hz.



FIGURE 3.13. Norman Ramsey (1915-2011), professeur de physique à Harvard qui reçut le prix Nobel de physique en 1989 pour la méthode dite des champs alternatifs séparés et ses applications au maser à hydrogène et aux horloges atomiques.

L'horloge est constituée d'un four électrique qui émet à travers un petit orifice un jet de césium gazeux dans une enceinte à vide. Les atomes sont d'abord préparés dans l'état de basse énergie E_1 à l'aide d'un tri magnétique qui élimine les atomes dans l'état E_2 , ceux-ci étant défléchis hors du jet. Le jet se propage ensuite dans un tube. Il traverse une cavité micro-onde excitée à une fréquence ν voisine de la fréquence de résonance atomique ν_0 où les ondes radio sont envoyées perpendiculairement au jet pour éviter l'effet Doppler (figure 3.14). Les atomes parcourent alors une certaine distance en ligne droite sans subir beaucoup de collisions avec les parois, évitant ainsi les problèmes des précédentes horloges à ammoniac. Plus ν est proche de ν_0 et plus nombreux sont les atomes qui absorbent le rayonnement micro-onde et passent de l'état de basse énergie E_1 à l'état de haute énergie E_2 . À la sortie, les atomes dans E_1 et E_2 sont à nouveau séparés magnétiquement et l'on détecte alors uniquement les atomes situés dans l'état E_2 . Plus le nombre des atomes détectés est grand et plus l'on est proche de la résonance $\nu = \nu_0$. Si l'on balaie la fréquence ν , on observe un signal de résonance, une courbe en cloche dont la largeur varie comme $1/T$, où T est la durée de la traversée de la cavité micro-onde par le jet. Un système d'asservissement ajuste en permanence la fréquence de l'oscillateur à quartz afin de maximiser le nombre d'atomes dans l'état E_2 . La fréquence de l'oscillateur est ainsi asservie

à la fréquence de la transition atomique. Plus la longueur de la cavité traversée est grande, plus $1/T$ et la largeur de la résonance sont petits, et donc plus la mesure est précise. L'encombrement limite la précision : pour des atomes de césium se déplaçant à une vitesse de 100 m.s^{-1} dans un tube de 1 mètre de long, $T = 0,01 \text{ s}$, ce qui correspond à une largeur de raie, ou encore à une dispersion de fréquence, de 100 Hz. Le facteur de qualité Q , égal au rapport entre la fréquence de résonance ν_0 et la dispersion de fréquence, est ainsi de l'ordre de 100 millions, une valeur beaucoup plus élevée que celle de l'oscillateur à quartz.

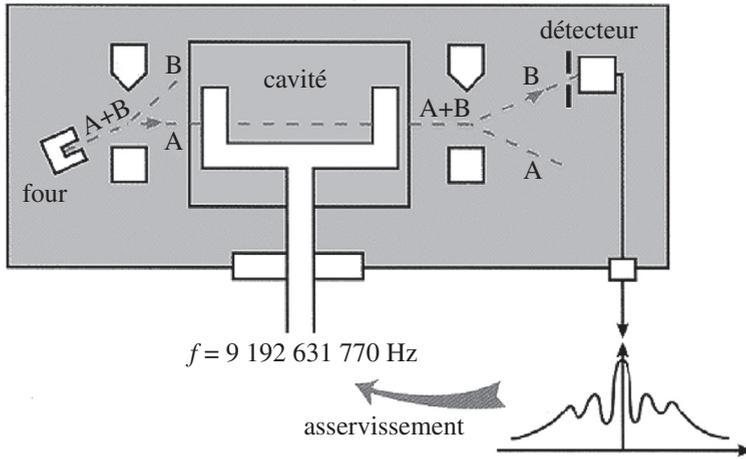


FIGURE 3.14. Principe de l'horloge à jet de césium. Le jet d'atomes passe deux fois à travers la cavité radiofréquence dite de Ramsey. Les états de l'atome A et B sont sélectionnés avant et après la sortie de la cavité. Le détecteur enregistre un signal d'interférence dont le pic central est d'autant plus étroit que le temps de transit des atomes est grand entre les deux zones d'interrogation radio-fréquence dans la cavité.

De nombreuses améliorations ont été apportées au principe de base de l'horloge à césium. L'une d'entre elles a consisté à remplacer le tri magnétique des états atomiques E_1 et E_2 par un procédé optique qui permet de mettre tous les atomes dans l'état E_1 à l'entrée de l'horloge sans en perdre une partie par élimination. La méthode s'appelle le pompage optique : elle consiste à jouer sur les états internes de spin des atomes à l'aide de faisceaux de lumière polarisée. Le pompage optique a fait considérablement progresser la précision des horloges atomiques dans les années 90.

Une autre méthode très puissante pour améliorer l'exactitude de l'horloge sans augmenter sa taille est due à Norman Ramsey, physicien qui reçut le prix Nobel en 1989 (figure 3.13). La méthode dite des *franges de Ramsey* consiste à utiliser deux zones d'interrogation radiofréquence séparées par une certaine distance L que le jet traverse successivement. On synchronise lors du passage des

atomes dans la première zone d'interrogation l'oscillateur local et l'oscillateur de référence (le système atomique). À la seconde interrogation, on mesure le déphasage de l'oscillateur local par rapport à l'oscillateur atomique. La double interrogation produit des interférences. La largeur du pic central est d'autant plus fine que le temps de vol $T = L/v$, où v est la vitesse des atomes entre les deux zones d'interaction, est plus grand (figure 3.14).

Il existe de nombreuses variantes des horloges à césium, de toutes les tailles et de toutes les formes. La figure 3.15 montre une horloge commerciale typique comme celles que l'on achète pour des usages courants de métrologie dans l'industrie ou dans le commerce. Ces horloges ont besoin d'être périodiquement recalées par rapport aux standards de temps fournis par les horloges primaires, comme nous l'expliquerons au chapitre 6.



FIGURE 3.15. Une horloge à césium standard de marque Hewlett Packard 5071A. La longueur de la cavité est de 30 cm, l'exactitude relative de 2×10^{-13} .

3.4 Les horloges à atomes froids

La précision des horloges à jets de césium est limitée par la taille de la cavité qui, pour des raisons pratiques, ne peut excéder une longueur de l'ordre du mètre. Elle est aussi limitée par les collisions des atomes entre eux et par le déplacement Doppler de la fréquence de résonance au moment de l'interrogation par la radiofréquence. On a rapidement compris dans les années 1980, au début du développement du refroidissement des atomes par laser, qu'il serait très intéressant que les horloges atomiques utilisent des atomes refroidis par laser

plutôt que des atomes à la température ordinaire. On pourrait ainsi allonger le temps d'interrogation, et donc la précision, tout en limitant les collisions et autres sources d'erreurs dues à des effets parasites comme l'effet Doppler.

3.4.1 Principe du refroidissement laser et du piégeage des atomes

Le domaine des atomes refroidis par laser s'est ouvert dans les années 1980. Il est toujours en pleine expansion et donne lieu à des découvertes spectaculaires en physique fondamentale comme en instrumentation. Les applications aujourd'hui les plus importantes concernent les horloges. Rappelons ici le principe du refroidissement d'un gaz par un faisceau laser. Il est fondé sur le cycle d'absorption/émission d'un photon par un atome qui passe d'un niveau E_1 à un niveau E_2 . Lors de l'absorption, l'atome encaisse l'impulsion \vec{k} portée par le photon et recule dans la direction opposée à celle du faisceau lumineux. Il retombe ensuite dans l'état E_1 en émettant un photon dans une direction aléatoire. Après un grand nombre de cycles de cette nature, le freinage de l'atome est considérable, l'atome est ralenti et sa vitesse peut ainsi descendre à quelques cm.s^{-1} , contre quelques centaines de m.s^{-1} à la température ordinaire. Pour comprendre le *refroidissement laser* du gaz, on peut se reporter à la figure 3.16 qui en illustre le principe dans un schéma simple à une dimension. La vitesse moyenne des atomes décroît, ainsi que la largeur de la distribution des vitesses qui caractérise la température. Une force de freinage résulte de l'application d'un faisceau laser légèrement décalé en fréquence par rapport à la résonance atomique. On atteint ainsi des températures situées dans le domaine du μK (microkelvin). Notons que des températures encore beaucoup plus basses peuvent être obtenues par

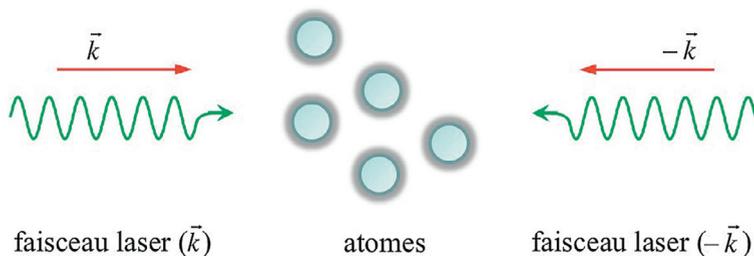


FIGURE 3.16. Principe du refroidissement laser par effet Doppler. Les atomes sont placés entre deux faisceaux laser décalés en fréquence par rapport à la résonance. A cause de l'effet Doppler, si un atome se déplace vers la droite il absorbe préférentiellement le faisceau qui vient de droite. De même s'il se déplace vers la gauche il absorbe le faisceau qui vient de gauche. Il en résulte une force de friction qui freine les atomes et les confine au centre de la figure.

des méthodes plus élaborées, mais elles ne sont pas pour l'instant employées pour les horloges.

La force de friction créée par le laser peut servir non seulement à abaisser la température mais aussi à piéger le nuage atomique confiné dans une enceinte à ultravide. En effet, si l'on applique un ensemble de six faisceaux lumineux dans les trois directions d'espace (figure 3.17), on peut confiner pour un laps de temps bref les atomes au point d'intersection des faisceaux : on parle de mélasse optique, les atomes étant comme englués dans la lumière. Les mélasses ne sont toutefois pas opérationnelles pour maintenir le piège pendant une longue durée.

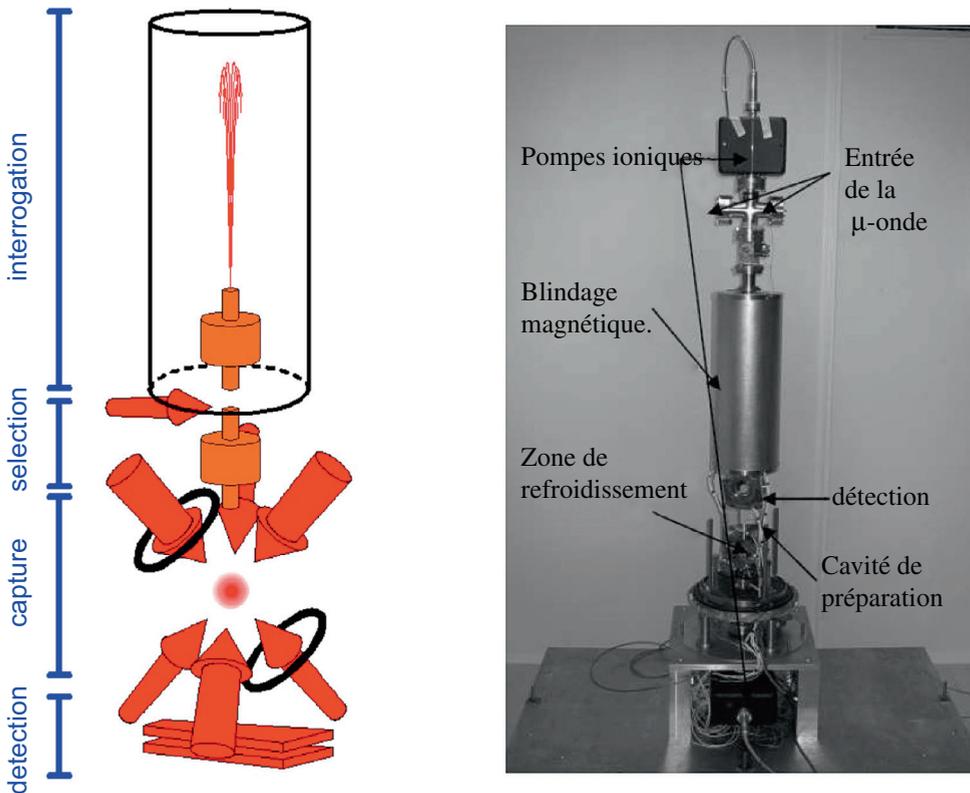


FIGURE 3.17. Le principe d'une horloge à fontaine d'atomes froids. Un jet d'atomes est envoyé verticalement à partir d'une zone de capture comprenant six faisceaux laser et des bobines de gradient de champ magnétique. Les atomes retombent et passent deux fois (à la montée et à la descente) à travers la cavité d'interrogation micro-onde. La détection s'opère sur la fluorescence des atomes qui retombent excités par un faisceau laser sonde. La température dans le piège est de l'ordre de $1 \mu\text{K}$, la vitesse dans la fontaine de $4 \text{ m}\cdot\text{s}^{-1}$. Le temps séparant les deux passages ascendant et descendant à travers la cavité micro-onde est de $0,5 \text{ s}$.

Un perfectionnement du piège à mélasse a été inventé dans les années 1980 : il consiste à rajouter aux six faisceaux lumineux un gradient de champ magnétique et à polariser circulairement les faisceaux. Faisant alors jouer le principe du pompage optique qui manipule les degrés de liberté internes des atomes, on montre qu'on crée ainsi une force de rappel sur les atomes. Un tel piège, dit *piège magnéto-optique* (ou MOT : *Magneto-Optic Trap*), est devenu l'outil de base de la physique atomique. Il permet de maintenir en lévitation dans l'ultravide un nuage froid typiquement de plusieurs millions d'atomes dans un volume dont les dimensions ne dépassent pas quelques millimètres à une température de l'ordre de quelques μK .

3.4.2 Les fontaines à atomes froids

Le piège magnéto-optique est le point de départ des horloges à atomes refroidis. Il n'est toutefois pas possible de construire l'horloge autour du piège, d'une part parce que les faisceaux laser du piège sont responsables d'un décalage de la fréquence atomique utilisée comme référence, d'autre part parce que la densité des atomes n'est pas stable à cause des pertes du piège. L'idée de la fontaine remédie à ces inconvénients : elle consiste à extraire des atomes froids du piège magnéto-optique en les poussant vers le haut verticalement par un coup de laser. Les atomes retombent ensuite vers le bas sous l'effet de la pesanteur, comme le jet d'eau d'une fontaine (figure 3.17). Ils traversent deux fois la zone d'interrogation radiofréquence, ce qui est bien conforme à la méthode interférométrique de Ramsey utilisée dans les horloges atomiques à jet à température ordinaire. L'avantage ici est que le temps de passage entre les deux zones d'interrogation est environ 100 fois plus long, et donc la précision de l'horloge est augmentée dans la même proportion, comme expliqué plus haut. La détection s'effectue sur la fluorescence des atomes que l'on excite avec un faisceau laser résonant après leur retombée.

La première horloge à fontaine de césium a été construite en France en 1995 au Laboratoire primaire temps fréquence (LPTF). Ce laboratoire a ensuite fusionné avec l'Horloge Atomique (LNHA) au sein du LNE-SYRTE à l'Observatoire de Paris, où l'horloge à fontaine a été développée sous la responsabilité d'André Clairon. La précision des horloges de ce type dépend de leur taille, qui limite la hauteur du jet de la fontaine. La figure 3.18 montre le signal obtenu. La fréquence de référence est fixée sur le pic central très fin de la figure d'interférences. Beaucoup de soin est apporté pour limiter les erreurs systématiques qui peuvent affecter l'exactitude de ces mesures. De très nombreux facteurs sont pris en considération. Ainsi, pour minimiser les déplacements de raie par l'effet Zeeman, l'ensemble du jet est placé dans un encombrant

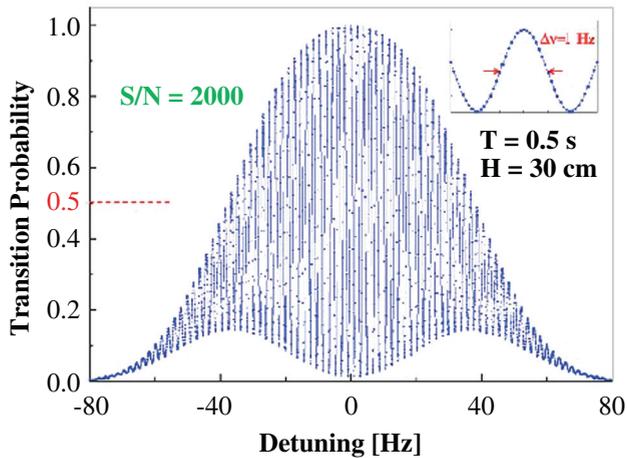
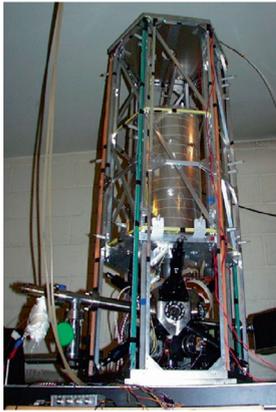


FIGURE 3.18. Le signal de détection est une figure d'interférence résultant de la double interrogation selon la méthode des franges de Ramsey obtenue avec une horloge à fontaine de césium à l'Observatoire de Paris. La fréquence est asservie sur le pic central dont la largeur est de l'ordre du hertz.

blindage magnétique consistant en plusieurs couches concentriques de μ -métal. Une des principales causes d'erreur est le *rayonnement du corps noir* : on sait en effet qu'un corps chauffé émet un rayonnement électromagnétique dont la fréquence moyenne est proportionnelle à la température. Pour éviter l'erreur que cela entraîne sur la fréquence atomique mesurée, il faudrait que tout l'ensemble de l'appareil soit parfaitement à la même température, ce qui est impossible en pratique. On se contente de climatiser le mieux possible la pièce dans laquelle est située l'horloge. Une autre erreur difficilement évitable résulte des collisions entre les atomes au cours de leur trajet dans la fontaine. Comme on ne peut pas trop diminuer la densité du jet sous peine de perdre en signal, on évalue cette erreur collisionnelle en faisant varier le flux dans la fontaine et en extrapolant à zéro les variations de fréquence mesurées en fonction de ce flux. Le bilan des erreurs systématiques d'une horloge à fontaine est très complexe, il comporte une vingtaine de causes estimées et minimisées, toutes réévaluées en permanence. On aboutit aujourd'hui pour les meilleures horloges à fontaine de césium à une exactitude relative de 10^{-15} . Les spécialistes aiment rendre ce chiffre plus concret en parlant d'une dérive de 1 seconde sur une durée de 30 millions d'années.

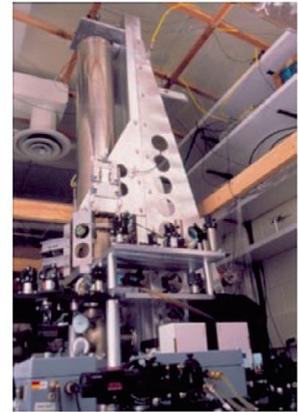
Tous les pays avancés construisent des horloges à fontaines d'atomes de césium 133 qui leur permettent de construire leur échelle de temps de référence. Toutes ces horloges de référence nationales sont reliées entre elles pour la définition du temps universel (figure 3.19).



BNM-SYRTE, FR



PTB, D



NIST, USA

FIGURE 3.19. À gauche, l'horloge à fontaine de césium du Bureau national de métrologie (BNM) à l'Observatoire de Paris qui sert de référence pour la France. Sa hauteur totale est de l'ordre de 3 m, celle du jet de 1 m. On aperçoit le blindage métallique qui protège le jet atomique des fluctuations du champ magnétique extérieur. A droite, des horloges du même type construites dans d'autres pays : au PTB à Braunschweig pour l'Allemagne, au National Institute of Science and Technology (NIST) à Boulder aux États-Unis.

3.4.3 Les horloges optiques à atomes froids

Pour aller plus loin en stabilité que les horloges à fontaine atomique, il fallait trouver une nouvelle idée, tout en gardant celle de choisir comme référence une transition entre deux niveaux d'un même atome. Passer dans le domaine optique en quittant les radiofréquences paraissait tentant, car en gagnant 3 ou 4 ordres de grandeur sur la fréquence de référence, on devait en principe augmenter l'exactitude de l'horloge dans les mêmes proportions. Cette perspective est restée longtemps inaccessible, mais les problèmes liés à cette nouvelle technologie ont été résolus il y a moins de dix ans.

La première difficulté est de mesurer directement une fréquence située dans une gamme très élevée – typiquement 10^{14} Hz –, pour laquelle il n'existe pas de détecteur électronique assez rapide. La solution est venue des systèmes dits à peigne de fréquences qui ont valu le prix Nobel à Theodor Hansch et John Hall en 2005. Ces appareils très versatiles, maintenant utilisés à grande échelle en métrologie, fournissent l'équivalent d'une règle graduée de fréquences très bien étalonnées issues d'un laser femtoseconde. On fait battre la fréquence du laser à mesurer avec la dent la plus proche du peigne. Le battement tombe dans le domaine des radiofréquences et il est mesuré de façon précise avec les détecteurs électroniques usuels.

La seconde difficulté tient au fait que la mesure doit porter sur des atomes directement refroidis dans un piège et non plus isolés dans le jet d'une fontaine. Pour les immobiliser, l'idée retenue a été de les figer dans un réseau optique. Les réseaux optiques sont des puits de potentiel périodiques créés par des ondes stationnaires résultant de faisceaux laser annexes, assez puissants pour que les atomes se piègent aux ventres de cette onde. Bien que la longueur d'onde des lasers constituant le réseau ne soit pas résonante avec celle des atomes étudiés, elle crée cependant un déplacement de la fréquence à mesurer. Une solution élégante à ce problème est due au physicien japonais Hidetoshi Katori qui a montré que, pour certains atomes comme le strontium, on peut trouver une longueur d'onde dite magique telle que le déplacement lumineux s'annule : la transition d'horloge mesurée est alors exactement celle que fournirait une assemblée d'atomes immobiles sans piège. Le principe de telles horloges optiques est représenté sur la figure 3.20.

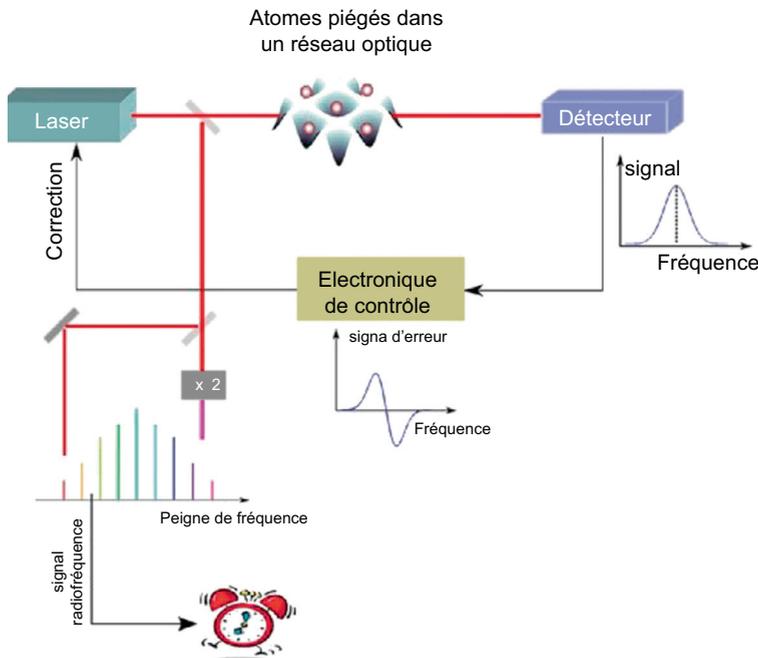


FIGURE 3.20. Principe de l'horloge optique à atomes froids. Le signal de référence est la fréquence de transition entre deux niveaux d'un atome, située dans le domaine optique et non plus radiofréquence comme pour les horloges à césium. Ce signal contrôle la fréquence d'un laser et non plus celle d'un oscillateur à quartz. Les atomes sont piégés dans un réseau optique créé par une onde lumineuse stationnaire. La fréquence optique du laser contrôlée par celle de l'atome est lue avec un dispositif dit à peigne de fréquences. Le signal de battement tombe dans le domaine radiofréquence.

Il existe plusieurs espèces atomiques pour lesquelles on a pu déterminer une longueur d'onde magique, par exemple le strontium, l'ytterbium ou le mercure. Les prototypes actuellement en développement dans de nombreux laboratoires donnent déjà des exactitudes relatives approchant 10^{-18} qui surpassent de loin celles des horloges à fontaine. Sur le même principe, certains laboratoires développent des *horloges optiques* utilisant des ions et non des atomes neutres. On opère toujours sur une transition entre niveaux d'énergie de la particule, qui ici est chargée puisque c'est un ion, c'est-à-dire un atome à qui il manque un électron. Comme précédemment l'ion est refroidi par laser. La différence avec les horloges à atomes tient essentiellement à la nature des pièges utilisés. On immobilise les ions dans des pièges électromagnétiques dont l'invention a valu le prix Nobel au physicien Wolfgang Paul en 1989. Le confinement est produit par des bobines ou des barres parcourues par un courant alternatif de haute fréquence dans le domaine des radiofréquences. L'avantage des ions sur les atomes neutres est que le nombre de particules utilisées est beaucoup plus réduit – on peut piéger, refroidir et détecter par fluorescence jusqu'à un ion unique. Les erreurs de l'horloge liées aux collisions entre particules sont ainsi considérablement réduites. En revanche, les micro-mouvements des ions liés au piégeage par des champs oscillants ne peuvent pas être évités, ce qui est un inconvénient que ne présentent pas les atomes neutres immobilisés dans les réseaux optiques. Ces horloges utilisent des ions Al^+ , Sr^+ , Yb^+ ou Hg^+ . Les horloges optiques à atomes en réseau et les horloges à ions sont en compétition, elles arrivent aujourd'hui aux mêmes ordres de grandeur d'exactitude relative, proche de 10^{-18} , soit une dérive d'une seconde sur l'âge de l'Univers. Et elles n'ont pas dit leur dernier mot!

3.5 Exactitude, stabilité, performances d'une horloge

3.5.1 Définition et mesure des qualités d'une horloge

On confond souvent *exactitude* et *stabilité* quand on qualifie les performances d'une horloge. Le signal délivré par une horloge est une fréquence. La fréquence mesurée $\nu(t)$ est égale à la fréquence de référence ν_0 de l'oscillateur, qui est ce que l'on cherche, modifiée par des erreurs de deux types : des fluctuations $y(t)$ qui varient dans le temps et des biais ε liés à des erreurs systématiques. On peut écrire :

$$\nu(t) = \nu_0(1 + \varepsilon + y(t)). \quad (3.4)$$

Les fluctuations sont responsables de l'instabilité de l'horloge, les biais de son inexactitude. Une horloge peut être instable ou inexacte ou en général un

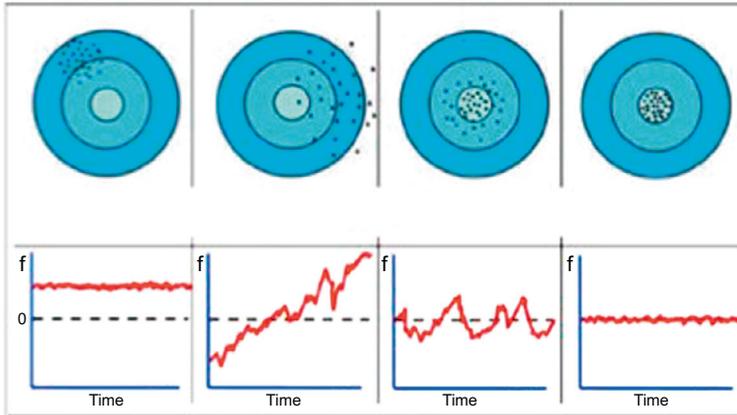


FIGURE 3.21. Illustration des qualités d'exactitude et de stabilité d'une horloge. De gauche à droite : signal d'horloge stable mais inexact, signal ni stable ni exact, signal exact en moyenne mais instable, signal stable et exact.

peu des deux. Ceci est illustré d'une façon imagée sur la figure 3.21. L'une ou l'autre ou bien l'ensemble des deux qualités sont nécessaires pour une horloge en fonction de son application. Pour la stabilité, on mesure le bruit global dans la comparaison entre deux horloges : le bruit mesuré est la somme quadratique des bruits de chaque horloge (en puissance de bruit). Si une horloge est bien plus stable que l'autre, on ne voit que le bruit de l'horloge la moins stable. Si les deux horloges sont identiques, on fait l'hypothèse qu'elles ont le même bruit. Pour l'exactitude, on procède différemment selon le type d'horloge considéré. Si c'est un étalon primaire de laboratoire, on évalue l'exactitude en cherchant à mesurer tous les effets systématiques qui dépendent de différents paramètres : soit on mesure un paramètre et on corrige le déplacement de fréquence correspondant (par exemple le déplacement Zeeman dû au champ magnétique) soit, si l'on a du mal à mesurer le paramètre de façon absolue (par exemple la densité atomique qui conditionne le déplacement collisionnel), on fait varier de façon contrôlée la valeur du paramètre, on mesure à chaque fois le déplacement de fréquence et on extrapole à valeur nulle du paramètre. Si c'est un étalon secondaire, on peut en principe suivre la même procédure que pour un étalon primaire mais, de façon plus simple, on compare directement sa fréquence à celle d'un étalon primaire. Cela fournit directement l'écart en fréquence et l'exactitude de l'horloge étudiée.

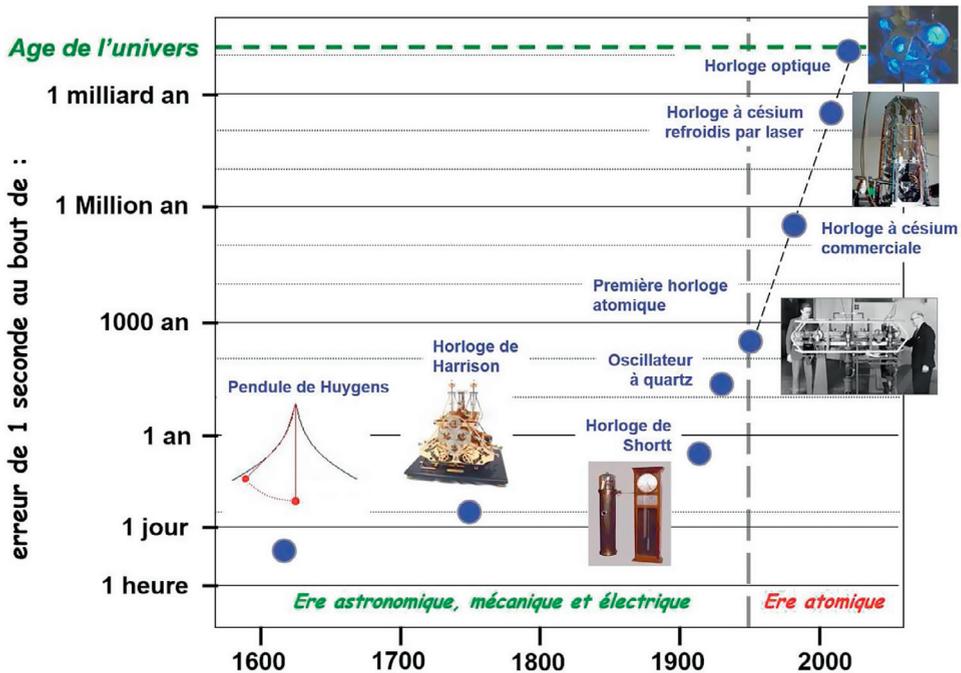


FIGURE 3.22. Évolution de l'exactitude des horloges des années 1600 à nos jours, définie en termes de dérive de 1 s au bout d'un temps variable : quelques heures à l'ère astronomique, quelques jours pour les horloges mécaniques, un an pour les doubles pendules, quelques années pour les oscillateurs à quartz. L'apparition des horloges atomiques marque une rupture et le début de progrès spectaculaires ; le refroidissement laser des atomes marque une nouvelle rupture. Les horloges actuelles dérivent d'une seconde sur l'âge de l'Univers. La figure 3.23 donne un état des lieux plus précis pour les progrès de l'exactitude des horloges au cours du dernier demi-siècle.

3.5.2 Amélioration de l'exactitude des horloges dans l'histoire

Les figures 3.22 et 3.23 résument les progrès obtenus dans l'exactitude des horloges depuis 1600. Depuis l'arrivée du refroidissement par laser des atomes et des ions, la précision des horloges est devenue très impressionnante. Les horloges à atomes neutres dans le domaine optique sont aujourd'hui les plus précises, concurrencées de près par les horloges à ions. Le perfectionnement des technologies utilisées laisse encore prévoir des progrès : une stabilité relative de l'ordre de 10^{-18} devrait être atteinte d'ici peu. Elle correspondra à une dérive de moins d'une seconde sur 30 milliards d'années, supérieure à l'âge de l'Univers (13,8 milliards d'années). Les applications potentielles de cette précision étaient encore inimaginables il y a seulement quelques années. Nous les évoquerons au chapitre 6. Il faut souligner que ces performances extrêmes sont pour l'instant

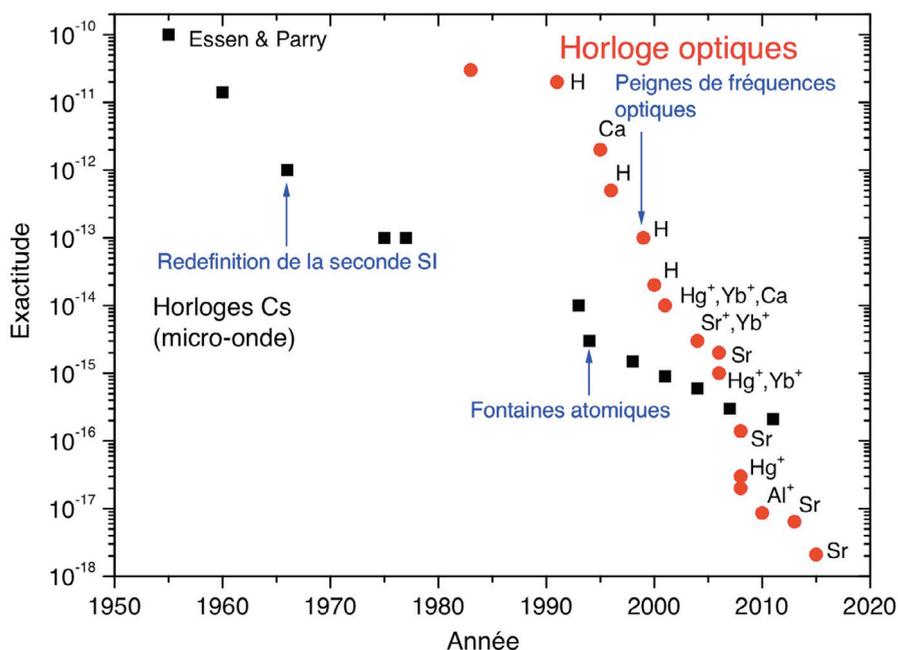


FIGURE 3.23. Évolution de l'exactitude des horloges atomiques depuis 60 ans. Les carrés noirs correspondent aux horloges à césium ; ils indiquent les progrès résultant du refroidissement des atomes dans les fontaines à atomes à partir des années 90. Les points rouges représentent les horloges fonctionnant dans le domaine optique pour différentes espèces d'atomes et d'ions.

obtenues avec des systèmes d'une grande complexité et peu maniables. Mais les progrès dans ces domaines technologiques sont très rapides et les instruments vont se simplifier et devenir plus compacts. Déjà des horloges à atomes froids de rubidium sont fabriquées commercialement par la société Muquans située à Bordeaux ; elles occupent tout compris un volume inférieur à 1 m^3 et ont une stabilité et une exactitude proches de 10^{-15} , sans précédent pour une horloge commerciale et très suffisante pour beaucoup d'applications métrologiques. Les horloges à atomes ou à ions sur puces vont aussi se répandre et conquérir des marchés. L'histoire des horloges est loin d'être achevée.

3.6 Définition de l'unité du temps : la seconde

Jusqu'ici nous avons étudié les différents types d'horloges, dont la précision a augmenté de façon spectaculaire au cours des temps. Les horloges mesurent des durées, elles doivent être calées les unes par rapport aux autres et aujourd'hui sur les horloges de référence les plus précises, elles-mêmes synchronisées

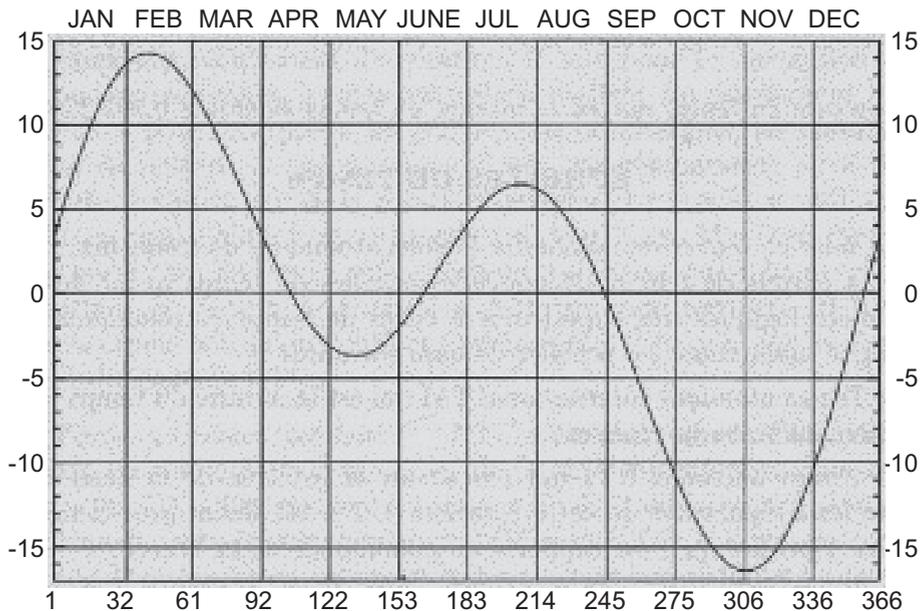


FIGURE 3.24. Fluctuations de la durée du jour au cours d'une année, mesurées en minutes. Les jours sont plus longs en novembre et plus courts en février. La cause principale de la fluctuation est liée à la précession des équinoxes.

entre elles comme nous le verrons ultérieurement. Ce qu'elles ne font pas, c'est fixer la date dans un calendrier et donner un repère en multiple d'une unité de temps. Le système des unités primaires de longueur, de poids et de temps – mais, comme nous le verrons au chapitre 4, l'unité de longueur est aujourd'hui une unité dérivée de celle de temps – a toujours été considéré comme d'une grande importance pour la vie pratique et tout particulièrement pour les relations économiques et commerciales entre pays. Le système est en redéfinition régulière et fixé mondialement par le Bureau international des poids et mesures (BIPM). La seconde est l'unité de temps, l'unité primaire dont la définition est aujourd'hui la plus précise de toutes. L'augmentation spectaculaire de sa précision a suivi les progrès des horloges atomiques, même si certaines communautés continuent à se fonder sur des mesures de nature astronomique, telles que la période des pulsars.

Les mesures d'origine astronomique ont été très longtemps utilisées pour la définition de la seconde. Jusqu'en 1956, la seconde a été définie comme la fraction $1/86\,400$ du jour solaire moyen. Cette définition a toujours posé des problèmes aux calendriers, compte tenu du fait déjà mentionné qu'il n'y a pas un nombre entier de jours dans une année définie par une révolution complète de la



FIGURE 3.25. Variations de la durée du jour au fil du temps sur deux siècles : elle fluctue mais s’allonge globalement.

Terre autour du Soleil. L’adoption du calendrier grégorien au XVI^e siècle, doté de corrections compliquées, a considérablement réduit le décalage du temps officiel par rapport aux saisons. Toutefois, une telle définition de l’unité de temps est intrinsèquement entachée d’imprécision. En effet, la rotation de la Terre autour du Soleil n’est pas totalement régulière. Elle varie principalement parce que l’orbite de la Terre autour du Soleil est elliptique et que l’axe de rotation de la Terre est incliné de 23° par rapport au plan qui contient l’orbite de la Terre autour du Soleil (figure 2.7). L’amplitude relative de la fluctuation est considérable : elle atteint 1%. Pour tenir compte des différences de la durée du jour entre novembre et février, on a dû définir le jour solaire moyen . La figure 3.24 montre les fluctuations autour de ce jour moyen au cours d’une année.

Ce temps solaire moyen peut être comparé à un autre temps : le temps sidéral, défini par la réapparition d’une étoile, qui est de 4 minutes plus court que le temps solaire. Il est aussi nettement plus constant ; la raison en est que les étoiles sont si éloignées de la Terre que l’ellipticité de la trajectoire de la Terre autour du Soleil peut être ignorée. Le temps sidéral est en fait plus facile à mesurer que le temps solaire, mais il n’a pas été utilisé pour la définition de la seconde. Notons que d’autres causes de fluctuations s’ajoutent aux irrégularités principales de la durée du jour déjà mentionnées. L’observation en a été faite par les astronomes royaux britanniques dès le début du XVII^e siècle. La vitesse de rotation de la Terre diminue progressivement sous l’effet de la friction liée aux marées, elles-mêmes dépendant de la Lune, ce qui se traduit par un retard d’environ 16 millisecondes tous les 1 000 ans (figure 3.25). D’autres fluctuations significatives résultent de la position des pôles Nord et Sud qui fluctue de quelques

mètres d'une année sur l'autre, sous l'effet de réarrangements internes dans la structure de la Terre. Enfin toutes sortes d'autres phénomènes sont susceptibles d'affecter ponctuellement la durée du jour tels que la météo, les fluctuations de l'atmosphère, voire les séismes.

Pour toutes ces raisons qui ont comme conséquence l'irrégularité de la rotation de la Terre autour de son axe, les prédictions pour l'apparition de certains phénomènes astronomiques se sont révélées ne pas coïncider exactement avec les observations, par exemple celle des orbites de la Lune ou des planètes. C'est pourquoi, à partir de 1956, une nouvelle définition plus précise et plus fiable de la seconde fut introduite, reliée aux événements astronomiques plutôt qu'à la rotation de la Terre. Le temps devint celui des Éphémérides. Sur un plan pratique, la seconde a été définie alors à partir de la durée de la rotation de la Terre autour du Soleil. La définition est plus précise que la précédente (elle comporte 12 chiffres significatifs). Par convention, la seconde est la fraction $1/315\,569\,259\,747$ de l'année dite « tropicale » 1900. Une année tropicale est égale à 365,2422 jours solaires, ou 366,2422 jours sidéraux. L'avantage est que ce temps éphéméride est uniforme. Mais il est vraiment peu pratique ! En outre il y a toujours des possibilités de fluctuations et des erreurs introduites par les irrégularités dans la rotation de la Terre autour du Soleil. Rien n'est immuable dans le système solaire, on sait même aujourd'hui qu'il est chaotique sur le long terme.

C'est pourquoi cette définition de la seconde, encore astronomique, n'a été employée que pendant une dizaine d'années. En 1967, elle a été définitivement abandonnée au profit d'une définition totalement différente faisant intervenir une donnée intrinsèque, universelle, reliée à une fréquence atomique. La définition actuelle de la seconde est le temps qui s'écoule entre 9 192 631 770 oscillations de l'atome de césium 133 « non perturbé » entre les deux niveaux de la structure hyperfine de son état fondamental. Cette définition est totalement indépendante des mouvements de la Terre. Elle s'est imposée avec le progrès des horloges atomiques. Elle pourrait encore changer prochainement, en ce sens qu'un autre atome pourrait remplacer le césium si l'usage d'horloges plus performantes que celles au césium se généralise. Néanmoins la définition restera atomique. Elle est la même en tout point du globe, ou plus exactement en tout point du géoïde (encadré 5.2). On construit un temps universel pour toute la Terre à partir des mesures faites par un ensemble d'horloges situées en divers points de la Terre. Nous y reviendrons au chapitre 6, quand nous aurons vu comment les horloges peuvent communiquer entre elles et se synchroniser.

4

Le temps de la relativité restreinte

En 1905, l'article fondateur d'Einstein introduisant la *relativité restreinte* bouleverse les conceptions newtoniennes d'espace et de temps. La relativité restreinte permet de traiter de mouvements où les vitesses ne sont pas négligeables par rapport à celle de la lumière, *en l'absence de gravitation*. C'est la théorie utilisée par exemple en physique des particules élémentaires, où les forces en jeu sont électromagnétiques, fortes et faibles, et où la gravitation ne joue aucun rôle. La *relativité générale*, introduite par Einstein en 1915, inclut la gravitation : c'est la théorie utilisée par les astrophysiciens.

4.1 Vitesse limite

Comme chacun sait, l'espace-temps newtonien n'est plus adapté lorsque les vitesses en jeu ne sont pas négligeables par rapport à celle de la lumière. Imaginons un voyageur venu d'une galaxie lointaine qui traverse le système solaire à très grande vitesse, par exemple le dixième de la vitesse de la lumière. Ce voyageur transporte avec lui une horloge, et il dispose aussi d'un radar qui lui permet d'explorer l'espace autour de sa trajectoire. Comment ce voyageur percevra-t-il l'espace-temps qui l'entoure, sa perception sera-t-elle conforme à celle de l'espace-temps newtonien ? Ainsi que nous allons le découvrir, cette perception sera très différente de la nôtre en raison de la vitesse finie de propagation des signaux : en effet, alors que dans la vision de Newton les vitesses peuvent être arbitrairement grandes, l'espace-temps relativiste est contraint par un postulat fondamental que nous énonçons ci-après.

Postulat fondamental.

Il existe une *vitesse limite* qu'aucune particule ne peut dépasser (figure 4.1). Cette vitesse limite est celle des signaux électromagnétiques se propageant dans le vide indépendamment de la façon dont ils ont été émis, et c'est en particulier la vitesse de la lumière dans le vide. C'est donc la vitesse des photons, ou « particules de lumière » et c'est aussi celle de toutes les particules de masse nulle. Enfin, les lignes d'univers des photons sont des droites, les *droites de lumière*, qui ont toutes la même inclinaison par rapport à la verticale : aucune direction d'espace n'est privilégiée, nous avons une situation *spatialement isotrope*. La vitesse de la lumière dans le vide, notée c , est par *définition* (\simeq représente « approximativement égal à ») :

$$c = 299\,792,458 \text{ km.s}^{-1} \simeq 300\,000 \text{ km.s}^{-1} = 3 \times 10^8 \text{ m.s}^{-1} .$$

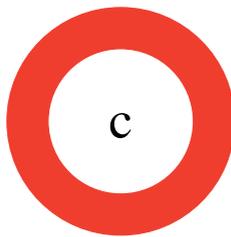


FIGURE 4.1. Un panneau de limitation de vitesse superflu.

Ce postulat mérite quelques commentaires. Tout d'abord, il n'existe pas de lien logique entre la première et la seconde phrase de son énoncé. La première phrase exprime une propriété géométrique de l'espace-temps, c'est certainement le point essentiel du postulat : en effet, nous verrons que l'existence d'une vitesse limite se traduit par une structure causale de l'espace-temps fondée sur les cônes de lumière, laquelle est fondamentalement différente de la structure newtonienne. Plus qu'une constante fondamentale, *la vitesse limite apparaît comme une caractéristique structurelle de l'espace-temps*. La seconde phrase exprime le fait que les photons ont une masse nulle. Si la masse des photons n'était pas nulle, ils se propageraient, tout comme les signaux électromagnétiques correspondants, à une vitesse inférieure à la vitesse limite ; de plus cette vitesse dépendrait de leur énergie ou, de façon équivalente, de leur fréquence. Comme il n'existe aucune indication en faveur d'une masse non nulle pour le photon, nous nous conformerons à l'usage et identifierons la vitesse limite et la vitesse de la lumière c .

Ensuite, étant donné que nous avons précédemment défini la seconde à l'aide de l'horloge à césium, la définition de la vitesse de la lumière fixe l'unité de longueur : c'est la distance parcourue par la lumière en une seconde,

environ 300 000 km. Cette définition signe l'arrêt de mort du mètre étalon. À notre échelle, celle du mètre, l'unité de temps commode est la nanoseconde (ns), ou milliardième de seconde ($1 \text{ ns} = 10^{-9} \text{ s}$). En une nanoseconde, la lumière parcourt 30 cm, ce qui est un point de repère très utile. Le système d'unités anglo-saxon donne un résultat particulièrement simple : en une nanoseconde, la lumière parcourt presque exactement *one foot* !

Renoncer aux mesures de longueur à l'aide d'un mètre au profit d'une mesure de temps est aujourd'hui chose courante. Le peintre en bâtiment au coin de la rue possède un télémètre laser qui lui permet de prendre les dimensions d'une pièce à rénover en se passant de son mètre à ruban. La police détermine la vitesse d'une voiture à l'aide de jumelles laser, qui sont des télémètres laser mesurant le temps d'aller-retour entre le policier et la voiture à des instants successifs, ce qui permet d'en déduire sa vitesse. La distance Terre-Lune est déterminée à un millimètre près en réfléchissant un rayon laser sur un miroir posé sur la Lune et en mesurant le temps d'aller-retour de l'impulsion laser. Les mesures de temps sont aujourd'hui les mesures les plus précises en physique, bien plus précises que les mesures directes de longueur, ce qui donne une raison supplémentaire pour abandonner le mètre comme unité de longueur, en fixant une fois pour toutes la valeur de c .

4.2 Temps propre

Un événement est défini par ses coordonnées d'espace et de temps, on peut donc lui faire correspondre un point de l'espace-temps ; par la suite nous identifierons souvent événement et point de l'espace-temps. Choisissons un point E de la ligne d'univers parcourue par notre voyageur extragalactique, représentée de la figure 4.2, et supposons qu'en ce point le voyageur émette des photons dans toutes les directions d'espace. D'après le postulat fondamental, les lignes d'univers de ces photons sont des demi-droites issues de E qui, en raison de l'isotropie, ont toutes la même inclinaison par rapport à la verticale et engendrent en E un demi-cône, appelé *cône de lumière futur* de E (figure 4.2). De même, supposons qu'un ensemble de photons arrive en E : leurs lignes d'univers engendrent aussi un demi-cône, le *cône de lumière passé* de E et la réunion des deux demi-cônes forme le *cône de lumière* de E . Dans la discussion de la figure 1.2, nous avons observé que la ligne d'univers d'un objet (dans ce cas un train) se propageant en ligne droite à vitesse constante était représentée dans l'espace-temps par une ligne droite, dont la pente était inversement proportionnelle à la vitesse ; cette pente était donc d'autant plus faible que la vitesse était grande. Comme la vitesse d'un objet massif est toujours inférieure à celle de la lumière, la pente en

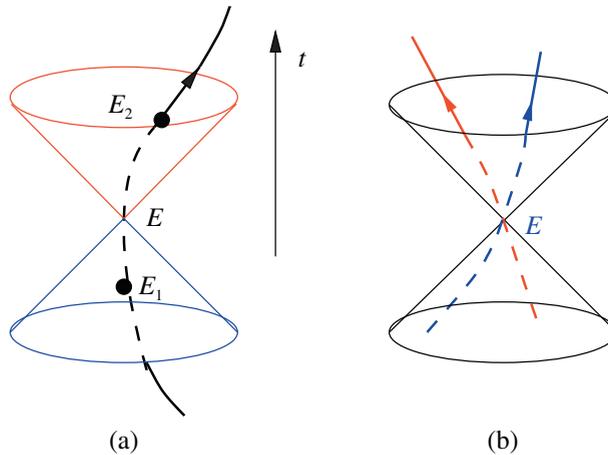


FIGURE 4.2. (a) Ligne d'univers d'un voyageur. On a représenté en rouge le cône de lumière futur de E et son cône passé en bleu. L'événement E_2 , situé à l'intérieur du cône futur de E , est postérieur à E , et E_1 , situé à l'intérieur du cône passé de E , est antérieur à E . (b) Indépendance des cônes de lumière par rapport aux lignes d'univers : le cône de lumière de E ne dépend que de E et pas des lignes d'univers passant par E . En termes plus intuitifs, la vitesse de la lumière ne dépend pas de la vitesse de la source.

chaque point de sa ligne d'univers sera toujours plus grande que celle des lignes d'univers des photons. Il en découle que la ligne d'univers d'un objet massif comme celle de notre voyageur passant par un point E sera toujours située à l'intérieur du cône de lumière issu de E . Dans un système d'unités où la vitesse de la lumière est prise égale à 1 (on dit souvent : « on pose $c = 1$ »), système que nous utilisons en général pour les graphiques, les droites de lumière qui engendrent les cônes de lumière sont inclinées à 45° par rapport à la verticale. De façon équivalente, on peut définir une coordonnée de temps x^0 par $x^0 = ct$, ou bien porter sur les graphiques ct en ordonnées.

Tous les cônes de lumière sont identiques et à chaque événement E est associé son cône de lumière. Comme les cônes de lumière sont des structures géométriques intrinsèques de l'espace-temps, ils sont indépendants de la ligne d'univers passant par E , et donc du voyageur (figure 4.2), et de plus ils sont spatialement isotropes : ils ne peuvent pas définir de direction privilégiée dans l'espace à trois dimensions. En revanche, ils permettent de définir les notions de passé et de futur : les événements situés sur la ligne d'univers du voyageur après E , comme l'événement E_2 sur la figure 4.2, sont postérieurs à E , tandis que les événements situés avant E sur cette ligne, comme E_1 , sont antérieurs à E .

Dans le cas newtonien, nous avons mesuré les distances sans prendre de précautions particulières, en utilisant des règles rigides. En ce qui concerne le temps, si nous avons admis que deux horloges éloignées pouvaient parfaitement

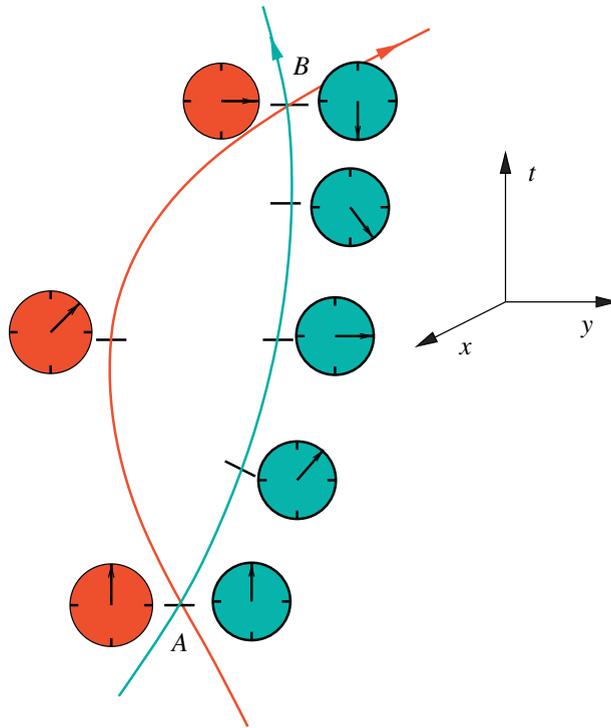


FIGURE 4.3. Croisement de deux lignes d'univers et temps propres de deux observateurs. La ligne d'univers et l'horloge du premier observateur sont dessinées en bleu, celles du second en rouge. Les deux horloges sont synchronisées en *A*, elles ne le sont plus en *B*.

indiquer des heures différentes, nous avons aussi supposé qu'il était toujours possible de les synchroniser sur une horloge de référence, une horloge maîtresse universelle donnant le temps absolu newtonien. Ce que nous souhaitons maintenant, c'est introduire la notion d'un observateur capable de mesurer temps et distances sans faire ce type d'hypothèse. En effet, la synchronisation des horloges, tout comme la notion de règle rigide, posent problème dès que l'on doit tenir compte de la vitesse finie de la lumière.

Revenons au voyageur parcourant le système solaire suivant la ligne d'univers de la figure 4.2 et, suivant l'usage, appelons-le désormais *observateur*. Cet observateur transporte une horloge atomique qui bat régulièrement le long de sa trajectoire. Si l'observateur parcourt la ligne d'univers d'un point (ou événement) *A* à un point *B*, son horloge va lui donner le temps écoulé entre l'événement *A* et l'événement *B* : l'intervalle de temps ainsi mesuré est appelé *intervalle de temps propre* (ou durée propre) pour l'observateur.

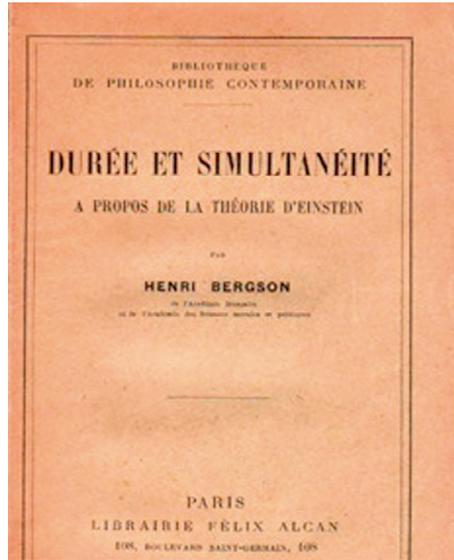


FIGURE 4.4. La couverture du livre de Bergson *Durée et simultanéité*.

Supposons maintenant que la ligne d'univers d'un second observateur croise celle du premier au point A (figure 4.3). Les deux observateurs peuvent décider de synchroniser leurs horloges lors de cette première rencontre, c'est-à-dire vérifier que leurs indications concordent, par exemple que les deux horloges indiquent midi. Les deux observateurs se croisent à nouveau en un point B . Leurs horloges seront-elles toujours synchronisées ? En général ce ne sera pas le cas : les intervalles de temps propre mesurés par les deux observateurs seront différents. Le temps propre écoulé dépend non seulement du point de départ et du point d'arrivée sur la ligne d'univers, *mais aussi en général du chemin choisi pour aller du point de départ au point d'arrivée*. En revanche, cette différence des temps propres est indépendante du type d'horloge utilisé et du mécanisme à la base de leur fonctionnement. Si les observateurs transportent chacun deux horloges différentes, par exemple une horloge à césium (§ 3.3.4) et un maser à hydrogène (§ 6.3.2), les intervalles de temps propre mesurés par chaque observateur seront identiques pour leurs deux horloges. Nous reviendrons ultérieurement sur la situation décrite sur la figure 4.3, illustrant le célèbre « *paradoxe des jumeaux* » : si deux jumeaux partent du même point A avec le même âge, en général ils se croiseront au point B avec un âge différent. Nous donnerons dans un cas simple l'expression explicite de la différence de temps propre au point B . *Nous avons donc rejeté la notion de temps absolu newtonien* : chaque observateur dispose d'un temps qui lui est propre, mesuré par une horloge qu'il

transporte au cours de ses déplacements, et ce temps est tout aussi valable que le temps mesuré par n'importe quel autre observateur utilisant sa propre horloge. Contrairement au cas newtonien, une durée n'est pas en général donnée par la différence de deux dates : seule la durée propre est physiquement pertinente. Cependant, si l'on connaît explicitement les lignes d'univers des deux observateurs de la figure 4.3, la relativité restreinte permet de relier le temps propre d'un des deux observateurs au temps propre de l'autre. Afin d'éviter toute ambiguïté, il est utile de souligner le point suivant : il est toujours possible d'étiqueter les points de l'espace-temps par une date t et une position, mais le temps t est ce que nous appellerons un *temps coordonnée* (ou fonction temporelle : voir Marc Lachièze-Rey, *Voyager dans le temps*). C'est le temps utilisé dans les figures, mais il n'est pas nécessairement et immédiatement relié à un temps physique, ou temps propre, mesurable par une horloge liée à un observateur et se déplaçant avec celui-ci. La distinction temps propre/temps-coordonnée est évidemment sans objet en physique newtonienne où le temps est absolu, mais elle est fondamentale, particulièrement en relativité générale.

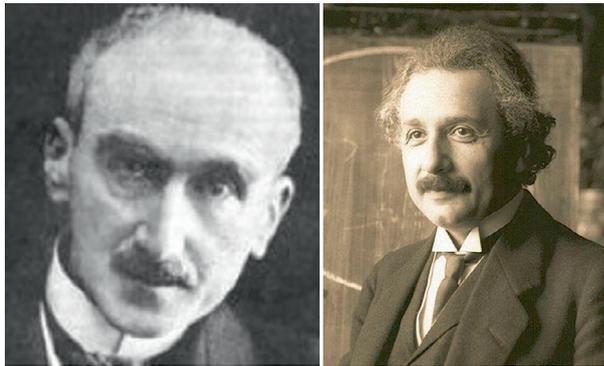


FIGURE 4.5. Henri Bergson (1859-1941) et Albert Einstein (1879-1955).

Le rejet d'un temps newtonien absolu est probablement la révolution conceptuelle la plus importante apportée par Einstein ; certains physiciens évoquent même une « disparition du temps ». Sans aller jusque-là, soulignons que cette révolution conceptuelle n'est généralement pas appréciée à sa juste valeur, y compris par nombre de physiciens. Poincaré avait en main tous les ingrédients de la relativité restreinte et en particulier les formules donnant la transformation du temps, mais il semble qu'il n'ait jamais voulu sauter le pas et abandonner la notion d'un temps absolu. Bien qu'Henri Bergson ait très sérieusement étudié la relativité à partir des articles originaux, les implications

contenues dans le graphique de la figure 4.3 étaient inconcevables pour ce philosophe, comme l'a montré le célèbre débat Bergson/Einstein organisé à Paris le 6 avril 1922 (figures 4.4 et 4.5). Le philosophe Gilles Deleuze résume ainsi l'attitude de Bergson :

« Bergson n'a jamais cru que la relativité fût fausse, mais seulement qu'elle était impuissante à constituer la philosophie du temps réel qui devait lui correspondre. »

Encore aujourd'hui, il est facile de trouver sur Internet des réfractaires au caractère relatif du temps, les « relativisto-sceptiques ». Cependant ceux-ci utilisent sans problème le GPS de leur voiture pour s'orienter, alors que si l'on ne tenait pas compte des effets illustrés sur la figure 4.3, ils n'arriveraient jamais à destination ! De plus, le GPS doit aussi prendre en compte un effet de relativité générale que nous examinerons au chapitre 5.

4.3 Espace-temps de Minkowski

Notre observateur de la figure 4.2 restait confiné à sa ligne d'univers, mais il serait logique qu'il s'intéresse aussi à son environnement, à ce qui se passe autour de lui. En particulier, la notion de simultanéité de deux événements situés en des points d'espace différents pose problème en l'absence d'un temps universel, et l'observateur aimerait savoir, lorsque son horloge en E marque une date t , quels sont les événements de son environnement qu'il peut considérer comme simultanés avec E . Plus généralement, comment peut-il dater des événements qui n'appartiennent pas à sa ligne d'univers ? La stratégie décrite ci-après a été conçue par Poincaré en 1900. Commençons par un exemple simple, où l'observateur est un peintre en bâtiment disposant d'une horloge qui mesure la largeur d'une pièce à rénover avec un télémètre laser. Le télémètre envoie au temps t_1 de l'horloge une impulsion laser qui suit un trajet horizontal dans l'espace ordinaire, est réfléchi par le mur en un point M et est reçue au temps t_2 . Comme le temps aller est identique au temps retour, le peintre va logiquement attribuer à la réflexion par le mur une date t qui est la moyenne arithmétique de t_1 et t_2

$$t = \frac{1}{2}(t_1 + t_2). \quad (4.1)$$

Plus précisément, il va conclure que l'impulsion laser a été réfléchi par le mur quand sa propre horloge indiquait le temps t de l'équation (4.1). Cependant, il faut bien se rendre compte que cette définition de la *simultanéité* n'est qu'une *convention*, qui est loin d'avoir le même statut que la simultanéité absolue de

la physique newtonienne. Le diagramme d'espace-temps de l'opération est dessiné sur la figure 4.6a. Ce schéma peut servir à synchroniser deux horloges, l'une en E et l'autre en M , sans avoir à connaître la distance entre les deux horloges : c'est la *synchronisation double voie*, examinée plus en détail au § 6.2. On remarque que cette synchronisation est indépendante de la vitesse du signal utilisé : il suffit que les trajets aller et retour du signal soient symétriques. Le diagramme 4.6a montre aussi que la distance L au mur est donnée simplement en fonction du temps de parcours $(t_2 - t_1)$:

$$L = \frac{c}{2}(t_2 - t_1). \quad (4.2)$$

Cependant le résultat dépend cette fois de la vitesse du signal. Pour mesurer les distances, nous avons remplacé les règles rigides par des échanges de photons. Le lecteur l'aura remarqué, nous faisons jouer un rôle privilégié au temps : l'unité de temps, la seconde, est l'unité de référence, alors que l'unité de longueur est une unité dérivée par l'intermédiaire de la vitesse de la lumière. De plus nous avons évité l'utilisation de règles pour quadriller l'espace et nous nous sommes repérés dans l'espace en mesurant uniquement des temps de départ et d'arrivée de photons. *Ce sont donc les mesures de temps qui sont fondamentales.*

Revenons à notre observateur initial, qui suit une ligne d'univers quelconque. Dans l'exemple précédent, nous avons pu identifier temps propre et temps coordonnée ordinaire t , car notre peintre en bâtiment était immobile dans le système d'axes où est défini le temps coordonnée ; toutefois une telle identification n'est pas possible dans le cas général. Nous allons introduire une notation spécifique en désignant par τ le temps propre d'un observateur tel qu'il est mesuré par son horloge. L'observateur est muni, en plus de son horloge, d'un radar qui envoie et détecte des impulsions électromagnétiques de très courte durée, de telle sorte qu'il peut mesurer avec précision l'intervalle de temps entre émission et réception. Le radar émet une impulsion à un instant τ_1 du temps propre de l'observateur, ou date τ_1 , cette impulsion est réfléchiée au point d'espace-temps M (M pour miroir) et revient vers l'observateur qui la détecte à la date τ_2 . Comme ci-dessus, nous allons décider *par convention* que l'événement M est simultané à un événement E situé sur la ligne d'univers si le temps propre τ de E est la moyenne arithmétique de τ_1 et τ_2 : $\tau = (\tau_1 + \tau_2)/2$. L'ensemble des événements simultanés à E forme un sous-espace de dimension trois de l'espace-temps, et donc une hypersurface, l'*hypersurface de simultanéité* de E . Comme il est difficile de dessiner dans un espace à quatre dimensions, nous avons représenté cette hypersurface dans un espace-temps à trois dimensions (figure 4.6b), où cette hypersurface est une surface ordinaire. Le lecteur qui a du mal à se représenter un espace à quatre dimensions ne doit pas s'inquiéter, les auteurs sont

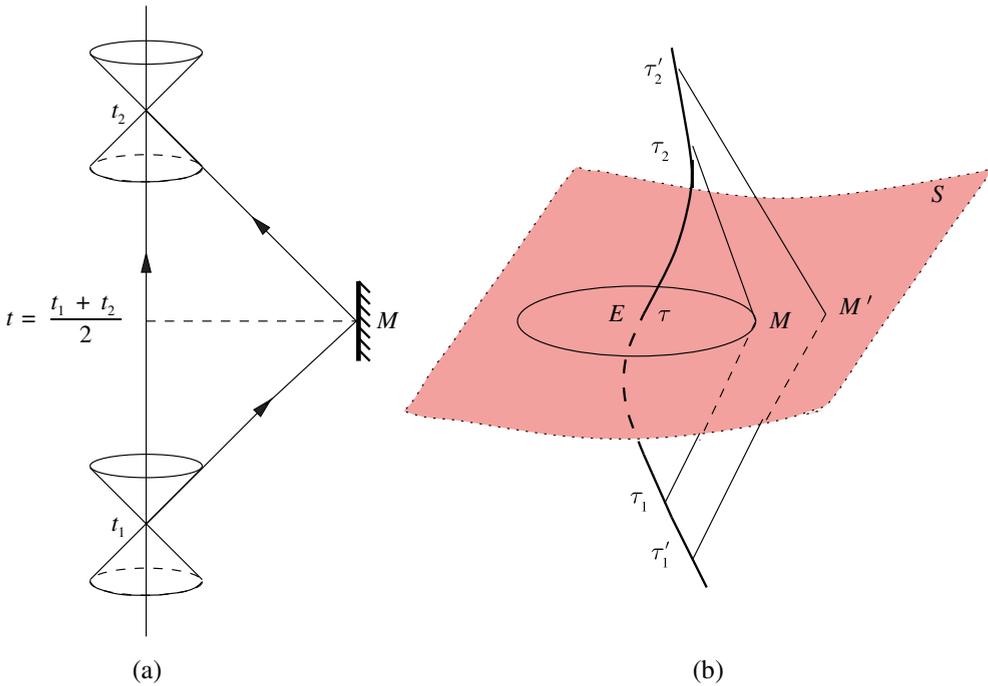


FIGURE 4.6. (a) Diagramme d'espace-temps pour définir la simultanéité. L'impulsion laser est réfléchi par le mur au point d'espace-temps M auquel l'observateur attribue par définition la date $t = (t_1 + t_2)/2$: c'est le principe de la synchronisation double voie. (b) Ligne d'univers d'un observateur et surface de simultanéité S (en rose) d'un événement E sur cette ligne, de date τ selon l'horloge de l'observateur, dans un espace-temps à trois dimensions. Les points M et M' appartiennent à cette surface si $(\tau_1 + \tau_2)/2 = (\tau'_1 + \tau'_2)/2 = \tau$.

dans le même cas ! Il suffit de visualiser les figures dans un espace-temps à trois dimensions, une de temps et deux d'espace, comme dans la figure 4.6b où l'hypersurface de simultanéité est une surface ordinaire. On peut même utiliser un espace-temps à deux dimensions, une de temps et une d'espace, comme dans la figure 1.2. La façon dont nous avons défini la simultanéité souligne une différence essentielle entre l'espace-temps ordinaire et l'espace-temps relativiste : la surface de simultanéité de E dépend de l'intégralité de la ligne d'univers, et pas uniquement de l'instant τ . Ce graphique illustre bien la notion d'univers-bloc : pour définir la simultanéité, nous avons besoin de l'histoire de l'observateur dans son intégralité. En relativité, il n'existe pas d'instant présent valable pour tous les observateurs. Deux événements simultanés pour un observateur ne le seront pas pour un autre se déplaçant par rapport au premier. Un exemple classique est celui d'un train, avec un passager situé au milieu de celui-ci, qui croise

un voyageur immobile sur le quai de la gare. Deux signaux lumineux sont émis, l'un depuis la tête du train, l'autre depuis sa queue, et le passager du train les reçoit au même instant. Il en conclut que les deux signaux ont été émis simultanément car ils ont parcouru la même distance, la moitié de la longueur du train. Le voyageur sur le quai les reçoit aussi au même instant, mais comme le signal émis depuis la queue du train a parcouru une distance plus grande que celui émis depuis sa tête, il conclut que le premier signal a été émis antérieurement au second : pour lui, l'émission des deux signaux n'est pas simultanée.

L'espace-temps muni de ses cônes de lumière forme *l'espace-temps de Minkowski* (figure 4.7). Dans l'espace-temps newtonien, les surfaces de simultanéité sont des (hyper-) plans $t = \text{constante}$ (figure 1.3) et ces surfaces de simultanéité sont universelles, tout comme le temps qui les définit. Elles ne dépendent que de l'événement, pas de la ligne d'univers de l'observateur. Dans l'espace de Minkowski, les surfaces de simultanéité dépendent de l'observateur, et le mieux que l'on puisse faire est d'associer un cône de lumière à chaque événement. Si un événement E_2 se trouve en dehors du cône de lumière de E comme dans la figure 4.7, alors l'ordre des temps de E et E_2 dépend de l'observateur. Cependant, si un événement E_3 se trouve à l'intérieur du cône futur de E par exemple, alors E_3 sera postérieur à E pour tout observateur. Ces observations permettent de définir une notion de causalité. En physique newtonienne, nous l'avons mentionné, un événement ne peut influencer que ceux qui lui sont postérieurs et inversement il ne peut avoir été influencé que par les événements antérieurs. En relativité, un événement ne peut influencer que ceux situés dans son cône futur et, inversement, il ne peut avoir été influencé que par les événements de son cône passé. De plus, il existe une troisième possibilité, illustrée sur la figure 4.7 : il ne peut exister aucune relation causale entre les événements E_1 et E_2 , car E_2 est situé hors du cône de lumière de E_1 . La structure causale de la relativité fondée sur les cônes de lumière est donc complètement différente de celle de la physique newtonienne, qui se confond avec sa structure chronologique.

4.4 Synchronisation des horloges et transformation du temps

Jusqu'à présent, nous n'avons pas éprouvé le besoin de spécifier le système de coordonnées, ou *référentiel*, dans lequel nous mesurons temps et distances. En physique il existe une classe de référentiels privilégiés dans lesquels les lois de la physique prennent une forme particulièrement simple, appelés *référentiels d'inertie*. Un référentiel d'inertie est une idéalisation de physicien, mais on peut en donner une très bonne approximation : un référentiel d'inertie est un référentiel

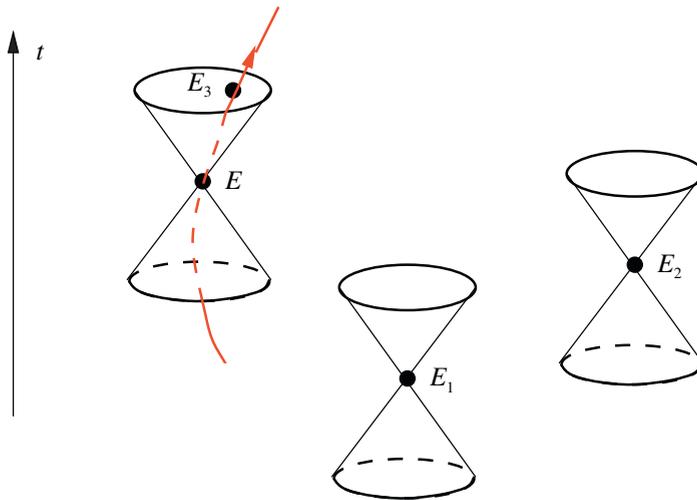


FIGURE 4.7. Espace-temps de Minkowski : à chaque point de l'espace-temps, ou événement, est attaché un cône de lumière. On a représenté un tel cône pour les événements E , E_1 et E_2 . Le point E_3 est situé à l'intérieur du cône futur de E et peut être influencé par E , mais le point E_2 à l'extérieur du cône de lumière de E_1 ne peut avoir aucune relation causale avec E_1 .

en chute libre. Par exemple la Station spatiale internationale (*International Spatial Station* ou ISS) est une excellente approximation d'un référentiel d'inertie, pourvu qu'elle n'ait pas de mouvement de rotation sur elle-même afin d'induire une gravité artificielle. Dans le langage courant, on invoque « l'impesanteur » : un individu flotte dans la cabine d'un ascenseur en chute libre tout comme un astronaute dans une station spatiale. On a en quelque sorte « effacé » la gravitation, et c'est la raison pour laquelle la description du mouvement d'un objet soumis à des forces autres que de gravitation sera particulièrement simple dans un référentiel en chute libre.

Cependant on peut se poser la question : une fois que l'on a déterminé un référentiel d'inertie, ce choix est-il unique ? En fait, tout référentiel se déplaçant par rapport à l'ISS suivant un mouvement sur une droite parcourue à vitesse constante, ou *mouvement rectiligne uniforme*, et sans mouvement de rotation sur lui-même, est aussi un référentiel d'inertie, en raison d'un principe fondamental de la physique, le *principe de relativité*.

Principe de relativité.

Si l'on ne peut pas se repérer par rapport à l'extérieur d'un référentiel d'inertie, par exemple si l'on ne peut pas regarder à l'extérieur de l'ISS, il est impossible de déceler le mouvement de ce référentiel et tous les référentiels d'inertie sont équivalents.

Sans aller dans l'espace, il suffit de s'imaginer dans un avion volant à son altitude de croisière en l'absence de turbulences : on peut déambuler dans l'avion, boire son café, etc., exactement comme sur la terre ferme, sans déceler un quelconque mouvement tant que l'on ne regarde pas par les hublots. En revanche, on s'aperçoit immédiatement de toute accélération de l'avion, par exemple quand le pilote amorce un virage. Le principe de relativité a été énoncé pour la première fois en 1632 par Galilée dans son *Dialogue sur les deux grands systèmes du monde*, en se plaçant dans un cadre mécanique. Ce principe va à l'encontre de la conception newtonienne d'un espace absolu, étant donné que l'espace associé à un référentiel se déplaçant à vitesse constante dans cet espace absolu est tout aussi valable. Dans le cadre de la physique du tout début du XX^e siècle, incluant mécanique et électromagnétisme, Poincaré fut le premier à énoncer ce principe de la façon la plus claire.

Pour définir complètement un référentiel d'inertie, nous avons besoin de lui associer un système d'axes permettant de repérer un point dans l'espace ainsi qu'un ensemble d'horloges synchronisées capable de donner le temps en chaque point d'espace ou, en pratique, dans le voisinage d'un ensemble de points quadrillant l'espace (figure 4.8). Toutefois, synchroniser des horloges est loin d'être évident en raison de la vitesse finie de la lumière. La méthode de synchronisation la plus simple est la suivante : une horloge de référence H_0 est

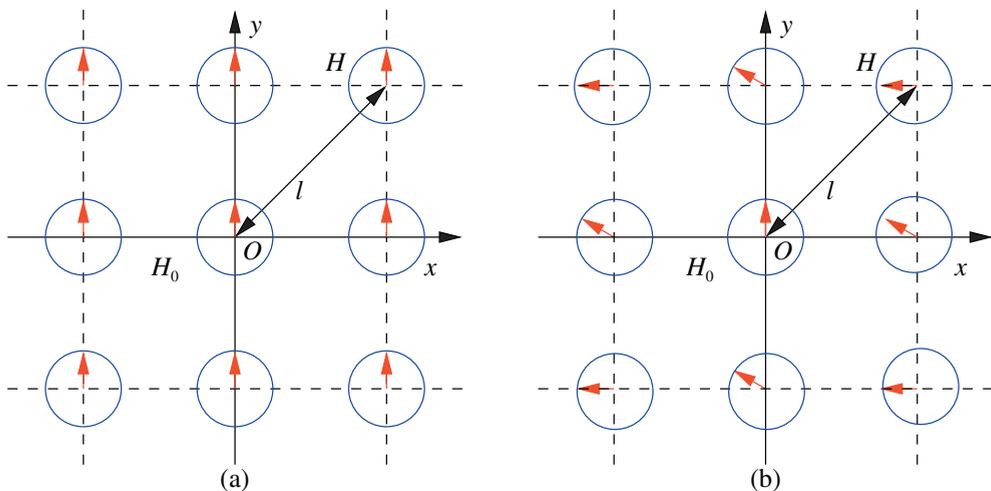


FIGURE 4.8. Synchronisation des horloges et quadrillage de l'espace. (a) Figure tracée dans un référentiel d'inertie : toutes les horloges sont synchronisées. (b) Horloges vues par un observateur en O . Si l'horloge H_0 de cet observateur indique $t = 0$, cet observateur voit une horloge H située à une distance l indiquer le temps $t = -l/c$.

localisée à l'origine des coordonnées. Au temps $t = 0$, elle envoie un signal lumineux vers l'horloge à synchroniser H , située à une distance l . L'horloge à synchroniser est bloquée sur le temps $t = l/c$, et quand l'observateur associé à cette horloge reçoit le signal, il la libère. C'est la méthode à « une voie » du § 6.2. Une autre méthode, légèrement plus compliquée, est la méthode à deux voies, qui présente l'avantage de ne pas dépendre d'une mesure préalable des distances. Elle suit le schéma de la figure 4.6a : au temps t_1 selon H , cette horloge envoie un signal lumineux vers H_0 , qui réfléchit le signal, lequel est reçu au temps t_2 . H_0 informe H du temps t (le temps t de H_0 !) auquel le signal a été réfléchi. Cette information met bien sûr un certain temps à arriver à H . Mais connaissant t , H peut être synchronisée sur H_0 grâce à (4.1) : $t = (t_1 + t_2)/2$.

L'étape suivante consiste à relier les temps mesurés dans deux référentiels d'inertie différents, par exemple celui de l'ISS et celui d'une fusée se déplaçant suivant une droite à vitesse constante v par rapport à l'ISS. Une horloge du référentiel de la fusée est fabriquée de la façon suivante : une impulsion lumineuse effectuée dans la fusée un aller-retour entre deux miroirs distants de $L/2$ disposés suivant la direction $O'y'$ perpendiculaire à la direction de la vitesse (figure 4.9). Dans la fusée, le trajet est parcouru par la lumière en un temps $t' = L/c$, autrement dit l'horloge ne bat pas la seconde mais L/c , du moins est-ce le point de vue d'un physicien de la fusée. Examinons le trajet de l'impulsion laser tel qu'il est vu par une physicienne de l'ISS. Pendant que la lumière effectue son aller-retour, la fusée a avancé de vt par rapport à l'ISS, où t est le temps mis par l'impulsion laser pour faire son aller-retour, *mesuré dans l'ISS*. Il faut bien noter que ce sont deux horloges *différentes* du réseau d'horloges associé à l'ISS qui vont mesurer la date de l'émission et celle de la réception de l'impulsion laser (figure 4.9). Dans le référentiel de l'ISS, la lumière parcourt un trajet plus long que dans la fusée, car il n'est pas très difficile de montrer que les distances perpendiculaires à la vitesse sont identiques dans la fusée et l'ISS. Comme la vitesse de la lumière est c dans tous les référentiels d'inertie, le temps de parcours t tel qu'il est mesuré dans l'ISS est *plus long* que dans la fusée.

Encadré 4.1 : le facteur de Lorentz.

Pour relier quantitativement le temps de parcours t' de l'impulsion lumineuse dans la fusée et celui t dans l'ISS, il suffit de remarquer que la distance parcourue dans l'ISS est donnée par le théorème de Pythagore, et par conséquent :

$$c^2 t^2 = L^2 + v^2 t'^2. \quad (4.3)$$

Compte tenu de $t' = L/c$, la relation entre t et t' en découle :

$$t = \frac{L/c}{\sqrt{1 - v^2/c^2}} = \gamma \frac{L}{c} = \gamma t'. \quad (4.4)$$

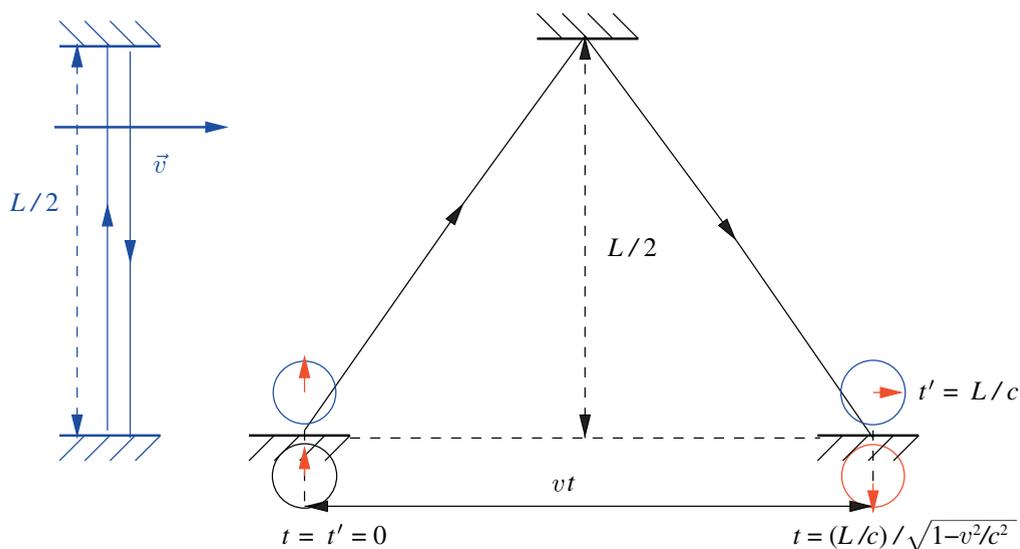


FIGURE 4.9. Horloge fixe dans une fusée à gauche et en mouvement par rapport à la Station spatiale internationale à droite ; le diagramme est tracé dans l'espace euclidien à trois dimensions, dont deux seulement sont représentées. Il est important de remarquer que c'est la *même* horloge (en bleu) qui mesure les instants de départ et d'arrivée de l'impulsion laser dans la fusée, mais ce sont deux horloges *différentes*, dessinées respectivement en noir et en rouge, qui mesurent les instants de départ et d'arrivée de l'impulsion dans l'ISS.

La quantité notée γ dans (4.4) est appelée *facteur de Lorentz*. Ce facteur est toujours plus grand que 1 (égal à 1 uniquement pour une vitesse nulle) : $\gamma \geq 1$. On constate que c'est le rapport v^2/c^2 qui mesure le caractère relativiste du mouvement. Si $v^2/c^2 \ll 1$, on retrouve la mécanique newtonienne, mais si cette condition n'est pas satisfaite, alors la relativité ne peut pas être ignorée. Cette condition jouera à nouveau un rôle important au chapitre 5.

La conclusion de cette analyse est que *l'intervalle de temps entre l'émission et la réception du signal lumineux est plus long dans l'ISS que dans la fusée*. Une physicienne de l'ISS qui observe la réception du signal lumineux voit la montre de son collègue de la fusée indiquer le temps $t' = L/c$, alors que l'horloge de l'ISS qui coïncide avec cette réception indique $t = \gamma(L/c)$. Autrement dit, elle constate que, de son point de vue, les horloges de la fusée battent trop lentement. N'a-t-on pas là une contradiction avec le principe de relativité, une possibilité de privilégier l'un des deux référentiels ? Ce n'est pas le cas, car un physicien de la fusée constate également que les horloges de l'ISS battent trop lentement ! Il faut souligner le point suivant : dans la fusée, l'intervalle de temps entre l'émission et la réception du signal lumineux est mesuré au *même* point

d'espace, alors que dans l'ISS cet intervalle, nous l'avons noté, est mesuré par des horloges situées en des points *différents*, $x = 0$ et $x = vt$, dessinées respectivement en noir et en rouge sur la figure 4.9. Nous avons défini précédemment le temps propre comme le temps mesuré par une horloge se déplaçant avec l'observateur. C'est le cas pour le physicien de la fusée : le battement de l'horloge est donné par l'intervalle de temps entre l'émission et la réception d'une impulsion lumineuse, et les deux événements se passent au *même endroit* dans la fusée : l'intervalle de temps $t' = L/c$ est donc un intervalle de temps propre. En revanche, pour la physicienne de l'ISS, les deux événements ne se passent pas au même endroit : le point d'émission et le point de réception sont séparés par la distance vt . En accord avec la définition précédente, un intervalle de temps entre deux événements se produisant au même point dans un référentiel d'inertie est appelé intervalle de temps propre, noté $\Delta\tau$. Si l'on compare le temps mesuré par un observateur immobile dans un référentiel d'inertie et celui mesuré par un observateur en mouvement par rapport à ce référentiel, l'équation (4.4) montre que le temps propre est toujours le plus court.

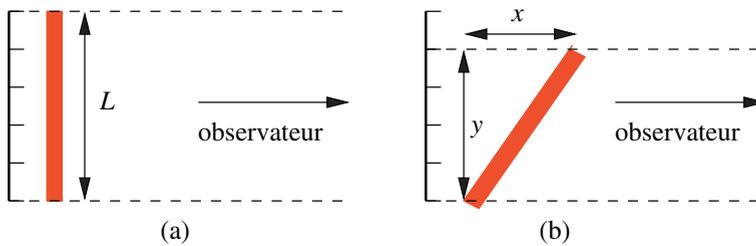


FIGURE 4.10. Effet de parallaxe. Afin de simplifier au maximum, nous avons supposé l'observateur à grande distance. Si la longueur de la règle est L , l'observateur mesure une longueur y dans le cas (b), et non L .

La transformation du temps entre deux référentiels inertiels peut être comprise comme résultant d'un effet de parallaxe (figure 4.10). Supposons que nous voulions mesurer dans l'espace ordinaire la longueur d'une règle à l'aide d'un mètre. Nous obtiendrons un résultat correct si la règle est parallèle au mètre (figure 4.10a), mais pas si la règle est inclinée par rapport à ce mètre (figure 4.10b). L'observateur mesure une longueur y plus courte que la « longueur propre » L , car, d'après le théorème de Pythagore,

$$L^2 = x^2 + y^2,$$

où x et y sont définis dans la figure 4.10. La situation est inversée pour un observateur qui mesure un intervalle de temps depuis un référentiel en mouvement

par rapport à celui où l'on mesure un temps propre. En effet, à cause du signe moins dans l'expression du temps propre (voir l'équation (4.3)) :

$$(\Delta\tau)^2 = (\Delta t)^2 - (\Delta x/c)^2,$$

l'intervalle de temps propre est plus court que l'intervalle Δt mesuré dans le référentiel en mouvement.

Encadré 4.2 : temps propre et métrique de Minkowski.

Soit deux événements quelconques séparés par un intervalle de temps Δt et par un intervalle d'espace Δx dans un certain référentiel. On peut construire avec Δt et Δx la distance de Minkowski, ou *métrique de Minkowski* :

$$c^2(\Delta\tau)^2 = c^2(\Delta t)^2 - (\Delta x)^2. \quad (4.5)$$

On remarquera la parenté avec le théorème de Pythagore (voir l'équation (4.3)), mais avec une différence cruciale : le terme $(\Delta x)^2$ est affecté d'un signe moins. Avec un signe plus, le second membre de (4.5) serait la distance euclidienne au carré dans l'espace (t, x) . Dans l'exemple ci-dessus, $\Delta\tau$ est bien l'intervalle de temps propre dans le référentiel de la fusée où $\Delta x = 0$ et cette propriété est générale : la quantité $(c^2(\Delta t)^2 - (\Delta x)^2)$ est la même dans tous les référentiels d'inertie, c'est le temps mesuré par une horloge lorsque $\Delta x = 0$. C'est ce que l'on appelle un invariant relativiste.

4.5 Le paradoxe des jumeaux

Nous allons maintenant détailler une conséquence particulièrement spectaculaire de la loi de transformation du temps, le paradoxe des jumeaux, qui a déjà été évoqué (figure 4.3). Imaginons la situation suivante : une astronaute intrépide, Alice, part pour un voyage dans la Galaxie à une vitesse égale à $4c/5 = 0,8c$, tandis que son frère jumeau Bob reste tranquillement sur la Terre (figure 4.11). Ce choix de la vitesse permet d'obtenir un facteur de Lorentz simple :

$$\gamma = \frac{1}{\sqrt{1 - 16/25}} = \frac{5}{3}.$$

Alice voyage 3 ans suivant son temps propre, soit 5 ans pour un observateur terrestre, et elle s'éloigne donc à $(4/5) \times 5 = 4$ années-lumière de la Terre. Elle fait ensuite demi-tour, repart en sens inverse toujours à la même vitesse, et rejoint son jumeau après un voyage qui pour elle a duré 6 ans, mais qui pour son frère jumeau a duré 10 ans. Si Alice a quitté la Terre à 20 ans, elle est âgée de 26 ans à son retour, mais Bob est lui âgé de 30 ans !

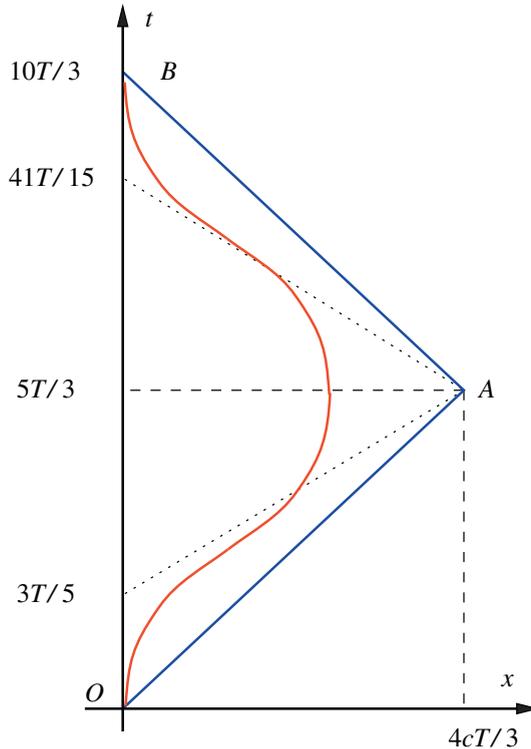


FIGURE 4.11. Le paradoxe des jumeaux. Alice s'éloigne de Bob avec une vitesse $v = 4c/5$, fait demi-tour en A et rejoint son frère jumeau Bob en B . La ligne d'univers de Bob est tracée en noir, celle d'Alice en bleu. Les lignes de simultanéité de A dans les deux référentiels associés à Alice sont tracées en pointillés. En rouge, une ligne d'univers d'Alice comportant uniquement des accélérations finies.

Le voyage des jumeaux est représenté sur le diagramme d'espace-temps de la figure 4.11, dessiné dans le référentiel de la Terre, qui est aussi celui de Bob. La ligne d'univers de Bob est l'axe des temps, étant donné qu'il reste immobile à $x = 0$. Celle d'Alice est composée de deux segments de droite, OA et AB : au point A , Alice fait brusquement demi-tour et entame son voyage retour vers son frère Bob qui se termine en B . Si son voyage a duré pour elle un temps $2T = 6$ ans (T pour l'aller et T pour le retour), pour Bob il a duré $2\gamma T = (10/3) \times 3 = 10$ ans. Le schéma de la figure 4.11 est critiquable, car non physique : au départ en O , au demi-tour en A et à son arrivée en B , Alice subit une accélération infinie. Un schéma plus satisfaisant est obtenu en se limitant à des accélérations finies, la ligne d'univers d'Alice étant alors représentée par la courbe en rouge. Le calcul de la différence d'âge est simple, mais cependant trop technique pour

être reproduit ici. Cette différence d'âge dépend quantitativement des détails du mouvement, mais qualitativement le résultat est toujours le même : le jumeau voyageur est toujours le plus jeune quand il rencontre son homologue sédentaire. Ce résultat est parfois appelé « paradoxe des jumeaux », mais il n'a en fait rien de paradoxal, car la situation des deux jumeaux n'est pas symétrique, un point que Bergson n'avait pas compris : le jumeau resté sur Terre ne subit aucune accélération, alors que le mouvement du jumeau voyageur est nécessairement accéléré et, comme nous l'avons observé, il est parfaitement possible de déceler un mouvement accéléré sans regarder au-dehors.

Le rôle des accélérations dans le paradoxe des jumeaux est souvent mal compris, et il vaut la peine de l'éclaircir. Tout d'abord certains auteurs, sous prétexte que l'on utilise des référentiels non inertiels, prétendent qu'il faut faire appel à la relativité générale. Rien n'est plus faux, la relativité restreinte sait parfaitement traiter les mouvements accélérés en l'absence de gravitation et il n'y a aucune raison d'introduire la gravitation ici. Ensuite, et bien que le paradoxe des jumeaux ne puisse pas exister sans qu'au moins un des référentiels soit accéléré, il ne faut surtout pas en déduire que la différence d'âge provient des accélérations ! Certes, nous l'avons noté, des horloges différentes peuvent réagir différemment aux accélérations, une horloge à balancier ou une horloge biologique peuvent réagir différemment d'une horloge atomique. Par exemple, dans le schéma de la figure 4.11 avec un demi-tour instantané, le jumeau voyageur ne va pas survivre aux accélérations infinies et à son retour sur Terre, il ne sera pas plus jeune que son jumeau, il sera mort ! La différence d'âge est donnée par une horloge idéale, insensible aux accélérations, et *cette différence ne dépend que de la vitesse sur le trajet.*

Bien que le résultat soit une conséquence incontournable de la relativité, il est tout de même un peu surprenant et il vaut la peine de mentionner plusieurs vérifications expérimentales. La vérification la plus spectaculaire du paradoxe des jumeaux vient de l'observation de muons dans un anneau de stockage où ils suivent une trajectoire circulaire. Le muon est une particule élémentaire très analogue à l'électron, mais il est environ 200 fois plus lourd et il est instable : il se désintègre en moyenne au bout de $2,2 \times 10^{-6}$ s ($2,2 \mu\text{s}$). Dans une expérience réalisée au CERN (Genève) en 1977, des muons de 3 GeV d'énergie, ce qui correspond à un facteur de Lorentz $\gamma \simeq 29,3$, sont stockés dans un anneau de 14 m de circonférence, où ils sont maintenus sur une trajectoire circulaire par un champ magnétique perpendiculaire à l'anneau. Une fois le muon inséré dans l'anneau de stockage, il tourne jusqu'au moment où l'on observe sa désintégration. On mesure alors une vie moyenne de $64,3 \mu\text{s}$, en excellent accord avec la prédiction de la relativité, aux erreurs expérimentales près : $2,2 \mu\text{s} \times 29,3 = 64,4 \mu\text{s}$. L'accélération des muons sur leur trajectoire est de 10^{20} m.s^{-2} , et on constate que cette

accélération n'a aucune influence sur la vie moyenne. Seule la vitesse intervient dans le résultat final.

L'expérience de Joseph Hafele et Richard Keating publiée en 1972 a consisté à embarquer des horloges atomiques à bord d'avions de ligne, à leur faire faire le tour de la Terre dans les deux sens et à les comparer à l'arrivée avec une horloge restée fixe sur la Terre. L'analyse complète de l'expérience est beaucoup plus complexe que dans le cas des muons, en particulier pour les deux raisons suivantes.

1. Un référentiel fixe par rapport à un point pris sur la surface de la Terre n'est pas d'inertie en raison de la rotation de la Terre sur elle-même, et il faut faire la comparaison avec un référentiel dont l'origine suit le mouvement du centre de la Terre et dont les axes sont fixes dans l'espace, car un tel référentiel, appelé *référentiel terrestre géocentré*, est lui approximativement inertiel.
2. Les horloges situées en altitude sont décalées par rapport à une horloge située au niveau de la mer, en raison du décalage vers le rouge gravitationnel que nous introduirons au chapitre suivant.

Une analyse complète tenant compte de ce deux effets montre que l'effet de transformation relativiste du temps est indispensable pour expliquer le décalage entre l'horloge fixe et les horloges voyageuses.

On doit également tenir compte de la transformation relativiste du temps si l'on veut synchroniser les horloges atomiques embarquées dans les satellites du GPS et les horloges fixes sur la Terre. La vitesse des satellites GPS par rapport à la Terre est d'environ 3,9 km/s, ce qui correspond à $v/c \simeq 1,3 \times 10^{-5}$ et à un facteur de Lorentz donné par :

$$\gamma - 1 \simeq v^2/2c^2 \simeq 0,85 \times 10^{-10}.$$

Cette valeur peut sembler ridiculement faible, mais au bout d'une heure, le décalage dû à la transformation du temps est d'environ $0,3 \mu\text{s}$ et en $0,3 \mu\text{s}$, la lumière parcourt 90 m, d'où une erreur de cet ordre sur la position donnée par le GPS si l'on ne fait pas la correction. L'analyse complète est plus complexe en raison des deux effets mentionnés ci-dessus. Nous y reviendrons au chapitre suivant.

Grâce à des horloges atomiques ultraprécises, les horloges optiques introduites au chapitre précédent, on a pu mettre en évidence la transformation du temps pour des vitesses typiques de la course d'un humain, de l'ordre de $10 \text{ m}\cdot\text{s}^{-1}$. Les horloges optiques à ion aluminium (un ion est un atome qui a perdu un ou plusieurs électrons) sont capables de mesurer le temps avec une précision relative de 10^{-17} , soit une dérive d'une seconde sur 3 milliards

d'années. Ces horloges reposent sur une transition électromagnétique entre deux niveaux d'énergie de l'ion, de fréquence $1,121 \times 10^{15}$ Hz, qui se situe dans le domaine optique. Cette fréquence est environ 10^5 fois plus élevée que celle utilisée par les horloges à césium, et cette différence de fréquence est à l'origine du gain en précision. Il est d'ailleurs vraisemblable que la définition de la seconde sera modifiée dans un avenir proche, étant donné le gain en précision apporté par les horloges optiques.

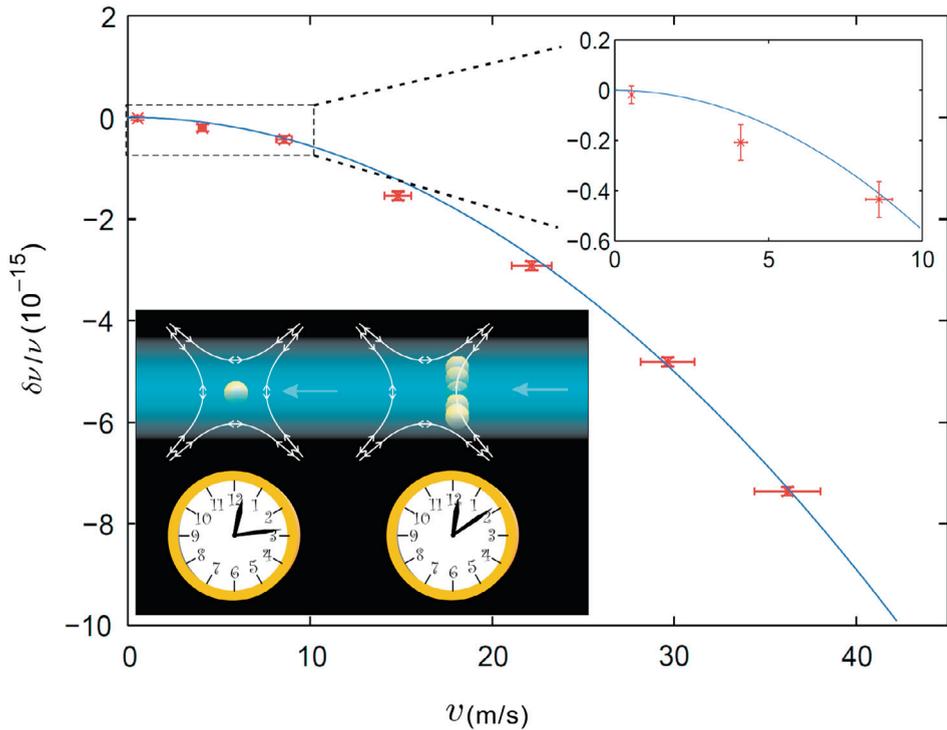


FIGURE 4.12. Vérification de la transformation relativiste du temps avec une horloge optique à ion aluminium. Les ions se déplacent dans le piège de droite avec une vitesse allant de 1 à 40 $\text{m}\cdot\text{s}^{-1}$: la fréquence ν du battement de l'horloge varie avec la vitesse des ions dans le piège. La comparaison avec l'horloge de gauche, correspondant à des ions immobiles dans le piège, permet de mettre en évidence la « dilatation du temps ». Adapté de C. Chou, D. Hume, T. Rosenband and D. Wineland, *Optical clocks and relativity*, *Science*, **329**, 1630 (2010). Courtoisie de David Wineland et du NIST.

Dans une horloge à ion aluminium, un ion unique est confiné dans un piège dont les dimensions sont de l'ordre du millimètre et il est manipulé par un laser (§ 3.4.3). En fonctionnement normal, l'ion responsable de la transition électromagnétique est pratiquement immobile dans le piège. Grâce à un champ de radiofréquences, on peut le mettre en mouvement et lui faire effectuer des

allers-retours dans le piège, à des vitesses allant de 1 à 40 m.s⁻¹. L'ion devient donc un jumeau voyageur, mais sur des distances très faibles, de l'ordre du millimètre. Une horloge avec un ion immobile (le jumeau sédentaire) est comparée à celle avec un ion voyageur : on voit très clairement l'horloge à ion voyageur retarder par rapport à celle à ion immobile. L'effet est visible expérimentalement pour des vitesses de quelques mètres par seconde (figure 4.12).

5

Le temps de la relativité générale

La relativité restreinte permet de traiter le mouvement de particules dont la vitesse n'est pas négligeable par rapport à celle de la lumière, pourvu qu'elles ne soient pas soumises à des forces de gravitation. La relativité générale, postérieure de dix ans à la relativité restreinte, permet d'inclure les forces de gravitation. La relativité générale repose sur le *principe d'équivalence* : la trajectoire d'un objet soumis uniquement à des forces de gravitation est indépendante de sa masse et de sa composition chimique, ce qui suggère que ces forces ont une origine géométrique. Une conséquence de ce principe est que la masse intervenant dans la deuxième loi de Newton, ou *masse inertielle*, est identique à la *masse gravitationnelle*, celle qui intervient dans la loi de la gravitation (encore une loi de Newton!). Dans le § 5.1, nous utilisons le principe d'équivalence pour montrer que les indications des horloges sont affectées par un champ gravitationnel, un phénomène directement relié au décalage vers le rouge gravitationnel. Le § 5.2 traite de cas où la gravité est faible, avec comme application principale le GPS, tandis que le § 5.3 introduit le cas de la gravité forte, avec comme exemple les trous noirs où les distorsions du temps sont particulièrement fortes. Enfin le § 5.4 examine brièvement le temps cosmologique.

5.1 Décalage vers le rouge gravitationnel

En relativité restreinte, les indications des horloges sont affectées par la vitesse, en relativité générale elles le sont aussi par la gravitation. Il est possible de s'en convaincre grâce à l'exemple illustré sur la figure 5.1 en utilisant la conservation

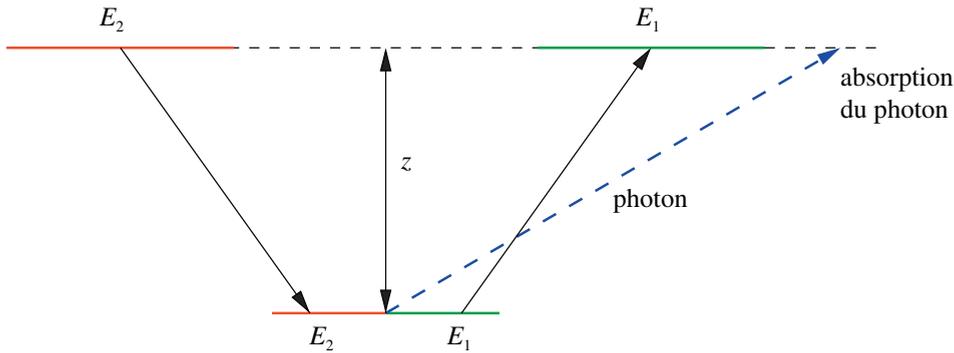


FIGURE 5.1. Décalage vers le rouge gravitationnel : ascension d'un photon dans le champ de pesanteur terrestre.

de l'énergie : un atome possède deux niveaux d'énergie E_1 (niveau fondamental) et E_2 (niveau excité), avec $E_2 > E_1$. Si l'atome est excité dans le niveau E_2 , il revient au niveau fondamental E_1 en émettant un photon d'énergie $E_\gamma = E_2 - E_1$ (figure 3.11). D'après la célèbre formule d'Einstein $E = mc^2$, l'atome dans le niveau E_1 (E_2) possède une masse $m_1 = E_1/c^2$ ($m_2 = E_2/c^2$). Par ailleurs, nous utiliserons le fait qu'il faut dépenser une énergie mgz pour porter une masse m de l'altitude zéro à l'altitude z ; g est l'accélération de la pesanteur, $g \simeq 9,81 \text{ m.s}^{-2}$ à Paris. Le principe d'équivalence intervient par le fait que la masse dans la relation d'Einstein, ou masse inertielle, est identique à celle donnant l'énergie gravitationnelle, ou masse gravitationnelle. Supposons que l'atome se trouve initialement dans le niveau E_2 à l'altitude z . Dans une première étape, on redescend lentement l'atome à l'altitude zéro. Au cours de cette descente, l'atome fournit une énergie E_2gz/c^2 , qui pourrait par exemple être utilisée pour monter une masse, et son énergie diminue donc de cette même quantité. Dans une deuxième étape, l'atome redescendu à l'altitude zéro émet un photon d'énergie E_γ , lequel est renvoyé vers le haut par un jeu de miroirs et est finalement absorbé à l'altitude z ; nous appelons E'_γ l'énergie absorbée. Dans une troisième étape l'atome, maintenant dans le niveau E_1 , est remonté à l'altitude z , ce qui coûte au milieu extérieur une énergie E_1gz/c^2 . Un bilan énergétique (encadré 5.1) montre que l'énergie du photon absorbé à l'altitude z est plus faible que celle du photon à l'altitude zéro. On peut interpréter intuitivement ce résultat en remarquant que le photon doit perdre de l'énergie en montant dans le champ de pesanteur terrestre : un photon est sensible à la gravité !

Encadré 5.1 : décalage vers le rouge gravitationnel.

Détaillons le bilan énergétique : par conservation de l'énergie, l'énergie finale $E_1 + E'_\gamma$ doit être égale à l'énergie initiale E_2 de l'atome plus l'énergie fournie par le milieu extérieur

$$E_1 + E'_\gamma = E_2 - E_2 \frac{gz}{c^2} + E_1 \frac{gz}{c^2},$$

soit

$$E'_\gamma = (E_2 - E_1) \left(1 - \frac{gz}{c^2}\right) = E_\gamma \left(1 - \frac{gz}{c^2}\right). \quad (5.1)$$

L'énergie E d'un photon est liée à sa fréquence ν par la relation de Planck-Einstein (3.3) $E = h\nu$, où h est la constante de Planck. On déduit de (5.1) la loi de variation de la fréquence dans le champ de pesanteur terrestre, où ν est la fréquence à l'altitude zéro et ν' celle à l'altitude z :

$$\nu' = \nu \left(1 - \frac{gz}{c^2}\right). \quad (5.2)$$

Ceci est le *décalage vers le rouge gravitationnel* : en effet, la fréquence ν' étant plus faible que ν , la fréquence se déplace vers le rouge dans le spectre lumineux.

Il est commode d'exprimer le décalage vers le rouge gravitationnel en utilisant la période T , qui est l'inverse de la fréquence : $T = 1/\nu$. La loi de transformation des périodes s'obtient en utilisant $(1 - \varepsilon)^{-1} \simeq 1 + \varepsilon$ si $\varepsilon \ll 1$:

$$T' = T \left(1 + \frac{gz}{c^2}\right), \quad (5.3)$$

où T est la période à l'altitude zéro et T' celle à l'altitude z . Cela veut dire (figure 5.2) que les horloges situées à basse altitude retardent par rapport à celles situées à une altitude plus élevée. Cet effet est très faible car $gz/c^2 \sim 10^{-16}$ pour $z = 1$ m, mais il a pu être vérifié expérimentalement dès 1960 par Robert Pound et Glenn Rebka sur une hauteur de 20 m. Il existe aujourd'hui des vérifications bien plus précises, et les horloges optiques (§ 3.4.3) permettent de détecter un décalage entre deux horloges dont l'altitude diffère de 30 cm (figure 5.3).

Revenons brièvement sur le problème des jumeaux : nous avons vu que l'horloge du jumeau sédentaire battait plus vite que celle du jumeau voyageur. Une horloge immobile, ou plus généralement une horloge attachée à un référentiel d'inertie, bat plus vite qu'une horloge dont la trajectoire comporte des accélérations. Dans le cas de la gravitation, si deux jumeaux, dont l'un est en chute libre, se croisent en un point d'espace-temps A avec le même âge et se rencontrent à nouveau en B , c'est le jumeau en chute libre qui sera le plus âgé en B .

Le décalage vers le rouge gravitationnel a un impact majeur sur la structure de l'espace-temps. Reprenons le schéma de la figure 5.2 : une horloge située à l'altitude zéro envoie des impulsions lumineuses brèves à intervalles réguliers $0, T, 2T, \dots, nT, \dots$ qui sont reçues à l'altitude z aux temps $T_0 = z/c$,

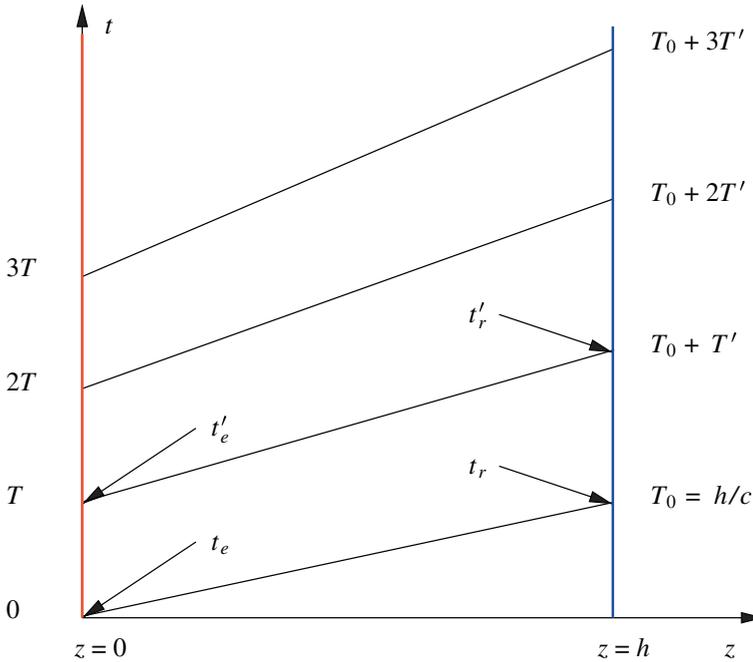


FIGURE 5.2. L'horloge à l'altitude $z = 0$ (rouge) émet des signaux aux temps $0, T, 2T \dots$. L'horloge à l'altitude z (bleu) les reçoit aux temps $T_0 = z/c, T_0 + T', T_0 + 2T' \dots$, avec $T' = T(1 + gz/c^2)$. Pour un observateur à l'altitude $z > 0$, l'horloge à $z = 0$ retarde. On notera que les temps propres mesurés en $z = 0$ et z le sont par deux horloges *différentes*, alors que t est un temps coordonnée unique.

$T_0 + T', T_0 + 2T' \dots, T_0 + nT', \dots$, où $T_0 = z/c$ est le temps nécessaire à la lumière pour parcourir la distance z et où T' est donné par (5.3). Essayons de représenter graphiquement cette suite d'événements dans l'espace-temps en leur attribuant des coordonnées de temps t et d'espace z . Soit t_e et t'_e les dates d'émission de deux impulsions lumineuses consécutives par l'horloge d'altitude zéro sur la figure 5.3; ces signaux sont reçus aux dates t_r et t'_r par l'horloge à l'altitude z . En raison de l'invariance par translation de temps, les intervalles de temps $(t'_r - t'_e)$ et $(t_r - t_e)$ doivent être égaux, ce qui implique $t'_r - t_r = t'_e - t_e = \Delta t$. Ce dernier résultat semble contredire le décalage vers le rouge gravitationnel $T' > T$ et l'expérience (voir la figure 5.2). En fait, il n'en est rien, car t est un temps coordonnée, un simple étiquetage de points de l'espace-temps et non un temps physique mesuré par une horloge. Nous pouvons toujours identifier Δt comme un intervalle de temps physique, ou temps propre, à l'altitude zéro car il est mesuré par une horloge unique : $\Delta t = T$, mais alors ce ne sera pas possible à l'altitude z car $T' \neq \Delta t$. Autrement dit, si vous voulons à la fois maintenir un étiquetage unique des points d'espace-temps et l'effet de décalage vers le rouge

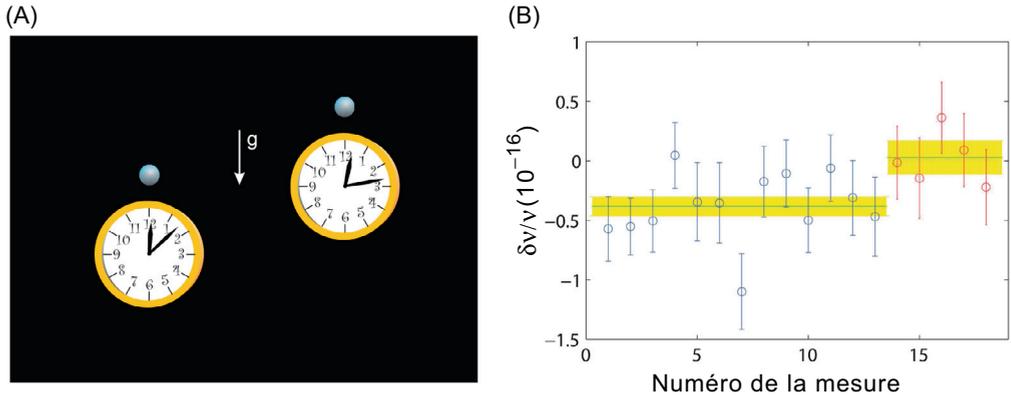


FIGURE 5.3. Mesure du décalage gravitationnel vers le rouge par une horloge optique à ion aluminium. L'horloge de droite est plus haute de 30 cm que celle de gauche et elle avance par rapport à cette dernière. Les points bleus correspondent à l'horloge de gauche, les points rouges à celle de droite. Adapté de C. Chou, D. Hume, T. Rosenband and D. Wineland, *Optical clocks and relativity*, *Science*, **329**, 1630 (2010). Courtoisie de David Wineland et du NIST.

gravitationnel, alors l'intervalle de temps propre $\Delta\tau$ à l'altitude z doit être relié à l'intervalle de temps coordonnée Δt par $\Delta\tau = \Delta t(1 + gz/c^2)$. Dans la situation de la figure 5.3, nous constatons que le temps ne s'écoule pas de façon unique dans une même pièce. Soulignons à nouveau que Δt est un intervalle de temps coordonnée, alors que T et T' sont des intervalles de temps propre, chacun étant mesuré par une horloge différente. Une conséquence de notre raisonnement est qu'il est impossible de tracer dans l'espace-temps un rectangle de côtés z suivant l'axe d'espace et de côtés Δt suivant l'axe du temps en utilisant des temps physiques, comme nous pourrions le faire dans un espace ordinaire, ou espace euclidien. L'espace-temps de la relativité générale n'est pas un espace euclidien, c'est un espace courbe.

5.2 Application au GPS

Pour expliquer le principe du GPS (*Global Positioning System*), nous allons partir d'un exemple simplifié où l'observateur peut se déplacer uniquement sur une droite (figure 5.4). Deux émetteurs de signaux électromagnétiques envoient à partir de satellites des signaux codés, de sorte que l'observateur, lorsqu'il reçoit le signal, peut obtenir à la fois l'information sur l'instant t_A où le signal a été émis et sur la position x_A de l'émetteur au moment de cette émission. Si l'observateur disposait d'une horloge atomique, il pourrait mesurer exactement le temps t de réception du signal et il en déduirait alors sa position x en fonction

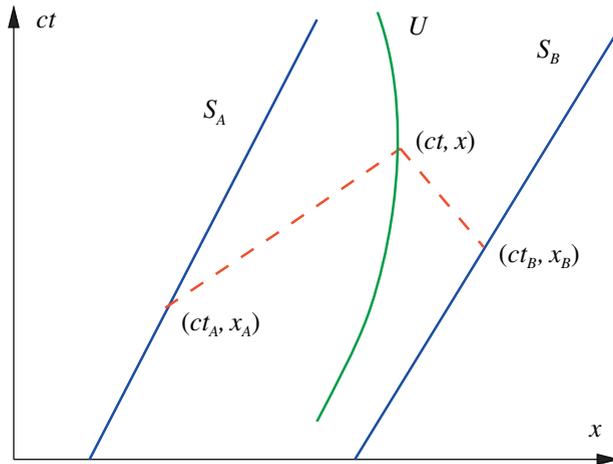


FIGURE 5.4. Explication simplifiée du GPS. Les lignes d'univers S_A et S_B des deux satellites sont les droites bleues, celle de l'observateur U est la courbe verte, et les lignes d'univers des photons sont représentées par des tirets rouges.

tout simplement du temps mis par la lumière pour voyager de l'émission en x_A à la réception en x :

$$x - x_A = c(t - t_A).$$

Mais ce n'est évidemment pas très commode de transporter avec soi une horloge atomique, par exemple dans une randonnée en montagne ! L'horloge du récepteur GPS n'est pas assez précise pour une mesure directe de la distance $(x - x_A)$ au point d'émission x_A , mais on s'en sort en utilisant deux satellites, car dans ce cas on dispose de deux équations pour deux inconnues, t et x . Les droites de lumière issues des deux émetteurs se croisent au point d'espace-temps (t, x) ; un calcul simple donne ces deux inconnues, et en particulier la position recherchée en fonction des instants d'émission t_A et t_B et des positions correspondantes x_A et x_B :

$$x = \frac{1}{2}[c(t_B - t_A) + (x_B + x_A)].$$

Si l'on passe d'un problème à une dimension (une droite) à un problème à trois dimensions de notre espace ordinaire, il faudra quatre satellites au lieu de deux pour déterminer la position d'un point sur la Terre, à savoir sa latitude, sa longitude et son altitude. C'est une constellation de 24 satellites qui envoie les signaux codés, de telle sorte qu'un point sur la Terre ait toujours en vision directe au moins quatre de ces satellites (figure 5.5). La position des satellites est

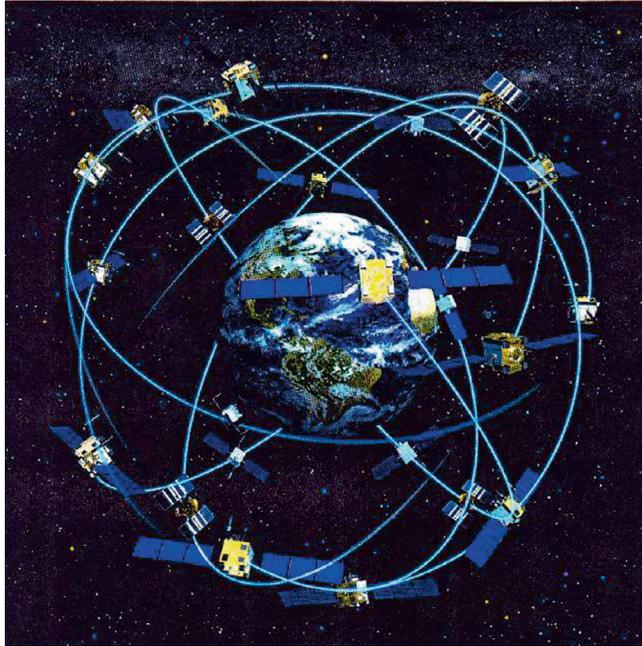


FIGURE 5.5 Constellation de satellites GPS.

connue à chaque instant par rapport à un référentiel dont l'origine est choisie au centre de la Terre. Le repérage à chaque instant de la position exacte des satellites est un problème simple en théorie, mais complexe en pratique, qui sort du cadre de cet ouvrage.

Quelle est la précision requise sur les horloges pour obtenir la position à un mètre près ? En une nanoseconde la lumière parcourt 30 cm et il faut donc limiter la dérive δt sur le temps à moins de 3 ns. La dérive δt des horloges atomiques à rubidium sur un intervalle de temps t est de l'ordre de $\delta t/t \simeq 10^{-13}$. On atteint une dérive $\delta t = 3$ ns en 10 heures environ : on doit donc régler les horloges des satellites au minimum toutes les 10 heures. Cela se fait au moyen d'un réseau d'horloges atomiques disposées sur la Terre, qui sont synchronisées entre elles et définissent un temps atomique international (TAI, § 6.5). Mais il faut en plus tenir compte de divers effets qui rendraient le GPS inopérant s'ils n'étaient pas corrigés. Nous allons estimer ci-dessous les dérives des horloges dues aux deux relativités. On montre en physique newtonnienne que le rayon r de l'orbite supposée circulaire d'un satellite (figure 5.6) et la vitesse v sur cette orbite sont reliés par :

$$\frac{v^2}{c^2} = \frac{GM}{c^2} \frac{1}{r}, \quad (5.4)$$

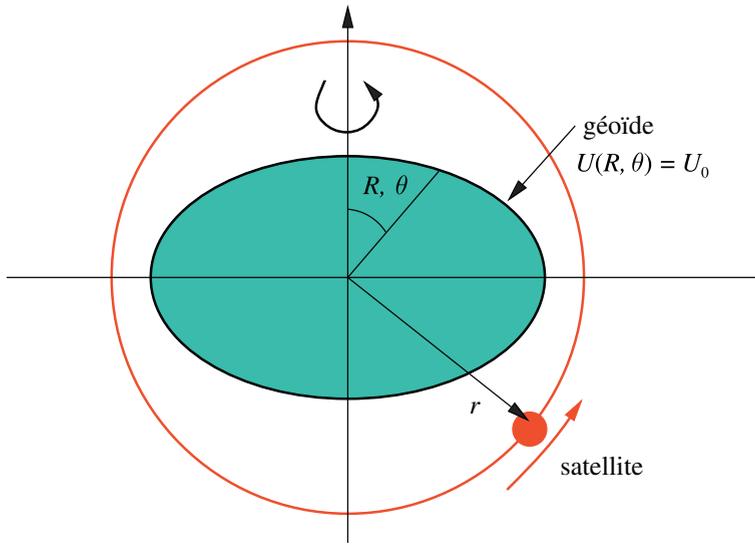


FIGURE 5.6. La Terre (en bleu) et l'orbite circulaire de rayon r d'un satellite (en rouge). Le potentiel gravitationnel de la Terre $U(R, \theta)$ est constant sur le géoïde. La déformation de la Terre par rapport à une sphère a été très exagérée.

où G est la constante de gravitation et M la masse de la Terre. On trouve numériquement que $GM/c^2 \simeq 4,4 \times 10^{-3} \text{ m}$.

On distingue deux types de corrections relativistes au GPS : la première vient de la relativité restreinte, la seconde de la relativité générale. Examinons d'abord l'effet de la relativité restreinte, la transformation relativiste du temps. Les satellites du GPS ont une période de 12 heures, une orbite de rayon $r = 4,16$ fois le rayon de la Terre, soit $r = 26\,561 \text{ km}$, et une vitesse $v \simeq 3,87 \text{ km/s}$, telle que $v/c \simeq 1,3 \times 10^{-5}$, d'où un facteur de Lorentz $\gamma - 1 \simeq 8,5 \times 10^{-11}$ qui donne la correction de relativité restreinte pour une horloge voyageant dans le satellite. Si aucune correction n'est appliquée, on atteint la limite fatidique des 3 ns en 30 s !

L'autre effet relativiste a pour origine la gravitation. Nous avons vu que la correction δt sur un intervalle de temps t entre l'altitude zéro et l'altitude z valait $\delta t/t = gz/c^2$, où gz est l'énergie qu'il faut dépenser pour monter une masse unité de l'altitude zéro à l'altitude z : cette énergie est le *potentiel gravitationnel* à l'altitude z lorsque l'accélération de la pesanteur est constante. Cette formule n'est valable que si z est très petit par rapport au rayon terrestre R : $z \ll R$. Dans le cas général, l'énergie que l'on doit dépenser pour monter une masse unité depuis la surface de la Terre jusqu'à un point situé à une distance r du centre est $GM(1/R - 1/r)$ dans le cas d'une Terre parfaitement

sphérique : $U(r) = -GM/r$ est le potentiel gravitationnel à une distance r du centre. Mais pour relier le temps propre d'une horloge terrestre au temps propre d'une horloge dans un satellite, il faut tenir compte d'une part de ce que la Terre n'est pas exactement sphérique, elle a la forme d'un ellipsoïde aplati, d'autre part qu'elle tourne sur elle-même, les deux phénomènes n'étant pas indépendants. La formule (5.5) reliant les temps propres est donnée explicitement et commentée dans l'encadré 5.2 ; avec les valeurs numériques données ci-dessus pour l'orbite des satellites, elle montre que l'horloge du satellite bat plus vite que l'horloge terrestre par un facteur $4,47 \times 10^{-10}$, ce qui se traduit par un décalage de $38,6 \mu\text{s}$ par jour. Pour démontrer la formule (5.5), il est commode de passer par l'intermédiaire du référentiel terrestre géocentré, dont l'origine est prise au centre de la Terre et dont les axes gardent une direction fixe dans l'espace : ils ne suivent pas la rotation de la Terre. Il faut observer qu'aucune horloge ne pourrait mesurer le temps dans ce référentiel : nous avons là un exemple évident d'un temps coordonnée, un temps abstrait utile comme intermédiaire pour étiqueter les points de l'espace-temps, mais qui ne correspond à aucune mesure physique.

Encadré 5.2. Temps propre dans un satellite et sur la Terre.

Si l'on veut être précis, il est nécessaire de tenir compte d'une part de ce que la Terre n'est pas exactement sphérique, d'autre part de la rotation de la Terre sur elle-même. Le potentiel gravitationnel $U(R, \theta)$ à une distance R du centre de la Terre et à une latitude θ est approximativement :

$$U(R, \theta) = -\frac{GM}{R} \left[1 - \alpha \cos^2 \theta \right],$$

où α est le rapport de la différence entre le rayon équatorial $R_{\text{Eq}} (\theta = \pi/2)$ et le rayon au pôle, $R_{\text{pôle}} (\theta = 0)$ divisée par R_{Eq}

$$\alpha = \frac{R_{\text{Eq}} - R_{\text{pôle}}}{R_{\text{Eq}}} \simeq 0,0034.$$

À la surface de la Terre, U doit être constant quelle que soit la latitude, $U(R, \theta) = U_0$, sinon la croûte terrestre se déformerait ; cela implique que le rayon de la Terre dépend de θ . L'altitude zéro est définie par le niveau moyen des océans et la surface correspondante est appelée *géoïde*. Le *temps atomique international* (TAI) est défini par une horloge atomique sur le géoïde. En revanche, l'accélération de la pesanteur à la surface de la Terre, g , dépend elle de la latitude. Elle vaut $9,832 \text{ m.s}^{-2}$ au pôle et $9,780 \text{ m.s}^{-2}$ à l'équateur. Le fonctionnement d'une horloge atomique est régi par U et deux horloges atomiques (ou deux montres à quartz) à l'altitude zéro au pôle et à l'équateur resteront synchronisées car la valeur de U est la même. En revanche, une horloge à balancier retardera à l'équateur par rapport à une horloge du même type

au pôle, car les battements dépendent de g . Les intervalles de temps propre dans un satellite d'orbite circulaire de rayon r et celui d'une horloge atomique à la surface de la Terre sont liés par :

$$\Delta\tau_{\text{satellite}} = \left(1 - \frac{1}{c^2} \left[\frac{3}{2} \frac{GM}{r} + U_0 \right] \right) \Delta\tau_{\text{horloge}} . \quad (5.5)$$

Le facteur $3/2$ du terme associé au satellite se décompose en $1/2$ venant de la relativité restreinte et 1 venant de la relativité générale.

5.3 Le temps en gravité forte

Nous avons donné dans l'équation (5.4) l'expression du rapport v^2/c^2 dans le cas d'une orbite circulaire de rayon r : ce rapport mesure le caractère relativiste du problème physique considéré. Il dépend de la constante de gravitation G , de la masse M et du rayon de l'orbite. Cet rapport a été obtenu dans le cadre de la mécanique de Newton et il n'est valable *a priori* que s'il est suffisamment petit : $v^2/c^2 \ll 1$. Lorsque $GM/(rc^2) \sim 1$, où \sim signifie « de l'ordre de », les effets de la relativité ne peuvent plus être ignorés. Pour fixer les idées, prenons le cas d'une étoile à neutrons : la masse d'une étoile à neutrons est de l'ordre de la masse du Soleil et son rayon r_0 de l'ordre de 10 km, ce qui correspond à une densité considérable, supérieure à celle d'un noyau atomique. Si $GM/(r_0c^2) \sim 1$, les effets relativistes vont devenir très importants au voisinage de cette étoile. Il est commode d'exprimer le rapport v^2/c^2 en fonction du rayon de Schwarzschild r_S de l'étoile à neutrons :

$$r_S = \frac{2GM}{c^2} \quad \text{soit} \quad \frac{v^2}{c^2} = \frac{r_S}{2r} , \quad (5.6)$$

où le facteur 2 dans l'expression de r_S résulte d'un calcul de relativité générale effectué pour la première fois en 1916 par l'astronome Karl Schwarzschild. La formule précédente est valable pour toute masse M , pas seulement pour une étoile à neutrons ; pour une telle étoile, le rapport r_S/r_0 est voisin de l'unité, environ 0,5. Les effets relativistes seront importants pour des orbites telles que $r_S/r \sim 1$, et ce rapport est aussi une mesure de l'intensité de la gravité, qui est très forte au voisinage d'une étoile à neutrons. Au contraire, on pourra considérer que la gravité est faible si le rayon de l'orbite est suffisamment grand, tel que $r_S/r \ll 1$. Dans le cas de la Terre, nous l'avons vu, $r_S \simeq 8,8 \times 10^{-3}$ m, et la gravité est toujours très faible, puisque le rayon minimum des orbites est le rayon terrestre, $r_0 \simeq 6,4 \times 10^6$ m, soit $r_S/r_0 \sim 1,4 \times 10^{-9}$. La gravité est également

faible dans le cas du Soleil, dont le rayon de Schwarzschild vaut 3×10^3 m, alors que son rayon est de 7×10^8 m, soit $r_S/r_0 \simeq 4 \times 10^{-6}$.

Compte tenu de la densité extrêmement élevée de l'étoile à neutrons, le décalage vers le rouge gravitationnel atteint des valeurs très importantes. Si grâce à des rétrofusées un astronaute se maintient immobile au voisinage de la surface de cette étoile à une distance r de son centre – en supposant qu'il survive à une gravité aussi considérable – et envoie vers la Terre des signaux périodiques de période T , on montre en relativité générale que la périodicité T' des signaux reçus par un observateur terrestre est alors :

$$T' = \frac{T}{\sqrt{1 - r_S/r}}. \quad (5.7)$$

Pour fixer les idées, prenons $r_0 = 2r_S$, ce qui est une valeur typique pour une étoile à neutrons, et supposons l'astronaute juste au-dessus de la surface, $r \gtrsim r_0$. La valeur observée de la périodicité des signaux est alors multipliée par $\sqrt{2} \simeq 1,41$. Pour un observateur terrestre, les mouvements de l'astronaute sont ralentis de 40% !

La formule précédente montre que le décalage vers le rouge gravitationnel tend vers l'infini lorsque $r \rightarrow r_S$: le temps de l'observateur terrestre se dilate indéfiniment, les mouvements de l'astronaute lui semblent figés. Dans le cas d'une étoile à neutrons, r est toujours supérieur à r_0 et on ne peut évidemment pas atteindre le rayon de Schwarzschild. Mais on peut atteindre r_S dans le cas des *trous noirs*, où l'on observe un cas extrême de distorsion relativiste du temps. En effet, contrairement à une étoile, un trou noir n'a pas de surface matérielle : un trou noir à symétrie sphérique, ou trou noir de Schwarzschild, est entouré d'un *horizon* qui est une sphère dont le centre coïncide avec celui du trou noir et dont le rayon est égal au rayon de Schwarzschild r_S . Cet horizon est une surface immatérielle, rien de particulier ne se passe quand on le traverse, du moins si on néglige tout effet quantique. La propriété essentielle de l'horizon est de piéger les rayons lumineux (ou les photons), et plus généralement tout type de particules situées à l'intérieur de cet horizon. Aucun photon, aucune particule massive ne peut traverser l'horizon pour s'échapper à l'infini : traverser l'horizon n'est possible que dans un seul sens, de l'extérieur vers l'intérieur du trou noir, sans espoir de retour (figure 5.7).

Si, en utilisant des rétrofusées, un astronaute se maintient immobile à une distance r du centre du trou noir et émet des signaux périodiques de périodicité T , la périodicité des signaux reçus par un observateur terrestre est donnée par la formule (5.7). Comme dans le cas de l'étoile à neutrons, l'observateur terrestre constate que les mouvements de l'astronaute lui paraissent ralentis, et cela d'autant plus que ce dernier s'approche de l'horizon. Finalement,

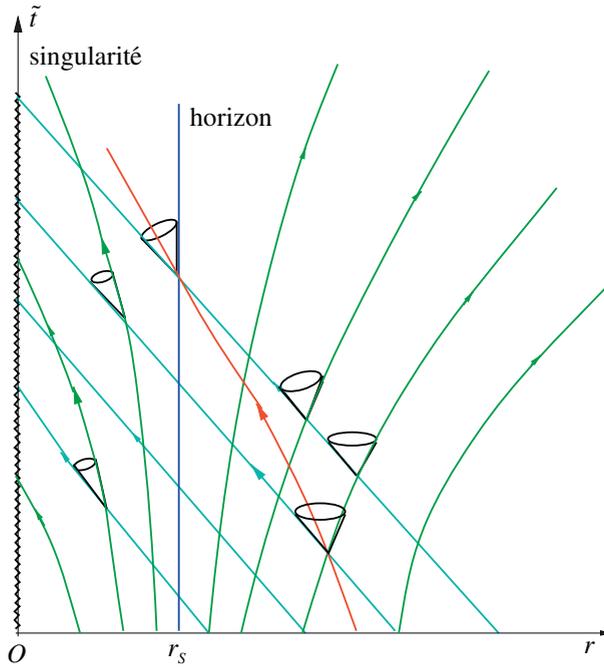


FIGURE 5.7. Lignes d'univers de photons dans des coordonnées particulières, celles d'Eddington-Finkelstein (\tilde{t}, r) : l'espace-temps est supposé à deux dimensions, une de temps \tilde{t} et une d'espace r . Les lignes d'univers des photons sortants sont tracées en vert et celles des photons entrants en bleu. On remarque un cône de lumière tangent à l'horizon. Les lignes d'univers des photons entrants sont des droites et les cônes de lumière sont inclinés par rapport à la verticale. Pour $r < r_S$, l'inclinaison des cônes de lumière piège les photons émis en un point $r < r_S$ à l'intérieur de l'horizon. La courbe en rouge représente la ligne d'univers d'un astronaute en chute libre vers le centre du trou noir.

le décalage vers le rouge gravitationnel tend vers l'infini lorsque r tend vers r_S . Si l'astronaute éteint ses rétrofusées et se laisse tomber en chute libre vers le centre du trou noir, il atteindra l'horizon en un temps fini, par exemple quelques secondes selon sa montre avec un choix convenable de r , mais pour l'observateur terrestre, la chute vers l'horizon dure un temps infini. Une fois passé l'horizon, aucun signal ne peut plus atteindre la Terre, mais l'astronaute ne ressent rien de particulier à la traversée de l'horizon. Cependant il finira déchiré par les forces de marée en s'approchant du centre du trou noir.

Le schéma des cônes de lumière permet de comprendre la structure d'un trou noir. Pour simplifier, nous considérons uniquement les rayons lumineux radiaux, c'est-à-dire d'une part les rayons qui convergent vers le centre du trou noir en se propageant dans le sens des r décroissants, ou rayons entrants, d'autre part ceux qui sont émis suivant une droite passant par le centre et se propageant

dans le sens des r croissants pour $r > r_S$, ou rayons sortants. Le schéma de la figure 5.7 est tracé dans un plan (\tilde{t}, r) , où r est la coordonnée radiale et \tilde{t} une coordonnée de temps. Il faut préciser que \tilde{t} est une coordonnée de temps et pas du tout un temps propre mesuré par une horloge. De fait, la figure est tracée avec des coordonnées appelées coordonnées d'Eddington-Finkelstein, qui sont commodes mais tout à fait arbitraires. La coordonnée radiale r n'est pas la distance au centre, mais en revanche l'aire d'une sphère de rayons r est bien $4\pi r^2$. Les photons entrants se propagent suivant des droites sans ressentir d'aucune manière la traversée de la ligne $r = r_S$. En revanche, si les photons se propagent de gauche à droite, ou « photons sortants », s'échappent sans problème à l'infini lorsqu'ils sont émis avec une coordonnée $r > r_S$, ceux qui sont émis avec $r < r_S$ sont piégés à l'intérieur de la région $r < r_S$ en raison de l'inclinaison des cônes de lumière et rejoignent le centre du trou noir. Il en est de même pour toute particule massive, dont la ligne d'univers doit se trouver à l'intérieur de ces cônes. Une fois passé $r = r_S$, ni les particules massives ni les photons ne peuvent s'échapper à l'infini : l'horizon, représenté dans ce diagramme par la droite $r = r_S$, est une ligne de non-retour. Cet horizon est tissé par des rayons lumineux radiaux particuliers, ceux qui d'une part ne peuvent pas s'échapper à l'infini, d'autre part ne rejoignent pas le point $r = 0$ et font donc du surplace. Toutefois cet énoncé doit être pris avec précaution, car il n'est valable que si l'on utilise les coordonnées de la figure 5.7 qui, rappelons-le, n'ont pas de signification physique directe. Un astronaute en chute libre radiale dans un trou noir mesurerait une vitesse de la lumière égale à c , qu'il soit situé à l'extérieur, au passage ou à l'intérieur de l'horizon. On lit parfois qu'en relativité générale la vitesse de la lumière ne vaut pas toujours c , mais c'est incorrect ! Afin de simplifier l'exposé, nous nous sommes limités ci-dessus aux trous noirs sphériques, mais il existe des trous noirs non sphériques, les trous noirs en rotation de Kerr, dont la structure est plus complexe. Par exemple, l'horizon des trous noirs de Kerr n'est pas défini par leur rayon de Schwarzschild.

Encadré 5.3 : métrique.

Nous avons introduit la métrique de Minkowski dans l'encadré 4.2. Cette métrique, qui définit une « distance » dans cet espace, est celle d'un espace-temps plat, alors que, nous l'avons vu, l'espace-temps de la relativité générale est courbe. On peut tenir compte de cette courbure en introduisant une métrique sur l'espace-temps. Nous pouvons par exemple reproduire les résultats du § 5.1 en écrivant une métrique qui se réduit à celle de Minkowski (4.5) en l'absence de gravitation ($g = 0$) :

$$c^2(\Delta\tau)^2 = c^2 \left(1 + \frac{2gz}{c^2} \right) (\Delta t)^2 - (\Delta z)^2. \quad (5.8)$$

En effet, nous retrouvons immédiatement la relation entre l'intervalle de temps propre $\Delta\tau$ et l'intervalle de temps-coordonnée Δt lorsque $\Delta z = 0$ pour deux événements se passant au même point d'espace :

$$\Delta\tau = \left(1 + \frac{g_z}{c^2}\right) \Delta t.$$

Dans le cas d'une symétrie sphérique, la métrique engendrée par la présence d'une masse M , qui est l'analogie du potentiel gravitationnel newtonien de cette masse, est donnée par la *métrique de Schwarzschild* dans un système de coordonnées particulier (t, r)

$$c^2(\Delta\tau)^2 = c^2 \left(1 - \frac{r_S}{r}\right) (\Delta t)^2 - \left(1 - \frac{r_S}{r}\right)^{-1} (\Delta r)^2. \quad (5.9)$$

Le temps propre d'un observateur situé en r est relié au temps coordonnée t par :

$$\Delta\tau = \Delta t \sqrt{1 - r_S/r},$$

ce qui explique (5.7). La forme (5.9) de la métrique n'est valable que pour $r > r_S$. Pour écrire une métrique valable pour tout r positif, il faut utiliser un autre système de coordonnées, par exemple celui (\bar{t}, r) d'Eddington-Finkelstein.

En résumé, la relativité générale conduit en gravité forte à des distorsions du temps qui peuvent être considérables, et même infinies dans le cas d'un trou noir. Le choix d'une coordonnée de temps, et plus généralement d'un système de coordonnées pour repérer un point de l'espace-temps, est largement arbitraire et souvent dicté par des considérations de commodité. Il faut se souvenir que les coordonnées ne sont rien d'autre que des *étiquetages* des points de l'espace-temps, et le temps coordonnée t n'est pas un temps physique, mesuré par une horloge. Les seuls temps physiques sont les temps propres mesurés par des observateurs utilisant des horloges qui leur sont liées, et ils peuvent être reliés au temps coordonnée grâce à la métrique (encadré 5.3).

5.4 Le temps cosmologique

La cosmologie contemporaine est fondée sur le *principe cosmologique*, lequel postule que l'Univers est homogène et isotrope : il a les mêmes propriétés en tout point et dans toutes les directions. Si on lisse les inhomogénéités évidentes (étoiles, galaxies) en faisant une moyenne sur des volumes suffisamment grands, ce principe est compatible avec les observations, et il est essentiel pour l'application de la relativité générale à la cosmologie. En effet, on montre mathématiquement qu'en raison de ce principe l'espace-temps peut être décomposé en une famille d'espaces à trois dimensions étiquetés par un paramètre que l'on

peut appeler le temps t , le *temps cosmologique*. On découpe donc l'espace-temps en une série de « tranches » Π_t où le temps prend une valeur constante. De plus, les espaces à trois dimensions homogènes et isotropes sont de trois types, et trois seulement :

1. l'espace euclidien à trois dimensions ;
2. la sphère S^3 à trois dimensions dans un espace à quatre dimensions (la sphère ordinaire est notée S^2) ;
3. l'hyperboloïde H^3 à trois dimensions dans un espace à quatre dimensions.

Le cas le plus simple, qui est semble-t-il réalisé par la Nature, est celui de l'espace euclidien ordinaire, et nous nous y restreindrons. La figure 5.8 montrant le découpage de l'espace-temps en tranches est dessinée dans un espace-temps à trois dimensions, où les tranches qui sont donc à deux dimensions sont des plans ordinaires.

Considérons la ligne d'univers d'une galaxie (ou plus exactement d'un amas de galaxies) qui traverse le plan Π_t au temps t (figure 5.8) : étant soumise

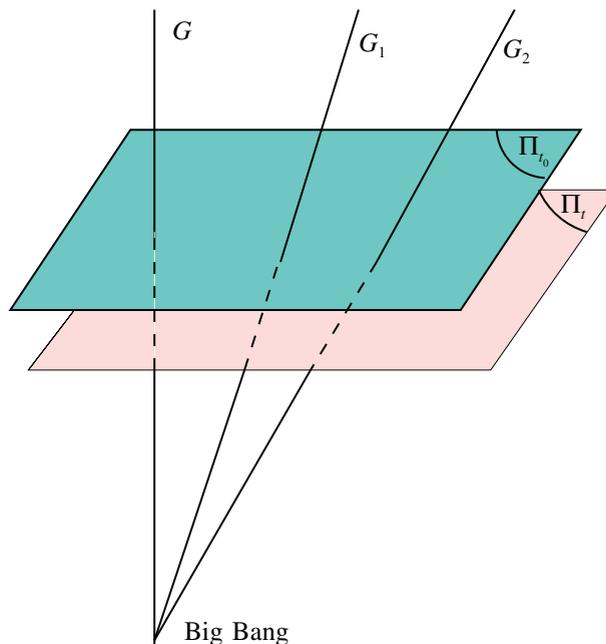


FIGURE 5.8. Plans Π_t et lignes d'univers (géodésiques) de trois galaxies (ou plus exactement amas de galaxies) G , G_1 et G_2 . Les lignes d'univers des galaxies sont orthogonales aux plans $t = \text{cste}$ pour la métrique de la relativité générale choisie dans cette figure. On observe l'expansion de l'Univers entre t et t_0 : les galaxies s'éloignent les unes des autres en partant du Big Bang.

uniquement à des forces de gravitation, elle est en chute libre ou, en termes plus techniques, elle suit une géodésique de l'espace-temps. Pour chaque galaxie, on est libre de choisir comme coordonnée de temps son temps propre $t = \tau$ et, en raison de l'homogénéité, ce temps peut être choisi identique pour toutes les galaxies. Ce temps est celui mesuré par l'horloge d'un observateur voyageant avec la galaxie, et on peut se le représenter intuitivement comme le temps indiqué par une horloge démarrant à $t = 0$ à l'instant du Big Bang : en raison du principe cosmologique, il existe une définition naturelle du temps. La notion « d'instant zéro » sera cependant réexaminée ci-dessous.

Le temps cosmologique t que nous venons d'introduire est bien défini mathématiquement en raison du principe cosmologique, mais quelle est sa signification physique ? Pour des temps pas trop reculés, on peut imaginer que ce temps est simplement celui indiqué par une horloge atomique voyageant avec notre galaxie. Le modèle du Big Bang chaud aujourd'hui universellement adopté prédit que si l'on remonte dans le temps, la température de l'Univers devient de plus en plus élevée. Elle est par exemple de 3 000 K environ 380 000 ans après le Big Bang, et aucune horloge ne pourrait résister quand on s'approche du temps $t = 0$ où la température tend vers l'infini. Par commodité, on introduit un « âge de l'Univers » et donc un « instant zéro » pour le temps. Mais en fait l'instant $t = 0$ ne fait pas partie de l'histoire de l'Univers : cet instant correspond à une singularité de la métrique où la courbure de l'espace-temps devient infinie. On peut d'ailleurs utiliser d'autres coordonnées de temps, par exemple le logarithme $\ln t$ qui tend vers $-\infty$ aux débuts du Big Bang. De plus, toute théorie du Big Bang pour des temps antérieurs à $t \sim 10^{-11}$ s, même largement médiatisée, est au minimum spéculative (inflation), voire hautement spéculative (univers quantiques, multivers).

L'introduction d'un temps cosmologique conduit à se poser à nouveau la question d'un référentiel privilégié « naturel », d'un retour à l'espace-temps newtonien. Supposons qu'un point O dans un des amas de galaxies de la figure 5.8 soit immobile par rapport au fond diffus cosmologique. Ce n'est pas le cas de notre galaxie, qui se déplace à environ 600 km/s par rapport à ce fond diffus, ce qui se montre grâce à l'effet Doppler. On pourrait alors choisir O comme origine des coordonnées d'espace d'un référentiel et le temps propre associé à ce point comme temps universel. Ce serait une sorte de retour à l'« éther » avec un espace et un temps absolus mais, comme dans le cas du référentiel terrestre géocentré mentionné ci-dessus, aucune horloge ne serait susceptible de mesurer ce temps coordonné, ou fonction temporelle. Il serait certainement possible de se servir de ce référentiel comme intermédiaire pour traiter tous les problèmes de physique classique, validant ainsi le point de vue de Poincaré qui pensait que conserver l'éther était une commodité dont on pouvait ou non se passer.

Cependant, une telle stratégie serait extrêmement malcommode, et surtout elle obscurcirait considérablement l'intuition physique. De plus, en théorie relativiste des champs quantiques, il est essentiel de conserver l'invariance relativiste explicite, et utiliser un tel référentiel rendrait les calculs inextricables, voire impossibles.

6

Le standard de temps universel

6.1 Connaître le temps sur la Terre

Dans la plupart des activités humaines, il est nécessaire de connaître le temps, avec une précision variable selon l'activité, mais qui peut être très grande dans certains cas. Il est clair aujourd'hui, dans notre société d'une complexité croissante avec ses réseaux de communication planétaires, qu'un standard de temps plus ou moins universel est indispensable. Établir ce *temps universel* a été une entreprise relativement récente, à cheval sur la fin du XIX^e et le début du XX^e siècle. Les méthodes utilisées pour ce faire sont variées et en progrès constant. Elles font l'objet du présent chapitre, qui donne aussi une idée de leur utilisation dans la vie courante, en recherche fondamentale et en astronomie.

Rappelons tout d'abord que la Terre est divisée en 24 zones de standard de temps, les *fuseaux horaires*, dont chacun s'étend sur environ 15 degrés de largeur en longitude. La zone zéro est centrée sur une ligne, le *méridien de Greenwich*, qui s'étend du nord au sud en passant par Greenwich en Grande-Bretagne, où fut situé l'un des plus prestigieux observatoires du monde occidental qui abrite aujourd'hui un superbe musée de l'histoire du temps. Le découpage en zones est en réalité assez complexe et en constante évolution, parfois conflictuelle, compte tenu des évidentes implications économiques de la délimitation des territoires et de leurs frontières. Il est ainsi possible de gagner un jour en faisant le tour du monde par l'ouest, comme Phileas Fogg dans *Le Tour du Monde en 80 jours* de Jules Verne. Notons aussi que certains pays adoptent deux heures différentes pour l'été et pour l'hiver à des fins d'économie d'énergie ; d'autres au contraire gardent la même heure toute l'année, ce qui n'est pas non plus sans causer des problèmes ; l'important reste que le temps soit bien défini pour une nation

entière, ou bien pour des fractions d'un grand territoire comme celui des États-Unis ou de la Russie.

Dans notre vie quotidienne, le temps peut signifier l'instant où un événement se produit, ou bien la durée qui s'écoule entre le début et la fin d'un processus. La mesure d'une durée en un point fixe sur la Terre s'effectue avec les montres ou les horloges décrites au chapitre 3. Elle peut atteindre une très grande précision. Rappelons que les meilleures horloges dans les observatoires peuvent mesurer l'écoulement du temps avec une précision allant jusqu'à 17 chiffres significatifs ! Pourtant cette précision n'est pas suffisante en elle-même pour ne pas rater un train si notre montre n'est pas synchronisée sur la pendule de la gare de départ, par exemple Paris. Et nous ne pouvons prévenir de notre heure d'arrivée en gare de Vladivostok que si les horloges en France et en Russie sont synchronisées, en tenant compte évidemment du décalage entre les fuseaux horaires. Ceci suppose le transfert du temps entre les horloges à distance et la synchronisation de ces horloges de plus en plus précises qui ont été décrites au chapitre 3.

Nous ne nous rendons pas toujours bien compte à quel point la connaissance du temps joue un rôle central dans nos activités. Nous portons en général une montre à quartz, qui dérive de moins d'une seconde par jour, à moins que, comme les jeunes générations, nous n'utilisions plus pour lire l'heure que nos téléphones portables dont l'exactitude est encore bien supérieure. Tous les appareils électroménagers de notre intérieur, téléviseurs programmables, capteurs de toutes les sortes, etc. comportent des horloges. Les systèmes de télécommunication par Internet implantés dans tous les foyers incluent, comme nous le verrons plus loin, des horloges qui déclenchent les impulsions laser se propageant dans les fibres et synchronisant les réseaux.

Les applications de la mesure du temps sont innombrables. Parmi les plus gros utilisateurs on compte les systèmes de navigation maritime et la coordination sur Terre des réseaux de circulation, qui nécessitent une très bonne synchronisation des horloges le long des trajets. Les atterrissages des avions sont pilotés à quelques mètres près, ce qui implique un chronométrage ultra-précis fourni par le GPS (*Global Positioning System*), maintenant amélioré par le système EGNOS (*European Geostationary Navigation Overlay Service*), inspiré du GPS différentiel, mais étendu localement à un ensemble de stations dont les données sont corrélées entre elles. Nous reviendrons plus loin sur le GPS et sur tout le fonctionnement du domaine spatial, qui repose de façon cruciale sur les mesures de temps et dont les applications civiles sont de plus en plus nombreuses. Dans des domaines scientifiques plus spécialisés, comme la métrologie et la spectroscopie haute résolution des atomes et des molécules, une bonne référence de temps est nécessaire pour caler les mesures de fréquence des outils laser. Citons enfin

l'astronomie et l'astrophysique, à l'origine des méthodes de repérage du temps, qui sont aujourd'hui de grandes utilisatrices des techniques les plus performantes pour leurs mesures et pour leurs instruments.

Moins connues du grand public mais extrêmement importantes sont les applications de la métrologie du temps au secteur industriel. Les produits de toute sorte qui circulent entre régions et pays ont besoin d'accréditations, qui sont des références garantissant les prestations sur une chaîne de distribution. La calibration des produits, par exemple en poids ou en volume, se réfère à un système de mesures étroitement dépendant de la précision du *Système d'unités international* (SI), le plus largement utilisé au monde. Celui-ci est fixé et redéfini tous les quatre ans à la conférence générale des poids et mesures à Paris. L'unité connue avec le plus de précision est l'unité de temps, la seconde. La définition de l'unité de longueur, le mètre, repose sur celle du temps, puisque depuis 1983 le mètre est la distance parcourue par la lumière dans le vide en une fraction $1/299\,792\,458$ de seconde. Rappelons que la vitesse de la lumière dans le vide est la même en tout point, ce que l'on sait depuis l'expérience de Michelson et Morley en 1887 montrant que la vitesse de la lumière était indépendante du mouvement de la Terre. Pour étalonner les longueurs, on a donc besoin d'une horloge. Plus cette horloge est précise, plus l'est aussi le calibrage des produits. On sera peut-être surpris que ce besoin de calibrage soit l'une des motivations principales pour les pays qui commencent à s'équiper d'une horloge atomique de haute précision à atomes froids. Mentionnons enfin la folie qui semble s'être emparée du système bancaire international, où les traders, ou plutôt leurs ordinateurs programmés à l'avance, effectuent des opérations financières dont les dates sont fixées avec une précision de l'ordre de la milliseconde, voire moins ! Il importe donc que leurs horloges soient synchronisées sur l'ensemble du globe. Les horloges les plus exactes étant en nombre limité, il faut pouvoir les utiliser comme instruments de référence et transférer leur exactitude à des horloges moins performantes.

6.2 Les transferts temps/fréquence (T/F)

6.2.1 Caractéristiques générales des liens T/F

Parallèlement à la construction d'horloges le plus exactes possible se pose la question du *transfert du temps*, ou de la fréquence qui lui est reliée, d'un lieu à un autre avec le minimum de dégradation de la précision des horloges ainsi connectées. Pour comparer deux horloges à distance, il faut établir un lien dit T/F (temps/fréquence), pour lequel les méthodes utilisées sont variées et ont évolué dans le temps. Le lien peut être optique au niveau de la Terre, à travers

des câbles ou des fibres optiques. Il peut aussi s'effectuer en construisant des horloges de haute précision transportables, comme l'horloge à atomes froids de césium construite par l'Observatoire de Paris, qui a été déplacée en camion vers des laboratoires de métrologie à Munich et a été comparée aux horloges sur place. Mais cette dernière solution n'est pas généralisable. Très souvent aujourd'hui le lien entre deux horloges sur Terre est aérien et utilise le rayonnement électromagnétique dans des gammes de longueur d'onde diverses.

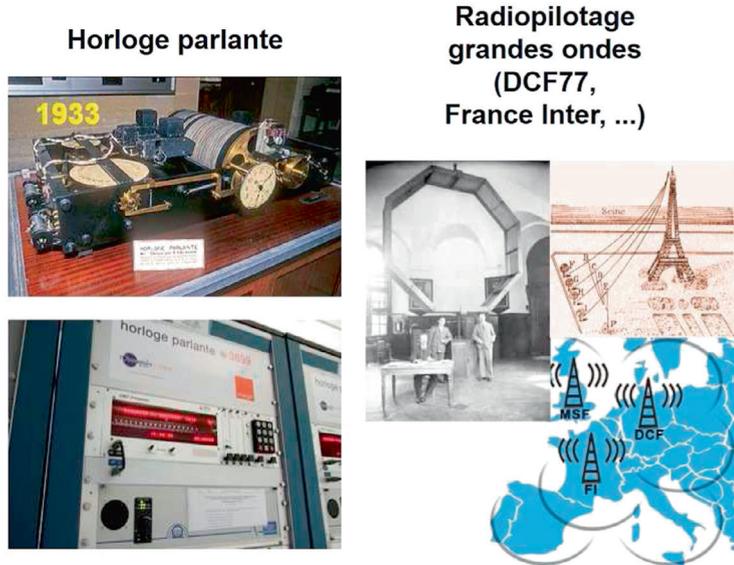


FIGURE 6.1. L'horloge parlante existe depuis 1933. Le transfert du temps s'effectue par voie aérienne dans le domaine radio. L'incertitude dans la synchronisation est de l'ordre de quelques millisecondes. Les émissions radiophoniques sont diffusées sur le même principe dans le domaine des grandes ondes, France-Inter à 162 kHz à partir d'Allouis dans le Cher, la radio allemande DCF77 à 77 kHz à partir de Mainflingen.

Le domaine radio, celui des grandes longueurs d'onde, est utilisé pour les émissions radiophoniques (France Inter émet à 162 kHz). Il a longtemps été émis à partir de la tour Eiffel (figure 6.1). Le domaine micro-onde est le plus utilisé pour l'ensemble des connexions par satellite ; il permet une précision plus grande que le domaine radio. Enfin le domaine des fréquences laser autorise une précision encore supérieure pour les transferts T/F, mais il est encore peu utilisé.

Le transfert de l'information entre deux horloges distantes peut s'effectuer de deux façons : par transfert de la fréquence ou transfert du temps. D'une part, on peut simplement comparer les fréquences de deux horloges distantes l'une de l'autre. On parle alors de syntonisation. Le procédé suppose que le lien n'introduit pas de fluctuations sur la phase du signal transmis. D'autre part, si l'on veut

transférer la mesure du temps d'une horloge à une autre à une certaine distance, il s'agit de synchronisation. Celle-ci implique non seulement la stabilité du lien entre les horloges, mais aussi la connaissance de la valeur absolue du délai de la transmission entre les deux horloges qui échangent des signaux. Le transfert dans ce second cas est plus difficile que dans le premier. Pour le transfert du temps entre deux horloges distantes, le lien peut être unidirectionnel : c'est la *méthode simple voie* (ou à une voie), qui suppose la connaissance du délai de propagation. Plus élaborée est la *méthode double voie* (ou à deux voies) avec aller-retour du signal (figure 4.6a), ce qui permet de s'affranchir de la connaissance du délai, mais suppose une bonne symétrie entre les parcours aller et retour. La figure 6.2a montre le double lien entre deux horloges sur la Terre situées en deux points qui ont un même satellite en vision directe. Cette méthode double voie ne s'applique pas au GPS, où les signaux sont envoyés à partir d'horloges situées sur différents satellites (voir ci-après).

Les signaux échangés sont des impulsions ou des signaux continus. Pour des signaux pulsés, il faut mesurer le délai entre l'impulsion envoyée et celle qui est reçue à distance (figure 6.2b). Les signaux continus sont eux enregistrés sur une onde porteuse à fréquence fixe élevée, de l'ordre de quelques GHz. Pour bien identifier le signal et ne pas le confondre avec d'autres signaux de même fréquence issus d'autres sources, on imprime sur la même porteuse, en plus du signal utile, un code de type bruit aléatoire (figure 6.2c). Ceci permet par décodage la reconnaissance du signal correct par le récepteur, à la condition que l'utilisateur connaisse ce code. Par exemple, dans le cas du GPS, ce système permet de ne pas mélanger les informations reçues par un même récepteur au sol en provenance des différents satellites.

6.2.2 Quelle heure est-il ?

L'information sur l'heure à l'usage du public est transmise grâce à la technique simple voie et au radio-pilotage par ondes hertziennes. L'horloge parlante existe depuis 1933. On peut obtenir l'heure par radio ou par téléphone. L'émission provient du laboratoire SYRTE à l'Observatoire de Paris où se trouve la référence pour le temps en France piloté par l'horloge à atomes froids de césium décrite au chapitre 3. La précision de la synchronisation est de l'ordre de dix millisecondes.

Il existe aujourd'hui une méthode alternative à l'horloge parlante qui permet d'obtenir l'heure sur son ordinateur en utilisant Internet. Elle utilise le système mondial NTP (*Network Time Protocol*), qui est un protocole permettant de synchroniser l'horloge locale d'ordinateurs sur une référence d'heure, par l'intermédiaire d'un réseau informatique et un algorithme approprié. Sa première version date de septembre 1985, la version 4 de 2004, la version 3 est

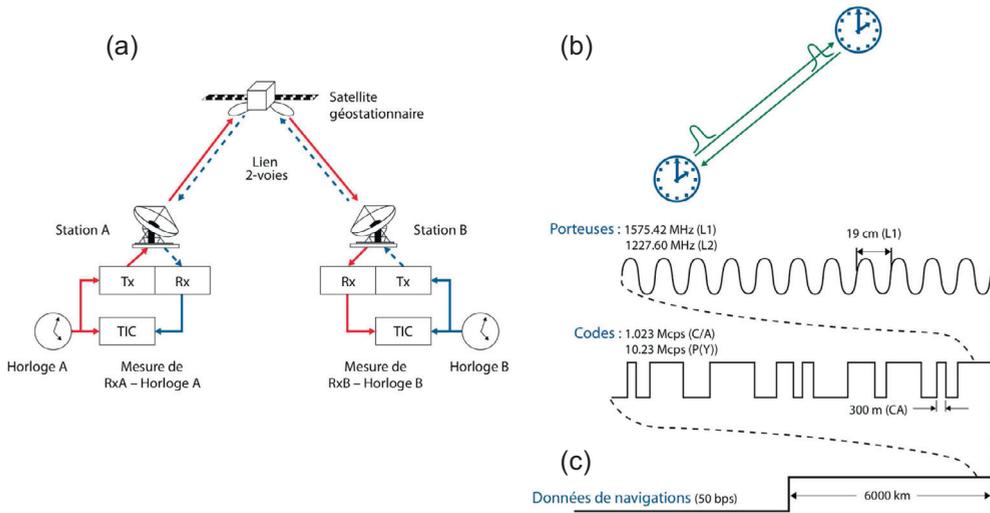


FIGURE 6.2. Différentes méthodes pour le transfert du temps. (a) Synchronisation entre deux horloges qui ont un même satellite en vision directe avec transfert deux voies, qui a d'une part l'avantage d'éviter la mesure du délai de propagation, d'autre part permet de corriger partiellement les erreurs introduites par la traversée de l'atmosphère. (b) Transfert simple voie d'une horloge à une autre avec des signaux pulsés ; la mesure du délai de propagation est alors requise. (c) Transfert simple voie d'un signal continu porté par une onde dans le domaine des courtes longueurs d'onde (fréquence de l'ordre du GHz), superposé sur la même porteuse à un signal de bruit aléatoire mais codé qui permet un repérage de l'origine du signal (utilisé pour les satellites du GPS).

la plus utilisée. Les améliorations sont constantes et dépendent aussi largement des utilisateurs. Les signaux transitent par les réseaux fibrés d'Internet à travers différentes strates de complexité croissante à l'échelle du monde entier. La première strate comporte 230 serveurs, la seconde strate 4 500, dont environ 1 000 aux États-Unis, qui se synchronisent sur les serveurs primaires. L'arborescence est verticale. La méthode est un système du type deux voies asymétriques (figure 6.3). Les horloges intermédiaires, dites secondaires, sont connectées de façon multiple déterminée par un algorithme. Les erreurs dans les étapes de la synchronisation sont de l'ordre de la milliseconde. Chaque nœud du serveur peut émettre en permanence une information sur l'heure comme l'horloge parlante. Il peut aussi répondre localement aux questions d'un nœud client, avec interopérabilité entre serveur et client. L'architecture de ce système est très souple, extensible et robuste, mais c'est aux utilisateurs d'y contribuer. Les perfectionnements possibles pour l'exactitude de l'ensemble dépendront beaucoup de l'amélioration des allers-retours entre deux horloges secondaires qui devraient être mieux symétrisés.

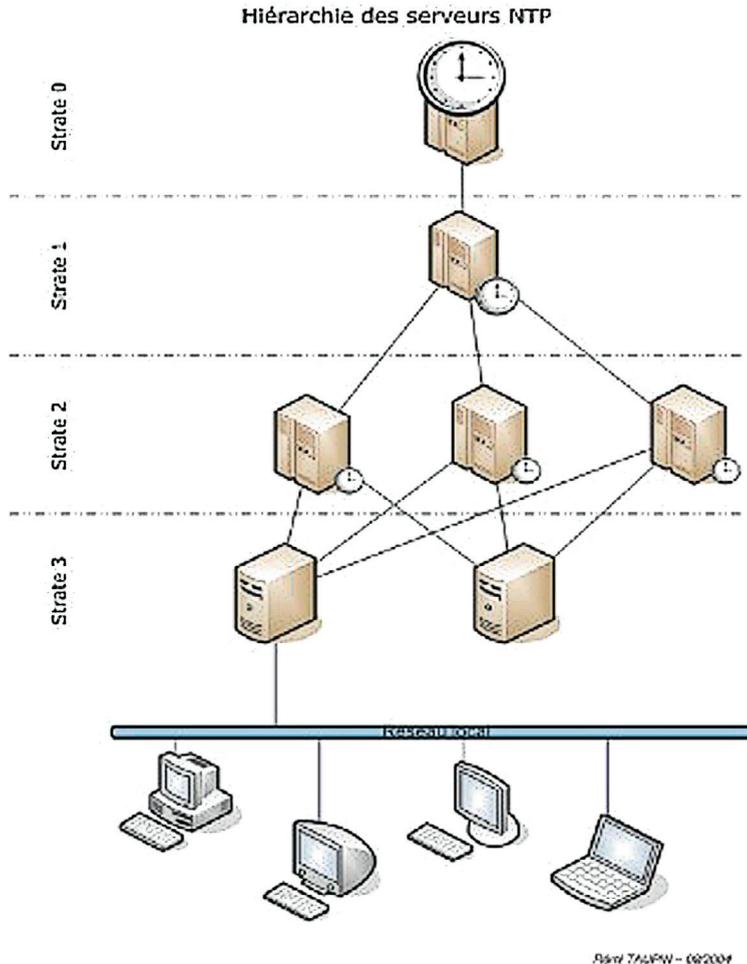


FIGURE 6.3. Distribution de l'heure pour les ordinateurs par l'intermédiaire d'Internet. Strate 0 : horloge maîtresse de référence. Strate 1 : horloge primaire synchronisée sur l'horloge maîtresse. Strates suivantes : horloges secondaires. Les traits représentent les connexions informatiques entre les horloges secondaires.

6.2.3 Le transfert T/F par satellites

Les méthodes précises de transfert T/F utilisent les satellites. Comme nous l'avons vu ci-dessus, elles peuvent être à une voie comme pour le GPS où les signaux partent de satellites, ou plus rarement à deux voies, ce qui évite d'avoir à mesurer des durées de propagation des signaux. La figure 6.4a montre comment peut s'opérer le transfert T/F entre deux horloges sur Terre, ayant en vision

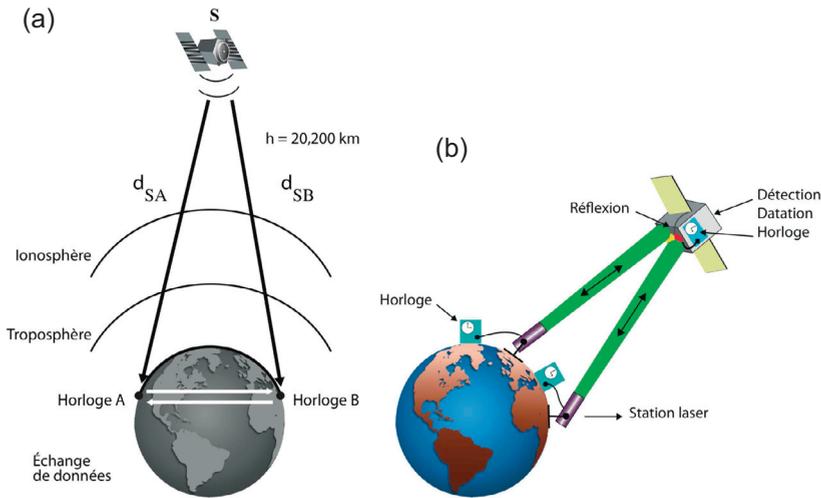


FIGURE 6.4. Transfert temps/fréquence entre deux horloges sur la Terre ayant un même satellite en vision directe. (a) La transmission est effectuée par l'intermédiaire d'ondes du domaine micro-onde (fréquence de l'ordre du GHz). (b) Le lien est optique, le transfert beaucoup plus précis qu'en (a).

directe un même satellite, ce dernier aussi pourvu d'une horloge. On mesure le temps T_S sur le satellite, T_A sur l'horloge A et T_B sur l'horloge B. On détermine $T_A - T_S$ en détectant en A le signal de temps envoyé par micro-onde depuis le satellite et en mesurant le délai de propagation entre le satellite et chacune des horloges. On procède de même avec l'horloge B et l'on en déduit $T_A - T_B$. On utilise en outre le GPS pour localiser les horloges qui échangent sur Terre des informations sur leur position. Les délais mesurés sont de l'ordre de la nanoseconde (10^{-9} seconde) et la précision relative du transfert de la fréquence est dans la gamme de 10^{-15} . Rappelons que la méthode alternative plus précise est à deux voies, avec aller-retour du signal de synchronisation entre chaque horloge et le satellite (figure 6.2a), ce qui permet de s'affranchir de la mesure du délai de propagation. Toutefois elle est plus complexe et de toute façon n'est pas applicable dans tous les cas ; par exemple elle ne fonctionne pas pour relier des horloges en des points de la Terre qui n'ont pas un même satellite en vision directe.

La méthode peut être encore améliorée si l'on utilise un lien dans le domaine optique plutôt que micro-onde. Le transfert T/F par l'intermédiaire de la lumière est en effet beaucoup plus précis, avec une erreur dans la gamme des picosecondes (10^{-12} seconde). La figure 6.4b illustre le transfert T/F entre les deux horloges sur Terre ayant en vision directe le satellite JASON. Il s'agit d'une mission du CNES initiée en 2009 qui doit relier l'Observatoire de Paris et

l'Observatoire de la Côte d'Azur. Cette voie très prometteuse en est encore au stade de l'expérimentation. Elle implique nombre de développements, en particulier de l'électronique associée à la détection des signaux qui doit pouvoir suivre, sans distorsion et sans rajouter du bruit, la précision extrême des signaux échangés. La figure 6.5 montre le faisceau laser de 400 W tiré vers le satellite JASON depuis le laboratoire du SYRTE à l'Observatoire de Paris. Le retour du faisceau a une puissance de quelques nanowatts.



FIGURE 6.5. Faisceau laser tiré du laboratoire SYRTE à l'Observatoire de Paris en direction du satellite JASON. Photographie de Pascal Blondet.

Il faut tenir compte des perturbations entraînées par la traversée de l'atmosphère en raison des effets de freinage du faisceau par les électrons et autres particules chargées qui entourent la Terre dans l'ionosphère. Des modèles permettent de les corriger en moyenne avec une précision de 50% environ. Pour obtenir une meilleure correction du délai lié à l'ionosphère, on peut utiliser simultanément deux fréquences différentes : le délai parasite étant inversement proportionnel au carré de la fréquence, on élimine cette erreur avec une efficacité de 99,9%. La troposphère engendre aussi des erreurs importantes que l'on corrige avec des modèles.

Il existe d'autres sources d'erreurs de nature plus fondamentale qui affectent la mesure de la durée du transfert entre le satellite et un récepteur au sol : celles liées à la relativité. Elles ont été évoquées au chapitre 5. Elles sont de deux types. Il faut d'abord prendre en compte l'effet dit (improprement) de dilatation des

temps si deux objets A et B sont en mouvement l'un par rapport à l'autre : c'est l'effet de relativité restreinte (RR). C'est le cas pour un satellite qui gravite sur son orbite autour de la Terre. Le décalage relatif de fréquence $\Delta\nu/\nu$ se calcule à partir de l'équation (4.4) :

$$\frac{\Delta\nu}{\nu} \Big|_{\text{RR}} \simeq \frac{v^2}{2c^2}, \quad (6.1)$$

où $v = v_A - v_B$ est la vitesse relative des deux objets, dans ce cas la Terre et le satellite, et c la vitesse de la lumière. Le second type d'erreur, dû à la relativité générale (RG), est appelé décalage vers le rouge gravitationnel, ou souvent *red shift* gravitationnel ; il est lié à la différence entre les potentiels de gravitation U_A et U_B des deux systèmes. Le déplacement relatif des fréquences est donné à partir de l'équation (5.2) :

$$\frac{\Delta\nu}{\nu} \Big|_{\text{RG}} \simeq \frac{1}{c^2}(U_A - U_B). \quad (6.2)$$

Il faut donc faire les corrections correspondantes pour toutes les mesures de transfert temps/fréquence. Ceci est particulièrement important pour toutes les applications à la navigation et pour le système de positionnement GPS universellement utilisé. Nous en donnons les ordres de grandeur plus loin.

6.2.4 Comment fonctionne le GPS ?

Nous venons de voir que le GPS (*Global Positioning System*) est un outil indispensable pour la mesure du temps, par exemple pour le transfert T/F entre deux horloges situées en des points différents de la Terre. Le GPS a été introduit brièvement au § 5.2 et nous donnons ci-dessous quelques détails complémentaires. De façon plus générale, la comparaison des horloges à l'échelle mondiale s'effectue avec le système GNSS (*Global Navigation Satellite System*) dont fait partie le GPS, initialement conçu et mis au point aux États-Unis à des fins de défense stratégique. Son importance pour la navigation sur Terre, sur mer et dans l'espace, ainsi que pour ses innombrables applications dans la vie courante, nous invite à donner plus de détails sur son fonctionnement.

Le système utilise l'émission d'ondes dans le domaine micro-onde émises à partir d'une constellation de satellites, qui sont reçues par des récepteurs basés sur la Terre (figure 5.5). La méthode est fondée sur la triangulation entre les signaux reçus à partir des satellites qui orbitent autour de la Terre en 12 h environ. Le principe consiste à mesurer la durée de propagation d'un signal électromagnétique entre chaque satellite et le récepteur. Pour un satellite situé par exemple à 25 000 km du centre de la Terre, la mesure du délai de propagation du signal à la vitesse de la lumière donne une localisation du récepteur sur une sphère ; avec

deux satellites, le récepteur sera localisé sur l'intersection de deux sphères, donc sur un cercle ; avec trois satellites, il sera localisé sur deux points, dont un seul est à retenir ; avec un quatrième, on corrige certaines erreurs et on réduit le flou dans la localisation. Le nombre minimum de satellites requis pour déterminer le positionnement du récepteur est de quatre, trois pour les trois coordonnées spatiales et le quatrième pour la coordonnée de temps. Un plus grand nombre de satellites est requis pour la couverture du monde entier, afin que chaque point sur la Terre ait au moins 4 satellites en vision directe. En réalité un système du type GNSS fait appel à un grand nombre de satellites, avec des programmes complexes pour réduire les erreurs (figure 6.6).

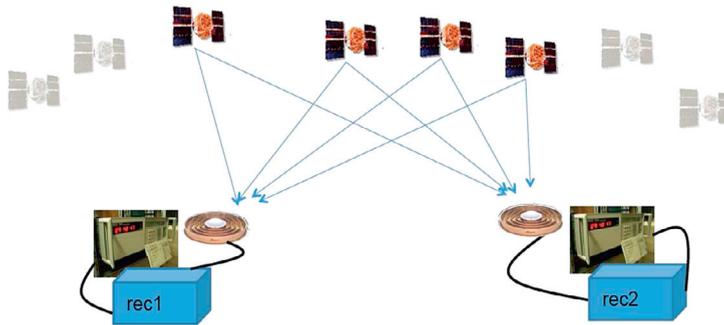


FIGURE 6.6. Principe du système de positionnement GNSS dont fait partie le GPS. Un ensemble de satellites, quatre au minimum, envoie des signaux dans le domaine micro-onde qui arrivent sur un même détecteur au sol sur la Terre.

Le GPS des États-Unis utilise 24 satellites et toujours 4 d'entre eux au moins sont visibles de n'importe quel point sur la Terre. Le système russe GLONASS en possède 30 en constellation et 24 opérationnels. Le système européen GALILEO, en cours de construction sous la responsabilité de l'ESA (*European Space Agency*), devrait également en développer 24 d'ici quelques années ; pour l'instant, 6 sont opérationnels. Notons aussi l'existence du système de navigation chinois COMPAS BEIDOU sur le même principe. Étant donné l'importance stratégique de ces méthodes de localisation, plusieurs autres grandes puissances envisagent de s'équiper d'un tel système. L'exactitude du positionnement par le GPS dépend de la précision de la mesure du délai entre l'émission par le satellite et sa réception sur Terre. Chacun des satellites est équipé d'une horloge commerciale à rubidium dont l'exactitude en fréquence est de l'ordre de 10^{-12} (figure 6.7).

En revanche, le récepteur est doté d'une horloge beaucoup moins précise. Pour pallier cet inconvénient, des astuces ont été inventées qui consistent à coder les signaux émis par chaque satellite, selon la méthode décrite au § 5.2 et illustrée sur la figure 6.2c. Chaque satellite utilise un code différent, le récepteur reconnaît

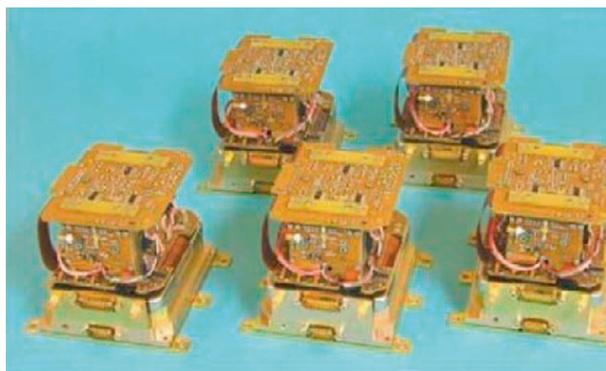


FIGURE 6.7. Horloges à rubidium en cellule qui équipent les satellites du système de positionnement européen GALILEO.

chaque code et décrypte les signaux significatifs ; le temps est ainsi mesuré de la même façon à chaque extrémité de la ligne de propagation.

Une des grandes difficultés rencontrées pour assurer la précision du GPS tient aux incertitudes dans la mesure des durées de propagation, compte tenu de la traversée de l'atmosphère par les ondes électromagnétiques, d'autant plus que les conditions varient en fonction de très nombreux paramètres. L'ionosphère est située entre 50 et 1 000 km de la Terre ; les particules chargées y introduisent un retard pouvant aller jusqu'à 150 ns et qui fluctue en phase avec l'activité solaire au cours de la journée et selon les périodes d'activité du Soleil. Rappelons que l'usage simultané de plusieurs fréquences permet de limiter considérablement les erreurs correspondantes. La troposphère à très basse altitude (entre 3 m et 30 m selon que le satellite émet près ou loin du zénith) introduit aussi des erreurs, en particulier parce que l'indice de réfraction de l'air dépend de l'humidité et donc de la météo. Des modèles adaptés s'efforcent d'en tenir compte.

Par ailleurs, de nombreuses stations au sol contrôlent la vitesse et la position des satellites, connue à 2,5 m près sur leur orbite, et corrigent les erreurs résultant de l'ellipticité des orbites, des distorsions de la forme de la Terre et des anomalies de gravitation liées à la Lune par exemple. Il est indispensable aussi d'effectuer les corrections relativistes décrites au chapitre 5 et mentionnées au § 6.2.4. Pour un satellite à 25 000 km du centre de la Terre tournant à la vitesse de $4 \text{ km}\cdot\text{s}^{-1}$, nous avons vu au chapitre 5 que l'effet relatif de dilation des temps était de l'ordre de 8×10^{-11} , avec comme résultat une erreur de $7 \mu\text{s}$ sur un jour. L'effet du décalage vers le rouge gravitationnel est de 5×10^{-10} et conduit à une erreur de $45 \mu\text{s}$ sur un jour. Ces corrections relativistes sont systématiquement

effectuées dans le calcul des temps de propagation des signaux. En l'absence de ces corrections, le GPS donnerait des localisations erronées de 11 km après une journée. Les relativisto-sceptiques feraient bien de se souvenir de la précision incroyable du GPS installé sur leur voiture ! En fin de compte, le GPS d'usage courant permet une localisation sur Terre avec une précision de l'ordre du mètre, qui peut se réduire au centimètre en géodésie si l'on moyenne sur une longue période en un point fixé. Rappelons qu'il s'agit d'une version dégradée du GPS mis au point aux États-Unis pour un usage strictement militaire, qui a une précision nettement supérieure. Les usages du GPS sont aujourd'hui innombrables dans la vie courante et pour la surveillance de l'environnement.

6.3 Une horloge sur la plateforme spatiale internationale (ISS)

6.3.1 L'horloge PHARAO

Dès les premières horloges à fontaine, l'idée de construire une horloge à atomes froids et de l'envoyer dans l'espace a été envisagée (§ 3.4.2). La première motivation était d'augmenter la précision de l'horloge en s'affranchissant de la gravité. Dans les fontaines, la hauteur du jet d'atomes lancés verticalement et qui retombent après un demi-tour est limitée par l'encombrement d'instruments situés à l'intérieur des observatoires. La fontaine à césium de l'Observatoire de Paris mesure 2 m de haut et le jet seulement 1 m environ. En revanche, en l'absence de pesanteur, on peut envisager de faire voyager les atomes froids avec une vitesse très réduite entre deux zones de Ramsey d'interrogation radio fréquence et d'augmenter ainsi la précision de la mesure, qui varie comme l'inverse du temps de propagation entre les deux zones. Le prototype de *l'horloge PHARAO* a été conçu sur ce principe (figure 6.8). Les atomes de césium sont refroidis à une température de l'ordre de $1 \mu\text{K}$ et piégés dans un piège magnéto-optique réalisé avec des faisceaux de diodes lasers. Ils sont extraits du piège, poussés d'un coup de laser vers les zones d'interrogation et se déplacent à une vitesse qui peut descendre jusqu'à $4 \text{ cm}\cdot\text{s}^{-1}$. Le temps d'interrogation T peut varier entre 0,5 et 5 s. La précision attendue peut être jusqu'à 10 fois meilleure que celle d'une fontaine sur Terre, mais l'amélioration est limitée par les contraintes imposées par la miniaturisation de l'ensemble de l'appareillage et par divers facteurs comme les microvibrations. La stabilité attendue pour cette horloge est de 10^{-13} sur une seconde, 3×10^{-16} sur 10 jours.

La réalisation de l'horloge PHARAO est un tour de force technologique. L'appareillage comporte la cellule représentée sur la figure 6.8, l'optique des

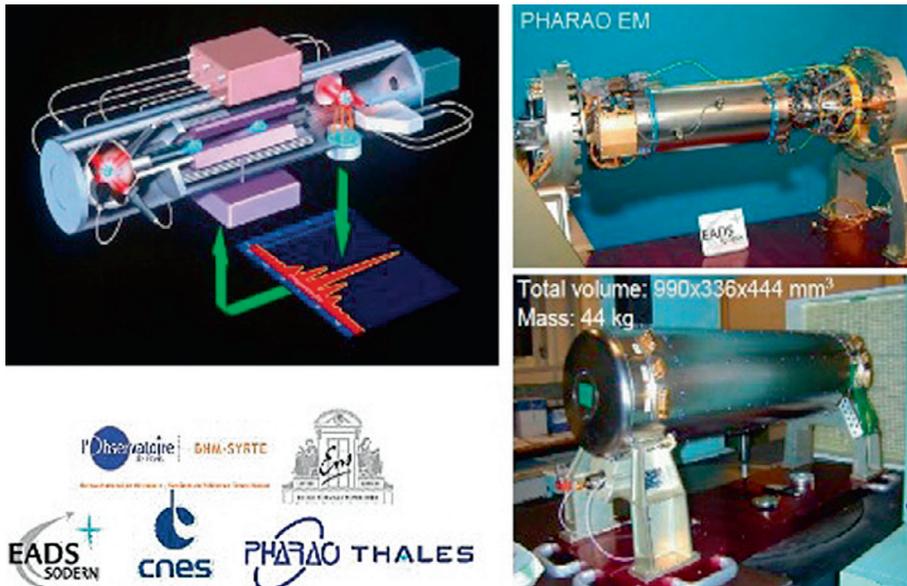


FIGURE 6.8. L'horloge PHARAO à atomes froids de césium qui sera embarquée sur la plateforme spatiale internationale. Sur la partie gauche de la figure on voit le schéma de la cavité : à gauche, les atomes sont confinés dans un piège magnéto-optique. Ils se déplacent en ligne droite à travers la partie centrale d'interrogation qui comporte deux zones de Ramsey distantes de 22 cm. À droite, ils sont interrogés optiquement et donnent en fluorescence un signal d'interférence.

lasers pour le piège et la détection, l'électronique et l'ordinateur qui pilote le tout. L'ensemble a été initialement conçu par le CNES en France et développé dans les laboratoires français. Le prototype a été testé pour des performances spatiales par le CNES à Toulouse. PHARAO deviendra la première horloge à atomes froids en orbite autour de la Terre. La figure 6.9 représente l'ensemble du système optique composé de 10 faisceaux laser issus de diodes semiconductrices stabilisées en fréquence.

6.3.2 La mission ACES

PHARAO est destiné à être arrimé sur le module COLUMBUS de l'*European Space Agency* (ESA) de la plateforme spatiale internationale ISS. Il sera placé sur la plateforme extérieure de COLUMBUS qui fait face à la Terre. ISS est à 400 km d'altitude et tourne autour de la Terre en une heure et demie environ. PHARAO représente la composante principale de la mission européenne ACES (*Atomic Clock Ensemble in Space*) : figure 6.9. Le lancement plusieurs fois différé devrait avoir lieu en 2017. ACES comportera, outre PHARAO, un maser réalisé

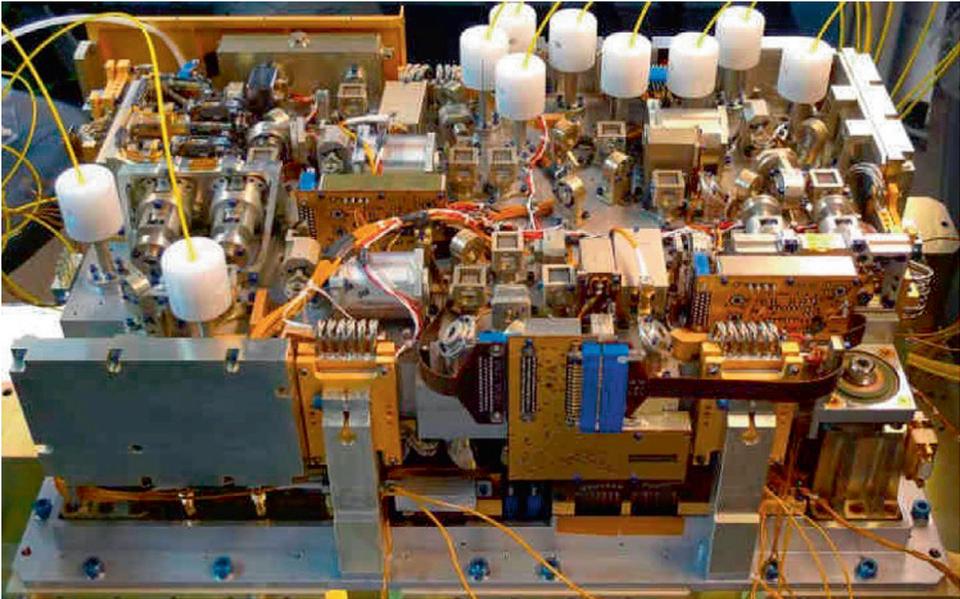


FIGURE 6.9. L'ensemble extrêmement compact du système optique alimentant la cellule de l'horloge PHARAO qui sera embarquée sur la plateforme spatiale internationale en 2017. Les lasers sont des diodes lasers semiconductrices stabilisées en fréquence. Tous les liens optiques sont câblés avec des fibres optiques. L'ensemble occupe un volume de 7 litres. Courtoisie de Philippe Laurent.

en Suisse nommé SHM (*Space Hydrogen Maser*). Nous donnons ci-après quelques détails sur le maser à hydrogène. Le terme MASER est un acronyme pour *Microwave Amplification by Stimulated Emission of Radiation*. Le maser a été inventé par Daniel Kleppner et Norman Ramsey en 1960 à l'université de Harvard. Le principe de fonctionnement est semblable à celui du LASER (*Light Amplification by Stimulated Emission of Radiation*), sauf que le rayonnement émis est du type micro-onde et non pas optique ou infra-rouge. La stabilité à court terme des masers, pour des durées inférieures à un jour, est meilleure que celle des étalons à césium, mais ils présentent une stabilité à long terme et une exactitude inférieures. On les couple souvent à des horloges à alcalins pour bénéficier des qualités des deux types d'instruments. Le maser à hydrogène utilise la transition hyperfine entre les niveaux $F = 1$ et $F = 0$ de l'atome d'hydrogène dans son état fondamental. Ceci correspond à la fréquence de 1,420 GHz. Son principe est décrit sur le schéma de la figure 6.10. De l'hydrogène moléculaire (H_2) est d'abord dissocié pour donner de l'hydrogène atomique H par un champ électrique généré par l'oscillateur haute-fréquence. Celui-ci est ensuite envoyé vers le sélecteur d'état où un tri est effectué à l'aide de champs magnétiques

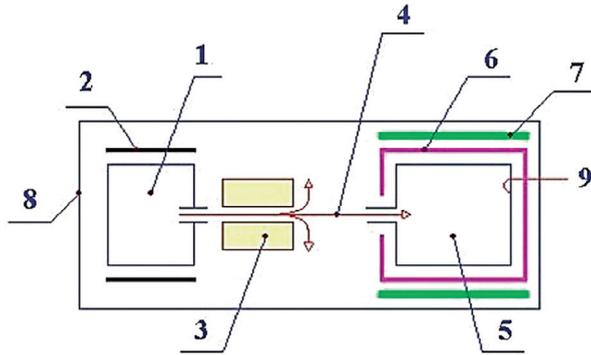


FIGURE 6.10. Schéma du maser à hydrogène qui utilise la transition hyperfine de l'atome d'hydrogène dans son état fondamental à la fréquence 1,42 GHz. (1) Ballon de dissociation (2) Oscillateur haute-fréquence (3) Sélecteur d'état (4) Flux utile d'atomes d'hydrogène (5) Ballon de stockage (6) Cavité micro-onde (7) Bobine de champ (8) Blindages magnétiques (9) Sortie haute fréquence.

non homogènes : les atomes dans l'état d'énergie supérieure sont focalisés sur l'axe de l'appareil, les atomes dans le niveau d'énergie inférieure sont déviés magnétiquement vers l'extérieur et éliminés. Les atomes utiles continuent leur chemin vers la cavité résonante où ils délivrent leur énergie par émission stimulée. Pour obtenir l'auto-oscillation, la probabilité de transition étant faible, il faut augmenter le temps d'interaction de l'atome avec le champ électromagnétique de la cavité. On utilise pour cela un ballon de stockage à l'intérieur duquel les atomes rebondissent durant environ une seconde tout en émettant une fréquence micro-onde. Cette fréquence est captée par la cavité micro-onde pour ensuite sortir du maser afin d'être utilisée pour asservir un oscillateur à quartz de haute qualité, comme dans le cas des horloges atomiques à césium décrites précédemment. Les influences magnétiques extérieures pouvant empêcher une telle émission, une bobine de champ crée un champ magnétique régulier, qui lui-même est protégé par une série de blindages magnétiques extérieurs. Les masers sont principalement utilisés en radioastronomie et en géodésie comme horloges pour la synchronisation de radiotélescopes à grande distance et pour la poursuite des sondes planétaires. Ils servent d'étalons primaires de fréquence.

La comparaison de fréquence effectuée à bord entre SHM et PHARAO sera un élément clé de l'évaluation de la précision et de la stabilité court/moyen terme de l'horloge au césium. Elle permettra d'identifier les conditions de fonctionnement optimales pour PHARAO et de sélectionner un bon compromis entre précision de fréquence et stabilité en combinant la grande stabilité à court terme du maser (10^{-15} en valeur relative sur un temps d'intégration de 1 000 s) avec la stabilité et la précision à long terme de PHARAO. Le maser pourra servir

d'oscillateur à l'horloge PHARAO et remplacer le quartz utilisé dans la plupart des horloges atomiques. La figure 6.11 représente l'ensemble des éléments rassemblés de façon compacte pour la mission ACES.

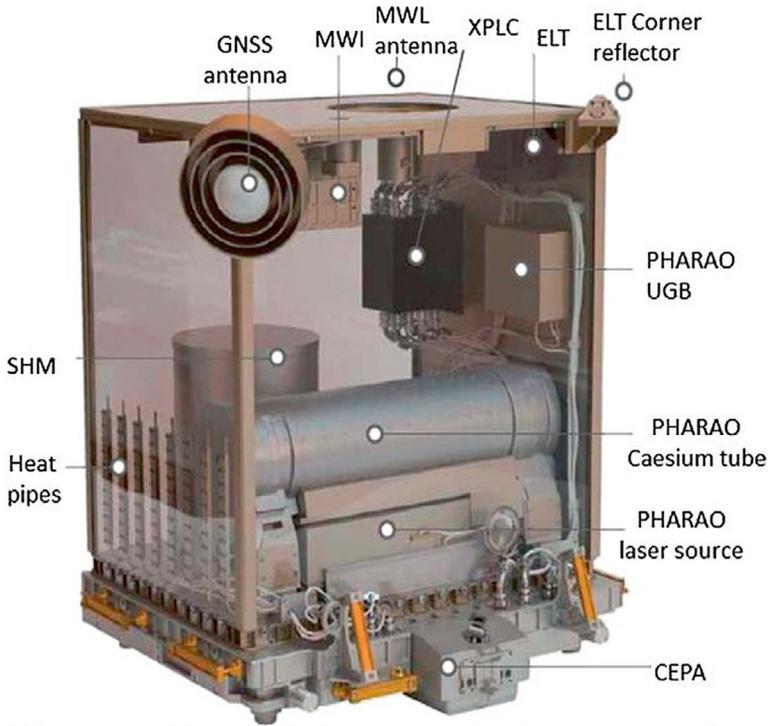


FIGURE 6.11. Éléments de la mission ACES qui sera embarquée sur la plateforme spatiale internationale ISS. L'ensemble est extrêmement compact, occupe 1m^3 , pèse 227 kg et consomme 450 W de puissance électrique. PHARAO est l'horloge à atomes froids de césium réalisée en France. SHM est le maser à hydrogène construit en Suisse. Courtoisie de Philippe Laurent.

PHARAO pilotera la fréquence du maser SHM et le signal maser synchronisera les signaux transférés aux récepteurs sur Terre. Trois méthodes de transfert seront employées simultanément en parallèle. Elles ont été spécialement conçues pour ne pas faire perdre le bénéfice de la grande précision des instruments embarqués. Deux des trois méthodes de transfert sont du type deux voies. La première utilisera un lien micro-onde dédié rétro-réfléchi aux performances qui surpasseront celles des technologies existantes par plus d'un ordre de grandeur. La deuxième méthode utilisera un lien dans le domaine optique, avec des impulsions laser émises du sol et rétro-réfléchies. La troisième sera simplement une connexion à l'échelle de temps fournie par GALILEO, car ACES servira aussi à la communauté GNSS comme horloge de référence. ACES déploiera au

sol les terminaux de réception dans sept laboratoires métrologiques nationaux situés aux États-Unis, en Europe et en Chine. Le SYRTE à l'Observatoire de Paris sera laboratoire pilote. On attend une déviation de la référence de temps de 0,3 ps (picosecondes) au bout de 300 secondes d'intégration, 23 ps à 10 jours. Il s'agit donc d'une échelle de temps très stable à court terme et à long terme. Cette référence de temps unique sera communiquée au réseau des horloges de la Terre, simultanément à toutes celles qui sont dans le champ de visibilité de l'ISS à un instant donné avec la périodicité de la rotation de la plateforme ISS (figure 6.12). La résolution de la comparaison des horloges sur Terre sera de 10^{-17} après quelques jours d'intégration, soit nettement supérieure au transfert du temps par les systèmes du type GPS. Assurer un meilleur transfert du temps à toutes les horloges terrestres est ainsi le but premier de la mission ACES. En outre, cette mission a été sélectionnée pour des objectifs de physique fondamentale, sur lesquels nous reviendrons à la fin de ce chapitre. Notons enfin que, en plus de son intérêt scientifique, elle aura un impact technologique important par la mise au point de composants fragiles, fiabilisés et miniaturisés aux spécifications de l'espace.

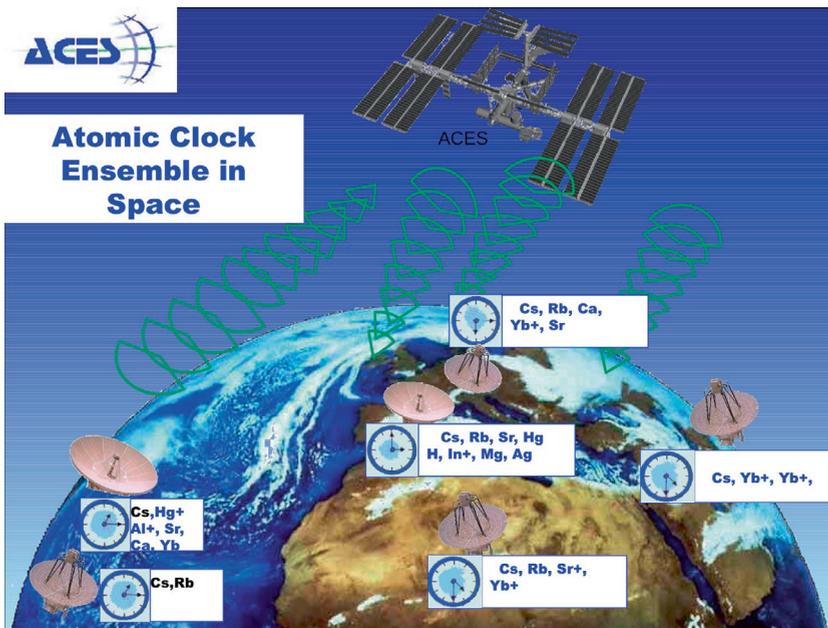


FIGURE 6.12. La mission ACES avec l'horloge à atomes froids PHARAO et le maser SHM sur la plateforme spatiale internationale donnera simultanément une référence de temps à toutes les horloges dans le champ de vision de la plateforme avec une précision supérieure à celle fournie par le GPS.

6.4 Le transfert temps/fréquence par fibre optique

En attendant la réussite de la mission ACES, les réseaux de satellites des systèmes GNSS offrent pour le moment le moyen principal de comparaison entre les fréquences de deux horloges atomiques installées à grande distance l'une de l'autre. Pourtant la méthode manque de précision, tout particulièrement en regard de celle atteinte aujourd'hui par les horloges atomiques ou à ions de nouvelle génération (§ 3.4.3), qui peuvent donner le temps avec 17 voire 18 chiffres significatifs, alors que le GPS est limité à 15 ou 16 au mieux. Les améliorations qu'apportera d'ici peu l'horloge PHARAO sur la plateforme internationale ne suffiront pas pour atteindre la précision des horloges actuelles.

On peut attendre beaucoup de méthodes très différentes émergeant depuis peu, qui pourraient bien un jour rendre caduc l'usage du spatial pour la métrologie. Elles consistent à propager des signaux de référence optiques dans des fibres optiques et relier un usager à l'horloge maîtresse, sans dégradation significative des performances métrologiques. Les spectroscopistes ont commencé à utiliser cette technique il y a une dizaine d'années en installant des fibres optiques dédiées à cet usage installées à l'intérieur des câbles où passent les fils du réseau téléphonique. En France, une fibre a été installée entre le laboratoire SYRTE à l'Observatoire de Paris, où se trouve l'horloge à césium de référence, le Laboratoire Kastler Brossel à Paris et le Laboratoire de physique des lasers à Villeteuse à 40 km. De même, le laboratoire PTB à Braunschweig près de Hanovre, où se trouve la référence de temps pour l'Allemagne, a été relié aux laboratoires de physique atomique à Munich. Ceci permet de calibrer la fréquence des lasers utilisés pour des expériences de très haute précision en spectroscopie atomique ou moléculaire. La bonne surprise a été de constater dès le départ que le transfert T/F s'effectuait avec une source de bruits additionnels faible et une relative insensibilité à l'environnement, étant donné que le câble est enterré. L'inconvénient de la méthode est le coût d'installation de la fibre dédiée et aussi celui de sa maintenance.

Une alternative fondée sur le même principe est en cours de développement. Il s'agit toujours de connexion par fibre optique avec des signaux lasers, mais l'idée nouvelle est de remplacer la fibre dédiée, spécialement installée entre deux laboratoires, par une fibre déjà existante dans le réseau Internet. Il n'y a donc plus d'installation supplémentaire coûteuse à opérer dans les câbles qui transportent les fils du téléphone. Le signal métrologique est constitué par un laser émettant à la longueur d'onde de $1,55 \mu\text{m}$, à l'intérieur de la bande utilisée pour les télécommunications. La fréquence du laser est référencée sur les horloges les plus stables à l'aide d'un peigne de fréquences (§ 3.4.3). Ce signal laser métrologique est injecté dans une des fibres qui transportent les données numériques

de télécommunication, à une fréquence légèrement différente pour ne pas les brouiller et selon les principes courants du multiplexage optique. Cependant ce dispositif n'est pas parfait sur de trop grandes distances. Il faut corriger les fluctuations sur la phase du signal qui résultent de la propagation. On ajoute donc un retour au trajet optique dans la fibre, avec un dispositif électronique constitué d'une boucle à verrouillage de phase (figure 6.13). Le signal métrologique est régulièrement corrigé et régénéré pour pallier les pertes, typiquement tous les 150 km. Le trajet est divisé en segments. Les stations de régénération au bout de chaque segment permettent d'effectuer une architecture en réseau avec distribution à des utilisateurs multiples. Leurs composants optiques et électroniques sont conçus pour être de bas coût, dans la perspective d'une utilisation à grande échelle.

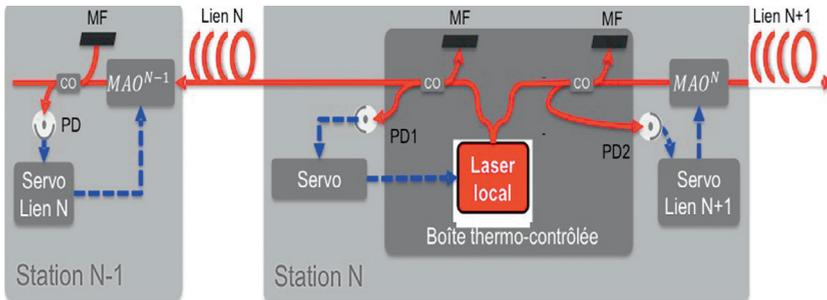


FIGURE 6.13. Transmission d'un standard de fréquence par lien optique à travers le réseau Internet. Ici schéma d'une station N de régénération du signal optique entre les deux segments $N - 1$ et $N + 1$ du réseau. Le laser local de la station est asservi en phase sur le signal arrivant par le lien amont et est injecté à la fois dans les fibres amont et aval pour permettre les corrections des deux liens. MOA : modulateur acousto-optique ; MF : miroir de Faraday ; CO : coupleur optique ; PD : photodiode. Courtoisie d'Anne Amy-Klein, Laboratoire de physique des lasers, Université Paris-Nord

Le procédé est développé en France depuis 2012 en utilisant le réseau Internet RENATER qui relie les centres universitaires. D'abord testé dans la région parisienne, il est maintenant effectif en quatre tronçons de Paris à Reims (270 km) et de Reims à Strasbourg (480 km) *via* Nancy (figure 6.14). La méthode devrait s'étendre prochainement à d'autres grandes universités et aussi à d'autres observatoires sur le territoire français.

Des études très minutieuses des sources de bruit et de l'efficacité de leur réjection ont abouti à d'excellents résultats qui dépassent de loin les prévisions. En effet la fréquence du signal transféré est identique à celle du signal injecté avec une incertitude relative de quelques 10^{-18} en une heure de moyennage et 10^{-19} au bout de quelques heures (figure 6.15). Ceci permet de comparer le

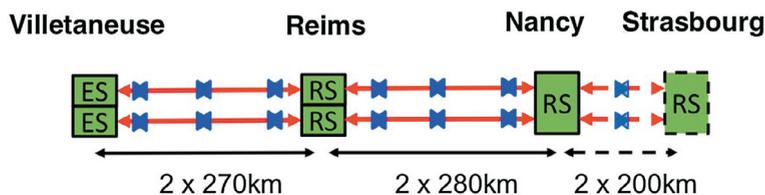


FIGURE 6.14. Liens optiques entre centres universitaires français par l'intermédiaire d'une fibre optique appartenant au réseau Internet. Le dispositif permet de transférer le standard de fréquence issu d'une horloge atomique de référence d'un point à un autre. Le signal optique provient d'un laser dans la bande de fréquence des télécommunications. Il est issu d'une station d'émission ES. Il effectue un aller-retour et traverse des stations de régénération RS.

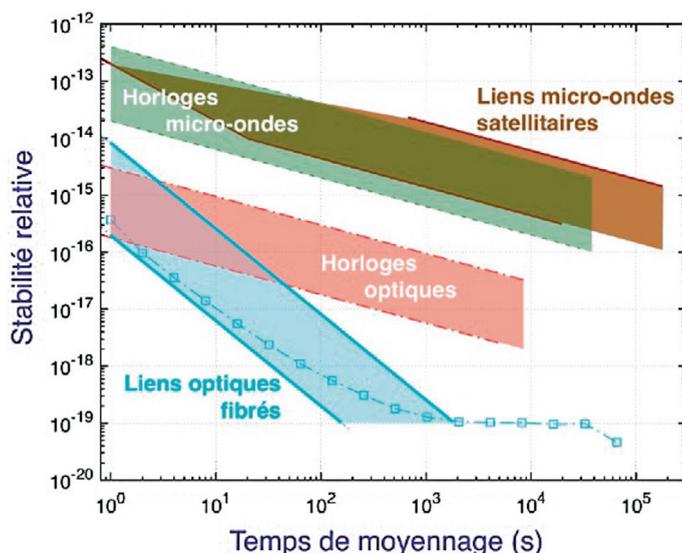


FIGURE 6.15. Comparaison des fréquences des horloges micro-ondes (les standards de fréquence), des horloges optiques (qui sont plus précises) utilisant les liens micro-ondes par satellites en fonction du temps de moyennage. En bleu, stabilité mesurée pour un aller-retour (1 100 km) entre Paris et Nancy avec un transfert par lien fibré via Internet. Courtoisie de Paul-Éric Pottie.

temps fourni par deux horloges distantes avec une précision 100 fois meilleure que par satellite en seulement 10 minutes.

Après ces succès tout récents, il est apparu intéressant de connecter directement par fibre optique les instituts nationaux européens. La première réussite est le lien qui vient d'être établi entre la France et l'Allemagne, avec raccord à Strasbourg. Le lien est réalisé par fibre Internet depuis le SYRTE à Paris jusqu'à Strasbourg, et par fibre dédiée directement du PTB à Braunschweig jusqu'à

Strasbourg. Les horloges optiques ultra-précises à atomes froids de strontium des deux observatoires, développées indépendamment, ont pu être ainsi directement comparées avec une incertitude inégalée de 5×10^{-17} , complètement inaccessible au GPS. Ce test était très important dans une perspective mondiale de définition du temps universel. Il a dans un premier temps validé les incertitudes des horloges. Outre la France et l'Allemagne, d'autres pays européens commencent à développer des liens optiques pour la métrologie. Ainsi, en Italie, une mesure très précise de la fréquence d'une horloge optique à ytterbium vient d'être faite *via* une fibre dédiée de 642 km reliant l'institut de métrologie italien (INRIM) à Turin et le laboratoire européen LENS à Florence. L'utilisation des fibres du réseau Internet pourrait se développer en Europe, en particulier en utilisant le réseau Internet GEANT qui relie l'ensemble des établissements d'enseignement supérieur et de recherche (figure 6.16). La méthode est généralisable sur la Terre à condition de pouvoir installer des stations de régénération régulières, à une distance de l'ordre de 150 km tout le long du trajet. Des pays comme la Chine développent des projets en ce sens. Toutefois l'extension à la traversée de l'Atlantique ne semble pas envisageable.

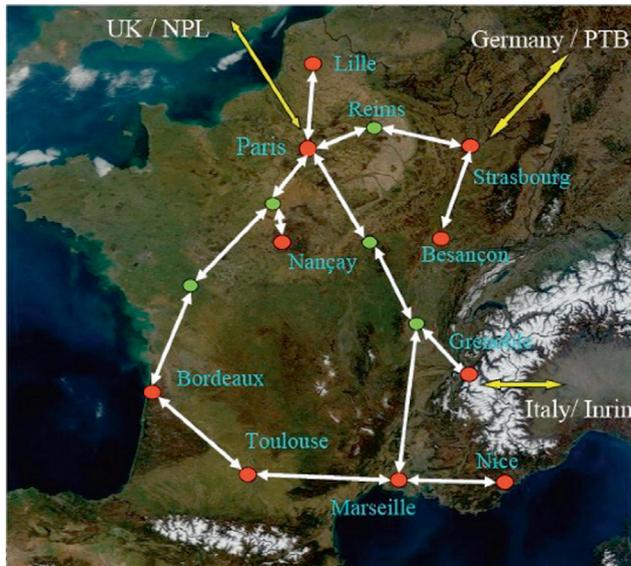


FIGURE 6.16. Le réseau de transfert T/F à travers la France par lien optique via une fibre du réseau Internet RENATER qui relie les centres universitaires en France. Les flèches blanches correspondent au projet REFIMEV. Le lien Paris-Reims-Strasbourg est déjà opérationnel. En jaune les liens prévus par fibre dédiée. Une connexion est maintenant établie avec le centre métrologique de l'Allemagne, le PTB. Courtoisie de Anne Amy-Klein, Laboratoire de Physique des Lasers à Villetaneuse.

En résumé, les liens optiques *via* Internet se sont développés très récemment et offrent un potentiel très supérieur à celui du transfert temps/fréquence par l'intermédiaire des satellites. Ils devraient permettre l'exploitation complète des performances de plus en plus exceptionnelles des horloges de nouvelle génération. Nous reviendrons sur les applications inédites qui en découlent dans le domaine de la physique très fondamentale et aussi en géophysique.

6.5 Le temps universel

En France, l'heure a été unifiée en 1891 en prenant comme référence l'heure repérée par le méridien de Paris (Observatoire de Paris). Depuis 1911, l'heure mondiale a été fixée à partir de l'Observatoire de Greenwich : c'est le temps du méridien de Greenwich (GMT). Depuis 1978, on utilise le *temps universel coordonné* UTC (*Universal Coordinate Time*) fondé sur la synchronisation des horloges les plus précises sur l'ensemble de la planète. C'est le BIPM (Bureau international des poids et mesures) qui établit l'échelle du *temps atomique international* TAI. Il est construit à partir de 400 horloges, pour beaucoup commerciales (des horloges à césium et des masers) mais aussi des horloges de laboratoire, réparties sur tous les continents. L'exactitude, mais aussi la stabilité à long terme sont assurées par les horloges de laboratoire. La calibration est effectuée sur 12 horloges dans le monde, qui en général ne fonctionnent pas toutes en même temps. L'Observatoire de Paris pilote à 40 %. Le TAI est une échelle de temps dont la stabilité en fréquence est de 4×10^{-16} à 30 jours et une exactitude de fréquence de 3×10^{-16} . À partir du TAI on définit le temps universel coordonné (UTC), qui est le même que le TAI mais avec des secondes intercalaires. UTC est à la base du temps légal. Il saute d'une seconde de temps en temps (par exemple en juin 2015). C'est une échelle de temps définie *a posteriori*.

6.6 Quelques applications fondamentales de la mesure du temps

6.6.1 La mesure du décalage gravitationnel

La localisation de l'horloge PHARAO à une grande distance de la Terre en dehors du champ de gravitation devrait fournir une possibilité de mesurer le déplacement gravitationnel prévu par la relativité générale d'Einstein (chapitre 5). Nous avons vu en effet que le décalage gravitationnel était donné en valeur relative par (6.2) : $(U_A - U_B)/c^2$, où U_A et U_B sont les potentiels gravitationnels sur la Terre en A et sur la plateforme en B. U varie comme $1/r$, où r est la

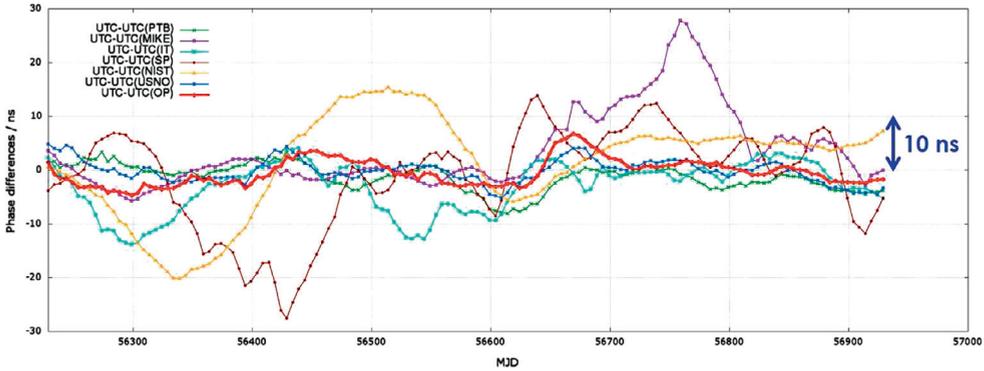


FIGURE 6.17. Le temps universel coordonné (UTC) déterminé dans chaque pays en fonction du temps en jours. Les unités en ordonnées sont la nanoseconde, en abscisses le temps en jours. Les dérivés sont les moins grandes pour l'UTC (PTB) en Allemagne (en vert), et pour l'UTC(OP) en France (en rouge). Chaque pays utilise l'UTC en temps réel pour définir son temps légal. L'UTC est déterminé localement : il y a UTC (NIST) pour les États-Unis, UTC (OP) pour la France, UTC (PTB) pour l'Allemagne. Le temps légal dans le pays x est égal à $UTC(x) + \text{le décalage du fuseau horaire} + \text{le décalage été/hiver}$. En France, le temps légal est $UTC(OP) + 1$ heure en hiver et 2 heures en été. Au PTB, le temps $UTC(PTB)$ a dérivé de moins de 1 ns sur 3 ans. La précision est moins bonne pour déterminer $UTC(x)$ dans les pays qui n'ont pas d'horloge atomique de laboratoire. C'est une des raisons pour lesquelles beaucoup de pays en développement veulent se construire des horloges atomiques. C'est le cas pour la Tunisie qui espère, avec l'aide de la France, mettre au point la première horloge à fontaine de l'Afrique.

distance au centre de la Terre. L'effet est de l'ordre de 10^{-16} par mètre. Avec une altitude de 400 km correspondant à l'éloignement moyen de l'ISS, on s'attend à mesurer un décalage gravitationnel de 4×10^{-11} . Ce décalage pourra être parfaitement mesuré compte tenu de la précision des horloges impliquées aux deux bouts de la chaîne et des performances du lien dédié entre elles. On devrait voir les horloges sur Terre retarder par rapport à celle sur la plateforme. La précision attendue pour ce test sera très grande, de 2×10^{-6} en valeur relative.

Sans constituer une vraie première, cette mesure surpassera considérablement, par près de deux ordres de grandeur, celle des tests précédents pour le même effet. En 1959, deux physiciens américains, Robert Pound et Glenn Rebka, avaient vérifié expérimentalement la prédiction d'Einstein du décalage vers le rouge gravitationnel sur une différence d'altitude de $22,6$ mètres, en utilisant une tour de l'université de Harvard. Rappelons aussi la mesure plus récente de l'expérience *Gravity Probe A* faite par la NASA en 1976 en suivant le temps mesuré par un maser embarqué dans une fusée. Aucun écart par rapport aux prévisions d'Einstein n'a jamais été observé jusqu'ici. Une surprise est-elle possible avec la mesure beaucoup plus précise que permettra l'horloge PHARAO ?

On ne peut l'exclure, car certaines théories prédisent des déviations par rapport aux prédictions de la relativité générale.

6.6.2 Les constantes fondamentales sont-elles constantes ?

La comparaison entre elles sur des temps longs des horloges terrestres a de plus un intérêt fondamental pour la physique. Ainsi les horloges à ions utilisent des ions différents Al^+ , Yb^+ ou Hg^+ . De même, les horloges à atomes utilisent des atomes variés, par exemple horloges à césium ou rubidium pour les fontaines, strontium, mercure ou ytterbium pour les horloges optiques. Pour chacune, la fréquence atomique utilisée pour stabiliser la fréquence de l'horloge est différente (9,2 GHz pour le Cs, 6,8 GHz pour le Rb par exemple) et elle dépend de façon différente des constantes fondamentales sur lesquelles sont fondées les lois de la physique. Si ces constantes ne sont pas vraiment des constantes indépendantes du temps, on doit pouvoir mesurer leur variation en comparant les horloges utilisant des atomes différents et en observant une dérive éventuelle des unes par rapport aux autres au cours du temps. Des comparaisons de cette nature ont déjà été faites depuis quelques années dans les grands observatoires en France, aux États-Unis, en Chine et au Japon, sans qu'une dérive ait pu être observée. Avec l'horloge sur la plateforme ou le lien par fibre optique, la précision pourra être améliorée pour ce test qui revêt une importance particulière pour la cosmologie.

6.6.3 Géodésie relativiste chronométrique

Le décalage vers le rouge gravitationnel, mesuré avec les horloges ultra-précises d'aujourd'hui comparées entre elles *via* l'horloge sur la plateforme ou par fibre optique, pourra trouver des applications encore non explorées en géophysique. En effet une précision de 10^{-18} comme celle des meilleures horloges actuelles permettra de détecter une différence de hauteur de 1 cm entre deux horloges sur la Terre, à condition que le transfert du temps entre les deux ait la même précision. On pourra ainsi déterminer le contour des géoïdes (figure 5.6) et suivre leur variation avec les marées, les cycles de la Lune et la planétologie, être sensible aux mouvements de l'intérieur de la Terre annonçant des séismes, etc. Les horloges ultra-précises sont potentiellement des outils précieux pour les études de géophysique.

6.7 Conclusion

Les horloges ultra-précises auront et ont déjà une utilité pour beaucoup d'autres applications non développées ici. Elles ont une importance fondamentale pour l'astronomie et l'astrophysique, en particulier pour la mesure du ralentissement de la rotation des pulsars. Elles sont indispensables pour la coordination des missions envoyées sur les planètes et pour les grandes expériences fondamentales comme la détection des ondes gravitationnelles avec le grand interféromètre spatial LISA d'ici quelques années. Plus concrètement, l'observation de la Terre à partir de l'espace, fondée sur le GPS, apporte de plus en plus d'informations dans les domaines de l'environnement, des évolutions climatiques, de la surveillance des infrastructures, de la prévision des catastrophes sismiques, etc. Or il est clair que le GPS fera encore des progrès, par exemple avec de meilleures horloges embarquées à atomes froids miniaturisées. Les horloges du futur, déjà arrivées à des précisions extrêmes, vont devenir bientôt plus compactes et plus aisées à manipuler. Le transfert du temps par lien optique finira pas remplacer les liens micro-ondes pour l'espace. Sur la Terre les liens fibrés *via* Internet vont se développer et apporter une précision inégalée. La mesure du temps et son transfert vont se perfectionner et révéler des potentialités fascinantes, tant pour la science fondamentale que pour les besoins sociétaux.

7

La flèche du temps

La plupart des processus de la vie courante se déroulent dans un sens déterminé : si l'on plonge un fer chauffé au rouge dans un récipient d'eau froide, les températures du fer et de l'eau finissent par s'égaliser. Une fois l'égalité des températures atteinte, on ne voit jamais le fer redevenir rouge et l'eau se refroidir, ce que permettrait pourtant la conservation de l'énergie. On dissout aisément du sel dans une casserole d'eau, mais on ne voit jamais des cristaux de sel se former spontanément dans la casserole et déssaliniser l'eau de mer, séparer le sel de l'eau, coûte de l'énergie. Si un apprenti cuisinier maladroit laisse tomber un œuf qui s'écrase sur le carrelage, il ne voit jamais l'œuf se reformer et remonter spontanément dans sa main. Nous vieillissons inexorablement et nous ne rajeunissons jamais spontanément, même si la chirurgie esthétique peut donner quelques illusions éphémères. Tous les événements de la vie courante se déroulent invariablement dans le même sens et on ne revient jamais en arrière : c'est l'*irréversibilité* des phénomènes physiques *macroscopiques*.

Pour fixer les idées, on peut estimer très qualitativement que l'échelle des phénomènes *microscopiques* se situe en dessous de 10 fois la taille atomique, soit environ 1 nanomètre ($1 \text{ nm} = 10^{-9} \text{ m}$), et l'échelle *macroscopique* au-dessus de 100 nanomètres, ou 0,1 micromètre ($1 \mu\text{m} = 10^{-6} \text{ m}$), en gros la limite de résolution d'un microscope optique. On peut aussi différencier une échelle de temps microscopique, typique de la durée d'une collision entre deux molécules, environ 1 picoseconde $= 10^{-12} \text{ s}$, et une échelle de temps macroscopique, le temps de parcours d'une molécule d'un gaz entre deux collisions successives, environ 1 nanoseconde. L'exemple de la figure 7.1 illustre cette irréversibilité visualisée sur des échelles de temps macroscopiques. Un fluide visqueux est contenu dans une cuvette rectangulaire, cette cuvette étant divisée en deux par une

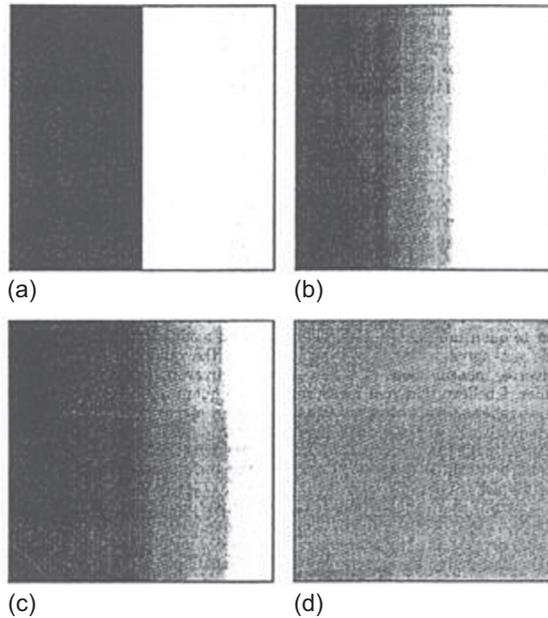


FIGURE 7.1. Écoulement d'un fluide visqueux dans une cuvette. L'intensité du gris représente le niveau du fluide dans la cuvette : le noir correspond à la hauteur maximale et le blanc à la hauteur minimale. Il est facile d'ordonner les figures dans le temps : (a) vient en premier, suivi de (b), ensuite de (c) et finalement de (d). D'après J. Lebowitz, *Time's arrow and Boltzmann's entropy*, Scholarpedia (2013).

paroi qui maintient initialement le fluide dans la partie de gauche de la cuvette. Lorsque l'on retire la paroi, le fluide s'écoule progressivement et se répartit sur l'ensemble de la cuvette. Si l'on nous présente quatre images (a, b, c, d) du processus, il est facile de les ordonner dans le temps : (a) vient en premier, suivi de (b), ensuite de (c) et finalement de (d) : $t_a < t_b < t_c < t_d$. On peut donner une version imagée de l'irréversibilité en filmant un processus de la vie courante et en projetant le film à l'envers : le film apparaîtra absurde, par exemple l'œuf écrasé va se reconstituer et remonter dans la main de l'apprenti cuisinier. Projeter un film à l'envers peut servir pour des effets spéciaux. Dans le film de Jean Cocteau *La Belle et la Bête* (exemple emprunté à Roger Balian), on voit des candélabres s'allumer au fur et à mesure qu'un acteur s'avance dans un couloir. En fait, tous les candélabres étaient initialement allumés, l'acteur marchait à reculons, et un assistant soufflait les flammes en suivant la progression – en marche arrière ! – de l'acteur. Le renversement du sens du temps effectué par une projection à l'envers sera largement commenté au § 7.4.

Bref, le temps semble s'écouler inexorablement toujours dans le même sens. Le concept fondamental qui régit le sens de l'écoulement du temps, souvent

appelé *flèche du temps*, est celui d'entropie : l'entropie d'un système isolé de toute influence extérieure ne peut que croître. Par exemple, l'entropie d'un mélange de sel et d'eau est supérieure à l'entropie de l'ensemble sel plus eau avant dissolution. C'est pourquoi la dissolution est un phénomène spontané, alors que la séparation exige une dépense d'énergie. L'objet de ce chapitre est d'expliquer aussi précisément que possible les différentes facettes du concept d'entropie et les raisons de sa croissance dans une évolution spontanée. Nous verrons que la mise en correspondance du sens d'évolution de l'entropie et du sens de l'écoulement du temps donne une réponse – partielle – à la question de la flèche du temps.

7.1 Le second principe et l'entropie

Le second principe de la thermodynamique, souvent simplement appelé *second principe*, a été déduit de l'analyse du fonctionnement des machines à vapeur, et plus généralement des machines thermiques, au cours de la première moitié du XIX^e siècle. Encore aujourd'hui, l'Académie des Sciences reçoit des schémas de machines prétendant réaliser le mouvement perpétuel. Il n'est pas toujours évident de trouver la faille dans le mécanisme proposé, car les schémas sont souvent d'une extrême complexité, mais on sait *a priori* que de telles machines ne peuvent pas fonctionner, car elles violent l'un des deux premiers principes de la thermodynamique. Le premier de ces principes nous dit qu'aucun mécanisme, quel qu'il soit, ne peut fournir plus d'énergie que celle que l'on a injectée dans le système. C'est un principe de conservation de l'énergie, qui en thermodynamique prend le nom de *premier principe*. Ce principe est absolument général, mais encore faut-il être capable d'identifier toutes les formes d'énergie. Une expérience célèbre de Joule a montré que la chaleur était une forme d'énergie : grâce à la rotation de pales provoquant des frottements visqueux dans un fluide, Joule fut capable non seulement de convertir de l'énergie mécanique en chaleur, mais également de mesurer la constante de conversion de l'énergie mécanique en chaleur, c'est-à-dire relier l'augmentation de température du fluide à l'énergie mécanique fournie. En revanche, dans les années 1930, la situation était beaucoup moins claire dans le cas de la radioactivité β , l'émission d'électrons par des noyaux atomiques radioactifs, car il semblait que l'énergie n'était pas conservée : l'énergie de l'état final après désintégration paraissait plus faible que celle de l'état initial. Ce fut Wolfgang Pauli qui sauva la situation en postulant correctement l'existence d'une particule inconnue et indétectable avec les techniques de l'époque, le *neutrino*. Ce neutrino emportait l'énergie qui semblait avoir disparu, et il fut effectivement mis

en évidence par une détection directe une trentaine d'années plus tard. Pour les physiciens, chimistes et biologistes, la conservation de l'énergie est un principe fondamental, qui ne souffre aucune exception.

Un mouvement perpétuel dont la faille réside dans la non-conservation de l'énergie, et donc dans la violation du premier principe, est appelé mouvement perpétuel de première espèce. Il existe un deuxième type de mouvement perpétuel hypothétique qui repose sur la violation du second principe, et dans ce cas la faille est moins évidente. On sait qu'il est facile de convertir *intégralement* de l'énergie mécanique en chaleur comme dans l'expérience de Joule. Un exemple familier est celui du freinage d'une voiture, où de l'énergie mécanique, dans ce cas l'énergie cinétique de la voiture $mv^2/2$, où m est la masse de la voiture et v sa vitesse, est convertie en chaleur dans les freins, et éventuellement aussi dans le frottement des pneus sur l'asphalte en cas de freinage brutal. Inversement, est-il possible de convertir *intégralement* de la chaleur en énergie mécanique utilisable ? On pourrait imaginer un navire tirant son énergie du refroidissement de l'eau de mer environnante et laissant derrière lui un sillage de glaçons, ce qui serait parfaitement compatible avec le premier principe. L'énergie extraite du refroidissement par un degré d'une couche d'un mètre de profondeur sur l'ensemble de la surface des océans suffirait à alimenter l'humanité en énergie pendant plusieurs années, ce qui serait une forme pratique de mouvement perpétuel, cette fois de deuxième espèce. Mais c'est maintenant le second principe qui interdit le fonctionnement de toute machine utilisant un tel processus.

En 1824, dans son ouvrage *Réflexions sur la puissance motrice du feu*, Sadi Carnot fut le premier à donner une analyse théorique fondamentale du fonctionnement d'une *machine thermique* (figure 7.2). Il observa qu'une telle machine devait utiliser au moins *deux* sources de chaleur ; dans un souci de simplification nous nous limiterons à deux sources, une source chaude à la température absolue T_1 et une source froide à la température absolue T_2 , avec $T_1 > T_2$. Ce que Carnot établit, c'est que le rendement d'une machine bitherme réversible ne dépend que de la température des sources, et que ce rendement est optimal. Pour ce faire, il utilise la conservation du « calorique » et l'impossibilité du mouvement perpétuel de première espèce. Rappelons que la *température absolue*, mesurée en kelvin (K), s'obtient en ajoutant 273 à la température ordinaire mesurée en degrés Celsius : ainsi la température absolue de fusion de la glace est de 273 K. Les physiciens idéalisent la notion de source de chaleur en admettant que celle-ci est suffisamment grande pour que sa température T varie de façon négligeable durant un cycle de la machine et qu'elle reste toujours à l'équilibre thermique. Une machine thermique fonctionne en effet suivant un cycle, c'est-à-dire qu'elle revient périodiquement à son état initial à la fin de chaque cycle. Au cours d'un cycle, la machine puise une quantité de chaleur Q_1 à la source chaude. L'énergie

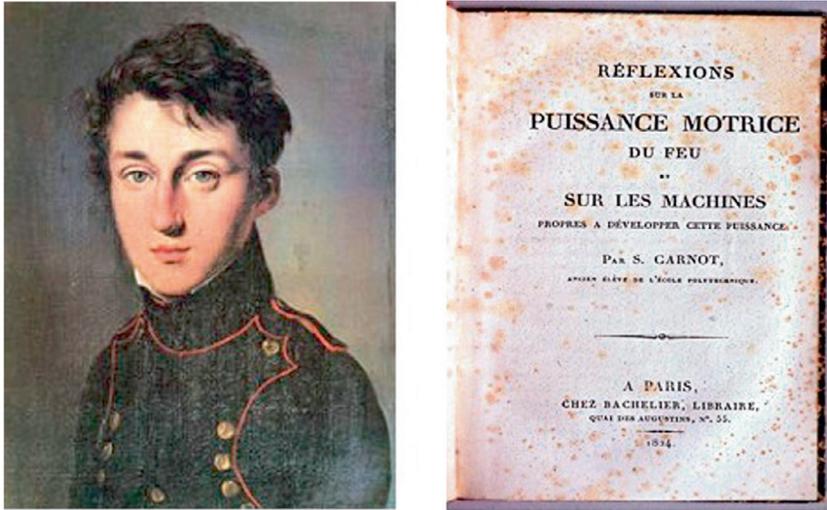


FIGURE 7.2. Sadi Carnot et la couverture de son livre.

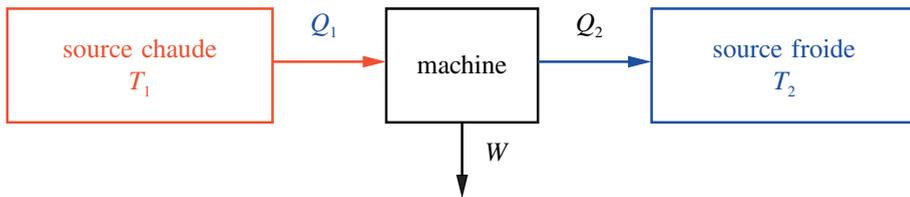


FIGURE 7.3. Schéma de fonctionnement d'une machine thermique : la machine puise une énergie Q_1 sous forme de chaleur dans la source chaude et en restitue Q_2 à la source froide. Par conservation de l'énergie, le travail fourni est $W = Q_1 - Q_2$.

maximale utilisable fournie par la machine, en général appelée travail et notée W , est nécessairement inférieure à Q_1 , car pour fonctionner la machine doit restituer une quantité de chaleur Q_2 à la source froide (figure 7.3) : la production d'un travail s'accompagne d'une inévitable « pollution thermique ». Dans la thermodynamique de Clausius et Thomson, en bonne partie inspirée par le traité de Carnot, le rapport W/Q_1 , appelé *rendement* de la machine, est limité par les températures des sources :

$$\frac{W}{Q_1} \leq 1 - \frac{T_2}{T_1}. \quad (7.1)$$

Le rendement est d'autant meilleur que le rapport des températures T_2/T_1 est petit. Autrement dit il faut s'efforcer de maximiser l'écart des températures entre

la source chaude et la source froide. Si $T_1 = T_2$ comme dans le cas du navire laissant derrière lui un sillage de glaçons, car le navire puise la chaleur dans la mer et la rejette dans cette même mer, alors le rendement est nul et le navire ne peut pas avancer. Un réfrigérateur ou un climatiseur sont des machines thermiques fonctionnant à l'envers (figure 7.4) : la machine puise une quantité de chaleur Q_2 dans le compartiment frigorifique du réfrigérateur ou dans la pièce à refroidir dans le cas du climatiseur. Elle en restitue Q_1 à la source chaude et reçoit un travail W .

L'équation (7.1) est un énoncé possible du second principe, mais il en existe un énoncé plus simple proposé par Clausius vers 1850, qui lui est équivalent. Considérons un système isolé composé de deux sources de chaleur à des températures différentes, par exemple un biberon de lait (trop) chaud dans une casserole d'eau froide, l'ensemble étant isolé thermiquement du milieu l'extérieur. Selon l'énoncé de Clausius, il est impossible qu'une quantité de chaleur positive soit transférée de la source froide vers la source chaude sans intervention extérieure. Autrement dit, la seule possibilité d'évolution spontanée est celle d'un transfert d'une quantité de chaleur positive de la source chaude vers la source froide : c'est l'eau de la casserole qui se réchauffe et le lait qui refroidit. Si l'on attend suffisamment longtemps, les températures de l'eau et du lait s'égalisent et on atteint l'équilibre thermique. Pour transférer de la chaleur depuis une source froide vers une source chaude, il faut fournir de l'énergie : c'est le principe du réfrigérateur ou du climatiseur. La source chaude est l'air extérieur, la source froide l'air de la pièce à refroidir, et le climatiseur fonctionne comme l'inverse d'une machine thermique. On trouve dans le commerce des « climatiseurs » supposés fonctionner en milieu fermé, à l'intérieur d'une pièce et sans contact avec l'extérieur, mais jamais un tel appareil ne pourra faire baisser la température de la pièce !

Poursuivant son étude du second principe et partant de considérations purement thermodynamiques, Clausius eut l'idée d'introduire une quantité dont

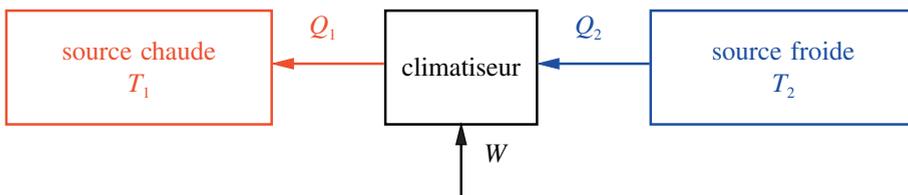


FIGURE 7.4. Schéma de fonctionnement d'un réfrigérateur ou d'un climatiseur. L'appareil puise une quantité de chaleur Q_2 dans le compartiment frigorifique ou la pièce à refroidir. Il en restitue Q_1 à la source chaude et reçoit un travail W . Par conservation de l'énergie, $W + Q_2 = Q_1$.

l'interprétation physique est longtemps restée mystérieuse, *l'entropie*, souvent notée S . Nous donnerons ultérieurement une interprétation microscopique de ce concept. Clausius montra qu'à tout état d'équilibre d'une source de chaleur à la température T on pouvait associer une entropie, et que cette entropie variait de Q/T si cette source recevait une quantité de chaleur Q , positive ou négative ; dans ce dernier cas, cela veut dire que le système fournit une quantité de chaleur positive $-Q$ au milieu extérieur. De plus, de sa formulation du second principe, Clausius déduisit que l'entropie d'un système isolé ne pouvait que croître. Reprenons en effet l'exemple du biberon de lait dans la casserole d'eau froide : pendant un intervalle de temps suffisamment petit pour que les températures de l'eau et du lait varient de façon négligeable, le biberon à la température T_1 (source chaude) cède à l'eau une quantité de chaleur $Q > 0$, et par conservation de l'énergie, l'eau de la casserole à la température T_2 (source froide) reçoit cette même quantité de chaleur Q . La variation d'entropie ΔS est :

$$\Delta S = Q \left(\frac{1}{T_2} - \frac{1}{T_1} \right) > 0, \quad (7.2)$$

car $T_1 > T_2$. De même, dans le fonctionnement d'une machine thermique et en appliquant le principe de Clausius sur un cycle, on obtient :

$$\Delta S = -\frac{Q_1}{T_1} + \frac{Q_2}{T_2} \geq 0. \quad (7.3)$$

Un exercice d'algèbre élémentaire utilisant la conservation de l'énergie $W = Q_1 - Q_2$ montre que (7.3) est équivalent à la loi de Carnot (7.1). Le meilleur rendement possible est obtenu lorsque l'inégalité dans (7.3) devient une égalité, c'est-à-dire lorsque la variation d'entropie est nulle. Lorsque c'est le cas, on dit que le fonctionnement de la machine thermique est réversible, et plus généralement on dit qu'un *processus est réversible* s'il est tel que l'entropie reste constante. En pratique, une machine thermique n'atteint jamais le rendement maximum car, si l'on voulait éviter toute dissipation par exemple par frottement, on devrait faire fonctionner la machine de façon infiniment lente.

Afin de définir précisément la notion d'entropie, il importe d'insister sur les notions de système isolé, de contrainte interne et d'équilibre thermodynamique, un équilibre tel qu'en particulier la température et la densité soient uniformes et indépendantes du temps dans tout le système. Deux exemples permettront d'illustrer ces notions. Dans le premier exemple (figure 7.5), deux fluides à des températures différentes sont contenus dans une enceinte divisée en deux compartiments par une cloison *adiabatique*, c'est-à-dire une cloison qui ne conduit pas la chaleur. L'enceinte est séparée du milieu extérieur par une paroi rigide et

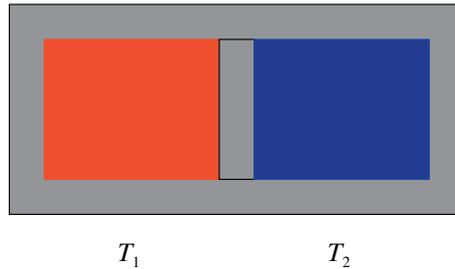


FIGURE 7.5. Deux fluides à des températures différentes sont contenus dans une enceinte divisée en deux compartiments par une cloison adiabatique, c'est-à-dire qui ne conduit pas la chaleur. L'enceinte est séparée du milieu extérieur par une paroi rigide et adiabatique, ce qui en fait un système isolé. Lorsque l'isolant thermique est retiré de la cloison interne, la chaleur peut s'écouler entre les deux compartiments et les températures s'égalisent : l'équilibre thermique s'établit. L'entropie de l'état final est supérieure à celle de l'état initial.

adiabatique, ce qui en fait un système isolé. Dans une première phase, on attend suffisamment longtemps pour que l'équilibre thermique soit atteint dans chacun des deux compartiments avec deux températures différentes. Lorsque l'isolant thermique est retiré de la cloison interne sans apporter d'énergie aux fluides, ce qui correspond à la levée d'une contrainte interne au système, la chaleur peut s'écouler librement entre les deux compartiments. On attend alors suffisamment longtemps pour que l'équilibre thermique soit atteint dans l'ensemble de l'enceinte dont la température devient uniforme : la température est la même dans les deux compartiments. L'énoncé précis de la croissance de l'entropie est le suivant : si on lève une contrainte interne à un système isolé, l'entropie de l'état d'équilibre final est supérieure ou égale à celle de l'état d'équilibre initial. Un exposé de la thermodynamique initié par Herbert Callen consiste à prendre comme postulat fondamental la croissance de l'entropie entre deux états d'équilibre après la levée d'une contrainte interne. On retrouve le second principe sous la forme usuelle comme conséquence de ce postulat, et plusieurs manuels récents utilisent cette structuration de l'exposé. Un autre cas sur lequel nous reviendrons plusieurs fois par la suite est l'« exemple magique » de la figure 7.6 : un gaz est contenu dans une enceinte de volume total V , isolée du milieu extérieur par une cloison adiabatique et rigide, divisée en deux compartiments de volume $V/2$ par une cloison. Le gaz se trouve initialement confiné dans le compartiment de gauche. Lorsqu'un trou est percé dans la cloison, ce qui correspond à la levée d'une contrainte interne au système, le gaz se répartit librement dans l'ensemble de l'enceinte : à l'équilibre, la densité du gaz est uniforme dans l'enceinte. L'entropie de l'état d'équilibre final est supérieure à celle de l'état d'équilibre initial.

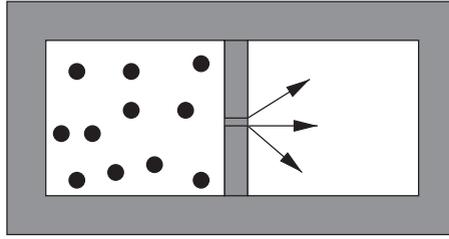


FIGURE 7.6. Un gaz est contenu dans une enceinte de volume total V divisée en deux compartiments de volume $V/2$ par une cloison rigide et adiabatique. Le gaz se trouve initialement dans le compartiment de gauche. Lorsque la cloison est retirée, le gaz se répartit dans l'ensemble de l'enceinte : à l'équilibre, la densité du gaz et sa température sont uniformes dans l'enceinte. L'entropie de l'état final est supérieure à celle de l'état initial.

Revenons un instant sur la notion d'équilibre thermodynamique en conclusion de ce § 7.1. Pour qu'un tel état d'équilibre soit atteint, il faut que toute la mémoire de l'histoire passée du système physique considéré ait été effacée car, si tel n'était pas le cas, l'état d'équilibre ne serait pas unique. Un état d'équilibre doit pouvoir être caractérisé par un petit nombre de paramètres. Dans le cas d'un système composé d'un seul type de molécules, l'entropie S ne doit dépendre que de l'énergie interne E du système, du volume V du compartiment qui le confine et du nombre N de molécules : $S(E, V, N)$. La condition d'effacement de toute mémoire est très contraignante. En effet, si un état d'équilibre est atteint très rapidement dans la plupart des cas, il existe des exceptions importantes. Par exemple un verre ordinaire, un carreau de fenêtre ou un vitrail de cathédrale sont des systèmes hors équilibre, qui se trouvent dans des états dits métastables. Ces états dépendent de la façon dont a été fabriqué le verre, par exemple la vitesse de refroidissement. Pour certains systèmes, atteindre l'équilibre peut prendre plusieurs siècles, voire plusieurs millénaires.

7.2 L'entropie de Boltzmann

Il faut bien admettre que le concept d'entropie tel qu'introduit par Clausius est mystérieux et abstrait, et plusieurs générations d'étudiants en thermodynamique – et parfois même leurs enseignants – ont transpiré en appliquant ce concept à des problèmes de physique ou de chimie. Dans ce § 7.2, nous allons donner une interprétation plus intuitive à l'entropie en examinant sa signification microscopique. L'idée de base est celle de configuration microscopique de l'état d'un système physique, ou *microétat*. Rappelons que l'échelle des phénomènes microscopiques se situe très grossièrement en dessous du nanomètre



FIGURE 7.7. Ludwig Boltzmann (1844-1906).

et celle des phénomènes macroscopiques au dessus de 100 nanomètres. Commençons par un exemple élémentaire. Supposons que nous disposions de deux cases, les deux compartiments de la figure 7.6, et que nous essayions de placer une par une et au hasard les molécules du gaz dans les deux compartiments supposés identiques. Avec une seule molécule, seulement deux possibilités s'offrent à nous : nous pouvons la placer dans le compartiment de droite ou dans celui de gauche. Si nous jouons à pile (P) ou face (F) afin de placer les molécules au hasard, P pour le compartiment de gauche et F pour celui de droite, les 2 possibilités sont P ou F . Avec deux molécules, nous avons 4 possibilités : les deux molécules dans le même compartiment, avec une probabilité $1/4$ pour chaque compartiment et une molécule dans chaque compartiment avec une probabilité $1/2$, ces 4 possibilités correspondant aux tirages :

PP, PF, FP, FF

Passons à 10 molécules en poursuivant notre jeu de pile ou face. Dans le tableau suivant, nous donnons deux chiffres entre parenthèses, le nombre de molécules à droite et à gauche, et ensuite le nombre de façons de réaliser cette combinaison : par exemple (7,3) : 120 veut dire 7 molécules à gauche, 3 à droite, et 120 façons de réaliser cette combinaison.

(10,0) : 1 (9,1) : 10 (8,2) : 45 (7,3) : 120 (6,4) : 210 (5,5) : 252

Il est inutile d'aller plus loin, car le nombre de combinaisons pour (4,6) est bien évidemment le même que pour (6,4). Dans cet exemple, le microétat est donné

par la spécification de la position des molécules, par exemple les molécules 1, 3, 4, 5, 9 à gauche et les molécules 2, 6, 7, 8, 10 à droite, ce qui correspond au tirage *PFPPFFFPF*. Le nombre total de microétats est de $2^{10} = 1024$. La probabilité d'observer la configuration (10,0) est de 2^{-10} , environ 0,1%, tandis que celle d'observer la configuration (5,5), où les molécules sont également distribuées dans les deux compartiments, est de $252/1024 \simeq 25\%$: cette configuration, qui est l'analogie du gaz réparti uniformément dans l'enceinte, est environ 250 fois plus probable que la configuration où toutes les molécules sont dans un même compartiment, et pourtant nous n'avons que 10 molécules ! Ces résultats sont illustrés sur l'histogramme de la figure 7.8.

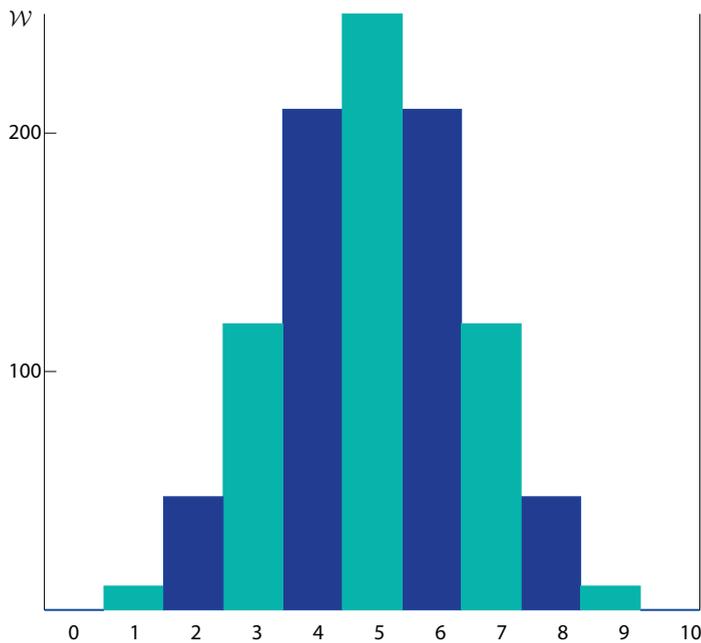


FIGURE 7.8. Histogramme donnant le nombre \mathcal{W} de configurations en fonction du nombre de molécules dans le compartiment de gauche.

Cet exemple montre que le nombre de configurations croît très rapidement avec le nombre N de molécules. Pour un système macroscopique typique, $N \sim 10^{23}$, le nombre d'Avogadro qui s'écrit comme un suivi de vingt-trois zéros, il n'est même pas envisageable d'écrire 2^N sous forme explicite. Afin d'éviter de manipuler des nombres énormes, au lieu d'écrire le nombre de configurations on passe par son logarithme, noté \ln . La propriété essentielle du logarithme est de transformer un produit en somme, $\ln(ab) = \ln a + \ln b$, d'où $\ln 2^N = N \ln 2$.

On le vérifie sur le cas simple $N = 3$

$$\ln 2^3 = \ln(2 \times 2 \times 2) = \ln 2 + \ln 2 + \ln 2 = 3 \ln 2.$$

L'entropie statistique est définie à un coefficient de proportionnalité près par le logarithme $\ln \mathcal{W}$ du nombre de configurations. Nous avons utilisé la notation \mathcal{W} pour le nombre de configurations afin d'une part de respecter les notations consacrées par l'usage et d'autre part d'éviter toute confusion avec le travail W . Si l'on considère deux systèmes indépendants ayant respectivement 2^{N_1} et 2^{N_2} configurations possibles, alors l'entropie totale sera la somme des entropies individuelles : l'entropie est définie comme une quantité *additive*.

Un exemple un peu plus complexe, mais plus proche d'un cas physique concret, est celui de la description microscopique d'un gaz contenu dans une enceinte. En première approximation, il est possible de négliger tous les effets quantiques et relativistes, et de décrire la dynamique des molécules du gaz en utilisant la mécanique newtonienne. La configuration microscopique du gaz est donnée par l'ensemble des N positions individuelles $\vec{r}_1, \dots, \vec{r}_N$ et des N vitesses $\vec{v}_1, \dots, \vec{v}_N$ des N molécules. Au lieu des vitesses, il est d'usage de se donner les impulsions $\vec{p} = m\vec{v}$, où m représente la masse des molécules. La logique de la mécanique newtonienne voudrait que l'on fixe les positions et vitesses initiales des molécules et que l'on calcule positions et vitesses ultérieures en fonction des forces s'exerçant sur les molécules, d'une part les forces entre les molécules et d'autre part les forces exercées par les parois de l'enceinte. Une telle stratégie est à l'évidence vouée à l'échec, car il est impossible de suivre le mouvement de chacune des 10^{23} molécules : le simple stockage des positions et des vitesses de toutes les molécules à un instant déterminé saturerait par plusieurs ordres de grandeur les mémoires des ordinateurs les plus performants. De plus, non seulement cette stratégie est vouée à l'échec, mais elle n'est pas pertinente : nous ne sommes pas intéressés par la trajectoire d'une molécule particulière, nous n'avons pas besoin d'une information aussi détaillée. Ce qui nous intéresse, ce sont les propriétés globales, ou macroscopiques, du gaz, par exemple sa pression et sa température. Nous allons définir un *macroétat* du gaz par la donnée de quelques informations macroscopiques : volume V de l'enceinte, énergie totale du gaz, qui est approximativement la somme des énergies cinétiques de toutes les molécules, pression, température, etc. Toutes ces quantités ne sont pas indépendantes : par exemple, dans le cas d'un gaz parfait où les forces entre molécules sont négligeables, pression P , température absolue T et volume V sont reliés par la loi des gaz parfaits :

$$PV = Nk_B T, \quad (7.4)$$

où k_B est la *constante de Boltzmann*, qui vaut $1,38 \times 10^{-23}$ Joules/degré. Comme dans le cas élémentaire étudié ci-dessus, le nombre de configurations microscopiques compatibles avec la donnée d'un macroétat, c'est-à-dire avec une information macroscopique, est énorme. On peut par exemple échanger la vitesse ou la position de deux molécules sans changer l'énergie totale du gaz, et donc le macroétat. Comme ci-dessus, appelons \mathcal{W} le nombre total de configurations microscopiques compatibles avec le macroétat et définissons l'*entropie de Boltzmann* S_B par :

$$S_B = k_B \ln \mathcal{W}. \quad (7.5)$$

En termes techniques, \mathcal{W} est le volume de l'espace des phases. La formule (7.5) est l'une des plus importantes de toute la physique et elle est gravée sur la tombe de Boltzmann à Vienne. Dans le cas de deux systèmes indépendants dont le nombre de configurations vaut respectivement \mathcal{W}_1 et \mathcal{W}_2 , le nombre total de configurations est $\mathcal{W} = \mathcal{W}_1 \mathcal{W}_2$ et, en raison des propriétés du logarithme, l'entropie totale est la somme des entropies individuelles : l'entropie de Boltzmann est définie comme une quantité additive. Il est tout à fait remarquable que l'on puisse montrer la propriété suivante : pour un système à l'équilibre, l'entropie de Boltzmann coïncide avec l'entropie thermodynamique de Clausius S à une constante additive près sans importance physique : $S_B = S + \text{constante}$.

Encadré 7.1. Configurations microscopiques d'un gaz.

Dans le cas de N molécules identiques contenues dans une enceinte de volume V , le nombre total de configurations est donné par une intégrale à $6N$ dimensions :

$$\mathcal{W} = \frac{1}{N! h^{3N}} \int d^3 r_1 \cdots d^3 r_N d^3 p_1 \cdots d^3 p_N, \quad (7.6)$$

où nous devons satisfaire les contraintes macroscopiques : les positions \vec{r}_i sont localisées à l'intérieur de l'enceinte et la somme des énergies cinétiques et potentielles des molécules doit être égale à l'énergie totale. Le facteur $N!$ tient compte de l'identité des molécules, tandis que h représente la constante de Planck, le seul souvenir de la physique quantique : en fait la formule (7.6) permet de faire le décompte du nombre d'états quantiques compatibles avec les contraintes macroscopiques. Si l'on peut négliger les interactions entre les molécules (approximation du gaz parfait) la somme des énergies cinétiques individuelles $\sum_i^N \vec{p}_i^2 / 2m$ est égale à l'énergie totale.

Dans le cas du gaz parfait, on montre à partir de (7.6) que \mathcal{W} est proportionnel au volume à la puissance N : $\mathcal{W} \propto V^N$. On retrouve alors une situation déjà décrite au début du § 7.2 : en effet, si le gaz passe du volume $V/2$ au volume V comme dans la figure 7.6, le nombre de configurations augmente par un facteur 2^N . Ce comptage des configurations va éclairer un exemple classique d'irréversibilité. Considérons un gaz parfait enfermé dans le compartiment de gauche de

volume $V/2$ d'une enceinte de volume total V (figure 7.6). Dans le macroétat initial, les positions des molécules sont contraintes à se trouver dans le compartiment de gauche. Levons la contrainte en retirant la cloison (ou en la perçant) entre les deux compartiments sans apporter d'énergie au gaz. Le nombre de configurations microscopiques est alors, nous l'avons vu, supérieur par un facteur 2^N au nombre de configurations du macroétat initial, et l'entropie de ce macroétat augmente de $Nk_B \ln 2$ par rapport à l'entropie initiale. La configuration où le gaz occupe l'ensemble de l'enceinte est plus probable par un facteur 2^N que celle où les molécules sont rassemblées dans le compartiment de gauche. De façon générale, la croissance de l'entropie s'explique par le fait que lever une contrainte augmente en règle générale le nombre de configurations microscopiques de façon considérable.

Le lecteur attentif aura remarqué une hypothèse implicite dans ce qui précède : nous avons supposé que toutes les configurations microscopiques compatibles avec un macroétat étaient également probables. C'est ce que nous avons supposé en distribuant les molécules au hasard. En d'autres termes, nous avons choisi une loi de probabilité présentant un biais minimum, une loi qui prend en compte toute l'information disponible mais seulement cette information. L'entropie est donc définie par le choix d'une loi de probabilité, le choix de considérer comme équiprobables toutes les configurations microscopiques accessibles compatibles avec l'état de nos connaissances macroscopiques sur le système. À cet état de connaissance, nous avons associé une loi de probabilité qui mesure notre ignorance. L'entropie apparaît comme une mesure de notre manque d'information sur le système, aussi appelé information manquante. Cette probabilité est à la fois objective et subjective : deux physiciens avec des états de connaissance différents attribueront des entropies différentes à un système physique, mais l'entropie sera la même si leur état de connaissance est identique. Il existe une différence de principe entre par exemple l'énergie et l'entropie. L'énergie d'un système physique est une propriété objective, mais pas l'entropie qui dépend de notre information. Ainsi on peut attribuer sans ambiguïté une énergie à un chat, il suffit de le peser et d'utiliser la relation d'Einstein $E = mc^2$, mais on ne peut pas lui attribuer une entropie, et de fait on ne peut même pas attribuer une entropie à une bactérie.

La définition de l'entropie d'un système physique à l'équilibre est non ambiguë, mais il n'en est pas de même pour un système hors équilibre. Par exemple, dans le cas de la figure 7.1, les états (a) et (d) sont des états d'équilibre pourvu que l'on ait attendu assez longtemps, mais pas les états intermédiaires (b) et (c). Dans un tel cas, il est souvent possible de fixer des caractéristiques macroscopiques et de définir proprement un macroétat. Dans l'exemple de la figure 7.1, on peut se fonder sur le niveau du fluide qui varie continûment dans la cuvette,

pourvu que l'écoulement soit suffisamment lent à l'échelle macroscopique. On peut alors admettre qu'un volume de fluide, grand à l'échelle microscopique mais petit à l'échelle macroscopique, est approximativement à l'équilibre et on peut donc lui attribuer une entropie. C'est une situation typique d'équilibre local, où température et densité par exemple sont définies localement, sur une échelle d'espace et de temps macroscopiques. L'entropie totale est alors obtenue en faisant la somme des entropies des petits volumes. Hors équilibre, cette entropie varie au cours du temps même dans une situation stationnaire, par exemple dans le cas de l'effet Joule où un courant continu constant dans le temps produit de l'entropie dans le fil conducteur. Nous avons également vu que l'entropie est fonction de l'information dont dispose l'observateur : mais, nous l'avons souligné, deux observateurs disposant de la même information seront d'accord sur l'entropie à attribuer au système. Le nombre de configurations microscopiques compatibles avec un macroétat sur lequel on dispose d'une information détaillée est plus faible que lorsque notre information est plus grossière, et l'entropie sera plus grande dans ce second cas. Dans l'exemple de la figure 7.6, notre information est plus précise si nous savons que toutes les molécules sont situées dans le compartiment de gauche que lorsque les molécules sont réparties dans l'ensemble de l'enceinte. Le fait que l'entropie dépende de l'information accessible à l'observateur sera illustré de façon spectaculaire au § 7.5.

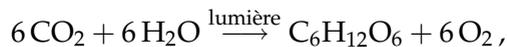
7.3 Énergie et entropie

Cette partie est marginale dans le livre et peut être omise en première lecture. En effet, notre sujet principal, le temps, n'y est pas mentionné, et notre objectif est d'expliquer, en mettant à profit le concept d'entropie, pourquoi ce que l'on appelle communément « crise de l'énergie » est fondamentalement une « crise de l'entropie ». Pour résumer la situation, on peut dire que ce qui nous intéresse n'est pas tant l'énergie que l'énergie utilisable, et cette énergie doit donc être accessible sous une forme de basse entropie. Par exemple, l'ensemble des océans contient une énergie thermique considérable, mais elle n'est pas utilisable car elle correspond à une entropie élevée. En revanche l'énergie des vagues, des courants ou des marées est utilisable car elle se présente sous une forme ordonnée.

En fin de compte on doit se poser la question : quelle est l'origine des états de basse entropie qui nous donnent de l'énergie sous une forme utilisable ? Commençons par le cas des êtres vivants, une bactérie ou bien nous-même. Si l'on fait le bilan énergétique global d'une personne adulte, ce bilan donne un

résultat nul : l'énergie stockée par cette personne est constante au cours du temps – sauf si elle prend du poids! –, l'énergie absorbée par une personne est exactement égale à celle qu'elle rejette dans son environnement. Pourquoi avons-nous donc besoin d'énergie pour fonctionner? En fait un être vivant est hautement organisé et il doit être capable de maintenir cette organisation, voire de l'améliorer, s'il veut survivre. En d'autres termes, il doit être capable de fabriquer des états de basse entropie. Or, nous l'avons vu, l'entropie d'un système isolé, que l'on peut dans ce cas assimiler à l'ensemble être vivant plus son environnement immédiat, ne peut que croître. Il faut donc que cet être vivant absorbe de l'énergie sous une forme correspondant à une basse entropie et la rejette dans son environnement sous une forme correspondant à une entropie élevée, afin que l'entropie de l'ensemble être vivant plus environnement immédiat augmente. Le réfrigérateur de la figure 7.4 fonctionne suivant ce principe : pour diminuer la température du compartiment frigorifique, et donc son entropie, il doit emprunter de l'énergie au milieu extérieur.

Un processus fondamental pour la vie est la *photosynthèse*, qui fournit l'essentiel de l'énergie utilisée par la vie sur la Terre : une plante absorbe du dioxyde de carbone (ou gaz carbonique), sépare le carbone de l'oxygène et utilise le carbone pour prospérer. Une réaction typique est celle qui transforme du dioxyde de carbone (CO₂) et de l'eau (H₂O) en glucose (C₆H₁₂O₆) et oxygène (O₂) :



en utilisant la lumière solaire. Ce processus correspond à une diminution d'entropie de la plante, qui doit se faire aux dépens de son environnement afin que l'entropie totale augmente. La photosynthèse utilise la lumière solaire sous forme de rayonnement visible de longueur d'onde moyenne 0,5 μm et rejette de l'énergie sous forme de rayonnement infrarouge de plus grande longueur d'onde, de fréquence plus faible. En raison de la relation de Planck-Einstein entre l'énergie E et la fréquence ν , $E = h\nu$, les photons infrarouges ont une énergie plus faible que les photons du visible. Pour une même valeur de l'énergie totale, l'entropie d'un ensemble de photons infrarouges sera plus grande que celle d'un ensemble de photons visibles, car l'énergie est distribuée sur un plus grand nombre de photons et donc sur un plus grand nombre de degrés de liberté, ce qui augmente le nombre de configurations microscopiques : on peut montrer que le rapport des entropies est égal à celui des longueurs d'onde. Ce résultat se comprend intuitivement : si l'on augmente la longueur d'onde des photons par un facteur 2, on divise par 2 leur énergie et il faut donc 2 fois plus de photons pour une énergie totale fixée d'un ensemble de photons. La puissance moyenne captée par la photosynthèse sur l'ensemble du globe avoisine

les 130 térawatts ($1 \text{ TW} = 10^{12} \text{ W}$), environ 6 fois la consommation énergétique de l'humanité. Les combustibles fossiles (charbon, pétrole, etc.) proviennent de plantes préhistoriques de sorte que la possibilité de récupérer de l'énergie sous une forme utilisable est à nouveau liée à la conversion dans un passé lointain de photons de haute énergie, des photons du visible, en photons de basse énergie, des photons infrarouges. Les mêmes considérations valent pour l'ensemble de la Terre : la Terre absorbe de l'énergie sous forme de photons de haute énergie émis par le Soleil et la rejette dans l'espace sous forme de photons de basse énergie, le bilan énergétique global étant nul : l'énergie stockée par la Terre doit rester constante dans le temps... sauf si sa température moyenne augmente en raison du changement climatique, mais l'effet éventuel est très faible comparativement à l'énergie totale absorbée. L'entropie des photons réémis par la Terre est environ 20 fois celle des photons reçus : ce facteur 20 n'est autre que le rapport entre la température de surface du Soleil, environ 6 000 K, et celle de la Terre, environ 300 K. En effet, on montre que la longueur d'onde moyenne émise par le rayonnement d'un corps noir est inversement proportionnelle à sa température et, en première approximation, la Terre et le Soleil peuvent être assimilés à des corps noirs en ce qui concerne leur rayonnement.

En dernière analyse, il faut comprendre les raisons qui nous permettent de disposer d'états de faible entropie. Si l'environnement de la Terre était en équilibre thermique, ce serait impossible. Mais cet environnement n'est pas en équilibre thermique, il existe une source chaude dans un environnement froid, le Soleil, qui est la source des photons visibles. Si l'on analyse l'origine de cette émission de photons, la conclusion est que le phénomène a pour origine ultime les forces de gravitation. C'est la gravitation qui a permis l'effondrement d'un nuage d'hydrogène gazeux, lequel a formé notre étoile. Cet effondrement a conduit à des températures suffisamment élevées pour déclencher des réactions thermonucléaires s'opposant à l'effondrement gravitationnel, ce qui permet au Soleil de fonctionner environ 10 milliards d'années. Nous reviendrons au § 7.4 sur le rôle fondamental joué par la gravitation dans les questions d'entropie à l'échelle cosmologique.

7.4 Réversibilité microscopique et irréversibilité macroscopique

Reprenons une fois de plus notre exemple magique, celui de l'expansion d'un gaz (figure 7.6). Lorsque la cloison séparant les deux compartiments est retirée, le gaz s'engouffre dans le compartiment de droite et remplit l'ensemble de l'enceinte de façon uniforme une fois l'équilibre atteint. Nous avons donné

la description microscopique du phénomène : le nombre de configurations microscopiques de l'état final est supérieur par un facteur énorme, de l'ordre de 2^N , au nombre de configurations initial. De façon équivalente, entre ces deux états, l'entropie a augmenté approximativement de $Nk_B \ln 2$. Nous disposons donc d'une explication plausible de l'expansion du gaz : le gaz s'engouffre dans le compartiment de droite car il est énormément plus probable pour les molécules de se trouver n'importe où dans l'enceinte une fois la cloison retirée, plutôt que de rester rassemblées dans le compartiment de gauche, la probabilité de cette dernière configuration étant de l'ordre de 2^{-N} , autant dire zéro.

Il semble donc que nous ayons pu établir de façon plausible une corrélation entre croissance de l'entropie et flèche du temps : l'évolution se fait du moins probable vers le plus probable. Mais avons-nous vraiment établi cette corrélation ? Proposé pour la première fois par Boltzmann à la fin du XIX^e siècle, cet argument fut contesté par des physiciens de renom comme Josef Loschmidt ou des mathématiciens comme Henri Poincaré et ensuite Ernst Zermelo. Plus récemment, il a aussi été critiqué par Karl Popper et Ilya Prigogine. Ces critiques reposent sur l'observation que les équations de la mécanique newtonienne sont invariantes par renversement du sens du temps, ou plus simplement renversement du temps. Ce concept n'étant pas intuitif, il est utile d'examiner un cas analogue mais plus simple, celui de la parité, qui consiste à observer l'évolution d'un système physique dans un miroir. De plus, nous verrons ultérieurement une relation où interviennent à la fois la parité et le renversement du temps. Si l'évolution vue dans un miroir représente celle d'un processus physiquement possible, on dit que la parité est conservée, ou que l'évolution est invariante par parité, ou encore que la parité est une *symétrie* de l'évolution. Cette définition semble absurde, car si vous êtes droitier, le miroir vous renvoie l'image d'un gaucher et, superficiellement au moins, la parité n'est pas conservée ! Mais être gaucher est physiquement possible, et un observateur qui ne sait pas si vous êtes droitier ou gaucher ne pourra pas distinguer entre la vision directe qu'il a de vous et celle de votre image dans un miroir. Les interactions (ou forces) de la vie quotidienne, les interactions électromagnétiques et gravitationnelles, sont invariantes par parité. Cependant, en 1956, Chen Nin Yang et Tsun Dao Lee, en s'appuyant sur l'examen de propriétés des mésons K , des particules dont la nature importe peu ici, prédirent que les interactions faibles, celles qui sont responsables de la radioactivité β , n'étaient pas invariantes par parité, une prédiction jugée absurde à l'époque par nombre de physiciens dont Wolfgang Pauli. La violation de la parité fut observée expérimentalement un an plus tard par Chien-Shiung Wu, connue y compris aux États-Unis sous le nom de « Madame » Wu. L'expérience consistait à observer la désintégration de noyaux atomiques de $^{60}\text{Cobalt}$ polarisés, c'est-à-dire de noyaux auxquels on peut associer un sens de

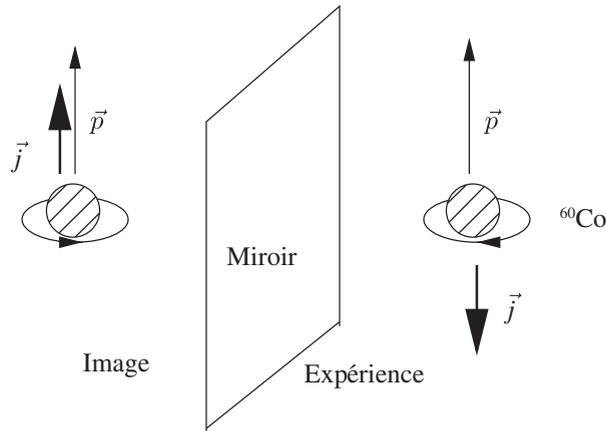


FIGURE 7.9. L'expérience de désintégration du cobalt polarisé et son image dans un miroir. \vec{p} représente l'orientation moyenne de l'émission des électrons et \vec{j} l'orientation de l'axe vertical.

rotation dans un plan perpendiculaire à l'axe vertical (figure 7.9). Cette désintégration conduit à un état final composé d'un état excité du noyau de $^{60}\text{Nickel}$, d'un électron (e^-) et d'un antineutrino ($\bar{\nu}$) :



Le sens de rotation des noyaux dans un plan perpendiculaire à l'axe vertical de la figure 7.9 permet de définir une orientation de cet axe vertical, noté \vec{j} sur la figure 7.9. Par exemple, on peut décider de l'orienter du bas vers le haut si le sens de rotation des noyaux est inverse de celui des aiguilles d'une montre posée dans un plan horizontal. Il s'agit bien évidemment d'une *convention*, et on aurait aussi bien pu choisir la convention contraire. On constate que les électrons sont émis préférentiellement vers le haut, donc dans la direction de l'axe vertical, et l'image de l'expérience dans un miroir ne représente pas un processus physiquement possible. En effet, dans le miroir, le sens de rotation est inversé, et il en est de même pour l'orientation de l'axe vertical, ce qui entraîne que dans le miroir les électrons sont émis préférentiellement dans la direction opposée à l'axe et non dans la direction de l'axe. Le processus examiné dans l'expérience n'est donc pas invariant par l'opération *parité* ; nous noterons P cette opération. Si l'on remplace dans l'expérience la matière par de l'antimatière, les noyaux de $^{60}\text{Cobalt}$ par des noyaux d'anticobalt, les électrons (e^-) par des positrons (e^+), (les antiparticules des électrons utilisés en médecine dans la tomographie électron-positron), et les antineutrinos ($\bar{\nu}$) par des neutrinos (ν), alors on retrouve - approximativement ainsi qu'on le verra ultérieurement - l'invariance

par parité : l'évolution vue dans le miroir serait celle de la désintégration de noyaux d'anticobalt, où les positrons de désintégration sont émis préférentiellement dans la direction opposée à l'axe. Si l'on note C l'opération qui transforme toute particule en son antiparticule, opération appelée conjugaison particule-antiparticule, la physique est – approximativement – invariante par l'opération combinée de la parité et de la conjugaison particule-antiparticule, une opération notée CP .

Revenons au renversement du temps en considérant par exemple la trajectoire d'une des molécules du gaz de la figure 7.6, prise entre deux collisions. On suppose cette molécule ponctuelle afin de simplifier le tracé de la figure 7.10 qui montre sa trajectoire entre l'instant $t = 0$ et l'instant $t = \tau$. À l'instant τ , on inverse les vitesses de toutes les molécules, y compris la vitesse de celle qui a été sélectionnée : si la molécule possédait une vitesse \vec{v}_1 au temps τ , sa vitesse devient $-\vec{v}_1$. Cette opération est bien sûr impossible à réaliser dans un gaz, il faudrait ce que l'on peut appeler un « démon de Loschmidt » pour l'effectuer, mais c'est possible dans une simulation numérique. Nous constatons alors que la molécule « revient sur ses pas », elle suit sa trajectoire initiale mais en sens inverse et se retrouve au temps $t = 2\tau$ à sa position initiale, mais avec une vitesse

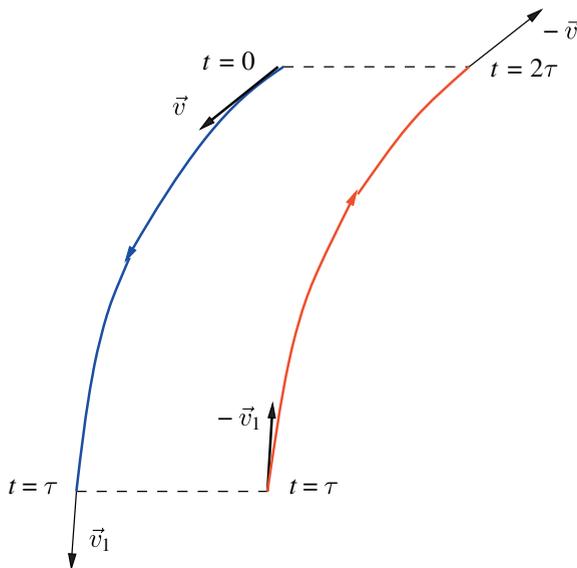


FIGURE 7.10. Trajectoire d'une molécule (en bleu) entre deux collisions, entre les instants $t = 0$ et $t = \tau$. Au temps $t = \tau$, on inverse la vitesse de toutes les molécules du gaz. La molécule considérée parcourt sa trajectoire en sens inverse (en rouge) et se retrouve au temps $t = 2\tau$ à sa position initiale avec une vitesse opposée à sa vitesse initiale.

opposée : si sa vitesse était \vec{v} , elle devient $-\vec{v}$. Si l'on observe une telle situation, alors la dynamique est dite invariante par renversement du temps. L'analogie du miroir qui nous avait permis d'illustrer l'opération parité est celle d'un film projeté à l'envers. Si nous filmons l'évolution d'un processus physique et que nous projetons le film à l'envers, nous dirons que l'évolution est invariante par renversement du temps si la projection du film donne l'image d'un processus physiquement possible. Remarquons que l'expression « renversement du temps » est quelque peu trompeuse. Ce n'est pas le temps que l'on inverse, mais comme dans l'exemple ci-dessus, c'est le sens de l'évolution d'un processus physique.

Les interactions de la vie quotidienne, électromagnétiques et gravitationnelles, sont invariantes par renversement du temps, mais il existe un type d'interaction, extrêmement faible, qui viole cette invariance. Cette interaction a été observée pour la première fois en 1964 dans la physique des mésons *K*. Mais cette observation de la violation du renversement du temps était indirecte : la violation observée était en fait celle de la symétrie *CP*. Cependant, il existe une invariance à laquelle tous les physiciens théoriciens croient fermement, c'est celle correspondant à l'invariance par le produit combiné de *C*, *P* et *T*, où *T* désigne le renversement du temps : c'est l'invariance *CPT*. En effet, si cette invariance n'était pas respectée, il serait tout simplement impossible de formuler une théorie des particules élémentaires compatible avec la relativité restreinte. Si l'invariance par l'opération *CP* est violée, alors nécessairement l'invariance par *T* doit l'être aussi, sinon l'invariance par l'opération *CPT* ne pourrait pas être respectée. Cette situation était relativement peu satisfaisante, car la preuve de la violation de l'invariance par renversement du temps dépendait d'une hypothèse théorique. Elle a perduré jusqu'en 2012, année où une expérience réalisée à Standford en Californie, utilisant des mésons *B* et dont le principe repose sur le célèbre mécanisme EPR (Einstein, Podolsky et Rosen), a permis de montrer directement la violation de l'invariance par renversement du temps, et aussi de donner des vérifications supplémentaires de l'invariance *CPT*. Ces interactions violant l'invariance par renversement du temps ont peut-être joué un rôle au tout début de l'Univers, dans la séparation entre la matière et l'antimatière, mais elles sont strictement sans influence à notre échelle, et d'ailleurs, même si elles étaient importantes, elles ne seraient d'aucun secours pour expliquer la flèche du temps.

Nous sommes donc confrontés au paradoxe suivant : étant donné que les interactions élémentaires entre molécules sont invariantes par renversement du sens du temps, autrement dit étant donné que la dynamique moléculaire ne contient aucun biais temporel, comment se fait-il que les processus de la vie courante soient irréversibles et que la projection à l'envers d'un film nous montre

des évolutions absurdes ? Le film projeté à l'envers ne donne pas l'impression de processus physiquement possibles, en contradiction apparente avec l'invariance par renversement du temps. Dans notre exemple magique, celui de l'expansion d'un gaz, une fois que les molécules se sont réparties dans l'enceinte, on ne les voit jamais se regrouper spontanément dans le compartiment de gauche, alors qu'un film projeté à l'envers montrerait ce regroupement. Cependant une telle évolution n'est pas strictement impossible, elle est simplement hautement improbable. La vision de Boltzmann est simple mais subtile : l'irréversibilité n'est pas incompatible avec l'invariance par renversement du temps, l'évolution obtenue en renversant le temps n'est pas strictement interdite, elle est seulement extrêmement improbable. Il faudrait attendre des temps largement supérieurs à l'âge de l'Univers avant d'observer une telle évolution. Il en est de même pour l'argument de Poincaré, fondé sur l'observation mathématique qu'un système fini et régi par une dynamique hamiltonienne doit nécessairement repasser arbitrairement près d'une situation initiale, ce que l'on appelle une *réurrence de Poincaré*. En raison de ces récurrences, les molécules du gaz de notre exemple magique devraient, un jour ou l'autre, se retrouver approximativement toutes dans le compartiment de gauche. C'est exact en théorie, mais il faudrait attendre un temps largement supérieur à l'âge de l'Univers pour observer cette récurrence, ce que les mathématiciens et les physiciens de la fin du XIX^e cités ci-dessus avaient parfaitement compris. Le point crucial dans l'argument de Boltzmann était d'ajouter une hypothèse probabiliste à la mécanique déterministe de Newton, ce qui était une sorte d'oxymore : c'est la raison pour laquelle cet argument eut beaucoup de difficulté à s'imposer auprès de ses contemporains, et il rencontre encore aujourd'hui des incompréhensions et des résistances. Si nous étions capables de renverser exactement les vitesses de toutes les molécules du gaz occupant l'ensemble de l'enceinte (figure 7.11), alors toutes les molécules regagneraient le compartiment de gauche en un temps fini. C'est impossible d'inverser les vitesses pour un gaz réel, mais possible dans une simulation numérique. On constate alors que de très petites erreurs au moment de l'inversion des vitesses font que le retour des molécules dans le compartiment de gauche ne s'effectue que partiellement. En effet la *dynamique est chaotique*, c'est-à-dire que des conditions initiales qui diffèrent de façon minime donnent des comportements ultérieurs très différents : c'est la *sensibilité aux conditions initiales*. Il est tentant de relier l'irréversibilité au chaos, comme cela a été proposé par exemple par Prigogine, mais le chaos n'a rien à voir avec l'irréversibilité. Le comportement de trois boules de billard peut parfaitement être chaotique, mais le film de leur mouvement projeté à l'envers semblera tout à fait plausible. L'irréversibilité est liée au très grand nombre de degrés de liberté, et un système de trois boules de billard ne peut pas exhiber un comportement irréversible.

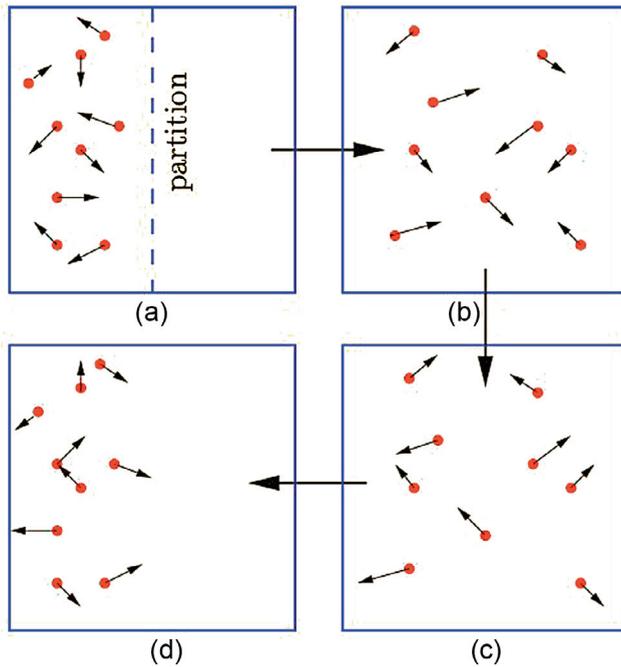


FIGURE 7.11. Expansion irréversible d'un gaz (a) → (b) et processus obtenu par renversement du sens du temps (c) → (d). (a) La cloison est retirée au temps $t = 0$. (b) Le gaz remplit uniformément l'enceinte au temps $t = \tau$. (c) Les vitesses sont inversées à $t = \tau$ par un « démon de Loschmidt ». (d) À $t = 2\tau$, le gaz est revenu dans le compartiment de gauche avec une inversion des vitesses des molécules !

Il semble donc que nous ayons réussi à identifier une flèche du temps grâce à l'irréversibilité des processus macroscopiques, ou grâce à la croissance de l'entropie, mais est-ce vraiment le cas ? Il reste un problème dont Boltzmann était parfaitement conscient : comment se fait-il que l'on puisse disposer initialement d'un état de basse entropie ? Dans un Univers en équilibre thermique, il n'existe pas de flèche du temps et un état de basse entropie ne pourrait provenir que d'une fluctuation, mais une explication par des fluctuations conduit à des paradoxes inextricables. Par exemple, ainsi que l'observe David Albert dans *Time and Chance*, il serait bien plus facile pour une fluctuation de fabriquer des faux documents et des faux souvenirs que de produire les événements réels dont ces documents et ces souvenirs sont censés être l'enregistrement. Il est beaucoup plus économique de créer un cerveau humain, le vôtre, où seraient imprimés dans ses neurones tous les événements qui se sont déroulés sous l'Empire romain, plutôt que de créer l'Empire romain lui-même avec tous les événements qui se sont déroulés sur plusieurs siècles. Un autre problème est que le décompte

des configurations ne comporte pas de biais temporel. Comme l'ont souligné par exemple Roger Penrose et Huw Price en suivant Boltzmann, si en nous appuyant sur ce décompte nous prédisons un état d'entropie élevé dans le futur, nous prédisons aussi un état d'entropie élevé dans le passé, une prédiction qui n'est pas exactement la bienvenue. La solution que propose Richard Feynman dans *Le caractère de la loi physique* est la suivante :

« Il est nécessaire d'ajouter aux lois physiques l'hypothèse que l'Univers a été, dans un sens technique, plus ordonné dans le passé qu'aujourd'hui ».

Autrement dit, il est nécessaire de disposer d'un mécanisme assurant que l'entropie de l'Univers a été très faible par le passé. Une explication plausible, mais non prouvée aujourd'hui, a été proposée par Roger Penrose. Elle repose sur une propriété unique des forces de gravitation, qui d'une part sont uniformément attractives et d'autre part ne sont limitées en intensité par aucun mécanisme de saturation. Sous l'effet de ces forces, la matière a tendance à se condenser en formant des objets de plus en plus massifs, en fin de compte des trous noirs (§ 5.3). En présence de forces de gravitation, et contrairement à l'intuition acquise avec des forces usuelles, l'entropie maximale ne correspond pas à une distribution de matière uniforme, mais à une condensation en objets massifs isolés les uns des autres, vraisemblablement des trous noirs. L'entropie d'un trou noir a été calculée par Jakob Bekenstein et Stephen Hawking, qui ont obtenu le résultat suivant dans le cas d'un trou noir sphérique :

$$S = \frac{\pi k_B A c^3}{2hG} \quad A = 4\pi r_S^2 = 16\pi \left(\frac{GM}{c^2} \right)^2, \quad (7.7)$$

où A est l'aire de l'horizon du trou noir, qui s'exprime en fonction de sa masse M ou de son rayon de Schwarzschild r_S . La formule (7.7) est remarquable : d'une part elle fait intervenir les trois constantes fondamentales de la physique, la vitesse de la lumière c , la constante de Planck h et la constante de gravitation G ; d'autre part elle montre que l'entropie d'un trou noir est proportionnelle à son aire et non au volume comme c'est le cas usuel. Les données récentes sur le fond diffus cosmologique montrent que 380 000 années après le Big Bang, donc au tout début de l'Univers, la distribution de matière était remarquablement uniforme, avec des fluctuations n'excédant pas 1/100 000 en valeur relative (figure 2.9). Cette distribution correspondait à un état de basse entropie et depuis l'entropie de l'Univers a crû continûment par condensation gravitationnelle. Des estimations qualitatives de Roger Penrose à partir de (7.7) montrent que l'entropie d'un ensemble de trous noirs serait *énormément* plus grande que celle contenue dans le fond diffus cosmologique. Reste à justifier le fait que

l'Univers se trouvait initialement dans un état extrêmement improbable, mais qu'une telle justification soit vraiment nécessaire fait encore aujourd'hui l'objet de débats d'experts.

7.5 Entropie et information

On dit souvent que l'entropie est une mesure du désordre. C'est un énoncé séduisant, mais il faut l'interpréter avec prudence. Considérons un jeu de bridge de 52 cartes rangé dans l'ordre canonique de l'as de pique au deux de trèfle. Ce jeu étalé sur une table va nous paraître parfaitement ordonné. Supposons qu'il soit battu par un prestidigitateur très habile capable de garder la mémoire de chacun de ses gestes. Le prestidigitateur va étaler les cartes sur la table dans un ordre qui nous semblera parfaitement aléatoire. Mais ce désordre, pour le prestidigitateur, contient un ordre caché : il sait comment il est parvenu à ce désordre en partant d'un ordre initial, un ordre qu'il sera capable de reconstituer en revenant en arrière. Notre notion de l'ordre d'un jeu de cartes bien rangé ou du désordre d'un jeu de cartes une fois battu est éminemment subjective. Comme l'avait déjà remarqué Maxwell :

« La confusion (dans le sens de l'aléatoire), tout comme le terme associé d'ordre, n'est pas une propriété des objets matériels en eux-mêmes, elle n'existe qu'en relation avec l'esprit qui les perçoit. »

Réexaminons l'expansion d'un gaz sur la figure 7.11 à la lumière de cette discussion. Pour un observateur qui ne connaît pas l'histoire de cette expansion, la situation de la figure 7.11b semblera parfaitement désordonnée, et cet observateur attribuera à l'état du gaz une entropie élevée. Mais un autre observateur au fait de l'histoire du gaz remarquera que cette situation est issue d'une configuration très particulière, celle où le gaz était initialement confiné dans le compartiment de gauche. La probabilité de la configuration 7.11b est exactement la même que celle de la configuration 7.11a, étant donné que les N molécules ont suivi une trajectoire déterminée par les lois de Newton, et chaque configuration de 7.11b est en correspondance biunivoque avec une configuration de 7.11a. La probabilité de la configuration 7.11b est très faible, de l'ordre de 2^{-N} . Mais, contrairement au prestidigitateur capable de suivre chaque carte, nous ne pouvons pas suivre chaque molécule et la configuration 7.11b contient des corrélations microscopiques entre positions et vitesses des molécules reflétant la spécificité de l'état initial, corrélations auxquelles nous n'avons pas accès. Dans l'expansion du gaz, toute l'information contenue dans ces corrélations a été en pratique perdue : ces corrélations sont absentes pour un état typique du gaz occupant toute l'enceinte. C'est pourquoi, si nous disposons uniquement

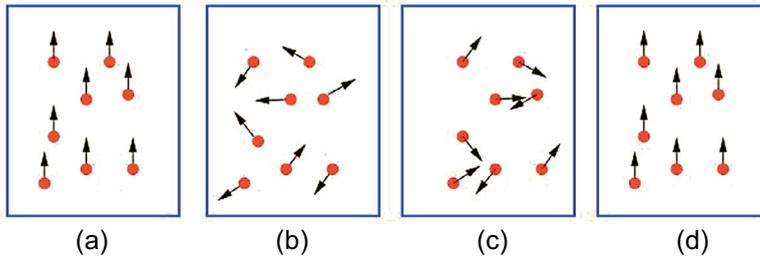


FIGURE 7.12. Une expérience d'écho de spin. (a) Tous les spins sont alignés au temps $t = 0$. (b) Les spins évoluent dans un champ magnétique aléatoire dans l'espace mais constant dans le temps. (c) Au temps $t = \tau$, les spins subissent une rotation 180° . (d) Au temps $t = 2\tau$, on retrouve l'alignement initial des spins.

de l'information macroscopique suivante : le gaz occupe tout le volume de l'enceinte, nous devons attribuer à la situation 7.11b une entropie élevée. L'entropie d'un système physique est donc intimement liée à l'information que nous avons sur ce système.

Dans la cas de l'expansion d'un gaz, il est impossible de reproduire l'analogie de ce que peut faire le prestidigitateur qui, connaissant l'ordre caché sous le désordre apparent, est capable de revenir à l'ordre canonique initial. Cependant il existe plusieurs exemples en physique où cela est possible. Le cas le plus connu est celui de l'expérience *d'écho de spin* qui, loin d'être une curiosité académique, possède des applications importantes en RMN (résonance magnétique nucléaire) et en IRM (imagerie par résonance magnétique). Le schéma d'une expérience d'écho de spin est illustré sur la figure 7.12. En RMN comme en IRM, on utilise les propriétés du moment magnétique de noyaux atomiques – d'où l'adjectif nucléaire (N), supprimé dans le cas de l'IRM afin de rester dans le politiquement correct – : ces moments magnétiques sont associés au mouvement de rotation des noyaux sur eux-mêmes, appelé spin, que l'on peut représenter graphiquement par une petite flèche. En effet, ainsi que nous l'avons vu sur la figure 7.9, on peut associer une orientation dans l'espace, une flèche, à un sens de rotation. On part d'un état initial où tous les spins d'un ensemble sont orientés suivant la verticale (figure 7.12a). Un champ magnétique constant perpendiculaire au plan de la figure et appliqué pendant un intervalle de temps τ les fait tourner d'un certain angle autour de la direction du champ. Comme ce champ magnétique et le milieu où se trouvent les spins présentent des inhomogénéités incontrôlables, l'angle de rotation va varier d'un spin à l'autre de façon aléatoire (figure 7.12b). Au temps τ , un champ magnétique oscillant vertical appliqué pendant une durée appropriée sur un intervalle de temps très petit par rapport à τ fait tourner tous les spins de 180° : figure 7.12c. Cette opération est

l'analogie de l'inversion des vitesses de la figure 7.11. Sous l'influence du champ horizontal, les spins vont maintenant tourner en sens inverse du sens initial, et comme les inhomogénéités n'ont pas varié, ils vont au bout d'un intervalle de temps τ retrouver exactement leur orientation verticale initiale (figure 7.12d).

Un observateur qui ne connaît pas l'histoire du système et auquel on présente la configuration 7.12b y verra un ensemble de spins complètement désordonnés et attribuera à l'ensemble de ces spins une entropie élevée. Mais le physicien qui effectue l'expérience sait que cette configuration apparemment désordonnée présente un ordre caché et, pour lui, l'état du système est de basse entropie. Tout comme le prestidigitateur reconstitue l'ordre caché des cartes à partir d'un jeu en apparence désordonné, ce physicien, en effectuant une rotation des spins de 180° et en appliquant un champ magnétique horizontal pendant une durée τ , peut reconstituer l'ordre initial des spins. Nous vérifions à nouveau que l'entropie que nous attribuons à un système physique dépend de façon cruciale de l'information que nous avons sur lui.

8

Renversement du temps et ondes

Nous avons vu au chapitre 7 que, bien que les équations de la physique microscopique et de la mécanique soient parfaitement réversibles, l'apparente irréversibilité du monde macroscopique s'explique avant tout en raisonnant en termes de probabilité. Si le comportement d'un gaz en expansion paraît irréversible, c'est que pour les nombreuses molécules de ce gaz la probabilité de se retrouver plus tard dans leur état initial est infinitésimale. En partant d'un état relativement ordonné (un état de basse entropie où toutes les molécules sont situées dans un petit compartiment), le gaz se retrouve plus tard, une fois la cloison retirée, dans des états beaucoup plus désordonnés et il faudrait attendre un temps extrêmement grand (le temps de Poincaré) pour avoir une chance de retrouver toutes les particules dans leurs positions initiales.

Le problème qui nous intéresse ici n'est pas de rester passif devant l'évolution naturelle d'un système de molécules, mais de s'interroger avec un regard d'ingénieur et de physicien sur la possibilité de fabriquer un dispositif qui puisse permettre de réaliser sur commande des expériences que l'on fait évoluer dans les deux sens : de l'ordre vers le désordre et ensuite de façon parfaitement symétrique du désordre vers l'ordre. Est-il possible de fabriquer un dispositif qui contrôle notre système de particules pour pouvoir inverser la dynamique de ces particules sur commande ?

La clef d'un tel dispositif a été donnée par le physicien Josef Loschmidt dans ses discussions avec Boltzmann. Il proposa une expérience de pensée que l'on peut visualiser en introduisant des petits démons, baptisés démons de Loschmidt par analogie avec les démons de Maxwell, qui seraient capables à un instant donné de mesurer les positions et les vitesses de toutes les particules du gaz, de les arrêter, puis dans un deuxième temps de renverser exactement toutes les

vitesse des particules (figure 7.11). Une telle opération qui consiste à préparer le système de particules dans un nouvel état initial aurait comme conséquence un renversement parfait de la dynamique des particules qui n'auraient pas d'autre possibilité que de retracer pas à pas leur vie passée, c'est-à-dire de parcourir leur trajectoire en sens inverse. Pour un système de particules indestructibles, refaire leur trajectoire en sens inverse c'est, en quelque sorte, revivre leur passé !

Il faut remarquer ici que, pour effectuer cette opération, les démons de Loschmidt ont besoin à la fois d'acquérir de l'information sur le système de particules (en mesurant l'état du système à un instant donné) et d'injecter de l'énergie et de la quantité de mouvement à toutes ces particules pour les faire repartir dans des directions opposées. Il est important de noter qu'en intégrant dans l'étude du système total la présence des démons (leurs appareillages de mesure et de contrôle), le second principe de la thermodynamique sera toujours bien respecté. Le retour du système de particules dans un état de basse entropie est compensé par la croissance de l'entropie (acquisition d'information) dans l'appareillage de mesure et de contrôle.

La fabrication de ces petits démons de Loschmidt pose néanmoins un problème majeur, comme nous l'avons vu au chapitre 7. Pour un système formé d'un grand nombre de particules, la moindre erreur dans la mesure des vitesses et des positions ou dans la préparation du nouvel état initial fait intervenir le caractère chaotique de la dynamique. C'est ce qui s'observe facilement dans une simulation numérique des trajectoires d'un grand nombre de particules régies par la dynamique newtonienne. Car, dans un ordinateur, on traite les positions et les vitesses des particules avec une précision finie qui dépend du nombre de bits sur lesquels les signaux sont échantillonnés. On se heurte alors au problème du *chaos déterministe* et au concept de *dynamique non linéaire*. La moindre erreur dans la mesure ou dans la préparation du système de particules s'amplifie de façon exponentielle avec le temps. Typiquement, une erreur relative de mesure de 10^{-15} suffit, sur une machine travaillant avec une précision de 64 bits, pour perdre l'évolution réversible après 10 collisions. L'erreur initiale est tellement amplifiée qu'après quelques collisions les particules ne retrouvent plus leurs anciens partenaires. Cette sensibilité exponentielle aux erreurs fixe les limites d'une telle expérience de renversement des trajectoires. C'est l'aspect non linéaire de la dynamique d'un grand système de particules qui est en cause. C'est le même phénomène qui interdit aussi de prévoir le climat une semaine à l'avance. Devant les limites imposées par la dynamique non linéaire, il semble donc que, bien que la physique microscopique soit réversible, on ne pourra jamais, avec nos moyens limités, renverser le temps dans une expérience de physique à grand nombre de particules.

Heureusement, ce qui vient d'être dit ne concerne que le monde des particules classiques régies par la dynamique non linéaire. Si l'on s'intéresse à l'évolution d'une onde au cours du temps (plutôt que d'une particule), la situation est beaucoup plus favorable. Or les ondes se rencontrent dans tous les domaines de la physique, depuis la mécanique quantique en passant par l'électromagnétisme ou par l'acoustique, et la propagation des ondes s'effectue généralement en régime linéaire, ce qui implique qu'une petite erreur de mesure ou de préparation de l'état initial d'un champ ondulatoire n'est pas amplifiée au cours de l'évolution de l'onde.

8.1 Acoustique et renversement du temps

C'est d'abord dans le domaine de l'acoustique que les premières expériences de renversement du temps ont été réalisées à la fin des années 1980. On sait que les ondes acoustiques se propagent dans des systèmes de particules dont le nombre est extraordinairement grand. Les systèmes de particules qui nous intéressent en acoustique peuvent être des gaz, des liquides ou des solides. À côté des mouvements d'agitation désordonnés de chaque particule associés à la température ambiante, il existe des mouvements collectifs de ces particules qui sont décrits par les ondes acoustiques. Lorsqu'une personne parle, le mouvement vibratoire de ses cordes vocales se répercute sur les molécules de l'air et de proche en proche se répand un mouvement de ces molécules autour de leur position d'équilibre. Ce qui est important, c'est qu'il s'agit de mouvements d'ensemble, où toutes les molécules bougent en phase sur des distances qui sont caractérisées par la *longueur d'onde* de l'onde acoustique. Sur une échelle de l'ordre de quelques dixièmes de longueur d'onde, c'est l'ensemble des particules qui bouge de façon coordonnée et non pas chaque particule indépendamment des autres. Le nombre de degrés de liberté nécessaire pour décrire complètement ces mouvements collectifs de particules devient très raisonnable. Ces mouvements d'origine acoustique peuvent être considérés, dans la plupart des cas, comme réversibles. Tant que les mouvements mécaniques associés à ces ondes ne se transforment pas en mouvement désordonné à petite échelle (chaleur), le processus peut être considéré comme adiabatique et donc invariant par renversement du temps. On peut associer à ces mouvements acoustiques des variations spatio-temporelles de la densité volumique du fluide à laquelle est associée une pression acoustique $p(\vec{r}, t)$, où \vec{r} est un vecteur dont les composantes sont x, y, z et t désigne le temps. Ce sont des fonctions continues de l'espace et du temps qui vérifient une équation linéaire aux dérivées partielles du deuxième ordre. Dans un milieu hétérogène où la vitesse du son $c(\vec{r})$ dépend de la position, l'équation

vérifiée par la pression acoustique s'écrit :

$$\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2} - \frac{1}{c^2(\vec{r})} \frac{\partial^2 p}{\partial t^2} = 0. \quad (8.1)$$

C'est le type même d'une équation invariante par renversement du temps, où la dérivée par rapport au temps n'apparaît que sous la forme d'une dérivée seconde. Cela implique que si l'on trouve une solution du type $p(\vec{r}, t)$, alors il existe toujours une seconde solution $p(\vec{r}, -t)$ duale de la première, car la dérivée seconde de $p(\vec{r}, -t)$ est identique à celle de $p(\vec{r}, t)$. Quelles que soient les déformations subies par une onde acoustique rayonnée par une source dans un milieu complexe et hétérogène caractérisé par une célérité dépendant de la position $c(\vec{r})$ (réfraction, diffusion multiple, réverbération), il existe alors toujours en théorie une onde duale capable de parcourir en sens inverse tous les chemins complexes et qui converge exactement à la source.

Les acousticiens ont aussi la chance de disposer de deux instruments très performants, qui peuvent être reproduits à des milliers d'exemplaires, le microphone et le haut-parleur. Au passage d'une onde acoustique, un microphone génère un courant électrique proportionnel à la composante, le long de la direction de propagation, de la vitesse des particules interceptées par sa membrane, alors que réciproquement la membrane d'un haut-parleur se met en mouvement avec une vitesse proportionnelle au courant électrique qu'on lui applique. Elle communique cette vitesse aux molécules de l'air avec laquelle elle est en contact. En superposant localement les deux fonctions microphone puis haut-parleur, on réalise ainsi une espèce de petit démon de Loschmidt. Or nous disposons de petits composants, les *transducteurs piézoélectriques*, qui remplissent cette double fonction (transducteurs réversibles). La piézoélectricité, propriété découverte par Pierre Curie, est un phénomène réversible qui se manifeste dans certains cristaux pour lesquels une contrainte appliquée engendre l'apparition d'un courant électrique, et dans lesquels réciproquement l'injection d'un courant électrique déforme le cristal, ainsi que nous l'avons vu au § 3.3.2.

Fort de ces dispositifs, une expérience idéale de renversement du temps à la *Loschmidt* pourrait s'envisager de la façon suivante. Il s'agirait d'abord de mesurer à un instant donné dans tout le volume de propagation les vitesses induites par l'onde acoustique. Pour cela, on peut imaginer disposer à l'intérieur de tout le volume de très nombreux transducteurs piézoélectriques répartis sur une grille tridimensionnelle dont le pas serait typiquement d'une demi-longueur d'onde. Ces transducteurs pourraient, à un instant t , mesurer les vitesses d'origine acoustique, puis plus tard, lorsque l'onde acoustique initiale aurait disparu, activer à un nouvel instant leur mode haut-parleur afin de générer des mouvements particuliers de vitesse inversée. Bien entendu, pour de nombreuses

raisons, cette expérience de pensée à la Loschmidt est parfaitement irréaliste. Il faudrait d'une part disposer de transducteurs capables de répondre et d'agir de façon instantanée, ce qui nécessiterait des bandes passantes infinies. Il faudrait d'autre part avoir en chaque point de la grille trois petits transducteurs piézo-électriques orientés dans les trois directions afin de recréer, par superposition, les trois composantes des vecteurs vitesses.

8.2 Principe de Huygens et miroir à retournement temporel (MRT)

C'est ici qu'un théorème fondamental de la physique des ondes apporte une énorme simplification. Les ondes ont une propriété tout à fait remarquable, mise en évidence pour la première fois par Christian Huygens. Il suffit de connaître l'évolution temporelle d'une onde sur les points d'une surface à deux dimensions pour pouvoir prévoir son évolution dans tout le volume entouré par cette surface. Une information bidimensionnelle sur un champ ondulatoire suffit pour retrouver la valeur du champ dans les trois dimensions. C'est le *principe de Huygens* et ses variantes (principe Huygens-Fresnel et théorème intégrale de la diffraction de Kirchhoff-Helmholtz) qui sont à la base du concept d'holographie développé dans les années 1960 par Dennis Gabor. C'est en exploitant cette remarquable propriété que le concept de *miroir à retournement temporel* (MRT) a été proposé dans les années 1980. Il suffit de tapisser les parois d'une surface entourant le volume d'intérêt de transducteurs piézoélectriques orientés vers l'intérieur (figure 8.1) et de les relier chacun à une mémoire électronique de grande taille au moyen de convertisseurs A/D et D/A (analogiques/digitaux). Toute l'information utile sur les détails du champ ondulatoire peut être capturée si l'on échantillonne spatialement la surface avec des transducteurs séparés les uns des autres d'une demi-longueur d'onde. En outre, si les transducteurs sont situés dans le champ lointain des sources, il n'est pas nécessaire de mesurer les trois composantes du champ des vitesses car, dans les liquides et dans les gaz, les ondes acoustiques observées en champ lointain ont un comportement longitudinal (elles sont localement planes), la vitesse des particules du fluide est toujours dirigée suivant la direction de propagation de l'onde. Les transducteurs sont alors orientés pour mesurer la bonne composante de la vitesse et pour réinjecter plus tard un mouvement de vitesse inversée.

Dans une expérience de renversement du temps acoustique, quelque part à l'intérieur du MRT, une source émet, pendant un *temps fini*, une onde acoustique, sonore ou ultrasonore. C'est dans le domaine ultrasonore que les premiers dispositifs ont été réalisés, les longueurs d'onde ultrasonores étant

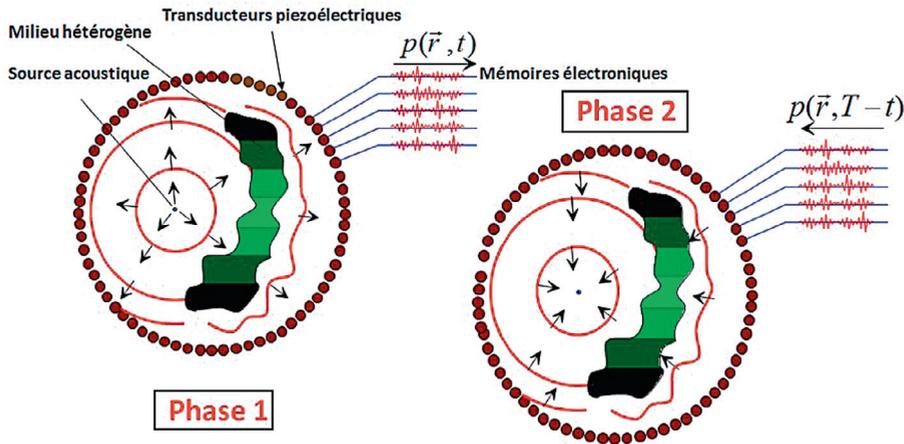


FIGURE 8.1. Dans une première étape, une onde acoustique est rayonnée par une source. Elle se propage à travers un milieu hétérogène et elle est captée par l'ensemble des transducteurs piézoélectriques qui entourent le volume d'intérêt. Ils jouent d'abord le rôle de microphones. Les signaux mesurés sont convertis sous forme digitale et stockés dans des mémoires électroniques. Dans une deuxième étape, tous les signaux captés sont réémis dans une chronologie inversée par les transducteurs qui jouent maintenant le rôle de haut-parleurs. L'onde totale ainsi réémise se refocalise parfaitement sur sa source initiale. Cependant cette onde de type convergente ne s'arrête pas sur la source mais, du fait de la conservation d'énergie, s'effondre au point source et se transforme alors en une onde divergente. C'est l'interférence entre cette onde divergente et l'onde convergente qui crée une tache focale dont la dimension dépend des longueurs d'onde utilisées.

millimétriques, alors que les longueurs d'onde sonores sont plutôt décimétriques. Cette onde divergente va mettre en vibration les molécules du milieu. Lorsque l'onde acoustique atteint les transducteurs, ils fonctionnent d'abord comme des microphones et les signaux électriques résultants sont enregistrés, pendant toute la durée de l'expérience, dans des mémoires électroniques. On enregistre ainsi, au cours du temps, à toutes les positions des microphones, l'évolution de la vitesse des particules (en fait la composante de la vitesse suivant l'axe des microphones). Une fois que l'intérieur de la cavité est devenu silencieux, on passe à la phase de renversement du temps. Il s'agit alors de relire l'ensemble des mémoires électroniques en une sorte de verlan. Plus exactement, on inverse la chronologie des signaux acoustiques reçus par chaque microphone. Les signaux arrivés les derniers sont relus en premier et vice versa. Chaque microphone passe alors en mode haut-parleur et est alimenté, pendant un temps suffisamment long, par les mémoires électroniques. L'ensemble de ces haut-parleurs émet une distribution d'ondes suivant une chronologie inversée et, du fait de la réversibilité des lois de la dynamique, ces ondes interfèrent pour créer une onde acoustique résultante qui n'a pas d'autre possibilité que

de revivre pas à pas toutes les étapes de la vie passée de l'onde initiale pour converger finalement sur la source.

Comparée à une expérience à la *Loschmidt* où un renversement *instantané* des vitesses est pratiqué dans le volume, l'expérience qui vient d'être décrite est singulièrement différente. On utilise ici un nombre bien plus restreint de démons de Loschmidt, mais ce sont des démons d'un genre nouveau, qui possèdent chacun une mémoire. Une des trois dimensions d'espace a été remplacée par la dimension temporelle en équipant les démons d'une mémoire.

Ce qui vient d'être décrit concerne les ondes acoustiques, mais ces raisonnements peuvent être généralisés à d'autres types d'ondes qui n'ont pas nécessairement besoin d'un support matériel fait de particules pour se propager. La seule condition nécessaire est que les équations qui décrivent l'onde soient invariantes par renversement du temps, comme c'est le cas des ondes électromagnétiques dans les milieux non dissipatifs, des vagues à la surface de l'eau ou des ondes de matière quantiques décrites par l'équation de Schrödinger. Toutefois, comme il faut aussi disposer d'antennes réversibles pour ces ondes, l'extension la plus facile de ce concept concerne les ondes électromagnétiques du type micro-onde, car on dispose aisément d'antennes qui mesurent et émettent les différentes composantes des champs électriques et magnétiques. Là encore, si ces antennes sont situées assez loin des sources, on peut simplifier l'expérience en utilisant des antennes qui ne mesurent pas tous les degrés de liberté du champ électromagnétique.

8.3 Causalité et limite de diffraction

Dans ces différentes expériences, on fabrique ainsi physiquement, et quelle que soit la complexité du milieu de propagation, une onde capable de converger pas à pas vers sa source initiale. C'est ce que les physiciens appellent la *solution avancée* du problème du rayonnement d'une source. C'est la solution duale de la solution divergente habituelle qui est dite *solution retardée* et qu'on observe dans la nature.

On sait en effet que, lorsque les physiciens essayent de prévoir le rayonnement d'une source, la symétrie par renversement du temps de l'équation des ondes pose un très sérieux problème. Il existe en effet, mathématiquement, toujours deux solutions possibles : l'une pour laquelle le champ rayonné par la source se répand dans l'espace et n'atteint les points d'observation qu'après le moment où la source a été allumée : c'est l'onde divergente décrite précédemment. Cette solution dite retardée nous paraît naturelle car elle vérifie le principe de causalité. Il existe aussi une deuxième solution qui ne respecte pas la

causalité : la solution « avancée » où le champ ondulatoire traverse l'espace avant même que la source ne soit allumée ! Bien que cette solution nous paraisse parfaitement absurde, il faut savoir que de grands physiciens comme Richard Feynman et John Archibald Wheeler ont essayé d'interpréter avec un certain succès la théorie du rayonnement électromagnétique en tenant compte à la fois des solutions avancées et retardées.

Dans le cas d'une expérience effectuée dans un *univers causal* comme le nôtre, un miroir à retournement temporel fabrique, en temps différé, une solution qui ressemble à pas mal d'égards à la solution avancée. On réalise physiquement le film de l'onde *retardée* à l'envers.

Mais une remarque importante s'impose concernant la réalisation de la solution avancée. Le champ réémis par la cavité à retournement temporel ressemble pendant un certain temps à une onde convergeant vers la source, mais *une onde ne sait pas s'arrêter* (elle a toujours une célérité) et, lorsque l'onde convergente atteint l'endroit où était située la source, elle s'effondre et repart en marche arrière sous la forme d'une onde divergente en raison de la conservation de l'énergie. On observe alors autour de la source l'interférence entre l'onde convergente et cette onde divergente. Ce n'est plus tout à fait l'onde avancée qui est créée, mais c'est *l'onde avancée moins l'onde retardée* : en effet l'onde divergente repart avec un signe opposé à l'onde convergente. Si l'onde qui est utilisée dans l'expérience est sinusoïdale, on démontre alors très simplement qu'autour du point source se crée une interférence constructive entre ces deux ondes sur une zone dont le rayon est égal à la demi-longueur d'onde utilisée. C'est la *limite de diffraction*. Dans le cas où la source initiale émet un rayonnement bref de large contenu spectral (un grand nombre de longueurs d'onde), le champ « retourné » focalise alors sur cette source sous la forme d'une tache dont les dimensions sont au mieux de l'ordre de la moitié de la plus petite des longueurs d'ondes émises.

Le raisonnement que nous venons d'exposer permet, à partir d'un simple argument de réversibilité, de comprendre d'une façon très physique ce qu'on appelle la limite de diffraction en imagerie. On sait que la résolution spatiale des images d'un objet dépend toujours des longueurs d'onde utilisées. En imagerie, le *critère de Rayleigh* spécifie qu'on peut résoudre sur l'image deux points séparés d'une distance égale à la largeur de la *tache focale* du système d'imagerie. Pour un dispositif d'imagerie classique utilisant une lentille, la taille de la tache focale dépend de la longueur d'onde et de l'ouverture angulaire sous laquelle est vue la lentille depuis le point focal. Ici, on n'a pas affaire à une lentille mais à un miroir d'un genre particulier qui renvoie, par retournement temporel, le rayonnement émis par l'objet (la source) sur l'objet lui-même. On réalise ainsi non pas une image virtuelle de l'objet, comme le ferait un miroir classique, mais une image réelle de l'objet sur lui-même. Celle-ci n'est pas parfaite car l'image de chaque

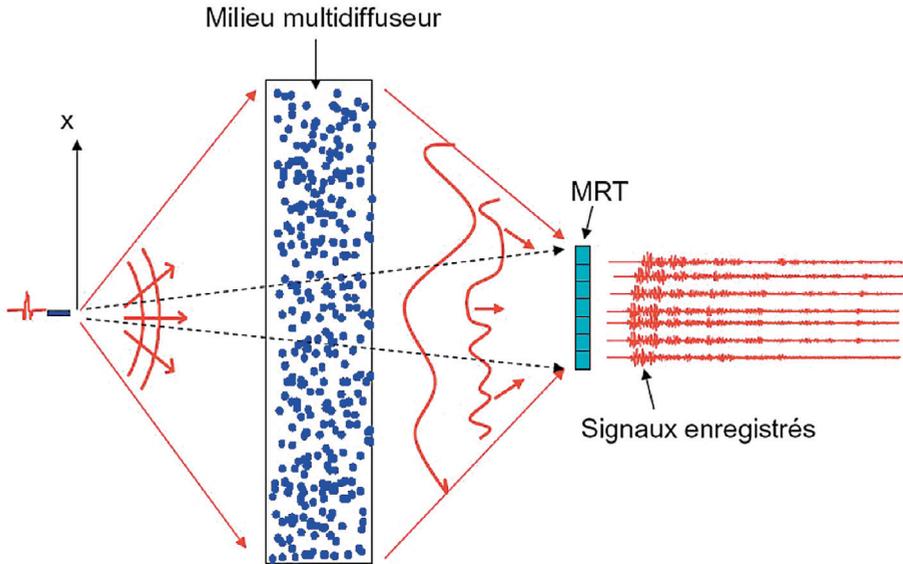


FIGURE 8.2. Schéma de l'expérience en présence d'un milieu multidiffuseur. Le MRT est situé à droite en position verticale et la source initiale est à gauche de la figure. Une impulsion ultrasonore est émise par la source. Le signal acoustique ayant traversé la forêt de tiges métalliques est capté par les 96 éléments du MRT. Dans une deuxième étape, les signaux sont réémis dans une chronologie temporelle inversée vers la forêt de tiges. On note que, vue de la source, l'ouverture angulaire du milieu diffuseur est beaucoup plus grande que celle du MRT.

point de l'objet est en fait une tache dont la dimension est la demi-longueur d'onde.

8.4 Complexité, codage spatio-temporel et degrés de liberté

La réalisation d'un miroir à retournement temporel entourant complètement la source n'est cependant pas très réaliste. Elle nécessiterait de très nombreux transducteurs disposés dans toutes les directions de l'espace en regard de la source. En pratique, on se restreint à des MRT qui n'occupent qu'une ouverture angulaire limitée. En testant les premiers MRT dans le domaine ultrasonore, les chercheurs ont eu une excellente surprise. Pour des signaux à spectre large, plus le milieu dans lequel les ondes se propagent est complexe (diffusion multiple entre des obstacles ou réverbération multiple entre des parois), plus le nombre de transducteurs nécessaires pour refocaliser efficacement une onde devient faible. Dans bien des cas, un seul transducteur suffit et bien que le miroir à retournement temporel devienne alors très petit, il est capable de réaliser

des taches focales dont la dimension est toujours de l'ordre de la demi-longueur d'onde, en violation apparente des lois de diffraction des antennes ! Cet effet a été observé dans les scénarios les plus divers, qui vont de la propagation d'ultrasons de longueur d'onde millimétrique sur des distances de plusieurs décimètres, à la propagation en mer sur plusieurs centaines de kilomètres d'ondes sonores et à la propagation d'ondes électromagnétiques centimétriques sur plusieurs kilomètres. Un MRT agit comme une antenne qui utilise la complexité du milieu environnant pour apparaître beaucoup plus grande qu'elle ne l'est en réalité, et ce ne sont plus la taille ou le nombre d'éléments de l'antenne qui fixent la résolution spatiale mais les propriétés de l'environnement.

Une première expérience dans le domaine ultrasonore illustre bien ce concept. On s'intéresse à la propagation d'un paquet d'ondes ultrasonores à travers un milieu diffuseur de structure aléatoire. Un petit miroir à retournement temporel de 40 mm formé de 96 transducteurs ultrasonores est immergé dans une cuve d'eau (figure 8.2) et il fonctionne autour d'une fréquence centrale de 3 MHz (longueur d'onde $\lambda \sim 0,5$ mm). Entre ce réseau et une très petite source, on introduit un milieu désordonné de grande largeur, formé de 2000 tiges de métal qui résonnent autour de la fréquence de 3 MHz. C'est un milieu particulièrement épouvantable pour les ultrasons. Dans la première phase de l'expérience, la source ponctuelle émet un signal ultrasonore très bref dont la durée est de 1 microseconde (figure 8.3a). Ce train d'ondes ultrasonores se propage à travers la forêt de tiges qui le modifie complètement. Seule une très faible partie de l'onde initiale se propage sans diffusion, alors que la plus grande partie de l'onde initiale se divise en de très nombreuses ondelettes à chaque collision avec les tiges. Le processus peut se répéter de nombreuses fois avec, comme résultante, à la sortie du milieu, un signal ultrasonore très long (300 microsecondes) qui a l'allure d'un bruit complètement aléatoire (figure 8.3b : on parle de coda). Les 96 transducteurs enregistrent ces signaux dont la durée typique est de l'ordre de 300 microsecondes. Après renversement du temps dans les mémoires électroniques, les 96 bruits acoustiques sont réémis dans le milieu. On observe alors le signal ultrasonore qui se repropage à travers l'ensemble des tiges et on constate qu'à l'emplacement de la source initiale le signal ultrasonore résultant dure maintenant 1 microseconde, c'est-à-dire qu'il a exactement retrouvé sa durée initiale (figure 8.3c). Par ailleurs, en explorant latéralement le champ autour de ce point, on constate que la taille de la tache focale obtenue à travers le milieu diffuseur est beaucoup plus fine (figure 8.4) que celle observée dans une expérience de retournement qui serait effectuée en milieu homogène (l'eau uniquement). Pour la géométrie de notre expérience, le gain en résolution est de 35 ! Il s'explique par la différence énorme entre les ouvertures angulaires sous lesquelles sont vus le miroir et le milieu diffuseur par la source.

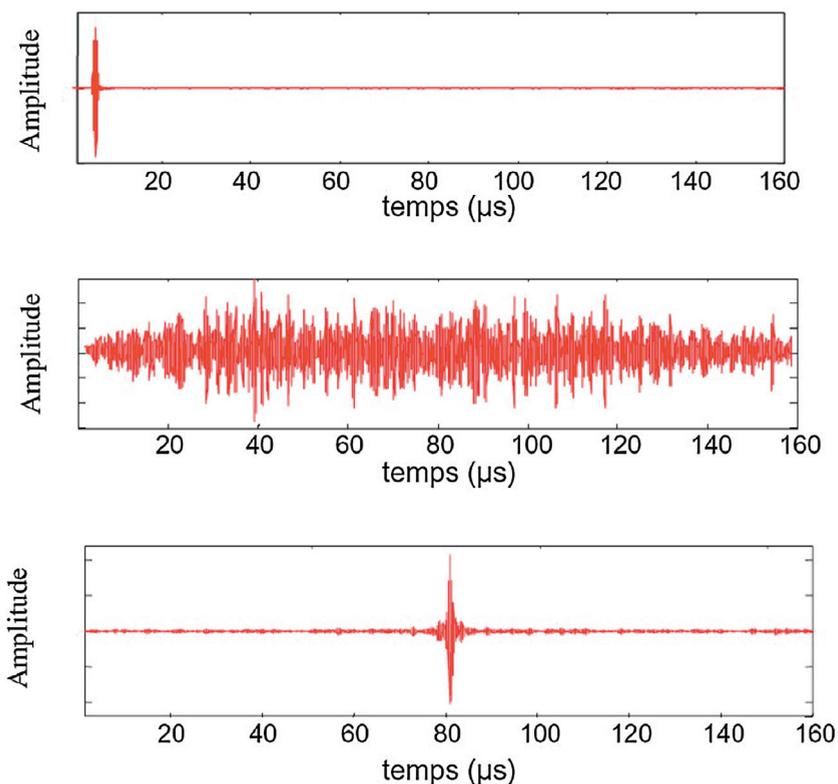


FIGURE 8.3. Figure du haut : signal reçu en provenance de la source par un des transducteurs du MRT en l'absence de milieu multi-diffuseur (à travers l'eau). Figure du milieu : signal reçu (la coda) par le même transducteur en présence du milieu multi-diffuseur. Figure du bas : impulsion réceptionnée par la source (un hydrophone) après retournement temporel et réémission des signaux du MRT à travers le milieu multi-diffuseur. On note la recompression temporelle. Le signal retrouve sa durée d'origine. On parle de focalisation temporelle.

Dans l'expérience effectuée en milieu homogène, on retrouve le résultat classique : la taille de la tache focale dépend de l'ouverture angulaire sous laquelle est vu le miroir du point focal. En présence du milieu diffuseur, c'est maintenant, après retournement temporel, l'ensemble des tiges du milieu qui devient la source d'un rayonnement cohérent focalisant sur le point initial : c'est donc cette nouvelle ouverture angulaire qui fixe la taille de la tache focale et non plus la taille du miroir. C'est l'ensemble du milieu diffuseur qui se comporte comme une lentille. . .

Plus étonnant encore, on peut réduire le nombre d'éléments du miroir. Au lieu d'utiliser les 96 éléments à retournement temporel, on peut n'utiliser qu'un unique élément pour former un miroir à retournement temporel quasi ponctuel.

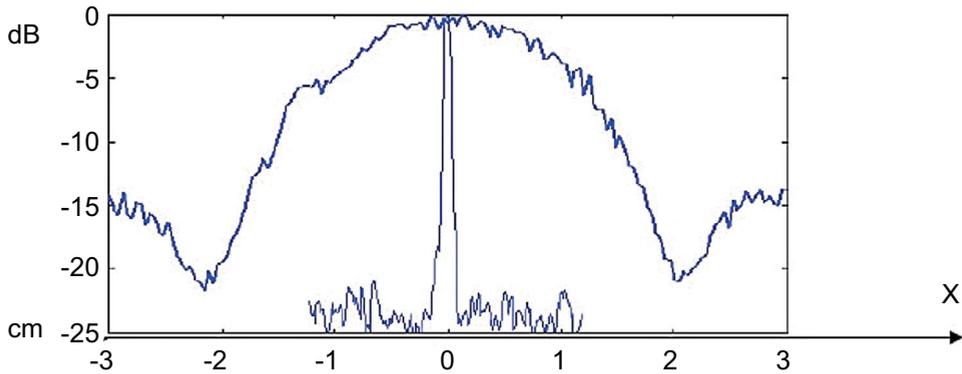


FIGURE 8.4. Formes de la tache focale. Les courbes représentent au temps focal la pression acoustique le long de l'axe x . L'abscisse $x = 0$ correspond à la position de la source initiale (figure 8.2). La tache focale la plus étroite correspond au retournement temporel à travers le milieu diffuseur. La tache focale la plus large correspond à la même expérience de retournement temporel effectuée dans l'eau sans milieu diffuseur. On note une diminution spectaculaire de la largeur de la tache focale d'un facteur 35 qui correspond au rapport des ouvertures angulaires du MRT et du milieu diffuseur vues du point source.

On observe toujours une tache focale de même dimension avec cependant un niveau de lobes parasites plus grand. Un tel résultat peut sembler incompréhensible pour qui est habitué à raisonner sur des ondes monochromatiques comme en optique cohérente. En effet, renverser le sens du temps revient pour une onde sinusoïdale (décrite par 2 nombres : une amplitude A et une phase ϕ) à renvoyer cette onde en conjuguant sa phase (remplacer ϕ par $-\phi$). Or, quelle que soit sa phase, une onde monochromatique émise par une unique antenne de très petite taille (une onde sphérique divergente) ne focalise jamais en traversant un milieu diffuseur. Au contraire, elle donne lieu, après la traversée du milieu, à un champ aléatoire dont les propriétés statistiques sont décrites par la théorie des tavelures, ou *speckles* (figure 8.5). Dans un plan d'observation situé en dehors du milieu diffuseur, le module (ou l'intensité) de ce champ a un profil aléatoire qui contient de nombreuses tavelures : les *grains de speckle* dont la taille caractéristique dépend de la longueur d'onde et de l'ouverture angulaire sous laquelle est vu l'ensemble des tiges. Ces figures sont l'analogie des franges qui résultent de l'interférence de deux ondes planes, mais ici il s'agit de l'interférence de très nombreuses ondes de phase et d'amplitude aléatoires.

Comment se fait-il donc que le code temporel émis par l'unique antenne du MRT refocalise à travers un milieu diffuseur, alors qu'une antenne unique fonctionnant en mode monochromatique en est incapable ? La réponse se trouve dans la bande passante des signaux que l'on traite dans les expériences de retournement temporel. Lorsque l'on utilise un rayonnement initial très bref

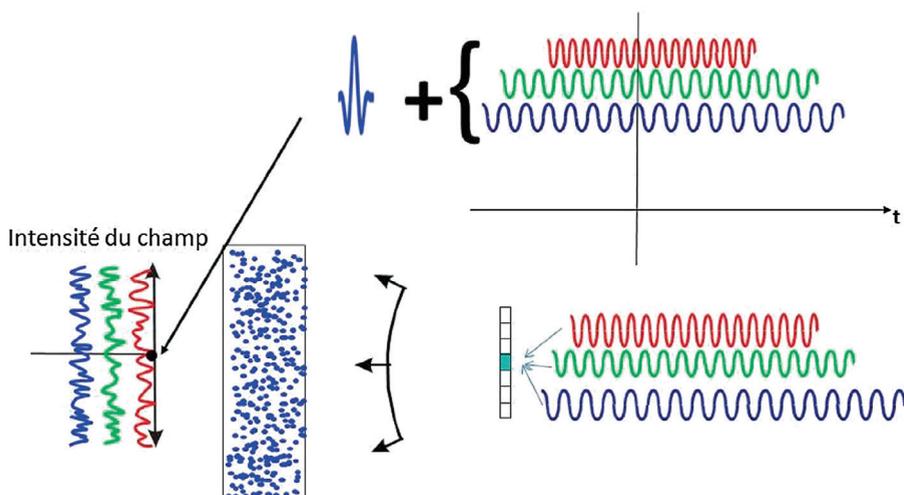


FIGURE 8.5. On a représenté ici l'émission d'un seul transducteur du MRT (en bleu) lorsqu'il fonctionne avec des signaux à large spectre qu'on a décomposés uniquement sur trois composantes spectrales, oscillant à des fréquences différentes représentées par des couleurs différentes. Si le transducteur n'émettait qu'une fréquence (une sinusoïde), l'intensité du champ observé dans le plan de la source serait aléatoire (une figure de speckle d'une couleur correspondant à une fréquence). Il n'y aurait aucune focalisation. On a représenté ici trois fréquences individuelles émises bien décorréelées, c'est-à-dire donnant des figures de speckle décorréelées. En revanche, quand le transducteur du MRT envoie exactement le retourné temporel du signal représenté sur la figure 8.3b, il ajuste en fait parfaitement les trois phases de chacune des sinusoïdes de couleurs différentes, pour qu'après la propagation dans le milieu diffuseur, elles arrivent toutes les trois en phase au temps focal au point source (ce qu'on peut observer en haut à droite de la figure). Il en résulte alors une décompression temporelle du signal, qui sera d'autant meilleure qu'il y aura de nombreuses fréquences décorréelées dans le spectre. L'addition des sinusoïdes de différentes fréquences est constructive uniquement au point source et au temps focal.

qui se propage à travers un milieu diffusant, le signal capté par le MRT s'étale sur une durée beaucoup plus longue tout en conservant un spectre large. Dans l'opération de retournement temporel, les phases de chacune des composantes spectrales sont conjuguées et bien qu'aucune des composantes spectrales ne refocalise à la source, elles arrivent toutes parfaitement en phase à la source (figure 8.5). On retrouve alors au point source l'addition en phase de toutes les sinusoïdes de fréquences différentes et d'amplitudes quelconques, ce qui donne un signal très intense à un temps particulier : le temps focal (figure 8.5). En revanche, pour les points situés en dehors de la source à une distance de l'ordre d'un grain de speckle, les composantes spectrales arrivent avec des phases différentes. On peut évaluer la qualité de la focalisation en étudiant le nombre de bandes de fréquence décorréelées qui couvrent le spectre total des signaux. À deux bandes de fréquence successives « décorréelées » du spectre (que l'on a

représentées par des couleurs différentes sur la figure 8.5) correspondent en effet deux figures de speckle décorréliées. Dans l'expérience décrite, le nombre de ces bandes de fréquence est de l'ordre de 300. C'est le rapport entre la durée du signal d'émission ($1 \mu\text{s}$) et la durée du signal dispersé ($300 \mu\text{s}$) et c'est ce qui explique la qualité de la focalisation. Quant à la taille de la tache focale au temps focal, elle est égale à la taille d'un grain de speckle.

Cette expérience nous permet d'entrevoir un nouveau concept très riche : celui des degrés de libertés spatio-temporels d'un champ ondulatoire à large spectre ou plus exactement des degrés de liberté spatio-fréquentiels (une transformée de Fourier fait passer des variables de temps à celles de fréquence). Dans un milieu diffuseur, le champ enregistré en un unique point de l'espace contient beaucoup d'informations fréquentielles : les degrés de liberté fréquentiels. Le code temporel réémis par cet unique point permet de réaliser une focalisation spatio-temporelle au point source. Cette focalisation est très différente de celle que l'on réalise dans un milieu homogène pour des signaux monochromatiques avec une lentille ou un réseau d'antennes. En effet, pour focaliser une onde monochromatique, il faut de nombreuses antennes disposées sur une grande surface angulaire. C'est le concept de diversité spatiale qui donne lieu au point focal à une concentration d'énergie stationnaire. Ce qui est montré en milieu diffuseur pour des signaux à large spectre est très différent : l'onde ne focalise autour du point source que pendant un instant très bref autour du temps focal. Une seule antenne exploitant une large diversité fréquentielle, et donc beaucoup de degrés de liberté fréquentiels, est suffisante pour focaliser. A chaque point de l'espace est associé un code temporel unique. C'est une sorte de cryptographie spatio-temporelle que réalise le milieu diffuseur. Une fois connu l'ensemble des codes correspondant à tous les points d'une zone d'intérêt, on peut soit focaliser sur chacun des points pour par exemple y concentrer des impulsions très intenses, soit y envoyer des informations de façon très discrète. On peut aussi utiliser cette banque de codes pour localiser des sources quelconques ou pour faire l'image d'une source complexe à partir du signal enregistré par une unique antenne.

Cette expérience de recompression spatio-temporelle d'un signal multidifusé illustre bien l'efficacité étonnante du renversement temporel pour les ondes à spectre large. Elle montre aussi la très faible sensibilité aux erreurs commises dans l'expérience. Les signaux enregistrés et retournés dans les mémoires sont échantillonnés par des convertisseurs A/D qui sont soumis aux erreurs de quantification. À partir du moment où la fréquence d'échantillonnage temporelle est suffisante (de l'ordre de 4 fois la fréquence maximale du spectre), on peut se permettre des erreurs de quantification très importantes. Avec un convertisseur A/D fonctionnant sur 8 bits, on fait une erreur d'une fraction $1/256$ sur

l'amplitude des signaux. Avec un convertisseur 1 bit, l'erreur sur l'amplitude des signaux devient énorme (on remplace le signal par +1 quand il est positif et par -1 quand il est négatif). Malgré tout, l'expérience de retournement temporel fonctionne aussi bien en utilisant des convertisseurs 1 bit ou 8 bits. Curieusement, un miroir à retournement temporel fonctionnant sur 1 bit de quantification permet de réaliser une expérience de renversement du temps de très bonne qualité pourvu que le milieu traversé soit très diffusant et que la bande passante des signaux soit assez grande. Pourtant, le milieu formé d'une répartition aléatoire de tiges s'apparente en mécanique statistique à ce qu'on appelle un « gaz de Lorentz », qui est un milieu chaotique. Le comportement dynamique d'une bille qui serait envoyée dans cette structure serait typiquement chaotique. Une petite erreur sur la vitesse de lancement d'une bille dans ce système de tiges entraînerait après un certain nombre de rebonds une modification exponentielle de la trajectoire. Pour renverser temporellement la trajectoire d'une bille en renversant la vitesse de la bille, il faudrait une précision infinie sur la mesure de cette vitesse, ce qui est impossible. Cette sensibilité aux conditions initiales interdit la réalisation d'une expérience de renversement du temps en mécanique, alors que la propagation des ondes dans un tel milieu est infiniment plus robuste du fait de la linéarité de l'équation des ondes.

8.5 Communication sous-marine et objets tactiles

Dans la description précédente, on a constaté que la présence d'un milieu diffusant permettait d'augmenter significativement l'ouverture apparente d'un petit miroir à retournement temporel. Le succès des techniques de retournement temporel vient du fait qu'il existe de nombreuses situations dans la vie courante où l'on observe le même type d'effet. C'est le cas chaque fois qu'une onde est générée au sein d'un milieu entouré de murs réfléchissants (on parle de cavité réfléchissante ou de guide d'onde). On est alors dans une situation analogue à celle qui se produit au sein d'un kaléidoscope, ou lorsque l'on observe un petit objet : cet objet semble se répéter de façon périodique suivant certaines directions. De même, pour un observateur situé au sein d'une cavité réfléchissante ou à l'intérieur des parois d'un guide d'onde, chaque transducteur d'un petit MRT apparaît comme accompagné d'une multitude de transducteurs virtuels créés par les parois réfléchissantes. On a ainsi l'impression d'avoir un grand MRT formé de beaucoup de transducteurs. Une belle expérience conduite en mer par l'équipe de William Kuperman du SCRIPPS Institute of Oceanography montre cet effet spectaculaire. Un miroir acoustique à retournement temporel de 98 m de long, formé de 30 transducteurs fonctionnant autour d'une fréquence

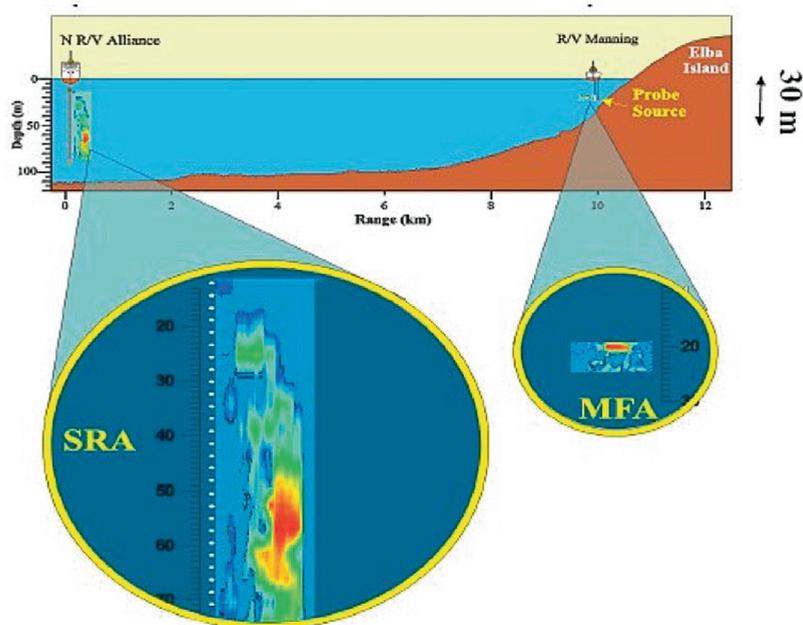


FIGURE 8.6. Figure du haut : photo d'une antenne acoustique à retournement temporel de 90 m de long réalisée par les chercheurs du *Scripps Institution of Oceanography*. Figure du bas : cette antenne a été immergée au large de l'île d'Elbe. La mer se comporte comme un guide d'onde de forme compliquée limitée par le fond et la surface. Pour une longueur d'onde de 50 cm, on obtient à une distance de 10 km une tache focale de 30 cm de largeur à mi-hauteur (c'est le signal observé dans la vignette MFA). C'est pratiquement la limite de diffraction.

centrale de 3 kHz (longueur d'onde de 50 cm) a été immergé au large de l'Île d'Elbe dans une zone de profondeur de l'ordre de 100 m (figure 8.6). C'est un véritable guide d'onde acoustique dans lequel les ondes réverbèrent entre le fond de la mer et la surface de l'eau. Une source sonore immergée à 10 km de distance du MRT émet un son bref qui est capté par les transducteurs du MRT comme une succession d'impulsions sonores étalées dans le temps par la réverbération. Une opération de retournement temporel refocalise cette onde sur sa source sous la forme d'un signal aussi bref que le signal initial et focalisé sur une tache de 30 cm, c'est-à-dire pratiquement sur une demi-longueur d'onde. En modulant le signal envoyé par le MRT, on focalise un message sonore à 10 km de distance qui sera uniquement entendu par un plongeur situé dans la tache focale ! Cette technique a donné lieu à de nombreux développements dans le domaine des communications sous-marines et récemment une équipe japonaise a réalisé des communications sur plus de 1 000 km par retournement temporel.

Peut-on encore maintenir une bonne focalisation en réduisant les 30 transducteurs du MRT à un unique transducteur ? Nous avons vu qu'à travers un milieu diffuseur la solution consistait à travailler avec des signaux à très large bande pour manipuler un grand nombre de degrés de libertés fréquentiels. On peut se demander quelles sont les configurations géométriques qui permettent d'obtenir le plus de degrés de liberté fréquentiels dans une bande passante donnée. On trouve alors que les cavités réverbérantes sont plus intéressantes que les guides d'onde. On sait que lorsqu'on étudie un champ ondulatoire enfermé dans une cavité, ce champ peut toujours s'écrire comme la superposition d'un certain nombre de « modes » dont les fréquences sont parfaitement définies. Pour chaque mode, tous les points du milieu vibrent avec la même fréquence temporelle mais avec des phases et des amplitudes différentes. Certains points vibrent avec une certaine phase, d'autres vibrent en opposition de phase. Certains points situés aux nœuds des modes ne vibrent pas du tout, alors que ceux situés aux ventres de vibration vibrent beaucoup. À chaque mode est associée une figure de vibration et la distance la plus petite entre un nœud et un ventre est égale à la demi-longueur d'onde des ondes acoustiques vibrant à la fréquence du mode. De plus, deux modes de fréquence différente ont des figures d'oscillations indépendantes. On se retrouve alors dans une situation analogue à celle observée en milieu diffuseur avec des figures de tavelures correspondant chacune à une fréquence. Si une source unique, située dans une cavité au point A, émet une impulsion brève à large spectre, elle peut exciter de nombreux modes de cette cavité. À chaque mode correspond une figure stationnaire d'interférence complexe avec des nœuds et des ventres dont les positions dépendent des fréquences modales et de la géométrie de la cavité (figure 8.7, partie haute). Un simple transducteur situé en un point quelconque

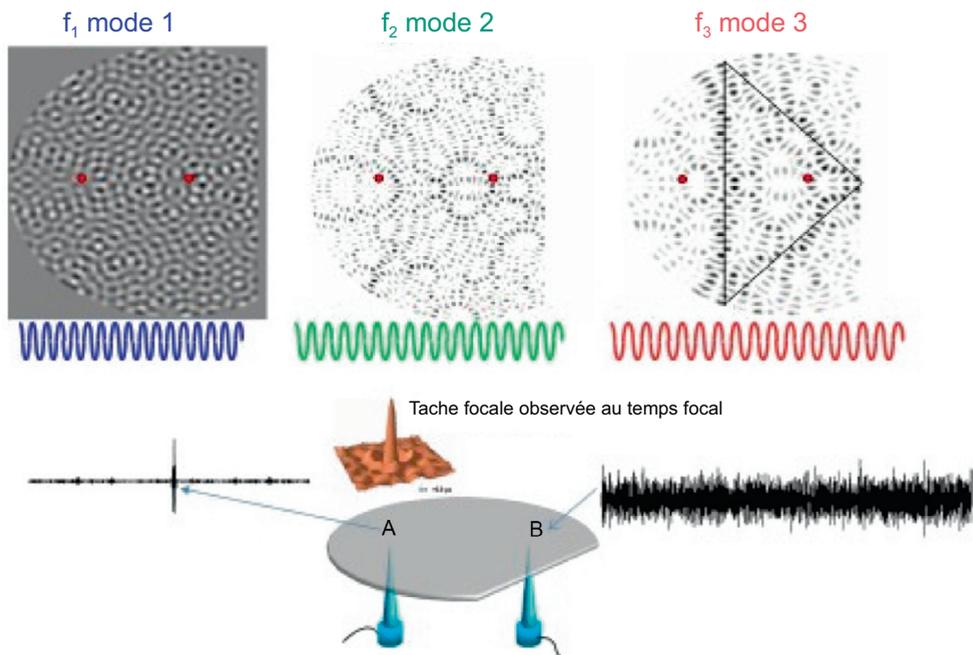


FIGURE 8.7. Sur la partie haute de la figure, on observe trois figures modales générées à partir d'une cavité pour des fréquences modales différentes. Si le transducteur situé au point B émet une de ces trois fréquences (les trois sinusoïdes de couleurs différentes), le champ dans la cavité a des nœuds et des ventres répartis sur les trois figures (la distance entre nœuds et ventres adjacents est d'une demi-longueur d'onde). Sur la partie basse de la figure, le transducteur B émet maintenant un code temporel compliqué et long (le retourné temporel du signal qu'il a reçu quand le transducteur A a émis une impulsion brève). On observe alors au point A un champ ondulatoire qui s'est recompressé temporellement. Sur la vignette au-dessus du point A, on a représenté l'allure du champ ondulatoire au temps focal (celui de la recompression maximale). La tache focale a une dimension latérale de l'ordre d'une demi-longueur d'onde.

B de la cavité peut alors servir de MRT. Du fait de la réverbération dans la cavité, il mesure un signal s'étalant sur un temps très grand qui, après retournement temporel, sera réémis dans la cavité et refocalisera au point A. La refocalisation spatio-temporelle est d'autant plus efficace que le nombre de modes excités par la source est grand. Tous les modes réémis par le point B se retrouvent en phase au point A où l'on observera au temps focal un signal très intense. Quant à la dimension de la tache focale, elle est de l'ordre de la plus petite distance qui sépare un nœud d'un ventre des différents modes, c'est-à-dire la demi-longueur d'onde des fréquences les plus élevées du spectre.

Cette expérience fonctionne avec tous les types de cavité, pourvu que la dissipation des ondes soit assez faible, mais elle est plus efficace pour des cavités dont la géométrie n'est pas régulière. En effet, on évite ainsi les effets

de « dégénérescence » du spectre qui existe dans une cavité à géométrie régulière et qui associe à une fréquence modale plusieurs figures d'interférence possibles. Cette dégénérescence est levée dans une cavité sans symétrie pour laquelle à une fréquence ne correspond qu'une figure d'interférence. Une expérience effectuée avec des ondes élastiques dans un billard (cavité à deux dimensions) de silicium ayant la forme de la lettre D illustre ce principe. Une impulsion mécanique ultrasonore de $1 \mu\text{s}$ de durée est appliquée sur la surface de la plaque de silicium qui génère des ondes élastiques guidées dans la plaquette (ondes de Lamb). On enregistre avec un petit transducteur situé au point B le champ réverbéré qui dure pratiquement 30 ms. On réémet ce signal après retournement temporel et on observe que l'onde élastique ainsi rayonnée refocalise au point source sous la forme d'une impulsion temporelle de $1 \mu\text{s}$ de durée. La magie opère toujours et la tache focale a bien une dimension de l'ordre de la demi-longueur d'onde.

Une application spectaculaire de cette expérience concerne la domotique. Comment transformer un objet solide quelconque (une table, une porte, un vase, un mur. . .) en un clavier intelligent ? Il suffit de penser à l'impact d'un doigt sur la surface d'un objet solide quelconque. Cet impact génère une onde élastique qui est réverbérée à l'intérieur de l'objet. L'enregistrement de cette signature sonore, sur un temps long, par un unique microphone (un accéléromètre) contient toute l'information nécessaire pour localiser la source de l'impact. Comme on l'a vu précédemment, un retournement temporel refocaliserait ce signal sur une tache focale d'une demi-longueur d'onde. On peut alors diviser la surface d'un objet en autant de zones distinctes de surface élémentaire $(\lambda/2)^2$. Pour réaliser un clavier tactile, il suffit, dans un premier temps, de stocker en mémoire toutes les signatures temporelles de chacun des points d'impact. Puis, dans un deuxième temps, lorsque l'on tape à un endroit de l'objet, on compare la signature de cet impact à la banque de données (figure 8.8). Une opération d'intercorrélation du nouveau signal avec l'ensemble des signatures mémorisées initialement revient mathématiquement à réaliser dans l'ordinateur l'opération de retournement temporel. Les longueurs d'onde des ondes élastiques qui sont générées dans la plupart des objets sont de l'ordre du cm, ce qui permet d'obtenir des claviers virtuels très pratiques. Une start-up, *Sensitive Object*, a été créée par Ros-Kiri Ing pour développer cette technologie et a connu un grand succès, puis a été rachetée par une grande société américaine TYCO-Electronics avant de devenir la société Elo-Touch. Elle développe différentes applications de ce principe.

Dans le même ordre d'idées, les chercheurs de l'US Army et de l'Institut franco-allemand de St Louis ont développé des méthodes de localisation par retournement temporel de tireurs embusqués (*snipers*) dans les villes. La réverbération rend en effet très difficile le repérage de la source du tir.

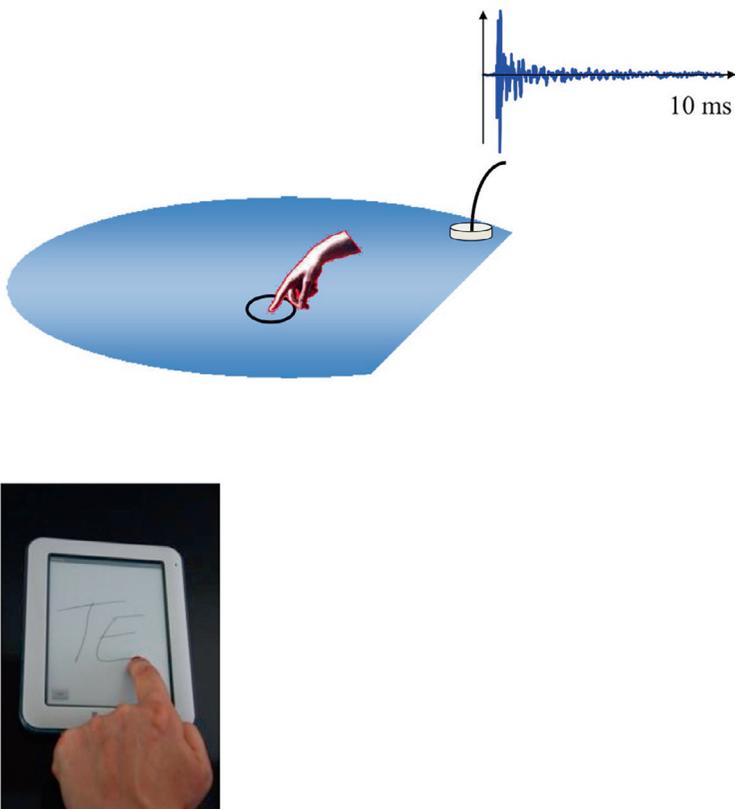


FIGURE 8.8. Figure du haut : l'impact d'un doigt génère un signal bref qui est capté sous la forme d'un code temporel par un accéléromètre placé sur une table. Ce code contient assez d'information pour permettre de retrouver la position de l'impact. Figure du bas : le bruit que fait un doigt en se déplaçant sur une surface produit une onde élastique qui se réverbère dans l'objet et qui est captée par l'accéléromètre. Ce signal est comparé à une banque de signaux préalablement mémorisés. On reconstruit ainsi en temps réel la trajectoire du doigt pixel par pixel si la tablette a été préalablement étalonnée.

En utilisant un modèle de propagation du son dans une ville dont on connaît la géométrie des bâtiments, ils ont montré qu'on pouvait retrouver la position du tireur avec un faible nombre de microphones.

8.6 Super-résolution et résonateurs sub-longueur d'onde

Nous venons de voir comment la présence d'un milieu diffuseur ou de parois réfléchissantes permettait à une petite antenne de se transformer en une étonnante lentille spatio-temporelle, capable de focaliser pendant un temps bref des

signaux sur une dimension de l'ordre de la longueur d'onde. Est-il possible d'utiliser le même principe pour focaliser sur des dimensions beaucoup plus petites ? C'est le problème de la super-résolution qui intéresse beaucoup de communautés et dont les applications potentielles sont nombreuses. La réponse est positive, mais pour cela il faut que la source sur laquelle on veut focaliser ne soit plus située dans un milieu homogène comme dans les deux expériences précédentes, mais au contraire enfouie ou située très près d'un milieu fortement hétérogène. La présence de ce milieu autour de la source crée un spectre de modes très différents de ceux que nous avons observés précédemment et qui peuvent avoir des échelles d'oscillation spatiales bien plus fines que les longueurs d'ondes utilisées. Pour obtenir un tel résultat, il faut un milieu formé d'une collection de petits résonateurs fortement couplés entre eux. Ces résonateurs doivent avoir des dimensions beaucoup plus petites que la longueur d'onde des ondes qui les mettent en résonance. On parle de *résonateurs sub-longueur d'onde* qui, du fait de leur petite dimension, peuvent être disposés en nombre dans une zone de dimension petite devant la longueur d'onde. Ils peuvent être répartis soit de façon désordonnée, soit de façon périodique. On réalise alors une structure (une sorte de méta-matériau) qui ressemble soit à une grosse molécule formée d'atomes identiques couplés entre eux, soit à un petit morceau de cristal. On sait en effet qu'un atome individuel se comporte comme un résonateur excité par des photons de certaines fréquences. La longueur d'onde de ces photons est bien supérieure à la taille de l'atome. Lorsque l'on étudie une molécule formée de plusieurs atomes couplés, on observe un spectre de modes bien plus riche que celui d'un atome particulier et à chaque fréquence modale correspond une structure spatiale particulière d'oscillations des atomes. C'est l'analogue de ce que l'on observe en mécanique en couplant par exemple de nombreux oscillateurs mécaniques (pendules) de même fréquence individuelle : l'ensemble de ces résonateurs possède des modes collectifs de fréquences très variées et aux fréquences les plus élevées correspondent des oscillations entre proches voisins en opposition de phase ; cela implique que l'échelle spatiale des oscillations de ces modes puisse être très fine, aussi petite que la distance entre deux atomes (ou résonateurs) voisins. Ces échelles étant beaucoup plus petites que la longueur d'onde de la lumière ou des ondes mécaniques qui excitent ces modes, on comprend alors qu'une antenne à large bande (à retournement temporel) capable d'exciter une collection de ces modes sub-longueur d'onde puisse créer des concentrations de lumière ou d'énergie mécanique sur des taches focales sub-longueur d'onde. Plus le nombre de résonateurs couplés est grand, plus l'effet sera spectaculaire. C'est dans le domaine des micro-ondes et de l'acoustique que les chercheurs de l'Institut Langevin ont réalisé les premières

démonstrations spectaculaires de ce principe. Cette technique a été étendue récemment en optique par une équipe coréenne.

Dans une première expérience, on s'intéresse à la focalisation d'ondes électromagnétiques du type micro-ondes utilisées en télécommunications dans la bande WiFi. La fréquence centrale d'utilisation est ici 2,4 GHz correspondant à une longueur d'onde de 12 cm. La source de rayonnement utilisée est un petit fil de cuivre que l'on insère au sein d'une structure faite de centaines de petits fils de cuivre, parallèles entre eux, répartis latéralement de façon aléatoire et très proches les uns des autres (mm). Ces fils ont tous la même longueur choisie pour résonner individuellement à 2,4 GHz (figure 8.9). En se couplant entre

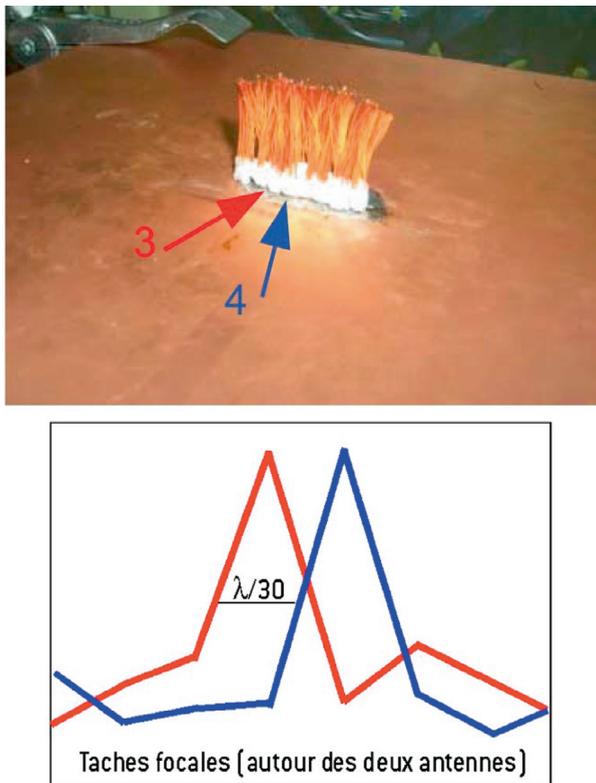


FIGURE 8.9. Un ensemble de petit fils de cuivre répartis de façon aléatoire (résonateurs électromagnétiques) permet de coder les signaux émis par deux antennes (3 et 4) sur la base de tous les modes électromagnétiques de cette structure résonante. Ces modes oscillent sur des distances sub-longueur d'onde et ils sont régénérés par retournement temporel des signaux reçus par une antenne localisée en champ lointain. Sur la figure du bas, on observe les taches focales de taille millimétrique $\lambda/30$ autour de chaque antenne source.

eux, ils créent un spectre de modes collectifs très riche qui contient de nombreuses fréquences s'étalant de quelques centaines de MHz à 2,4 GHz. Après émission par la source au sein de cette structure d'une impulsion large bande de durée 10 ns, une antenne située en champ lointain reçoit alors un signal très long qui résulte de la superposition de tous les modes excités. Un retournement temporel de ce signal électromagnétique suffit à refocaliser à la source, mais cette fois la tache focale mesurée a une dimension de $\lambda/30$. Elle résulte de la superposition cohérente de tous les modes sub-longueur d'onde régénérés par retournement temporel qui oscillent sur des dimensions de l'ordre de la distance entre deux fils, de l'ordre du mm. Des taches focales de dimension millimétrique dans le domaine des télécommunications sont intéressantes car en disposant non pas uniquement d'une seule antenne au sein du milieu de résonateurs, mais de plusieurs d'entre elles séparées de quelques millimètres, on peut apprendre à partir d'une ou de plusieurs antennes situées en champ lointain à envoyer des messages différents focalisés en parallèle sur chacune des antennes du milieu. On peut ainsi exploiter la complexité du milieu pour augmenter la quantité d'information reçue dans un volume donné ! À partir de ces recherches, la société *Time-Reversal Communications* a été créée pour exploiter ces idées dans le domaine des télécommunications discrètes et sécurisées à très haut débit.

Ce même principe de focalisation sub-longueur d'onde s'applique aussi en acoustique et il a été illustré en fabriquant un matériau tout à fait étonnant. Il s'agit de regrouper sous la forme d'un petit cristal 49 canettes de soda (figure 8.10). On sait qu'en acoustique on peut fabriquer des résonateurs sub-longueur d'onde appelés *résonateurs de Helmholtz*. Une bouteille en verre à petit goulot résonne pour une longueur d'onde acoustique nettement supérieure à la taille de la bouteille. On aurait pu imaginer utiliser une caisse de telles bouteilles pour faire un matériau acoustique intéressant, mais c'est oublier la forte dissipation des ondes acoustiques dans les bouteilles. Une solution bien plus intéressante consiste à utiliser des canettes cylindriques trouées du type canettes de soda. Ce sont des résonateurs dont le diamètre est de 6,6 cm et qui résonnent individuellement à 420 Hz, ce qui correspond à une longueur d'onde dans l'air de 80 cm. En fabriquant une structure cristalline de 7×7 canettes avec une faible séparation entre canettes, on réalise un milieu qui peut vibrer sous la forme de 49 modes d'oscillations différents dont les fréquences s'étalent de 340 Hz à 420 Hz. En utilisant une émission acoustique brève de 10 ms centrée sur 400 Hz et localisée près d'une canette, on observe en champ lointain une onde sonore résultant de la superposition de tous les modes. La refocalisation par retournement temporel est, comme en micro-ondes, très efficace, et on refocalise l'onde sonore sur une dimension de l'ordre de $\lambda/8$. En fait, les modes de fréquence spatiale

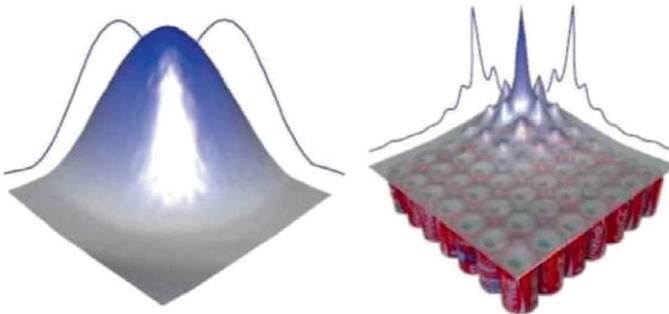
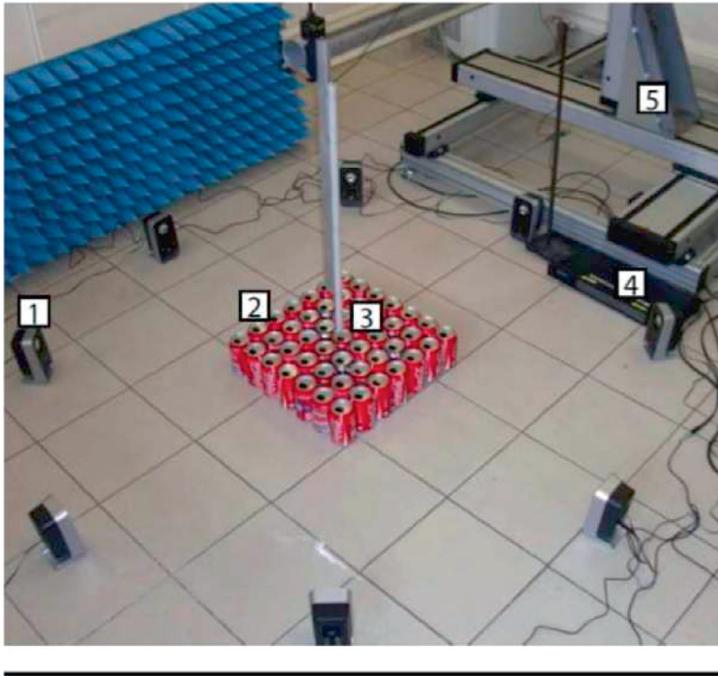


FIGURE 8.10. Expérience de retournement temporel sonore en présence d'un cristal de 49 canettes de soda. L'ensemble des modes du cristal est excité par retournement temporel à partir d'un miroir fait de 8 microphones et de haut-parleurs. Les modes du cristal sont étalés sur des fréquences qui vont de 340 à 420 Hz alors que la fréquence de résonance d'une canette individuelle est de 420 Hz. En bas, on compare la tache focale en présence des canettes (en bas à droite) dont la dimension est de l'ordre de $\lambda/25$, à une focalisation sans les canettes qui donne une tache de l'ordre de $\lambda/2$, à savoir 40 cm.

élevée sont de plus assez fortement atténués par la dissipation et l'opération de retournement temporel n'est alors plus optimale. On peut renforcer le poids de

ces modes avant de les réémettre par une méthode de retournement temporel itératif développée à l'Institut Langevin. On obtient alors une tache focale de dimension $\lambda/25$.

8.7 Imagerie ultrarapide et retournement temporel

Si la nature a sélectionné différents types d'ondes pour aider les êtres vivants à se faire une image de leur environnement, c'est fondamentalement parce que l'équation des ondes est réversible et qu'on peut ainsi reconstruire l'image d'un objet à partir du champ rayonné par cet objet. Les objets dégagent aussi des odeurs ou de la chaleur, mais reconstruire l'image d'un objet à partir de ces informations est très difficile car l'équation que vérifient ces quantités (l'équation de la diffusion) n'est pas réversible. Les fondements de la théorie de l'imagerie reposent sur la réversibilité des ondes. Pour passer du concept de miroir à retournement temporel à celui d'un système d'imagerie, il y a cependant une difficulté majeure. Un MRT fabrique physiquement une image réelle des objets sur eux-mêmes alors qu'un système d'imagerie doit faire une image de l'objet dans un deuxième milieu : l'ordinateur, son écran, la rétine de l'œil. . . Si le milieu initial est homogène (célérité uniforme, sans obstacle), il existe un dispositif très pratique : la lentille, qui est capable de transformer toute onde divergente provenant des points d'un objet en onde convergente (suivant une loi stigmatique). Mais ici les ondes convergentes focalisent, non pas dans le milieu initial, mais dans un milieu adjacent dans lequel est disposée une rétine qui visualise l'image. On peut fabriquer des lentilles pour tout type d'ondes (optique, micro-onde, et ultrasonore) mais elles présentent de nombreuses aberrations. Pour les ondes ultrasonores et les micro-ondes, on dispose d'une autre alternative. On peut fabriquer un *dispositif à retournement temporel*, formé d'un réseau de transducteurs, qui après avoir enregistré les signaux provenant de l'objet, les retournent temporellement et les propagent dans un deuxième milieu d'observation. Ce concept présente beaucoup d'analogies avec l'holographie mais ici, contrairement à l'optique, on dispose de transducteurs linéaires et instantanés qui peuvent enregistrer toute la complexité temporelle de signaux à large spectre.

On peut appliquer ce concept à l'imagerie des tissus mous du corps humain. On sait que pour les ultrasons, les tissus se comportent comme de l'eau chargée de petits diffuseurs. Dans de tels milieux les ultrasons se propagent sous forme d'ondes de compression dont la célérité est pratiquement uniforme (autour de 1520 m.s^{-1}). Connaissant cette célérité, on peut effectuer directement dans un ordinateur simulant la propagation des ondes les phases de réémission et de

propagation des signaux renversés. C'est ainsi que les chercheurs de l'Institut Langevin ont réalisé, il y a une quinzaine d'années, le premier échographe ultrarapide capable de délivrer 10 000 images par seconde du corps humain. Un réseau de plusieurs centaines de transducteurs piézoélectriques « insonifie » le milieu par un front d'onde ultra-sonore très large. En tenant compte du temps d'aller-retour des ultrasons, on peut répéter les tirs à une cadence proche de 10 000 Hz. À chaque tir, les échos d'une large zone de tissus sont captés par les transducteurs et stockés dans une très grande mémoire électronique, puis une *opération de retournement temporel* et de repropagation est effectuée par des processeurs ultra-rapides. Ce mode d'imagerie ultra-rapide est très différent de celui des échographes classiques, où le milieu est éclairé par un faisceau très fin déplacé électroniquement le long de la zone à imager, qui donnent des cadences inférieures à 100 images par seconde.

Est-il vraiment utile d'observer l'intérieur du corps humain à des cadences de plusieurs milliers d'images par seconde, alors que la plupart des cycles du corps humain ont des périodes bien plus lentes, de l'ordre de la seconde pour le cycle cardiaque et de quelques secondes pour le cycle respiratoire ? La réponse est positive car nous avons pu filmer pour la première fois, à l'intérieur du corps, la propagation d'ondes élastiques lentes : les ondes de cisaillement qui mettent en mouvement toutes les petites structures diffusantes des tissus. Ces ondes parcourent les tissus à des vitesses lentes (entre 1 et 10 m.s⁻¹). Elles ne se propagent qu'aux fréquences sonores basses (de 10 Hz à 500 Hz), au delà elles sont trop atténuées par la viscosité et leur célérité dépend directement du module d'élasticité des tissus (module de cisaillement). Elles ont de grandes longueurs d'onde (centimétriques), mais elles ressentent le même module que celui auquel est sensible le médecin pendant l'acte de palpation. À partir du film ultrarapide de la propagation de ces ondes, on peut localement déduire leur célérité qui donne directement la valeur de l'élasticité des tissus avec la résolution millimétrique des ultrasons. C'est le premier dispositif d'*imagerie multi-ondes* où le contraste de l'image finale est obtenu grâce à l'onde de cisaillement, mais où la résolution spatiale est obtenue grâce aux ondes ultrasonores millimétriques.

À cet échographe ultrarapide est associée une méthode de génération d'ondes de cisaillement dans les tissus. Un micro-déplacement (microséisme) est induit à distance dans une petite zone des tissus par la pression de rayonnement ultrasonore associée à un faisceau focalisé de longue durée (1 ms). En déplaçant ce microséisme à une vitesse supersonique (pour les ondes de cisaillement), on induit un « bang » sonique qui génère une onde de cisaillement suffisamment intense pour parcourir l'organe à étudier. La société *SuperSonic Imagine*, créée par les chercheurs de l'Institut Langevin, commercialise cet échographe ultrarapide, l'*Aixplorer*, dont les applications médicales sont très nombreuses,

comme la détection des cancers du sein, de la prostate, du foie et le diagnostic des maladies musculaires, cardiaques et vasculaires, etc.

8.8 Thérapie

Une autre application médicale très prometteuse du retournement temporel est la thérapie du cerveau (nécrose thermique) par ultrasons focalisés. Il s'agit de focaliser des ultrasons de grande intensité pendant plusieurs secondes (régime sinusoïdal) à travers la boîte crânienne en corrigeant les très fortes aberrations de la propagation induites par l'os. Bien qu'il existe des effets de dissipation dans l'os et que la phase finale de focalisation s'accompagne d'effets non linéaires et dissipatifs, la technique du retournement temporel permet de compenser avec une grande précision l'ensemble des effets de la boîte crânienne. Avec un

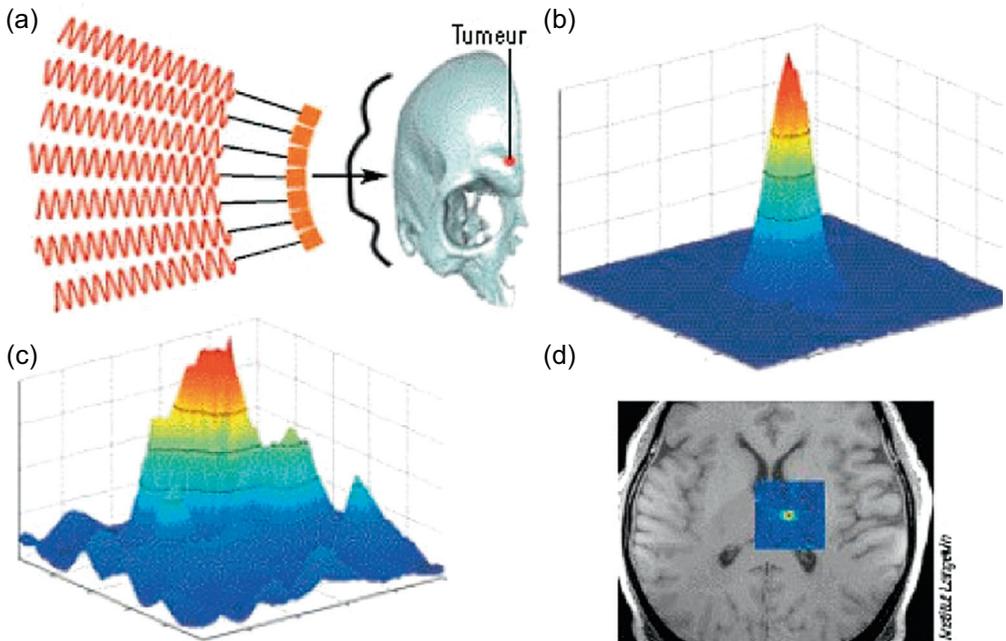


FIGURE 8.11. On peut brûler à distance une zone du cerveau à travers la boîte crânienne en focalisant des ultrasons de grande énergie par retournement temporel à partir d'un réseau de transducteurs piézo-électriques (a). La tache focale obtenue par retournement temporel à travers le crâne (b) est comparée à la tache focale (c) obtenue sans correction. La méthode testée sur l'animal et sur des cadavres montre qu'on peut atteindre localement des températures de 65 °C (au point rouge) qui suffisent pour détruire les tissus du cerveau.

premier prototype de 300 transducteurs, fonctionnant autour de 1 MHz, des puissances acoustiques du kW.cm^{-2} sont atteintes facilement derrière le crâne sur des zones millimétriques et sont suffisantes pour élever la température des tissus à 65 °C (température de coagulation des protéines). Initialement, le procédé nécessitait l'implantation d'un hydrophone dans le cerveau pour servir de source de référence : des nécroses très précises ont été obtenues. Puis pour éviter l'utilisation d'un hydrophone, une nouvelle approche dite de source virtuelle a consisté, après avoir obtenu l'image 3D du crâne par scanner X, à bâtir un modèle informatique de la propagation ultrasonore à l'intérieur du crâne et à simuler l'opération de retournement temporel de signaux provenant de sources virtuelles situées en n'importe quel point du cerveau. Cette approche a été testée avec succès *in vivo* sur l'animal puis sur des cadavres dans lesquels ont été réalisées des nécroses des tissus du cerveau avec une résolution millimétrique (figure 8.11). Un système de 512 transducteurs est actuellement en phase de test pour une évaluation clinique à l'Institut du cerveau et de la moelle (ICM) à Paris.

8.9 Sonar et radar à retournement temporel

Si les expériences décrites précédemment illustrent l'intérêt des miroirs à retournement temporel, l'application à la détection de cibles est aussi très prometteur. C'est le domaine du sonar et du radar. Le problème que l'on cherche à résoudre est celui de la détection ou de l'imagerie de cibles réfléchissantes situées derrière ou dans des milieux de propagation complexes. Les milieux dans lesquels les ultrasons se propagent sont souvent hétérogènes et les faisceaux sont distordus au cours de la propagation. En contrôle non destructif par ultrasons des objets solides (en aéronautique et dans l'industrie nucléaire), les pièces à contrôler sont généralement immergées. La forme de l'interface entre la pièce solide et le liquide de couplage complique la propagation ultrasonore et limite actuellement le niveau de détection des petits défauts. En acoustique sous-marine, les phénomènes de convection, les variations de la vitesse du son en fonction de la profondeur et la réverbération acoustique entre le fond et la surface de la mer sont aussi des sources très importantes de distorsion pour les sonars.

Comment, dans les cas énumérés ci-dessus, exploiter le principe du retournement temporel pour détecter et focaliser sur une ou plusieurs cibles réfléchissantes ? Il suffit pour cela d'utiliser le réseau de transducteurs formant le miroir à retournement temporel suivant la séquence suivante. Dans un premier temps, on illumine le milieu à explorer au moyen d'une impulsion ultrasonore brève. Dans le cas où le milieu contient une unique cible réfléchissante, on sélectionne l'écho réfléchi par cette cible au moyen d'une fenêtre temporelle et on renverse

temporellement cet écho avant de le réémettre. La cible réfléchissante, une fois illuminée, se comporte comme une source acoustique secondaire sur laquelle le front d'onde réémis focalise de façon optimale, même à travers un milieu de propagation hétérogène.

En présence de plusieurs cibles, le problème se complique. Si l'on éclaire au moyen d'une impulsion brève un milieu contenant plusieurs cibles de réflectivités différentes, l'opération de retournement des échos provenant de ces cibles génère plusieurs fronts d'onde d'amplitudes différentes qui convergent vers chacune des cibles. Si les réflecteurs sont séparés les uns des autres d'une distance supérieure à la limite de résolution du miroir, celui-ci crée les images réelles acoustiques des cibles sur elles-mêmes sans interférences parasites. Le front d'onde le plus intense éclaire la cible de plus forte réflectivité et les fronts d'ondes d'amplitudes plus basses éclairent les autres cibles. On peut alors itérer et après plusieurs itérations le processus converge vers la cible la plus brillante. Une fois le faisceau orienté sur cette cible, l'itération ne modifie plus le résultat. On obtient un premier *invariant de l'opération de retournement temporel*. Dans un deuxième temps, on peut soustraire ce front d'onde de l'ensemble des échos pour ne plus éclairer cette cible et une nouvelle itération permet de sélectionner la seconde cible par ordre de brillance. On trouve alors un second invariant de l'opération de retournement temporel ... Cette approche permet de détecter très rapidement l'ensemble des cibles d'un milieu et de trouver tous les fronts d'ondes optimaux focalisant sur chacune des cibles. Le gros intérêt de cette méthode est de permettre la correction optimale de toutes les aberrations. Une autre approche très puissante a été développée à l'Institut Langevin par C. Prada. Plutôt que d'effectuer physiquement toutes ces itérations dans le milieu, on exploite la linéarité des opérations décrites ci-dessus pour effectuer toutes ces opérations en logiciel. On mesure dans un premier temps l'ensemble des réponses échographiques croisées entre les éléments d'un réseau de transducteurs, puis on fabrique une matrice de retournement temporel, dont les vecteurs propres sont les invariants recherchés et dont les valeurs propres donnent la réflectivité des cibles. C'est la méthode DORT qui signifie : *diagonalisation de l'opérateur de retournement temporel*.

À partir de ces principes, on peut développer des méthodes de détection de cibles et dans certains cas de leur destruction. C'est ce qui a été fait dans le domaine de la lithotritie (destruction de calculs rénaux) et du contrôle non destructif. Des systèmes de focalisation-poursuite et de destruction de calculs rénaux très précis ont été fabriqués qui, à chaque instant, focalisent une onde de choc de grande amplitude (1 000 bars) sur les calculs qui bougent avec la respiration du patient. De la même façon, ces solutions sont très intéressantes en imagerie sous-marine des petits fonds où les cibles (mines, plongeurs) sont

pratiquement indétectables à cause de la réverbération. Des démonstrateurs ont été fabriqués avec le financement de la DGA (Direction générale de l'armement) qui permettent de détecter toutes les cibles présentes dans des hauts fonds jusqu'à plusieurs centaines de mètres de distance.

8.10 Conclusion

La conclusion de ce chapitre est que le concept de retournement temporel des ondes est étonnamment robuste et qu'il s'adapte remarquablement bien aux milieux les plus complexes. Mieux encore, il profite de cette complexité pour apporter des solutions originales en télécommunications, en imagerie, en thérapie, en localisation de sources. C'est un concept tellement robuste qu'on peut même se demander si la nature ne l'exploite pas aussi. . .

Bibliographie

Audoin, Claude et Guinot, Bernard. *The measurement of time : time frequency and the atomic clock*, Cambridge University Press, Cambridge 2001.

Bretenaker, Fabien et Treps, Nicolas. *Le laser*, « Introduction à », EDP Sciences, 2010.

Fink, Mathias. *Time reversed acoustics*, *Physics Today*, **20**, 34-40, 1997.

Fink, Mathias. Les miroirs à retournement temporel, *Pour la Science*, 66-73, 2000.

Galison, Peter *L'empire du temps : les horloges d'Einstein et les cartes de Poincaré*, folios essais Paris, 2005.

Jespersen, James et Fitz-Randolph, Jane. *From sundials to atomic clocks* (deuxième édition révisée), Dover Publications, Mineola, New York, 2014.

Klein, Étienne. *Le facteur temps sonne toujours deux fois*, Flammarion, Paris, 2007.

Lachièze-Rey, Marc. *Voyager dans le temps : la physique moderne et la temporalité*, Éditions du Seuil, Paris, 2013.

Le Bellac, Michel. *Les relativités : espace, temps, gravitation*, « Introduction à » EDP Sciences, Paris, 2015.

Mosk, A.P., Lagendijk, G., Leroose, G., Fink, M. *Controlling waves in space and time for imaging and focusing in complex media*, *Nature Photonics*, **6**, 283-292, 2012.

Penrose, Roger. *L'esprit, l'ordinateur et les lois de la physique*, Dunod, Paris, 1998.

Pour la Science numéro spécial, *Le temps. Est-il une illusion ?*, Belin, Paris, 2010. Voir en particulier Craig Callender, *Le temps est-il une illusion ?*; Roger Balian, *Le paradoxe de l'irréversibilité*; Christophe Salomon, *Mesurer le temps avec les atomes*.

Pour la Science numéro spécial, *L'univers des ondes. Quand la physique dépasse ses limites*, Belin, Paris, 2011. Voir en particulier Mathias Fink, *La complexité : un atout pour le retournement temporel des ondes* ; Sylvain Gigan, *Voir à travers le brouillard*.

Price, Huw. *Time's arrow and Archimedes' point*, Oxford University Press, Oxford, 1996.

Rovelli, Carlo. *Et si le temps n'existait pas ?* Dunod, Paris, 2014

Ruelle, David. *Hasard et chaos*, Poches Odile Jacob, Paris, 1991.

Zeh, Hans-Dieter. *The physical basis of the direction of time*, Springer, Berlin, 2001

Index

A

accélération de la pesanteur 26, 74
ACES : *Atomic Clock Ensemble in Space*
105
adiabatique 123, 145
atomes froids 35, 103
attoseconde 12

C

causalité 8, 61, 150
chaos déterministe 144
cône de lumière 53, 84
cône de lumière futur 53
cône de lumière passé 53
conjugaison particule–antiparticule
135
constante de Boltzmann 128
constante de gravitation 11, 80, 140
constante de Planck 11, 31, 75, 140
critère de Rayleigh 150

D

date 6
décalage vers le rouge gravitationnel
75, 83, 100, 114
démon de Loschmidt 136, 143
diagramme d'espace-temps 7

dispositif à retournement temporel
167
droite de lumière 52
dynamique chaotique 138
dynamique non-linéaire 144

E

écho de spin 141
effet Doppler 33, 37
entropie 118, 122
entropie de Boltzmann 129
entropie d'un trou noir 140
entropie thermodynamique 129
équilibre thermique 122, 125
espace-temps 5, 53
espace-temps de Minkowski 61
espace-temps newtonien 5
éther 88
étoile à neutrons 82
événement 6, 53
évolution temporelle 147
exactitude 43

F

facteur de Lorentz 64
facteur de qualité 28
femtoseconde 12
flèche du temps 4, 118

fontaine atomique 32
franges de Ramsey 36
fréquence 8, 25
fréquence de résonance 29
fuseau horaire 2, 91, 114

G

géoïde 81
GPS : *Global Positioning System* 77, 92,
100

H

hertz (Hz) 8, 25
horloge à atomes froids 35
horloge à césium 8, 33
horloge à fontaine 39
horloge à quartz 29
horloge atomique 7, 32, 46
horloge (atomique) optique 8, 42, 70, 77
horloge PHARAOH 104
horizon (d'un trou noir) 83
hypersurface de simultanéité 59

I

imagerie multi-ondes 168
intervalle de temps propre 66, 86
ion 8, 43
irréversibilité 117
ISS : *International Spatial Station* 61, 103

L

ligne d'univers 7, 53
limite de diffraction 150
longueur d'onde 145

M

machine thermique 120
macroétat 128
masse gravitationnelle 73
masse inertielle 73
méridien de Greenwich 91
méthode double voie 95
méthode simple voie 95
métrique de Minkowski 67
métrique de Schwarzschild 86
milliseconde 15
miroir à retournement temporel
(MRT) 147
mouvement rectiligne uniforme 62

N

nanoseconde 12, 53, 117
neutrino 119
niveau d'énergie 31, 74

O

observateur 55
onde acoustique (ou sonore) 145,
151
onde ultrasonore 148, 168
opération de retournement
temporel 168, 171
oscillateur 24
oscillateur mécanique 26

P

paradoxe des jumeaux 55, 67
parité 134
pendule 26, 81
photosynthèse 132
photon 31

picoseconde 98, 117
piège magnéto-optique 38, 104
potentiel gravitationnel 80
premier principe 119
principe cosmologique 86
principe de Huygens 147
principe d'équivalence 73
principe de relativité 8, 62
processus réversible 123

R

rayon de Schwarzschild 82
rayonnement du corps noir 39, 133
récurrence de Poincaré 137
référentiel 61
référentiel d'inertie 61
référentiel terrestre géocentré 70, 81
refroidissement laser 37, 45
régime linéaire 145
relativité générale 51, 73
relativité restreinte 51
relation de Planck-Einstein 31, 75, 132
rendement 121
renversement du temps 136, 145, 152
réseau optique 41
résonateur de Helmholtz 164
résonateur sub-longueur d'onde 163
ressort 24
retournement temporel 147
réversibilité 117

S

seconde 8, 47
second principe 119
sensibilité aux conditions initiales
139, 157
simultanéité 8, 58
stabilité 43

solution avancée 149
solution retardée 149
speckle (ou tavelure) 154
symétrie 134
synchronisation 63, 95
synchronisation (ou méthode) double
voie 59, 95
synchronisation (ou méthode) simple
voie 63, 95

T

tache focale 150, 154, 160
tavelure 154
température absolue 120, 128
TAI : *Temps Atomique International*
1, 5, 81, 113
temps coordonnée 56
temps cosmologique 87
temps de Planck 11
temps focal 155, 160
temps (absolu) newtonien 1, 5, 55
temps propre 2, 55
temps universel 88, 91
transducteur piézo-électrique 146
transfert du temps 93
transition atomique 31
trou noir 83

U

univers bloc 3, 60
univers causal 150
UTC : *Universal Coordinate Time* 114

V

vitesse de la lumière 52
vitesse limite 52

LE TEMPS : MESURABLE, RÉVERSIBLE, INSAISSABLE ?

Depuis l'Antiquité, la nature du temps a fasciné nombre de grands penseurs. Cet ouvrage expose ce que la physique est capable de dire aujourd'hui sur le sujet.

La mesure du temps, ou plus exactement celle d'une durée, se fait grâce à des horloges atomiques dont l'exactitude peut atteindre une seconde sur plusieurs milliards d'années. Nous décrivons la façon dont s'effectue le transfert du temps qui permet la synchronisation d'horloges en différents points de la Terre ou de l'espace au milliardième de seconde près, ou même mieux. Les relativités, restreinte et générale, ont bouleversé notre conception du temps et ont un impact considérable sur certains problèmes de la vie quotidienne comme l'utilisation du GPS. On abandonne l'idée d'un temps absolu, le temps devient multiple et insaisissable, et peut-être même une illusion.

Enfin la flèche du temps, ou l'irréversibilité, implique que les phénomènes physiques se déroulent toujours dans un sens déterminé, en relation avec la croissance de l'entropie. Cependant il est possible dans certaines conditions d'échapper à cette contrainte et de construire un temps réversible grâce aux miroirs à retournement temporel, dont nous décrivons les nombreuses applications pratiques.

Mathias Fink occupe la chaire George Charpak à l'École Supérieure de Physique et de Chimie Industrielles de la Ville de Paris (ESPCI) Paris-Tech, où il a fondé l'Institut Langevin. Il est membre de l'Académie des sciences et de l'Académie des technologies.

Michel Le Bellac est professeur émérite à l'Université de Nice-Sophia Antipolis. Il est l'auteur de plusieurs manuels de physique à un niveau avancé et de deux livres dans la présente collection, Le monde quantique et Les relativités.

Michèle Leduc est directrice de recherche émérite CNRS au Laboratoire Kastler-Brossel et directrice de l'Institut Francilien de Recherche sur les Atomes Froids (IFRAF).

Isbn : 978-2-7598-1911-9



9 782759 819119

19 €

edp sciences
www.edpsciences.org

Création graphique : Béatrice Couëdel

La collection « **UNE INTRODUCTION À...** » se propose de faire connaître à un large public les avancées les plus récentes de la science. Les ouvrages sont rédigés sous une forme simple et pédagogique par les meilleurs experts français.