

Mécanique avancée et relativité restreinte

Éric Brunet¹, Jules Grucker², Jesper Jacobsen³

De grandes parties de ce cours ont été recopiées d'autres polys écrits entre autres par
Mathieu Bertin, Alice Sinatra et Paul Windey

Équipe enseignante 2024-2025 :

Dalil Brouri⁴ (TD Mono1), Éric Brunet¹ (Cours Mono), Arnaud Cassan⁵ (TD Mono2),
Michel Fioc⁶ (TD Maj2), Jules Grucker² (Cours Maj, 2^e partie),
Jesper Jacobsen³ (Cours Maj, 1^e partie), Maxime Lecomte⁷ (TD Mono3),
François Mallet⁸ (TD Maj3 et CMI), Éric Michel⁹ (TD Maj1).

3 décembre 2024

Ouvrages de références pour la mécanique avancée :

- Herbert Goldstein, *Mécanique classique*.
- Pour aller plus loin : Jean-Marcel Rax, *Mécanique analytique*.
- Et encore plus loin : Lev Landau & Evgeny Lifchitz, *Physique théorique — Mécanique*,
- Des exercices : Murray R. Spiegel, *Théorie et applications de la mécanique générale*.

Ouvrages de références pour la relativité :

- *Berkley Physics Course 1, Mechanics*, chapitres 10–13. (Il existe une version française.)
- *The Feynman lectures on Physics*, chapitres 15–17. (Il existe une version française. La version anglaise est accessible en ligne : <https://www.feynmanlectures.caltech.edu/>)

Les articles originaux d'Einstein (ils se lisent bien !) :

- *Zur Elektrodynamik bewegter Körper*, *Annalen der Physik* **17**, 891-921 (1905). La naissance de relativité restreinte ! Un des articles scientifiques les plus célèbres jamais écrits. On trouve des traductions anglaises et françaises (et dans de nombreuses autres langues).
- *Ist die Trägheit eines Körpers von seinem Energieinhalt abhängig ?* *Annalen der Physik* **18**, 639-641 (1905) où pour la première fois a été écrit $E = mc^2$ même si les notations n'étaient pas exactement celles-ci. On trouve facilement la traduction anglaise.

Livre de “vulgarisation” (même s'il n'y a rien de vulgaire dedans, bien au contraire) d'Einstein :

- *La Relativité* édition Payot. Einstein était un excellent pédagogue et on apprend énormément sur les concepts de la théorie en lisant ce bijou.

1. email : eric.brunet@sorbonne-universite.fr
2. email : jules.grucker@sorbonne-universite.fr
3. email : jesper.jacobsen@sorbonne-universite.fr
4. email : dalil.brouri@sorbonne-universite.fr
5. email : arnaud.cassan@sorbonne-universite.fr
6. email : michel.fioc@sorbonne-universite.fr
7. email : maxime.lecomte@sorbonne-universite.fr
8. email : francois.mallet@sorbonne-universite.fr
9. email : eric.michel@sorbonne-universite.fr

Table des matières

I	Mécanique avancée	1
1	Problème à deux corps	3
1.1	Lois de conservation et référentiel du centre de masse	3
1.2	Force centrale	4
1.2.1	Conservation du moment cinétique et loi des aires	5
1.2.2	Potentiel effectif	5
1.2.3	Étude qualitative des trajectoires	6
1.3	Problème de Kepler	8
1.3.1	Équation de la trajectoire	8
1.3.2	Premières propriétés des trajectoires kepleriennes	9
1.3.3	Trajectoire elliptique	10
1.3.4	Trajectoire hyperbolique	11
1.3.5	Potentiel répulsif en $1/r$ et diffusion de Rutherford	12
2	Équations de Lagrange	15
2.1	Les équations de Lagrange	15
2.1.1	Introduction	15
2.1.2	Notations et définitions	16
2.1.3	Les équations de Lagrange	16
2.2	Exemples	17
2.2.1	Une particule dans un potentiel	17
2.2.2	Le pendule pesant	18
2.2.3	Une perle coulissant sur une tige en rotation	19
2.2.4	Mobile sur un plan incliné	20
2.2.5	Le pendule au point d'attache mobile	22
2.2.6	Le double pendule pesant	23
2.3	Quelques propriétés du lagrangien	26
2.3.1	Dépendance de T et U en les coordonnées et vitesses généralisées	26
2.3.2	Additivité du lagrangien	28
2.3.3	Invariance de jauge	28
2.3.4	Impulsions généralisées et lois de conservation	29
2.3.5	Conservation de l'énergie	30
2.4	Première justification : le principe de moindre action	31
2.5	Deuxième justification : à partir des équations de Newton	33
2.5.1	Équations du mouvement	34
2.5.2	Prise en compte des différents types de force	34
2.5.3	Les forces de contrainte et le déplacement virtuel	35
2.5.4	Quelques exemples	36

3	Mouvement d'un solide	39
3.1	Généralités	39
3.1.1	Définition d'un corps solide	39
3.1.2	Description mathématique d'un solide	39
3.2	Énergie cinétique et moment d'inertie	41
3.2.1	Calcul de l'énergie cinétique et moment d'inertie	41
3.2.2	Exemple : le pendule rigide	42
3.2.3	Exemple : cylindre roulant sans glisser sur un plan incliné	43
3.3	Moment cinétique et axes principaux du solide	44
3.3.1	Calcul direct du moment cinétique	44
3.3.2	Axes principaux	45
3.3.3	Exemple	46
3.4	Un exemple difficile : la toupie	48
3.5	Tenseur d'inertie	49
3.5.1	Calcul général de l'énergie cinétique de rotation	50
3.5.2	Calcul général du moment cinétique	52
3.5.3	Notion de tenseur d'inertie	52
3.5.4	Axes principaux	53
II	Relativité restreinte	55
4	Postulats et transformations de Lorentz	57
4.1	Relativité galiléenne — rappels	57
4.1.1	Mécanique newtonienne et transformations de Galilée	57
4.1.2	Signification du postulat de relativité	59
4.2	La propagation de la lumière en mécanique classique	60
4.2.1	Nature ondulatoire de la lumière	60
4.2.2	Hypothèse de l'éther	61
4.2.3	L'expérience de Michelson-Morley	61
4.3	Postulats de la relativité restreinte	64
4.3.1	Postulats d'Einstein	64
4.3.2	Transformations de Lorentz	65
5	Conséquences des transformations de Lorentz	71
5.1	Relativité du temps et de la simultanéité	71
5.1.1	Définition de la simultanéité	71
5.1.2	Relativité de la simultanéité et donc du temps	72
5.2	Dilatation des durées et contraction des longueurs	74
5.2.1	Première approche	74
5.2.2	Plus de détail sur la dilatation des durées	75
5.2.3	Plus de détail sur la contraction des longueurs	77
5.3	Intervalle d'espace-temps Δs	78
5.3.1	Invariant relativiste Δs^2	78
5.3.2	Signe de Δs^2 et causalité	79
5.3.3	Cônes de lumière	80
6	Transformation des vitesses et effet Doppler	83
6.1	Composition des vitesses	83
6.2	Aberrations — Transformations des angles	84
6.2.1	Formule générale	85
6.2.2	Aberrations optiques	86

6.3	Effet Doppler	86
6.3.1	Effet Doppler sonore	87
6.3.2	Effet Doppler optique	88
6.3.3	Une application de l'effet Doppler optique : le <i>redshift</i> en astrophysique	89
7	Énergie et quantité de mouvement relativistes	93
7.1	De nouvelles expressions pour E et $\vec{p}$	93
7.1.1	Quantité de mouvement relativiste	94
7.1.2	Énergie cinétique relativiste	96
7.1.3	Énergie interne et énergie totale	96
7.1.4	Relation entre énergie et impulsion	98
7.1.5	Énergie et impulsion du photon.	98
7.2	Mécanique lagrangienne relativiste	99
7.3	Équivalence masse-énergie	101
7.3.1	Ordres de grandeur	101
7.4	Collisions et loi de conservations	102
7.4.1	Un exemple de réaction nucléaire : fission de l'uranium-235	102
7.4.2	Création de paire	103
	Appendice	105
A	Les lois de la mécanique newtonienne	107
A.1	Lois de la mécanique pour un point matériel	107
A.2	Ensemble de points matériels	108
A.2.1	Impulsion du système et centre de masse	109
A.2.2	Principe fondamental de la dynamique	109
A.2.3	Moment cinétique	110
A.2.4	Point d'application du poids ou d'une autre force extérieure	110
A.2.5	Théorème de l'énergie cinétique	111
A.2.6	Théorème de l'énergie mécanique	112
A.2.7	Référentiel du centre de masse et théorèmes de König	112

Première partie

Mécanique avancée

Chapitre 1

Problème à deux corps

Remarque : avant d'aborder ce chapitre, il est *fortement* conseillé de lire (et comprendre) l'appendice A, au moins jusqu'à la section A.2.2.

1.1 Lois de conservation et référentiel du centre de masse

Considérons deux corps, deux points matériels M_1 et M_2 , de masses respectives m_1 et m_2 , en interaction mutuelle. On pose $\vec{r}_1 = \overrightarrow{OM_1}$ et $\vec{r}_2 = \overrightarrow{OM_2}$ les positions de ces points dans le référentiel du laboratoire, supposé galiléen. Nous appellerons $\vec{F}_{12}$ la force que le corps 1 exerce sur le corps 2 et inversement $\vec{F}_{21}$ la force que le corps 2 exerce sur le corps 1. D'après la troisième loi de Newton, $\vec{F}_{12}$ et $\vec{F}_{21}$ ont le même module, la même direction et des sens opposés. De plus, la direction des forces d'interaction se confond avec la droite qui passe par les deux points :

$$\vec{F}_{12} = -\vec{F}_{21} \quad \text{et} \quad \vec{F}_{12} \wedge (\vec{r}_1 - \vec{r}_2) = \vec{0}. \quad (1.1)$$

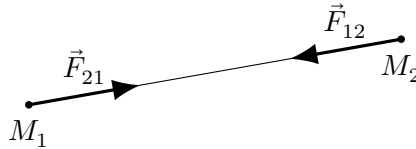


FIGURE 1.1 – Deux corps interagissent avec une force mutuelle $\vec{F} = \vec{F}_{12} = -\vec{F}_{21}$.

Le centre de masse C du système se situe sur le segment $[M_1 M_2]$, plus près du corps le plus lourd. Son vecteur position est

$$\vec{r}_c = \frac{m_1 \vec{r}_1 + m_2 \vec{r}_2}{m_1 + m_2}. \quad (1.2)$$

Nous supposons le système isolé, c'est à dire il n'y a pas de force extérieure agissant sur les deux corps. D'après les lois de Newton, le centre de masse de ce système isolé a une vitesse rectiligne uniforme dans le référentiel du laboratoire (supposé galiléen) :

$$\ddot{\vec{r}}_c = \vec{0} \quad \text{ce qui est équivalent à} \quad \dot{\vec{r}}_c = \overrightarrow{\text{cste}}. \quad (1.3)$$

Ceci implique que le *référentiel du centre de masse* $\mathcal{R}_c$, défini comme le référentiel d'origine C en translation par rapport au référentiel du laboratoire, est un référentiel galiléen.

On pose $\vec{r}_{1,c} = \overrightarrow{CM_1}$ et $\vec{r}_{2,c} = \overrightarrow{CM_2}$ les vecteurs positions des deux points matériels dans le référentiel $\mathcal{R}_c$. Ces positions s'expriment simplement en terme de la position relative $\vec{r}$

$$\vec{r} := \overrightarrow{M_1 M_2} = \vec{r}_2 - \vec{r}_1 = \vec{r}_{2,c} - \vec{r}_{1,c} \quad (1.4)$$

On trouve, après un rapide calcul,

$$\vec{r}_{1,c} = \vec{r}_1 - \vec{r}_c = -\frac{m_2}{m_1 + m_2} \vec{r} \quad \vec{r}_{2,c} = \vec{r}_2 - \vec{r}_c = \frac{m_1}{m_1 + m_2} \vec{r}. \quad (1.5)$$

La deuxième loi de Newton dans $\mathcal{R}_c$ appliquée à chacun des deux corps

$$m_1 \ddot{\vec{r}}_{1,c} = \vec{F}_{21} \quad m_2 \ddot{\vec{r}}_{2,c} = \vec{F}_{12} \quad (1.6)$$

donne alors deux fois la même équation

$$\frac{m_1 m_2}{m_1 + m_2} \ddot{\vec{r}} = \vec{F}_{12} \quad (1.7)$$

Nous allons introduire m la *masse réduite*

$$m \equiv \frac{m_1 m_2}{m_1 + m_2} \quad (1.8)$$

et la force $\vec{F} \equiv \vec{F}_{12}$. À cause de l'invariance par translation du système, $\vec{F}$ dépend seulement de la position relative $\vec{r}_2 - \vec{r}_1 = \vec{r}$. On obtient donc l'équation du mouvement relatif

$$\boxed{m \ddot{\vec{r}} = \vec{F}(\vec{r})} \quad (1.9)$$

À partir d'un problème à deux corps, nous avons obtenu deux problèmes à un corps. Le centre de masse a un mouvement rectiligne uniforme dans le référentiel du laboratoire, et l'équation (1.9) du mouvement relatif représente le mouvement d'un corps fictif de masse m soumis à une force extérieure $\vec{F}$. On dit que le problème à deux corps est séparable.

1.2 Force centrale

Nous allons supposer que la force $\vec{F}$ est centrale, c'est-à-dire que

- La force est dirigée selon $\vec{r}$
- La norme et le sens de la force ne dépendent que de r .

Dans le contexte du problème à deux corps, $\vec{r} = \vec{r}_2 - \vec{r}_1$ est la position relative des deux corps, et est donc aussi la position de la particule fictive. Cependant, tout ce qui est dit dans ce chapitre s'applique également à une particule réelle dans un champ de force centrale.

Une force centrale est forcément conservative, c'est-à-dire qu'il existe un potentiel $V(r)$ tel que¹

$$\boxed{\vec{F}(r) = -\vec{\nabla} V(r) = -\frac{dV(r)}{dr} \vec{u}_r} \quad \text{avec} \quad \vec{u}_r = \frac{\vec{r}}{r}. \quad (1.10)$$

Le potentiel V ne dépend que de r , et pas de θ ou de ϕ , parce que la force $\vec{F}$ a seulement une composante selon $\vec{u}_r$, et pas selon $\vec{u}_\theta$ ou $\vec{u}_\phi$. Si la force entre les deux corps est attractive, comme représenté sur le dessin en figure 1.1, on a $\frac{dV(r)}{dr} > 0 \forall r > 0$.

Dorénavant, nous nous concentrons sur les équations (1.9) et (1.10), qui décrivent le mouvement d'une particule fictive de position $\vec{r}$ et masse m dans un potentiel extérieur $V(r)$ qui dépend seulement de la distance r à un point immobile donné O , dit le centre du champ de force. Une fois ce problème résolu, on pourra remonter aux mouvements individuels des deux corps par rapport à leur centre de masse en utilisant les équations (1.5).

1. Nous avons utilisé dans (1.10) l'expression du gradient en coordonnées sphériques $\vec{\text{grad}} f(r, \theta, \phi) = \frac{\partial f}{\partial r} \vec{u}_r + \frac{1}{r} \frac{\partial f}{\partial \theta} \vec{u}_\theta + \frac{1}{r \sin \theta} \frac{\partial f}{\partial \phi} \vec{u}_\phi$

1.2.1 Conservation du moment cinétique et loi des aires

Appliquons le théorème du moment cinétique à la particule fictive en O :

$$\frac{d\vec{L}}{dt} = \vec{r} \wedge \vec{F} = \vec{0} \quad \Rightarrow \quad \vec{L} \text{ est constant.} \quad (1.11)$$

Le moment cinétique de la particule fictive est $\vec{L} = \vec{r} \wedge m\dot{\vec{r}}$. Comme $\vec{r}$ et $\dot{\vec{r}}$ sont à chaque instant orthogonaux à $\vec{L}$ qui est constant, on déduit que la trajectoire de la particule fictive se trouve dans un plan P orthogonal à $\vec{L}$. On peut ainsi réduire ultérieurement le nombre de degrés de liberté du système. Dans le plan P , on utilisera les coordonnées polaires. En rajoutant un axe z orthogonal à P orienté selon $\vec{L}$, on aura des coordonnées cylindriques. On considère pour l'instant le cas $\vec{L} \neq \vec{0}$. Dans ce cas² :

$$\vec{L} = r\vec{u}_r \wedge m\dot{\vec{r}} = r\vec{u}_r \wedge m(\dot{r}\vec{u}_r + r\dot{\theta}\vec{u}_\theta) = mr^2\dot{\theta}\vec{u}_z \quad \text{et donc} \quad \ell = \|\vec{L}\| = mr^2\dot{\theta} = \text{Cste.} \quad (1.12)$$

La constance du module ℓ de $\vec{L}$ a une interprétation géométrique simple : l'aire élémentaire dA balayée par $\vec{r}$ pendant dt est constante, comme montré en figure 1.2.

$$dA = \frac{r^2}{2}d\theta \quad \Rightarrow \quad \boxed{\frac{dA}{dt} = \frac{r^2}{2} \frac{d\theta}{dt} = \frac{\ell}{2m}} \quad (1.13)$$

C'est la *loi des aires*, ou seconde loi de Kepler.

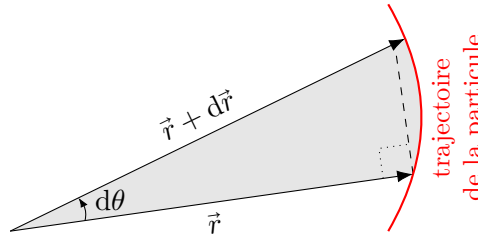


FIGURE 1.2 – Soit dA l'aire élémentaire balayée par $\vec{r}$ pendant dt , en gris sur la figure. L'aire du triangle (à gauche des pointillés) est $\frac{1}{2}r^2 \tan(d\theta) \approx \frac{1}{2}r^2 d\theta$. L'aire à droite des pointillés est négligeable (d'ordre $dr d\theta$). On a donc au premier ordre $dA = \frac{1}{2}r^2 d\theta$, et $dA/dt = \frac{1}{2}r^2 \dot{\theta} = \ell/(2m)$, qui est constant.

1.2.2 Potentiel effectif

Nous allons maintenant utiliser la conservation de l'énergie pour obtenir la solution complète du problème. La force étant conservative, dérivée du potentiel $V(r)$, l'énergie mécanique est conservée :

$$E = \frac{1}{2}mv^2 + V(r) = \text{const.} \quad (1.14)$$

En utilisant le fait que

$$\frac{1}{2}mv^2 = \frac{m}{2}\|\dot{r}\vec{u}_r + r\dot{\theta}\vec{u}_\theta\|^2 = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2) = \frac{m}{2}\left(\dot{r}^2 + \frac{\ell^2}{m^2r^2}\right) \quad \text{car} \quad \ell = mr^2\dot{\theta} \quad (1.15)$$

on se ramène à un problème purement radiale, à un seul degré de liberté, pour une particule fictive de masse m dans un potentiel effectif $U_{\text{eff}}(r)$

$$E = \frac{1}{2}m\dot{r}^2 + U_{\text{eff}}(r) \quad \text{avec} \quad U_{\text{eff}}(r) = V(r) + \frac{\ell^2}{2mr^2}. \quad (1.16)$$

2. Clairement, $\vec{L} = \vec{0}$ correspond à une trajectoire rectiligne pour la particule fictive.

1.2.3 Étude qualitative des trajectoires

Supposons pour simplifier que le potentiel $V(r)$ atteint une limite finie $V(\infty)$ pour r grand ; comme le potentiel est défini à une constante près, on s'arrange pour avoir $V(\infty) = 0$. On s'intéresse à deux situations :

Force attractive : si la force est toujours attractive entre les deux corps, alors $\frac{dV(r)}{dr} > 0$ et, puisque $V(\infty) = 0$, on a $V(r) < 0$.

Force répulsive : si la force est toujours répulsive entre les deux corps, alors $\frac{dV(r)}{dr} < 0$ et, puisque $V(\infty) = 0$, on a $V(r) > 0$.

(On exclut les forces qui sont attractives pour certaines valeurs de r et répulsives pour d'autres valeurs.)

Le potentiel effectif $U_{\text{eff}}(r)$ est la somme de $V(r)$, et du terme $\frac{\ell^2}{2mr^2}$ qui est positif et décroissant. Dans le cas répulsif, U_{eff} est donc la somme de deux termes positifs et décroissants, et est lui-même positif et décroissant. Dans le cas attractif, U_{eff} est la somme de deux termes de signes opposés et de sens de variation opposés. La situation est plus compliquée, mais on a typiquement une somme comme représentée sur la figure 1.3 : U_{eff} est d'abord décroissant, atteint un minimum négatif U_* pour un certain rayon r_* , puis croît jusqu'à 0 quand $r \rightarrow \infty$.

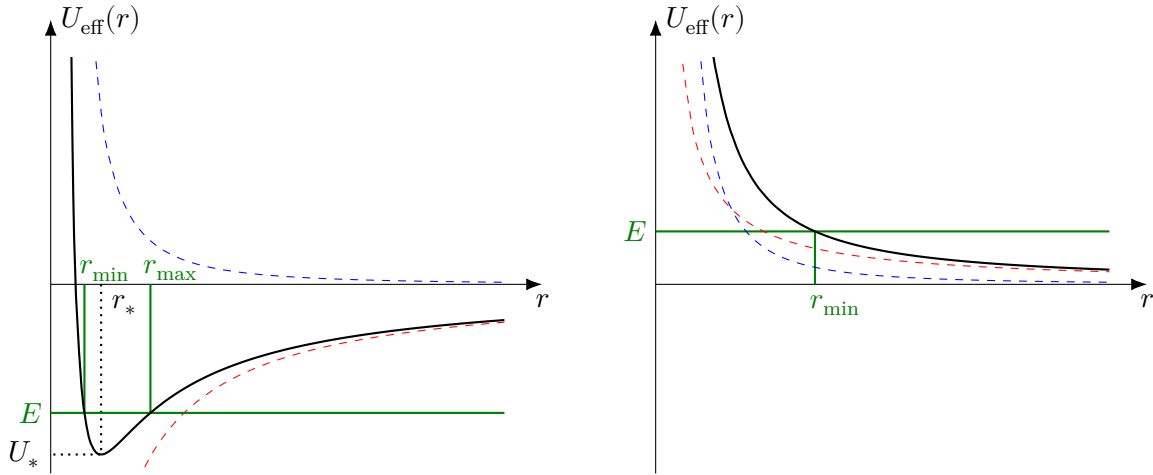


FIGURE 1.3 – Le potentiel effectif U_{eff} dans le cas attractif (à gauche) et dans le cas répulsif (à droite). Le potentiel effectif est la somme de deux termes, $V(r)$ (en pointillés rouges) et $\frac{\ell^2}{2mr^2}$ (en pointillés bleus). Dans le cas attractif, le potentiel effectif prend sa valeur minimal U_* pour une valeur de $r = r_*$. Si $E < 0$ (uniquement dans le cas attractif), le mouvement est limité aux valeurs de r comprises entre les *points de rebroussement* $r_{\min}$ et $r_{\max}$. Si $E > 0$ (dans le cas répulsif ou attractif), il n'y a qu'un point de rebroussement $r_{\min}$, et toutes les valeurs de r supérieures à $r_{\min}$ sont possibles.

Lorsque le système a une énergie E donnée, sa position r doit être telle que $U_{\text{eff}}(r) \leq E$, puisque $E - U_{\text{eff}}(r) = \frac{1}{2}m\dot{r}^2 \geq 0$, voir (1.16). On en déduit que

- $E \geq U_*$ dans le cas attractif, et $E > 0$ dans le cas répulsif.
- Lorsque $E = U_*$ (cas attractif uniquement), la seule valeur de r possible est $r = r_*$, et on a $\dot{r} = 0$. La trajectoire est un cercle de rayon r_* .
- Lorsque $U_* < E < 0$ (cas attractif uniquement), il existe deux valeurs de r appelées $r_{\min}$ et $r_{\max}$ pour lesquelles $U_{\text{eff}}(r) = E$, voir la figure 1.3. Ce sont les *points de rebroussement* pour le mouvement radial ; à ces points de rebroussement, la vitesse radiale $\dot{r}$ change de signe en passant par 0. La coordonnée r augmente ($\dot{r} > 0$) entre $r_{\min}$ et $r_{\max}$, puis diminue ($\dot{r} < 0$) de $r_{\max}$ à $r_{\min}$, puis augmente à nouveau, etc.

- Lorsque $E \geq 0$ (cas attractif ou répulsif), il n'y a qu'un seul point de rebroussement, $r_{\min}$. Si r est initialement grand et décroissant, il diminue jusqu'à $r_{\min}$ puis augmente à nouveau vers l'infini.

Supposons pour fixer les idées que, à $t = 0$ la particule est en $r_{\min}$. Pour $t > 0$, et jusqu'au moment où $r = r_{\max}$ (si on est dans le cas $E < 0$), on a $\dot{r} > 0$. On a donc, à partir de (1.16),

$$\dot{r} = \sqrt{\frac{2}{m}(E - U_{\text{eff}}(r))}. \quad (1.17)$$

Cette équation différentielle non-linéaire du premier ordre se résout par séparation des variables :

$$dt = \frac{dr}{\sqrt{\frac{2}{m}(E - U_{\text{eff}}(r))}}, \quad (1.18)$$

puis, après intégration, on obtient t en fonction de r :

$$t = \int_{r_{\min}}^r \frac{dr}{\sqrt{\frac{2}{m}(E - U_{\text{eff}}(r))}}. \quad (1.19)$$

(On rappelle que cette expression n'est valable que pour l'intervalle de temps où r va de $r_{\min}$ à $r_{\max}$, mais des expressions similaires peuvent être écrites pour les instants ultérieurs.)

En particulier, dans le cas $E < 0$, le temps nécessaire pour aller de $r_{\min}$ à $r_{\max}$ est

$$\Delta t = \int_{r_{\min}}^{r_{\max}} \frac{dr}{\sqrt{\frac{2}{m}(E - U_{\text{eff}}(r))}}. \quad (1.20)$$

Lorsque t atteint Δt , la distance r atteint $r_{\max}$ et la vitesse radiale $\dot{r}$ s'annule puis change de signe. Il faut modifier (1.17) en rajoutant un signe $-$. La trajectoire de r en fonction du temps est alors symétrique autour de Δt ; $r(\Delta t + t) = r(\Delta t - t)$. Le temps nécessaire pour aller de $r_{\max}$ à $r_{\min}$ est donc égal à Δt et, finalement, la valeur de r dans le cas $E < 0$ est une fonction périodique et symétrique du temps de période $2\Delta t$.

La trajectoire dans le plan P est-elle pour autant une courbe fermée ? Pour répondre à cette question, il faut regarder la variation de l'angle θ pendant ce temps. Supposons que $\ell > 0$. (En général, on s'arrange toujours pour orienter l'espace de manière à avoir $\ell > 0$. Nous ferons toujours cette hypothèse dans la suite.) On commence par remarquer que

$$\dot{\theta} = \frac{\ell}{mr^2} > 0 \quad (1.21)$$

L'angle θ grandit donc toujours, et la vitesse angulaire $\dot{\theta}$ ne s'annule jamais. Toujours par séparation des variables, on peut déterminer θ en fonction de r . Partant de (1.21), on a $d\theta = \frac{\ell}{mr^2} dt$, puis on utilise (1.18) et on intègre ; en supposant que $\theta = 0$ lorsque $r = r_{\min}$, on obtient

$$\theta = \int_{r_{\min}}^r \frac{\ell}{mr^2} \frac{dr}{\sqrt{\frac{2}{m}(E - U_{\text{eff}}(r))}} = \int_{r_{\min}}^r \frac{\frac{\ell}{r^2} dr}{\sqrt{2m(E - U_{\text{eff}}(r))}}, \quad (1.22)$$

expression uniquement valable jusqu'à ce que r vaille $r_{\max}$ si on est dans le cas $E < 0$. Lorsque r va de $r_{\min}$ à $r_{\max}$ (ce qui prend un temps Δt), l'angle augmente donc de

$$\Delta\theta = \int_{r_{\min}}^{r_{\max}} \frac{\frac{\ell}{r^2} dr}{\sqrt{2m(E - U_{\text{eff}}(r))}}. \quad (1.23)$$

Pendant la période $2\Delta t$ (temps pour que r aille de $r_{\min}$ à $r_{\max}$ et retour), l'angle augmente de $2\Delta\theta$. Si $2\Delta\theta$ est un multiple de 2π , alors le mouvement est périodique de période $2\Delta t$. Si $2\Delta\theta$ est de

la forme $2\pi \frac{m}{n}$ (où la fraction $\frac{m}{n}$ est irréductible), alors le mouvement est périodique de période $2\Delta t \times n$. Si $2\Delta\theta$ est de la forme $2\pi \times$ (nombre irrationnel), le mouvement n'est pas périodique.

Dans le cas où $E > 0$, le mouvement radial n'est pas limité, et on a seulement la contrainte $r \geq r_{\min}$. La particule peut venir de l'infini, s'approcher à une distance $r_{\min}$ du centre et repartir vers l'infini dans une direction différente. L'angle $\Delta\theta$ obtenu à partir de (1.23) en remplaçant $r_{\max}$ par $+\infty$ est alors l'angle entre la position de la particule à la distance minimale et la direction dans laquelle la particule repart vers l'infini. L'angle $2\Delta\theta$ est l'angle entre la direction d'approche (depuis l'infini) et la direction dans laquelle la particule repart vers l'infini.

1.3 Problème de Kepler

1.3.1 Équation de la trajectoire

Nous considérons dans cette section le problème de Kepler, qui correspond au cas de la force gravitationnelle

$$V(r) = -\frac{\alpha}{r} \quad \text{avec} \quad \alpha = Gm_1m_2 = GMm > 0 \quad (1.24)$$

où $M = m_1 + m_2$ est la masse totale, $m = \frac{m_1m_2}{m_1+m_2}$ est la masse réduite et G est la constante de gravitation universelle $[G] = M^{-1}L^3T^{-2}$ de valeur $G \simeq 6,674 \times 10^{-11} \text{ m}^3 \cdot \text{s}^{-2} \cdot \text{kg}^{-1}$ dans le Système International d'unités (SI).

Le potentiel effectif (1.16) est donc de la forme

$$U_{\text{eff}}(r) = -\frac{\alpha}{r} + \frac{\ell^2}{2mr^2} \quad (1.25)$$

En résolvant $\frac{dU_{\text{eff}}}{dr} = 0$, on peut calculer aisément la valeur r_* qui rend le potentiel effectif minimum et la valeur U_* du minimum correspondante

$$r_* = \frac{\ell^2}{m\alpha} \quad \text{et} \quad U_* = -\frac{1}{2} \frac{\alpha^2 m}{\ell^2} \quad (1.26)$$

En remplaçant $U_{\text{eff}}(r)$ dans la formule intégrale (1.22) pour l'angle θ , on obtient

$$\theta = \int_{r_{\min}}^r \frac{\frac{\ell}{r^2} dr}{\sqrt{2m \left(E + \frac{\alpha}{r} \right) - \frac{\ell^2}{r^2}}}. \quad (1.27)$$

Miraculeusement, cette intégrale se calcule. On remarque d'abord que le dénominateur est une fonction de $1/r$, et que le dr/r^2 du numérateur est justement, au signe près, la différentielle de $1/r$. On fait donc naturellement le changement de variable $x = 1/r$ pour obtenir

$$\theta = \int_{1/r_{\min}}^{1/r} \frac{-\ell dx}{\sqrt{2m(E + \alpha x) - \ell^2 x^2}}. \quad (1.28)$$

Le contenu de la racine carrée est un polynôme en x du second degré; on le met sous sa forme canonique, puis on factorise par $m^2\alpha^2/\ell^2$:

$$2m(E + \alpha x) - \ell^2 x^2 = -\ell^2 \left(x - \frac{m\alpha}{\ell^2} \right)^2 + 2mE + \frac{m^2\alpha^2}{\ell^2} = \frac{m^2\alpha^2}{\ell^2} \left[1 + \frac{2E\ell^2}{m\alpha^2} - \left(\frac{\ell^2}{m\alpha} x - 1 \right)^2 \right], \quad (1.29)$$

et donc

$$\theta = \int_{1/r_{\min}}^{1/r} \frac{-\frac{\ell^2}{m\alpha} dx}{\sqrt{1 + \frac{2E\ell^2}{m\alpha^2} - \left(\frac{\ell^2}{m\alpha} x - 1 \right)^2}}. \quad (1.30)$$

En introduisant le paramètre p et l'excentricité e de l'orbite :

$$\boxed{p = \frac{\ell^2}{m\alpha} = r_* \quad e = \sqrt{1 + \frac{2E\ell^2}{m\alpha^2}} = \sqrt{1 - \frac{E}{U_*}}} \quad (1.31)$$

on obtient

$$\theta = \int_{1/r_{\min}}^{1/r} \frac{-p \, dx}{\sqrt{e^2 - (px - 1)^2}} = \int_{1/r_{\min}}^{1/r} \frac{-\frac{p}{e} \, dx}{\sqrt{1 - \left(\frac{px-1}{e}\right)^2}}. \quad (1.32)$$

Cette intégrale se calcule à l'aide de la fonction arccos : on rappelle que $\arccos'(x) = -1/\sqrt{1-x^2}$, et donc $\frac{d}{dx} \arccos(ax+b) = -a/\sqrt{1-(ax+b)^2}$. On obtient donc

$$\theta = \arccos\left(\frac{p/r - 1}{e}\right) - \arccos\left(\frac{p/r_{\min} - 1}{e}\right). \quad (1.33)$$

Le deuxième terme du membre de droite est nul : en effet, les valeurs de $r_{\min}$ et $r_{\max}$ sont telles que $E = U_{\text{eff}}(r_{\min}) = U_{\text{eff}}(r_{\max})$. Cela signifie que la racine carrée dans (1.22) s'annule quand r vaut $r_{\min}$ ou $r_{\max}$. En suivant les étapes, on voit que la racine carrée dans (1.32) s'annule quand x vaut $1/r_{\min}$ ou $1/r_{\max}$. Comme $r_{\min}$ est plus petit que $r_{\max}$, cela implique que

$$\frac{p/r_{\min} - 1}{e} = +1 \quad \text{et} \quad \frac{p/r_{\max} - 1}{e} = -1, \quad \text{c'est-à-dire} \quad r_{\min} = \frac{p}{1+e} \quad \text{et} \quad r_{\max} = \frac{p}{1-e}. \quad (1.34)$$

Le deuxième terme dans le membre de droite de (1.33) est donc $\arccos(1) = 0$. Finalement, en prenant le cosinus des deux côtés, on obtient l'équation de la trajectoire :

$$\boxed{r = \frac{p}{e \cos \theta + 1}} \quad (1.35)$$

Exercices 1) Vérifier que cette solution vérifie la seconde loi de Newton, $m\ddot{\vec{r}} = -\frac{\alpha}{r^2}\vec{u}_r$, ce qui donne en projection $m(\ddot{r} - r\dot{\theta}^2) = -\frac{\alpha}{r^2}$. Indice : pour dériver par rapport au temps, il faut écrire $\frac{d}{dt} = \frac{d\theta}{dt} \times \frac{d}{d\theta} = \frac{\ell}{mr^2} \frac{d}{d\theta}$, voir (1.12). 2) Refaire le calcul dans le cas d'une force répulsive, où l'on a maintenant $V(r) = +\alpha/r$. (On trouve le même résultat, sauf que le $+1$ est remplacé par -1 .)

Une fois que l'on a la trajectoire (1.35) de la particule virtuelle, on peut remonter aux positions du problème à deux corps dans le référentiel $\mathcal{R}_c$ du centre de masse, voir (1.5) :

$$\vec{r}_{1,c} = -\frac{m_2}{m_1 + m_2} = -\frac{m}{m_1} \vec{r}, \quad \vec{r}_{2,c} = \frac{m_1}{m_1 + m_2} = \frac{m}{m_2} \vec{r}. \quad (1.36)$$

Chacun des deux corps a donc une trajectoire de la forme (1.35) autour du centre de masse, mais avec la valeur de p qui est modifiée. Dans le cas particulier où on aurait $m_1 \gg m_2$, comme dans le cas du Soleil et une planète telle la Terre, $\frac{m}{m_1} \simeq 0$ et $\frac{m}{m_2} \simeq 1$ si bien que $\vec{r}_{1,c} \simeq 0$ et $\vec{r}_{2,c} \simeq \vec{r}$.

1.3.2 Premières propriétés des trajectoires keplériennes

Le type de la trajectoire dépend de la valeur de l'excentricité e :

Si $e = 0$ (ce qui correspond à $E = U_*$), alors $r = p$ pour tout angle θ : **la trajectoire est circulaire.**

Si $e < 1$ (ce qui correspond à $U_* < E < 0$), alors $-1 < e \cos \theta < 1$ et le dénominateur dans (1.35) ne s'annule pour aucune valeur de θ : la valeur de r reste bornée. En fait, on peut facilement déterminer les valeurs extrémales de r : le minimum $r_{\min}$ est atteint quand $\cos \theta$ est maximal ($\cos \theta = 1$ pour $\theta = 0$), et le maximum $r_{\max}$ est atteint quand $\cos \theta$ est minimal ($\cos \theta = -1$ pour $\theta = \pi$). On retrouve les valeurs données dans (1.34) :

$$r_{\min} = \frac{p}{1+e}, \quad r_{\max} = \frac{p}{1-e}. \quad (1.37)$$

La trajectoire est une ellipse. Ce cas est étudié section 1.3.3.

Si $e > 1$ (ce qui correspond à $E > 0$), alors le dénominateur dans (1.35) s'annule pour $\cos \theta = -1/e$, c'est-à-dire pour $\theta = \pm \arccos(-1/e)$. Pour ces angles, la valeur r est infinie. Physiquement, l'objet "vient de l'infini" avec l'angle $-\arccos(-1/e)$, atteint la distance minimale $r_{\min} = p/(1+e)$ pour $\theta = 0$, puis repart vers l'infini avec l'angle $+\arccos(-1/e)$. **La trajectoire est une hyperbole.** Ce cas est étudié section 1.3.4.

Mathématiquement, le cas $e = 1$ correspond à une parabole (r est infini pour $\theta = \pi$), mais physiquement, e sera soit très légèrement inférieur à 1 (ellipse), soit très légèrement supérieur à 1 (hyperbole), mais jamais exactement égal à 1.

La famille de courbes décrites par (1.35) (cercle, ellipse, parabole, hyperbole) s'appelle *les coniques*, parce que ce sont les courbes que l'on obtient en prenant l'intersection d'un cône et d'un plan.

1.3.3 Trajectoire elliptique

Voici quelques caractérisations des ellipses, illustrées par la figure 1.4 :

— L'ensemble des points (x, y) tels que

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (1.38)$$

est une ellipse centrée sur l'origine $(0, 0)$, de grand axe horizontal $2a$ et de petit axe vertical $2b$. C'est, littéralement, un cercle aplati : si (X, Y) appartiennent au cercle de rayon a , alors $(x = X, y = \frac{b}{a}Y)$ appartient à l'ellipse dont on vient de donner l'équation. Par conséquent, sa surface est $\frac{b}{a} \times \pi a^2 = \pi ab$. On appelle a le demi-grand axe et b le demi-petit axe.

— Si on choisit deux points F et F' dans le plan et une longueur L supérieure à la distance FF' , l'ensemble de tous les points M tels que

$$FM + F'M = L \quad (1.39)$$

est une ellipse. Les points F et F' s'appellent les *foyers* de l'ellipse. En considérant les deux points de l'ellipse sur l'axe (FF') , on voit que l'on doit avoir $L = 2a$. Si on pose $2c = FF'$ et que l'on considère les deux points de l'ellipse sur la médiatrice de $[FF']$, on voit que l'on doit avoir $a^2 = b^2 + c^2$.

— En coordonnées polaires, l'équation de l'ellipse s'écrit $r = p/(e \cos \theta + 1)$, mais dans cette description **le centre du repère est le foyer F** , pas le centre de l'ellipse.

On peut relier les demi-axes a et b aux paramètres p et à l'excentricité e de l'ellipse en passant par les valeurs de $r_{\min}$ et $r_{\max}$ données dans (1.37) ; en effet, il est clair géométriquement que

$$2a = r_{\min} + r_{\max} = \frac{p}{1+e} + \frac{p}{1-e} = \frac{2p}{1-e^2} \quad \Rightarrow \quad \boxed{a = \frac{p}{1-e^2}} \quad (1.40)$$

et que

$$c = a - r_{\min} = \frac{p}{1-e^2} - \frac{p}{1+e} = \frac{pe}{1-e^2} = ea. \quad (1.41)$$

On obtient enfin b par Pythagore :

$$b^2 = a^2 - c^2 = a^2(1-e^2) = ap = \frac{p^2}{1-e^2} \quad \Rightarrow \quad \boxed{b = \frac{p}{\sqrt{1-e^2}}} \quad (1.42)$$

On a donc obtenu a et b en fonction de p et e . On peut inverser ces relations et obtenir p et e en fonction de a et b ; on trouve (exercice) :

$$p = \frac{b^2}{a}, \quad e = \sqrt{1 - \frac{b^2}{a^2}}. \quad (1.43)$$

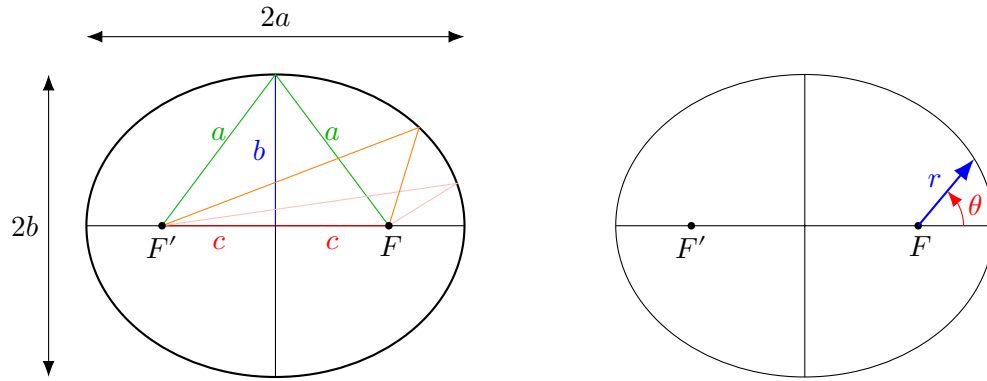


FIGURE 1.4 – Ellipse de demi-axes a et b , et position des foyers. À gauche : la longueur totale des lignes vertes, ou celle des lignes oranges, ou celle des lignes roses est égale à $2a$. En choisissant l'origine au centre de l'ellipse, l'équation cartésienne de l'ellipse est $x^2/a^2 + y^2/b^2 = 1$. À droite : en choisissant l'origine des axes sur le foyer F , l'équation de l'ellipse en coordonnées polaires est $r = p/(e \cos \theta + 1)$. Pour l'ellipse de la figure, on a $a = 2,5$ cm, $b = 2$ cm, $c = 1,5$ cm, $p = 1,6$ cm et $e = 0,6$.

Nous sommes maintenant en mesure de démontrer la *troisième loi de Kepler*, qui relie la période T de révolution au demi-grand axe a de l'orbite elliptique. La loi des aires (1.13) intégrée sur une période donne que l'aire de l'ellipse est $\frac{\ell}{2m}T$. Par ailleurs, l'aire de l'ellipse est πab . En écrivant l'égalité au carré et en utilisant l'égalité $p = \frac{\ell^2}{m\alpha}$ voir (1.31), on obtient

$$\pi^2 a^2 b^2 = T^2 \frac{\ell^2}{4m^2} = T^2 \frac{p\alpha}{4m} \quad (1.44)$$

Par ailleurs on a montré que $p = b^2/a$ et on sait (par définition de α) que $\alpha = GMm$. On a donc

$$\boxed{\frac{a^3}{T^2} = \frac{GM}{4\pi^2}} \quad (1.45)$$

Ici $M = m_1 + m_2$ est la masse totale. Encore une fois, dans le cas d'une planète du système solaire, M est approximativement égale à la masse du Soleil, et on voit que pour les différentes planètes du système solaire, le rapport entre le cube du demi-grand axe de l'orbite et le carré de la période est une constante ; c'est la troisième loi de Kepler.

1.3.4 Trajectoire hyperbolique

Considérons toujours le problème de Kepler avec le potentiel attractif (1.24), mais dans le cas où $E > 0$, et donc $e > 1$. Dans ce cas, la courbe décrite par l'équation (1.35) est une branche d'hyperbole qui contourne son foyer F , centre du champ de force.

Une branche d'hyperbole est une courbe avec deux asymptotes, voir figure 1.5. La courbe est symétrique, et le point de la courbe le plus proche du foyer est le *sommet* S de l'hyperbole, qui est sur l'axe de symétrie. On note ϕ_0 l'angle entre l'axe de symétrie et les asymptotes.

L'équation de la trajectoire en coordonnées polaires par rapport au point F est encore donnée par l'équation (1.35), à savoir $r = p/(e \cos \theta + 1)$, mais l'angle θ ne varie que dans un certain intervalle de valeurs : on a vu section 1.3.2 que θ varie entre $-\arccos(-1/e)$ et $+\arccos(-1/e)$. Par définition de ϕ_0 , on a par ailleurs que θ varie entre $-(\pi - \phi_0)$ et $\pi - \phi_0$. On a donc $\pi - \phi_0 = \arccos(-1/e)$, ce qui implique que

$$\cos \phi_0 = \frac{1}{e}. \quad (1.46)$$

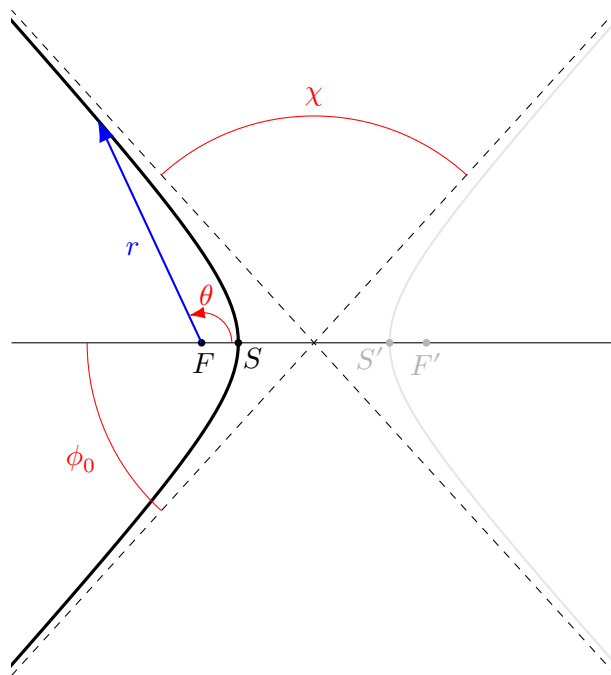


FIGURE 1.5 – Branche d’hyperbole tournant autour de son foyer F dans le cas d’un potentiel attractif en $1/r$. L’équation polaire de cette branche par rapport au foyer F est $r = p/(e \cos \theta + 1)$. Le sommet S est le point où r est minimal, obtenu pour $\theta = 0$. L’axe horizontal est un axe de symétrie, et ϕ_0 est l’angle entre cet axe et les asymptotes (en pointillés). L’angle $\chi = \pi - 2\phi_0$ est l’angle de déviation de la particule. La courbe représentée en gris clair est la deuxième branche de l’hyperbole ; elle ne nous intéresse pas pour le potentiel attractif.

On note χ l’angle de déflexion de la particule, c’est-à-dire l’angle entre la direction d’arrivée et la direction finale de la particule. On a clairement $\chi = \pi - 2\phi_0$. Par ailleurs, l’angle $\theta = 0$ correspond au périhélie (point où r est minimal), et on a encore $r_{\min} = \frac{p}{e+1}$.

Remarque : on peut caractériser une hyperbole de manière similaire aux caractérisations de l’ellipse dans la section précédente : l’équation d’une hyperbole en cartésien peut s’écrire $x^2/a^2 - y^2/b^2 = 1$, c’est l’ensemble de tous les points M tels que $F'M - FM = 2a$, etc., et on peut relier les paramètres a et b aux paramètres e et p . Dans ce cours, nous ne rentrons pas dans ces détails techniques.

1.3.5 Potentiel répulsif en $1/r$ et diffusion de Rutherford

Considérons maintenant le cas d’un potentiel répulsif en $1/r$:

$$V(r) = \frac{\alpha}{r}, \quad \alpha > 0. \quad (1.47)$$

C’est le potentiel électrostatique entre deux charges q_1 et q_2 de même signe ; on a alors $\alpha = q_1 q_2 / (4\pi\epsilon_0)$.

Comme illustré dans la partie droite de la figure 1.3, Le potentiel effectif $U_{\text{eff}}(r)$ est maintenant une fonction monotone décroissante, et l’énergie E est forcément positive.

On peut refaire les calculs de la section 1.3.1 et on trouve (comme déjà mentionné) une expression très similaire à (1.35), mais avec un -1 à la place du $+1$:

$$r = \frac{p}{e \cos \theta - 1} \quad (1.48)$$

Cette équation décrit la deuxième branche de l’hyperbole de la figure 1.5 du cas attractif ; voir la figure 1.6. L’angle θ varie maintenant entre $-\arccos(1/e)$ et $+\arccos(1/e)$. Par ailleurs, par

définition de ϕ_0 (angle entre l'axe de symétrie et les asymptotes), l'angle θ varie entre $-\phi_0$ et $+\phi_0$. On a donc $\phi_0 = \arccos(1/e)$ et, comme dans le cas attractif, $1/e = \cos(\phi_0)$. L'angle de déviation χ est lui aussi donné par la même formule que dans le cas précédent : $\chi = \pi - 2\phi_0$.

La distance minimale au centre de force F est obtenue pour $\theta = 0$ et vaut :

$$r_{\min} = \frac{p}{e - 1}. \quad (1.49)$$

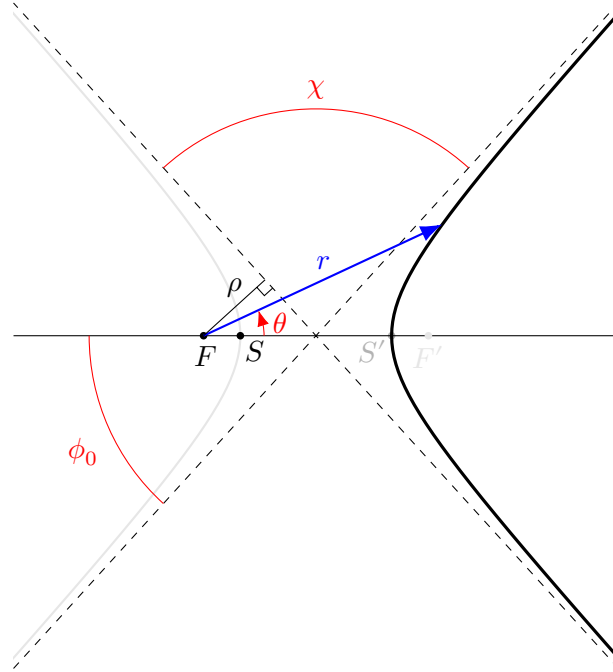


FIGURE 1.6 – Dans le cas d'un potentiel répulsif, c'est la branche de l'hyperbole extérieure au centre de force qui représente la trajectoire physique. χ est l'angle de déflexion d'une particule suivant cette trajectoire. (On imagine que la particule vient du coin inférieur droit et repart par le coin supérieur droit.) ρ est le *paramètre d'impact* : c'est la distance minimale entre le centre F et la particule si celle-ci n'avait pas été déviée.

On souhaite maintenant exprimer l'angle de déflexion en fonction de deux paramètres physiques importants :

La vitesse à l'infini v_∞ , qui est la norme de la vitesse de la particule lorsque r est très grand, c'est-à-dire longtemps avant l'approche de F ou longtemps après.

Le paramètre d'impact ρ , qui est la distance minimale qu'il y aurait eu entre la particule et le centre d'attraction si la particule n'avait pas été déviée, voir figure 1.6.

(Remarque : on peut aussi définir ces quantités dans le cas attractif.)

Lorsque la particule est infiniment loin du centre de la force, le potentiel est nul et l'énergie est purement cinétique. On a donc

$$E = \frac{1}{2}mv_\infty^2. \quad (1.50)$$

(Lorsque la particule se rapproche du centre, $V(r)$ est positif et, puisque E est conservée, la vitesse v est inférieure à v_∞ .)

Toujours lorsque la particule est infiniment loin du centre de la force, elle est sur l'asymptote à l'hyperbole et sa vitesse est dirigée selon cette asymptote. On en déduit que la norme ℓ du moment cinétique est donnée par

$$\ell = |\vec{r} \wedge m\vec{v}| = \rho mv_\infty. \quad (1.51)$$

On voit donc que, connaissant ρ , v_∞ et α , on peut calculer le paramètre p et l'excentricité e de l'hyperbole. Pour l'excentricité :

$$e = \sqrt{1 + \frac{2E\ell^2}{m\alpha^2}} = \sqrt{1 + \left(\frac{\rho m v_\infty^2}{\alpha}\right)^2} \quad (1.52)$$

Calculons l'angle ϕ_0 entre les asymptotes de l'hyperbole et l'axe de symétrie. On a :

$$\phi_0 = \arccos \frac{1}{e} = \arccos \frac{1}{\sqrt{1 + \left(\frac{\rho m v_\infty^2}{\alpha}\right)^2}}. \quad (1.53)$$

ou, plus simplement,

$$\boxed{\phi_0 = \arctan \frac{\rho m v_\infty^2}{\alpha}} \quad (1.54)$$

Pour cette dernière expression, on a utilisé la relation de trigonométrie $\arccos \frac{1}{\sqrt{1+x^2}} = \arctan x$ démontrée à l'aide d'un dessin figure 1.7.

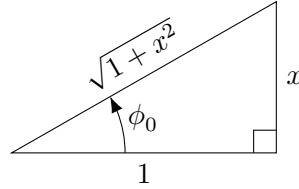


FIGURE 1.7 – Un triangle rectangle de petits côtés 1 et x , d'après Pythagore, une hypoténuse égale à $\sqrt{1+x^2}$. Avec l'angle ϕ_0 défini sur la figure, on a $\cos \phi_0 = 1/\sqrt{1+x^2}$ et $\tan \phi_0 = x$. On en déduit que $\phi_0 = \arccos \frac{1}{\sqrt{1+x^2}} = \arctan x$.

En 1909, Hans Geiger, Ernest Marsden et Ernest Rutherford ont fait une expérience importante en envoyant des particules alpha (des noyaux d'hélium) sur une feuille d'or. Les particules alpha interagissent avec les noyaux d'or selon une interaction électrostatique répulsive $V(r) = \alpha/r$ avec

$$\alpha = \frac{q_{\text{alpha}} q_{\text{noyau d'or}}}{4\pi\epsilon_0} = \frac{2e \times 79e}{4\pi\epsilon_0} = 3,6 \cdot 10^{-26} \text{ J} \cdot \text{m}. \quad (1.55)$$

(Avec $e = 1,6 \cdot 10^{-19} \text{ C}$ la charge élémentaire, $\epsilon_0 = 8,85 \cdot 10^{-12} \text{ F/m}$. Il y a 79 protons dans un atome d'or et 2 dans une particule alpha.)

Les particules alpha sont envoyées avec une énergie d'environ $5 \text{ MeV} \simeq 5 \cdot 10^6 \times 1,6 \cdot 10^{-19} = 8 \cdot 10^{-13} \text{ J}$. On a donc

$$m v_\infty^2 \simeq 1,6 \cdot 10^{-12} \text{ J}. \quad (1.56)$$

(Nota : on confond la masse réduite m avec la masse de la particule alpha parce que le noyau d'or est beaucoup plus lourd.)

En combinant ces deux résultats, on calcule $\alpha/(m v_\infty^2)$ et on obtient donc avec (1.54) :

$$\phi_0 \simeq \arctan \frac{\rho}{2,3 \cdot 10^{-14} \text{ m}}. \quad (1.57)$$

Pour la grande majorité des particules alpha envoyées, le paramètre d'impact ρ était grand devant $2,3 \cdot 10^{-14} \text{ m}$. L'angle ϕ_0 est alors l'arctangente d'un nombre très grand, et vaut presque $\pi/2$. L'angle de déviation $\chi = \pi - 2\phi_0$ est presque nul.

Cependant, une petite fraction des particules alpha furent très déviées, parfois à plus de 90° . Pour avoir $\chi \geq \frac{\pi}{2}$, il faut que $\phi_0 \leq \frac{\pi}{4}$, et donc $\rho \leq 2,3 \cdot 10^{-14} \text{ m}$ puisque $\tan \frac{\pi}{4} = 1$. Ce qui est intéressant, c'est que le paramètre d'impact est forcément plus grand que l'objet qui dévie la particule alpha. C'est à la suite de cette expérience que Rutherford a émis l'hypothèse que la charge positive d'un atome occupait un très faible espace, de l'ordre de 10^{-14} m ou plus petit : c'est ce qu'on appelle le noyau atomique.

Chapitre 2

Équations de Lagrange

Dans le formalisme lagrangien, les équations du mouvement prennent une forme scalaire, très élégante et compacte, ayant la propriété remarquable d'être invariante par changement de systèmes de coordonnées. Ce formalisme, qui se révèle fructueux dans de nombreux domaines de la physique dont la théorie quantique des champs et la relativité, est aussi précieux en mécanique car il permet d'absorber toutes les contraintes mécaniques d'un système dans un choix opportun de coordonnées indépendantes.

Dans ce chapitre, nous allons commencer par donner sans justification les équations de Lagrange (section 2.1), puis montrer comment les mettre en pratique sur plusieurs exemples (section 2.2). Enfin, nous donnerons quelques propriétés de ces équations (section 2.3) puis nous terminerons avec deux justifications des équations de Lagrange (sections 2.4 et 2.5).

2.1 Les équations de Lagrange

2.1.1 Introduction

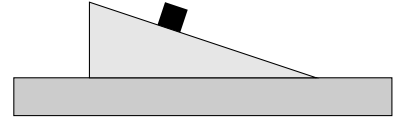
Pour repérer un point matériel dans l'espace, on a besoin de trois coordonnées, x , y et z . Pour repérer N points matériels, il faut $3N$ coordonnées, $x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_N, y_N$ et z_N .

Cependant, le choix des coordonnées cartésiennes n'est pas toujours le plus naturel, ni le plus pratique. On peut par exemple faire le choix d'utiliser les coordonnées sphériques ou cylindriques. En fait, le choix des coordonnées est libre tant qu'elles permettent de décrire entièrement le système que l'on considère.

Bien choisir les coordonnées est crucial quand le système est soumis à des contraintes ; prenons quelques exemples

- **Un solide** Il est inutile de donner les positions de chacun des atomes du solide ! Les atomes en question sont contraints de rester toujours positionnés de la même manière les uns par rapport aux autres, et il suffit de donner la position du centre de masse (trois coordonnées) et trois angles pour orienter le solide dans l'espace.
- **Un bateau voguant sur les océans** Pour donner la position de ce système, il serait ridicule d'utiliser les coordonnées cartésiennes (x, y, z) . Le bateau est contraint de rester à la surface de la mer et deux coordonnées suffisent, par exemple la longitude et la latitude, pour donner sa position. À ces deux coordonnées, il faut en ajouter une troisième pour donner le cap, c'est-à-dire l'orientation du bateau.
- **Le chariot d'une montagne russe** Ici, le chariot est contraint de rester sur la piste. Une seule coordonnée suffit à décrire sa position : la distance parcourue le long de la piste depuis le point de départ.

- **Un exercice de mécanique** On considère en deux dimensions un plan incliné pouvant glisser sur une table. On pose un solide (assimilable à un point matériel) sur le plan incliné.



Le plan est contraint de reposer sur la table, et le solide de reposer sur le plan. Le système peut être décrit par deux coordonnées : l'abscisse x du plan, et la distance ℓ entre le coin supérieur du plan et le solide.

Le choix des coordonnées est libre, mais il faut quand même faire attention : on sait que l'écriture du vecteur accélération est plus compliqué en coordonnées sphériques qu'en coordonnées cartésiennes !

La formulation lagrangienne des lois de la mécanique permet d'utiliser assez facilement n'importe quel système de coordonnées. En particulier, elle permet ainsi de prendre en compte de manière automatique toutes les contraintes.

2.1.2 Notations et définitions

On considère un système de N points matériels (ou de solides en translation). On pose

$$\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N, \quad (2.1)$$

les vecteurs position des N points.

- On note s le **nombre de degrés de libertés**, qui est le nombre minimal de coordonnées nécessaires pour décrire entièrement la position du système en tenant compte des éventuelles contraintes. Ainsi, un point matériel dans l'espace a trois degrés de liberté, qui peuvent être (x, y, z) ou (r, θ, ϕ) ou (ρ, z, ϕ) ou autre. Comme on l'a vu plus haut, un solide indéformable en a six, le système constitué d'un bloc glissant sur un plan incliné mobile en a deux.
- Pour un système à s degrés de libertés, on se donne s **coordonnées généralisées**

$$q_1, q_2, \dots, q_s \quad \text{où } s \leq 3N. \quad (2.2)$$

Ces coordonnées sont choisies de manière à caractériser entièrement la position du système. On peut donc écrire les vecteurs positions $\vec{r}_1, \dots, \vec{r}_N$ comme des fonctions des $q_1, \dots, q_s$ et (éventuellement) du temps.

- Lors du mouvement, les coordonnées généralisées dépendent du temps. On introduit les **vitesse généralisées** $\dot{q}_i = \frac{dq_i}{dt}$ et les **accélérations généralisées** $\ddot{q}_i = \frac{d^2q_i}{dt^2}$. Attention ! Comme les coordonnées généralisées n'ont pas forcément la dimension d'une longueur, les vitesses généralisées n'ont pas forcément la dimension d'une vitesse.
- On établit les **équations du mouvement**, qui sont des relations entre les coordonnées q_i , les vitesses $\dot{q}_i$, les accélérations $\ddot{q}_i$ et (éventuellement) du temps t .
- Enfin, on cherche à intégrer (c'est-à-dire résoudre) les équations du mouvement pour les conditions initiales $q_i(0)$ et $\dot{q}_i(0)$ qui nous intéressent, afin d'obtenir la trajectoire $q_i(t)$.

2.1.3 Les équations de Lagrange

On considère un système mécanique **sans forces magnétiques**, et **sans forces dissipatives**, c'est-à-dire des forces de frottement qui travaillent. (Les équations de Lagrange peuvent prendre en compte les forces dissipatives et les forces magnétiques au prix d'une complexité accrue. Dans ce cours, on ne considère que le cas simple.)

Voici la méthode pour obtenir les équations du mouvement

1. On choisit les coordonnées généralisées q_i .

2. On calcule l'énergie cinétique T du système en fonction des coordonnées généralisées q_i , des vitesses généralisées $\dot{q}_i$ et (éventuellement) du temps t :

$$T = T(q_1, \dots, q_s, \dot{q}_1, \dots, \dot{q}_s, t). \quad (2.3)$$

3. On calcule l'énergie potentielle d'interaction U du système en fonction des coordonnées généralisées et (éventuellement) du temps t . (L'énergie potentielle ne dépend pas des vitesses généralisées.)

$$U = U(q_1, \dots, q_s, t). \quad (2.4)$$

4. On forme le **lagrangien** $\mathcal{L}$ du système

$$\mathcal{L} = T - U. \quad (2.5)$$

(Attention au signe moins ! Le lagrangien est une quantité différente de l'énergie mécanique $T + U$.)

5. On écrit les équations de Lagrange, aussi appelées équations d'Euler-Lagrange :

$$\boxed{\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) = \frac{\partial \mathcal{L}}{\partial q_i} \quad \text{pour } i = 1, 2, \dots, s} \quad (2.6)$$

Remarque : on commence par dériver $\mathcal{L}$ par rapport aux coordonnées q_i (à droite) et par rapport aux vitesses généralisées $\dot{q}_i$ (à gauche). Pour faire ces dérivées il faut donc considérer les q_i et les $\dot{q}_i$ comme des variables indépendantes. Par contre, lorsque l'on dérive par rapport au temps, on considère les q_i et les $\dot{q}_i$ comme des fonctions du temps. (Par exemple, $\frac{d}{dt}(q_1^2 \dot{q}_2) = 2q_1 \dot{q}_1 \dot{q}_2 + q_1^2 \ddot{q}_2$.)

2.2 Exemples

2.2.1 Une particule dans un potentiel

On considère un point matériel de masse m soumis à une force $\vec{F}$ dérivant d'un potentiel U :

$$\vec{F} = -\overrightarrow{\text{grad}} U. \quad (2.7)$$

Choix des coordonnées généralisées Il n'y a aucune contrainte au mouvement de la particule, le nombre de degrés de liberté est 3 et on choisit naturellement les coordonnées cartésiennes de la particule : $q_1 = x$, $q_2 = y$ et $q_3 = z$.

Calcul de l'énergie cinétique La vitesse de la particule est le vecteur $\vec{v} = \dot{x} \vec{e}_x + \dot{y} \vec{e}_y + \dot{z} \vec{e}_z$ et son énergie cinétique est

$$T = \frac{1}{2} m \vec{v}^2 = \frac{1}{2} m (\dot{x}^2 + \dot{y}^2 + \dot{z}^2). \quad (2.8)$$

Calcul de l'énergie potentielle L'énergie potentielle est donnée dans l'énoncé : c'est $U = U(x, y, z)$. (Rappel : l'énergie potentielle est un champ, c'est-à-dire une fonction de l'espace).

Écriture du lagrangien

$$\mathcal{L} = T - U = \frac{1}{2} m (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - U(x, y, z) \quad (2.9)$$

Écriture des équations du mouvement Il y a $s = 3$ coordonnées, il faut donc écrire trois équations. On commence par la coordonnée x :

$$\frac{\partial \mathcal{L}}{\partial x} = -\frac{\partial U}{\partial x}, \quad \frac{\partial \mathcal{L}}{\partial \dot{x}} = m\dot{x}, \quad \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}} \right) = m\ddot{x}. \quad (2.10)$$

On obtient alors l'équation du mouvement à partir de l'équation de Lagrange (2.6) :

$$m\ddot{x} = -\frac{\partial U}{\partial x} = F_x, \quad (2.11)$$

où $F_x = -\frac{\partial U}{\partial x}$ est la composante selon x de la force $\vec{F}$, voir (2.7). On a retrouvé le principe fondamental de la dynamique projeté selon l'axe $\vec{e}_x$. On trouve de même $m\ddot{y} = F_y$ et $m\ddot{z} = F_z$.

2.2.2 Le pendule pesant

On considère un point matériel de masse m relié à un point fixe O par un fil idéal (sans masse, sans torsion, inextensible, ...) de longueur ℓ , oscillant dans un plan donné. Le point est soumis à son poids $m\vec{g}$, et la force de tension $\vec{F}$ du fil.

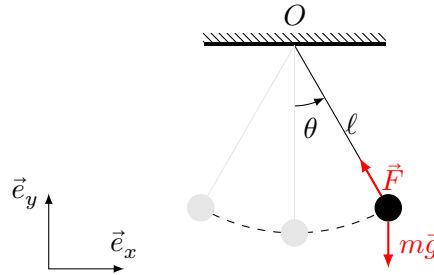


FIGURE 2.1 – Le pendule pesant

Choix des coordonnées généralisées Le mouvement de la masse est contraint de demeurer dans un plan, à une distance constante (égale à ℓ) du point O . Pour décrire sa position, il suffit d'une seule coordonnée. On choisit l'angle orienté θ par rapport à la verticale.

$$s = 1, \quad q_1 = \theta \quad (2.12)$$

(d'autres choix de la coordonnée sont possible ; celui-ci est le plus naturel.) Attention : ℓ , m , g sont des paramètres du problème, qui sont fixés une fois pour toute. Ce ne sont pas des coordonnées.

Calcul de l'énergie cinétique $\dot{\theta}$ est la vitesse angulaire du point matériel. Sa vitesse est donc $\ell\dot{\theta}$, et son énergie cinétique est

$$T = \frac{1}{2}m(\ell\dot{\theta})^2 = \frac{1}{2}m\ell^2\dot{\theta}^2. \quad (2.13)$$

Calcul de l'énergie potentielle La distance entre la masse et le plafond est $\ell \cos \theta$. En prenant l'origine des altitudes en O , l'altitude y est donc $y = -\ell \cos \theta$. (Attention au signe : on a pris un axe y orienté vers le haut.) L'énergie potentielle de pesanteur est alors

$$U = mgy = -mg\ell \cos \theta. \quad (2.14)$$

Nota : la force $\vec{F}$ n'intervient pas dans ce calcul. Son seul rôle est de garantir la contrainte (distance entre O et la masse constante et égale à ℓ) et, d'ailleurs cette force ne travaille pas puisqu'elle est toujours perpendiculaire au mouvement : il n'y a pas d'énergie potentielle associée.

Écriture du lagrangien

$$\mathcal{L} = T - U = \frac{1}{2}m\ell^2\dot{\theta}^2 + mg\ell \cos \theta. \quad (2.15)$$

Écriture des équations du mouvement Il n'y a qu'un degré de liberté (une seule coordonnée généralisée), et donc il n'y a qu'une équation à écrire. On a

$$\frac{\partial \mathcal{L}}{\partial \theta} = -mg\ell \sin \theta, \quad \frac{\partial \mathcal{L}}{\partial \dot{\theta}} = m\ell^2\dot{\theta}, \quad \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\theta}} \right) = m\ell^2\ddot{\theta}. \quad (2.16)$$

On obtient donc l'équation du mouvement à partir de l'équation de Lagrange (2.6) :

$$m\ell^2\ddot{\theta} = -mg\ell \sin \theta \quad (2.17)$$

ou, après simplification,

$$\ddot{\theta} + \frac{g}{\ell} \sin \theta = 0. \quad (2.18)$$

Remarques Les équations écrites ci-dessus sont valides quelle que soit l'amplitude des oscillations. Cependant, dans la limite des “petites oscillations”, on a $\theta \ll 1$ et l'on peut écrire

$$\cos \theta \simeq 1 - \frac{\theta^2}{2}, \quad \sin \theta \simeq \theta, \quad (2.19)$$

Le lagrangien et les équations du mouvement peuvent alors se mettre sous la forme

$$\mathcal{L} = \frac{1}{2}M\dot{\theta}^2 - \frac{1}{2}M\omega^2\theta^2 + \text{Cste}, \quad \ddot{\theta} + \omega^2\theta = 0, \quad (2.20)$$

où l'on a noté $M = m\ell^2$ et $\omega = \sqrt{g/\ell}$. Il est facile de voir que la constante dans le lagrangien n'a aucune influence sur les équations du mouvement, et on peut l'enlever pour simplifier. Les équations obtenues sont les mêmes que pour une masse M accrochée à un ressort de raideur $M\omega^2$. Un système décrit par des équations de la forme (2.20) s'appelle un *oscillateur harmonique* à une dimension. Les solutions peuvent s'écrire sous plusieurs formes équivalentes

$$x = c_1 \cos(\omega t) + c_2 \sin(\omega t) \quad \text{ou} \quad x = a \cos(\omega t + \phi) \quad \text{ou} \quad x = \text{Re}(Ae^{i\omega t}), \quad (2.21)$$

avec des constantes (c_1 et c_2 , ou a et ϕ , ou $A \in \mathbb{C}$) à déterminer selon les conditions initiales. ω est la *pulsation*, a est l'*amplitude*, ϕ est la *phase* et $A = ae^{i\phi}$ est l'*amplitude complexe*.

2.2.3 Une perle coulissant sur une tige en rotation

On considère une petite perle de masse m enfilée sur une tige sur laquelle la perle glisse sans frotter. La tige est maintenue en rotation avec une vitesse angulaire ω constante dans un plan horizontal.

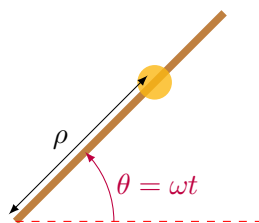


FIGURE 2.2 – La perle sur la tige en rotation

Choix des coordonnées généralisées La position de la perle est entièrement déterminée si on se donne la distance ρ entre la perle et le centre de rotation. On a donc $s = 1$ coordonnée généralisée.

Calcul de l'énergie cinétique Il y a deux composantes au vecteur vitesse de la perle. D'une part, lorsque ρ varie, la perle a une vitesse $\dot{\rho}$ le long de l'axe de la tige. D'autre part, la perle est entraînée par la tige et acquiert une vitesse $\rho\omega$ perpendiculaire à la tige. (C'est la vitesse du point de la tige où se trouve la perle.) Si on introduit les vecteurs polaires $\vec{e}_\rho$ et $\vec{e}_\theta$, la vitesse est

$$\vec{v} = \dot{\rho}\vec{e}_\rho + \rho\omega\vec{e}_\theta. \quad (2.22)$$

Une autre manière de comprendre ce vecteur vitesse est d'écrire le vecteur position

$$\vec{r} = \rho\vec{e}_\rho = \rho(\cos(\omega t)\vec{e}_x + \sin(\omega t)\vec{e}_y), \quad (2.23)$$

puis de dériver, en se souvenant que $\frac{d\vec{e}_\rho}{dt} = \dot{\theta}\vec{e}_\theta$ et que $\dot{\theta} = \omega$.

À partir du vecteur vitesse, il est facile d'obtenir l'énergie cinétique :

$$T = \frac{1}{2}m(\vec{v})^2 = \frac{1}{2}m(\dot{\rho}^2 + \rho^2\omega^2). \quad (2.24)$$

Calcul de l'énergie potentielle Il n'y a pas d'énergie potentielle dans ce système !

$$U = 0. \quad (2.25)$$

Remarque : on pourrait mettre $U = \text{Cste}$, ça ne changerait rien.

Écriture du lagrangien

$$\mathcal{L} = T - U = \frac{1}{2}m(\dot{\rho}^2 + \rho^2\omega^2). \quad (2.26)$$

Écriture des équations du mouvement Il n'y a qu'un degré de liberté (une seule coordonnée généralisée), et donc il n'y a qu'une équation à écrire. On a

$$\frac{\partial \mathcal{L}}{\partial \rho} = m\omega^2\rho, \quad \frac{\partial \mathcal{L}}{\partial \dot{\rho}} = m\dot{\rho}, \quad \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{\rho}}\right) = m\ddot{\rho}. \quad (2.27)$$

On obtient donc l'équation du mouvement à partir de l'équation de Lagrange (2.6) ; après simplification par m :

$$\ddot{\rho} = \omega^2\rho \quad (2.28)$$

Remarque : la tige exerce une force sur la perle, pour l'entraîner avec elle. On n'a pas besoin de se préoccuper de cette force qui n'est là que pour maintenir la contrainte.

2.2.4 Mobile sur un plan incliné

Un plan incliné de masse m_1 et faisant un angle θ glisse sans frottement sur une table horizontale. Sur ce plan, un mobile M de masse m_2 glisse sans frottement.

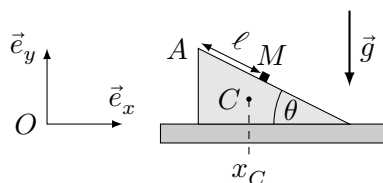


FIGURE 2.3 – Un mobile sur un plan incliné.

Choix des coordonnées généralisées Il faut repérer la position du plan incliné sur la table (une coordonnée), et la position du mobile sur le plan incliné (une coordonnée). On a donc $s = 2$ degrés de liberté. Choisissons par exemple comme première coordonnée $q_1 = x_C$, l'abscisse du centre de masse C du plan incliné, et $q_2 = \ell = AM$, la distance entre le coin supérieur gauche du plan incliné et le mobile M .

Calcul de l'énergie cinétique Dans ce genre de situation où il est un peu difficile d'écrire directement les vitesses, on commence par écrire les vecteurs positions des objets considérés. Soient $\vec{r}_1$ la position du centre de masse du plan incliné, et $\vec{r}_2$ la position du mobile. En posant y_C l'ordonnée du centre de masse C du plan incliné, on a, dans le repère $(O, \vec{e}_x, \vec{e}_y)$,

$$\begin{aligned}\vec{r}_1 &= \overrightarrow{OC} = x_C \vec{e}_x + y_C \vec{e}_y, \\ \vec{r}_2 &= \overrightarrow{OM} = \overrightarrow{OC} + \overrightarrow{CA} + \overrightarrow{AM} = x_C \vec{e}_x + y_C \vec{e}_y + \overrightarrow{CA} + \ell \cos \theta \vec{e}_x - \ell \sin \theta \vec{e}_y.\end{aligned}\quad (2.29)$$

Dans ces expressions, seules x_C et ℓ sont des variables (les coordonnées généralisées). Les quantités y_C , $\overrightarrow{CA}$, θ , $\vec{e}_x$ et $\vec{e}_y$ sont des constantes (des paramètres).

On dérive ces expressions par rapport au temps, afin d'obtenir les *vitesse réelles* $\dot{\vec{r}}_1$ et $\dot{\vec{r}}_2$ en fonction des *vitesse généralisées* $\dot{x}_C$ et $\dot{\ell}$.

$$\dot{\vec{r}}_1 = \dot{x}_C \vec{e}_x, \quad \dot{\vec{r}}_2 = \dot{x}_C \vec{e}_x + \dot{\ell}(\cos \theta \vec{e}_x - \sin \theta \vec{e}_y). \quad (2.30)$$

L'énergie cinétique est alors

$$T = \frac{1}{2}m_1(\dot{\vec{r}}_1)^2 + \frac{1}{2}m_2(\dot{\vec{r}}_2)^2 = \frac{1}{2}m_1\dot{x}_C^2 + \frac{1}{2}m_2(\dot{x}_C^2 + \dot{\ell}^2 + 2\dot{x}_C\dot{\ell}\cos\theta). \quad (2.31)$$

Calcul de l'énergie potentielle La seule force qui travaille est le poids et, donc, la seule énergie potentielle à prendre en compte est l'énergie potentielle de pesanteur. (Comme d'habitude, on peut ignorer la réaction de la table sur le plan incliné, la réaction du plan incliné sur le mobile, la réaction du mobile sur le plan incliné, etc. : toutes ces forces ne sont là que pour garantir les contraintes.)

Le plan incliné se déplace à altitude constante; son énergie potentielle de pesanteur est donc constante et n'intervient pas. L'altitude du mobile par rapport au point A (par exemple) est $-\ell \sin \theta$; l'énergie potentielle du système est donc

$$U = -m_2 g \ell \sin \theta. \quad (2.32)$$

Écriture du lagrangien

$$\mathcal{L} = T - U = \frac{1}{2}(m_1 + m_2)\dot{x}_C^2 + \frac{1}{2}m_2\dot{\ell}^2 + m_2\dot{x}_C\dot{\ell}\cos\theta + m_2g\ell\sin\theta. \quad (2.33)$$

Écriture des équations du mouvement Il y a $s = 2$ degrés de liberté, il y a donc deux équations du mouvement à écrire, une pour chaque coordonnée. Commençons par ℓ :

$$\frac{\partial \mathcal{L}}{\partial \ell} = m_2 g \sin \theta, \quad \frac{\partial \mathcal{L}}{\partial \dot{\ell}} = m_2 \dot{\ell} + m_2 \dot{x}_C \cos \theta, \quad \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\ell}} \right) = m_2 \ddot{\ell} + m_2 \ddot{x}_C \cos \theta. \quad (2.34)$$

On obtient la première équation du mouvement à partir de l'équation de Lagrange (2.6); après simplification par m_2 :

$$\ddot{\ell} + \ddot{x}_C \cos \theta = g \sin \theta. \quad (2.35)$$

On passe maintenant à la deuxième coordonnée :

$$\frac{\partial \mathcal{L}}{\partial x_C} = 0, \quad \frac{\partial \mathcal{L}}{\partial \dot{x}_C} = (m_1 + m_2)\dot{x}_C + m_2\dot{\ell}\cos\theta. \quad (2.36)$$

L'équation de Lagrange donne alors la seconde équation du mouvement :

$$\frac{d}{dt} \left[(m_1 + m_2) \dot{x}_C + m_2 \dot{\ell} \cos \theta \right] = 0. \quad (2.37)$$

Remarquer que, puisque $\partial \mathcal{L} / \partial x_C = 0$, on a préféré ne pas calculer la dérivée par rapport au temps, mais de la laisser explicitement dans l'équation du mouvement (2.37). Sous cette forme, elle s'intègre directement sous la forme

$$(m_1 + m_2) \dot{x}_C + m_2 \dot{\ell} \cos \theta = \text{Cste}. \quad (2.38)$$

2.2.5 Le pendule au point d'attache mobile

Considérons un pendule, de masse m et de longueur ℓ , dont le point d'attache bouge horizontalement selon une loi $X(t)$ imposée, comme illustré en figure 2.4.

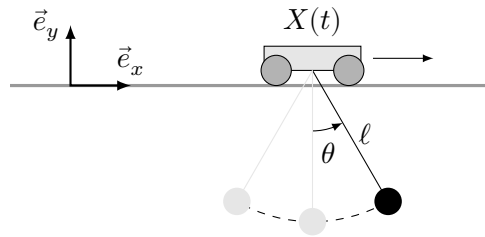


FIGURE 2.4 – Pendule dont le point d'attache bouge horizontalement selon une loi $X(t)$ imposée.

Choix des coordonnées généralisées Puisque la position $X(t)$ du point d'attache est une fonction connue du temps, l'angle θ suffit à déterminer complètement la position du pendule. Elle sera notre coordonnée lagrangienne.

Calcul de l'énergie cinétique Exprimons d'abord la position et la vitesse de la masse m en fonction de θ :

$$\begin{aligned} x &= X(t) + \ell \sin \theta, & \dot{x} &= \dot{X}(t) + \ell \dot{\theta} \cos \theta, \\ y &= -\ell \cos \theta, & \dot{y} &= \ell \dot{\theta} \sin \theta. \end{aligned} \quad (2.39)$$

On a donc pour l'énergie cinétique

$$T = \frac{1}{2} m (\dot{x}^2 + \dot{y}^2) = \frac{1}{2} m (\dot{X}^2 + 2\ell \dot{X} \dot{\theta} \cos \theta + \ell^2 \dot{\theta}^2) \quad (2.40)$$

Calcul de l'énergie potentielle On trouve la même expression (2.14) que pour le pendule simple :

$$U = mgy = -mg\ell \cos \theta \quad (2.41)$$

Écriture du lagrangien

$$\mathcal{L} = T - U = \frac{1}{2} m (\dot{X}^2 + 2\ell \dot{X} \dot{\theta} \cos \theta + \ell^2 \dot{\theta}^2) + mg\ell \cos \theta. \quad (2.42)$$

Écriture des équations du mouvement On a :

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \theta} &= -m\ell \dot{X} \dot{\theta} \sin \theta - mg\ell \sin \theta, \\ \frac{\partial \mathcal{L}}{\partial \dot{\theta}} &= m\ell \dot{X} \cos \theta + m\ell^2 \dot{\theta}, & \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\theta}} \right) &= m\ell (\ddot{X} \cos \theta - \dot{X} \dot{\theta} \sin \theta) + m\ell^2 \ddot{\theta}. \end{aligned} \quad (2.43)$$

On obtient donc l'équation du mouvement à partir de l'équation de Lagrange (2.6) ; le terme en $-m\ell\ddot{X}\dot{\theta}\sin\theta$ se simplifie et il reste, après division par $m\ell^2$,

$$\ddot{\theta} = -\frac{g}{\ell}\sin\theta - \frac{\ddot{X}}{\ell}\cos\theta \quad (2.44)$$

Remarquons que le $\dot{X}^2$ dans (2.42) n'a aucune influence dans la suite du calcul : on pourrait l'enlever du lagrangien sans changer les équations du mouvement. Plus de détails sur ce genre d'opérations seront donnés dans la section 2.3.3.

2.2.6 Le double pendule pesant

On considère le pendule double schématisé figure 2.5, où deux points matériels M_1 et M_2 de même masse m au bout de deux fils de même longueur ℓ oscillent dans le plan vertical xy sous l'action de la pesanteur $\vec{g} = -g\vec{e}_y$, le premier fil étant attaché à un point fixe O et le deuxième fil étant attaché à la première masse.

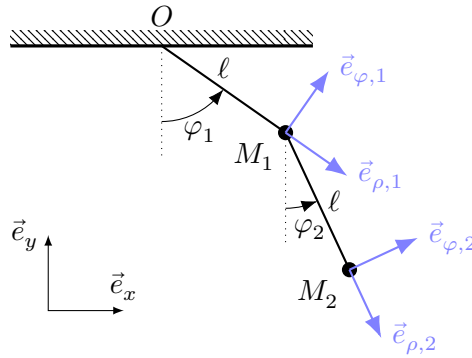


FIGURE 2.5 – Le double pendule pesant.

Choix des coordonnées généralisées Pour décrire entièrement la position du système, il suffit de connaître les angles φ_1 et φ_2 que font les deux fils par rapport à la verticale.

$$s = 2, \quad q_1 = \varphi_1, \quad q_2 = \varphi_2. \quad (2.45)$$

De manière assez naturelle, on travaille en coordonnées polaires, mais on a besoin de deux jeux de vecteurs unitaires, l'un associé à φ_1 et l'autre associé à φ_2 . (Voir la figure.) Les positions des deux masses sont alors

$$\overrightarrow{OM_1} = \ell \vec{e}_{\rho,1}, \quad \overrightarrow{OM_2} = \overrightarrow{OM_1} + \overrightarrow{M_1M_2} = \ell \vec{e}_{\rho,1} + \ell \vec{e}_{\rho,2}. \quad (2.46)$$

Calcul de l'énergie cinétique Les vecteurs vitesse $\vec{v}_1$ et $\vec{v}_2$ des deux masses s'obtiennent en dérivant les positions (2.46). En utilisant les propriétés des coordonnées polaires, on obtient

$$\vec{v}_1 = \ell \dot{\varphi}_1 \vec{e}_{\varphi,1}, \quad \vec{v}_2 = \ell \dot{\varphi}_1 \vec{e}_{\varphi,1} + \ell \dot{\varphi}_2 \vec{e}_{\varphi,2}. \quad (2.47)$$

Les normes des vitesses au carré s'obtiennent en prenant le produit scalaire de chaque vitesse avec elle-même. On obtient

$$v_1^2 = \ell^2 \dot{\varphi}_1^2, \quad v_2^2 = \ell^2 \left(\dot{\varphi}_1^2 + \dot{\varphi}_2^2 + 2\dot{\varphi}_1 \dot{\varphi}_2 \vec{e}_{\varphi,1} \cdot \vec{e}_{\varphi,2} \right). \quad (2.48)$$

Par ailleurs, on a

$$\vec{e}_{\varphi,1} \cdot \vec{e}_{\varphi,2} = \cos(\vec{e}_{\varphi,1}, \vec{e}_{\varphi,2}) = \cos(\vec{e}_{\rho,1}, \vec{e}_{\rho,2}) = \cos(\varphi_2 - \varphi_1). \quad (2.49)$$

(La notation $(\vec{u}, \vec{v})$ signifie ici “angle entre les vecteurs $\vec{u}$ et $\vec{v}$.” La première égalité vient de la définition du produit scalaire et du fait que les vecteurs $\vec{e}_{\varphi,1}$ et $\vec{e}_{\varphi,2}$ sont unitaires.)

On obtient finalement pour l'énergie cinétique

$$T = \frac{1}{2}mv_1^2 + \frac{1}{2}mv_2^2 = \frac{1}{2}m\ell^2(2\dot{\varphi}_1^2 + \dot{\varphi}_2^2 + 2\cos(\varphi_2 - \varphi_1)\dot{\varphi}_1\dot{\varphi}_2). \quad (2.50)$$

(Remarque : on peut aussi obtenir ce résultat en écrivant les coordonnées cartésiennes des deux points, $x_1 = \ell \sin \varphi_1$, $y_1 = -\ell \cos \varphi_1$ et $x_2 = \ell \sin \varphi_1 + \ell \sin \varphi_2$, $y_2 = -\ell \cos \varphi_1 - \ell \cos \varphi_2$, puis l'énergie cinétique sous la forme $T = \frac{1}{2}m(\dot{x}_1^2 + \dot{y}_1^2) + \frac{1}{2}m(\dot{x}_2^2 + \dot{y}_2^2)$. Exercice.)

Calcul de l'énergie potentielle En prenant l'origine des altitudes en O , l'altitude du point M_1 est $y_1 = -\ell \cos \varphi_1$, et celle du point M_2 est $y_2 = -\ell(\cos \varphi_1 + \cos \varphi_2)$. L'énergie potentielle est alors

$$U = mgy_1 + mgy_2 = -mg\ell(2\cos \varphi_1 + \cos \varphi_2). \quad (2.51)$$

Écriture du lagrangien

$$\mathcal{L} = T - U = \frac{1}{2}m\ell^2(2\dot{\varphi}_1^2 + \dot{\varphi}_2^2 + 2\cos(\varphi_2 - \varphi_1)\dot{\varphi}_1\dot{\varphi}_2) + mg\ell(2\cos \varphi_1 + \cos \varphi_2). \quad (2.52)$$

On va simplifier le problème en se mettant dans la limite de petites oscillations : $\varphi_1 \ll 1$, $\varphi_2 \ll 1$, et on développe : $\cos \varphi_i \simeq 1 - \frac{1}{2}\varphi_i^2$. Comme $\cos(\varphi_2 - \varphi_1)$ multiplie un terme $\dot{\varphi}_1\dot{\varphi}_2$ qui est déjà petit, on peut se contenter d'écrire $\cos(\varphi_2 - \varphi_1) \simeq 1$. On arrive alors à

$$\mathcal{L} = \frac{1}{2}m\ell^2(2\dot{\varphi}_1^2 + \dot{\varphi}_2^2 + 2\dot{\varphi}_1\dot{\varphi}_2) - \frac{1}{2}mg\ell(2\varphi_1^2 + \varphi_2^2). \quad (2.53)$$

Remarque : le développement des cos dans (2.52) donne une constante $3mg\ell$ que l'on n'a pas écrite dans (2.53) : cette constante n'influence pas les équations du mouvement, et on peut l'omettre.

Écriture des équations du mouvement Il y a deux degrés de liberté, et donc deux équations à écrire. On a, pour la première coordonnée φ_1 ,

$$\frac{\partial \mathcal{L}}{\partial \varphi_1} = -mg\ell \times 2\varphi_1, \quad \frac{\partial \mathcal{L}}{\partial \dot{\varphi}_1} = m\ell^2(2\dot{\varphi}_1 + \dot{\varphi}_2), \quad \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\varphi}_1} \right) = m\ell^2(2\ddot{\varphi}_1 + \ddot{\varphi}_2). \quad (2.54)$$

L'équation de Lagrange donne alors la première équation du mouvement ; après simplification par $m\ell^2$:

$$2\ddot{\varphi}_1 + \ddot{\varphi}_2 = -\omega_0^2 \times 2\varphi_1, \quad \text{avec} \quad \omega_0^2 = \frac{g}{\ell}. \quad (2.55)$$

De même, partant de l'équation de Lagrange pour φ_2 , on obtient :

$$\ddot{\varphi}_1 + \ddot{\varphi}_2 = -\omega_0^2 \times \varphi_2. \quad (2.56)$$

Remarques et résolution des équations Les équations (2.55) et (2.56) décrivent un *oscillateur harmonique à deux degrés de liberté* : des combinaisons linéaires des dérivées secondes sont égales à des combinaisons linéaires des variables. Le lagrangien d'un oscillateur harmonique s'écrit comme une *forme quadratique* des vitesses généralisées (c'est-à-dire une combinaison linéaire de termes de la forme $\dot{q}_i^2$ ou $\dot{q}_i\dot{q}_j$ moins une forme quadratique des positions (des termes de la forme q_i^2 ou q_iq_j). Si on définit le vecteur V de toutes les coordonnées, ici

$$V = \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix}, \quad (2.57)$$

on peut écrire les équations du mouvement d'un oscillateur harmonique avec deux matrices *symétriques* G et H sous la forme suivante

$$G\ddot{V} = -HV. \quad (2.58)$$

Dans l'exemple du double pendule, ces matrices sont données par

$$G = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \quad H = \omega_0^2 \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}. \quad (2.59)$$

En effet la première ligne de l'égalité vectorielle (2.58) donne (2.55), et la deuxième ligne (2.56). On remarque aussi que le lagrangien peut s'écrire sous la forme

$$L = \frac{1}{2} \dot{V} \cdot G \dot{V} - \frac{1}{2} V \cdot H V. \quad (2.60)$$

Pour résoudre un tel système, on cherche des *modes propres* ou *modes normaux*, c'est-à-dire des solutions qui oscillent comme un oscillateur simple, avec une seule pulsation et une seule phase, mais avec des amplitudes différentes pour les deux coordonnées :

$$V = V(t) = \begin{pmatrix} \varphi_1(t) \\ \varphi_2(t) \end{pmatrix} = \begin{pmatrix} a_1 \cos(\omega t + \phi) \\ a_2 \cos(\omega t + \phi) \end{pmatrix}. \quad (2.61)$$

On voit facilement que, pour un tel mode propre de pulsation ω ,

$$\ddot{V} = -\omega^2 V. \quad (2.62)$$

Les équations vectorielles du mouvement (2.58) deviennent alors

$$(\omega^2 G - H)V = 0. \quad (2.63)$$

En posant

$$\omega^2 = \lambda \omega_0^2, \quad (2.64)$$

et en utilisant les expressions (2.59) des matrices G et H , on a dans le cas qui nous intéresse

$$\omega^2 G - H = \omega_0^2 \begin{pmatrix} 2\lambda - 2 & \lambda \\ \lambda & \lambda - 1 \end{pmatrix}. \quad (2.65)$$

On sait qu'une équation de la forme $MV = 0$ (avec M une matrice donnée et V un vecteur à déterminer) n'a de solution non-nulle que si $\det(M) = 0$. En simplifiant par les ω_0 , on voit donc que λ doit vérifier l'équation

$$0 = (2\lambda - 2)(\lambda - 1) - \lambda^2 = \lambda^2 - 4\lambda + 2, \quad (2.66)$$

qui a deux solutions :

$$\lambda_+ = 2 + \sqrt{2}, \quad \lambda_- = 2 - \sqrt{2}, \quad (2.67)$$

correspondant à deux pulsations possibles, voir (2.64) :

$$\omega_+ = \sqrt{2 + \sqrt{2}} \omega_0, \quad \omega_- = \sqrt{2 - \sqrt{2}} \omega_0. \quad (2.68)$$

On voit donc qu'il n'y a que deux modes propres possibles, et on a pu calculer leurs pulsations. On dit que ω_+ et ω_- sont les *pulsations propres* du système.

À chacune de ces pulsations est associé un vecteur propre U tel que $(\omega^2 G - H)U = 0$. Un calcul rapide donne

$$U_+ = \begin{pmatrix} 1 \\ -\sqrt{2} \end{pmatrix}, \quad U_- = \begin{pmatrix} 1 \\ \sqrt{2} \end{pmatrix}. \quad (2.69)$$

Finalement, une expression de la forme (2.61) est solution si et seulement si ω est ω_+ ou ω_- , et si V est proportionnel à U_+ ou, respectivement, U_- :

$$V_+(t) = a_+ \begin{pmatrix} \cos(\omega_+ t + \phi_+) \\ -\sqrt{2} \cos(\omega_+ t + \phi_+) \end{pmatrix}, \quad V_-(t) = a_- \begin{pmatrix} \cos(\omega_- t + \phi_-) \\ \sqrt{2} \cos(\omega_- t + \phi_-) \end{pmatrix}, \quad (2.70)$$

avec a_+ , a_- , ϕ_+ , ϕ_- des constantes arbitraires. Ce sont les deux modes propres, illustrés figure 2.6. Pour le premier mode propre, la pulsation ω_+ est élevée et les deux angles φ_1 et φ_2 ont des signes opposés. Le deuxième mode propre a une pulsation ω_- plus faible et des angles de même signe.

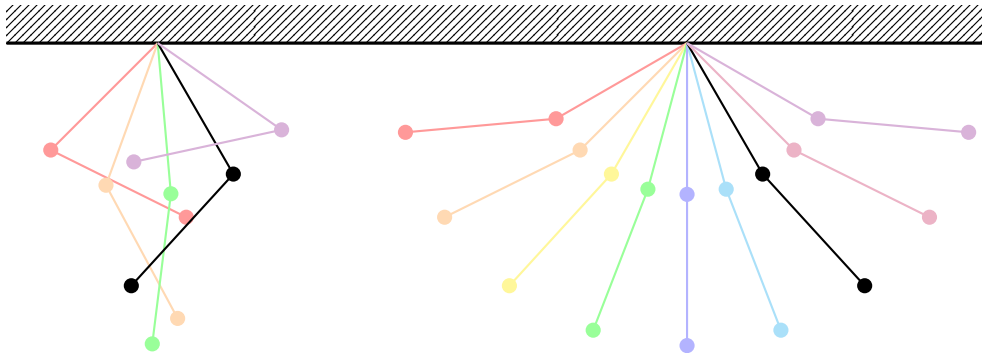


FIGURE 2.6 – Modes propres du pendule double. À gauche, $V_+(t)$, et à droite, $V_-(t)$

Les modes propres sont des solutions particulières ; la solution la plus générale est une combinaison linéaire des modes propres

$$V(t) = V_+(t) + V_-(t) \iff \begin{cases} \varphi_1(t) = a_+ \cos(\omega_+ t + \phi_+) + a_- \cos(\omega_- t + \phi_-), \\ \varphi_2(t) = -a_+ \sqrt{2} \cos(\omega_+ t + \phi_+) + a_- \sqrt{2} \cos(\omega_- t + \phi_-). \end{cases} \quad (2.71)$$

Il reste quatre paramètres (a_+ , a_- , ϕ_+ , ϕ_-) à ajuster à partir des quatre conditions initiales $\varphi_1(0)$, $\dot{\varphi}_1(0)$, $\varphi_2(0)$ et $\dot{\varphi}_2(0)$. Si a_+ et a_- sont tous les deux non-nuls (c'est-à-dire : si on n'est pas dans un mode propre), alors le mouvement n'est pas périodique puisque ω_+/ω_- n'est pas un nombre rationnel.

2.3 Quelques propriétés du lagrangien

2.3.1 Dépendance de T et U en les coordonnées et vitesses généralisées

On considère un système de N points matériels avec s degrés de liberté. Cela signifie que l'on peut déterminer les N vecteurs positions $\vec{r}_1, \dots, \vec{r}_N$ à l'aide des s coordonnées généralisées $q_1, \dots, q_s$, et du temps t :

$$\vec{r}_\alpha = \vec{r}_\alpha(q_1, \dots, q_s, t), \quad \text{pour } \alpha = 1, 2, \dots, N. \quad (2.72)$$

(Voir (2.23) pour un exemple où $\vec{r}$ dépend explicitement du temps t .)

Comme l'énergie potentielle U est une fonction des $\vec{r}_\alpha$ et du temps, on voit que l'on peut écrire cette énergie potentielle comme une fonction des q_i et du temps :

$$U = U(q_1, \dots, q_s, t). \quad (2.73)$$

Passons maintenant à l'énergie cinétique. Si on dérive par rapport au temps (2.72), on obtient les vitesses réelles $\dot{\vec{r}}_\alpha$ des particules. On n'oublie pas que les q_i dépendent du temps ; on a donc

une dérivée de fonctions composées :

$$\dot{\vec{r}}_\alpha = \frac{d\vec{r}_\alpha}{dt} = \left(\sum_{i=1}^s \frac{\partial \vec{r}_\alpha}{\partial q_i} \dot{q}_i \right) + \frac{\partial \vec{r}_\alpha}{\partial t}. \quad (2.74)$$

On rappelle que les quantités $\frac{\partial \vec{r}_\alpha}{\partial q_i}$ pour $i = 1, \dots, s$ et $\frac{\partial \vec{r}_\alpha}{\partial t}$ sont les $s + 1$ *dérivées partielles* de la fonction $\vec{r}_\alpha$. On rappelle que $\vec{r}_\alpha$ est une fonction des $s + 1$ variables $q_1, \dots, q_s, t$, et qu'il y a une dérivée partielle par variable, obtenue en dérivant par rapport à cette variable tout en gardant les autres variables constantes. Ces dérivées partielles sont encore des fonctions des q_i et de t .

La quantité $\frac{d\vec{r}_\alpha}{dt}$ est la *dérivée totale* par rapport au temps de la fonction $\vec{r}_\alpha$ prise le long de la trajectoire. Il ne faut pas confondre $\frac{d\vec{r}_\alpha}{dt}$ et $\frac{\partial \vec{r}_\alpha}{\partial t}$:

- Lors du mouvement, entre t et $t + dt$, la position $\vec{r}_\alpha$ augmente de $d\vec{r}_\alpha = \frac{d\vec{r}_\alpha}{dt} dt$ parce que les $s + 1$ variables de la fonction $\vec{r}_\alpha$ changent : t augmente de dt (évidemment) et les q_i augmentent de $dq_i = \frac{\partial q_i}{\partial t} dt$.
- Si on suppose que t augmente de dt sans que les q_i changent (ce qui n'est pas une situation physique), alors $\vec{r}_\alpha$ augmente de $\frac{\partial \vec{r}_\alpha}{\partial t} dt$.

Formellement, on peut interpréter (2.74) en disant que $\dot{\vec{r}}_\alpha$ est une fonction des q_i , des $\dot{q}_i$ et du temps t . En écrivant explicitement toutes les variables :

$$\dot{\vec{r}}_\alpha = \dot{\vec{r}}_\alpha(q_1, \dots, q_s, \dot{q}_1, \dots, \dot{q}_s, t) = \left(\sum_{i=1}^s \frac{\partial \vec{r}_\alpha}{\partial q_i}(q_1, \dots, q_s, t) \dot{q}_i \right) + \frac{\partial \vec{r}_\alpha}{\partial t}(q_1, \dots, q_s, t). \quad (2.75)$$

Remarquons que $\dot{\vec{r}}_\alpha$ est une fonction *affine* des vitesses généralisées, et que (2.75) implique

$$\frac{\partial \dot{\vec{r}}_\alpha}{\partial \dot{q}_i} = \frac{\partial \vec{r}_\alpha}{\partial q_i}. \quad (2.76)$$

(On a dérivé $\dot{\vec{r}}_\alpha$ par rapport à $\dot{q}_i$, en supposant toutes les positions q_j fixées, les autres vitesses $\dot{q}_j$ fixées et le temps t fixé.) Ce résultat servira plus tard.

Remarquons que si les $\vec{r}_\alpha$ ne dépendent pas explicitement du temps, c'est-à-dire si $\frac{\partial \vec{r}_\alpha}{\partial t} = 0$, alors la vitesse est une fonction *linéaire* des $\dot{q}_i$ et son carré une fonction *quadratique* des $\dot{q}_i$:

$$\dot{\vec{r}}_\alpha^2 = \sum_{i=1}^s \sum_{j=1}^s \frac{\partial \vec{r}_\alpha}{\partial q_i} \cdot \frac{\partial \vec{r}_\alpha}{\partial q_j} \dot{q}_i \dot{q}_j. \quad (2.77)$$

En particulier, l'énergie cinétique $T = \frac{1}{2} \sum_\alpha m_\alpha \dot{\vec{r}}_\alpha^2$ est une fonction quadratique des $\dot{q}_i$ et peut s'écrire :

$$T = \frac{1}{2} \sum_{\alpha=1}^N m_\alpha (\dot{x}_\alpha^2 + \dot{y}_\alpha^2 + \dot{z}_\alpha^2) = \frac{1}{2} \sum_{i,k=1}^s a_{ik}(q) \dot{q}_i \dot{q}_k \quad (2.78)$$

avec

$$a_{ik} = \sum_{\alpha=1}^N m_\alpha \left(\frac{\partial x_\alpha}{\partial q_i} \frac{\partial x_\alpha}{\partial q_k} + \frac{\partial y_\alpha}{\partial q_i} \frac{\partial y_\alpha}{\partial q_k} + \frac{\partial z_\alpha}{\partial q_i} \frac{\partial z_\alpha}{\partial q_k} \right)$$

Remarquons que les coefficients a_{ik} sont symétriques.

En résumé, on retiendra que :

- U dépend des q_i et de t ,
- T dépend des q_i , des $\dot{q}_i$ et de t ,
- Si les relations entre les q_i et les $\vec{r}_\alpha$ ne fait pas intervenir t , alors T est une fonction quadratique des $\dot{q}_i$.

2.3.2 Additivité du lagrangien

Le lagrangien est une grandeur *additive*. Cela veut dire que pour un système composé de deux parties A et B , décrites séparément par deux fonctions de Lagrange $\mathcal{L}_A$ et $\mathcal{L}_B$, en l'absence d'interaction entre les deux parties, si par exemple on les éloigne suffisamment pour rendre leur interaction négligeable, la fonction de Lagrange du système total tend vers la somme : $\mathcal{L}_A + \mathcal{L}_B$

$$\mathcal{L}_{A+B} \xrightarrow{|r_A - r_B| \rightarrow \infty} \mathcal{L}_A + \mathcal{L}_B. \quad (2.79)$$

Ceci implique que le système A sera décrit par des variables qui sont indépendantes de celles du système B duquel il serait infiniment éloigné. On peut remarquer que la situation n'est pas la même en mécanique quantique, où pour deux systèmes intriqués même infiniment éloignés, on ne peut pas décrire les deux systèmes localement, de façon indépendante.

2.3.3 Invariance de jauge

À partir de la forme des équations d'Euler-Lagrange (2.6), on voit que si on multiplie le lagrangien par une constante, ou si l'on y ajoute une constante, on ne change pas les équations du mouvement.

Une propriété supplémentaire un peu moins triviale, dite invariance de jauge, est la suivante. On ne change pas les équations du mouvement si l'on rajoute au lagrangien une quantité Δ qui peut s'écrire comme la dérivée totale par rapport au temps d'une fonction quelconque $F(q, t)$ des coordonnées et du temps (mais pas des vitesses)

$$\mathcal{L}(q, \dot{q}, t) \text{ est équivalent à } \mathcal{L}(q, \dot{q}, t) + \Delta(q, \dot{q}, t) \quad \text{si } \Delta(q, \dot{q}, t) = \frac{d}{dt} [F(q, t)]. \quad (2.80)$$

(On a écrit q au lieu de $q_1, \dots, q_s$ et $\dot{q}$ au lieu de $\dot{q}_1, \dots, \dot{q}_s$ pour alléger les équations.)

En particulier, on peut toujours ajouter (ou enlever) une fonction $\Delta(t)$ (qui ne dépend que du temps) à un lagrangien sans changer les équations du mouvement. En effet, on peut toujours écrire $\Delta(t) = \frac{dF(t)}{dt}$ où $F(t)$ est une primitive de Δ , et on est dans le cas (2.80). Par exemple, dans le lagrangien (2.42) du pendule accroché à un chariot, on a déjà vu que l'on pouvait enlever le terme $\frac{1}{2}m\dot{X}^2$ du lagrangien sans changer les équations du mouvement.

La propriété d'invariance de jauge est une conséquence immédiate du principe de moindre action que nous verrons section 2.4. On pourrait la démontrer directement, mais le calcul est un peu compliqué¹.

1. Pour montrer la propriété (2.80), nous vérifions que les contributions du terme supplémentaire $\Delta(q, \dot{q}, t)$ au membre de gauche et au membre de droite des équations de Lagrange (2.6) sont identiques. Commençons par expliciter $\Delta(q, \dot{q}, t)$:

$$\Delta(q, \dot{q}, t) = \frac{d}{dt} [F(q_1, \dots, q_s, t)] = \frac{\partial F}{\partial t} + \sum_j \frac{\partial F}{\partial q_j} \dot{q}_j. \quad (2.81)$$

À droite de l'équation (2.6) pour la coordonnée q_i nous avons :

$$\frac{\partial \Delta}{\partial q_i} = \frac{\partial^2 F}{\partial q_i \partial t} + \sum_j \frac{\partial^2 F}{\partial q_i \partial q_j} \dot{q}_j. \quad (2.82)$$

À gauche, en remarquant avec (2.81) que $\frac{\partial \Delta}{\partial \dot{q}_i} = \frac{\partial F}{\partial q_i}$:

$$\frac{d}{dt} \left[\frac{\partial \Delta}{\partial \dot{q}_i} \right] = \frac{d}{dt} \left[\frac{\partial F}{\partial q_i} \right] = \frac{\partial^2 F}{\partial t \partial q_i} + \sum_j \frac{\partial^2 F}{\partial q_j \partial q_i} \dot{q}_j, \quad (2.83)$$

ce qui est la même chose que (2.82). Les contributions de Δ se simplifient donc, et (2.80) est démontré.

2.3.4 Impulsions généralisées et lois de conservation

Si un lagrangien $\mathcal{L}(q, \dot{q}, t)$ ne dépend pas d'une des coordonnées généralisées q_i , alors $\frac{\partial \mathcal{L}}{\partial q_i} = 0$ et, d'après les équations de Lagrange (2.6), on obtient

$$\frac{d}{dt} \left[\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right] = 0. \quad (2.84)$$

Autrement dit, la quantité

$$p_i := \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \quad (2.85)$$

est une quantité conservée. On appelle p_i l'impulsion généralisée associée à la coordonnée q_i .

Considérons quelques exemples simples.

- On considère une particule évoluant dans un potentiel qui ne dépend que de l'altitude z (par exemple, le potentiel associé à la pesanteur). On choisit comme coordonnées généralisées les coordonnées cartésiennes x , y et z . Le lagrangien est alors

$$\mathcal{L} = \frac{1}{2}m[\dot{x}^2 + \dot{y}^2 + \dot{z}^2] - U(z). \quad (2.86)$$

Comme $\mathcal{L}$ ne dépend ni de x , ni de y , on voit que les impulsions généralisées p_x et p_y sont conservées, avec

$$p_x = \frac{\partial \mathcal{L}}{\partial \dot{x}} = m\dot{x}, \quad p_y = \frac{\partial \mathcal{L}}{\partial \dot{y}} = m\dot{y}. \quad (2.87)$$

Dans ce cas, ces deux impulsions généralisées correspondent aux projections du vecteur impulsion habituel $\vec{p} = m\vec{v}$ selon les axes $\vec{e}_x$ et $\vec{e}_y$. Ces quantités sont bien conservées, puisque la force $-\vec{\text{grad}} U$ subie par la particule est dirigée selon $\vec{e}_z$.

- Dans l'exemple 2.2.4 du mobile sur un plan incliné, le lagrangien obtenu était indépendant de la coordonnée x_C et on avait déjà remarqué que l'impulsion généralisée associée, donnée par (2.38), était conservée. En fait, cette quantité est égale à la projection de l'impulsion totale du système selon l'axe $\vec{e}_x$.
- Considérons maintenant le cas d'une particule contrainte à vivre dans un plan et que l'on repère par ses coordonnées polaires (r, θ) . On suppose que cette particule est soumise à un potentiel central, c'est-à-dire qui ne dépend que de r . On rappelle que la vitesse en polaire est $\vec{v} = \dot{r}\vec{e}_r + r\dot{\theta}\vec{e}_\theta$. Le lagrangien est donc

$$\mathcal{L} = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - U(r). \quad (2.88)$$

Comme $\mathcal{L}$ ne dépend pas de θ , l'impulsion généralisée p_θ associée est conservée :

$$p_\theta := \frac{\partial \mathcal{L}}{\partial \dot{\theta}} = mr^2\dot{\theta} = \text{Cste}. \quad (2.89)$$

Cette quantité est en fait le moment cinétique de la particule selon l'axe $\vec{e}_z$, voir (1.12).

On peut généraliser ce type de raisonnement de la manière suivante : considérons un lagrangien écrit en coordonnées cartésiennes : $\mathcal{L} = \mathcal{L}(\dot{x}_1, \dot{y}_1, \dot{z}_1, \dot{x}_2, \dots, \dot{z}_N, x_1, y_1, z_1, x_2, \dots, z_N)$. Le lagrangien dépend effectivement de tous les x_i , et donc les impulsions généralisées $p_{x,i} = \partial \mathcal{L} / \partial \dot{x}_i$ ne sont pas conservées. On a quand même le résultat suivant :

Si le lagrangien est invariant par translation selon l'axe des x (c'est-à-dire si $\mathcal{L}$ est inchangé si on remplace tous les x_i par $x_i + \varepsilon$), alors l'impulsion totale selon x est conservée.

(Et idem pour les axes y et z .) En effet, pour ε petit, on a

$$\mathcal{L}(\dots, x_i + \varepsilon, \dots) = \mathcal{L}(\dots, x_i, \dots) + \sum_i \varepsilon \frac{\partial \mathcal{L}}{\partial x_i}. \quad (2.90)$$

L'hypothèse faite implique alors que $\sum_i \frac{\partial \mathcal{L}}{\partial x_i} = 0$. Mais les équations de lagrange donnent $\frac{\partial \mathcal{L}}{\partial x_i} = \frac{d}{dt} p_{i,x}$, et donc

$$\frac{d}{dt} \sum_i p_{i,x} = 0. \quad (2.91)$$

L'impulsion totale selon x est conservée, comme annoncé. On montre de même que si le système est invariant par rotation autour d'un axe, alors le moment cinétique total projeté sur cet axe est conservé.

2.3.5 Conservation de l'énergie

Un des axiomes fondamentaux de la mécanique est que *le temps est uniforme*. Cela signifie que l'évolution d'un système isolé ne dépend pas de la date à laquelle le système a été mis en place. Autrement dit, le temps t ne peut pas apparaître explicitement dans les équations du mouvement. Pour étudier ce genre de systèmes, on suppose dans cette section que le lagrangien ne dépend pas explicitement du temps :

$$\mathcal{L} = \mathcal{L}(q, \dot{q}),$$

et non $\mathcal{L} = \mathcal{L}(q, \dot{q}, t)$. Cette propriété est automatique pour les systèmes isolés, mais elle est aussi réalisée dans d'autres situations : par exemple, pour un système soumis à une force extérieure indépendante du temps, comme le champ de pesanteur.

$\mathcal{L}$ ne dépend pas explicitement du temps, et donc la dérivée partielle $\frac{\partial \mathcal{L}}{\partial t}$ est nulle. La dérivée totale $\frac{d\mathcal{L}}{dt}$ n'est cependant pas nulle, car $\mathcal{L}$ dépend de q et $\dot{q}$ qui dépendent du temps. On a :

$$\frac{d\mathcal{L}}{dt} = \sum_{i=1}^s \left(\frac{\partial \mathcal{L}}{\partial q_i} \dot{q}_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \ddot{q}_i \right). \quad (2.92)$$

(Si le lagrangien avait dépendu explicitement du temps, il aurait fallu rajouter un terme $\frac{\partial \mathcal{L}}{\partial t}$.)

En utilisant les équations de Lagrange, on peut aussi écrire

$$\frac{d\mathcal{L}}{dt} = \sum_{i=1}^s \left(\dot{q}_i \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) + \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \ddot{q}_i \right) = \sum_{i=1}^s \frac{d}{dt} \left(\dot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) \quad (2.93)$$

d'où l'on obtient

$$\frac{d}{dt} \left(\sum_{i=1}^s \dot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \mathcal{L} \right) = 0. \quad (2.94)$$

On appelle *hamiltonien*, noté H , la quantité entre parenthèses :

$$H = \sum_{i=1}^s \dot{q}_i \frac{\partial \mathcal{L}}{\partial \dot{q}_i} - \mathcal{L} = \sum_{i=1}^s \dot{q}_i p_i - \mathcal{L}. \quad (2.95)$$

(On rappelle que $p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}$ est l'*impulsion généralisée* associée à la variable q_i .) Comme $\frac{dH}{dt} = 0$, la quantité H est une constante du mouvement.

Calculons maintenant H en faisant l'hypothèse supplémentaire que les relations entre les coordonnées généralisées q_i et les vecteurs position $\vec{r}_\alpha$ ne font pas intervenir explicitement le temps : $\vec{r}_\alpha = \vec{r}_\alpha(q_1, \dots, q_s)$, et non $\vec{r}_\alpha = \vec{r}_\alpha(q_1, \dots, q_s, t)$. On a vu vers la fin de la section 2.3.1 que, dans ces conditions, l'énergie cinétique T était une *fonction quadratique des vitesses généralisées*

$\dot{q}_i$, c'est-à-dire une somme de termes de la forme $a_{i,j}(q)\dot{q}_i\dot{q}_j$. En particulier, la fonction T doit vérifier, pour tout λ , l'égalité

$$T(q_1, \dots, q_s, \lambda\dot{q}_1, \dots, \lambda\dot{q}_s) = \lambda^2 T(q_1, \dots, q_s, \dot{q}_1, \dots, \dot{q}_s). \quad (2.96)$$

Si on dérive cette égalité par rapport à λ , puis que l'on pose $\lambda = 1$, on arrive à

$$\sum_{i=1}^s \dot{q}_i \frac{\partial T}{\partial \dot{q}_i} = 2T. \quad (2.97)$$

Comme $\mathcal{L} = T - U$ et que U ne dépend pas des $\dot{q}_i$, on a $\frac{\partial \mathcal{L}}{\partial \dot{q}_i} = \frac{\partial T}{\partial \dot{q}_i}$. On obtient donc, avec (2.97),

$$H = 2T - (T - U) = T + U, \quad (2.98)$$

c'est-à-dire que H est l'énergie mécanique du système.²

2.4 Première justification : le principe de moindre action

On considère un système mécanique sans dissipation, et on se choisit des coordonnées généralisées. On note $q = (q_1, q_2, \dots, q_s)$ l'ensemble de toutes les coordonnées généralisées, et $\dot{q} = (\dot{q}_1, \dot{q}_2, \dots, \dot{q}_s)$ l'ensemble de toutes les vitesses généralisées.

On cherche à répondre à la question suivante :

$$\text{Sachant que le système est à la position } q^{(1)} \text{ au temps } t_1 \text{ et à la position } q^{(2)} \text{ au temps } t_2 > t_1, \text{ quelle est sa trajectoire } q(t) \text{ pour } t \text{ entre } t_1 \text{ et } t_2 ? \quad (2.100)$$

(Remarque : ce n'est pas la question que l'on pose habituellement en mécanique, qui serait plutôt : sachant que le système est à la position $q^{(1)}$ et à la vitesse $\dot{q}^{(1)}$ au temps t_1 , quelle est sa trajectoire $q(t)$ pour $t > t_1$?)

Selon le *principe de moindre action*, la réponse à la question (2.100) est qu'il doit exister une *fonctionnelle* $S[q]$ de la trajectoire, appelée action, telle que $S[q]$ est extrémale (souvent minimale) sur la trajectoire réelle pour tous les choix de q tels que $q(t_1) = q^{(1)}$ et $q(t_2) = q^{(2)}$.

L'action S est une fonctionnelle, c'est-à-dire qu'elle prend comme argument toute la fonction q ; il est habituel de souligner ce point en utilisant des crochets plutôt que des parenthèses pour son argument : $S[q]$ plutôt que $S(q)$. Si q^* est la trajectoire réelle, et que δq est une petite perturbation de cette trajectoire, alors dire que $S[q]$ est extrémale pour q^* revient à dire que $S[q^* + \delta q] - S[q^*]$ est toujours du même signe (positif si S est minimale, par exemple), quel que soit δq si δq est petit. Ceci implique que

$$\delta S = S[q^* + \delta q] - S[q^*] = \mathcal{O}(\delta q^2). \quad (2.101)$$

En effet, si cette différence était d'ordre δq , on pourrait changer le signe de $S[q^* + \delta q] - S[q^*]$ en changeant le signe de δq .

Remarque importante : dans (2.101), il ne faut considérer que les perturbations telles que

$$\delta q(t_1) = \delta q(t_2) = 0, \quad (2.102)$$

2. Remarque : la différentielle de $\mathcal{L}$ est $d\mathcal{L} = \sum_i (\frac{\partial \mathcal{L}}{\partial q_i} dq_i + \frac{\partial \mathcal{L}}{\partial \dot{q}_i} d\dot{q}_i) = \sum_i (\dot{p}_i dq_i + p_i d\dot{q}_i)$. À partir de (2.95), on obtient que la différentielle de H est $dH = \sum_i (\dot{q}_i dp_i + p_i d\dot{q}_i) - d\mathcal{L}$. En remplaçant $d\mathcal{L}$ par son expression, on voit que les termes en $d\dot{q}_i$ se simplifient et on obtient

$$dH = \sum_i (\dot{q}_i dp_i - \dot{p}_i dq_i). \quad (2.99)$$

On voit que si on écrit H en fonction des variables q_i et p_i (plutôt qu'en fonction de q_i et $\dot{q}_i$), alors les équations du mouvement se mettent sous la forme élégante et symétrique suivante : $\dot{q}_i = \frac{\partial H}{\partial p_i}$ et $\dot{p}_i = -\frac{\partial H}{\partial q_i}$. Ce sont les *équations de Hamilton*.

afin que la trajectoire $q^* + \delta q$ aille bien de $q^{(1)}$ à $q^{(2)}$, comme la trajectoire q^* .

Beaucoup de problèmes en physique se résolvent en minimisant ou en maximisant une quantité, et donc c'est une hypothèse raisonnable de postuler qu'il doit exister une action S qui permet de résoudre les problèmes de mécanique par une opération de minimisation. Par exemple, un siècle avant la mécanique analytique formulée par Lagrange, Fermat avait reformulé l'optique géométrique en disant que la trajectoire suivie par un rayon lumineux allant de A à B était celle qui minimisait le *chemin optique*; l'action est une sorte de généralisation de ce principe à la mécanique.

On admet que l'action doive se mettre sous la forme

$$S[q] = \int_{t_1}^{t_2} \mathcal{L}(q(t), \dot{q}(t), t) dt \quad (2.103)$$

pour une certaine fonction (pas fonctionnelle) $\mathcal{L}(q, \dot{q}, t)$ des coordonnées q , des vitesses généralisées $\dot{q}$ et du temps; c'est en effet la forme la plus générale qui n'aboutit pas à des résultats absurdes. Cette fonction $\mathcal{L}$ est bien évidemment le lagrangien.

Établissons maintenant les équations du mouvement à partir de (2.101). Pour simplifier nous considérons un système à un seul degré de liberté, si bien que la position du système à un instant donné t est caractérisée par la valeur d'une seule coordonnée $q(t)$. On a

$$\delta S = S[q^* + \delta q] - S[q^*] = \int_{t_1}^{t_2} \mathcal{L}(q^*(t) + \delta q(t), \dot{q}^*(t) + \delta \dot{q}(t), t) dt - \int_{t_1}^{t_2} \mathcal{L}(q^*(t), \dot{q}^*(t), t) dt \quad (2.104)$$

On a supposé δq petit. En développant au premier ordre, à t fixé, on a

$$\mathcal{L}(q^*(t) + \delta q(t), \dot{q}^*(t) + \delta \dot{q}(t), t) = \mathcal{L}(q^*(t), \dot{q}^*(t), t) + \frac{\partial \mathcal{L}}{\partial q} \times \delta q(t) + \frac{\partial \mathcal{L}}{\partial \dot{q}} \times \delta \dot{q}(t) + \mathcal{O}(\delta q^2). \quad (2.105)$$

Dans cette expression, les deux dérivées sont prises de manière implicite au point $(q^*(t), \dot{q}^*(t), t)$.

En reportant dans (2.104), on obtient

$$\delta S = \int_{t_1}^{t_2} \frac{\partial \mathcal{L}}{\partial q} \times \delta q(t) dt + \int_{t_1}^{t_2} \frac{\partial \mathcal{L}}{\partial \dot{q}} \times \delta \dot{q}(t) dt + \mathcal{O}(\delta q^2). \quad (2.106)$$

Dans cette expression, $\delta \dot{q}$ est la variation de la vitesse et est égale à

$$\delta \dot{q}(t) = \frac{d}{dt} \delta q(t). \quad (2.107)$$

On peut donc intégrer par parties la deuxième intégrale de (2.106)

$$\delta S = \int_{t_1}^{t_2} \frac{\partial \mathcal{L}}{\partial q} \times \delta q(t) dt + \left[\frac{\partial \mathcal{L}}{\partial \dot{q}} \times \delta q(t) \right]_{t_1}^{t_2} - \int_{t_1}^{t_2} \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}} \right) \times \delta q(t) dt + \mathcal{O}(\delta q^2). \quad (2.108)$$

Le terme intégré est nul, car on a supposé $\delta q(t_1) = \delta q(t_2) = 0$, voir (2.102). En regroupant les deux intégrales restantes, on obtient

$$\delta S = \int_{t_1}^{t_2} \left[\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}} \right) \right] \delta q(t) dt + \mathcal{O}(\delta q^2). \quad (2.109)$$

La fonction $\delta q(t)$ est une fonction quelconque, en particulier on pourrait prendre $\delta q(t)$ nulle partout sauf dans une région aussi petite que l'on veut de la trajectoire. Par conséquent (2.109) ne peut conduire à (2.101) que si le terme entre crochets s'annule pour tous les temps t . On voit donc que la trajectoire réelle doit vérifier

$$\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}} \right) = 0, \quad (2.110)$$

ce qui est l'équation d'Euler-Lagrange (2.6) à un seul degré de liberté. On vérifie que le cas général (à plusieurs degrés de liberté) s'obtient de la même manière.

Faisons quelques remarques pour conclure :

- On a retrouvé les équations de Lagrange, mais il reste à justifier que l'on a $\mathcal{L} = T - U$. Cela se fait tout simplement en vérifiant que l'on retrouve bien les équations de Newton avec cette expression.
- L'action d'une trajectoire est donc l'intégrale sur cette trajectoire de la différence entre l'énergie cinétique et l'énergie potentielle. On voit bien que, pour une trajectoire donnée, la valeur de l'action ne dépend pas des coordonnées choisies, même si la forme du lagrangien en dépend. C'est pour cette raison que les équations de Lagrange sont bien indépendantes du système de coordonnées.
- Le principe de moindre action prend facilement en compte les contraintes : on cherche la trajectoire qui est extrémale parmi toutes les trajectoires qui vérifient ces contraintes. Il suffit alors de choisir un système de coordonnées où les contraintes sont automatiquement respectées pour résoudre le problème.
- La propriété d'invariance de jauge (section 2.3.3) est évidente si on utilise l'action : soient deux lagrangiens $\mathcal{L}$ et $\mathcal{L}'$ qui ne diffèrent que par l'ajout de la dérivée totale par rapport au temps d'une fonction $F(q, t)$ des coordonnées et du temps :

$$\mathcal{L}'(q, \dot{q}, t) = \mathcal{L}(q, \dot{q}, t) + \frac{d}{dt} [F(q, t)], \quad (2.111)$$

et posons S et S' les actions correspondant à ces deux lagrangiens. On a alors

$$S'[q] - S[q] = \int_{t_1}^{t_2} \frac{d}{dt} [F(q, t)] dt = [F(q, t)]_{t_1}^{t_2} = F(q^{(2)}, t_2) - F(q^{(1)}, t_1). \quad (2.112)$$

La différence entre les deux actions est constante, donc une trajectoire qui est extrémale pour l'une est extrémale pour l'autre, et les deux lagrangiens sont donc équivalents. De même, la multiplication de $\mathcal{L}$ par une constante change l'action, mais ne change pas les équations du mouvement.

2.5 Deuxième justification : à partir des équations de Newton

Dans cette section, nous allons retrouver les équations d'Euler-Lagrange à partir des équations de Newton. Nous verrons comment les forces de contraintes disparaissent en utilisant le *principe du travail virtuel de d'Alembert*.

On considère, comme d'habitude, un système composé de N points matériels de positions $\vec{r}_\alpha$, avec $\alpha = 1, 2, \dots, N$. Posons

$$\vec{F}_\alpha = \text{résultante de toutes les forces agissant sur le point matériel } \alpha. \quad (2.113)$$

(Ces forces incluent toutes les forces agissant sur α , aussi bien les *forces extérieures*, exercées par des objets extérieurs au système, que les *forces intérieures*, exercées par d'autres éléments du système.)

Les équations du mouvement, d'après la deuxième loi de Newton, sont alors

$$m_\alpha \ddot{\vec{r}}_\alpha = \vec{F}_\alpha, \quad (2.114)$$

avec m_α la masse du point matériel α .

On se choisit des coordonnées généralisées $q_1, \dots, q_s$. Les coordonnées réelles $\vec{r}_\alpha$ s'écrivent en fonction des coordonnées généralisées et du temps :

$$\vec{r}_\alpha = \vec{r}_\alpha(q_1, q_2, \dots, q_s, t). \quad (2.115)$$

2.5.1 Équations du mouvement

Pour un i et α fixés, on prend le produit scalaire de (2.114) avec $\partial \vec{r}_\alpha / \partial q_i$:

$$m_\alpha \ddot{\vec{r}}_\alpha \cdot \frac{\partial \vec{r}_\alpha}{\partial q_i} = \vec{F}_\alpha \cdot \frac{\partial \vec{r}_\alpha}{\partial q_i}. \quad (2.116)$$

On va exprimer le membre de gauche de (2.116) en fonction de l'énergie cinétique de la particule α . Pour cela, on utilisera les deux propriétés suivantes, valables pour tout vecteur $\vec{r}$ fonction de q et t :

$$\frac{\partial \dot{\vec{r}}}{\partial \dot{q}_i} = \frac{\partial \vec{r}}{\partial q_i}, \quad \frac{d}{dt} \left(\frac{\partial \vec{r}}{\partial q_i} \right) = \frac{\partial \dot{\vec{r}}}{\partial q_i}. \quad (2.117)$$

La première égalité a déjà été vue en (2.76) ; démontrons la deuxième. Tout comme $\vec{r}$, la dérivée partielle $\frac{\partial \vec{r}}{\partial q_i}$ dépend des q_i et de t , mais pas des $\dot{q}_i$. On peut donc écrire

$$\frac{d}{dt} \left(\frac{\partial \vec{r}}{\partial q_i} \right) = \sum_{b=1}^s \left[\frac{\partial}{\partial q_b} \left(\frac{\partial \vec{r}}{\partial q_i} \right) \dot{q}_b \right] + \frac{\partial}{\partial t} \left(\frac{\partial \vec{r}}{\partial q_i} \right), \quad (2.118)$$

qui est la même formule pour la dérivée que (2.74), mais appliquée à $\frac{\partial \vec{r}}{\partial q_i}$ plutôt qu'à $\vec{r}_\alpha$. On peut alors changer l'ordre des dérivées secondes et conclure

$$\frac{d}{dt} \left(\frac{\partial \vec{r}}{\partial q_i} \right) = \sum_{b=1}^s \left[\frac{\partial}{\partial q_i} \left(\frac{\partial \vec{r}}{\partial q_b} \right) \dot{q}_b \right] + \frac{\partial}{\partial q_i} \left(\frac{\partial \vec{r}}{\partial t} \right) = \frac{\partial}{\partial q_i} \left(\sum_{b=1}^s \frac{\partial \vec{r}}{\partial q_b} \dot{q}_b + \frac{\partial \vec{r}}{\partial t} \right) = \frac{\partial \dot{\vec{r}}}{\partial q_i}. \quad (2.119)$$

On peut maintenant transformer le membre de gauche de (2.116) de la façon suivante :

$$m_\alpha \ddot{\vec{r}}_\alpha \cdot \frac{\partial \vec{r}_\alpha}{\partial q_i} = m_\alpha \left[\frac{d}{dt} \left(\dot{\vec{r}}_\alpha \cdot \frac{\partial \vec{r}_\alpha}{\partial q_i} \right) - \dot{\vec{r}}_\alpha \cdot \frac{d}{dt} \left(\frac{\partial \vec{r}_\alpha}{\partial q_i} \right) \right] \quad (2.120)$$

$$= m_\alpha \left[\frac{d}{dt} \left(\dot{\vec{r}}_\alpha \cdot \frac{\partial \dot{\vec{r}}_\alpha}{\partial \dot{q}_i} \right) - \dot{\vec{r}}_\alpha \cdot \frac{\partial \dot{\vec{r}}_\alpha}{\partial q_i} \right] \quad (2.121)$$

$$= \frac{1}{2} m_\alpha \left[\frac{d}{dt} \left(\frac{\partial (\dot{\vec{r}}_\alpha^2)}{\partial \dot{q}_i} \right) - \frac{\partial (\dot{\vec{r}}_\alpha^2)}{\partial q_i} \right] \quad (2.122)$$

$$= \frac{d}{dt} \left(\frac{\partial T_\alpha}{\partial \dot{q}_i} \right) - \frac{\partial T_\alpha}{\partial q_i}, \quad (2.123)$$

où nous avons utilisé (2.117) et où $T_\alpha = \frac{1}{2} m_\alpha \dot{\vec{r}}_\alpha^2$ est l'énergie cinétique de la particule α . En sommant sur tous les α , on fait apparaître l'énergie cinétique totale du système $T = T_1 + T_2 + \dots + T_N$, et on obtient finalement, avec (2.116),

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_i} \right) - \frac{\partial T}{\partial q_i} = Q_i. \quad \text{avec } Q_i := \sum_{\alpha} \vec{F}_\alpha \cdot \frac{\partial \vec{r}_\alpha}{\partial q_i}. \quad (2.124)$$

La quantité Q_i s'appelle la *force généralisée* associée à la coordonnée i . Le résultat (2.124) est valable pour tous les systèmes, y compris avec des forces de frottement ou des forces électromagnétiques.

2.5.2 Prise en compte des différents types de force

On rappelle que $\vec{F}_\alpha$ est la résultante de toutes les forces s'appliquant sur le point matériel α . On décompose cette force en trois contributions : les forces qui dérivent d'un potentiel, les forces de contrainte, et les autres forces (magnétisme, frottement, ...) :

$$\vec{F}_\alpha = \vec{F}_\alpha^{\text{potentiel}} + \vec{F}_\alpha^{\text{contraintes}} + \vec{F}_\alpha^{\text{autres}}. \quad (2.125)$$

Cette décomposition entraîne une décomposition associée pour les Q_i :

$$Q_i = Q_i^{\text{potentiel}} + Q_i^{\text{contraintes}} + Q_i^{\text{autres}}. \quad (2.126)$$

Examinons ces contributions les unes après les autres

Les forces qui dérivent d'un potentiel Ce sont les forces pour lesquelles il existe un potentiel $U = U(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N, t)$ qui dépend des positions mais non des vitesses tel que

$$\vec{F}_\alpha^{\text{potentiel}} = -\vec{\nabla}_{\vec{r}_\alpha} U \quad \text{pour tout } \alpha \in \{1, 2, \dots, N\}. \quad (2.127)$$

La notation $\vec{\nabla}_{\vec{r}_\alpha}$ signifie que l'on prend le gradient par rapport à la variable $\vec{r}_\alpha$. Si U ne dépend pas explicitement du temps, les forces $\vec{F}_\alpha$ sont dites conservatives.

Réécrivons U comme une fonction des coordonnées généralisées

$$U(q_1, \dots, q_s, t) = U(\vec{r}_1, \dots, \vec{r}_N, t) = U(\vec{r}_1(q_1, \dots, q_s, t), \dots, \vec{r}_N(q_1, \dots, q_s, t), t), \quad (2.128)$$

et en dérivant par rapport à q_i , on obtient

$$-\frac{\partial U}{\partial q_i} = -\sum_{\alpha=1}^N \vec{\nabla}_{\vec{r}_\alpha} U \cdot \frac{\partial \vec{r}_\alpha}{\partial q_i} = \sum_{\alpha=1}^N \vec{F}_\alpha^{\text{potentiel}} \cdot \frac{\partial \vec{r}_\alpha}{\partial q_i} = Q_i^{\text{potentiel}}. \quad (2.129)$$

(Cette écriture justifie l'appellation de «force généralisée» pour Q_i .)

Les forces de contraintes Le résultat important est que

$$Q_i^{\text{contraintes}} = 0 \quad \text{pour tout } i. \quad (2.130)$$

Les explications sur cette propriété des forces de contraintes sont données dans la section 2.5.3. Admettons-là pour l'instant pour conclure.

Le terme Q_i^{autres} Ce terme est à traiter au cas par cas.

En utilisant (2.129) et (2.130) dans (2.124), on obtient

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_i} \right) - \frac{\partial T}{\partial q_i} = -\frac{\partial U}{\partial q_i} + Q_i^{\text{autres}} \quad \Leftrightarrow \quad \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_i} \right) = \frac{\partial (T - U)}{\partial q_i} + Q_i^{\text{autres}}. \quad (2.131)$$

On peut alors construire le *lagrangien*

$$\mathcal{L} = T - U. \quad (2.132)$$

Le lagrangien dépend des q_i et des $\dot{q}_i$ (cf. T) et parfois explicitement du temps t . Comme U ne dépend pas des vitesses généralisées $\dot{q}_i$, on obtient

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) = \frac{\partial \mathcal{L}}{\partial q_i} + Q_i^{\text{autres}}, \quad (2.133)$$

Dans le cas où les seules forces sont des forces de contraintes ou des forces dérivant d'un potentiel, on a $Q_i^{\text{autres}} = 0$ et on retrouve les équations d'Euler-Lagrange (2.6).

2.5.3 Les forces de contrainte et le déplacement virtuel

On considère le système à un instant t donné, et on imagine que l'on change les coordonnées généralisées q_i d'une petite quantité dq_i , sans changer la valeur de t . Alors, tous les vecteurs positions augmentent de

$$\delta \vec{r}_\alpha = \sum_{i=1}^s \frac{\partial \vec{r}_\alpha}{\partial q_i} dq_i. \quad (2.134)$$

Ces quantités, pour $\alpha = 1, \dots, N$, s'appellent les *déplacements virtuels* des particules du système. En variant le choix des dq_i , les $\delta \vec{r}_\alpha$ représentent tous les déplacements infinitésimaux autorisés par les contraintes à un instant donné.

Les forces $\vec{F}_\alpha^{\text{contraintes}}$ s'ajustent continuellement pour garantir que les contraintes sont respectées. Par contre elle ne favorisent les déplacements virtuels autorisés ni dans un sens, ni dans l'autre, Cela signifie que le travail *virtuel* global (correspondant au déplacement virtuel) doit être nul :

$$\sum_{\alpha=1}^N \vec{F}_\alpha^{\text{contraintes}} \cdot \delta \vec{r}_\alpha = 0 \quad \text{pour tous les choix de } dq_i. \quad (2.135)$$

C'est le *principe du travail virtuel de d'Alembert*. En utilisant la définition (2.134) des $\delta \vec{r}_\alpha$ dans (2.135), on obtient

$$\begin{aligned} 0 &= \sum_{\alpha=1}^N \vec{F}_\alpha^{\text{contraintes}} \cdot \delta \vec{r}_\alpha \\ &= \sum_{\alpha=1}^N \vec{F}_\alpha^{\text{contraintes}} \cdot \left(\sum_{i=1}^s \frac{\partial \vec{r}_\alpha}{\partial q_i} dq_i \right) \\ &= \sum_{i=1}^s \left(\sum_{\alpha=1}^N \vec{F}_\alpha^{\text{contraintes}} \cdot \frac{\partial \vec{r}_\alpha}{\partial q_i} \right) dq_i \\ &= \sum_{i=1}^s Q_i^{\text{contraintes}} dq_i, \end{aligned} \quad (2.136)$$

où l'on a utilisé la définition (2.116) de Q_i .

Cette égalité devant être valable pour tous les choix des dq_i , on voit que le principe du travail virtuel de d'Alembert (2.135) redonne bien la condition (2.130) : $Q_i^{\text{contraintes}} = 0$ pour tout i .

Faisons deux remarques importantes :

- Dans la condition (2.135), c'est la somme sur tous les α qui est nulle, mais il peut arriver que certains termes soient non-nuls.
- Lorsque le système évolue pendant un instant dt , les coordonnées généralisées évoluent de dq_i , et les positions réelles $\vec{r}_\alpha = \vec{r}_\alpha(q_1, \dots, q_s, t)$ évoluent de

$$d\vec{r}_\alpha = \left(\sum_{i=1}^s \frac{\partial \vec{r}_\alpha}{\partial q_i} dq_i \right) + \frac{\partial \vec{r}_\alpha}{\partial t} dt. \quad (2.137)$$

C'est le déplacement réel. Si on le divise par dt , on retrouve la vitesse, voir (2.74). La seule différence entre le déplacement réel $d\vec{r}_\alpha$ et le déplacement virtuel $\delta \vec{r}_\alpha$ est la présence du terme $(\partial \vec{r}_\alpha / \partial t) dt$. Lorsque les fonctions $\vec{r}_\alpha(q_1, \dots, q_s, t)$ ne dépendent pas explicitement du temps, alors $d\vec{r}_\alpha = \delta \vec{r}_\alpha$, le travail virtuel est égal au travail réel, et le principe de d'Alembert se simplifie en : le travail de toutes les forces de contraintes est nul. *A contrario*, si les fonctions $\vec{r}_\alpha(q_1, \dots, q_s, t)$ dépendent explicitement du temps, il peut arriver que les forces de contraintes aient un travail non nul.

2.5.4 Quelques exemples

On vérifie le principe du travail virtuel de d'Alembert sur trois des exemples de la section 2.2, voir figure 2.7.

Pour le pendule La position du pendule s'écrit $\vec{r} = \ell \vec{e}_\rho$ (avec $(\vec{e}_\rho, \vec{e}_\varphi)$ les vecteurs unitaires polaires). Cette position dépend de la coordonnée généralisée θ puisque le vecteur $\vec{e}_\rho$ dépend de θ , mais elle ne dépend pas explicitement du temps t . On a donc ici $d\vec{r} = \delta \vec{r}$, et le travail virtuel est égal au travail réel.

La force de contrainte est la force de tension $\vec{F}$ du fil, qui s'ajuste continuellement pour maintenir le fil à la longueur constante ℓ . On voit bien que le travail de la force de contrainte

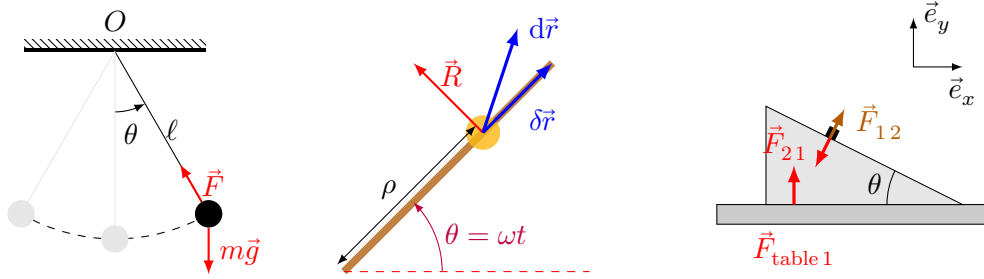


FIGURE 2.7 – Trois des exemples de la section 2.2.

est nul, puisque le déplacement est toujours perpendiculaire à la force. Le principe de d'Alembert est respecté.

Pour la perle sur la tige La position de la perle est $\vec{r} = \rho \vec{e}_\rho$ (avec $(\vec{e}_\rho, \vec{e}_\varphi)$ les vecteurs unitaires polaires). Cette position dépend de la coordonnée généralisée ρ et du temps t , puisque le vecteur $\vec{e}_\rho$ dépend de l'angle ωt entre l'horizontale et la tige. On est dans la situation où $d\vec{r}$ et $\delta\vec{r}$ sont différents : le vecteur $\delta\vec{r}$ décrit le déplacement le long de la tige dû aux changements de ρ , alors que le vecteur $d\vec{r}$ prend également en compte le mouvement de la tige :

$$\delta\vec{r} = d\rho \vec{e}_\rho, \quad d\vec{r} = d\rho \vec{e}_\rho + \rho \omega dt \vec{e}_\varphi. \quad (2.138)$$

(On a représenté ces deux vecteurs sur la figure 2.7.)

La force de contrainte est la force $\vec{R}$ exercée par la tige sur la perle qui s'ajuste continuellement pour maintenir la perle sur la tige. Cette force est perpendiculaire à la tige ; on a donc :

$$\vec{R} \cdot \delta\vec{r} = 0. \quad (2.139)$$

Le travail virtuel de la force de contrainte est bien nul, le principe de d'Alembert est respecté. Cependant, le travail réel $\vec{R} \cdot d\vec{r}$ n'est pas nul : en pratique, la tige fournit de l'énergie à la perle pour l'éjecter.

Pour le plan incliné+mobile Il y a dans ce système $N = 2$ objets, le plan incliné (de position $\vec{r}_1$) et le mobile (de position $\vec{r}_2$). Leurs positions sont des fonctions des $s = 2$ coordonnées généralisées, l'abscisse x_C du centre de masse du plan, et la distance ℓ entre le mobile et le coin supérieur gauche du plan. Les expressions de $\vec{r}_1$ et $\vec{r}_2$ en fonction de ℓ et x_C ne font pas intervenir le temps, voir (2.29). On est donc dans la situation où le déplacement virtuel est égal au déplacement réel :

$$d\vec{r}_1 = \delta\vec{r}_1, \quad d\vec{r}_2 = \delta\vec{r}_2. \quad (2.140)$$

Le plan incliné (objet 1), ressent deux forces de contrainte, l'une étant la réaction normale de la table $\vec{F}_{\text{table 1}}$ et l'autre la réaction du mobile $\vec{F}_{21}$. Le mobile (objet 2) ressent comme unique force de contrainte la réaction du plan incliné $\vec{F}_{12}$:

$$\vec{F}_1^{\text{contraintes}} = \vec{F}_{\text{table 1}} + \vec{F}_{21}, \quad \vec{F}_2^{\text{contraintes}} = \vec{F}_{12}. \quad (2.141)$$

Ces forces font en sorte que le plan reste sur la table et le mobile sur le plan. On les a représentées sur la figure 2.7.

Le travail (réel ou virtuel, c'est ici la même chose) des forces de contraintes est donc

$$\delta W = (\vec{F}_{\text{table 1}} + \vec{F}_{21}) \cdot d\vec{r}_1 + \vec{F}_{12} \cdot d\vec{r}_2. \quad (2.142)$$

Il est clair que $\vec{F}_{\text{table 1}}$ ne travaille pas : cette force est verticale, mais agit sur un point qui se déplace horizontalement. Cependant, les deux autres forces de contrainte travaillent :

par exemple, on voit bien sur la figure 2.7 que la force $\vec{F}_{21}$ n'est pas perpendiculaire au mouvement du plan incliné.

Selon le principe des actions réciproques, on a $\vec{F}_{21} = -\vec{F}_{12}$. Le travail des forces de contrainte se réduit donc à

$$\delta W = \vec{F}_{12} \cdot (d\vec{r}_2 - d\vec{r}_1) = \vec{F}_{12} \cdot d(\vec{r}_2 - \vec{r}_1). \quad (2.143)$$

À partir de l'expression (2.29) de $\vec{r}_1$ et $\vec{r}_2$ en fonction des coordonnées généralisées x_C et ℓ , on a

$$\vec{r}_2 - \vec{r}_1 = \overrightarrow{CA} + \ell(\cos \theta \vec{e}_x - \sin \theta \vec{e}_y), \quad d(\vec{r}_2 - \vec{r}_1) = d\ell(\cos \theta \vec{e}_x - \sin \theta \vec{e}_y). \quad (2.144)$$

Mais le vecteur $\cos \theta \vec{e}_x - \sin \theta \vec{e}_y$ est le vecteur unitaire dans la direction descendante de la pente. Il est perpendiculaire à $\vec{F}_{21}$, et le travail total des forces de contraintes est nul. Le principe de d'Alembert est respecté.

Chapitre 3

Mouvement d'un solide

3.1 Généralités

3.1.1 Définition d'un corps solide

Un *corps solide* est simplement défini comme un système de points matériels à distances fixes les uns des autres. Nous considérons des corps solides parfaitement rigides et excluons toute élasticité ou déformation de ce solide. Le corps solide constitue donc un cas très particulier d'un système de points matériels. En effet même si un corps solide est constitué d'un grand nombre de points matériels, la description de son mouvement ne requiert que la connaissance (en fonction du temps) de la position de l'un de ses points et de l'orientation du solide autour de ce point. On voit que tout corps solide possède six degrés de liberté, et les six coordonnées choisies pour décrire la position (et le mouvement) du solide peuvent être choisies d'un grand nombre de manières différentes.

3.1.2 Description mathématique d'un solide

On peut considérer un solide comme un ensemble discret de N points matériels de masse m_α et de position $\vec{R}_\alpha$, avec $\alpha = 1, \dots, N$. On peut également considérer un solide comme un milieu continu, caractérisé par une masse volumique $\rho(\vec{R})$. Toutes les quantités intéressantes peuvent alors s'écrire comme une somme sur les particules α , ou comme une intégrale sur le volume du solide, le passage de l'un à l'autre se faisant en remplaçant les masses des particules par $\rho dV = dm$, la masse contenue dans un élément de volume dV ; en coordonnées cartésiennes par exemple, on a $\rho dV = \rho(x, y, z) dx dy dz$. Ainsi, la masse totale M du solide ou la position $\vec{R}_C$ de son centre de masse peuvent s'écrire :

$$M = \sum_{\alpha} m_{\alpha} = \int \rho dV = \int dm, \quad \vec{R}_C = \frac{1}{M} \sum_{\alpha} m_{\alpha} \vec{R}_{\alpha} = \frac{1}{M} \int dV \rho \vec{R} = \frac{1}{M} \int dm \vec{R}. \quad (3.1)$$

(Beaucoup de livres préfèrent mettre des intégrales triples $\iiint$ plutôt que des intégrales simples $\int$ pour rappeler qu'on est en trois dimensions. C'est une affaire de goût.)

Pour décrire le mouvement d'un point P du solide on utilise d'une part le vecteur $\vec{R} = \overrightarrow{OP}$, où O est le centre du référentiel d'étude. On va également utiliser le vecteur $\vec{r} = \overrightarrow{AP}$, où A est un point donné du solide. Il sera pratique plus tard de prendre A égal au centre de masse, mais pour l'instant A est un point quelconque lié au solide. On a évidemment :

$$\vec{R} = \overrightarrow{OP}, \quad \vec{R}_A = \overrightarrow{OA}, \quad \vec{r} = \overrightarrow{AP}, \quad \Rightarrow \quad \vec{R} = \vec{R}_A + \vec{r}. \quad (3.2)$$

On considère un petit mouvement du solide. Si on suppose, dans un premier temps, que le point A est immobile, alors le mouvement est une rotation autour d'un certain axe passant par A et

porté par un vecteur unitaire $\vec{n}$. Supposons que la rotation soit infinitésimale et d'angle $d\varphi$; pour décrire cette rotation, on définit $d\vec{\varphi}$ le *vecteur rotation infinitésimal* par la formule suivante :

$$d\vec{\varphi} = d\varphi \vec{n}. \quad (3.3)$$

(Le sens de la rotation décrit par un vecteur $d\vec{\varphi}$ est donné par la règle du tire-bouchon : le solide tourne dans le sens selon lequel on devrait tourner un tire-bouchon pour qu'il avance dans le sens pointé par $d\vec{\varphi}$.)

Dans le cas général où A n'est pas immobile, on peut décrire un mouvement infinitésimal du solide comme une translation infinitésimale $d\vec{R}_A$ du point A plus une rotation infinitésimale $d\vec{\varphi}$.

$$\text{mouvement infinitésimal du solide} = \begin{cases} \text{translation infinitésimale du point } A \\ + \\ \text{rotation infinitésimale autour d'un axe passant par } A \end{cases}$$

Une formule importante de ce chapitre est

$$d\vec{r} = d\vec{\varphi} \wedge \vec{r}, \quad \text{et donc} \quad d\vec{R} = d\vec{R}_A + d\vec{\varphi} \wedge \vec{r}. \quad (3.4)$$

Pour comprendre l'expression de $d\vec{r}$, on décompose $\vec{r}$ en un vecteur parallèle et un vecteur perpendiculaire à $\vec{n}$: $\vec{r} = \vec{r}_{\parallel} + \vec{r}_{\perp}$ (voir la figure 3.1). Lors de la rotation autour de $\vec{n}$, le vecteur $\vec{r}_{\parallel}$ ne change pas, tandis que $\vec{r}_{\perp}$ change de direction mais pas de longueur et reste orthogonal à $\vec{n}$. Sa variation $d\vec{r}_{\perp}$ est donc à la fois orthogonale à $\vec{n}$ et à $\vec{r}_{\perp}$ (puisque $0 = d(\vec{r}_{\perp}^2) = 2\vec{r}_{\perp} \cdot d\vec{r}_{\perp}$). Comme par ailleurs la norme de $d\vec{r}_{\perp}$ est $\|\vec{r}_{\perp}\|d\varphi$, on arrive à $d\vec{r}_{\perp} = d\vec{\varphi} \wedge \vec{r}_{\perp}$. Finalement, comme $d\vec{r} = d\vec{r}_{\perp}$ (car $d\vec{r}_{\parallel} = 0$) et que $d\vec{\varphi} \wedge \vec{r} = d\vec{\varphi} \wedge \vec{r}_{\perp}$ (car $\vec{r}_{\parallel}$ est colinéaire à $d\vec{\varphi}$), on arrive bien à $d\vec{r} = d\vec{\varphi} \wedge \vec{r}$.

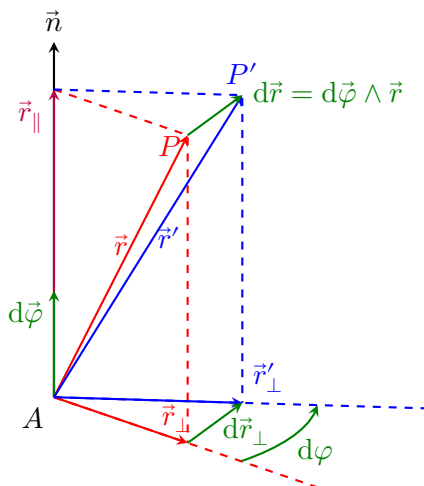


FIGURE 3.1 – Lors de la rotation infinitésimale de vecteur $d\vec{\varphi} = d\varphi \vec{n}$, le point P de vecteur $\vec{r} = \overrightarrow{AP}$ est transformé en le point P' de vecteur $\vec{r}' = \overrightarrow{AP'}$. On décompose $\vec{r}$ en $\vec{r}_{\parallel}$ (parallèle à $d\vec{\varphi}$, et donc à $\vec{n}$) et $\vec{r}_{\perp}$ (perpendiculaire à $d\vec{\varphi}$). Le vecteur $\vec{r}_{\parallel}$ est inchangé lors de la rotation, alors que le vecteur $\vec{r}_{\perp}$ subit une rotation d'angle $d\varphi$ dans le plan passant par A perpendiculaire à $\vec{n}$.

En divisant (3.4) par dt et en introduisant la vitesse $\vec{V}$ du point P , la vitesse $\vec{V}_A$ du point A et le vecteur vitesse angulaire $\vec{\Omega}$ par

$$\frac{d\vec{R}}{dt} = \vec{V}, \quad \frac{d\vec{R}_A}{dt} = \vec{V}_A, \quad \frac{d\vec{\varphi}}{dt} = \vec{\Omega}, \quad (3.5)$$

on obtient la relation

$$\boxed{\frac{d\vec{r}}{dt} = \vec{\Omega} \wedge \vec{r}, \quad \text{et donc} \quad \vec{V} = \vec{V}_A + \vec{\Omega} \wedge \vec{r}} \quad (3.6)$$

La vitesse $\vec{V}_A$ du point A caractérise le mouvement de translation du solide à un instant donné, et la vitesse angulaire $\vec{\Omega}$ caractérise sa rotation. (Remarque et exercice : on peut montrer que le vecteur vitesse angulaire $\vec{\Omega}$ ne dépend pas du point A choisi.).

3.2 Énergie cinétique et moment d'inertie

3.2.1 Calcul de l'énergie cinétique et moment d'inertie

À partir de l'expression (3.6), écrivons l'énergie cinétique du solide considéré d'abord comme un ensemble discret de points matériels, puis comme un continuum

$$T = \sum_{\alpha} \frac{1}{2} m_{\alpha} V_{\alpha}^2 = \frac{1}{2} \int dm V^2 = \frac{1}{2} \int dm [V_A^2 + 2\vec{V}_A \cdot (\vec{\Omega} \wedge \vec{r}) + \|\vec{\Omega} \wedge \vec{r}\|^2] \quad (3.7)$$

Remarquons que $\vec{V}_A$ et $\vec{\Omega}$ ne dépendent pas du point courant peuvent sortir de l'intégrale. Ainsi, le premier terme est simplement $\int dm V_A^2 = V_A^2 \int dm = M V_A^2$ avec M la masse totale du solide..

Pour le deuxième terme, on trouve

$$\int dm \vec{V}_A \cdot (\vec{\Omega} \wedge \vec{r}) = \vec{V}_A \cdot \left(\vec{\Omega} \wedge \int dm \vec{r} \right). \quad (3.8)$$

Ce terme est nul si $\vec{V}_A = \vec{0}$, bien sûr, ou si $A = C$, le centre de masse du solide. En effet, par définition du centre de masse, on a $\int dm \vec{r} = \vec{0}$ si $A = C$. **On se placera toujours dans l'un de ces deux cas où ce second terme est nul.**

Pour le troisième terme, introduisons θ l'angle entre $\vec{r}$ et $\vec{\Omega}$, et $r_{\perp}$ la distance entre P et l'axe de rotation, c'est-à-dire l'axe porté par $\vec{\Omega}$ et passant par A , voir figure 3.2.

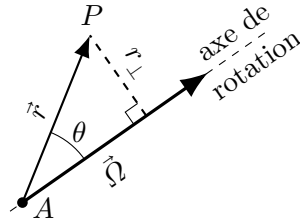


FIGURE 3.2 – Définitions de θ et $r_{\perp}$.

On a $\|\vec{\Omega} \wedge \vec{r}\|^2 = \Omega^2 r^2 \sin^2 \theta$. Mais $r_{\perp} = r \sin \theta$, et donc

$$\|\vec{\Omega} \wedge \vec{r}\|^2 = \Omega^2 r_{\perp}^2. \quad (3.9)$$

On arrive alors à

$$\frac{1}{2} \int dm (\vec{\Omega} \wedge \vec{r})^2 = \frac{1}{2} \int dm r_{\perp}^2 \Omega^2 = \frac{1}{2} I_{\Delta} \Omega^2 \quad (3.10)$$

avec

$$I_{\Delta} = \int dm r_{\perp}^2. \quad (3.11)$$

Dans cette expression, on a noté Δ l'axe passant par A et parallèle à $\vec{\Omega}$, et I_{Δ} est le **moment d'inertie du solide par rapport à l'axe Δ** . Dans l'expression de I_{Δ} , $r_{\perp}$ est la distance entre le point courant et l'axe Δ .

On obtient finalement deux expressions pour l'énergie cinétique du solide

$$\text{si } C \text{ est le centre de masse} \quad T = \frac{1}{2}MV_C^2 + \frac{1}{2}I_\Delta\Omega^2 \quad (\text{avec } \Delta \text{ passant par } C), \quad (3.12)$$

$$\text{si } A \text{ est un point immobile} \quad T = \frac{1}{2}I_{\Delta'}\Omega^2 \quad (\text{avec } \Delta' \text{ passant par } A). \quad (3.13)$$

Dans la première expression, le premier terme $\frac{1}{2}MV_C^2$ représente l'énergie cinétique de translation et le second terme $\frac{1}{2}I_\Delta\Omega^2$ l'énergie cinétique de rotation.

3.2.2 Exemple : le pendule rigide

On considère un pendule rigide constitué d'une tige fine homogène de masse M et de longueur ℓ , qui oscille par rapport à l'une de ses extrémités qui est maintenue fixe, voir figure 3.3. À tout instant, la position du pendule est caractérisée par son angle orienté θ par rapport à la verticale descendante. Le vecteur rotation $\vec{\Omega}$ est perpendiculaire au plan (Oxy) du mouvement, et en particulier perpendiculaire à la tige. Il est donné par $\vec{\Omega} = \dot{\theta}\vec{e}_z$.

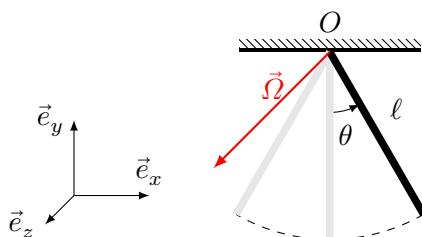


FIGURE 3.3 – Le pendule rigide

Ici, le point O est immobile. On note Δ' l'axe passant par O et porté par $\vec{e}_z$, et on calcule $I_{\Delta'}$. En notant que la masse d'un élément dr de la tige est $dm = \lambda dr$ avec $\lambda = \frac{M}{\ell}$ la masse linéique, l'expression (3.11) donne :

$$I_{\Delta'} = \int_0^\ell \frac{M}{\ell} dr r^2 = \frac{M}{\ell} \times \left[\frac{\ell^3}{3} - \frac{0^3}{3} \right] = \frac{M\ell^2}{3}. \quad (3.14)$$

L'expression dans (3.13) donne alors :

$$T = \frac{1}{2}I_{\Delta'}\Omega^2 = \frac{M\ell^2}{6}\Omega^2. \quad (3.15)$$

On peut aussi utiliser (3.12) et considérer le centre de masse C de la tige, qui se trouve évidemment en son milieu. Soit Δ l'axe parallèle à $\vec{e}_z$ et passant par C ; l'expression (3.11) donne :

$$I_\Delta = \int_{-\ell/2}^{\ell/2} \frac{M}{\ell} dr r^2 = \frac{M}{\ell} \times \left[\frac{(\ell/2)^3}{3} - \frac{(-\ell/2)^3}{3} \right] = \frac{M\ell^2}{12}. \quad (3.16)$$

Par ailleurs, la vitesse du point C est (en norme) $V_C = \Omega\ell/2$, et (3.12) donne :

$$T = \frac{1}{2}MV_C^2 + \frac{1}{2}I_\Delta\Omega^2 = \frac{1}{2}M\left(\Omega\frac{\ell}{2}\right)^2 + \frac{1}{2} \times \frac{M\ell^2}{12}\Omega^2 = M\ell^2\Omega^2\left(\frac{1}{8} + \frac{1}{24}\right) = \frac{1}{6}M\ell^2\Omega^2. \quad (3.17)$$

On trouve évidemment la même expression.

Terminons l'exercice ; l'altitude du centre de masse (au milieu de la tige) est $y = -\frac{\ell}{2}\cos\theta$. L'énergie potentielle est donc

$$U = -\frac{Mg\ell}{2}\cos\theta, \quad (3.18)$$

et, en se souvenant que $\Omega = |\dot{\theta}|$, le lagrangien est

$$\mathcal{L} = \frac{M\ell^2}{6}\dot{\theta}^2 + \frac{Mg\ell}{2}\cos\theta. \quad (3.19)$$

Les équations de Lagrange donnent, après division par $M\ell^2/3$,

$$\ddot{\theta} = -\frac{3g}{2\ell}\sin\theta. \quad (3.20)$$

Pour des petites oscillations, $\sin\theta \simeq \theta$ et on retrouve un oscillateur harmonique de pulsation $\omega = \sqrt{\frac{3g}{2\ell}}$.

3.2.3 Exemple : cylindre roulant sans glisser sur un plan incliné

Considérons un cylindre homogène de rayon a et masse M qui roule sur un plan incliné d'angle α comme représenté figure 3.4. On note X la distance parcourue depuis le début de l'expérience, et θ l'angle dont a tourné le cylindre depuis le début de l'expérience.

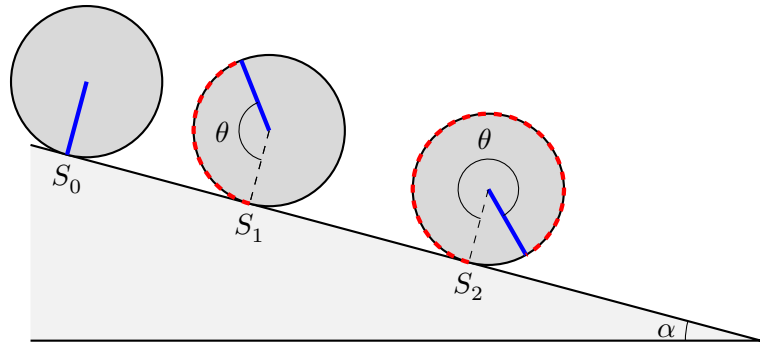


FIGURE 3.4 – Un cylindre homogène de rayon a et masse M roule sans vitesse initiale sur un plan incliné d'angle α . On a représenté le cylindre dans sa configuration initiale (en haut) et à deux autres instants. Les points S_0 , S_1 , S_2 sont les points de contact entre le cylindre et la pente ; pour se repérer, on a tracé (en bleu) sur le cylindre le rayon entre le centre et le point de contact initial S_0 . On pose X la distance parcourue depuis le début de l'expérience ; pour le cylindre du milieu, on a $X = S_0S_1$.

La vitesse du centre de masse est $\dot{X}$, et la vitesse de rotation du cylindre est $\dot{\theta}$. En utilisant (3.12), l'énergie cinétique du solide s'écrit donc

$$T = \frac{1}{2}M\dot{X}^2 + \frac{1}{2}I_{\Delta}\dot{\theta}^2, \quad (3.21)$$

où I_{Δ} est le moment d'inertie par rapport à l'axe Δ qui est l'axe de révolution du cylindre. Calculons maintenant I_{Δ} . On se place (naturellement) en coordonnées cylindriques (r, θ, z) avec O au centre du cylindre et l'axe z selon l'axe de révolution du cylindre. On rappelle que l'élément de volume est $dV = r dr d\theta dz$. Soit h la hauteur du cylindre ; son volume est $\pi a^2 h$, et sa masse volumique est $M/(\pi a^2 h)$. La masse correspondant à un élément de volume est alors

$$dm = \frac{M}{\pi a^2 h} dV = \frac{M}{\pi a^2 h} r dr d\theta dz. \quad (3.22)$$

La distance à l'axe est justement la coordonnée r du système. On obtient donc

$$I_{\Delta} = \int dm r^2 = \int_0^a dr \int_0^{2\pi} d\theta \int_{-\frac{h}{2}}^{\frac{h}{2}} dz \frac{M}{\pi a^2 h} r \times r^2 = \frac{M}{\pi a^2 h} \times \frac{a^4}{4} \times 2\pi \times h = \frac{Ma^2}{2}. \quad (3.23)$$

Pour relier X et θ , on a besoin d'une hypothèse supplémentaire. On va supposer que le cylindre *roule sans glisser*. Avec cette hypothèse, on a alors

$$X = a\theta. \quad (3.24)$$

En effet, entre les deux premiers cylindres de la figure 3.4, la distance parcourue $X = S_0S_1$ doit être égale, en l'absence de glissement, à la longueur représentée en pointillés rouges sur le cylindre du milieu, et cette longueur est justement $a\theta$.

Avec cette contrainte du «roulement sans glissement», le système n'a qu'un seul degré de liberté et, si on choisit X comme variable, en utilisant $\theta = X/a$ et $\dot{\theta} = \dot{X}/a$, on trouve pour l'énergie cinétique

$$T = \frac{1}{2}M\dot{X}^2 + \frac{1}{2} \times \frac{Ma^2}{2}\dot{\theta}^2 = \frac{1}{2}M\dot{X}^2 + \frac{M}{4}\dot{X}^2 = \frac{3}{4}M\dot{X}^2. \quad (3.25)$$

Lorsque le cylindre se déplace d'une distance X , son altitude décroît de $X \sin \alpha$. En prenant l'origine des altitudes à la position initiale, l'altitude du centre de masse est donc $-X \sin \alpha$ et l'énergie potentielle est

$$U = -Mg \sin \alpha X. \quad (3.26)$$

Le lagrangien est donc

$$\mathcal{L} = \frac{3}{4}M\dot{X}^2 + Mg \sin \alpha X. \quad (3.27)$$

et les équations du mouvement sont, après simplification,

$$\ddot{X} = \frac{2}{3}g \sin \alpha. \quad (3.28)$$

On remarque que le rayon a n'intervient pas dans ces équations. Pour un objet qui glisse sans rouler sur une pente d'angle α , on aurait eu la même équation sans le $2/3$.

3.3 Moment cinétique et axes principaux du solide

3.3.1 Calcul direct du moment cinétique

On rappelle que le moment cinétique d'un ensemble de points P_α de masses m_α par rapport à un point fixe O est

$$\vec{L}_O = \sum_{\alpha} \vec{R}_\alpha \wedge (m_\alpha \dot{\vec{R}}_\alpha) \quad \text{avec } \vec{R}_\alpha = \overrightarrow{OP}_\alpha \quad (3.29)$$

Le théorème du moment cinétique dit que la dérivée temporelle de $\vec{L}_O$ est la somme des moments des forces extérieures par rapport au point O :

$$\frac{d}{dt} \vec{L}_O = \vec{\mathcal{M}}_O = \sum_{\alpha} \vec{R}_\alpha \wedge \vec{F}_{\alpha, \text{ext}}. \quad (3.30)$$

(Plus de détails sont donnés dans la section A.2.3 de l'appendice A.)

On note C le centre de masse de l'ensemble de points, et on écrit $\vec{R}_\alpha = \vec{R}_C + \vec{r}_\alpha$ (ce qui implique $\dot{\vec{R}}_\alpha = \dot{\vec{R}}_C + \dot{\vec{r}}_\alpha$) où $\vec{R}_C = \overrightarrow{OC}$ et $\vec{r}_\alpha = \overrightarrow{CP}_\alpha$. On remplace dans (3.29); après développement, on obtient quatre termes, mais un petit calcul montre que deux de ces termes sont nuls, et il reste

$$\vec{L}_O = \vec{R}_C \wedge M\dot{\vec{R}}_C + \vec{L}' \quad \text{avec } \vec{L}' = \sum_{\alpha} \vec{r}_\alpha \wedge (m_\alpha \dot{\vec{r}}_\alpha). \quad (3.31)$$

C'est le *premier théorème de König*, à démontrer en exercice. (Voir aussi l'appendice A.) Le terme $\vec{R}_C \wedge M\dot{\vec{R}}_C$ représente le moment cinétique qu'aurait le système s'il était un point matériel positionné en son centre de masse. Le terme $\vec{L}'$ est le moment cinétique due à la rotation du

système autour de son centre de masse. $\vec{L}'$ est également le moment cinétique du système dans le référentiel barycentrique.

Le résultat qui précède est vrai pour n'importe quel ensemble de points. Calculons maintenant $\vec{L}'$ dans le cas particulier où le système est un solide. On peut alors écrire $\dot{\vec{r}}_\alpha = \vec{\Omega} \wedge \vec{r}_\alpha$ et on obtient, en remplaçant la somme par une intégrale

$$\vec{L}' = \sum_\alpha \vec{r}_\alpha \wedge (m_\alpha \vec{\Omega} \wedge \vec{r}_\alpha) = \int dm \vec{r} \wedge (\vec{\Omega} \wedge \vec{r}) = \int dm \vec{r} \wedge (\vec{\Omega} \wedge \vec{r}_\perp). \quad (3.32)$$

Dans la dernière expression, on a écrit $\vec{r} = \vec{r}_\parallel + \vec{r}_\perp$, avec $\vec{r}_\parallel$ parallèle à $\vec{\Omega}$ et $\vec{r}_\perp$ perpendiculaire à $\vec{\Omega}$, de sorte que $\vec{\Omega} \wedge \vec{r} = \vec{\Omega} \wedge \vec{r}_\perp$.

On rappelle l'égalité $\vec{a} \wedge (\vec{b} \wedge \vec{c}) = (\vec{a} \cdot \vec{c})\vec{b} - (\vec{a} \cdot \vec{b})\vec{c}$. On obtient alors

$$\vec{L}' = \int dm [(\vec{r} \cdot \vec{r}_\perp) \vec{\Omega} - (\vec{r} \cdot \vec{\Omega}) \vec{r}_\perp]. \quad (3.33)$$

Mais $\vec{r} \cdot \vec{r}_\perp = 0$. On reconnaît alors dans le premier terme l'expression (3.11) du moment cinétique I_Δ par rapport à l'axe Δ passant par le centre de masse et parallèle à $\vec{\Omega}$:

$$\vec{L}' = I_\Delta \vec{\Omega} - \int dm (\vec{r} \cdot \vec{\Omega}) \vec{r}_\perp. \quad (3.34)$$

Le deuxième terme dans (3.34) est un vecteur perpendiculaire à $\vec{\Omega}$. On voit donc que :

En général, le moment cinétique $\vec{L}'$ du solide dans le référentiel barycentrique n'est pas colinéaire à l'axe de rotation $\vec{\Omega}$.

3.3.2 Axes principaux

Parfois, le deuxième terme dans (3.34) est nul, le moment cinétique $\vec{L}'$ est colinéaire à l'axe de rotation $\vec{\Omega}$ et est donné par $\vec{L}' = I_\Delta \vec{\Omega}$. En particulier, cela arrive si la rotation $\vec{\Omega}$ se fait selon un axe de symétrie du solide. («axe de symétrie» peut avoir plusieurs sens ; ici, « Δ est un axe de symétrie» veut dire que si on remplace toutes les coordonnées des points $\vec{r} = \vec{r}_\parallel + \vec{r}_\perp$ par $\vec{r}' = \vec{r}_\parallel - \vec{r}_\perp$, alors le solide est inchangé.)

On appelle *axe principal* du solide un axe Δ passant par le centre de masse tel que le deuxième terme dans (3.34) est nul, et donc tel que $\vec{L}' = I_\Delta \vec{\Omega}$.

Tous les axes de symétrie sont des axes principaux, mais il peut aussi exister des axes principaux qui ne sont pas des axes de symétrie.

Nous allons admettre les résultats suivants (des éléments de démonstration sont donnés dans la section 3.5

- Tout solide a au moins trois axes principaux orthogonaux deux à deux. On note $\Delta_1, \Delta_2, \Delta_3$ ces trois axes, et I_1, I_2, I_3 les moments d'inertie correspondants, appelés «moments principaux» du solide.
- Si le solide a un axe de symétrie, on peut choisir Δ_1, Δ_2 ou Δ_3 selon cet axe.
- Pour une rotation $\vec{\Omega}$ d'axe quelconque, on peut décomposer le vecteur $\vec{\Omega}$ de la façon suivante :

$$\vec{\Omega} = \vec{\Omega}_1 + \vec{\Omega}_2 + \vec{\Omega}_3, \quad \text{avec } \vec{\Omega}_i \text{ parallèle à } \Delta_i. \quad (3.35)$$

On a alors

$$\vec{L}' = I_1 \vec{\Omega}_1 + I_2 \vec{\Omega}_2 + I_3 \vec{\Omega}_3, \quad T = \frac{1}{2} I_1 \vec{\Omega}_1^2 + \frac{1}{2} I_2 \vec{\Omega}_2^2 + \frac{1}{2} I_3 \vec{\Omega}_3^2 \quad (3.36)$$

(En particulier, si $\vec{\Omega}$ est selon Δ_1 (par exemple), alors $\vec{\Omega}_1 = \vec{\Omega}$, $\vec{\Omega}_2 = \vec{\Omega}_3 = \vec{0}$ et on obtient, $\vec{L}' = I_1 \vec{\Omega}$ et $T = \frac{1}{2} I_1 \vec{\Omega}^2$

3.3.3 Exemple

On considère une boîte dont la base est un carré de côté a avec quatre pieds en ses coins qui reposent sur une table horizontale. Dans la boîte, un moteur fait tourner un axe vertical, sortant du centre de la boîte, à une vitesse angulaire Ω . Sur cet axe est fixée en son centre une tige de masse M et de longueur ℓ , mais qui est inclinée d'un angle α avec l'horizontale. La masse du bloc moteur (=boîte+pieds+moteur+axe vertical) étant M' , la masse totale du système est $M + M'$. Le centre de masse C est aligné avec l'axe en rotation, voir figure 3.5.

On utilise le repère fixe $(O, \vec{e}_x, \vec{e}_y, \vec{e}_z)$ avec $\vec{e}_z$ vertical ascendant et $\vec{e}_x, \vec{e}_y$ selon les côtés du carré, et O sur la table à la verticale de C . On utilise également la base orthonormée $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ liée à la tige en rotation : $\vec{e}_3$ est dans l'axe de la tige, $\vec{e}_1$ est dans le plan vertical contenant la tige et $\vec{e}_2$ est horizontal.

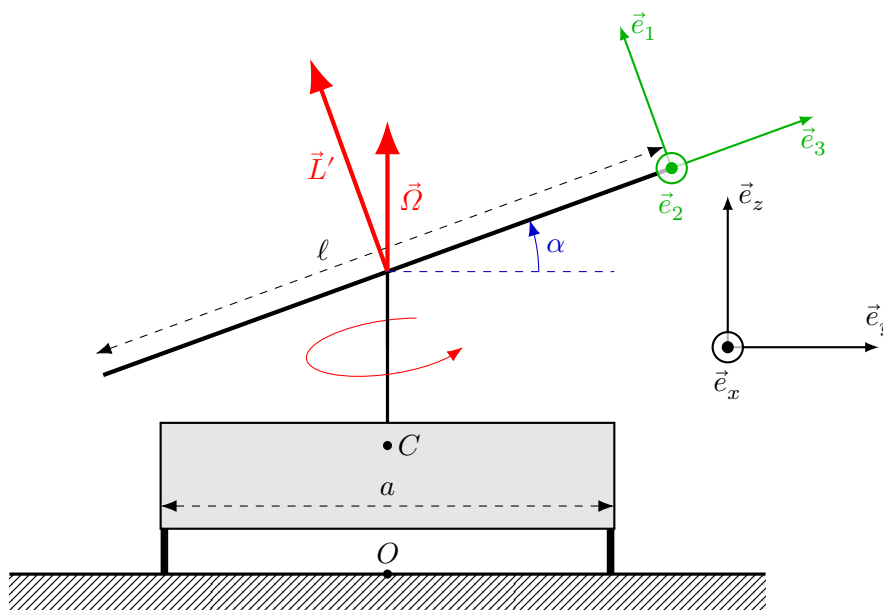


FIGURE 3.5 – Le système étudié vu au moment où la tige inclinée est dans le plan (Oyz) .

Question : à l'instant où la tige est dans le plan (Oyz) (comme sur la figure), quelles sont les forces de réaction s'exerçant sur les pieds de la boîte ?

On note $\vec{R}_{\text{gauche}} = R_{\text{gauche}} \vec{e}_z$ la force de réaction totale s'exerçant sur les deux pieds de gauche (sur la figure, on ne voit qu'un pied à gauche parce qu'ils sont l'un derrière l'autre), et $\vec{R}_{\text{droite}} = R_{\text{droite}} \vec{e}_z$ la force de réaction totale s'exerçant sur les deux pieds de droite. Comme le centre de masse est immobile, et que les seules forces extérieures sont le poids et les réactions de la table sur les pieds, l'équilibre en translation de la boîte implique

$$R_{\text{gauche}} + R_{\text{droite}} = (M + M')g. \quad (3.37)$$

(Rappel : les forces exercées par le moteur sont des forces internes, qui n'interviennent pas dans le principe fondamental de la dynamique.)

On cherche maintenant à appliquer le théorème du moment cinétique. Le moment cinétique de la boîte est nul, celui de l'axe vertical est négligeable, reste le moment cinétique de la tige ; on le calcule en utilisant (3.36) mais, pour cela, on doit identifier les axes principaux Δ_1 , Δ_2 et Δ_3 de la tige, et calculer les moments principaux I_1 , I_2 et I_3 correspondants. Dans le cas de la tige, le calcul est très simple.

- Clairement, par symétrie, l'axe de la tige est un axe principal, et tout axe perpendiculaire à la tige (et passant par son centre) est également un axe principal. On peut donc prendre, avec nos notations, Δ_1 parallèle à $\vec{e}_1$, Δ_2 parallèle à $\vec{e}_2$ et Δ_3 parallèle à $\vec{e}_3$.

- Comme toute la masse de la tige est sur l'axe Δ_3 , le moment I_3 est nul. En effet, dans l'intégrale (3.11), on a $r_\perp = 0$ pour tous les points.
- Les moments I_1 et I_2 sont clairement égaux par symétrie. En fait, on les a déjà calculé dans la section 3.2.2 sur le pendule rigide, voir (3.16). On a donc

$$I_1 = I_2 = \frac{M\ell^2}{12}, \quad I_3 = 0. \quad (3.38)$$

Le vecteur vitesse angulaire est $\vec{\Omega} = \Omega \vec{e}_z$, mais pour appliquer (3.36) on a besoin de l'écrire dans le repère $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$. On a $\vec{e}_z = \cos \alpha \vec{e}_1 + \sin \alpha \vec{e}_3$, et donc

$$\vec{\Omega} = \vec{\Omega}_1 + \vec{\Omega}_2 + \vec{\Omega}_3 \quad \text{avec} \quad \vec{\Omega}_1 = \Omega \cos \alpha \vec{e}_1, \quad \vec{\Omega}_2 = \vec{0}, \quad \vec{\Omega}_3 = \Omega \sin \alpha \vec{e}_3. \quad (3.39)$$

Avec (3.36) on obtient alors

$$\vec{L}' = I_1 \vec{\Omega}_1 + I_2 \vec{\Omega}_2 + I_3 \vec{\Omega}_3 = \frac{M\ell^2}{12} \Omega \cos \alpha \vec{e}_1. \quad (3.40)$$

D'après (3.31), le moment cinétique du système par rapport à n'importe quel point fixe est égal à $\vec{L}'$ puisque le point C est immobile. En particulier, $\vec{L}_O = \vec{L}'$.

On remarque que le moment cinétique $\vec{L}'$ n'est pas colinéaire au vecteur vitesse angulaire. Par ailleurs, $\vec{L}'$ dépend du temps (puisque le vecteur $\vec{e}_1$ tourne avec le système), alors que le vecteur $\vec{\Omega}$ est constant.

Un petit calcul (à faire en exercice) donne

$$\frac{d}{dt} \vec{e}_1 = \vec{\Omega} \wedge \vec{e}_1 = \Omega \sin \alpha \vec{e}_2, \quad (3.41)$$

et donc

$$\frac{d\vec{L}'}{dt} = \frac{M\ell^2}{12} \Omega^2 \cos \alpha \sin \alpha \vec{e}_2. \quad (3.42)$$

Cette quantité doit être égale à la somme $\vec{\mathcal{M}}$ des moment des forces extérieures. Comme la somme des forces extérieures est nulle, la somme des moments ne dépend pas du point de référence, mais on la calcule à partir du point O . On trouve (à faire en exercice)

$$\vec{\mathcal{M}} = \frac{a}{2} (R_{\text{droite}} - R_{\text{gauche}}) \vec{e}_x + \frac{a}{2} (R_{\text{derrière}} - R_{\text{devant}}) \vec{e}_y. \quad (3.43)$$

Avec R_{devant} la projection selon $\vec{e}_z$ de la somme des deux forces de réaction des pieds de devant, et idem pour $R_{\text{derrière}}$.

Les expressions (3.42) et (3.43) sont vraies à tout instant. À l'instant de la figure, quand la tige est dans le plan (Oyz) , on a $\vec{e}_2 = \vec{e}_x$. On en déduit que, à cet instant,

$$R_{\text{derrière}} = R_{\text{devant}}, \quad \text{et} \quad \frac{a}{2} (R_{\text{droite}} - R_{\text{gauche}}) = \frac{M\ell^2}{12} \Omega^2 \cos \alpha \sin \alpha. \quad (3.44)$$

et on peut en déduire R_{droite} et R_{gauche} avec l'aide de (3.37).

Le calcul donne que si $M\ell^2 \Omega^2 \cos \alpha \sin \alpha > 6a(M + M')g$, alors $R_{\text{gauche}} < 0$ ce qui n'est pas possible (à moins que les pieds soient collés à la table). On en déduit que la boîte doit bouger, et certains pieds doivent décoller, si Ω est trop grand.

3.4 Un exemple difficile : la toupie

On considère une toupie constitué d'une sphère de masse M et d'une tige de masse négligeable ; voir la figure 3.6 pour les notations. La distance entre une extrémité de la tige et le centre de la sphère est notée $BC = \ell$. On suppose que le point B glisse sur le sol sans frotter.

Ce problème peut se résoudre en appliquant le principe fondamental de la dynamique et le théorème de l'énergie cinétique, mais il est (légèrement) plus simple à traiter avec les équations de Lagrange.

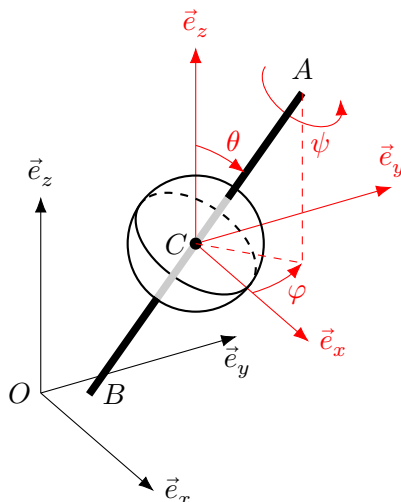


FIGURE 3.6 – Toupie sphérique.

Pour repérer l'orientation de la toupie dans l'espace, on utilise les *angles d'Euler*¹. Les deux premiers angles d'Euler θ et φ sont les angles habituels des coordonnées sphériques : ils repèrent ici l'extrémité A de la toupie dans le repère $(C, \vec{e}_x, \vec{e}_y, \vec{e}_z)$. (Les trois vecteurs $\vec{e}_x, \vec{e}_y, \vec{e}_z$ sont fixes, ils ne sont pas liés à la toupie.) Le troisième angle d'Euler est l'angle ψ , qui mesure de combien la toupie a tourné autour de son axe (par rapport à un point de départ arbitraire).

Les trois angles d'Euler donnent l'orientation de la toupie. Il reste à donner la position. *A priori*, on aurait besoin de donner trois nombres, disons les coordonnées X, Y, Z du point C . Mais à cause de la contrainte que B touche le sol, on sait que $Z = \ell \cos \theta$. On a donc juste besoin de donner X et Y .

On a donc finalement $s = 5$ variables : θ, φ, ψ, X et Y . Écrivons le lagrangien, en commençant par les termes faciles.

L'énergie potentielle de pesanteur est

$$U = MgZ = Mg\ell \cos \theta. \quad (3.45)$$

L'énergie cinétique de translation est

$$T_{\text{trans}} = \frac{1}{2}M(\dot{X}^2 + \dot{Y}^2 + \dot{Z}^2) = \frac{1}{2}M(\dot{X}^2 + \dot{Y}^2 + \ell^2 \sin^2 \theta \dot{\theta}^2). \quad (3.46)$$

Reste à calculer l'énergie cinétique de rotation.

On néglige la masse de la tige ; tout se passe comme si la toupie était une sphère homogène. Clairement, pour une sphère homogène, tous les axes sont des axes de symétrie, et donc tous les axes sont des axes principaux. On a donc, tout simplement,

$$T_{\text{rot}} = \frac{1}{2}I\vec{\Omega}^2, \quad (3.47)$$

1. Il y a plusieurs variantes pour définir les angles d'Euler, on peut rencontrer des définitions légèrement différentes.

où I , calculé en TD, est le moment cinétique de la sphère homogène par rapport à n'importe quel axe passant par son centre.

Il reste à calculer le vecteur vitesse angulaire $\vec{\Omega}$.

- Lorsque φ varie, cela correspond à une rotation le long de l'axe $\vec{e}_z$ et donc à un vecteur rotation $\dot{\varphi} \vec{e}_z$.
- Lorsque ψ varie, cela correspond à une rotation autour de l'axe $\vec{e}_r$ et donc à un vecteur rotation $\dot{\psi} \vec{e}_r$. (On note $(\vec{e}_r, \vec{e}_\theta, \vec{e}_\varphi)$ les vecteurs habituels des coordonnées sphériques, non représentées sur la figure.)
- Lorsque θ varie, cela correspond à une rotation autour de l'axe $\vec{e}_\varphi$ et donc à un vecteur rotation $\dot{\theta} \vec{e}_\varphi$. (Celui-là est peut-être plus difficile à voir, mais la rotation associée à l'angle θ se fait bien dans le plan contenant $\vec{e}_r$ et $\vec{e}_\theta$, et donc d'axe $\vec{e}_\varphi$, perpendiculaire à ce plan.)

En combinant ces trois contributions, on obtient le vecteur rotation

$$\vec{\Omega} = \dot{\varphi} \vec{e}_z + \dot{\psi} \vec{e}_r + \dot{\theta} \vec{e}_\varphi, \quad (3.48)$$

et

$$\vec{\Omega}^2 = \dot{\varphi}^2 + \dot{\psi}^2 + \dot{\theta}^2 + 2\dot{\varphi}\dot{\psi} \vec{e}_z \cdot \vec{e}_r = \dot{\varphi}^2 + \dot{\psi}^2 + \dot{\theta}^2 + 2\cos\theta \dot{\varphi}\dot{\psi} \quad (3.49)$$

(Remarque : attention au double produit ; les trois vecteurs $\vec{e}_z, \vec{e}_r, \vec{e}_\varphi$ ne forment pas un repère orthonormé. On a quand même $\vec{e}_z \cdot \vec{e}_\varphi = 0$ et $\vec{e}_r \cdot \vec{e}_\varphi = 0$, et donc il ne reste que le double produit avec $\vec{e}_r \cdot \vec{e}_z$.)

En combinant ce résultat avec les précédents, on peut enfin écrire le lagrangien de la toupie :

$$\mathcal{L} = \frac{1}{2}M(\dot{X}^2 + \dot{Y}^2 + \ell^2 \sin^2 \theta \dot{\theta}^2) + \frac{1}{2}I(\dot{\varphi}^2 + \dot{\psi}^2 + \dot{\theta}^2 + 2\cos\theta \dot{\varphi}\dot{\psi}) - Mgl \cos \theta. \quad (3.50)$$

On remarque que ce lagrangien dépend des cinq vitesses généralisées $\dot{X}, \dot{Y}, \dot{\varphi}, \dot{\psi}$ et $\dot{\theta}$, mais ne dépend que d'une seule coordonnée généralisée, θ . D'après les résultats de la section 2.3.4, cela signifie que les quatre impulsions généralisées $p_X = \partial\mathcal{L}/\partial\dot{X}$, $p_Y = \partial\mathcal{L}/\partial\dot{Y}$, $p_\varphi = \partial\mathcal{L}/\partial\dot{\varphi}$ et $p_\psi = \partial\mathcal{L}/\partial\dot{\psi}$ sont des quantités conservées au cours du mouvement. Après simplification par M ou I , on trouve donc les quatre constantes du mouvement suivantes :

$$\frac{p_X}{M} = \dot{X}, \quad \frac{p_Y}{M} = \dot{Y}, \quad \frac{p_\psi}{I} = \dot{\psi} + \dot{\varphi} \cos \theta, \quad \frac{p_\varphi}{I} = \dot{\varphi} + \dot{\psi} \cos \theta. \quad (3.51)$$

Il n'est pas étonnant que $\dot{X}$ et $\dot{Y}$ soient des constantes du mouvement ; si la toupie (sans frottement, rappelons-le), est lancée sur la table, elle conserve sa vitesse initiale. Le fait que $\dot{\psi} + \dot{\varphi} \cos \theta$ et $\dot{\varphi} + \dot{\psi} \cos \theta$ soient des quantités conservées est moins évident.

Écrivons maintenant l'équation de Lagrange associée à la variable θ ; après simplification :

$$M\ell^2 \sin \theta \cos \theta \dot{\theta}^2 + (M\ell^2 \sin^2 \theta + I)\ddot{\theta} = (Mgl - I\dot{\varphi}\dot{\psi}) \sin \theta. \quad (3.52)$$

Cette équation est horriblement compliquée, mais on voit qu'il existe des solutions où θ est une constante à condition que

$$\dot{\varphi}\dot{\psi} = \frac{Mg\ell}{I} \quad (3.53)$$

Plus la toupie tourne vite sur elle-même ($\dot{\psi}$ grand), plus le mouvement de précession ($\dot{\varphi}$) est faible.

3.5 Tenseur d'inertie

(Les notions abordées dans cette section ne seront pas demandées en examen.)

3.5.1 Calcul général de l'énergie cinétique de rotation

Reprenons le calcul de l'énergie cinétique d'un solide en rotation à partir de (3.7). Pour le point A , on choisit de prendre le centre de masse C , ce qui fait que le deuxième terme est nul, et on obtient

$$T = \frac{1}{2}MV_C^2 + T_{\text{rot}} \quad \text{avec} \quad T_{\text{rot}} = \frac{1}{2} \int dm \|\vec{\Omega} \wedge \vec{r}\|^2. \quad (3.54)$$

Dans cette expression, on avait écrit $\|\vec{\Omega} \wedge \vec{r}\|^2 = \Omega^2 r^2 \sin^2 \theta = \Omega^2 r_{\perp}^2$, voir (3.9), et on avait ainsi obtenu (3.11) et (3.12).

Nous allons choisir un repère orthonormé $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ liée au solide, c'est-à-dire telle que les vecteurs tournent avec le solide (voir la figure 3.7 pour un exemple), et nous allons calculer T_{rot} dans ce repère. Pour l'instant, on choisit le repère $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ de manière arbitraire, sans se préoccuper des axes principaux du solide.

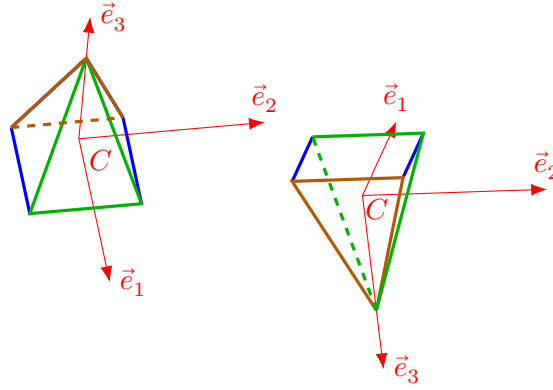


FIGURE 3.7 – Une pyramide en mouvement, à deux instants de sa trajectoire. Les axes $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ sont liés au solide ; par exemple, l'axe $\vec{e}_3$ va du centre C vers la pointe de la pyramide, et l'axe $\vec{e}_1$ sort par la face verte.

Écrivons $\vec{r}$ et $\vec{\Omega}$ dans le repère $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$:

$$\vec{r} = x_1 \vec{e}_1 + x_2 \vec{e}_2 + x_3 \vec{e}_3, \quad \vec{\Omega} = \Omega_1 \vec{e}_1 + \Omega_2 \vec{e}_2 + \Omega_3 \vec{e}_3. \quad (3.55)$$

On pourrait calculer en coordonnées le vecteur $\vec{r} \wedge \vec{\Omega}$ et remplacer dans (3.54), mais les calculs sont plus simples si on fait la transformation suivante : avec θ l'angle entre $\vec{r}$ et $\vec{\Omega}$, on a $\|\vec{\Omega} \wedge \vec{r}\|^2 = \Omega^2 r^2 \sin^2 \theta = \Omega^2 r^2 (1 - \cos^2 \theta) = \Omega^2 r^2 - \Omega^2 r^2 \cos^2 \theta = \Omega^2 r^2 - (\vec{\Omega} \cdot \vec{r})^2$, et donc

$$T_{\text{rot}} = \frac{1}{2} \int dm \left[\Omega^2 r^2 - (\vec{\Omega} \cdot \vec{r})^2 \right]. \quad (3.56)$$

Il ne reste plus qu'à remplacer Ω^2 , r^2 et $\vec{\Omega} \cdot \vec{r}$ par leurs expressions en fonction des coordonnées :

$$\Omega^2 = \Omega_1^2 + \Omega_2^2 + \Omega_3^2, \quad r^2 = x_1^2 + x_2^2 + x_3^2, \quad \vec{\Omega} \cdot \vec{r} = \Omega_1 x_1 + \Omega_2 x_2 + \Omega_3 x_3. \quad (3.57)$$

On peut aussi utiliser un symbole somme :

$$\Omega^2 = \sum_{i=1}^3 \Omega_i^2, \quad r^2 = \sum_{i=1}^3 x_i^2, \quad \vec{\Omega} \cdot \vec{r} = \sum_{i=1}^3 \Omega_i x_i. \quad (3.58)$$

En fait, ce genre de calcul est habituellement mené en utilisant la *convention d'Einstein* qui consiste à supprimer les symboles $\sum$. Si un indice apparaît deux fois dans un terme, on considère que, implicitement, on somme sur cet indice et, alors, au lieu d'écrire (3.58), on écrit plus simplement :

$$\Omega^2 = \Omega_i^2, \quad r^2 = x_i^2, \quad \vec{\Omega} \cdot \vec{r} = \Omega_i x_i \quad (\text{avec la convention d'Einstein}). \quad (3.59)$$

Dans ces expressions, le nom de l'indice est arbitraire ; au lieu d'écrire $\vec{\Omega} \cdot \vec{r} = \Omega_i x_i$, on pourrait aussi bien écrire $\vec{\Omega} \cdot \vec{r} = \Omega_j x_j$ ou $r^2 = x_j^2$. Il faut juste faire attention à ne pas réutiliser le même indice dans un produit. Ainsi, on peut écrire

$$(\vec{\Omega} \cdot \vec{r})^2 = (\Omega_i x_i)(\Omega_j x_j) = x_i x_j \Omega_i \Omega_j, \quad \Omega^2 r^2 = (\Omega_i \Omega_i)(x_j x_j) = x_j x_j \Omega_i \Omega_i, \quad (3.60)$$

De manière implicite, dans les deux cas, on somme sur i et sur j . Il était important d'utiliser j plutôt que i dans un facteur ; si on ne l'avait pas fait, on aurait obtenu $x_i x_i \Omega_i \Omega_i$, et on n'aurait pas su ce que cela voulait dire !

Dans la première expression de (3.60), on a le facteur $\Omega_i \Omega_j$. Pour des raisons qui seront claires plus loin, on aimerait avoir le même facteur dans la deuxième expression de (3.60). On peut le faire en introduisant le *symbole de Kronecker*

$$\delta_{ij} = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{si } i \neq j. \end{cases} \quad (3.61)$$

On peut alors écrire

$$\Omega_i = \delta_{ij} \Omega_j, \quad (3.62)$$

(implicitement, on somme sur les j dans le terme de droite parce qu'il y en a deux, mais on ne somme pas sur les i , ni à gauche ni à droite). Puis :

$$\Omega^2 = \Omega_i^2 = \delta_{ij} \Omega_i \Omega_j, \quad (3.63)$$

(implicitement, on somme sur les i dans l'expression du milieu, et on somme sur les i et les j à droite).

Finalement, à partir de (3.56), on obtient l'expression suivante de T_{rot} :

$$T_{\text{rot}} = \frac{1}{2} \int dm [\delta_{ij} \Omega_i \Omega_j x_k^2 - x_i x_j \Omega_i \Omega_j]. \quad (3.64)$$

(Implicitement, le premier terme dans les crochets est sommé sur i , j et k , et le deuxième terme est sommé sur i et j .) On peut alors mettre $\Omega_i \Omega_j$ en facteur,

$$T_{\text{rot}} = \frac{1}{2} \Omega_i \Omega_j \int dm (\delta_{ij} x_k^2 - x_i x_j), \quad (3.65)$$

et donc :

$$T_{\text{rot}} = \frac{1}{2} I_{ij} \Omega_i \Omega_j \quad \text{avec } I_{ij} = \int dm (\delta_{ij} x_k^2 - x_i x_j). \quad (3.66)$$

L'objet I_{ij} a deux indices qui varient de 1 à 3, et peut être représenté par une matrice 3×3 . On l'appelle la *matrice d'inertie*. Cette matrice est clairement symétrique ($I_{ij} = I_{ji}$) et s'écrit, explicitement,

$$I = \begin{pmatrix} \int dm (x_2^2 + x_3^2) & - \int dm x_1 x_2 & - \int dm x_1 x_3 \\ - \int dm x_2 x_1 & \int dm (x_1^2 + x_3^2) & - \int dm x_2 x_3 \\ - \int dm x_3 x_1 & - \int dm x_3 x_2 & \int dm (x_1^2 + x_2^2) \end{pmatrix}. \quad (3.67)$$

Ainsi, par exemple, si le solide tourne autour de l'axe $\vec{e}_3$, c'est-à-dire si $\vec{\Omega} = \Omega \vec{e}_3$, l'énergie cinétique de rotation est $\frac{1}{2} I_{33} \Omega^2$ avec $I_{33} = \int dm (x_1^2 + x_2^2)$. On reconnaît l'expression (3.11) de I_{Δ} puisque $x_1^2 + x_2^2$ est justement le terme $r_{\perp}^2$, carré de la distance à l'axe de rotation, qui est ici l'axe $\vec{e}_3$.

Comme les vecteurs $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ sont liés au solide, la matrice d'inertie ne dépend pas de la position ou de l'orientation du solide ; on peut la calculer une bonne fois pour toute (pour un solide donnée) et l'utiliser pour calculer T_{rot} pour un vecteur $\vec{\Omega}$ quelconque.

3.5.2 Calcul général du moment cinétique

On peut calculer le moment cinétique intrinsèque $\vec{L}'$ en utilisant le même repère $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ que dans la partie précédente. On repart de l'expression (3.32), et on utilise l'égalité $\vec{a} \wedge (\vec{b} \wedge \vec{c}) = (\vec{a} \cdot \vec{c})\vec{b} - (\vec{a} \cdot \vec{b})\vec{c}$. On obtient

$$\vec{L}' = \int dm \vec{r} \wedge (\vec{\Omega} \wedge \vec{r}) = \int dm [\vec{r}^2 \vec{\Omega} - (\vec{r} \cdot \vec{\Omega})\vec{r}]. \quad (3.68)$$

Le moment cinétique est un vecteur, que l'on projette sur le même repère orthonormé $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ attaché au solide que l'on avait utilisé pour $\vec{\Omega}$ et $\vec{r}$: on avait écrit $\vec{\Omega} = \Omega_1 \vec{e}_1 + \Omega_2 \vec{e}_2 + \Omega_3 \vec{e}_3$ et $\vec{r} = x_1 \vec{e}_1 + x_2 \vec{e}_2 + x_3 \vec{e}_3$; maintenant, on écrit aussi $\vec{L}' = L'_1 \vec{e}_1 + L'_2 \vec{e}_2 + L'_3 \vec{e}_3$.

On projette (3.68) sur l'axe $\vec{e}_i$:

$$L'_i = \int dm [\vec{r}^2 \Omega_i - (\vec{r} \cdot \vec{\Omega})r_i]. \quad (3.69)$$

(Il n'y a aucune somme sur i .) Puis, en utilisant la convention d'Einstein,

$$L'_i = \int dm [x_k^2 \Omega_i - x_j \Omega_j x_i] = \int dm [\delta_{ij} x_k^2 - x_i x_j] \Omega_j = I_{ij} \Omega_j, \quad (3.70)$$

où l'on a utilisé (3.62) et où on a reconnu l'expression de I_{ij} , voir (3.66). On voit donc que la matrice d'inertie I_{ij} permet de calculer à la fois l'énergie cinétique et le moment cinétique d'un solide :

$$T_{\text{rot}} = \frac{1}{2} I_{ij} \Omega_i \Omega_j, \quad L'_i = I_{ij} \Omega_j. \quad (3.71)$$

En particulier, on remarque que

$$\vec{L}' \cdot \vec{\Omega} = L'_i \Omega_i = \Omega_i I_{ij} \Omega_j = 2T_{\text{rot}}. \quad (3.72)$$

3.5.3 Notion de tenseur d'inertie

Dans cette section, nous avons commencé par choisir de manière arbitraire un repère orthonormé $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ lié au solide. Si nous choisissons un autre repère orthonormé $(\vec{e}'_1, \vec{e}'_2, \vec{e}'_3)$, alors le vecteur rotation $\vec{\Omega}$ aura des composantes $\tilde{\Omega}_i$ différentes dans cette nouvelle base. On aura également une nouvelle matrice d'inertie $\tilde{I}_{ij}$ mais, évidemment, l'énergie cinétique et le moment cinétique demeurent inchangés.

$$T_{\text{rot}} = \frac{1}{2} I_{ij} \Omega_i \Omega_j = \frac{1}{2} \tilde{I}_{ij} \tilde{\Omega}_i \tilde{\Omega}_j, \quad \vec{L}' = L'_i \vec{e}_i = I_{ij} \Omega_j \vec{e}_i = \tilde{I}_{ij} \tilde{\Omega}_j \vec{e}_i = \tilde{L}'_i \vec{e}_i. \quad (3.73)$$

Des quantités telles que T_{rot} , $\vec{L}'$, $\vec{\Omega}$ sont des quantités intrinsèques, indépendantes du système de coordonnées, alors que les Ω_i , les L'_i ou les I_{ij} dépendent du système de coordonnées. On dit que les Ω_i représentent le vecteur $\vec{\Omega}$ dans un certain repère.

De la même manière, la matrice I_{ij} n'est que l'écriture dans un certain repère d'un objet intrinsèque noté I ou parfois $\underline{I}$ qu'on appelle le *tenseur d'inertie*. On écrira alors, sans faire référence au système de coordonnées,

$$\vec{L}' = I\vec{\Omega}, \quad T_{\text{rot}} = \frac{1}{2} \vec{\Omega} \cdot (I\vec{\Omega}). \quad (3.74)$$

(Attention, I n'est pas un scalaire, $\vec{L}'$ n'est pas colinéaire à $\vec{\Omega}$ et $\vec{\Omega} \cdot (I\vec{\Omega})$ n'est pas égal à $I\vec{\Omega}^2$.)

Mathématiquement, le tenseur d'inertie I représente une application linéaire qui transforme le vecteur $\vec{\Omega}$ en un vecteur $\vec{L}'$. (L'application est linéaire parce que $I(a\vec{\Omega}_1 + b\vec{\Omega}_2) = aI\vec{\Omega}_1 + bI\vec{\Omega}_2$.)²

2. Attention, le vocabulaire autour des tenseurs dépend des auteurs : certains préfèrent appeler «opérateur d'inertie» l'objet I dans $\vec{L}' = I\vec{\Omega}$ qui transforme un vecteur en un autre vecteur, et réservent le terme de tenseur à une application qui transforme deux (ou plus...) vecteurs en un scalaire, notant $I(\vec{A}, \vec{B})$ ou $I\vec{A}\vec{B}$ la quantité $I_{ij}A_iB_j$; ainsi, avec cette notation, on pourrait écrire $T_{\text{rot}} = \frac{1}{2} I\vec{\Omega}\vec{\Omega}$.

Dans un repère orthonormé $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ donné, cette application linéaire s'écrit à l'aide d'une matrice I_{ij} qui transforme les coordonnées du vecteur $\vec{\Omega}$ pour donner les coordonnées du vecteur $\vec{L}'$ dans le même repère.

On a vu que la matrice I_{ij} était symétrique ; on dit alors que le tenseur I est symétrique, ce qui implique (à faire en exercice) $\vec{A} \cdot (I\vec{B}) = (I\vec{A}) \cdot \vec{B}$.

3.5.4 Axes principaux

On a dit que Δ était un axe principal si, lors d'une rotation autour de Δ , le moment cinétique était colinéaire au vecteur rotation, et donc dirigé selon Δ . Autrement dit, si on note $\vec{\Omega} = \Omega \vec{n}$ avec $\vec{n}$ un vecteur directeur de Δ , alors $\vec{L}' = L' \vec{n}$. On a alors

$$L' \vec{n} = I(\Omega \vec{n}) \quad \text{c'est-à-dire} \quad I\vec{n} = \frac{L'}{\Omega} \vec{n}. \quad (3.75)$$

on dit alors que $\vec{n}$ est un vecteur propre du tenseur I , dont la valeur propre correspondante est L'/Ω . On utilise le même vocabulaire que pour les matrices parce que c'est le même concept. D'ailleurs, dans un repère orthonormé $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ donné, le vecteur $\vec{n}$ écrit dans ce repère est un vecteur propre de la matrice I_{ij} qui représente le tenseur I dans ce repère.

Un théorème d'algèbre très important affirme que

Toute matrice réelle symétrique est diagonalisable, avec des vecteurs propres orthogonaux deux à deux.

(Le fait que la matrice est réelle est important, mais il existe une version de ce théorème pour des matrices complexes qui a une grande importance en mécanique quantique.)

Cela implique (il faut prendre le temps d'y réfléchir un peu...) que l'on peut choisir un repère orthonormé $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ tel que les trois vecteurs de la base sont des vecteurs propres du tenseur. (C'est-à-dire : tels que ces trois vecteurs donnent les directions de trois axes principaux du solide.)

Autrement dit, il existe un repère orthonormé $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ et trois scalaires I_1, I_2, I_3 tels que

$$I\vec{e}_i = I_i \vec{e}_i, \quad \text{ou encore} \quad I_{ij} = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix}. \quad (3.76)$$

On retrouve alors facilement les expressions (3.36).

Deuxième partie

Relativité restreinte

Chapitre 4

Postulats de la relativité restreinte et transformations de Lorentz

4.1 Relativité galiléenne — rappels

Cette première partie consiste en un rappel des notions de changements de référentiel et de la relativité en mécanique dite newtonienne ou classique. Ces notions sont *a priori* déjà connues, et doivent être maîtrisées pour la suite du cours.

4.1.1 Mécanique newtonienne et transformations de Galilée

On appelle un référentiel d’inertie, ou galiléen, un référentiel dans lequel le mouvement des corps obéit au principe d’inertie, à savoir que si aucune force ne s’applique à une masse m donnée, celle-ci est immobile ou se meut à vitesse constante.

La mécanique Newtonienne suppose que de tels référentiels existent. Dans ces référentiels s’applique le **principe fondamental de la dynamique (PFD)**, voir **appendice A**.

Un autre postulat de la mécanique classique est ce qu’on appelle le principe de relativité :

Postulat de relativité :

Tout référentiel en translation uniforme par rapport à un référentiel d’inertie est également un référentiel d’inertie.

Les lois de la physique sont identiques dans tous les référentiels d’inertie.

On peut associer à chaque référentiel un jeu de coordonnées d’espace (par exemple (x, y, z)) et de temps (t) qui serviront à décrire des phénomènes physiques dans le référentiel associé (par exemple le mouvement d’une masse donnée) et on peut passer d’une description dans un référentiel à une description dans un autre en changeant les coordonnées avec lesquelles le système est décrit.

On appelle “transformations” les opérations appliquées à ces jeux de coordonnées pour passer d’un référentiel à un autre, c’est-à-dire exprimer les coordonnées du référentiel 1 en fonction de celles du référentiel 2.

En *mécanique newtonienne*, il est admis que les transformations reliant deux référentiels en translation uniforme qui satisfont au postulat de relativité sont les *transformations de Galilée* que nous rappelons ci dessous.

Soient $\mathcal{R}$ un référentiel et $\mathcal{R}'$ un autre référentiel animé d’une vitesse $\vec{V}_{\mathcal{R}'/\mathcal{R}}$ par rapport à $\mathcal{R}$. On associe à chacun de ses référentiels les jeux de coordonnées (x, y, z, t) et (x', y', z', t') . On choisit arbitrairement les axes x et x' comme étant confondus, y et y' et z et z' comme étant parallèles. En terme de vecteur unitaires, on peut alors écrire que $\vec{u}_x = \vec{u}_{x'}$, $\vec{u}_y = \vec{u}_{y'}$ et $\vec{u}_z = \vec{u}_{z'}$. On écrit alors la vitesse de déplacement de $\mathcal{R}'$ par rapport à $\mathcal{R}$ comme $\vec{V}_{\mathcal{R}'/\mathcal{R}} = V\vec{u}_x$, où V est

une grandeur positive (*cf.* figure 1.1). Enfin, on considère que à $t = 0$, les origines O et O' de $\mathcal{R}$ et $\mathcal{R}'$ sont confondues.

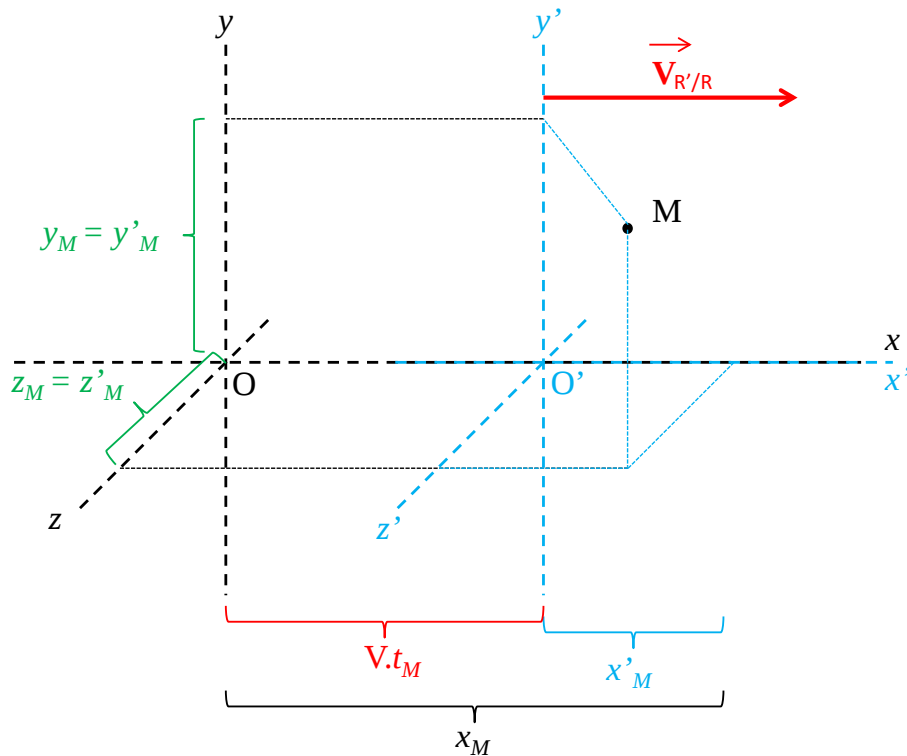


FIGURE 4.1 – Schéma à l'instant t , des coordonnées d'un point M quelconque dans les deux référentiels $\mathcal{R}$ et $\mathcal{R}'$.

On considère un *événement* : une lampe s'allume, M. Dupont s'endort, etc. Tout événement est caractérisé par sa position spatiale de coordonnées (x, y, z) , et par sa date t , dans un référentiel $\mathcal{R}$. On notera donc (x, y, z, t) les coordonnées spatio-temporelles de cet événement. Dans un autre référentiel $\mathcal{R}'$, les coordonnées du même événement seront différentes ; notons les (x', y', z', t') .

Exprimons maintenant (x', y', z', t') en fonction de (x, y, z, t) dans le cadre de la mécanique newtonienne ; tout d'abord, le temps s'écoule de la même façon quel que soit le référentiel considéré, et donc $t = t'$. Ensuite, $\mathcal{R}'$ s'étant déplacé uniquement suivant x , on peut immédiatement conclure que $y' = y$ et que $z' = z$. Enfin, On peut noter que, pendant le temps t le référentiel $\mathcal{R}'$ s'est déplacé suivant l'axe x de $V \times t$ car $V = \text{cste}$. Ainsi, $x' = x - Vt$.

En résumé, pour tout événement, on a :

$$\begin{cases} x' = x - Vt \\ y' = y \\ z' = z \\ t' = t \end{cases} \quad (\text{ce résultat est incorrect en relativité restreinte})$$

Ces équations sont appelées les **transformations de Galilée**. On peut alors exprimer la vitesse du point M dans les deux référentiels et ainsi établir les transformations des vitesses.

Dans $\mathcal{R}$:

$$\vec{v}_M = v_x \vec{u}_x + v_y \vec{u}_y + v_z \vec{u}_z = \frac{dx}{dt} \vec{u}_x + \frac{dy}{dt} \vec{u}_y + \frac{dz}{dt} \vec{u}_z$$

De même, dans $\mathcal{R}'$, par définition :

$$\vec{v}'_M = \frac{dx'}{dt'} \vec{u}_x + \frac{dy'}{dt'} \vec{u}_y + \frac{dz'}{dt'} \vec{u}_z$$

On peut alors utiliser les transformations de Galilée pour exprimer $\vec{v}'_M$ en fonction de $\vec{v}_M$:

$$\begin{aligned} \frac{dy'}{dt'} &= \frac{dy}{dt} \quad \text{et} \quad \frac{dz'}{dt'} = \frac{dz}{dt} \\ \frac{dx'}{dt'} &= \frac{d}{dt}(x - Vt) = \frac{dx}{dt} - V \quad \text{car} \quad V = \text{cste} \end{aligned}$$

Ainsi :

$$\vec{v}'_M = \frac{dx}{dt} \vec{u}_x + \frac{dy}{dt} \vec{u}_y + \frac{dz}{dt} \vec{u}_z - V \vec{u}_x$$

$$\vec{v}'_M = \vec{v}_M - \vec{V}_{\mathcal{R}'/\mathcal{R}} \quad (\text{incorrect en relativité restreinte})$$

On trouve le **principe d'additivité des vitesses de Galilée**, valable en mécanique classique et qui jouera un rôle essentiel dans la remise en cause des transformations de Galilée par la théorie de la relativité restreinte.

Enfin, on peut vouloir transformer l'accélération du point M en utilisant les transformations de Galilée. On trouve :

$$\vec{a}'_M = \frac{d\vec{v}'_M}{dt'} = \frac{d\vec{v}_M}{dt} - \frac{d\vec{V}_{\mathcal{R}'/\mathcal{R}}}{dt} = \vec{a}_M - \vec{0} \quad \text{car } \vec{V}_{\mathcal{R}'/\mathcal{R}} \text{ est constant,}$$

et donc

$$\vec{a}'_M = \vec{a}_M \quad (\text{incorrect en relativité restreinte}).$$

L'accélération est donc invariante (c'est la même dans $\mathcal{R}$ ou $\mathcal{R}'$) par les transformations de Galilée.

4.1.2 Signification du postulat de relativité

Dans le postulat de relativité, que veut-on dire par identiques dans “lois de la physique *identiques* dans deux référentiels $\mathcal{R}$ et $\mathcal{R}'$ en translation uniforme” ?

On veut dire que les lois de la physique et donc les équations les décrivant s'écrivent exactement de la même manière dans le système primé et dans le système non primé.

Une conséquence de cela est qu'il n'est pas possible à un observateur A enfermé dans un vaisseau spatial sans fenêtre de faire des expériences qui lui permettent de déterminer s'il est en mouvement uniforme par rapport à un observateur B situé en dehors du vaisseau. En ouvrant les fenêtres du vaisseau, l'observateur A pourra certainement voir s'il est en mouvement par rapport à l'observateur B mais ce mouvement est nécessairement relatif car il n'y a aucune expérience permettant de trancher si c'est A qui bouge par rapport à B ou B qui bouge par rapport à A. Le principe de relativité stipule qu'il n'y a pas de vitesse absolue, c'est-à-dire de vitesse par rapport à un référentiel privilégié.

En mécanique newtonienne, on admet que les transformations de Galilée décrivent correctement les changements de référentiels et que le PFD est valable dans les référentiels d'inertie.

Si on reprend le cas des deux référentiels précédents, et que l'on applique le PFD à un point M de masse m , on peut écrire dans $\mathcal{R}$ (qu'on suppose être un référentiel d'inertie) :

$$\vec{F}_{\text{ext}} = m \vec{a}_M.$$

D'après le principe de relativité on doit donc écrire dans $\mathcal{R}'$:

$$\vec{F}'_{\text{ext}} = m \vec{a}'_M.$$

(On a supposé que la masse était identique dans les deux référentiels.)

Or, on a montré plus haut que

$$\vec{a}'_M = \vec{a}_M.$$

On peut écrire que

$$\vec{F}'_{\text{ext}} = \vec{F}_{\text{ext}}.$$

La force qui s'exerce sur M est identique (en norme et en direction) dans les référentiels $\mathcal{R}$ et $\mathcal{R}'$. C'est tout du moins la prédiction de la mécanique newtonienne. Prédiction que l'on peut tester expérimentalement.

4.2 Problèmes liés à la propagation de la lumière en mécanique classique

4.2.1 Nature ondulatoire de la lumière

En 1864, Maxwell publie son étude sur la nature ondulatoire de la lumière, décrite comme une onde électromagnétique. Par exemple, dans le vide, la lumière peut être décrite par un champ électrique $\vec{E}$ et un champ magnétique $\vec{B}$, tous les deux oscillants à la même fréquence dans une direction transverse à la direction de propagation de l'onde que l'on repère par le vecteur d'onde $\vec{k}$ (cf. figure 1.2).

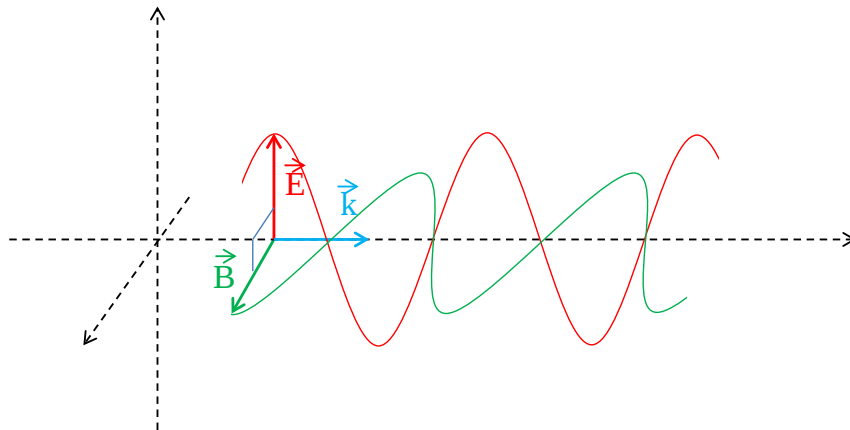


FIGURE 4.2 – Représentation d'une onde électromagnétique dans le vide.

L'onde est régie par quatre équations, les *équations de Maxwell* qui couplent les champs $\vec{E}$ et $\vec{B}$ entre eux et avec les propriétés du milieu dans lequel l'onde se propage (permittivité électrique, perméabilité magnétique, densité de charges libres... etc...). Au final, pour décrire l'onde lumineuse, on se contente souvent de décrire le champ $\vec{E}$, sachant que $\vec{B}$ pourra être déduit de $\vec{E}$ grâce aux équations de Maxwell. **Ce n'est pas le but du cours de rentrer dans le détail des lois de l'électromagnétisme, qui sont vues par ailleurs dans une autre UE. On rappellera juste ici quelques résultats que l'on confrontera à la relativité galiléenne et au principe de relativité.**

L'un des résultats de la théorie de Maxwell est que l'onde électromagnétique se propage, dans le vide, à la célérité, notée c , donnée par :

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}} \approx 3 \cdot 10^8 \text{ m} \cdot \text{s}^{-1},$$

où ϵ_0 est la permittivité du vide et μ_0 la perméabilité du vide, deux grandeurs associées au milieu de propagation, et donc indépendantes du référentiel de description.

Ce qui est remarquable ici, c'est que, contrairement aux ondes mécaniques (vibrations sur une corde, onde sonore...), l'onde électromagnétique, donc la lumière, peut se propager dans le vide.

4.2.2 Hypothèse de l'éther

Si on se place dans le cadre des transformations de Galilée, l'onde lumineuse ne peut se propager à la vitesse c que dans un unique référentiel d'inertie $\mathcal{R}_0$. Dans tout autre référentiel d'inertie $\mathcal{R}$ animé d'une vitesse V par rapport à $\mathcal{R}_0$, le principe d'additivité des vitesses s'applique et donc la vitesse de propagation doit obligatoirement être différente de c . Comme la vitesse de propagation de la lumière dans le vide est une conséquence directe de la formulation des équations de Maxwell, cela implique que ces équations, et ainsi les lois de l'électromagnétisme, ne sont vraies que dans le référentiel $\mathcal{R}_0$. C'est en désaccord direct avec le principe de relativité énoncé plus haut.

Ainsi, à la fin du XIX^e siècle, a été posée l'hypothèse de l'éther. L'éther serait un milieu hypothétique dans lequel se déplacerait la lumière à la vitesse c . Maxwell postule ainsi que la lumière serait le résultat de la déformation élastique de ce milieu, un peu à l'image d'une onde mécanique (par exemple une déformation transversale d'une corde, ou encore une onde sonore qui est une déformation élastique de l'air se propageant). **L'hypothèse de l'éther implique que les relations de Maxwell ne seraient valables que dans le référentiel où l'éther est immobile.** En particulier, ce n'est que dans ce référentiel que la vitesse de la lumière est égale à c . On introduit donc la notion de référentiel absolu, ce qui brise là encore le principe de relativité où tous les référentiels d'inertie sont censés être équivalents.

4.2.3 L'expérience de Michelson-Morley

L'hypothèse de l'éther a été testée expérimentalement. L'un des premiers test expérimentaux est l'expérience de Michelson-Morley (1887), dont le but était de mesurer la vitesse de rotation de la Terre dans le référentiel de l'éther.

L'expérience est la suivante : dans le référentiel terrestre, un faisceau lumineux est séparé à l'aide d'une lame séparatrice en deux faisceaux (notés (1) et (2)), orthogonaux l'un à l'autre. Chaque faisceau est réfléchi par un miroir et renvoyé vers la lame séparatrice après avoir chacun parcouru la même distance $2d$, voir figure 4.3.

Les deux faisceaux peuvent interférer au moment de leur recroisement sur la lame séparatrice, et ainsi l'intensité lumineuse finale mesurée en sortie I dépend du déphasage entre les faisceaux (1) et (2), c'est-à-dire de la différence de temps de trajet entre les faisceaux (1) et (2) au moment où ils se recombinaient sur la séparatrice. Sans poser les calculs, on peut en effet comprendre que l'intensité finale de l'onde totale dépend du retard entre les deux ondes (1) et (2) que l'on va supposer sinusoïdales, cf. figure 4.4. L'interféromètre de Michelson-Morley permet ainsi, par la mesure de l'intensité lumineuse en sortie, de mesurer le retard de la propagation de la lumière entre les trajets (1) et (2).

Dans le problème, on s'arrange à ce que le faisceau (1) soit dans la direction de rotation de la Terre, et le faisceau (2) dans une direction transverse.

Pour prévoir les résultats de cette expérience, on va se placer dans le cadre des transformations de Galilée, et de l'hypothèse de l'éther. L'éther étant un référentiel absolu, on considère que la Terre se déplace dans l'éther suivant la vitesse constante $\vec{V}_{T/e}$. Cela revient également à dire que c'est l'éther qui se déplace par rapport à la Terre avec la vitesse constante $\vec{V}_{e/T} = -\vec{V}_{T/e}$ qui est en module la vitesse de rotation de la Terre autour du Soleil. On a donc un problème à deux référentiels d'inertie : $\mathcal{R}_T$ le référentiel terrestre que l'on considère immobile, et dans lequel la lumière se propage à la vitesse *a priori* inconnue v_T , et le référentiel $\mathcal{R}_e$ qui est l'éther, et dans lequel la vitesse de la lumière en module est fixe et vaut $v_e = c$.

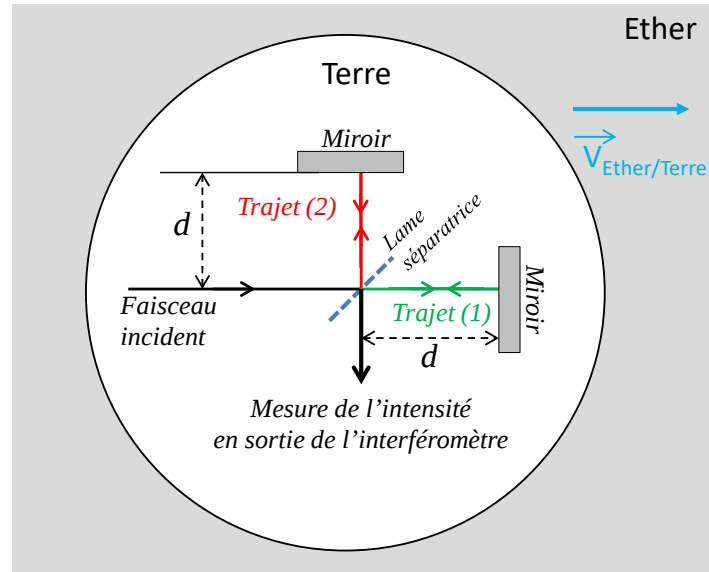


FIGURE 4.3 – Schéma de principe de l'expérience de Michelson-Morley.

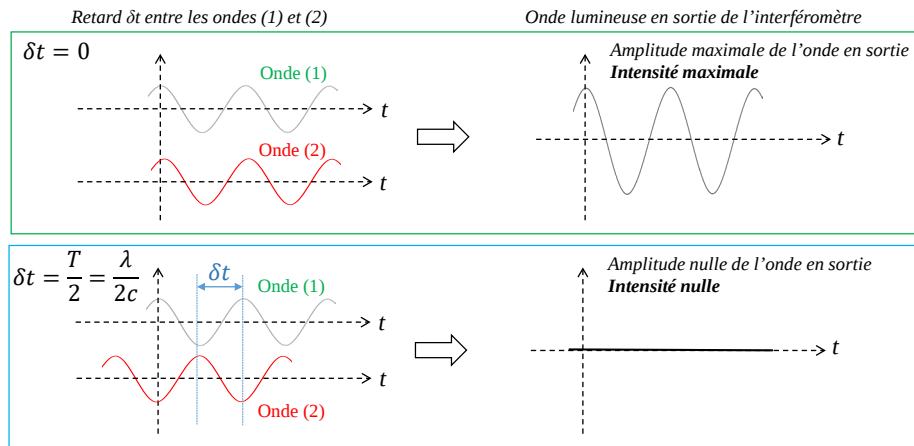


FIGURE 4.4 – l'onde lumineuse en sortie de l'interféromètre est issue de la somme des deux ondes ayant parcouru les trajets (1) et (2). l'amplitude de cette onde, et donc son intensité, dépend donc du retard de propagation δt entre les deux trajets. On représente ici deux cas extrêmes où les deux ondes arrivent en même temps (intensité maximale), ou avec un retard entre les deux de $\frac{T}{2}$ (intensité minimale).

Dans la suite, on va déterminer les temps d'aller-retour mis par les faisceaux lumineux afin de parcourir les trajets (1) et (2) de la figure 4.3, **dans le référentiel Terre** $\mathcal{R}_T$, en utilisant uniquement les transformations de Galilée et donc le principe d'additivité des vitesses entre les référentiels $\mathcal{R}_T$ et $\mathcal{R}_e$, c'est-à-dire :

$$\vec{v}_T = \vec{v}_e + \vec{V}_{e/T}$$

Trajet (1) Dans ce cas, les vitesses $\vec{v}_T$ et $\vec{V}_{e/T}$ sont colinéaires. On obtient donc, pour l'aller, que :

$$v_{Ta} = v_e + V_{e/T} = c + V_{e/T}$$

Pour le retour, on a :

$$v_{Tr} = v_e - V_{e/T} = c - V_{e/T}$$

Donc, au final, le temps d'une aller-retour sur le trajet (1) est :

$$t_1 = \frac{d}{v_{T_a}} + \frac{d}{v_{T_r}} = \frac{2dc}{c^2 - V_{e/T}^2}$$

Trajet (2) Dans le cas du trajet (2), les vecteurs $\vec{v}_T$ et $\vec{V}_{e/T}$ sont orthogonaux. Comme $\vec{v}_T = \vec{v}_e + \vec{V}_{e/T}$, on peut représenter schématiquement le vecteur $\vec{v}_e$, par exemple pour l'aller (cf. figure 1.4). En utilisant Pythagore :

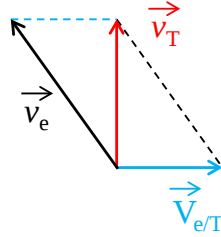


FIGURE 4.5 – Vitesses à l'aller du trajet 2

$$v_T^2 = v_e^2 - V_{e/T}^2 = c^2 - V_{e/T}^2$$

$$v_T = \sqrt{c^2 - V_{e/T}^2}$$

La situation est la même pour le retour. On obtient alors le temps de parcours total sur le trajet (2) :

$$t_2 = \frac{2d}{v_T} = \frac{2d}{\sqrt{c^2 - V_{e/T}^2}}$$

Bilan Ainsi, dans le cadre des transformations de Galilée, les temps de parcours des trajets (1) et (2) sont différents. Si on calcule la différence de temps de trajet entre (1) et (2) :

$$\delta t = t_1 - t_2 = 2d \left[\frac{c}{c^2 - V_{e/T}^2} - \frac{1}{\sqrt{c^2 - V_{e/T}^2}} \right]$$

$$\delta t = \frac{2d}{c^2 - V_{e/T}^2} \left(c - \sqrt{c^2 - V_{e/T}^2} \right)$$

Comme $c > V_{e/T}$, alors on a que $\delta t > 0$. Le trajet (1) est donc le trajet qui prend le plus de temps. L'appareil de Michelson-Morley devrait donc, dans le cadre de la mécanique classique, pouvoir mesurer une intensité totale en sortie qui correspond à ce retard.

La figure 4.6 montre deux figures de l'article original de Michelson et Morley paru dans the American Journal of Science, volume **34**, page 333 en 1887. La figure de gauche est un dessin représentant leur interféromètre monté sur une pierre massive le tout flottant sur un bain de mercure qui permet de faire tourner le dispositif. La figure de droite montre les résultats qu'ils ont obtenus. Ils sont spectaculaires et sans appel ! En traits pleins, le déplacement des franges d'interférences (la quantité δt déterminée plus haut multipliée par c) mesurée en fonction de l'orientation d'un des bras de l'interféromètre. En traits pointillés, 1/8 du déplacement théoriquement attendu par théorie éther et transformation galiléenne des vitesses ! Les résultats expérimentaux (traits pleins) et théoriques (pointillés) sont absolument incompatibles ! Depuis

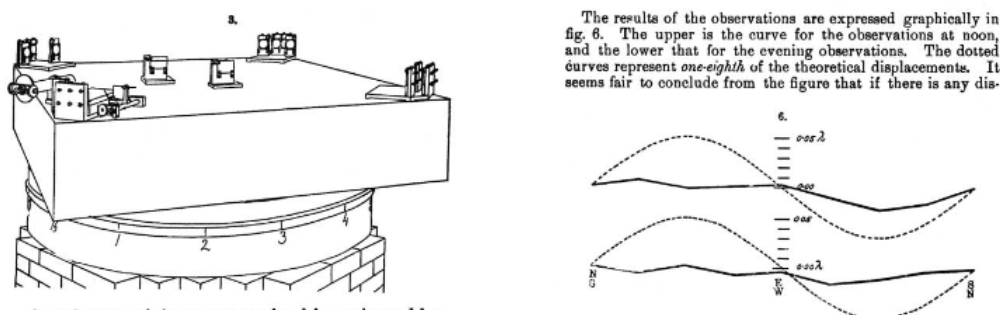


FIGURE 4.6 – Figures issues de l'article original de Michelson et Morley (*Am. J. Sci.*, **34**,333 (1887)). A gauche, schéma de l'interféromètre. A droite les résultats de l'expérience : déplacement des franges (*côt*) en unités d'une longueur d'ondes de référence (raie jaune du sodium) en fonction de l'orientation d'un des bras de l'interféromètre Nord/Sud Est/Ouest Sud/Nord. Traits pleins : la mesure, Traits pointillés : 1/8 du déplacement attendu dans le modèle éther et transformation galiléenne des vitesses.

cette expérience pionnière, ce constat a été vérifié par beaucoup d'autre expériences comparables réalisées plus tard (et même de nos jours avec des instruments de très haute sensibilité).

Et c'est là où il faut bien comprendre que la physique est par nature une science expérimentale. Aussi belle et cohérente que semble être une théorie et dans notre cas celle de l'éther et de la relativité galiléenne, elle n'est physiquement pas acceptable si elle ne rend pas compte, si elle ne prédit pas, les résultats expérimentaux observés.

La démarche du physicien est donc de remettre en cause les fondements (les postulats) de la théorie invalidée par l'expérience et d'en proposer de nouveaux et de construire une autre théorie capable de rendre compte des faits expérimentaux.

C'est ce que va faire Einstein en proposant sa théorie de la relativité restreinte que nous exposons dans la suite de ce cours.

4.3 Postulats de la relativité restreinte

4.3.1 Postulats d'Einstein

En 1905 dans son célèbre article *Zur Elektrodynamik bewegter Körper* (Sur l'électrodynamique des corps en mouvement), *Annalen der Physik* **17**, 891-921 (1905), Einstein pose les postulats sur lesquels il fonde la théorie de la relativité restreinte. Il réaffirme le postulat de relativité et élève au rang de postulat l'invariance de la vitesse de la lumière dans tous les référentiels d'inertie. De ce fait, les lois de l'électromagnétisme de Maxwell deviennent universelles et ne dépendent plus d'un référentiel en particulier et les transformations de Galilée sont remises en cause.

Postulat de relativité :

Tout référentiel en translation uniforme par rapport à un référentiel d'inertie est également un référentiel d'inertie.

Les lois de la physique sont identiques dans tous les référentiels d'inertie.

Postulat d'invariance de la vitesse de la lumière

La vitesse de la lumière dans le vide est la même dans tous les référentiels en translation uniforme les uns par rapport aux autres.

C'est ce deuxième postulat qui est conceptuellement radicalement nouveau car il remet en

cause les transformations de Galilée en s'opposant au principe d'additivité galiléenne des vitesses solidement établi à l'époque.

Il faut donc trouver de nouvelles transformations permettant de passer de la description dans un référentiel donné à un autre.

Remarque importante :

La théorie de la relativité restreinte ne s'applique que dans les cas de changements de référentiels en translation uniforme les uns par rapport aux autres. Le cas des référentiels accélérés fait intervenir une théorie plus élaborée qui est la théorie de la relativité générale. Ainsi, les interactions gravitationnelles ne rentrent pas dans le cadre de la relativité restreinte, et font donc exception au premier postulat de la relativité restreinte.

4.3.2 Transformations de Lorentz

Afin de satisfaire aux postulats d'Einstein, on doit établir de nouvelles transformations des coordonnées pour passer d'un référentiel d'inertie à un autre. Ces transformations doivent obéir à plusieurs contraintes.

D'une part, elles doivent laisser invariante la vitesse de la lumière dans le vide c . La vitesse est définie comme un rapport d'une distance et d'une durée. Comme les changements de référentiels en mouvement font nécessairement intervenir un changement des coordonnées spatiales, cela implique donc que la coordonnée temporelle est également modifiée.

D'autre part, elles doivent garantir que, les référentiels d'inertie étant tous équivalents, la transformation d'un référentiel $\mathcal{R}$ vers $\mathcal{R}'$ est la même que celle de $\mathcal{R}'$ vers $\mathcal{R}$.

Événement spatio-temporel

Nous avons déjà mis en avant le fait que les relations de passage entre un référentiel $\mathcal{R}$ et un référentiel $\mathcal{R}'$ en translation uniforme doivent modifier à la fois les coordonnées spatiales et la coordonnée temporelle.

Un "événement" est quelque chose qui se passe à un endroit de l'espace et à un temps donné.

Pour caractériser un événement il faut donc donner 4 nombres réels, trois qui caractérisent la position spatiale de l'événement et qui ont la dimension d'une longueur et un qui caractérise sa position temporelle et qui a la dimension d'un temps. On parle ainsi d'**événement spatio-temporel**. Ces 4 nombres repèrent un point de ce que l'on appelle l'**espace-temps** qui est un "espace" (au sens mathématique du terme) à 4 dimensions. La valeur de ces nombres dépend du référentiel dans lequel est repéré l'événement.

Par exemple, on considère l'événement :

E_1 "Antoine Dupont aplatit le ballon au point M de l'en-but néo-zélandais".

On considère deux référentiels en translation uniforme. Le référentiel $\mathcal{R}$ dont le centre des coordonnées peut être par exemple le point O du centre du terrain et le référentiel $\mathcal{R}'$ d'un avion passant au dessus du stade qui est en translation uniforme par rapport à $\mathcal{R}$ et dont le centre des coordonnées O' est situé sur le siège du pilote.

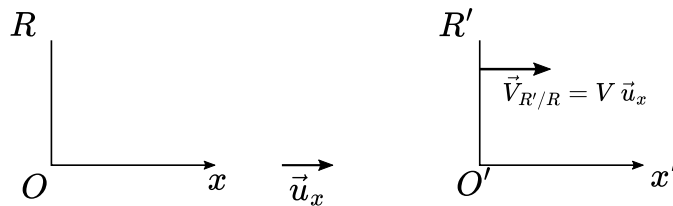
Les coordonnées de E_1 dans $\mathcal{R}$ sont $(x_1 = x_M, y_1 = y_M, z_1 = z_M, t_1)$.

Les coordonnées de E_1 dans $\mathcal{R}'$ sont $(x'_1 = x'_M, y'_1 = y'_M, z'_1 = z'_M, t'_1)$.

En règle générale, $x'_1 \neq x_1$ (par exemple $x_1 = -7.8$ m et $x'_1 = +14.9$ m), $y'_1 \neq y_1$, $z'_1 \neq z_1$ et (c'est la grande nouveauté de la théorie de la relativité restreinte) $t'_1 \neq t_1$.

Formulation des transformations dites de "Lorentz".

On introduit deux référentiels, chacun à une dimension d'espace pour simplifier les expressions. Un référentiel $\mathcal{R}$ et un deuxième $\mathcal{R}'$ est animé d'une vitesse *constante* par rapport à $\mathcal{R}$ $\vec{V}_{\mathcal{R}'/\mathcal{R}} = V\vec{u}_x$, voir la figure 4.7 où l'on a fait une représentation des référentiels pour $V > 0$.

FIGURE 4.7 – Représentation des deux référentiels $\mathcal{R}$ et $\mathcal{R}'$.

On notera x et t les coordonnées d'un événement dans $\mathcal{R}$, et x' et t' les coordonnées de ce même événement dans $\mathcal{R}'$. On appellera O et O' les origines de chacun des référentiels. Dans les deux repères, on définit l'origine des temps et de l'espace de manière à ce que l'événement qui a pour coordonnées $(x = 0, t = 0)$ dans $\mathcal{R}$ ait pour coordonnées $(x' = 0, t' = 0)$ dans $\mathcal{R}'$. Autrement dit, on s'arrange pour que les origines O et O' soient confondus à la date $t = 0$ dans $\mathcal{R}$ et $t' = 0$ dans $\mathcal{R}'$.

La transformation la plus générale possible pour passer de $\mathcal{R}$ à $\mathcal{R}'$ s'écrit

$$\begin{cases} x' = g(x, t) \\ t' = h(x, t), \end{cases} \quad (4.1)$$

où les fonctions g et h dépendent également de la vitesse V . Cependant, un argument simple basé sur l'homogénéité de l'espace¹ implique que la transformation est, en fait, linéaire. On écrit donc

$$\begin{cases} x' = ax + bt, \\ t' = ex + ft, \end{cases} \quad (4.2)$$

où les coefficients a , b , e et f dépendent de V .

Dans la suite, on va chercher à exprimer les coefficients a , b , e et f afin de satisfaire aux contraintes que l'on a exprimées plus haut.

Par définition, la vitesse d'un objet suivant x est dans $\mathcal{R}$:

$$v_x = \frac{dx}{dt}$$

Dans $\mathcal{R}'$:

$$v'_x = \frac{dx'}{dt'} = \frac{d(ax + bt)}{d(ex + ft)} = \frac{a dx + b dt}{e dx + f dt} = \frac{a \frac{dx}{dt} + b}{e \frac{dx}{dt} + f} = \frac{av_x + b}{ev_x + f}$$

On va utiliser cette relation sur quelques points particuliers (O et O').

1. Dans $\mathcal{R}$, considérons une expérience où l'on suit le mouvement d'une particule subissant des forces exercées par une machine. La particule part de $(x = 0, t = 0)$ et arrive en (x_1, t_1) . La même particule vue dans $\mathcal{R}'$ part de $(x' = 0, t' = 0)$ et arrive en (x'_1, t'_1) avec $x'_1 = g(x_1, t_1)$ et $t'_1 = h(x_1, t_1)$.

Maintenant, on recommence la même expérience, mais décalée dans l'espace et le temps : la particule part de (X, T) et arrive en $(X + x_1, T + t_1)$; en effet, on a également décalé la machine, donc la particule est dans le même environnement et suit la même trajectoire, mais décalée dans l'espace de X et dans le temps de T . Pour la même expérience décalée dans $\mathcal{R}'$, la particule part de (X', T') avec $X' = g(X, T)$ et $T' = h(X, T)$, et arrive en $(X' + x'_1, T' + t'_1)$ parce que, là encore, la trajectoire dans $\mathcal{R}'$ est simplement décalée. Mais on doit également avoir $X' + x'_1 = g(X + x_1, T + t_1)$ et $T' + t'_1 = h(X + x_1, T + t_1)$. Finalement, on obtient que

$$g(X + x_1, T + t_1) = g(X, T) + g(x_1, t_1), \quad h(X + x_1, T + t_1) = h(X, T) + h(x_1, t_1).$$

Ces relations, valables pour tout X , T , x_1 et t_1 , impliquent que g et h sont linéaires.

Vitesse de O' Dans $\mathcal{R}$, O' est animé de la vitesse V . Dans $\mathcal{R}'$, O' est immobile et donc sa vitesse est nulle. Si on remplace dans l'équation précédente :

$$0 = \frac{a \times V + b}{e \times V + f},$$

ce qui donne :

$$aV + b = 0 \quad (4.3)$$

Vitesse de O Dans $\mathcal{R}$, O est immobile est donc sa vitesse est nulle. Dans $\mathcal{R}'$, O se déplace à la vitesse $-V$ (vers les x' négatifs).

$$-V = \frac{a \times 0 + b}{e \times 0 + f},$$

ce qui donne :

$$-V = \frac{b}{f} \quad (4.4)$$

Invariance de c Si $v_x = c$ dans $\mathcal{R}$, alors $v'_x = c$ dans $\mathcal{R}'$:

$$c = \frac{ac + b}{ec + f},$$

et on obtient

$$ac + b = ec^2 + fc \quad (4.5)$$

Bilan Des équations (4.3) et (4.4), on tire :

$$f = a, \quad b = -Va.$$

En injectant cela dans l'équation (4.5), on arrive à :

$$e = -\frac{V}{c^2}a$$

On peut ainsi réécrire les transformations en fonction du seul paramètre a :

$$\begin{cases} x' &= a(x - Vt) \\ t' &= a\left(t - \frac{V}{c^2}x\right) \end{cases}$$

Pour déterminer a , on utilise le fait qu'une transformation de $\mathcal{R}'$ vers $\mathcal{R}$, s'écrit de la même façon que celle de $\mathcal{R}$ vers $\mathcal{R}'$, à condition de changer la vitesse V en $-V$; en effet : $\vec{V}_{\mathcal{R}/\mathcal{R}'} = -\vec{V}_{\mathcal{R}'/\mathcal{R}} = -V\vec{u}_x$. On peut donc écrire²

$$\begin{cases} x &= a(x' + Vt') \\ t &= a\left(t' + \frac{V}{c^2}x'\right) \end{cases}$$

Ainsi, si on effectue deux transformations à la suite de $\mathcal{R}$ vers $\mathcal{R}'$ puis de $\mathcal{R}'$ vers $\mathcal{R}$, on doit retomber sur le jeu de coordonnées de départ dans $\mathcal{R}$. Par exemple, pour x :

$$x = a(x' + Vt') = a\left[a(x - Vt) + Va\left(t - \frac{V}{c^2}x\right)\right] = a^2\left[\left(1 - \frac{V^2}{c^2}\right)x\right] \quad \text{et donc } a = \frac{1}{\sqrt{1 - \frac{V^2}{c^2}}}$$

2. N'oublions pas que a dépend de la vitesse entre les deux référentiels. Plus précisément, il dépend de la *norme* de cette vitesse, parce que toutes les directions dans l'espace sont équivalentes. Comme les vitesses $V\vec{u}_x$ et $-V\vec{u}_x$ ont la même norme, on a bien le *même* coefficient a dans les équations passant de (x, t) à (x', t') et celles passant de (x', t') à (x, t) .

On obtient alors un jeu de transformations à une dimension spatiale et une dimension temporelle entre les référentiels $\mathcal{R}$ et $\mathcal{R}'$ qui garantit la validité des postulats d'Einstein.

$$\begin{cases} x' &= \gamma(x - Vt) \\ t' &= \gamma\left(t - \frac{V}{c^2}x\right) \end{cases} \quad \text{avec } \gamma = \frac{1}{\sqrt{1 - \frac{V^2}{c^2}}}. \quad (4.6)$$

Pour déterminer la loi de transformation de la coordonnée spatiale y , on considère, dans le référentiel $\mathcal{R}$, un flash lumineux partant de $(x = 0, y = 0, t = 0)$ et arrivant en (x_1, y_1, t_1) . Dans le référentiel $\mathcal{R}'$, le même flash lumineux part de $(x' = 0, y' = 0, t' = 0)$ et arrive en (x'_1, y'_1, t'_1) .

La vitesse de la lumière étant la même (en norme!) dans $\mathcal{R}$ et $\mathcal{R}'$, on a :

$$c^2 = \frac{x_1^2 + y_1^2}{t_1^2} = \frac{x_1'^2 + y_1'^2}{t_1'^2}, \quad (4.7)$$

qui conduit à :

$$x_1^2 + y_1^2 - c^2 t_1^2 = x_1'^2 + y_1'^2 - c^2 t_1'^2. \quad (4.8)$$

En utilisant les transformations des coordonnées x'_1 et t'_1 données par les équations (4.6), on arrive à :

$$x_1^2 + y_1^2 - c^2 t_1^2 = \gamma^2(x_1 - Vt_1)^2 + y_1'^2 - c^2 \gamma^2(t_1 - Vx_1/c^2)^2 \quad (4.9)$$

Après quelques lignes de calculs que vous pouvez faire pour vous entraîner et en remarquant que $\gamma^2(1 - \frac{V^2}{c^2}) = 1$, on montre que :

$$y_1' = y_1 \quad (4.10)$$

Pour la coordonnée spatiale y , la transformation qui garantit la validité des postulats d'Einstein est simple :

$$y' = y \quad (4.11)$$

Un raisonnement équivalent sur la coordonnée spatiale z conduit à :

$$z' = z \quad (4.12)$$

Les transformations qui permettent de passer du référentiel $\mathcal{R}$ au référentiel $\mathcal{R}'$ en translation uniforme par rapport $\mathcal{R}$ avec la vitesse $\vec{V}_{\mathcal{R}'/\mathcal{R}} = V\vec{u}_x$ en accord avec les postulats de la relativité restreinte s'appellent les transformations de Lorentz. En posant que les origines des coordonnées de ces référentiels O et O' sont confondues l'instant $t = t' = 0$, elles s'écrivent :

$$\begin{cases} x' &= \gamma(x - Vt) \\ y' &= y \\ z' &= z \\ t' &= \gamma\left(t - \frac{V}{c^2}x\right) \end{cases} \quad (4.13)$$

avec le facteur de Lorentz :

$$\gamma = \frac{1}{\sqrt{1 - \frac{V^2}{c^2}}} > 1$$

Les transformations de Lorentz sont le résultat le plus important du cours. Tous les résultats dont nous parlerons par la suite sont des conséquences plus ou moins directes de ces transformations! Leur signification profonde est déroutante même pour le physicien expérimenté! Mais leur application mathématique est simple et tout problème de relativité restreinte peut se traiter en se rapportant à ces transformations.

Les transformations inverse qui permettent de trouver les coordonnées dans $\mathcal{R}$ quand on connaît celles dans $\mathcal{R}'$, se trouvent en faisant dans (4.13) les substitutions $x' \leftrightarrow x$, $y' \leftrightarrow y$, $z' \leftrightarrow z$, $t' \leftrightarrow t$ et $V \rightarrow -V$. On arrive à :

$$\begin{cases} x &= \gamma(x' + Vt') \\ y &= y' \\ z &= z' \\ t &= \gamma(t' + \frac{V}{c^2}x) \end{cases} \quad (4.14)$$

Les transformations de Lorentz s'écrivent sous la forme donnée dans l'encadré si on suppose que O et O' sont confondus aux temps $t = t' = 0$. C'est le plus commode et le plus parlant. On peut cependant bien évidemment écrire les transformations de Lorentz dans le cas général où l'événement $(x = 0, y = 0, z = 0, t = 0)$ dans $\mathcal{R}$ a pour coordonnées (x'_0, y'_0, z'_0, t'_0) dans $\mathcal{R}'$:

$$\begin{cases} x' - x'_0 &= \gamma(x - Vt) \\ y' - y'_0 &= y \\ z' - z'_0 &= z \\ t' - t'_0 &= \gamma(t - \frac{V}{c^2}x) \end{cases} \quad (4.15)$$

Chapitre 5

Conséquences des transformations de Lorentz : simultanéité, transformations des longueurs et des durées

5.1 Relativité du temps et de la simultanéité

Le point le plus curieux et le plus intrigant des transformations de Lorentz (4.13) est certainement le fait que $t \neq t'$. Les transformations de Lorentz impliquent que la coordonnée de temps n'est plus commune à tous les référentiels en translation uniforme. Chaque référentiel possède sa propre échelle de temps, le temps s'écoule différemment dans les référentiels $\mathcal{R}$ et $\mathcal{R}'$. Il ne fait aucun doute que cette notion est difficile à intégrer lorsqu'on la rencontre pour la première fois mais c'est bien ce que disent les transformations de Lorentz !

Une conséquence de cela est la suivante :

Si $\mathcal{R}$ et $\mathcal{R}'$ sont deux référentiels inertiels en mouvement l'un par rapport à l'autre avec une vitesse selon $\vec{u}_x$, alors deux événements simultanés dans $\mathcal{R}$ mais ayant des abscisses x différentes **ne peuvent pas** être simultanés dans $\mathcal{R}'$.

5.1.1 Définition de la simultanéité

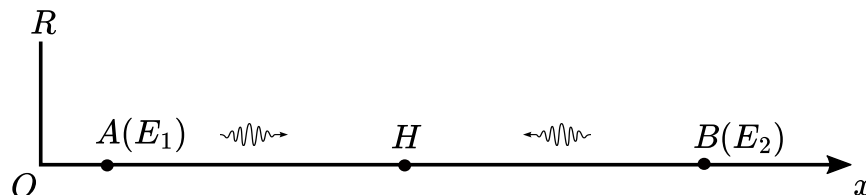
Deux événements E_1 et E_2 ayant lieu dans des endroits distincts sont simultanés dans un référentiel donné s'ils ont lieu au même instant dans ce référentiel. On pourrait se satisfaire de cette phrase comme définition de la simultanéité mais elle manque de précision pour le physicien.

Le physicien veut répondre à la question : «comment puis-je mesurer, c'est-à-dire réaliser une expérience, qui me permette de savoir si deux événements sont simultanés ou non ? De manière plus générale, comment puis-je déterminer la date d'un événement donné ?»

Prenons un exemple : depuis son observatoire dans la montagne, un physicien observe un éclair frappant une tour située dans la vallée. Le physicien, qui dispose d'une horloge ultra-précise, veut savoir la date précise à laquelle l'éclair a frappé. Pour cela, il relève la date t_0 à laquelle il a reçu le flash de l'éclair, et il soustrait le temps de propagation de la lumière, à savoir : (distance entre lui et la tour)/(vitesse de la lumière). (Il pourrait aussi mesurer la date $\tilde{t}_0$ à laquelle il reçoit le bruit du tonnerre, et soustraire le temps de propagation du son, mais c'est plus difficile parce que la vitesse de propagation du son dépend de la température, du vent, etc.). Si un autre physicien sur la montagne d'en face fait la même mesure, les deux physiciens trouveront la même date pour l'événement «l'éclair frappe la tour».

Dans un référentiel donné $\mathcal{R}$ muni d'une origine des temps, il est donc possible de **définir sans ambiguïté la date d'un événement**, et de la mesurer. En particulier, il est possible de **déterminer si deux événements sont simultanés** ou non, et il est possible d'avoir des horloges identiques immobiles dans le référentiel qui soient synchronisées.

Une manière simple et directe de déterminer si deux événements E_1 et E_2 sont simultanés dans un référentiel donné est la suivante : on place un observateur H au milieu du segment liant le point A , coordonnée spatiale de l'événement E_1 , et le point B , coordonnée spatiale de l'événement E_2 .



À l'instant où a lieu l'événement E_1 , un flash lumineux part de A en direction de H . À l'instant où a lieu l'événement E_2 , un flash lumineux part de B en direction de H . Les événements E_1 et E_2 sont simultanés dans $\mathcal{R}$ si les flashes lumineux issus de A et B arrivent simultanément en H .

Avec ce protocole, on peut non seulement de tester expérimentalement si deux événements sont simultanés mais on peut aussi synchroniser des horloges entre elles et ainsi de définir le temps dans le référentiel $\mathcal{R}$.

Imaginons que nous prenions deux horloges de même constitution et qui donc *battent au même rythme* et qu'on les amène aux points A et B considérés précédemment.

Soit E_3 l'événement "l'aiguille de l'horloge situé en A pointe sur midi" et E_4 l'événement "l'aiguille de l'horloge situé en B pointe sur midi". Si l'observateur H situé au milieu des coordonnées spatiales de E_3 et E_4 reçoit au même instant les photons issus des événements E_3 et E_4 (c'est-à-dire qu'il voit *simultanément* les aiguilles des horloges situées en A et B pointer sur midi), alors **ces horloges sont synchronisées dans $\mathcal{R}$** .

En suivant ce protocole, on peut synchroniser dans $\mathcal{R}$ autant d'horloges que l'on veut, en autant qu'on veut de points fixes de $\mathcal{R}$.

À mesure que le "temps passe", les aiguilles des différents horloges se décalent *au même rythme* et pointent donc simultanément dans des directions communes.

On a ainsi défini le temps dans $\mathcal{R}$. C'est l'indication de la position des aiguilles des horloges *situées en tout point de l'espace* où l'on veut dater un événement.

5.1.2 Relativité de la simultanéité et donc du temps

Une fois la simultanéité définie arrive naturellement la question suivante : si deux événements sont simultanés dans un référentiel $\mathcal{R}$, le sont-ils dans un référentiel $\mathcal{R}'$ qui est en translation uniforme par rapport à $\mathcal{R}$?

On considère le cas suivant : deux lampes sont immobiles dans $\mathcal{R}$ aux points d'abscisses $x_1 = -L$ et $x_2 = +L$. Un observateur H est situé au point O origine des coordonnées ($x_H = 0$)

Soient E_1 l'événement "la lampe de gauche émet un flash" et E_2 l'événement "la lampe de droite émet un flash".

H , situé au milieu des deux lampes, sait quel événement a lieu en premier dans $\mathcal{R}$ en observant quel flash arrive en premier. Imaginons que H constate que les flashes issus de E_1 et E_2 lui arrivent au même instant ; d'après ce qui précède, H conclut que les événements E_1 et E_2 sont

simultanés. En redéfinissant l'origine des temps dans $\mathcal{R}$ de manière à ce que la date (commune !) de ces événements soit $t = 0$, les coordonnées des deux événements dans $\mathcal{R}$ sont donc

$$E_1^{\mathcal{R}}(x = -L, t = 0), \quad E_2^{\mathcal{R}}(x = +L, t = 0). \quad (5.1)$$

Supposons de plus que, à cet instant $t = 0$ où les photons sont émis dans $\mathcal{R}$, passe au point O un observateur H' situé à l'origine des coordonnées O' ($x'_{H'} = x'_{O'} = 0$) d'un référentiel $\mathcal{R}'$ en translation uniforme par rapport à $\mathcal{R}$ avec une vitesse $\vec{V}_{\mathcal{R}'/\mathcal{R}} = V\vec{u}_x$.

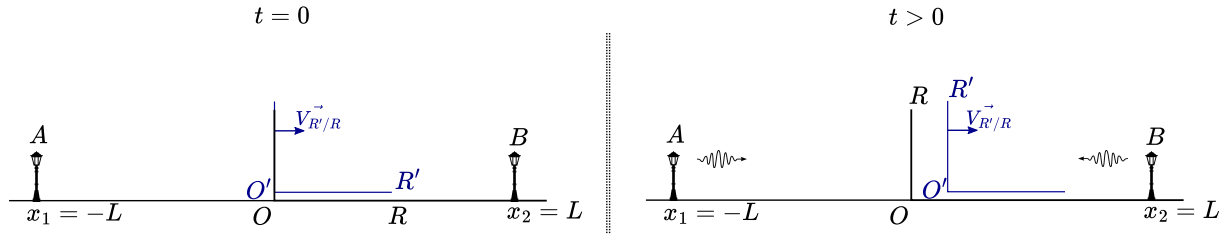


FIGURE 5.1 – Relativité de la simultanéité. Gauche : dans $\mathcal{R}$, les lampes A et B émettent simultanément deux photons à $t = 0$. Représentation de la situation à $t > 0$ dans $\mathcal{R}$.

Aux événements E_1 et E_2 de coordonnées spatiales $x_1 = -L$ et $x_2 = +L$ dans $\mathcal{R}$ correspondent des coordonnées spatiales x'_1 et x'_2 dans $\mathcal{R}'$ qui sont nécessairement symétriques par rapport à O' , c'est-à-dire $x'_1 = -x'_2$, parce que les transformations pour passer d'un référentiel à un autre sont linéaires.¹

H' ($x'_{H'} = x'_{O'} = 0$) est donc situé au milieu du segment $[x'_1, x'_2]$ et il peut donc déterminer expérimentalement si les événements E_1 et E_2 sont simultanés dans $\mathcal{R}'$. Pour cela, il faut simplement que H' mesure les instants auxquels les photons issus des deux lampes lui arrivent. Si ces photons lui arrivent en même temps, E_1 et E_2 sont simultanés dans $\mathcal{R}'$, et sinon E_1 et E_2 ne sont pas simultanés dans $\mathcal{R}'$.

En raisonnant dans le référentiel $\mathcal{R}$, il est facile de voir que H' étant en train d'avancer vers la lampe B au moment où le photon est émis de B et étant en train de s'éloigner la lampe A au moment où le photon est émis de A , il recevra d'abord le photon issu de la lampe B puis celui issu de la lampe A .

H' doit donc conclure que dans $\mathcal{R}'$ les événements E_1 et E_2 ne sont pas simultanés !

$$t'_1 \neq t'_2 \quad (\text{dans } \mathcal{R}'). \quad (5.2)$$

Cette conclusion est bien en accord avec les transformations de Lorentz qui prédisent que :

$$t'_1 = \gamma(t_1 - Vx_1/c^2) = +\gamma V \frac{L}{c^2}, \quad t'_2 = \gamma(t_2 - Vx_2/c^2) = -\gamma V \frac{L}{c^2},$$

qui donnent bien $t'_1 \neq t'_2$ et, plus précisément, $t'_2 < t'_1$: dans $\mathcal{R}'$, l'événement E_2 a lieu avant l'événement E_1 .

1. On peut vérifier directement que les transformations de Galilée ou de Lorentz donnent bien des coordonnées x'_1 et x'_2 symétriques. La transformation de Galilée donne en effet

$$x'_1 = x_1 - Vt_1 = x_1 = -L, \quad x'_2 = x_2 - Vt_2 = x_2 = +L.$$

(On rappelle que les dates des deux événements sont nulles : $t_1 = t_2 = 0$.) La transformation de Lorentz donne quant à elle :

$$x'_1 = \gamma(x_1 - Vt_1) = -\gamma L, \quad x'_2 = \gamma(x_2 - Vt_2) = +\gamma L.$$

Dans les deux cas, on a bien $x'_1 = -x'_2$.

5.2 Dilatation des durées et contraction des longueurs

La simultanéité des événements étant une notion relative (qui dépend du référentiel), la durée séparant deux événements dépend donc elle aussi du référentiel. Nous allons voir qu'il en va de même pour les longueurs des objets, qui dépendent elles aussi du référentiel.

5.2.1 Première approche

Nous allons préciser ces notions sur l'exemple d'un vaisseau spatial partant du système solaire vers une étoile voisine, voir figure 5.2.

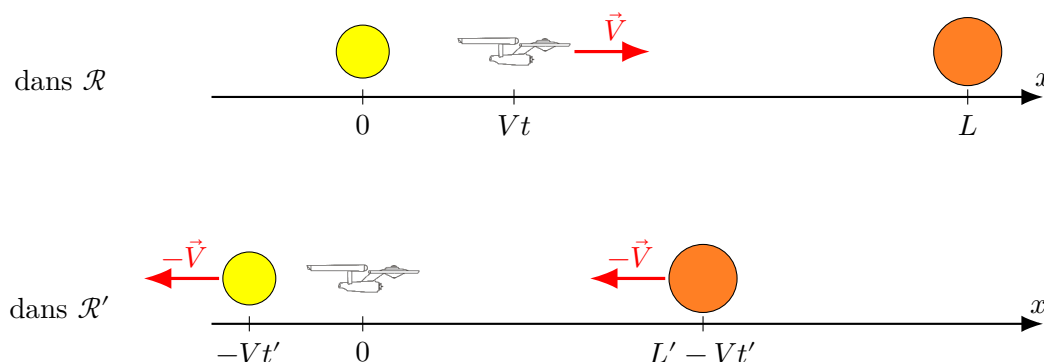


FIGURE 5.2 – Un voyage sidéral depuis le Soleil (à gauche) vers une étoile voisine (à droite). La figure du dessus est à un instant t dans le référentiel $\mathcal{R}$ où les étoiles sont fixes, alors que celle du dessous est à un instant t' dans le référentiel $\mathcal{R}'$ où le vaisseau est fixe. La distance entre les étoiles est L dans $\mathcal{R}$ et L' dans $\mathcal{R}'$.

Dans le référentiel héliocentrique $\mathcal{R}$, le point de départ du voyage (le Soleil) est fixe à la position $x = 0$, et le point d'arrivée (l'étoile voisine) est également fixe à la position $x = L$. La distance L entre les deux étoiles est ici **mesurée dans le référentiel où ces deux étoiles sont immobiles**. On dit, par définition que c'est la **distance propre** entre les deux étoiles.

Toujours dans le référentiel $\mathcal{R}$, on choisit l'origine des temps de manière à ce que le départ du voyage ait lieu à l'instant $t = 0$. Le vaisseau arrive à l'instant $t = T$ à destination, où T est donc la durée du voyage dans le référentiel $\mathcal{R}$. (Si les voyageurs envoient un signal "Nous sommes arrivés" au terme de leur voyage, ce signal atteint donc la Terre à la date $T + L/c$, ce qui permet aux terriens, connaissant L , d'en déduire T .)

On a donc deux événements importants : le départ et l'arrivée du vaisseau, dont les coordonnées dans $\mathcal{R}$ sont

$$E_{\text{départ}}^{\mathcal{R}}(x = 0, t = 0), \quad E_{\text{arrivée}}^{\mathcal{R}}(x = L, t = T).$$

On suppose que le vaisseau se déplace à vitesse constante V dans le référentiel $\mathcal{R}$. (Il y a peut-être une phase d'accélération au départ, et une phase de décélération à la fin, mais on suppose que ces phases sont d'une durée courte et négligeable devant la durée totale du voyage.) Clairement, on a

$$V = L/T.$$

On note $\mathcal{R}'$ le référentiel lié au vaisseau. En appliquant les transformations de Lorentz (4.13)

$$E_{\text{départ}}^{\mathcal{R}'}(x' = 0, t' = 0), \quad E_{\text{arrivée}}^{\mathcal{R}'}(x' = 0, t' = T') \quad \text{avec } T' = \frac{1}{\gamma}T.$$

En effet, pour $E_{\text{arrivée}}$, on a $x' = \gamma(L - VT) = 0$ et $t' = \gamma(T - VL/c^2) = \gamma T(1 - V^2/c^2) = \gamma T/\gamma^2$. Le résultat $x' = 0$ était évident : la coordonnée spatiale de l'événement «le vaisseau arrive» dans

le référentiel $\mathcal{R}'$ est forcément la position du vaisseau qui, par définition, est $x' = 0$. Le résultat $T' = T/\gamma$ est plus intéressant : **le voyage du point de vue du vaisseau dure un temps T' inférieur à T** . Ce temps T' serait effectivement le temps affiché par l'horloge du vaisseau, et le temps du voyage ressenti par les voyageurs. Ce temps T' , mesurée par une horloge dans le référentiel où elle est immobile, s'appelle le **temps propre**, ou **durée propre**. Le temps T du voyage mesuré dans le référentiel $\mathcal{R}$ où les voyageurs sont en mouvement est alors un temps “impropre”.

Revenons maintenant à la distance entre les deux étoiles, mais du point de vue des voyageurs. Dans le référentiel $\mathcal{R}'$, le vaisseau est immobile et les étoiles se déplacent à la vitesse V . (Le Soleil s'éloigne à la vitesse V , et l'étoile voisine se rapproche à vitesse V). À l'instant $t' = 0$, le soleil est juste à côté du vaisseau. À l'instant $t' = T'$, c'est l'étoile voisine qui est à côté du vaisseau. Puisqu'il faut (de leur point de vue) un temps T' pour que les étoiles à la vitesse V passent devant eux, du Soleil à l'étoile voisine, les voyageurs en déduisent que la distance entre les astres est $L' = VT' = VT/\gamma = L/\gamma$. Cette distance L' est une distance impropre, puisqu'elle est mesurée dans un référentiel où les deux extrémités de la distance sont en mouvement.

En résumé

- La **durée propre**, ou **temps propre**, entre deux événements est la durée mesurée dans un référentiel où les deux événements sont à la même position, et où la mesure est faite par une horloge qui est immobile à cette position.
- La durée entre deux événements dans tout autre référentiel est une **durée impropre**. Si γ est le facteur de Lorentz correspondant au passage entre ces deux référentiels, alors

$$(\text{durée impropre}) = \gamma(\text{durée propre}), \quad \text{et donc} \quad (\text{durée impropre}) > (\text{durée propre})$$

On parle de **dilatation de la durée** (sous entendu : par rapport à la durée propre).

- La **longueur propre** d'un objet rigide (une règle, un train, l'écartement entre deux étoiles fixes) est la longueur de l'objet mesurée dans le référentiel où il est immobile.
- La longueur d'un objet mesurée dans un référentiel où il est en mouvement est une **longueur impropre**. Si l'objet est en mouvement dans l'axe de sa longueur et si γ est le facteur de Lorentz correspondant au passage entre ces deux référentiels, alors

$$(\text{longueur impropre}) = \frac{1}{\gamma}(\text{longueur propre}), \quad \text{et donc} \quad (\text{longueur impropre}) < (\text{longueur propre})$$

On parle de **contraction des longueurs** (sous entendu : par rapport au longueurs propres).

5.2.2 Plus de détail sur la dilatation des durées

Pour mieux comprendre pourquoi le temps s'écoule de manière différente entre deux référentiels, on peut considérer l'exemple suivant où l'on prend deux horloges, l'une immobile et l'autre en mouvement à la vitesse $\vec{V} = V\vec{u}_x$ dans un référentiel $\mathcal{R}$. Appelons A l'horloge immobile et B l'horloge en mouvement. On note $\mathcal{R}'$ le référentiel lié à l'horloge en mouvement.

On va supposer que ce sont des horloges un peu particulières : il s'agit de deux miroirs parallèles espacés d'une distance d , et entre lesquels un photon fait des allers-retours dans une direction orthogonale à $\vec{V}$. Les durées sont alors mesurées par le nombre d'allers-retours qu'aura fait le photon.

Pour un observateur fixe dans $\mathcal{R}$, le temps d'un aller-retour pour l'horloge A est $\tau_A = \frac{2d}{c}$. Pour l'horloge B , les miroirs étant en déplacement constant dans $\mathcal{R}$, le photon doit parcourir une distance d_B plus grande que d dans $\mathcal{R}$, voir figure 5.3. Or, sa vitesse est toujours c ; l'observateur dans $\mathcal{R}$ mesure alors un temps caractéristique $\tau_B = \frac{2d_B}{c} > \tau_A$.

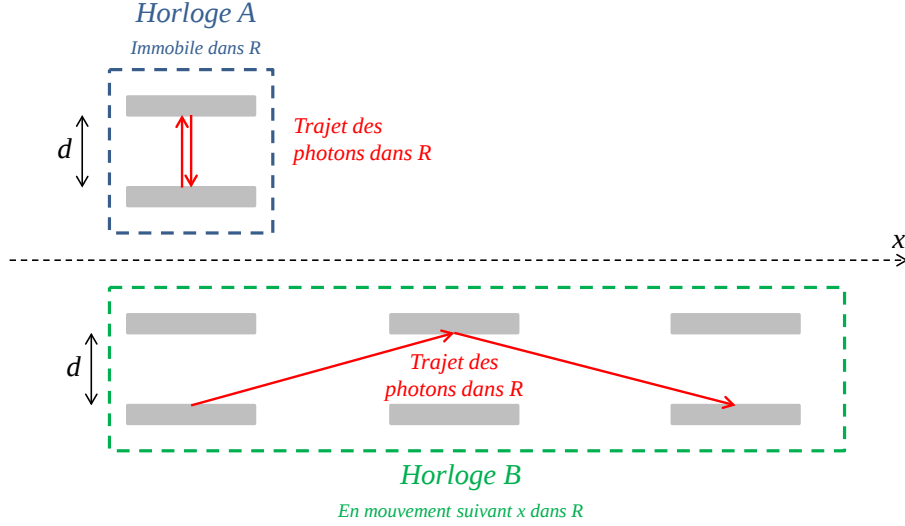


FIGURE 5.3 – Horloges immobile A et mobile B dans $\mathcal{R}$. Les trajets que les photons doivent emprunter pour faire des allers-retours entre les miroirs sont représentés.

Les deux horloges sont identiques, et le problème est symétrique ; si maintenant on considère un observateur fixe dans $\mathcal{R}'$, cet observateur doit mesurer dans ce référentiel des temps caractéristiques τ'_A et τ'_B tels que $\tau'_B = \tau_A$ (temps d'une horloge immobile dans le référentiel de mesure) et $\tau'_A = \tau_B$ (temps d'une horloge en mouvement). On a donc simultanément $\tau_B > \tau_A$ et $\tau'_A > \tau'_B$! La durée mesurée par une horloge au repos dans le référentiel de mesure est toujours inférieure à celle donnée par une horloge en mouvement. Si on mesure le temps en comptant une succession de durées égales (principe de l'horloge), alors le temps s'écoule plus lentement dans le référentiel où l'horloge est en mouvement que dans le référentiel où l'horloge est au repos.

La durée propre d'un aller-retour de photon est la durée mesurée dans le référentiel où l'horloge est au repos ; il s'agit donc de la durée τ_A pour l'horloge A (mesurée dans $\mathcal{R}$) ou τ'_B pour l'horloge B (mesurée dans $\mathcal{R}'$) : durée propre = $\tau_A = \tau'_B$. L'autre durée, $\tau'_A = \tau_B$, mesurée pour chaque horloge dans le référentiel où elle est en mouvement, est la durée impropre. On retrouve bien (durée impropre) $>$ (durée propre).

Calculons maintenant la relation entre durées propre et impropre pour deux événements quelconques. On considère deux événements E_1 et E_2 tels que, dans le référentiel $\mathcal{R}$, les événements ont lieu au même endroit :

$$E_1^{\mathcal{R}}(x, t_1), \quad E_2^{\mathcal{R}}(x, t_2) \quad (\text{c'est le même } x \text{ pour } E_1^{\mathcal{R}} \text{ et } E_2^{\mathcal{R}}). \quad (5.3)$$

(Dans l'exemple des horloges, E_1 pourrait être l'événement "le photon quitte le miroir du bas de l'horloge A" et E_2 l'événement "le photon revient au miroir du bas", mais le raisonnement que nous allons faire est vrai pour n'importe quel paire d'événements.) Le temps $\tau = t_2 - t_1$ est la durée propre entre les deux événements.

Dans le référentiel $\mathcal{R}'$ en mouvement à la vitesse $\vec{V} = V\vec{u}_x$ par rapport à $\mathcal{R}$, les dates t'_1 et t'_2 des deux événements sont, en utilisant les transformations de Lorentz,

$$t'_1 = \gamma\left(t_1 - \frac{Vx}{c^2}\right), \quad t'_2 = \gamma\left(t_2 - \frac{Vx}{c^2}\right). \quad (5.4)$$

La durée impropre entre les deux événements est donc

$$\tau' = t'_2 - t'_1 = \gamma(t_2 - t_1) = \gamma\tau, \quad (5.5)$$

comme annoncé.²

2. On peut retrouver le même résultat dans l'exemple des horloges par un calcul direct. On a déjà obtenu

5.2.3 Plus de détail sur la contraction des longueurs

Dans un référentiel $\mathcal{R}$ donné, il est facile de mesurer la longueur d'un objet immobile : on mesure la position x_A de l'extrémité gauche, la position x_B de l'extrémité droite, et la longueur est $L = x_B - x_A$. Une manière pratique d'effectuer ces mesures est d'utiliser, tout simplement une règle. Cette longueur L mesurée dans le référentiel où l'objet est au repos est la longueur propre de l'objet.

Mesurer la longueur d'un objet en mouvement est plus délicate à formaliser. Considérons donc un référentiel $\mathcal{R}'$ de vitesse $\vec{V}_{\mathcal{R}'/\mathcal{R}} = V\vec{u}_x$ par rapport à $\mathcal{R}$. Dans ce référentiel $\mathcal{R}'$, notre objet (qui est immobile dans $\mathcal{R}$ avec une longueur propre L) a une vitesse $-V\vec{u}_x$, voir figure 5.4.

On place un observateur immobile dans $\mathcal{R}'$, à l'origine O' de ce référentiel. Cet observateur voit passer devant lui le point A de l'objet à l'instant t'_A et le point B à l'instant t'_B . Le temps de passage $\tau' = t'_B - t'_A$ est le temps propre, puisqu'il est mesuré par un observateur immobile. Cet observateur, qui a vu que l'objet mettait un temps τ' à passer devant lui à la vitesse V , conclut que la longueur de l'objet est $L' = V\tau'$.

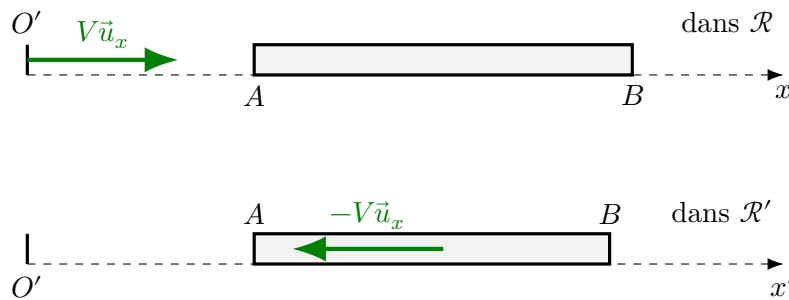


FIGURE 5.4 – Principes de mesure de la longueur de l'objet AB dans les référentiels $\mathcal{R}$ et $\mathcal{R}'$. Dans $\mathcal{R}$, l'objet est immobile et le point O' se déplace vers la droite, alors que dans $\mathcal{R}'$, le point O' est immobile et l'objet se déplace vers la gauche.

Plaçons maintenant un observateur immobile dans $\mathcal{R}$. Cet observateur pourrait simplement utiliser une règle pour mesurer la longueur L de l'objet (immobile dans ce référentiel), mais il décide de faire autrement. Il observe le point O' qui, de son point de vue se déplace vers la droite à la vitesse V . Pour l'observateur dans le référentiel $\mathcal{R}$, ce point passe devant les extrémités A et B de l'objet aux temps t_A et t_B . Le temps $\tau = t_B - t_A$ que met le point O' à passer devant l'objet est un temps impropre (les événements “ O' passe devant A ” et “ O' passe devant B ” n'ont pas lieu dans $\mathcal{R}$ au même endroit). L'observateur dans $\mathcal{R}$ conclut que l'objet a une longueur $L = V\tau$.

Les deux observateurs en $\mathcal{R}$ et $\mathcal{R}'$ ont donc respectivement trouvé des longueurs L et L' données par

$$L = V\tau, \quad L' = V\tau'.$$

On rappelle que τ' est la durée propre entre les événements “ O' passe devant A ” et “ O' passe devant B ”, et que τ est la durée impropre entre les mêmes événements. On a donc

$$\tau = \gamma\tau', \quad \text{ce qui implique} \quad L' = \frac{1}{\gamma}L.$$

Dans cet exemple, L est la longueur propre et L' la longueur impropre. On a bien retrouvé le résultat annoncé.

$\tau_A = 2d/c$ pour la durée propre de l'aller-retour, calculons-en la durée impropre. Dans le référentiel où l'horloge est en mouvement à la vitesse $\vec{V} = V\vec{u}_x$, Le photon montant a un vecteur vitesse de norme c et orienté par rapport à l'horizontal avec un angle θ : $\vec{v}_{\text{photon}} = c \cos \theta \vec{u}_x + c \sin \theta \vec{u}_y$.

Le temps que met le photon à atteindre la plaque du haut est donc $d/(c \sin \theta)$, et la durée (impropre) de l'aller-retour est $\tau_B = \tau'_A = 2d/(c \sin \theta) = \tau_A / \sin \theta$. Par ailleurs, on a $c \cos \theta = V$, puisque le photon reste en face du miroir. On a donc $\sin \theta = \sqrt{1 - \cos^2 \theta} = \sqrt{1 - (V/c)^2} = 1/\gamma$. On retrouve bien $\tau'_A = \gamma\tau_A$.

Pour l'instant, nous n'avons raisonné que sur des longueurs qui étaient définies uniquement sur la coordonnée impliquée dans le changement de référentiel (c'est-à-dire la coordonnée suivant $\vec{V}$, à savoir x dans les exemples précédents). Si on généralise à trois dimensions d'espace, la longueur propre d'un objet quelconque s'écrit dans le référentiel $\mathcal{R}(x, y, z, t)$ où il est immobile :

$$L_{\text{propre}} = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$$

Dans tout autre référentiel $\mathcal{R}'(x', y', z', t')$ en mouvement suivant x par rapport à $\mathcal{R}$, seul les coordonnées de la longueur suivant x est modifiée. On appliquera donc la contraction des longueurs seulement à Δx . La même longueur mesurée dans $\mathcal{R}'$ est donc :

$$L_{\text{impropre}} = \sqrt{\Delta x'^2 + \Delta y'^2 + \Delta z'^2} = \sqrt{\left(\frac{\Delta x}{\gamma}\right)^2 + \Delta y^2 + \Delta z^2}$$

5.3 Intervalle d'espace-temps Δs

5.3.1 Invariant relativiste Δs^2

En mécanique newtonienne, la durée séparant deux événements ou la longueur d'un objet sont invariantes par changement de référentiel. Nous venons de voir que cela n'est plus vraie en relativité restreinte.

Nous nous posons donc la question suivante : existe-t-il une quantité impliquant les coordonnées spatio-temporelles (coordonnées d'espace-temps) de deux événements E_1 et E_2 qui soit invariante dans le cadre de la relativité restreinte, c'est-à-dire invariante par transformations de Lorentz ?

Nous allons voir qu'une telle quantité existe.

Soient $\mathcal{R}$ et $\mathcal{R}'$ deux référentiels d'inertie en mouvement relatif avec la vitesse V constante suivant x . Dans $\mathcal{R}$ on note (x_1, t_1) les coordonnées de E_1 et (x_2, t_2) les coordonnées de E_2 . (x'_1, t'_1) et (x'_2, t'_2) sont donc les coordonnées dans $\mathcal{R}'$.

On calcule $\Delta t' = t'_2 - t'_1$ et $\Delta x' = x'_2 - x'_1$ en fonction de $\Delta t = t_2 - t_1$ et $\Delta x = x_2 - x_1$:

$$\begin{aligned}\Delta x' &= x'_2 - x'_1 = \gamma(x_2 - Vt_2) - \gamma(x_1 - Vt_1) = \gamma(\Delta x - V\Delta t) \\ \Delta t' &= t'_2 - t'_1 = \gamma\left[\left(t_2 - \frac{V}{c^2}x_2\right) - \left(t_1 - \frac{V}{c^2}x_1\right)\right] = \gamma\left(\Delta t - \frac{V}{c^2}\Delta x\right)\end{aligned}$$

On peut faire apparaître c afin d'obtenir des relations similaires :

$$\begin{aligned}\Delta x' &= \gamma\left(\Delta x - \frac{V}{c}c\Delta t\right) \\ c\Delta t' &= \gamma\left(c\Delta t - \frac{V}{c}\Delta x\right)\end{aligned}$$

Si on combine $\Delta x'^2$ et $c^2\Delta t'^2$, on peut faire apparaître le facteur γ^2 :

$$c^2\Delta t'^2 - \Delta x'^2 = \gamma^2\left(c^2\Delta t^2 + \frac{V^2}{c^2}\Delta x^2 - 2\Delta x\frac{V}{c}c\Delta t - \Delta x^2 - \frac{V^2}{c^2}c^2\Delta t^2 + 2c\Delta t\frac{V}{c}\Delta x\right)$$

$$c^2\Delta t'^2 - \Delta x'^2 = \gamma^2\left(1 - \frac{V^2}{c^2}\right)(c^2\Delta t^2 - \Delta x^2)$$

$$c^2\Delta t'^2 - \Delta x'^2 = c^2\Delta t^2 - \Delta x^2$$

On appelle cette grandeur le **carré de l'intervalle spatio-temporel (ou d'espace-temps)** entre les deux événements $E_1(x_1, t_1)$ et $E_2(x_2, t_2)$ et on la note Δs^2 :

$$\Delta s^2 = c^2 (t_2 - t_1)^2 - (x_2 - x_1)^2$$

Δs^2 est un **invariant relativiste**, c'est-à-dire que, quels que soient les référentiels d'inertie $\mathcal{R}$ et $\mathcal{R}'$, on aura toujours :

$$\Delta s'^2 = \Delta s^2$$

Concrètement, si on mesure dans $\mathcal{R}$ l'intervalle spatio-temporel Δs^2 et qu'il vaut -8 m^2 , alors sa mesure dans $\mathcal{R}'$ donnera aussi $\Delta s'^2 = -8 \text{ m}^2$.

À trois dimensions d'espace x , y et z , on peut généraliser l'écriture du carré de l'intervalle spatio-temporel :

$$\Delta s^2 = (c\Delta t)^2 - (\Delta x)^2 - (\Delta y)^2 - (\Delta z)^2$$

La quantité Δs est quant-à-elle appelée l'intervalle d'espace-temps entre les événements E_1 et E_2 .

5.3.2 Signe de Δs^2 et causalité

Le carré de l'intervalle spatio-temporel est un invariant relativiste : tout raisonnement fondé sur la valeur de cet invariant est donc valide quelque soit le référentiel. On verra dans la suite que, pour comparer deux événements E_1 et E_2 , on peut se baser sur le signe du Δs^2 qui les sépare.

Dans la suite, pour simplifier les notations, on s'intéresse à des événements ayant une seule coordonnée spatiale x et on considère la quantité :

$$\Delta s^2 = c^2 \Delta t^2 - \Delta x^2 \quad (5.6)$$

Si $\Delta s^2 < 0$

Dans ce cas, c'est la partie spatiale du carré de l'intervalle Δx^2 qui donne le signe à l'expression : **on parle d'intervalle de type "espace"**.

Les événements E_1 et E_2 ne peuvent pas être reliés par lien de "cause à effet". En effet, l'information voyageant entre les deux événements devrait se propager à la vitesse moyenne :

$$v = \frac{\Delta x}{\Delta t}$$

Comme $\Delta s^2 < 0$, cela implique que $v > c$, ce qui est interdit par les postulats de la relativité restreinte. Ainsi, aucune information ne peut être échangée entre les événements E_1 et E_2 et donc **les deux événements ne peuvent pas être reliés par des relations de cause à effet**.

Δt^2 peut être nul dans ce type d'intervalle. Ainsi, on peut trouver un référentiel où les événements E_1 et E_2 sont simultanés. On peut de même intervertir l'ordre temporel entre E_1 et E_2 par changement de référentiel. Δx^2 , par contre, ne peut pas être nul dans le cas $\Delta s^2 < 0$: il n'existe pas de référentiel d'inertie où les deux événements E_1 et E_2 ont lieu au même endroit.

Si $\Delta s^2 > 0$

Cette fois, c'est la partie temporelle de carré de l'intervalle Δt^2 qui impose le signe de Δs^2 . **On parle d'intervalle de type "temps"**.

Δx^2 peut être nul dans ce type d'intervalle. Ainsi, on peut trouver un référentiel où les événements E_1 et E_2 peuvent avoir lieu au même endroit. Δt^2 , par contre, ne peut pas être nul

dans ce type d'intervalle : on ne peut trouver aucun référentiel dans lequel les deux événements sont simultanés.

Montrons maintenant qu'il est impossible de trouver un référentiel d'inertie qui inverse l'ordre temporel de E_1 et E_2 . Puisque les deux événements sont séparés par un $\Delta s^2 > 0$, alors il existe un référentiel $\mathcal{R}$ dans lequel $\Delta x=0$ (avec $\Delta t \neq 0$). Dans n'importe quel référentiel $\mathcal{R}'$ en mouvement par rapport à $\mathcal{R}$, on peut écrire $\Delta t'$ l'intervalle temporel entre les deux événements, en fonction de Δt grâce aux transformations de Lorentz :

$$\Delta t' = \gamma \left(\Delta t - \frac{V}{c^2} \Delta x \right) = \gamma \Delta t.$$

Or ici, $\gamma > 1$, donc, quelque soient $\mathcal{R}$ et $\mathcal{R}'$, le signe de Δt est le même que celui de $\Delta t'$: on voit bien que les changements de référentiel ne peuvent pas, dans le cas $\Delta s^2 > 0$, modifier l'ordre temporel entre les événements E_1 et E_2 .

Dire que les événements E_1 et E_2 peuvent être reliés par relation de cause à effet dans ce cas revient à estimer, comme précédemment, quelle vitesse moyenne v devrait avoir une information pour passer de E_1 à E_2 . Dans le cas où $\Delta s^2 > 0$:

$$v = \frac{\Delta x}{\Delta t} < c$$

Ceci est évidemment autorisé par la relativité restreinte : **les deux événements séparés par un intervalle $\Delta s^2 > 0$ peuvent être reliés par cause à effet**. Mais, comme on ne peut en aucune façon trouver un référentiel dans lequel l'ordre temporel entre E_1 et E_2 s'inverse, alors, on ne peut dans le cadre de la relativité restreinte se retrouver dans des cas de paradoxe où la cause aurait lieu après la conséquence.

Si $\Delta s^2 = 0$

Le cas limite où $\Delta s^2 = 0$ correspond à deux événements qui pourraient être reliés uniquement par une information voyageant à la vitesse $v = \frac{\Delta x}{\Delta t} = c$, c'est-à-dire par un signal lumineux. **On parle donc d'intervalle de type "lumière"**.

5.3.3 Cônes de lumière

Choisissons, dans un référentiel $\mathcal{R}$ particulier, un événement E_0 quelconque. On peut alors délimiter des zones de l'espace-temps pouvant contenir tous les événements $E(x, y, z, t)$ susceptibles d'influencer ou d'être influencé par E_0 (zone où $\Delta s^2 > 0$) et une zone où l'ensemble des événements E ne peut influencer ou être influencé par l'événement E_0 (zone où $\Delta s^2 < 0$). Ces zones d'espace-temps sont alors délimitées par les lignes de lumières, c'est-à-dire les lignes correspondantes à $\Delta s^2 = 0$.

Pour simplifier la représentation, on va fixer E_0 comme l'origine du référentiel ($x_0 = y_0 = z_0 = 0$ et $t_0 = 0$). Si on ne considère qu'une seule dimension d'espace x , dans un espace (x, t) , les lignes de lumières auront l'équation :

$$\Delta s^2 = 0$$

$$c^2(t - t_0)^2 - (x - x_0)^2 = 0$$

$$x^2 = c^2 t^2$$

$$x = \pm ct$$

On obtient alors la figure 5.5. On a ainsi fait apparaître deux zones $\Delta s^2 > 0$, une vers les t positifs et une vers les t négatifs. Ces zones sont appelées "cône de lumière futur" et "cône

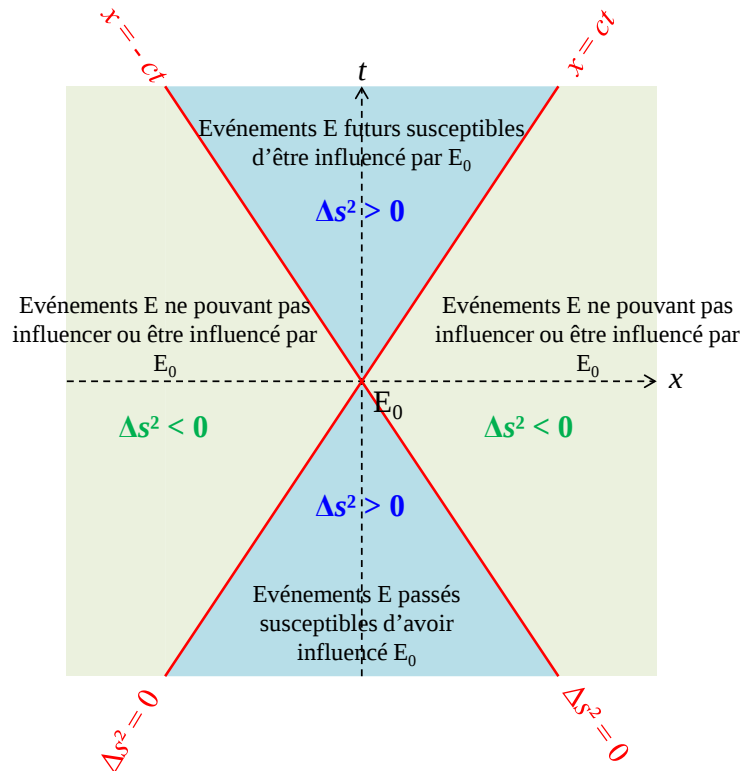


FIGURE 5.5 – Cône de lumière à une dimension d'espace

de lumière passé” de l'événement E_0 , et contiennent l'ensemble des événement pouvant être influencé par E_0 ou pouvant influencer E_0 .

L'appellation “cône” se justifie dans le cas où on ajoute des dimensions d'espace. À deux dimensions d'espace, les zones limitrophes $\Delta s^2 = 0$ sont définies dans l'espace (x, y, t) par :

$$c^2(t - t_0)^2 - (x - x_0)^2 - (y - y_0)^2 = 0$$

$$ct = \pm \sqrt{x^2 + y^2}$$

On reconnaît ici l'équation d'un cône. Une représentation graphique des cônes de lumière passé et futur est donc possible dans l'espace (x, y, t) .

À trois dimensions d'espace, il n'est plus possible de représenter les cônes de lumière : cela nécessiterait une représentation à quatre dimensions (x, y, z, t) . On parle cependant toujours, un peu abusivement, de “cône”.

Chapitre 6

Transformation des vitesses et effet Doppler

6.1 Composition des vitesses

On a vu dans le chapitre 4 que la loi de composition des vitesses issues des transformations de Galilée n'est pas acceptable car elle ne respect pas l'invariance de c par changement de référentiel. Nous allons maintenant établir les lois relativistes de composition des vitesses par changement de référentiel d'inertie à l'aide des transformations de Lorentz.

Considérons deux référentiels : un référentiel $\mathcal{R}(x, y, z, t)$ immobile, et un autre référentiel $\mathcal{R}'(x', y', z', t')$ animé dans $\mathcal{R}$ de la vitesse $\vec{V}$ constante. On choisira les axes x et x' confondus, de telle façon à ce que $\vec{V} = V\vec{u}_x$ avec V une constante positive. Enfin, on choisit l'origine des temps t et t' telle que, à $t = t' = 0$, les origines des coordonnées des deux référentiels soient confondues.

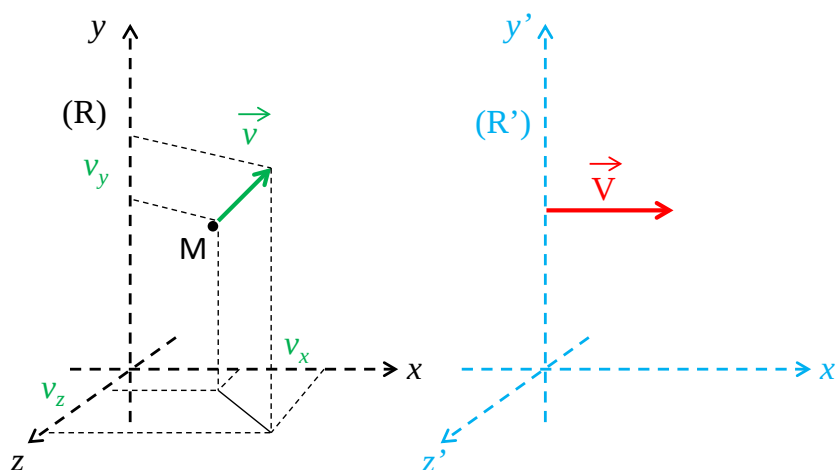


FIGURE 6.1 – Dessin dans $\mathcal{R}$ du mouvement et des coordonnées de la vitesse du point M.

Soit un point $M(x, y, z)$ animé d'une vitesse $\vec{v}$ quelconque dans $\mathcal{R}$. On cherche à exprimer $\vec{v}'$ sa vitesse dans $\mathcal{R}'$ grâce aux transformations de Lorentz.

Dans $\mathcal{R}$, par définition, on peut exprimer les composantes de $\vec{v}$:

$$v_x = \frac{dx}{dt} \quad v_y = \frac{dy}{dt} \quad v_z = \frac{dz}{dt}$$

Dans $\mathcal{R}'$, on fera la distinction entre la transformation de la composante de $\vec{v}'$ suivant x' , (donc dans la direction de $\vec{V}$) et suivant y' et z' (donc dans les directions orthogonales à $\vec{V}$).

Composante de $\vec{v}'$ dans la direction de $\vec{V}$

$$v'_x = \frac{dx'}{dt'} = \frac{d[\gamma(x - Vt)]}{d[\gamma(t - \frac{V}{c^2}x)]}$$

Comme V est une constante :

$$v'_x = \frac{\gamma(dx - Vdt)}{\gamma(dt - \frac{V}{c^2}dx)} = \frac{\frac{dx}{dt} - V}{1 - \frac{V}{c^2}\frac{dx}{dt}}$$

Soit :

$$\boxed{v'_x = \frac{v_x - V}{1 - \frac{Vv_x}{c^2}}} \quad (6.1)$$

Composante de $\vec{v}'$ dans les directions orthogonales à $\vec{V}$

$$\begin{aligned} v'_y &= \frac{dy'}{dt'} = \frac{dy}{d[\gamma(t - \frac{V}{c^2}x)]}, & v'_z &= \frac{dz'}{dt'} = \frac{dz}{d[\gamma(t - \frac{V}{c^2}x)]} \\ v'_y &= \frac{\frac{dy}{dt}}{\gamma(1 - \frac{V}{c^2}\frac{dx}{dt})}, & v'_z &= \frac{\frac{dz}{dt}}{\gamma(1 - \frac{V}{c^2}\frac{dx}{dt})}, \end{aligned}$$

et donc :

$$\boxed{v'_y = \frac{v_y}{\gamma(1 - \frac{Vv_x}{c^2})} \quad v'_z = \frac{v_z}{\gamma(1 - \frac{Vv_x}{c^2})}} \quad (6.2)$$

Attention, il ne faut pas confondre ici V qui est la composante de la vitesse entre $\mathcal{R}$ et $\mathcal{R}'$ suivant l'axe considéré (ici x), et v_x , v_y et v_z qui sont les coordonnées de la vitesse d'un point matériel dans le référentiel $\mathcal{R}$. On veillera en particulier au signe de V (si $\mathcal{R}'$ va dans le sens des x négatifs, alors on pourra changer dans les expressions précédentes V par $-V$ afin de garder le paramètre V positif).

Enfin, on peut vérifier sur un exemple que ces transformations laissent bien c invariante. Supposons que le point M se déplace à la vitesse $\vec{v} = c\vec{u}_x$ dans le référentiel $\mathcal{R}$; alors, dans le référentiel $\mathcal{R}'$ allant à $\vec{V} = V\vec{u}_x$ par rapport à $\mathcal{R}$, sa vitesse $\vec{v}'$ est :

$$\begin{aligned} v'_y &= v'_z = 0 \\ v'_x &= \frac{v_x - V}{1 - \frac{Vv_x}{c^2}} = \frac{c - V}{1 - \frac{V}{c}} = c \frac{1 - \frac{V}{c}}{1 - \frac{V}{c}} = c. \end{aligned}$$

On retrouve bien que $\|\vec{v}'\| = c$.

6.2 Aberrations — Transformations des angles

La loi de composition des vitesses permet de déterminer simplement la loi de transformations des angles en passant de $\mathcal{R}$ à $\mathcal{R}'$. Ce phénomène connu sous le nom d'aberrations est très bien illustré par la figure 6.2 issue du cours de physique de Berkeley. Un homme au repos sous son parapluie n'est pas mouillé par la pluie qui tombe selon l'axe de son parapluie. Il se met en mouvement et se mouille. Dans le référentiel lié à l'homme, la pluie a maintenant une composante horizontale de vitesse non nulle et l'angle sous lequel la pluie arrive est changé.

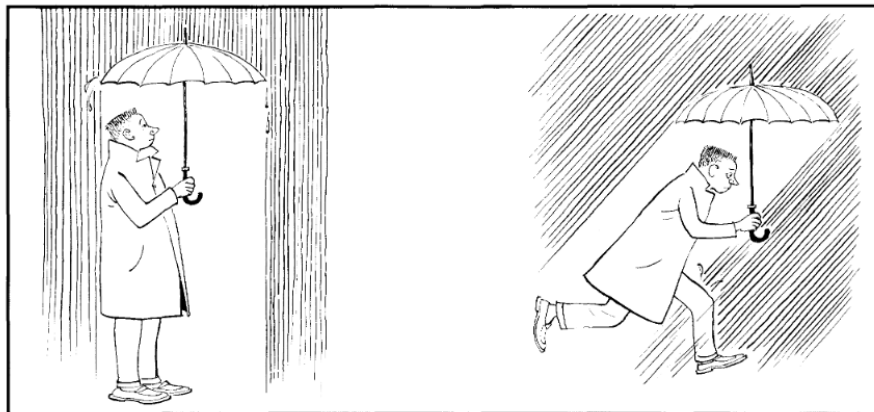


FIGURE 6.2 – Figure issue de *Berkeley Physics Course vol.1 : Mechanics* (Kittel/Knight/Ruderman). Légende originale (in english of course) : A homely example of aberration. This student is caught in rain coming straight down. If he stands still under his umbrella, he keeps dry. But If he runs for it... he gets wet. In his new reference frame the rain has horizontal velocity $-\mathbf{v}$, where $\mathbf{v}$ is his velocity with respect to the ground.

6.2.1 Formule générale

Nous allons considérer le cas suivant. Dans un référentiel $\mathcal{R}$ on émet depuis l'origine des coordonnées un signal qui se propage avec le vecteur vitesse $\vec{v}$ faisant un angle θ avec l'axe (Ox) , voir figure 6.5. On cherche à déterminer l'angle θ' que fait le vecteur $\vec{v}'$ avec l'axe (Ox') dans le référentiel $\mathcal{R}'$ qui est en translation uniforme par rapport à $\mathcal{R}$ avec la vitesse $\vec{V}_{\mathcal{R}'/\mathcal{R}} = V\vec{u}_x$.

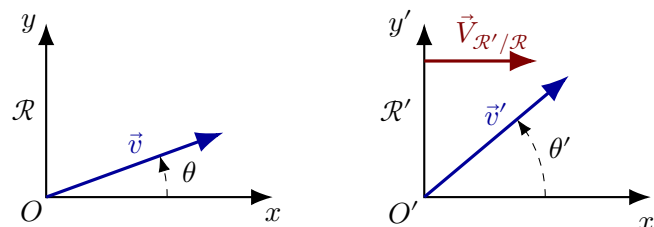


FIGURE 6.3 – Notations pour le calcul des formules de transformations des angles

Dans $\mathcal{R}$, on a :

$$\tan \theta = \frac{v_y}{v_x} \quad (6.3)$$

et dans $\mathcal{R}'$

$$\tan \theta' = \frac{v'_y}{v'_x} \quad (6.4)$$

En appliquant la loi de composition des vitesses on trouve :

$$\tan \theta' = \frac{v'_y}{v'_x} = \frac{v_y}{\gamma(1 - v_x V/c^2)} \frac{(1 - v_x V/c^2)}{v_x - V} = \frac{v_y}{\gamma v_x (1 - V/v_x)} \quad (6.5)$$

avec $\gamma = 1/\sqrt{1 - (V/c)^2}$

En remarquant que $\frac{v_y}{v_x} = \tan \theta$ et que $v_x = v \cos \theta$, on arrive à une formule des transformation des angles :

$$\tan \theta' = \frac{\tan \theta}{\gamma(1 - \frac{V}{v \cos \theta})} \quad (6.6)$$

En exercice, vous pouvez montrer que les transformations de Galilée mènent à une formule de transformation des angles (incorrecte) qui diffèrent de l'expression (6.6) par l'absence du facteur γ .

6.2.2 Aberrations optiques

Un cas d'intérêt particulier est celui où le signal émis en O est un signal lumineux. Dans ce cas, $v = c$ et en notant $\beta = V/c$, on arrive à la formule de transformations des angles pour les signaux lumineux (aberrations *optiques*) :

$$\tan \theta' = \frac{\tan \theta}{\gamma(1 - \frac{\beta}{\cos \theta})} = \frac{\sin \theta}{\gamma(\cos \theta - \beta)} \quad (6.7)$$

On laisse le soin à l'étudiant curieux de montrer les formules suivantes qui sont équivalentes à l'équation (6.7) :

$$\cos \theta' = \frac{\cos \theta - \beta}{1 - \beta \cos \theta}, \quad \sin \theta' = \frac{\sin \theta}{\gamma(1 - \beta \cos \theta)}. \quad (6.8)$$

Sur la partie gauche de la figure 6.4 sont représentées les angles θ' en fonction des angles θ pour $\beta = 0,9$. On voit que pour des angles $0 < \theta < \pi$, les angles θ' se rapprochent de π .

Dans le cas où $\beta \rightarrow 1$, quels que soient les directions d'émission des rayons lumineux dans $\mathcal{R}$ (à part $\theta = 0$), un observateur dans $\mathcal{R}'$ voit les rayons se propager au voisinage de la direction $-\vec{u}'_x$, voir partie droite de la figure. C'est donc que les rayons arrivent de face.

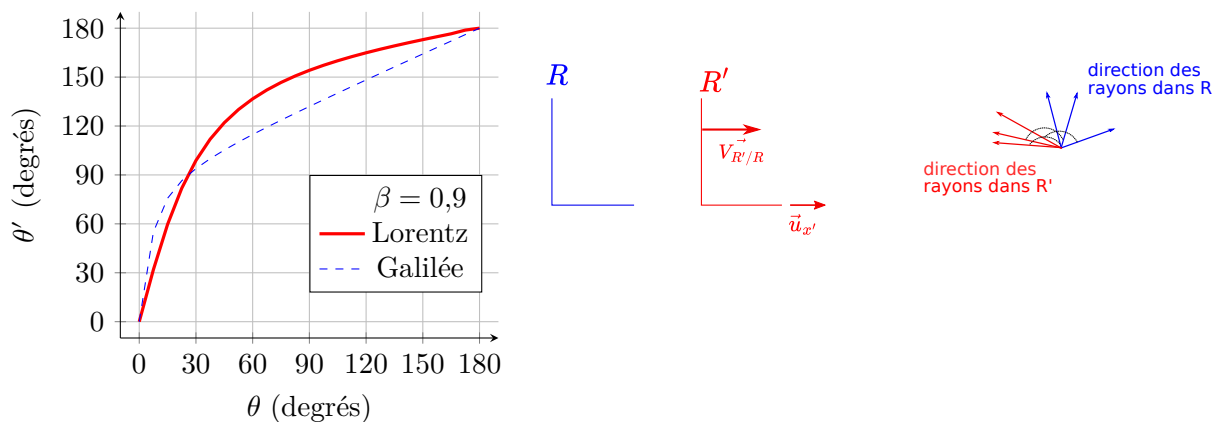


FIGURE 6.4 – Partie gauche : $\theta' = f(\theta)$ pour $\beta = 0,9$. Traits pleins : résultat (correct) obtenu à partir des transformations de Lorentz. Traits pointillés : résultat (incorrect) donné par les transformations de Galilée. Partie droite : lorsque $\beta \rightarrow 1$, pour des rayons émis environ isotropiquement dans $\mathcal{R}$, la direction des rayons dans $\mathcal{R}'$ se concentrent dans la direction $-\vec{u}'_x$.

En particulier, un observateur dans vaisseau spatial en direction des étoiles lointaines qui accélère soudainement dans la direction des étoiles les verra se concentrer vers l'avant du faisceau et certainement pas fuir vers l'arrière.

6.3 Effet Doppler

Si la source d'un signal périodique et l'observateur de ce signal sont en mouvement relatif, la fréquence du signal observé est différente de la fréquence du signal émis. Ce phénomène appelé "effet Doppler" est observé dans la vie de tous les jours. Par exemple, un observateur sur le bord

de la route perçoit le son de la sirène d'une ambulance qui s'approche de lui comme plus aigu que le son entendu par le conducteur de l'ambulance.

Nous voulons déterminer la relation existant entre les fréquences émises et observées dans le cadre de la relativité restreinte.

Pour cela, nous considérons nos deux référentiels $\mathcal{R}$ et $\mathcal{R}'$ avec $\vec{V}_{\mathcal{R}'/\mathcal{R}} = V\vec{u}_x$ avec V une constante positive. On choisit l'origine des temps t et t' telle que, à $t = t' = 0$, les origines des coordonnées des deux référentiels soient confondues.

6.3.1 Effet Doppler sonore

Intéressons-nous tout d'abord à un effet Doppler sonore. Nous ne traiterons que de l'effet Doppler longitudinal (ou colinéaire), c'est-à-dire lorsque la direction de propagation du signal périodique coïncide avec la direction du vecteur vitesse $\vec{V}_{\mathcal{R}'/\mathcal{R}} = V\vec{u}_x$.

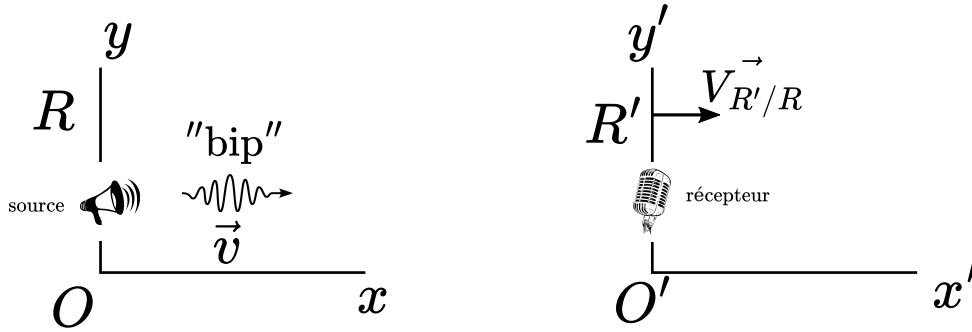


FIGURE 6.5 – Notations pour le calcul transformations des fréquences : effet Doppler

Une source sonore (une langue, un haut parleur) est fixe dans $\mathcal{R}$ en O et un détecteur sonore (une oreille, un microphone) est fixe dans $\mathcal{R}'$ en O' .

On considère les événements suivants :

- E_0** : la source (en $x = 0$) émet un “bip” sonore vers les x positifs à $t = 0$. Les coordonnées de E_0 dans $\mathcal{R}$ sont $(x_0 = 0, t_0 = 0)$.
- E_1** : le récepteur (en $x' = 0$) reçoit le premier “bip” émis par la source. Mais à $t = 0$ le récepteur est au niveau de la source car $x_O = x_{O'}$ et donc il reçoit instantanément ce premier “bip”. Les coordonnées de E_1 dans $\mathcal{R}$ sont $(x_1 = 0, t_1 = 0)$; en fait, E_0 et E_1 sont le même événement.
- E_2** : la source (en $x = 0$) émet un second “bip” sonore vers les x positifs à $t = T$. Les coordonnées de E_2 dans $\mathcal{R}$ sont $(x_2 = 0, t_2 = T)$.
- E_3** : le récepteur (en $x' = 0$) reçoit le second “bip” émis par la source. On note (x_3, t_3) les coordonnées de E_3 dans $\mathcal{R}$ qu'on cherche à déterminer.

Pour déterminer les coordonnées de E_3 , on notera v la vitesse de propagation du signal (ici du son dans un milieu, l'air par exemple). On fait l'hypothèse que l'air est au repos dans $\mathcal{R}$ et que $v > V^1$.

Le deuxième bip sonore part de la source à $t = T$ et se propage vers les x à la vitesse v . Son équation de propagation dans $\mathcal{R}$ est donc :

$$x_{\text{bip}}(t) = v(t - T) \quad (6.9)$$

L'équation de propagation dans $\mathcal{R}$ du récepteur fixe en O' est :

1. Que la vitesse de translation du référentiel soit supérieure à celle de la propagation du signal ($V > v$) est une possibilité pour le son dans un milieu (choc supersonique des avions supersoniques) mais pas pour la lumière dans le vide.

$$x_{\text{rec}}(t) = Vt \quad (6.10)$$

Pour déterminer t_3 , il suffit de résoudre l'équation $x_{\text{bip}}(t_3) = x_{\text{rec}}(t_3)$. Puis x_3 suit immédiatement car $x_3 = Vt_3$. On trouve :

$$t_3 = T \frac{v}{v - V} \quad \text{et} \quad x_3 = VT \frac{v}{v - V} \quad (6.11)$$

t_3 est le temps auquel arrive le deuxième bip dans $\mathcal{R}$. Calculons t'_3 , le temps auquel arrive le deuxième bip dans $\mathcal{R}'$, pour un observateur fixe dans $\mathcal{R}'$:

$$t'_3 = \gamma(t_3 - Vx_3/c^2) = \gamma(t_3 - VVt_3/c^2) = t_3/\gamma, \quad (6.12)$$

avec $\gamma = 1/\sqrt{1 - (V/c)^2}$.

Le premier bip est reçu par le récepteur en O' à $t'_1 = t'_0 = 0$. La durée qui sépare la réception des deux bips en O' est donc :

$$T' = t'_3 - t'_1 = t_3/\gamma - 0 = \frac{1}{\gamma} T \frac{v}{v - V} \quad (6.13)$$

T est la période du signal émis dans $\mathcal{R}$. T' est la période du signal détecté dans $\mathcal{R}'$. $\nu = 1/T$ est la fréquence du signal émis, $\nu' = 1/T'$ étant la fréquence du signal détecté.

On trouve donc que les fréquences émises et reçues sont liées par la relation :

$$\nu' = \gamma \nu \frac{v - V}{v} = \nu \gamma \left(1 - \frac{V}{v}\right). \quad (6.14)$$

Cette relation entre la fréquence sonore perçue par l'observateur ν' et celle émise par la source ν correspond au cas où le récepteur sonore est en mouvement par rapport au milieu de propagation et la source fixe. On laissera à l'étudiant consciencieux le soin de déterminer la formule pour l'effet Doppler sonore dans le cas où la source est en mouvement par rapport au milieu et le récepteur fixe.

6.3.2 Effet Doppler optique

Le bip périodique émis par la source peut être lumineux et non sonore. Par exemple, une source fixe dans $\mathcal{R}$ émet à fréquence ν un flash lumineux. Le signal se propage alors à la vitesse $v = c$ et contrairement au cas sonore, *il peut se propager dans le vide*.

Nous nous intéressons à ce cas-là, c'est-à-dire à l'effet Doppler pour un signal lumineux se propageant dans le vide.

On reprend rigoureusement le même raisonnement que précédemment en posant $v = c$. On trouve que la fréquence lumineuse ν' perçue par l'observateur fixe en O' est liée à la fréquence lumineuse ν émise par l'observateur fixe en O par la relation :

$$\nu' = \nu \gamma \left(1 - \frac{V}{c}\right) = \nu \frac{1 - \beta}{\sqrt{1 - \beta^2}} = \nu \sqrt{\frac{1 - \beta}{1 + \beta}}, \quad (6.15)$$

où $\beta = V/c$. La formule (6.15) a été établie dans le cas où le récepteur lumineux *s'éloigne* et la source lumineuse s'éloignent l'un de l'autre. Dans ce cas, la fréquence perçue par l'observateur est inférieure à celle émise par la source : $\sqrt{(1 - \beta)/(1 + \beta)} < 1$.

Si le récepteur et la source se rapprochent l'un de l'autre alors il faut changer V en $-V$ dans le raisonnement que nous avons suivi et on trouve :

$$\nu' = \nu \sqrt{\frac{1 + \beta}{1 - \beta}} \quad (\text{source et récepteur se rapprochant}) \quad (6.16)$$

Dans ce cas, la fréquence perçue par l'observateur est plus grande que celle émise par la source : $\sqrt{(1+\beta)/(1-\beta)} > 1$.

Deux remarques importantes concernant l'effet Doppler optique :

- La lumière se propageant dans le vide et non dans un milieu comme les ondes sonores, il n'y a pas de différence, pour l'effet Doppler optique, entre les cas “source en mouvement” et “récepteur en mouvement”. Les deux configurations mènent au même résultat et la seule chose qui compte est de savoir si la source et le récepteur se rapprochent ou s'éloignent l'un de l'autre.
- les fréquences ν et ν' des formules (6.15) et (6.16) peuvent être celles d'une onde électromagnétique. La longueur d'onde d'une onde électromagnétique monochromatique étant reliée à sa fréquence par la relation $\lambda = \frac{c}{\nu}$, les longueurs d'onde observées dans $\mathcal{R}'$ sont reliées à celles émises dans $\mathcal{R}$ (où la source est au repos) par :

$$\lambda' = \lambda \sqrt{\frac{1+\beta}{1-\beta}} \quad (\text{source et récepteur s'éloignant}). \quad (6.17)$$

$$\lambda' = \lambda \sqrt{\frac{1-\beta}{1+\beta}} \quad (\text{source et récepteur se rapprochant}). \quad (6.18)$$

6.3.3 Une application de l'effet Doppler optique : le *redshift* en astrophysique

Le décalage de fréquence entre le référentiel d'une source lumineuse et le référentiel de l'observateur est quelque chose de régulièrement rencontré en astrophysique. Beaucoup d'observations astronomiques ne consistent pas uniquement qu'à faire des images des objets (astres, galaxies, ...). La majorité du temps, la lumière qui nous parvient est analysée spectralement, c'est-à-dire que l'on analyse la structure en longueur d'onde (ou fréquence) du rayonnement observé (en d'autres mots on analyse les spectres de la lumière détectée). On constate alors que les longueurs d'onde associées à l'émission ou l'absorption d'éléments connus sont différentes de celle que l'on obtient en faisant le spectre expérimental de ces mêmes éléments en laboratoire.

Pour exprimer ce décalage de longueur d'onde, on utilise le **redshift** Z . Il est défini comme suit :

$$Z = \frac{\lambda_{\text{obs}} - \lambda_{\text{p}}}{\lambda_{\text{p}}}$$

où λ_{obs} est la longueur d'onde de la transition observée, et λ_{p} la longueur d'onde propre de cette même transition, c'est-à-dire mesurée dans un référentiel où la source est au repos. En pratique λ_{p} est déterminée en laboratoire par des mesures dites de spectroscopie.

- Si $Z > 0$, alors $\lambda_{\text{obs}} > \lambda_{\text{p}}$: le *redshift* positif se traduit donc par un décalage vers le rouge, correspondant à un éloignement entre source et récepteur.
- Si $Z < 0$, alors $\lambda_{\text{obs}} < \lambda_{\text{p}}$: le *redshift* négatif se traduit donc par un décalage vers le bleu, correspondant à un rapprochement entre source et récepteur.

Les *redshifts* observés peuvent être de différentes origines, que nous allons maintenant passer en revue.

Le *redshift* dû à l'expansion de l'univers

Des observations conduisent à dire que l'univers observable est en expansion. Cela revient à dire que plus l'objet observé est lointain (donc vieux car la lumière a mis plus de temps à nous parvenir), et plus les longueurs d'onde qui nous parviennent sont décalées vers le rouge car, entre

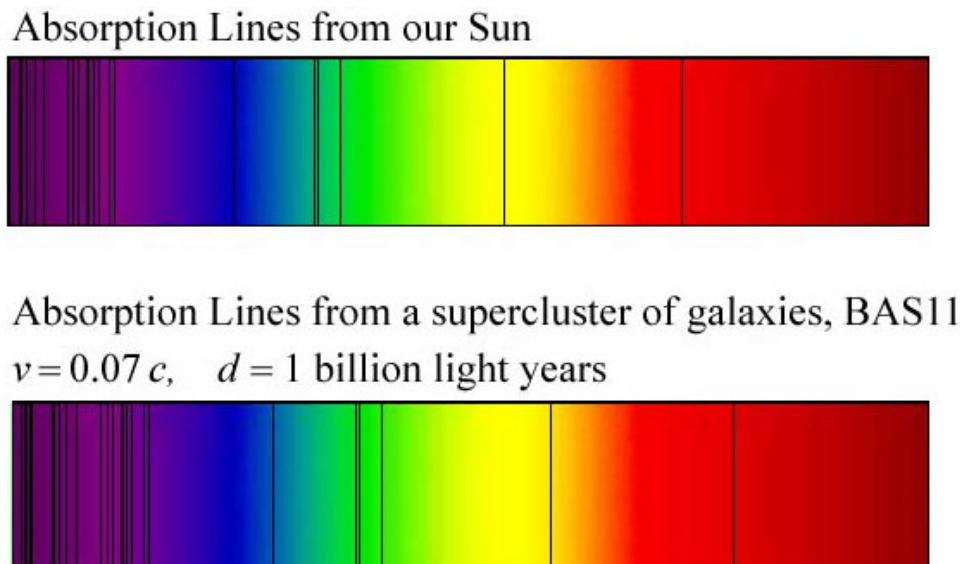


FIGURE 6.6 – Simulation du spectre d’absorption de notre soleil et d’un amas de galaxie lointaine. Toutes les raies d’absorption de l’amas de galaxie sont décalées vers le rouge par rapport à celle de notre étoile locale. L’amas de galaxie s’éloigne de nous. Image et simulation : Harold T. Stokes, Department of Physics and Astronomy, Brigham Young University, Utah, USA

temps, nous nous sommes beaucoup éloigné de cet objet. On traduit cet effet d’expansion par **la loi empirique de Hubble** qui est une conséquence directe de la formule de l’effet Doppler :

$$Z = H_0 \frac{d}{c}$$

où d est la distance nous séparant de l’objet en question, et H_0 est un paramètre appelé “constante” de Hubble. Ce n’est en réalité pas une constante à proprement parlé car, comme l’expansion de l’univers s’accélère, la vitesse entre l’objet et la Terre évolue en fonction du temps, et donc la valeur de H_0 dépend de la vitesse d’expansion à un instant donné.

Quoiqu’il en soit, comme la distance entre l’objet observé et la Terre évolue en fonction du temps, **on observera les objets les plus anciens (donc les plus lointain) à des grands *redshifts* Z . L’expansion induit toujours des *redshift* positifs. La mesure de Z permet donc de dater les objets astrophysiques.**

Exemples de valeurs de *redshifts* :

- Plus ancien rayonnement détecté sur Terre : **le fond diffus cosmologique (ou Cosmic Microwave Background CMB)** : $Z \approx 1000$. Ce rayonnement date de l’époque où l’univers primordial (environ 380 000 ans après le bigbang), initialement ionisé (gaz d’électrons et de protons) devient principalement neutre (électrons liés aux protons), ce qui a conduit le rayonnement à pouvoir être émis librement dans toutes les directions de l’espace. Initialement, ce rayonnement était dans le domaine des rayons X, mais on l’observe maintenant dans le domaine micro-onde/centimétrique ($0,1 < \lambda_{\text{obs}} < 100$ cm) à cause de l’expansion.
- Premières étoiles : on ne les détecte pas, mais les modèles cosmologiques prédisent un *redshift* de $Z \approx 30$.
- Plus vieilles galaxies détectées : on les observe avec un *redshift* $Z \approx 8$.

Le *redshift* cinétique

Ce *redshift* est directement dû au mouvement relatif des astres entre eux, et par rapport à la Terre. Il peut-être positif pour des objets s’éloignant, ou négatif pour des objets se rapprochant :

c'est très exactement l'effet Doppler tel qu'il est décrit plus tôt dans ce chapitre.

Un très bon exemple est le *redshift* observé sur la lumière qui nous provient de la galaxie d'Andromède : on observe toujours un décalage vers le bleu ($Z < 0$) d'environ $Z \simeq -0,001$. Cette observation nous a conduit à déterminer qu'en effet la galaxie se rapproche de nous (à une vitesse de l'ordre de 300 km/s), et rentrera en collision avec la voie lactée d'ici quelques millions d'années.

Souvent, ce *redshift* s'ajoute à celui dû à l'expansion. Ainsi, on peut observer le mouvement relatif, par exemple de gaz atomique chaud, dans une galaxie à haut *redshift*, en comparant le *redshift* des étoiles ("fixes") et celui du gaz.

Le *redshift* gravitationnel

On ne s'étendra pas trop sur ce point qui relève purement de la relativité générale. Dans le cadre de cette théorie, on montre qu'une masse très importante induit également une dilatation des durées, et donc des fréquences. Cela se traduit également par un *redshift*, souvent observé pour les objets à proximité d'amas stellaires, galactique, ou de trous noirs (centre galactique, par exemple).

Chapitre 7

Énergie et quantité de mouvement relativistes

7.1 De nouvelles expressions pour E et $\vec{p}$

En mécanique, la quantité de mouvement (ou impulsion) totale et l'énergie totale d'un système isolé sont deux quantités conservées. En mécanique classique, la quantité de mouvement d'une particule est $\vec{p} = m\vec{v}$, et son énergie (quand elle est isolée) est $E = \frac{1}{2}m\vec{v}^2$. En relativité restreinte, ces expressions doivent être changées : on va montrer dans ce chapitre que, pour respecter les principes de la relativité restreinte, il faut maintenant écrire

$$\boxed{\vec{p} = \gamma_v m \vec{v}, \quad E = \gamma_v m c^2 \quad \text{avec} \quad \gamma_v = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}} \quad (7.1)$$

Remarques :

- La vitesse v qui intervient dans γ_v est la vitesse de la particule dans le référentiel de travail ; ce n'est pas, comme dans le facteur de Lorentz γ , la vitesse d'un référentiel par rapport à un autre référentiel.
- La norme $\|\vec{p}\|$ de la quantité de mouvement et l'énergie E divergent lorsque la vitesse v tend vers c : il est difficile pour une particule massive d'avoir une vitesse v proche de c , et il est impossible que sa vitesse atteigne c , puisqu'elle aurait alors une énergie infinie.
- À vitesse nulle ($v = 0$), la particule a une énergie “de repos”, ou énergie interne, $E_{\text{int}} = mc^2$. On écrit parfois $E = E_{\text{int}} + E_c$ avec $E_c = (\gamma_v - 1)mc^2$ l'énergie cinétique de la particule, mais il est souvent plus simple de raisonner directement avec l'énergie totale E .
- Lorsque $v \ll c$, on peut écrire $\gamma_v \simeq 1 + \frac{v^2}{2c^2}$ et on trouve, au premier ordre intéressant, $\vec{p} \simeq m\vec{v}$ et $E \simeq mc^2 + \frac{1}{2}mv^2$. On retrouve dans ce développement les expressions classiques de la quantité de mouvement et de l'énergie cinétique.
- La masse m de la particule dans (7.1) est indépendante de sa vitesse v .¹
- Le cas des particules de masse nulle $m = 0$ et en particulier du photon est à part. Un photon possède une énergie et une quantité de mouvement mais celles-ci ne sont pas données par les formules (7.1).

1. Dans certains vieux livres, on définit parfois une masse au repos m_0 et une masse en mouvement $m_v = \gamma_v m_0$, ce qui permet d'écrire $\vec{p} = m_v \vec{v}$ et $E = m_v c^2$. Dans ces vieux livres, il est alors dit que la “masse” m_v augmente avec la vitesse. Dans ce cours, on utilisera exclusivement l'approche moderne, où la masse m est indépendante de la vitesse.

Lorsque le système n'est pas isolée, la quantité de mouvement varie. Par définition, la dérivée de la quantité de mouvement est la force totale extérieure $\vec{F}_{\text{ext}}$:

$$\boxed{\frac{d\vec{p}}{dt} = m \frac{d(\gamma_v \vec{v})}{dt} = \vec{F}_{\text{ext}}} \quad (7.2)$$

C'est le **principe fondamental de la dynamique** adapté à la mécanique relativiste. Dans la limite classique $v \ll c$ on a $\gamma_v \simeq 1$ et on retrouve l'expression classique $m\vec{a} = \vec{F}_{\text{ext}}$ avec $\vec{a} = d\vec{v}/dt$ l'accélération. On voit cependant que l'accélération $\vec{a}$ n'est pas une quantité très intéressante en mécanique relativiste.

Nous allons d'abord vérifier sur quelques exemples que les expressions (7.1) satisfont au principe de relativité, puis nous montrerons comment on peut obtenir ces expressions à partir de la mécanique lagrangienne. Enfin, nous en explorerons plusieurs conséquences.

7.1.1 Quantité de mouvement relativiste

Dans le référentiel d'étude $\mathcal{R}_0$, on considère une collision élastique entre deux particules identiques de masse m qui se déplacent l'une vers l'autre à une vitesse initiale v . Après la collision, les particules s'éloignent l'une de l'autre avec, en norme, la même vitesse v , dans une direction différente de la direction incidente. On choisit les axes $\vec{e}_x$ et $\vec{e}_y$ selon les axes de symétrie de la figure de collision, de façon à ce que la particule 1 progresse toujours dans la direction des x croissants, voir figure 7.1.

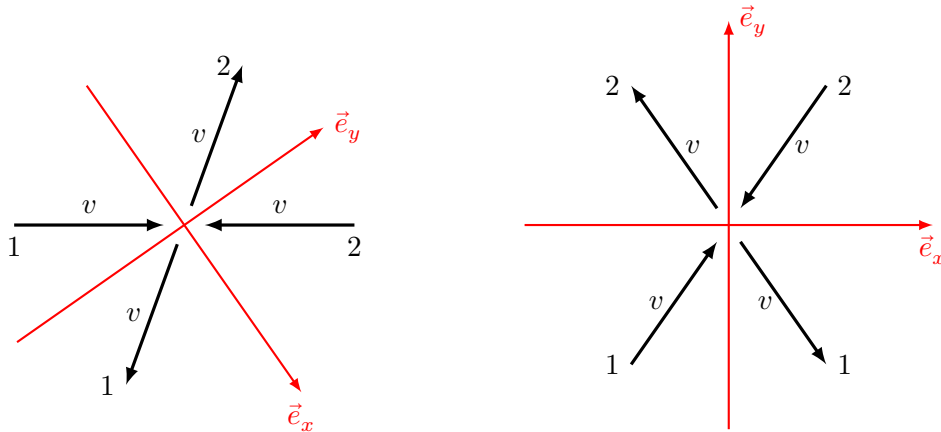


FIGURE 7.1 – Une collision entre deux particules identiques, notées 1 et 2. La figure de droite est identique à celle de gauche après rotation.

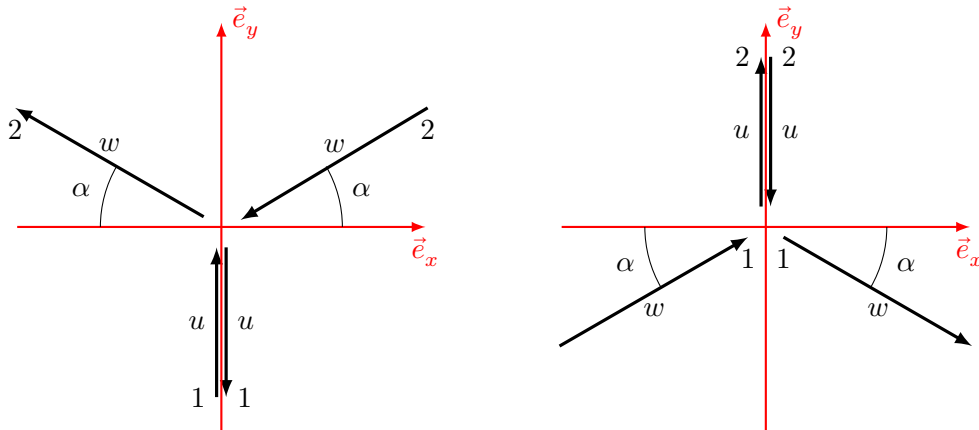
On considère le référentiel $\mathcal{R}$ qui est en translation uniforme selon $\vec{e}_x$ par rapport à $\mathcal{R}_0$ de manière à ce que la particule 1 ait un mouvement selon $\vec{e}_y$ dans $\mathcal{R}$. (C'est-à-dire que $\mathcal{R}$ est lié à un observateur se déplaçant selon $\vec{e}_x$ pour rester "à la verticale" de la particule 1.) On appelle u et w les normes des vitesses des deux particules dans $\mathcal{R}$, u étant la vitesse verticale, et α l'angle que fait la particule 2 avec l'axe $\vec{e}_x$, voir la partie gauche de la figure 7.2.

Notons $\vec{v}_{1,i}$ et $\vec{v}_{2,i}$ les vecteurs vitesse initiaux (avant la collision) des deux particules, et $\vec{v}_{1,f}$ et $\vec{v}_{2,f}$ les vecteurs vitesse finaux (après le choc). On a, avec les notations de la figure,

$$\vec{v}_{1,i} = \begin{pmatrix} 0 \\ u \end{pmatrix}, \quad \vec{v}_{2,i} = \begin{pmatrix} -w \cos \alpha \\ -w \sin \alpha \end{pmatrix}, \quad \vec{v}_{1,f} = \begin{pmatrix} 0 \\ -u \end{pmatrix}, \quad \vec{v}_{2,f} = \begin{pmatrix} -w \cos \alpha \\ w \sin \alpha \end{pmatrix}. \quad (7.3)$$

En utilisant l'expression $\vec{p} = \gamma_v m \vec{v}$, voir (7.1), la conservation de la quantité de mouvement s'écrit, après simplification par m ,

$$\gamma_u \vec{v}_{1,i} + \gamma_w \vec{v}_{2,i} = \gamma_u \vec{v}_{1,f} + \gamma_w \vec{v}_{2,f}. \quad (7.4)$$

FIGURE 7.2 – La même collision dans les référentiels $\mathcal{R}$ (à gauche) et $\mathcal{R}'$ (à droite).

On n'utilise pas, pour l'instant, l'expression de γ_v donnée dans (7.1), on va la retrouver par le calcul. En projetant (7.4) sur l'axe des x , on obtient $-\gamma_w w \cos \alpha = -\gamma_w w \cos \alpha$, ce qui n'est pas très intéressant. En projetant sur l'axe des y , on obtient, après simplification,

$$\gamma_u u = \gamma_w w \sin \alpha. \quad (7.5)$$

Pour aller plus loin, il nous faut une autre relation entre u , w et α . On va l'obtenir en faisant un changement de référentiel. On considère donc le référentiel $\mathcal{R}'$ ayant une vitesse $-w \cos \alpha \vec{e}_x$ par rapport à $\mathcal{R}$. Dans ce référentiel, c'est la particule 2 qui a un mouvement selon $\vec{e}_y$. ($\mathcal{R}'$ est le référentiel lié à un observateur se déplaçant selon $\vec{e}_x$ pour rester "à la verticale" de la particule 2.) Les formules (6.1) et (6.2) de changement de vitesse donnent, dans $\mathcal{R}'$,

$$\vec{v}'_{1,i} = \begin{vmatrix} w \cos \alpha \\ \frac{u}{\gamma} \end{vmatrix}, \quad \vec{v}'_{2,i} = \begin{vmatrix} 0 \\ (\dots) \end{vmatrix}, \quad \vec{v}'_{1,f} = \begin{vmatrix} w \cos \alpha \\ -\frac{u}{\gamma} \end{vmatrix}, \quad \vec{v}'_{2,f} = \begin{vmatrix} 0 \\ (\dots) \end{vmatrix} \quad \text{avec } \gamma = \frac{1}{\sqrt{1 - \frac{w^2 \cos^2 \alpha}{c^2}}}. \quad (7.6)$$

Ici, γ est le facteur de Lorentz associée au changement de référentiel. Les points de suspension remplacent des expressions compliquées que l'on ne va pas utiliser.

Mais la collision dans le référentiel $\mathcal{R}'$ est la même que dans le référentiel $\mathcal{R}$, à part que les rôles des particules 1 et 2 ont été inversées. (Dans le référentiel $\mathcal{R}'$, la particule 1 a la vitesse w et la particule 2 la vitesse u !) En comparant $\vec{v}'_{2,i}$ et $\vec{v}'_{1,i}$, on en déduit que l'on a

$$\frac{u}{\gamma} = w \sin \alpha. \quad (7.7)$$

Avec (7.5), on en déduit que

$$\gamma_u = \frac{\gamma_w}{\gamma} = \gamma_2 \sqrt{1 - \frac{w^2 \cos^2 \alpha}{c^2}}. \quad (7.8)$$

Rappelons que γ_u est une fonction de u et γ_w la même fonction appliquée à w . Dans la limite où u est très petit, on doit avoir $\gamma_u \simeq 1$: c'est la limite classique. On doit aussi avoir $\alpha \simeq 0$, voir (7.5). On trouve alors, dans cette limite,

$$1 = \gamma_w \sqrt{1 - \frac{w^2}{c^2}}. \quad (7.9)$$

La formule générale pour γ_v est donc forcément celle donnée dans (7.1). Il faudrait alors, pour bien faire, vérifier que cette formule fonctionne dans (7.8) même lorsque u n'est pas petit, c'est-à-dire que l'on doit avoir

$$\frac{1}{\sqrt{1 - \frac{u^2}{c^2}}} = \frac{1}{\sqrt{1 - \frac{w^2}{c^2}}} \times \sqrt{1 - \frac{w^2 \cos^2 \alpha}{c^2}}, \quad (7.10)$$

sachant que l'on a aussi (7.7). Cette vérification est laissée en exercice.

7.1.2 Énergie cinétique relativiste

On peut retrouver l'expression relativiste de l'énergie cinétique E_c en écrivant le théorème de l'énergie cinétique ;

$$dE_c = \delta W = \vec{F}_{\text{ext}} \cdot d\vec{r} = \vec{F}_{\text{ext}} \cdot \vec{v} dt = m \frac{d(\gamma_v \vec{v})}{dt} \cdot \vec{v} dt = m d(\gamma_v \vec{v}) \cdot \vec{v}. \quad (7.11)$$

(δW est le travail infinitésimal fourni par la force $\vec{F}_{\text{ext}}$. On a utilisé le principe fondamental de la dynamique (7.2) pour exprimer $\vec{F}_{\text{ext}}$.)

En utilisant les propriétés des différentielles, on a

$$dE_c = m d(\gamma_v \vec{v}) \cdot \vec{v} = m [d(\gamma_v v^2) - \gamma_v \vec{v} \cdot d\vec{v}]. \quad (7.12)$$

(c'est l'idée à la base de l'intégration par parties...) On remarque que

$$d\left(\frac{1}{\gamma_v}\right) = d\left(\sqrt{1 - \frac{v^2}{c^2}}\right) = \frac{-2\vec{v} \cdot d\vec{v}/c^2}{2\sqrt{1 - \frac{v^2}{c^2}}} = -\frac{1}{c^2} \gamma_v \vec{v} \cdot d\vec{v}, \quad (7.13)$$

ce qui donne une expression pour $\gamma_v \vec{v} \cdot d\vec{v}$. On reporte dans (7.12) pour obtenir

$$dE_c = m d\left[\gamma_v v^2 + \frac{c^2}{\gamma_v}\right]. \quad (7.14)$$

Mais

$$\gamma_v v^2 + \frac{c^2}{\gamma_v} = \gamma_v \left(v^2 + \frac{c^2}{\gamma_v^2}\right) = \gamma_v \left(v^2 + c^2 \left(1 - \frac{v^2}{c^2}\right)\right) = \gamma_v c^2. \quad (7.15)$$

On obtient donc finalement

$$dE_c = d(\gamma_v m c^2), \quad \text{et donc} \quad E_c = \gamma_v m c^2 + \text{Cste}. \quad (7.16)$$

On choisit la constante de manière à avoir $E_c = 0$ lorsque $v = 0$, et on obtient

$$\boxed{E_c = (\gamma_v - 1) m c^2} \quad (7.17)$$

7.1.3 Énergie interne et énergie totale

Pour établir l'expression relativiste de l'énergie interne, puis de l'énergie totale, d'une particule, on va procéder de manière similaire à ce que l'on a fait pour la quantité de mouvement. On pose que, pour un système isolé, l'énergie totale se conserve lors d'une transformation. Dans le cas d'une collision inélastique, une partie de l'énergie cinétique est convertie en énergie interne.

On considère deux particules identiques de masse m , allant toutes deux à la vitesse v , et faisant une collision frontale. À l'issue de la collision, toute l'énergie cinétique est absorbée et les deux particules fusionnent pour créer une particule plus grosse, et immobile ; voir la partie gauche de la figure 7.3.

Un premier résultat, crucial, qui sortira du calcul est que

À partir de la collision de deux particules de masse m , on obtient une particule dont la masse *n'est pas* égale à $2m$.

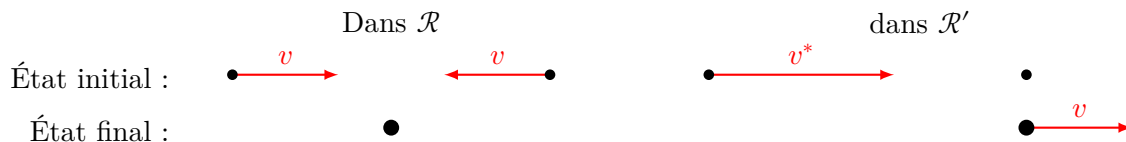


FIGURE 7.3 – Une collision inélastique entre deux particules identiques (en haut) qui se combinent pour former une particule plus grosse (en bas). À gauche, la collision est vue dans le référentiel $\mathcal{R}$ où les deux particules incidentes ont la même vitesse v (en norme). À droite, elle est vue dans le référentiel $\mathcal{R}'$ où la particule de droite est immobile.

On va poser m' la masse de la particule produite, et on va bien voir lors du calcul qu'il faudra prendre $m' \neq 2m$.

On pose E_{int} l'énergie interne d'une particule incidente, et E'_{int} l'énergie interne de la particule produite. L'énergie totale de chacune des deux particules incidentes est $E_{\text{int}} + mc^2(\gamma_v - 1)$, le deuxième terme étant l'énergie cinétique. La particule produite, immobile, n'a pas d'énergie cinétique et son énergie totale est donc égale à E'_{int} , son énergie interne. Le bilan d'énergie s'écrit donc

$$2[E_{\text{int}} + mc^2(\gamma_v - 1)] = E'_{\text{int}}. \quad (7.18)$$

On considère maintenant la même collision dans le référentiel $\mathcal{R}'$, qui se déplace vers la gauche à la vitesse v par rapport au référentiel de départ. Dans $\mathcal{R}'$, la particule de droite est initialement immobile; son énergie est par définition E_{int} . La particule de gauche se déplace initialement à la vitesse

$$v^* = \frac{2v}{1 + v^2/c^2}, \quad (7.19)$$

voir (6.1); son énergie est donc $E_{\text{int}} + mc^2(\gamma_{v^*} - 1)$. Enfin, la particule obtenue lors de la collision a une vitesse v ; son énergie est donc $E'_{\text{int}} + m'c^2(\gamma_v - 1)$. Le bilan d'énergie écrit dans $\mathcal{R}'$ donne donc

$$2E_{\text{int}} + mc^2(\gamma_{v^*} - 1) = E'_{\text{int}} + m'c^2(\gamma_v - 1). \quad (7.20)$$

En faisant la différence entre (7.20) et (7.18), on élimine les énergies internes E_{int} et E'_{int} . Après division par c^2 , on obtient

$$m(\gamma_{v^*} - 2\gamma_v + 1) = m'(\gamma_v - 1). \quad (7.21)$$

Calculons γ_{v^*} :

$$\gamma_{v^*} = \left[1 - \frac{1}{c^2} \left(\frac{2v}{1 + v^2/c^2} \right)^2 \right]^{-1/2} = \left[\frac{(1 + v^2/c^2)^2 - 4v^2/c^2}{(1 + v^2/c^2)^2} \right]^{-1/2} \quad (7.22)$$

$$= \left[\frac{(1 - v^2/c^2)^2}{(1 + v^2/c^2)^2} \right]^{-1/2} = \frac{1 + v^2/c^2}{1 - v^2/c^2} \quad (7.23)$$

$$= \frac{2}{1 - v^2/c^2} - 1 = 2\gamma_v^2 - 1. \quad (7.24)$$

On a donc $\gamma_{v^*} - 2\gamma_v + 1 = 2\gamma_v^2 - 2\gamma_v = 2\gamma_v(\gamma_v - 1)$. Dans (7.21), on obtient alors après simplification

$$m' = 2m\gamma_v, \quad (7.25)$$

qui est supérieur à $2m$. Remarquons que la masse de la particule obtenue dépend de la vitesse v des particules incidentes !

On cherche maintenant à déterminer les énergies internes. Dans (7.18), on développe et on passe le terme avec γ_v de l'autre côté; en utilisant (7.25), on obtient

$$2[E_{\text{int}} - mc^2] = E'_{\text{int}} - m'c^2. \quad (7.26)$$

Pour que cette équation soit vérifiée quels que soient les choix de m et de m' , la seule solution est de prendre

$$E_{\text{int}} = mc^2 \quad (7.27)$$

On comprends avec cette expression que la masse d'une particule (ou en fait, d'un système quelconque) est directement proportionnelle à son énergie interne. La raison pour laquelle la masse m' de l'exemple ci-dessus augmente avec la vitesse v d'impact est maintenant claire : plus v est importante, plus il y a d'énergie cinétique à absorber, et donc plus l'énergie interne de la particule produite est grande et, finalement, plus sa masse m' est grande.

Cette expression de l'énergie interne montre que l'on ne peut séparer les notions de masse et d'énergie d'un objet : on parle alors d'*équivalence masse-énergie*. On verra dans la suite que cette équivalence a d'importantes conséquences, notamment car elle suppose que l'on peut “transformer” la masse en énergie (et réciproquement) comme c'est le cas par exemple dans les réactions nucléaires.

Pour une particule isolée (non soumise à un potentiel extérieur), l'énergie totale E est la somme de l'énergie cinétique (7.17) et de l'énergie interne (7.27). On trouve bien, comme annoncée,

$$E = \gamma_v mc^2 \quad (7.28)$$

7.1.4 Relation entre énergie et impulsion

En partant des expressions de E , $\vec{p}$ et γ , on arrive après un petit calcul (à faire en exercice) à la relation importante suivante :

$$E^2 = m^2 c^4 + p^2 c^2 \quad (7.29)$$

Une manière intéressante de lire cette relation est la suivante : l'énergie et l'impulsion d'une particule dépendent du référentiel choisi, mais la quantité $(E/c)^2 - p^2$ n'en dépend pas. C'est la même situation que pour les écarts spatiaux et temporels entre deux événements : Δr et Δt dépendent du référentiel, mais $(c\Delta t)^2 - \Delta r^2$ n'en dépend pas.

En fait, on peut montrer avec un peu de calcul que E/c et $\vec{p}$ se transforment selon la *même transformation de Lorentz* que les coordonnées ct et $\vec{r}$ lors d'un changement de référentiel : si $\mathcal{R}'$ se déplace à la vitesse $V = \beta c$ selon l'axe des x par rapport à $\mathcal{R}$, alors

$$\begin{cases} E/c = \gamma(E'/c + \beta p'_x) & \text{à comparer à } ct = \gamma(ct' + \beta x'), \\ p_x = \gamma(p'_x + \beta E'/c) & \text{à comparer à } x = \gamma(x' + \beta ct'), \\ p_y = p'_y & \text{à comparer à } y = y', \\ p_z = p'_z & \text{à comparer à } z = z'. \end{cases} \quad (7.30)$$

De même que (ct, x, y, z) est le **quadrivecteur temps-position**, on appelle $(E/c, p_x, p_y, p_z)$ le **quadrivecteur énergie-impulsion**.

7.1.5 Énergie et impulsion du photon.

Le photon est un quanta d'énergie lumineuse, certains parlent d'une “particule ou d'un corpuscule de lumière”.

La masse du photon est nulle $m = 0$.

Le photon est toujours en mouvement. Sa vitesse dans vide dans **tous** les référentiels est c . Une conséquence de cela est que le **référentiel du photon** (où le photon dans le vide est au repos) **n'existe pas**.

La mécanique quantique associe au photon une énergie :

$$\boxed{E = h\nu} \quad (7.31)$$

où $h = 6,626\,070\,15 \cdot 10^{-34} \text{ J} \cdot \text{s}$ est la constante de Planck et ν la fréquence de l'onde électromagnétique associée au photon dans un référentiel donné.

La relation (7.29) se simplifie pour les photons. Elle devient :

$$E = pc \quad \text{si } m = 0, \quad (7.32)$$

ce qui permet de définir l'impulsion du photon :

$$\boxed{\vec{p} = \frac{E}{c} \vec{u} = \frac{h\nu}{c} \vec{u}} \quad (7.33)$$

où $\vec{u}$ est un vecteur unitaire dans la direction de propagation de l'onde lumineuse associée au photon dans un référentiel donné.

Puisqu'un photon a une quantité de mouvement non nulle, la lumière peut “pousser” la matière.

7.2 Mécanique lagrangienne relativiste

Dans cette section, on va déterminer le lagrangien relativiste $\mathcal{L}(\vec{r}, \vec{v})$ pour une particule, puis retrouver, à partir de ce lagrangien, les expressions (7.1).

On rappelle le principe de moindre action, voir section 2.4 : dans le référentiel (galiléen) $\mathcal{R}$ d'étude, parmi toutes les trajectoires envisageables entre le point $\vec{r}_1$ au temps t_1 et le point $\vec{r}_2$ au temps t_2 , la trajectoire physique $\vec{r}(t)$ est celle qui extrémise (souvent : minimise) l'action

$$S[r] = \int_{t_1}^{t_2} dt \mathcal{L}(\vec{r}(t), \dot{\vec{r}}(t)). \quad (7.34)$$

Ici, il faut penser à une trajectoire comme à une longue succession d'événements infiniment proches : il y a l'événement de départ E_1 , de coordonnées $(\vec{r}_1, t_1)$ dans $\mathcal{R}$, puis l'événement suivant dont les coordonnées dans $\mathcal{R}$ sont $(\vec{r}(t_1 + dt), t_1 + dt)$, etc., jusqu'à l'événement final E_2 dont les coordonnées dans $\mathcal{R}$ sont $(\vec{r}_2, t_2)$.

Pour satisfaire au principe de relativité, le problème à résoudre est le suivant : comment choisir $\mathcal{L}$, et donc S , de manière à ce que la trajectoire physique soit indépendante du référentiel, c'est-à-dire passe par la même succession d'événements, quel que soit le référentiel. Bien sûr, les coordonnées $(\vec{r}(t), t)$ des événements dépendront du référentiel, mais pas les événements eux-mêmes.

Pour garantir le principe de relativité, la solution la plus simple, et qui s'avère correspondre à la réalité, est de faire en sorte que *l'action S soit un invariant relativiste*, c'est-à-dire que S ne dépende pas du référentiel. Ainsi, la succession d'événements qui minimise S sera automatiquement indépendante du référentiel.

Pour une particule libre, la seule manière d'obtenir une action qui soit un invariant relativiste et qui s'écrive comme une intégrale le long de la trajectoire est la suivante :

$$S[\text{trajectoire}] = K \int_{E_1}^{E_2} ds, \quad (7.35)$$

où K est une constante à déterminer, E_1 et E_2 sont respectivement les événements de départ et d'arrivée, et où ds est *l'intervalle d'espace temps* infinitésimal entre deux événements successifs de la trajectoire, voir section 5.3.

La signification de (7.35) est la suivante :

- On découpe la trajectoire en une succession d'événements infiniment proches, partant de E_1 et arrivant en E_2 .
- On calcule les intervalles d'espace temps infinitésimaux ds entre deux événements successifs le long de la trajectoire, et on somme tous ces intervalles.
- On multiplie par K , à déterminer.

Comme les intervalles d'espace temps infinitésimaux ds sont des invariants relativistes, on obtient automatiquement une action S indépendante du référentiel.

On se place maintenant dans un référentiel $\mathcal{R}$ donné. Dans ce référentiel, les événements composant la trajectoire sont de la forme $(\vec{r}(t), t)$. L'intervalle d'espace temps infinitésimal ds entre deux événements successifs s'obtient alors en écrivant

$$(ds)^2 = c^2(dt)^2 - (d\vec{r})^2 = c^2(dt)^2 \left[1 - \frac{1}{c^2} \left(\frac{d\vec{r}}{dt} \right)^2 \right] = c^2(dt)^2 \left[1 - \frac{(\dot{\vec{r}})^2}{c^2} \right]. \quad (7.36)$$

On prend la racine carrée et on reporte dans (7.35) :

$$S[\vec{r}(t)] = Kc \int_{t_1}^{t_2} dt \sqrt{1 - \frac{(\dot{\vec{r}})^2}{c^2}} \quad (7.37)$$

Par identification avec (7.34), on obtient le lagrangien relativiste de la particule libre :

$$\mathcal{L}(\vec{r}, \vec{v}) = Kc \sqrt{1 - \frac{\vec{v}^2}{c^2}}. \quad (7.38)$$

Pour le lagrangien de la particule soumise à une force extérieure, dans les cas simples, il faudra comme d'habitude enlever le potentiel $U(\vec{r})$, c'est-à-dire écrire $\mathcal{L} = \mathcal{L}_{\text{libre}} - U(\vec{r})$ avec $\mathcal{L}_{\text{libre}}$ donné par (7.38).²

Déterminons maintenant K . Dans la limite classique $v \ll c$, on a $\sqrt{1 - v^2/c^2} \simeq 1 - v^2/(2c^2)$, et donc $\mathcal{L} \simeq Kc - Kv^2/(2c)$. La quantité Kc est une constante qui ne modifie pas les équations du mouvement et peut être ignorée. Il reste $-Kv^2/(2c)$ qui doit être, dans cette limite classique, égal à l'énergie cinétique de la particule. On en déduit que $K = -mc$ et donc, finalement,

$$\boxed{\mathcal{L}(\vec{r}, \vec{v}) = -mc^2 \sqrt{1 - \frac{\vec{v}^2}{c^2}} - U(\vec{r})} \quad (7.39)$$

C'est le lagrangien relativiste de la particule soumise à une force extérieure "simple". On remarque que ce lagrangien *n'est pas* de la forme $T - U$ avec T l'énergie cinétique de la particule.

À partir du lagrangien, on peut calculer l'impulsion $\vec{p}$ par la formule générale, voir la section 2.3.4 :

$$p_x = \frac{\partial \mathcal{L}}{\partial v_x}, \quad p_y = \frac{\partial \mathcal{L}}{\partial v_y}, \quad p_z = \frac{\partial \mathcal{L}}{\partial v_z}. \quad (7.40)$$

Un petit calcul, à faire en exercice, montre que l'on retrouve bien l'expression de $\vec{p}$ dans (7.1), à savoir $\vec{p} = \gamma_v m \vec{v}$. Pour une particule soumise à un potentiel $U(\vec{r})$, les équations de Lagrange $\frac{d}{dt} p_i = \frac{\partial \mathcal{L}}{\partial x_i}$ redonnent bien le principe fondamentale de la dynamique (7.2).

2. Afin que l'action S reste un invariant relativiste, il faudra que $U(\vec{r})$ se transforme de manière adéquate lors d'un changement de référentiel. On se rendra alors compte que la quantité à enlever, qui ne sera plus un potentiel, devra alors dépendre également de la vitesse... En fait, ce n'est pas étonnant : si le système est un électron se déplaçant autour de charges fixes, alors l'électron est soumis à un potentiel $U(\vec{r}) = -eV(\vec{r})$, avec V le potentiel électrostatique. Dans un autre référentiel, les charges qui étaient fixes sont maintenant en mouvement, il y a un champ magnétique $\vec{B}$ et une force $-e\vec{v} \wedge \vec{B}$ qui dépend de la vitesse.

On a vu dans la section 2.3.5 que la quantité $H = \sum_i \dot{q}_i p_i - \mathcal{L}$ étaient conservée si le lagrangien ne dépendait pas explicitement du temps. Cette conclusion reste vraie en mécanique relativiste. Pour une particule dans un champ, cela donne

$$H = \dot{x} p_x + \dot{y} p_y + \dot{z} p_z - \mathcal{L} = \vec{v} \cdot \vec{p} - \mathcal{L} = \gamma_v m v^2 - \mathcal{L}. \quad (7.41)$$

Mais on peut réécrire (7.39) en $\mathcal{L} = -\gamma_v m c^2 (1 - v^2/c^2) - U = -\gamma_v m c^2 + \gamma_v m v^2 - U$. On obtient donc

$$H = \gamma_v m c^2 + U. \quad (7.42)$$

C'est l'expression de l'énergie d'une particule dans un champ U . En particulier, le premier terme $\gamma_v m c^2$ est bien l'énergie de la particule libre, voir (7.1).

Pour un système de particules, le Lagrangien sera de la forme $\sum_i -m_i c^2 \sqrt{1 - v_i^2/c^2} - U(\vec{r}_1, \dots, \vec{r}_N)$. Les conclusions de la section 2.3.4 restent valides : l'énergie totale est conservée ; si le lagrangien est invariant par translation, alors la quantité de mouvement totale est conservée, etc.

7.3 Équivalence masse-énergie

Cette partie traite de la conséquence de l'existence de l'énergie au repos et de l'équivalence masse-énergie prévue par la relativité restreinte.

7.3.1 Ordres de grandeur

On peut associer la masse m d'un objet à son énergie interne $E_{\text{int}} = m c^2$. En ordre de grandeur, on peut calculer à quelle énergie équivalente correspondent quelques masses usuelles.

Échelle macroscopique, $m = 1 \text{ g}$: Si on fait le calcul, on peut associer à cette masse une énergie au repos de $E_{\text{int}} = 10^{-3} \times (3 \cdot 10^8)^2 = 9 \cdot 10^{13} \text{ J} \approx 2,5 \cdot 10^7 \text{ kW h}$. À titre de comparaison, c'est ce que produit une centrale nucléaire en 24 heures, et c'est la consommation électrique moyenne de 10 000 personnes pendant un an.

La masse au repos peut donc être considérée comme une très importante réserve d'énergie équivalente, pour peu que l'on puisse l'extraire par des processus de désintégration.

Échelle microscopique, $m = 1 \text{ u}$: On rappelle que l'unité de masse atomique u est, en gros, la masse d'un proton ou d'un neutron. Précisément, elle est définie comme un douzième de la masse du carbone 12, et vaut $1 \text{ u} \simeq 1,66 \cdot 10^{-27} \text{ kg}$. Cette masse correspond à une énergie au repos de $E_{\text{int}}(1 \text{ u}) \simeq 1,66 \cdot 10^{-27} \times (3 \cdot 10^8)^2 \simeq 1,5 \cdot 10^{-10} \text{ J} \simeq 1 \text{ GeV}$. (On rappelle que l'électron-volt, eV, est défini comme l'énergie cinétique d'un électron accéléré par une différence de potentiel de 1 V, soit $1 \text{ eV} \simeq 1,6 \cdot 10^{-19} \text{ J}$.) Un calcul plus précis donne

$$E_{\text{int}}(1 \text{ u}) = 1 \text{ u } c^2 = 931,5 \text{ MeV}. \quad (7.43)$$

Ainsi, souvent, on trouve les masses des particules exprimées non pas en unité de masse atomiques, mais en énergie équivalente. Pour information, voici les masses de quelques particules subatomiques en u et en eV/c^2 :

Electron	$m_e \simeq 5,487 \cdot 10^{-3} \text{ u}$	$\simeq 511,1 \text{ keV}/c^2$
Proton	$m_{\text{proton}} \simeq 1,0075 \text{ u}$	$\simeq 938,5 \text{ MeV}/c^2$
Neutron	$m_{\text{neutron}} \simeq 1,0087 \text{ u}$	$\simeq 939,6 \text{ MeV}/c^2$
Muon μ^+	$m_{\mu} \simeq 0,113 \text{ u}$	$\simeq 105,6 \text{ MeV}/c^2$
Boson de Higgs H^0	$m_{H^0} \simeq 134 \text{ u}$	$\simeq 125 \text{ GeV}/c^2$

7.4 Collisions et loi de conservations

Lors d'une collision entre particules, **la quantité de mouvement totale** (de l'ensemble des particules, c'est-à-dire d'un système isolé) et **l'énergie totale se conservent**.

Les expressions que nous avons établies pour la quantité de mouvement et l'énergie permettent de concilier cette exigence de conservation avec les postulats de la relativité restreinte.

Une conséquence de la théorie est que **la masse n'est pas nécessairement conservée lors de collisions entre particules** !

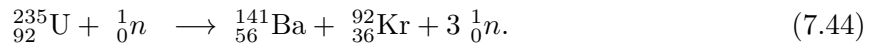
7.4.1 Un exemple de réaction nucléaire : fission de l'uranium-235

À titre d'exemple, nous étudions une réaction nucléaire d'importance industrielle de premier plan puisque c'est une des réactions ayant lieu dans les centrales nucléaires permettant la production d'électricité.

Un neutron entrant en collision avec un noyau d'uranium-235 peut fissionner ce dernier en deux noyaux plus légers tout en émettant des neutrons.

On rappelle qu'un noyau donné se note : A_ZX , où A est le *nombre de masse*, c'est-à-dire le nombre de nucléons (protons et neutrons) composant le noyau. Z est le *nombre de charges*, c'est-à-dire dans le cas d'un noyau, le nombre de protons. Z donne la nature de l'élément dont il est question. Le nombre de neutrons est alors donné par $N = A - Z$. Deux noyaux ayant le même Z , mais des A différents sont deux isotopes du même élément.

Une des réactions de fission de l'uranium 235 est la suivante :



(On note 1_0n ou, plus simplement n pour un neutron.)

Cette réaction conserve le nombre de masse A et le nombre de charge Z .

Elle conserve aussi la quantité de mouvement totale, c'est-à-dire que dans un référentiel $\mathcal{R}$ donné on a :

$$\vec{p}({}^{235}\text{U}) + \vec{p}(n) = \vec{p}({}^{141}\text{Ba}) + \vec{p}({}^{92}\text{Kr}) + \vec{p}(1^{\text{er}} n) + \vec{p}(2^{\text{e}} n) + \vec{p}(3^{\text{e}} n). \quad (7.45)$$

L'énergie totale est également conservée si bien qu'en notant (dans $\mathcal{R}$) E_1 l'énergie totale du noyau d'uranium et du neutron absorbé, et E_2 l'énergie totale des noyaux de Baryum et de Krypton et des trois neutrons émis, on a :

$$E_1 = E_2. \quad (7.46)$$

Si on sépare l'énergie totale en une partie d'énergie cinétique E_c et une partie d'énergie interne mc^2 où m est la masse totale des particules au repos, on arrive à :

$$E_{c_1} + m_1c^2 = E_{c_2} + m_2c^2, \quad (7.47)$$

qui conduit à :

$$E_{c_2} - E_{c_1} = (m_1 - m_2)c^2. \quad (7.48)$$

En pesant les différentes particules au repos, on se rend compte que $m_1 > m_2$ ($m_1 - m_2 \simeq 0,185 \text{ u}$), et que donc :

$$E_{c_2} - E_{c_1} = (m_1 - m_2)c^2 \simeq 0,185 \text{ u}c^2 \simeq 172,4 \text{ MeV}. \quad (7.49)$$

Cette réaction crée donc une variation d'énergie cinétique qu'on retrouve dans les produits de la réaction.

Cette énergie cinétique peut être transférée, par exemple, à des molécules d'eau qu'on aurait mis dans l'environnement proche des réactifs. L'eau recevant cette énergie peut ensuite se vaporiser et faire tourner une turbine. C'est sur ce principe que fonctionnent les centrales nucléaires générant la majeure partie de l'électricité produite en France.

Il faut bien comprendre que *cette réaction n'a pas créé d'énergie*. Elle a transformé de l'énergie de masse en énergie cinétique.

7.4.2 Création de paire

La création de paire est un effet purement relativiste : il s'agit de la génération spontanée de masse, en l'occurrence la paire d'un électron e^- et d'un positron e^+ , à partir d'un photon suffisamment énergétique (qui est annihilé pendant le processus). On va étudier cette création de paire et mettre en évidence à quelles conditions elle peut avoir lieu, en se basant sur les lois de conservation.

Impossibilité de la création de paire dans le vide à partir d'un seul photon

On pourrait imaginer qu'un photon isolé puisse spontanément se transformer en une paire électron impulsion, comme sur la figure 7.4, mais c'est impossible.

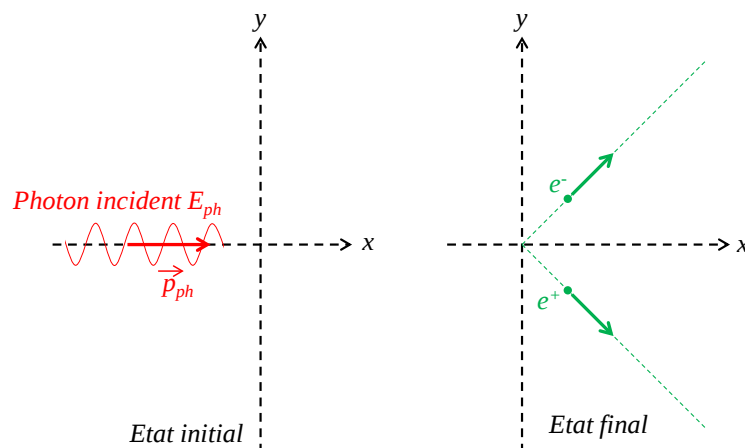


FIGURE 7.4 – Schéma de principe de la création de paire dans le vide

En effet, dans cette situation, on pourrait trouver un référentiel galiléen $\mathcal{R}'$ où les deux particules produites ont une impulsion totale nulle. (C'est le référentiel de centre de masse.) Par conservation de l'impulsion, le photon incident devrait également avoir une impulsion nulle dans ce référentiel, ce qui est absurde.

Création de paire dans la matière

La réaction de création de paire à partir d'un photon peut avoir lieu en présence d'un corps annexe de masse M , voir la figure 7.5.

On suppose que la masse M est initialement au repos avant la collision avec le photon, puis est animé d'une quantité de mouvement $\vec{p}_M$ après le choc. Ainsi, son énergie avant collision est son énergie au repos Mc^2 , et son énergie après collision peut s'exprimer en fonction de p_M par $E_M = \sqrt{p_M^2 c^2 + M^2 c^4}$. Si on applique la conservation de l'énergie totale au système (photon + masse M) :

$$E_{ph} + Mc^2 = \gamma_- mc^2 + \gamma_+ mc^2 + \sqrt{p_M^2 c^2 + M^2 c^4},$$

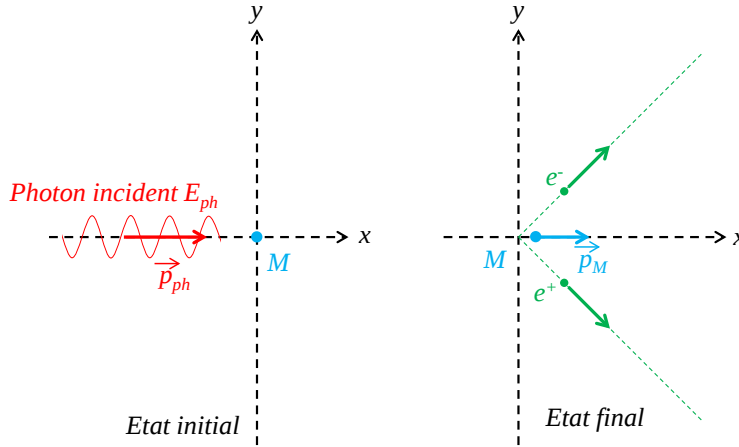


FIGURE 7.5 – Schéma de principe de la création de paire dans la matière

où γ_+ est le facteur γ du positron (animé d'une vitesse $\vec{v}_+$) et γ_- est le facteur γ de l'électron (animé d'une vitesse $\vec{v}_-$).

Au seuil de la réaction ($E_{ph} = E_S$), les électrons/positrons sont par définition formés au repos, et donc leur énergie vaut mc^2 . Ainsi, au seuil, on peut écrire :

$$E_S + Mc^2 = 2mc^2 + \sqrt{p_M^2 c^2 + M^2 c^4}$$

Si on applique maintenant la conservation de la quantité de mouvement au seuil de la réaction, on trouve :

$$\vec{p}_{ph} = \frac{E_S}{c} \vec{u}_x = \vec{p}_M$$

C'est-à-dire que la quantité de mouvement initiale du photon a été transférée à M . On voit ainsi que la présence du corps M permet à la réaction de respecter la conservation de $\vec{p}_{tot}$.

On peut donc injecter l'expression de p_M dans l'équation de la conservation de l'énergie pour déterminer E_S :

$$E_S + Mc^2 = 2mc^2 + \sqrt{E_S^2 + M^2 c^4}$$

On vérifie que cette équation a une solution dès que $M > 2m$; en pratique la masse M est souvent un atome, et donc $M \gg m$. Dans ce cas, on peut montrer que $E_S \approx 2mc^2$. Ainsi, **la création de paire est un processus qui peut avoir lieu dans la matière, à partir du moment où le photon a une énergie supérieure ou égale à $2mc^2$** . En pratique, $m = 511 \text{ keV}/c^2$, et donc on observera la création spontanée de paire électron/positron pour des photons d'énergie $E_{ph} \geq 1,02 \text{ MeV}$.

Remarques : On peut également trouver d'autres processus associés à la création de paire.

- Réaction inverse : Dans ce cas, le processus provient de la recombinaison d'un électron et du positron, pour former un photon. Dans ce cas encore, la conservation de la quantité de mouvement impose la présence d'un troisième corps, et la réaction ne peut pas se faire dans le vide.
- Création de paire par choc photon-photon : Dans ce cas, la paire électron/positron est générée lors d'une collision entre deux photons suffisamment énergétiques. Ce processus peut avoir lieu dans le vide. En cosmologie, on considère qu'il s'agit d'une des premières réactions (dans les premières millisecondes du Big Bang), ayant conduit à l'apparition des premières particules massives dans l'univers.

Appendice

Appendice A

Les lois de la mécanique newtonienne

Dans cet appendice, nous rappelons les lois de Newton pour un point matériel et montrons comment elles se généralisent à un système de points.

A.1 Lois de la mécanique pour un point matériel

Pour un point matériel α , on note m_α sa masse, M_α sa position, et $\vec{F}_\alpha$ la résultante des forces appliquées sur ce point α . On note également $\vec{r}_\alpha = \overrightarrow{OM}_\alpha$ son vecteur position, où O est un point fixe que l'on se donne dans le référentiel d'étude. La vitesse et l'accélération du point matériel sont alors $\dot{\vec{r}}_\alpha$ et $\ddot{\vec{r}}_\alpha$. On rappelle la définition de *l'impulsion* ou *quantité de mouvement* $\vec{p}_\alpha$ d'un point matériel :

$$\vec{p}_\alpha = m_\alpha \dot{\vec{r}}_\alpha. \quad (\text{A.1})$$

On rappelle les lois de Newton pour un point matériel :

N1 : principe d'inertie Il existe des référentiels, dits *inertiels* ou *galiléens*, dans lesquels le mouvement d'un point matériel isolé ou pseudo-isolé est rectiligne et uniforme.

N2 : principe fondamental Dans un référentiel galiléen, on a

$$\dot{\vec{p}}_\alpha = m_\alpha \ddot{\vec{r}}_\alpha = \vec{F}_\alpha. \quad (\text{A.2})$$

N3 : actions réciproques Si un point matériel α exerce une force $\vec{F}_{\alpha\beta}$ sur un point matériel β , alors le point matériel β exerce sur le point α une force $\vec{F}_{\beta\alpha}$ opposée à $\vec{F}_{\alpha\beta}$. De plus, ces forces sont dirigées selon l'axe reliant les deux points matériels :

$$\vec{F}_{\beta\alpha} = -\vec{F}_{\alpha\beta}, \quad \text{ces forces étant colinéaires à } \vec{r}_\beta - \vec{r}_\alpha = \overrightarrow{M_\alpha M_\beta}. \quad (\text{A.3})$$

On rappelle également trois théorèmes importants :

Théorème du moment cinétique

$$\frac{d}{dt} [\overrightarrow{OM}_\alpha \wedge \vec{p}_\alpha] = \frac{d}{dt} [\overrightarrow{OM}_\alpha \wedge (m_\alpha \dot{\vec{r}}_\alpha)] = \overrightarrow{OM}_\alpha \wedge \vec{F}_\alpha, \quad (\text{A.4})$$

où $\overrightarrow{OM}_\alpha \wedge \vec{p}_\alpha = \overrightarrow{OM}_\alpha \wedge (m_\alpha \dot{\vec{r}}_\alpha)$ est le *moment cinétique du point matériel α par rapport au point O* et $\overrightarrow{OM}_\alpha \wedge \vec{F}_\alpha$ est le *moment des forces exercées sur α par rapport au point O* .

Théorème de l'énergie cinétique

$$d\left[\frac{1}{2}m_\alpha(\dot{\vec{r}}_\alpha)^2\right] = \vec{F}_\alpha \cdot d\vec{r}_\alpha, \quad (\text{A.5})$$

où $\frac{1}{2}m_\alpha(\dot{\vec{r}}_\alpha)^2$ est l'*énergie cinétique du point matériel α* et où $\vec{F}_\alpha \cdot d\vec{r}_\alpha$ est le *travail (infinitésimal) de la force $\vec{F}_\alpha$ pendant le déplacement infinitésimal $d\vec{r}_\alpha$* .

Théorème de l'énergie mécanique Supposons que l'on puisse écrire

$$\vec{F}_\alpha = \vec{F}_\alpha^{\text{conservatif}} + \vec{F}_\alpha^{\text{dissipatif}}, \quad \text{avec } \vec{F}_\alpha^{\text{conservatif}} = -\nabla U(\vec{r}_\alpha), \quad (\text{A.6})$$

(on dit que la force conservative $\vec{F}_\alpha^{\text{conservatif}}$ dérive du potentiel U .) Alors

$$d\left[\frac{1}{2}m_\alpha(\dot{\vec{r}}_\alpha)^2 + U(\vec{r}_\alpha)\right] = \vec{F}_\alpha^{\text{dissipatif}} \cdot d\vec{r}_\alpha. \quad (\text{A.7})$$

Le terme dans le crochet est *l'énergie mécanique* du point matériel.

A.2 Ensemble de points matériels

Définitions et notations

On considère un système constitués de N points matériels, indexés par les lettres grecques (α, β , etc.) On garde les mêmes notations que dans la section précédente : m_α est la masse du point matériel α , M_α sa position, $\vec{r}_\alpha$ son vecteur position, $\vec{F}_\alpha$ la résultante des forces s'exerçant sur ce point, $\vec{p}_\alpha$ l'impulsion de ce point.

On va pouvoir écrire les propriétés du système comme des sommes sur les points matériels α . Par exemple, la première relation très simple donne la masse m_{tot} du système :

$$m_{\text{tot}} = \sum_{\alpha=1}^N m_\alpha. \quad (\text{A.8})$$

Ici, N peut aussi bien être petit, égal à 2 ou 3 (pour étudier quelques points matériels en interaction), ou très grand, d'ordre 10^{20} ou plus, pour étudier de la matière macroscopique : les N points matériels du système représentent alors tous les atomes constituant la matière.

Lorsque l'on étudie de la matière macroscopique (N grand), il est souvent plus pratique d'adopter une démarche où l'on traite la matière comme étant continue. Dans cette approche, on note

— L'élément de volume

$$dV = d^3\vec{r}. \quad (\text{A.9})$$

C'est un volume tout petit à l'échelle macroscopique, mais contenant néanmoins beaucoup d'atomes, localisé autour de la position courante $\vec{r}$: cet élément de volume est adapté aux coordonnées choisies : en cartésien, $\vec{r} = (x, y, z)$ et $dV = dx dy dz$ est le volume de la région obtenue en faisant varier l'abscisse entre x et $x + dx$, l'ordonnée entre y et $y + dy$ et la cote entre z et $z + dz$.

— L'élément de masse dm , et la masse volumique locale $\rho(\vec{r}, t)$

$$dm = \rho(\vec{r}, t) dV, \quad (\text{A.10})$$

dm est la masse contenue dans le volume dV et $\rho(\vec{r}, t)$ la masse volumique du système autour du point $\vec{r}$ à l'instant t . Dans cette approche, (A.8) devient

$$m_{\text{tot}} = \iiint dm = \iiint \rho(\vec{r}, t) dV. \quad (\text{A.11})$$

(L'intégrale se fait sur le volume total du système. Nous écrivons en bleu les formules correspondant à cette approche continue.)

A.2.1 Impulsion du système et centre de masse

L'impulsion totale $\vec{p}_{\text{tot}}$ du système est égale à la somme des impulsions des points matériels :

$$\vec{p}_{\text{tot}} = \sum_{\alpha} \vec{p}_{\alpha} = \sum_{\alpha} m_{\alpha} \dot{\vec{r}}_{\alpha} = \sum_{\alpha} m_{\alpha} \frac{d}{dt} \overrightarrow{OM}_{\alpha} = \frac{d}{dt} \left[\sum_{\alpha} m_{\alpha} \overrightarrow{OM}_{\alpha} \right]. \quad (\text{A.12})$$

(Les sommes sur α vont évidemment de 1 à N , mais on ne l'écrit plus pour alléger.)

On introduit le *centre de masse* C du système par la relation

$$m_{\text{tot}} \overrightarrow{OC} = \sum_{\alpha} m_{\alpha} \overrightarrow{OM}_{\alpha} \quad (\text{A.13})$$

On vérifie (en exercice) que la définition C donnée dans (A.13) est indépendante du choix du point O . Autrement dit, la relation (A.13) reste valable si on remplace le point O par n'importe quel point. En particulier, en remplaçant O par C , on obtient

$$\sum_{\alpha} m_{\alpha} \overrightarrow{CM}_{\alpha} = \vec{0}. \quad (\text{A.14})$$

Posons $\vec{r}_C = \overrightarrow{OC}$ le vecteur position du centre de masse. En reportant (A.13) dans (A.12), on obtient

$$\vec{p}_{\text{tot}} = \sum_{\alpha} m_{\alpha} \dot{\vec{r}}_{\alpha} = m_{\text{tot}} \dot{\vec{r}}_C \quad (\text{A.15})$$

Autrement dit,

L'impulsion totale d'un ensemble de points est égale à l'impulsion qu'aurait un unique point matériel de masse m_{tot} situé au centre de masse C .

Réécrivons les relations importantes dans l'approche continue. On a besoin d'introduire $\vec{v}(\vec{r}, t)$ qui est la vitesse moyenne de la matière dans le volume dV autour du point $\vec{r}$ au temps t .

$$\vec{p}_{\text{tot}} = \iiint \vec{v}(\vec{r}, t) dm, \quad m_{\text{tot}} \vec{r}_C = \iiint \vec{r} dm \quad (\text{A.16})$$

Il reste vrai (bien sûr) que $\vec{p}_{\text{tot}} = m_{\text{tot}} \dot{\vec{r}}_C$, mais c'est un peu plus difficile à démontrer sous cette forme...

A.2.2 Principe fondamental de la dynamique

Dérivons deux fois (A.13), et appliquons la seconde loi de Newton à chacun des points matériels

$$m_{\text{tot}} \ddot{\vec{r}}_C = \sum_{\alpha} m_{\alpha} \ddot{\vec{r}}_{\alpha} = \sum_{\alpha} \vec{F}_{\alpha}. \quad (\text{A.17})$$

On décompose la force totale s'exerçant sur le point α de la manière suivante :

$$\vec{F}_{\alpha} = \vec{F}_{\text{ext } \alpha} + \sum_{\beta \neq \alpha} \vec{F}_{\beta \alpha}, \quad (\text{A.18})$$

où $\vec{F}_{\text{ext } \alpha}$ est la force exercée par des objets extérieurs au système sur le point α , et où $\vec{F}_{\beta \alpha}$ est la force exercée par le point β sur le point α . La somme dans (A.18) est donc la contribution totale de toutes les forces intérieures (venant du système lui-même) agissant sur le point α . On a alors

$$\sum_{\alpha} \vec{F}_{\alpha} = \sum_{\alpha} \vec{F}_{\text{ext } \alpha} + \sum_{\alpha} \sum_{\beta \neq \alpha} \vec{F}_{\beta \alpha} = \sum_{\alpha} \vec{F}_{\text{ext } \alpha}. \quad (\text{A.19})$$

En effet, le terme avec une double somme est nul :

$$\begin{aligned}
 \sum_{\alpha} \sum_{\beta \neq \alpha} \vec{F}_{\beta\alpha} = & \vec{F}_{21} + \vec{F}_{31} + \vec{F}_{41} + \dots \\
 & + \vec{F}_{12} + \vec{F}_{32} + \vec{F}_{42} + \dots \\
 & + \vec{F}_{13} + \vec{F}_{23} + \vec{F}_{43} + \dots \\
 & + \vec{F}_{14} + \vec{F}_{24} + \vec{F}_{34} + \dots \\
 & + \dots
 \end{aligned} \tag{A.20}$$

Mais, d'après le principe des actions réciproques, on a $\vec{F}_{21} + \vec{F}_{12} = \vec{0}$, $\vec{F}_{31} + \vec{F}_{13} = \vec{0}$, etc. On voit que tous les termes de la double somme s'annulent deux à deux.

Finalement,

$$m_{\text{tot}} \ddot{\vec{r}}_C = \sum_{\alpha} m_{\alpha} \ddot{\vec{r}}_{\alpha} = \sum_{\alpha} \vec{F}_{\text{ext},\alpha}. \tag{A.21}$$

Le mouvement du centre de masse d'un système est celui qu'aurait un point matériel de masse m_{tot} soumis à la somme des forces extérieures agissant sur le système.

A.2.3 Moment cinétique

Posons $\vec{L}_O$ le moment cinétique total du système par rapport à un point O :

$$\vec{L}_O = \sum_{\alpha} \overrightarrow{OM}_{\alpha} \wedge (m_{\alpha} \dot{\vec{r}}_{\alpha}) = \sum_{\alpha} \vec{r}_{\alpha} \wedge (m_{\alpha} \dot{\vec{r}}_{\alpha}) = \iiint \vec{r} \wedge \vec{v}(\vec{r}, t) dm. \tag{A.22}$$

On dérive par rapport au temps pour obtenir le théorème du moment cinétique d'un ensemble de points matériels :

$$\frac{d}{dt} \vec{L}_O = \sum_{\alpha} [\dot{\vec{r}}_{\alpha} \wedge (m_{\alpha} \dot{\vec{r}}_{\alpha}) + \vec{r}_{\alpha} \wedge (m_{\alpha} \ddot{\vec{r}}_{\alpha})] \tag{A.23}$$

Le premier terme dans la somme est nul (produit vectoriel entre deux vecteurs colinéaires), et dans le deuxième terme, on applique le principe fondamental de la dynamique :

$$\frac{d}{dt} \vec{L}_O = \sum_{\alpha} \overrightarrow{OM}_{\alpha} \wedge \vec{F}_{\alpha} = \sum_{\alpha} \overrightarrow{OM}_{\alpha} \wedge \vec{F}_{\text{ext},\alpha} + \sum_{\alpha} \sum_{\beta \neq \alpha} \overrightarrow{OM}_{\alpha} \wedge \vec{F}_{\beta\alpha}, \tag{A.24}$$

où on a réutilisé la décomposition (A.18). Ici aussi, le terme avec la double somme est nul. Vérifions simplement que le terme $(\alpha = 1, \beta = 2)$ et le terme $(\alpha = 2, \beta = 1)$ ont une somme nulle :

$$\overrightarrow{OM}_1 \wedge \vec{F}_{21} + \overrightarrow{OM}_2 \wedge \vec{F}_{12} = \overrightarrow{OM}_1 \wedge \vec{F}_{21} - \overrightarrow{OM}_2 \wedge \vec{F}_{21} = (\overrightarrow{OM}_1 - \overrightarrow{OM}_2) \wedge \vec{F}_{21} = \overrightarrow{M_2 M_1} \wedge \vec{F}_{21}. \tag{A.25}$$

Mais la force $\vec{F}_{21}$ est alignée avec $\overrightarrow{M_2 M_1}$, et donc le produit vectoriel est nul.

Finalement :

$$\frac{d}{dt} \vec{L}_O = \sum_{\alpha} \overrightarrow{OM}_{\alpha} \wedge \vec{F}_{\text{ext},\alpha}, \tag{A.26}$$

La dérivée du moment cinétique total du système par rapport à un point O fixe est égal à la somme des moments des forces extérieures s'appliquant sur le système.

A.2.4 Point d'application du poids ou d'une autre force extérieure

Considérons la contribution du poids sur le système. Chaque point matériel ressent la force $m_{\alpha} \vec{g}$, et le poids total est

$$\overrightarrow{\text{poids}} = \sum_{\alpha} m_{\alpha} \vec{g} = m_{\text{tot}} \vec{g} \tag{A.27}$$

(On a utilisé le fait que $\vec{g}$, étant constant, pouvait se factoriser.)

Le moment total du poids par rapport au point O est

$$\vec{\mathcal{M}}_O(\overrightarrow{\text{poids}}) = \sum_{\alpha} \overrightarrow{OM}_{\alpha} \wedge (m_{\alpha} \vec{g}) = \left(\sum_{\alpha} m_{\alpha} \overrightarrow{OM}_{\alpha} \right) \wedge \vec{g} = m_{\text{tot}} \overrightarrow{OC} \wedge \vec{g} = \overrightarrow{OC} \wedge (m_{\text{tot}} \vec{g}). \quad (\text{A.28})$$

Dans (A.21) et (A.26), on obtient donc, en notant $\vec{F}'_{\text{ext},\alpha}$ les forces extérieures autres que le poids,

$$m_{\text{tot}} \ddot{\vec{r}}_C = m_{\text{tot}} \vec{g} + \sum_{\alpha} \vec{F}'_{\text{ext},\alpha}, \quad \frac{d}{dt} \vec{L}_O = \overrightarrow{OC} \wedge (m_{\text{tot}} \vec{g}) + \sum_{\alpha} \overrightarrow{OM}_{\alpha} \wedge \vec{F}'_{\text{ext},\alpha}. \quad (\text{A.29})$$

On voit donc que l'on peut remplacer toutes les forces $m_{\alpha} \vec{g}$ agissant sur tous les points matériels du système par une unique force $m_{\text{tot}} \vec{g}$ s'appliquant au centre de masse C .

On peut considérer que le poids s'applique uniquement en C .

Ce type de raisonnement peut être mené (plus difficilement...) avec d'autres forces qui agissent sur plusieurs points matériels en même temps. Pour prendre trois exemples :

- Dans un solide immergé, La poussée d'Archimède est la résultante de toutes les forces de pression qui s'exercent sur la surface du solide. On peut remplacer toutes ces forces par une seule force s'appliquant au centre du volume immergé.
- La force aérodynamiques de portance (ou celle de traînée) agit également sur tous les éléments de la surface d'un système, mais elle peut être remplacée par une seule force s'appliquant en un point appelé le centre de poussée.
- La force de réaction d'un support sur un système agit sur tous les points matériels du système en contact avec le support, mais on peut remplacer toutes ces actions par une seule force s'appliquant en un certain point de la *surface de sustentation*.

A.2.5 Théorème de l'énergie cinétique

Posons T l'énergie cinétique totale du système :

$$T = \sum_{\alpha} \frac{1}{2} m_{\alpha} (\dot{\vec{r}}_{\alpha})^2 = \iiint \frac{1}{2} v(\vec{r}, t)^2 dm. \quad (\text{A.30})$$

Le théorème de l'énergie cinétique s'écrit alors, pour un système de points,

$$dT = \sum_{\alpha} \vec{F}_{\alpha} \cdot d\vec{r}_{\alpha} \quad (\text{A.31})$$

On pourrait essayer de décomposer $\vec{F}_{\alpha}$ en forces extérieures et forces intérieures, comme dans (A.18), mais ça ne sert à rien car *les travaux des forces intérieures ne s'annulent pas*. Par exemple, la travail total des forces $\vec{F}_{12}$ et $\vec{F}_{21}$ est

$$\vec{F}_{12} \cdot d\vec{r}_2 + \vec{F}_{21} \cdot d\vec{r}_1 = \vec{F}_{12} \cdot (d\vec{r}_2 - d\vec{r}_1). \quad (\text{A.32})$$

Cette quantité n'est, en général, pas nulle. On garde donc (A.31), avec les forces intérieures, et le théorème de l'énergie cinétique s'énonce :

Dans un système de points, la variation d'énergie cinétique est égale à la somme des travaux de toutes les forces, y compris les forces intérieures.

A.2.6 Théorème de l'énergie mécanique

Supposons que l'on puisse écrire

$$\vec{F}_\alpha = \vec{F}_\alpha^{\text{conservatif}} + \vec{F}_\alpha^{\text{dissipatif}}, \quad \text{avec } \vec{F}_\alpha^{\text{conservatif}} = -\vec{\nabla}_{\vec{r}_\alpha} U(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N). \quad (\text{A.33})$$

(La notation $\vec{\nabla}_{\vec{r}_\alpha}$ correspond au gradient en supposant que la position $\vec{r}_\alpha$ varie alors que toutes les autres positions sont fixes.) Si on considère U comme une fonction des $3N$ variables $x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_N, y_N, z_N$, la différentielle de U s'écrit, par définition,

$$dU = \sum_{\alpha} \left[\frac{\partial U}{\partial x_{\alpha}} dx_{\alpha} + \frac{\partial U}{\partial y_{\alpha}} dy_{\alpha} + \frac{\partial U}{\partial z_{\alpha}} dz_{\alpha} \right]. \quad (\text{A.34})$$

Mais le terme dans la somme est $\vec{\nabla}_{\vec{r}_\alpha} U \cdot d\vec{r}_\alpha$, et on a donc l'écriture plus compacte

$$dU = \sum_{\alpha} \vec{\nabla}_{\vec{r}_\alpha} U \cdot d\vec{r}_\alpha = - \sum_{\alpha} \vec{F}_\alpha^{\text{conservatif}} \cdot d\vec{r}_\alpha. \quad (\text{A.35})$$

En utilisant cette expression dans le théorème de l'énergie cinétique, on obtient finalement le théorème de l'énergie mécanique :

$$d[T + U] = \sum_{\alpha} \vec{F}_\alpha^{\text{dissipatif}} \cdot d\vec{r}_\alpha. \quad (\text{A.36})$$

Dans un système de points, la variation d'énergie mécanique est égale à la somme des travaux des forces dissipatives, y compris les forces dissipatives intérieures.

Attention, il y a une énergie potentielle U pour tout le système, qui dépend de toutes les positions. Il ne faut pas écrire une énergie potentielle par particule ! Par exemple, si $N = 3$ points matériels sont en interaction gravitationnelle, on a

$$U(\vec{r}_1, \vec{r}_2, \vec{r}_3) = G \frac{m_1 m_2}{\|\vec{r}_2 - \vec{r}_1\|} + G \frac{m_1 m_3}{\|\vec{r}_3 - \vec{r}_1\|} + G \frac{m_2 m_3}{\|\vec{r}_3 - \vec{r}_2\|}. \quad (\text{A.37})$$

A.2.7 Référentiel du centre de masse et théorèmes de König

Le *référentiel barycentrique* ou *référentiel du centre de masse du système* est le référentiel en *translation* par rapport au référentiel d'étude (et donc, sans rotation) dans lequel le centre de masse C est immobile. Si $(O, \vec{e}_x, \vec{e}_y, \vec{e}_z)$ est un repère lié au référentiel d'étude, alors $(C, \vec{e}_x, \vec{e}_y, \vec{e}_z)$ est un repère lié au référentiel barycentrique.

Attention ! À moins que C n'ait un mouvement rectiligne uniforme, ce référentiel n'est pas galiléen.

On a $\overrightarrow{OM}_\alpha = \overrightarrow{OC} + \overrightarrow{CM}_\alpha$, c'est-à-dire $\vec{r}_\alpha = \vec{r}_C + \overrightarrow{CM}_\alpha$. En dérivant par rapport au temps, on obtient

$$\dot{\vec{r}}_\alpha = \dot{\vec{r}}_C + \vec{v}'_\alpha \quad \text{avec} \quad \vec{v}'_\alpha = \overrightarrow{CM}_\alpha. \quad (\text{A.38})$$

C'est la loi de composition des vitesses dans un changement de référentiel. Ici, $\vec{v}'_\alpha$ est la vitesse du point matériel α dans le référentiel barycentrique, c'est-à-dire la vitesse du point matériel α par rapport au point C .

On introduit L'_C , le moment cinétique par rapport à C dans le référentiel barycentrique :

$$\vec{L}'_C = \sum_{\alpha} \overrightarrow{CM}_\alpha \wedge (m_{\alpha} \vec{v}'_{\alpha}), \quad (\text{A.39})$$

et T' , l'énergie cinétique dans le référentiel barycentrique :

$$T' = \sum_{\alpha} \frac{1}{2} m_{\alpha} (\vec{v}'_{\alpha})^2. \quad (\text{A.40})$$

Les théorèmes de König permettent de relier $\vec{L}_O$ à $\vec{L}'_C$ et T à T' .

Premier théorème de König

Partant de (A.22), on a

$$\begin{aligned}\vec{L}_O &= \sum_{\alpha} \overrightarrow{OM}_{\alpha} \wedge (m_{\alpha} \dot{\vec{r}}_{\alpha}) = \sum_{\alpha} (\overrightarrow{OC} + \overrightarrow{CM}_{\alpha}) \wedge (m_{\alpha} \dot{\vec{r}}_{\alpha}) \\ &= \overrightarrow{OC} \wedge \sum_{\alpha} m_{\alpha} \dot{\vec{r}}_{\alpha} + \sum_{\alpha} \overrightarrow{CM}_{\alpha} \wedge (m_{\alpha} \dot{\vec{r}}_{\alpha}) = \overrightarrow{OC} \wedge (m_{\text{tot}} \dot{\vec{r}}_C) + \sum_{\alpha} \overrightarrow{CM}_{\alpha} \wedge (m_{\alpha} \dot{\vec{r}}_{\alpha})\end{aligned}\quad (\text{A.41})$$

On a utilisé (A.15). On se concentre maintenant sur le second terme ; en utilisant (A.38),

$$\sum_{\alpha} \overrightarrow{CM}_{\alpha} \wedge (m_{\alpha} \dot{\vec{r}}_{\alpha}) = \sum_{\alpha} \overrightarrow{CM}_{\alpha} \wedge (m_{\alpha} \vec{v}_C + m_{\alpha} \vec{v}'_{\alpha}) = \sum_{\alpha} m_{\alpha} \overrightarrow{CM}_{\alpha} \wedge \vec{v}_C + \sum_{\alpha} \overrightarrow{CM}_{\alpha} \wedge (m_{\alpha} \vec{v}'_{\alpha}).$$

Le premier terme est nul : il suffit de factoriser par $\vec{v}_C$ et d'utiliser (A.14). Le deuxième terme est $\vec{L}'_C$, voir (A.39). On obtient donc

$$\vec{L}_O = \overrightarrow{OC} \wedge (m_{\text{tot}} \dot{\vec{r}}_C) + \vec{L}'_C \quad (\text{A.42})$$

C'est le premier théorème de König :

Le moment cinétique d'un système de points est égal au moment cinétique d'un point matériel de masse m_{tot} situé en C , plus le moment cinétique du système par rapport à C dans le référentiel barycentrique.

Second théorème de König

Remarquons tout d'abord que l'impulsion dans le référentiel barycentrique est nulle. En utilisant (A.38), on a

$$\sum_{\alpha} m_{\alpha} \vec{v}'_{\alpha} = \sum_{\alpha} m_{\alpha} \dot{\vec{r}}_{\alpha} - \sum_{\alpha} m_{\alpha} \dot{\vec{r}}_C = \vec{0}. \quad (\text{A.43})$$

En effet, les deux termes sont égaux à $m_{\text{tot}} \dot{\vec{r}}_C$, voir (A.15).

On passe maintenant à l'énergie cinétique ; partant de (A.30) et en utilisant (A.38), on a

$$T = \sum_{\alpha} \frac{1}{2} m_{\alpha} (\dot{\vec{r}}_{\alpha})^2 = \sum_{\alpha} \frac{1}{2} m_{\alpha} (\dot{\vec{r}}_C + \vec{v}'_{\alpha})^2 = \sum_{\alpha} \frac{1}{2} m_{\alpha} (\dot{\vec{r}}_C)^2 + \sum_{\alpha} \frac{1}{2} m_{\alpha} (\vec{v}'_{\alpha})^2 + \sum_{\alpha} m_{\alpha} \dot{\vec{r}}_C \cdot \vec{v}'_{\alpha}. \quad (\text{A.44})$$

Dans le premier terme, on met $\frac{1}{2} (\dot{\vec{r}}_C)^2$ en facteur et la somme sur α restante donne m_{tot} . Le deuxième terme est l'énergie cinétique dans le référentiel barycentrique, voir (A.40). On voit que le troisième terme est nul en mettant $\dot{\vec{r}}_C$ en facteur et en utilisant (A.43).

Finalement, on obtient le second théorème de König,

$$T = \frac{1}{2} m_{\text{tot}} (\dot{\vec{r}}_C)^2 + T'. \quad (\text{A.45})$$

L'énergie cinétique d'un système de points est égal à l'énergie cinétique d'un point matériel de masse m_{tot} situé en C , plus l'énergie cinétique du système dans le référentiel barycentrique.