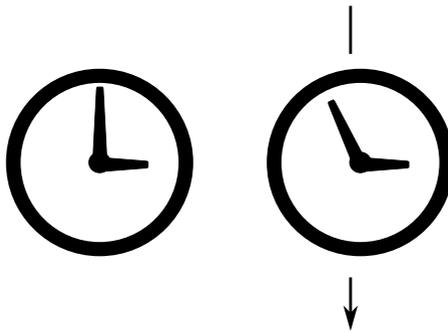


# Théorie de la relativité

Bahram HOUCHMANDZADEH



<http://houchmandzadeh.net/cours/Relativite/reltavite.html>

Première version : 12 Juillet 2011.  
Dernière mise à jour : 18 octobre 2016

# Table des matières

<b>1. Les faits expérimentaux.</b>	<b>5</b>
<b>2. La géométrie de l'espace temps.</b>	<b>9</b>
2.1. Géométrie et le concept du mouvement. . . . .	9
2.2. Espace affine. . . . .	10
2.3. L'espace physique. . . . .	12
2.4. Visite guidé. . . . .	17
2.4.1. vitesse de $R_1$ dans le repère $R_2$ . . . . .	17
2.4.2. Addition des vitesses. . . . .	18
2.4.3. Retard des horloges embarquées. . . . .	19
2.4.4. Simultanéité. . . . .	20
2.4.5. Causalité. . . . .	21
2.4.6. équation d'onde. . . . .	22
2.5. Problèmes. . . . .	24
2.5.1. L'expérience de Rossi et Hall, 1941. . . . .	24
2.5.2. L'expérience de Hafele & Keating, 1972. . . . .	24
2.5.3. L'expérience de Trouton-Noble,1901 . . . . .	25
<b>3. Dynamique relativiste : formulation à l'ancienne.</b>	<b>26</b>
3.1. Vitesse et impulsion à 4 dimensions. . . . .	26
3.2. Chocs élastiques entre particules. . . . .	31
3.3. Le concept de Force. . . . .	31
3.4. L'équation du mouvement dans un repère donné. . . . .	34
3.5. Transformation des forces. . . . .	37
3.6. Application simple : le champ magnétique. . . . .	40
<b>4. Formulation lagrangienne de la physique et minimisations.</b>	<b>41</b>
4.1. Géometrie classique, formulation variationnelle et géodésiques. . . . .	46
<b>5. Les formes linéaires, les vecteurs, les tenseurs et les autres.</b>	<b>50</b>
5.1. Définitions. . . . .	50
5.2. Association bi-univoque entre une forme et un vecteur à travers une expression quadratique. . . . .	53

5.3. Covariant et contravariant. . . . .	55
5.4. Formes quadratiques invariantes et groupes de transformation. . . . .	56
5.5. La dérivation. . . . .	57
5.6. Retour sur la formulation lagrangienne des problèmes de géométries. . . . .	59
5.7. Les tenseurs. . . . .	59
<b>6. Dynamique relativiste : formulation lagrangienne.</b>	<b>63</b>
6.1. Lagrangien relativiste. . . . .	64
6.2. Couplage d'une particule au champ. . . . .	66
6.3. Couplage au champ électromagnétique, formulation invariante. . . . .	66
6.4. La force de Laplace. . . . .	71
<b>7. Électrodynamique.</b>	<b>72</b>
7.1. Formulation variationnelle des champs. . . . .	72
7.2. Formulation invariante du champs électromagnétique. . . . .	75
7.3. Les équations de Maxwell. . . . .	76
7.4. Interaction d'une particule avec son propre champ. . . . .	76
<b>8. Électrodynamique et l'action à distance.</b>	<b>77</b>
<b>9. La formulation de la relativité à l'aide des formes différentielles.</b>	<b>80</b>
9.1. Les 1-formes. . . . .	80
9.2. Intégration des formes différentielles. . . . .	81
9.3. Algèbre des formes différentielles. . . . .	82
9.4. La dérivation extérieure et l'intégration. . . . .	86
9.5. Le déterminant et l'opérateur de Hodge. . . . .	88
9.5.1. Volume et déterminant. . . . .	88
9.5.2. Hodge et norme. . . . .	89
9.6. Le lagrangien et les équations d'euler-lagrange. . . . .	90
<b>10. Mécanique quantique relativiste : l'équation de Dirac.</b>	<b>91</b>
10.1. Equation de Klein-Gordon. . . . .	91
10.2. Equation de Dirac. . . . .	91
10.3. Couplage au champ. . . . .	91
10.4. Electrodynamique quantique. . . . .	91
<b>A. Les dérivées totales et partielles.</b>	<b>92</b>
<b>B. Un peu plus de rigueur : géométrie affine.</b>	<b>94</b>

**C. L'expérience de Michelson-Morley.**

**95**

# 1. Les faits expérimentaux.

L'essence de la théorie de la relativité est extrêmement simple : supposez que vous prenez le train Grenoble-Valence ; au moment où le train quitte la gare, vous regardez votre montre qui indique 12h, exactement comme l'horloge de la gare. Arrivé à Valence, vous vérifiez à nouveau l'horloge de la gare qui indique 13h et votre montre, qui indique  $13h - \epsilon$  : votre montre a pris du retard. Ceci n'est pas un défaut de votre montre, tout processus que vous auriez utilisé pour mesurer le temps aurait indiqué le même retard.

Ce que nous venons de décrire est un fait expérimental ; la seule raison pourquoi nous avons mis si longtemps à le découvrir<sup>1</sup> est que nos instruments de mesure n'étaient pas suffisamment précis jusqu'à la fin du XIX<sup>ème</sup> siècle. L'expérience que nous venons de décrire a été effectuée dans les années 1950, où un avion embarquant une horloge atomique<sup>2</sup> a fait le tour de la terre et on a pu mesurer exactement le retard  $\epsilon$  en comparant l'horloge embarquée et l'horloge restée sur place : le retard est exactement celui prédit par la théorie que nous allons visiter<sup>3</sup>. Une autre façon de mesurer le retard de l'horloge embarqué était de mesurer le temps de désintégration de certaines particules élémentaires (comme les muons) : plus ces particules vont vite, plus le temps de désintégration est long. L'explication est la même que ci-dessus : le temps de désintégration dépend des processus embarqués sur la particules, et ces "horloges" prennent du retard par rapport aux horloges fixes du laboratoire.

En réalité, de nombreuses preuves expérimentales s'étaient accumulées dès la fin du XIX<sup>ème</sup> siècle, bien avant l'expérience des horloges atomiques. Le fait le plus marquant était la mesure de la vitesse des ondes lumineuses. Supposez que vous jetez une pierre dans l'eau et regardez les ondes de surface s'écarter de vous. La vitesse  $v$  de ces ondes dépend seulement des propriétés

---

1. L'énoncé est bien sûr due à Einstein, 1905.

2. Les horloges étaient basées sur des oscillations mécaniques jusque là ; les premières horloges basées sur la mesure des fréquences de résonance des oscillations atomiques sont apparues après la seconde guerre. De nos jours, la précision relative de ces horloges sont de l'ordre de  $10^{-10}$  et les progrès et les recherches pour les améliorer sont très actifs. La seconde est maintenant est définie de cette façon.

3. Le retard comprend également la contribution gravitationnelle dont nous parlerons à la fin de ce cours.

## 1. Les faits expérimentaux.

du liquide et de l'environnement (comme l'accélération de gravité  $g$ , la profondeur du canal  $h$ , la tension de surface, ...). Supposez maintenant que vous glissez à la surface de l'eau et qu'à nouveau, vous provoquez des ondes se surfaces : cette fois, vous verrez les ondes (dans la direction avant) s'écarter de vous avec une vitesse  $v'$  plus faible<sup>4</sup> ; si votre propre vitesse est  $v_p$ , alors  $v' = v - v_p$ . Cela d'ailleurs est une façon efficace de mesurer votre vitesse par rapport à l'eau<sup>5</sup>. A la fin du XIXème, des scientifiques<sup>6</sup> ont mesuré la vitesse des ondes lumineuses lors d'une telle expériences, et ont constaté que l'observateur fixe et l'observateur mobile mesureraient *exactement* la même vitesse. Ceci ne devrait pas être étonnant : d'après ce que nous venons de dire plus haut, les deux observateurs ne mesurent pas le temps de façon identiques. Le fait est encore plus précis : les deux observateurs mesurent tous les deux la même onde sphérique. Nous calculerons cela en détail.

Einstein, suivant les pas de Galilé, a formulé cela précisément : aucune expérience physique ne peut distinguer deux systèmes en mouvement relatif à vitesse constante l'un par rapport à l'autre (systèmes galiléens). C'est ce qu'on l'on appelle *le principe de la relativité*. Supposez que vous êtes dans la cale d'un bateau en mouvement par rapport à la terre à vitesse constante. Toutes les expériences physiques que vous feriez dans cette cale auront exactement les mêmes résultats que celles de votre collègue resté à terre. Vous pouvez par exemple mesurer la période  $T$  de l'oscillation d'une pendule de longueur  $\ell$  et mesurer l'accélération de la gravité  $g = \ell/T^2$  ; votre collègue resté à terre mesurera exactement la même valeur de  $g$  par la même expérience. Vous auriez pu faire d'autres expériences comme mesurer la distribution des impulsions après un choc<sup>7</sup>, l'écoulement de l'eau par un robinet, ... Aucune de ces expériences ne peut vous permettre de distinguer si vous bougez par rapport à la terre ou non, et à fortiori, à quelle vitesse. Einstein à poussé cette idée au bout : *aucune expérience* veut dire vraiment *aucune*, y compris la mesure de la vitesse des ondes lumineuses. Cela l'a emmené à calculer le retard que doivent prendre les horloges embarquées pour que la

---

4. et les ondes vers l'arrière avec une vitesse plus grande. Les ondes s'écartant de vous ne se plus circulaire mais elliptique.

5. Une variante de cette expérience était pratiqué jusqu'à très tard par les marins pour mesurer la vitesse d'un bateau : faites des noeuds régulièrement espacés sur une corde, fixez une extrémité de la corde à votre bateau et jetez là à l'eau. Comptez combien de noeuds reste à la surface, c'est votre vitesse, mesuré en "noeuds".

6. Ceci est appelé l'expérience de Michelson-Morley (vers 1885) , voir l'appendice XXX. La mesure des vitesses était fait par des techniques d'interférométrie ; l'interféromètre développé à l'occasion, appelé Michelson, est une bête noire des étudiants lors des travaux pratiques d'optique.

7. Cela s'appelle plus communément un billard

## 1. Les faits expérimentaux.

vitesse de la lumière  $c$  soit une constante pour tous les observateurs galiléens. Une fois formulée cette énoncée, les scientifiques ont accumulées les preuves expérimentales de l'existence de ce retard.

La théorie de la relativité a été adoptée par la communauté scientifique très rapidement, pratiquement dès son apparition, et sans attendre l'accumulation des preuves expérimentales. D'autres théories, formulées par des scientifiques admirés par Einstein (comme Lorentz) existaient pour expliquer la constance de la vitesse de la lumière. Si la théorie d'Einstein a triomphé si rapidement, c'est qu'elle était d'une très grande beauté. Le concept de beauté est aussi cher aux scientifiques qu'aux artistes<sup>8</sup> et probablement aussi flous pour les uns que pour les autres. La beauté de la théorie d'Einstein venait du fait que L'énoncé d'un principe très simple, la relativité, avait des conséquences extrêmement profond, permettant d'unifier des domaines qui paraissaient disparate. L'exemple le plus frappant est l'existence des champs magnétiques.

Depuis les travaux d'Oersted, d'Ampère et de Faraday (aux alentours de 1820), les scientifiques savaient que deux forces de nature très différentes, l'électricité et le magnétisme, pouvaient influencer l'une sur l'autre : un courant pouvait défléchir l'aiguille d'une boussole *et vice et versa*. La dynamique de ces forces a été formulée par Maxwell (vers 1860) et s'appelle désormais l'électromagnétisme qui calcule précisément comment des variations temporelles d'un de ces champs peuvent être une source pour l'autre champs : nous avons donc deux forces différentes, mais capable de se "transformer" l'une dans l'autre. La théorie d'Einstein permet de voir que cela est une illusion de notre esprit, que ces deux champs sont la même chose mais que l'on mesure différemment dans deux repères en mouvement relatif. Pour voir cela, considérons un courant dans un fil rectiligne infini<sup>9</sup> : un courant dans un fil veut dire des charges positifs fixes (par rapport au laboratoire) et des charges négatives mobiles. Dans le repère fixe des charges positifs, il existe un champs électrique radial ; dans le repère mobile des charges négatives, il existe également un champ électrique radial. Bien sûr, personne ne mesure un champs directement, on ne mesure que des forces ; d'ailleurs, personne ne mesure de forces non plus directement, mais l'effet de ces forces sur des charges par exemple à travers des déplacements durant des temps données (d'où la dimension des forces :  $MLT^{-2}$ ). L'horloge de l'observateur mobile

---

8. et doit probablement impliquer les mêmes processus neuronaux.

9. Pour le physicien, l'infini est un autre nom pour dire très grand (l'infini potentiel) qui nous simplifie les calculs. Les mathématiciens ont inventé le concept du vrai infini (en acte) et se battent depuis pour résoudre tous les paradoxes associés à ce concept ; quelques uns finissant à l'asile.

## 1. Les faits expérimentaux.

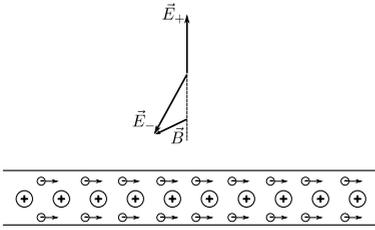


FIGURE 1.1. – Le champs magnétique n’est que le résultant des champs électriques des charges fixes et mobiles mesuré dans le repère fixe.. En réalité, la figure telle que dessiner ici est inexacte, nous verrons la forme exacte un peu plus loin ; l’idée cependant est à retenir.

ne mesure pas le temps de façon identique à celui de l’observateur fixe ; à fortiori donc, les deux observateurs ne vont pas mesurer les forces de la même façon : la force qui pour l’observateur mobile était radiale, quand mesurée par l’observateur fixe, comporte une composante orthoradiale. Ignorant la théorie de la relativité, les scientifiques ont donné à cette composante le nom de “champ magnétique” (Fig. 1.1) Quand nous avanceront un peu dans les mathématiques de la relativité, nous calculerons précisément cette effet.

La beauté avec laquelle la relativité permet d’unifier deux champs a marqué les physiciens ; depuis lors, beaucoup sont persuadé que toutes les forces de la nature ne sont qu’une seule est même chose et qu’il faut juste inventer La belle théorie qui permet de les unifier. Ainsi, les forces à très petite distance qui gouvernent l’intérieur des noyaux atomiques ont été unifiées à travers les théories électrofaible et le Modèle Standard. Le graal des physiciens est d’y incorporer les forces gravitationnelles.

Bon, il est temps maintenant de développer la théorie et faire des calculs prédictifs.

## 2. La géométrie de l'espace temps.

### 2.1. Géométrie et le concept du mouvement.

La géométrie étudie la relation entre des points, des courbes, des surfaces,... Considérons par exemple un espace à deux dimensions où les points sont repérés par  $(x, y)$ . La solution de l'équation différentielle

$$dy/dx = -x/y \tag{2.1}$$

est le cercle  $C : x^2 + y^2 = R^2$ , où  $R$  est une constante d'intégration. L'équation différentielle veut seulement dire que deux points infiniment proche  $P(x, y)$  et  $P'(x + dx, y + dy)$  appartiennent à la courbe si les petites variations  $dx$  et  $dy$  obéissent à la relation (2.1). Le cercle est la donnée de l'ensemble des points  $P$  où l'on peut établir une telle relation entre points proches.

Par contre, si nous avons dit qu'une particule d'abscisse  $x$  obéit à l'équation

$$dx/dt = -t/x$$

notre esprit commence à imaginer une particule *parcourant* l'axe  $x$  avec une loi donnée du temps  $t$  solution de l'équation différentielle. Pour notre esprit, être en mouvement est de disparaître en un un point à un moment pour réapparaître en un point proche *un instant* après, tout cela en conservant son identité. Ce concept de mouvement a préoccupé les philosophes depuis l'antiquité; en ce qui concerne la formulation du monde physique, ce concept n'est absolument pas pertinent et ne sert qu'à nous obscurcir les raisonnements. Toute la physique peut se formuler simplement à travers des formulations purement géométriques où l'on établit des relations entre points proches dans un espace à quatre dimensions (que parfois on appelle espace-temps). Ainsi, une particule pour nous est juste une courbe dans l'espace, muni de quelques scalaires qu'on appelle masse, charge, ...<sup>1</sup>. Une fois les lois énoncées géomé-

---

1. que l'on peut imaginer, si l'on veut, comme des épaisseurs de trait

## 2. La géométrie de l'espace temps.

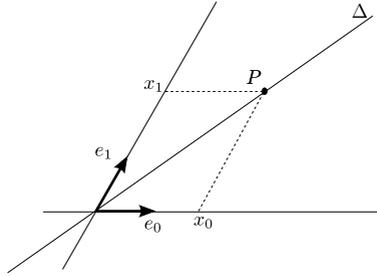


FIGURE 2.1. – Points et droites dans un espace à deux dimensions.

triquement, nous pouvons les réinterpréter dans notre conception humaine du mouvement.

### 2.2. Espace affine.

Donnons nous un espace à deux dimensions, la généralisation à plus sera triviale. Quand nous disons que nous avons un espace à deux dimensions, cela veut dire que nous pouvons associer à chaque point de l'espace, de façon bi-univoque, un couple de nombres réels  $(x_0, x_1)$ . Pour faire cela, nous nous donnons deux vecteurs  $\mathbf{e}_0, \mathbf{e}_1$  et un point  $O$  qu'on appelle origine. Nous disons que pour atteindre le point  $P$  en partant de l'origine, il nous faut un vecteur  $x_0\mathbf{e}_0 + x_1\mathbf{e}_1$  et nous appelons les deux réels  $(x_0, x_1)$  les coordonnées du point  $P$ , comme cela est montré dans la figure 2.1. Comme vous pouvez le remarquer, nous n'avons pas tracé les deux vecteurs comme d'habitude perpendiculaire l'un à l'autre, ceci pour la bonne raison que nous n'avons pas encore défini ce que le mot perpendiculaire ou orthogonal veut dire. Nous ne l'avons pas défini parce que nous ne savons pas trop comment le définir. En effet, prenons un peu d'avance, pour nous  $x_0$  sera la coordonnée temporelle et  $x_1$  la coordonnée spatiale, comment choisir des vecteurs dans ses deux directions *perpendiculaire*? La question n'a pas beaucoup de sens, et surtout, nous n'en avons pas besoin. De façon encore plus général, nous n'avons pas encore définie ce qu'est un angle dans notre espace, donc oubliez ces notions pour le moments<sup>2</sup>.

---

2. Ces concepts appartiennent à l'espace métrique, un espace affine que l'on enrichit par la donnée d'un "produit scalaire". Par contre, le concept de "parallèle" est bien défini en géométrie affine.

## 2. La géométrie de l'espace temps.

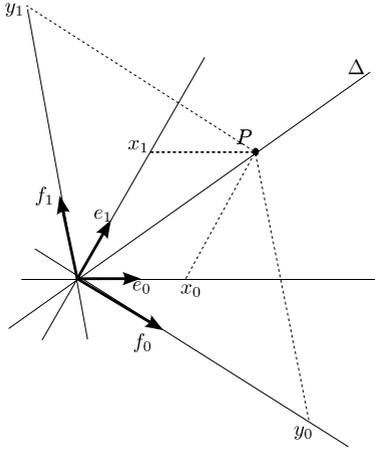


FIGURE 2.2. – Autres choix de vecteurs de la base.

Une droite  $\Delta$  passant par l'origine est donnée par une relation linéaire entre les deux coordonnées :  $a_0x_0 + a_1x_1 = 0$ . La pente de la courbe, si l'on veut privilégier la coordonnée  $x_0$  est définie par

$$p = dx_1/dx_0 = -a_0/a_1$$

. Nous pouvons prendre cela comme la définition de la droite passant par l'origine : deux points proche  $P(x_0, x_1)$  et  $P'(x_0 + dx_0, x_1 + dx_1)$  appartiennent à la droite  $\Delta$  si les petites variations obéissent à la relation énoncée. Remarquons également que nous avons deux droites particulières :  $a_0 = 0$  correspond à une droite portée par le vecteur  $\mathbf{e}_0$  (puisque  $x_1 = 0$ ) ;  $a_1 = 0$  correspond à une droite portée par le vecteur  $\mathbf{e}_1$  (puisque  $x_0 = 0$ ). Par la suite, nous appellerons un repère dans lequel l'équation d'une droite s'écrit  $x_1 = 0$  le *repère propre de cette droite*.

Evidemment, notre choix des deux vecteurs  $\mathbf{e}_0, \mathbf{e}_1$  est arbitraire, nous aurions pu choisir deux autres vecteurs  $\mathbf{f}_0, \mathbf{f}_1$  et repérer les points de l'espace par des coordonnées  $(y_0, y_1)$  (Figure 2.2). Nous pouvons écrire les nouveaux vecteurs comme des combinaisons linéaires des anciens

$$\mathbf{f}_0 = \alpha_{00}\mathbf{e}_0 + \alpha_{10}\mathbf{e}_1$$

$$\mathbf{f}_1 = \alpha_{01}\mathbf{e}_0 + \alpha_{11}\mathbf{e}_1$$

Ce qui nous donne directement la relation entre les anciennes coordonnées

## 2. La géométrie de l'espace temps.

et les nouvelles pour un point quelconque :

$$x_0 = \alpha_{00}y_0 + \alpha_{01}y_1 \quad (2.2)$$

$$x_1 = \alpha_{10}y_0 + \alpha_{11}y_1 \quad (2.3)$$

L'équation de la droite  $\Delta$  dans ces nouvelles coordonnées devient pareillement, par simple remplacement linéaire :

$$\Delta : (a_0\alpha_{00} + a_1\alpha_{10})y_0 + (a_0\alpha_{01} + a_1\alpha_{11})y_1 = 0$$

Le choix des vecteurs de la base sont entièrement arbitraire : n'importe quel couple de vecteurs, du moment qu'ils ne sont pas collinéaires, et donc n'importe quelle choix des coefficients  $\alpha_{ij}$  de transformation (du moment que leur déterminant n'est pas nul) conviennent pour repérer les points de l'espace. Nous pouvons par exemple chercher les nouveaux vecteurs tel que une droite  $\Delta$  donnée soit définie par l'équation

$$y_1 = 0$$

où autrement dit, le nouveau repère soit le repère propre de la droite  $\Delta$ . Nous voyons que dans ce cas, nous aurions du choisir les coefficients  $\alpha_{ij}$  de façon à ce que

$$a_0\alpha_{00} + a_1\alpha_{10} = 0 \quad (2.4)$$

Cela nous donne une relation, tandis que nous avons quatre coefficients à déterminer. Cela veut simplement dire que le choix n'est pas unique, du moment que l'on prend le vecteur  $\mathbf{f}_0$  parallèle à la droite  $\Delta$ , (et peu importe le vecteur  $\mathbf{f}_1$ ), nous avons un repère propre de cette droite.

### 2.3. L'espace physique.

Bien, il nous faut maintenant faire la connexion entre l'espace affine défini plus haut et notre espace physique. Comme nous l'avons dit en passant, l'indice 0 correspond pour nous aux coordonnées temporelles et l'indice 1 (et plus) aux coordonnées spatiales. Soit  $R_1$  le repère porté par les vecteurs  $\mathbf{e}_0, \mathbf{e}_1$ . La droite  $\Delta$  désigne un mouvement à vitesse constante par rapport à ce repère, et la pente que nous avons noté par  $p$  n'est autre que la vitesse  $v$ . Le repère  $R_2$  propre de la droite  $\Delta$  (où l'équation de celle ci est  $y_1 = 0$ ) est le repère "mobile" par rapport à  $R_1$  (si l'on veut utiliser le concept du mouvement). Supposez que  $R_1$  est le repère de la gare où un train passe ; vous êtes assis dans ce train. Si on choisit l'origine des temps quand vous passez

## 2. La géométrie de l'espace temps.

devant la gare, l'équation de votre mouvement dans le repère  $R_1$  (de la gare) est  $x_1 = v.x_0$ . Dans le repère associé à vous, vous êtes bien sûr immobile et votre coordonnée spatiale ne change pas  $y_1 = 0$ .

Dans chaque repère, nous disposons de deux instruments de mesure : une règle pour mesurer les distances et une horloge pour mesurer le temps. Nous supposons qu'un observateur peut arrêter<sup>3</sup> le temps et mesurer tranquillement la distance entre deux points. Tous les observateurs de tous les repères ont embarqué des instruments identiques. D'ailleurs, prenons un peu d'avance, comme nous savons que la vitesse de la lumière est une constante, nous mesurerons dorénavant le temps et la distance dans les mêmes unités. Notre choix d'unités revient à avoir  $c = 1$ .

Soit maintenant un événement quelconque (un flash instantané par exemple). La question est : si nous connaissons les coordonnées de cet événement dans le repère  $R_1$  (l'événement est arrivé au temps  $x_0$  à l'abscisse  $x_1$ ), peut-on en déduire les coordonnées de cet événement dans le repère  $R_2$ ? Autrement dit, quel sont les coefficients  $\alpha_{ij}$  de passage d'un repère à un autre? Le choix de ces coefficients doit nous permettre d'expliquer l'ensemble des faits expérimentaux détaillés au premier chapitre. Nous n'avons plus la liberté de choisir n'importe comment nos repères : les repères *constructibles* par nos instruments de mesures sont soumis aux contraintes des faits expérimentaux.

La réponse à cette question a été donné par Lorentz<sup>4</sup> vers 1895 : Soient deux points quelconques de l'espace  $P$  et  $P'$  infiniment proche ; il existe une quantité scalaire (qu'on appellera pseudo-distance) qui doit être conservée lors des changements de repères. Supposons que dans le repère  $R_1$ , les deux points sont repérés par les coordonnées  $(x_0, x_1)$  et  $(x_0 + dx_0, x_1 + dx_1)$  ; la pseudo distance  $ds$  entre ces deux points est définie par

$$ds^2 = dx_0^2 - dx_1^2 \quad (2.5)$$

Quand on change de repère pour passer aux coordonnées  $y$ , nous devons trouver la même pseudo-distance entre les deux points :

$$ds^2 = dy_0^2 - dy_1^2 \quad (2.6)$$

Avant d'aller plus loin, prenons un peu d'avance et voyons pourquoi Lorentz a fait cette proposition. Supposons que l'observateur du repère  $R_2$  (vous assis

---

3. Il existe mille façon expérimentale d'implémenter une mesure de distance équivalente sans arrêter le temps, nous y viendront plus tard.

4. Beaucoup de physiciens (comme par exemple Larmor) ont contribué à cette transformation, une question très "chaude" à la fin du XIXème. Le nom de "transformation de Lorentz" a été figé par Poincaré en 1905.

## 2. La géométrie de l'espace temps.

dans le train) provoque un flash lumineux quand il passe devant la gare et observe l'onde *sphérique* s'écarter de lui à vitesse  $c$ . L'observateur du repère  $R_1$  doit également observer une onde sphérique<sup>5</sup> s'écartant à la vitesse  $c$  de lui-même (ceci est le fait expérimental établi par Michelson-Morley). L'équation de l'onde s'écrit, dans le repère  $R_1$  quelque chose comme

$$\frac{\partial^2 f}{\partial x_0^2} - \frac{\partial^2 f}{\partial x_1^2} = 0$$

qui est comme vous remarquez, une sorte de carré de dérivé par rapport à  $x_0$  moins une sorte de carré de dérivé par rapport à  $x_1$ . Comme nous allons le voir sous peu, la proposition de Lorentz permet de conserver exactement la même forme pour l'équation d'onde dans le repère  $R_2$  et avoir

$$\frac{\partial^2 f}{\partial y_0^2} - \frac{\partial^2 f}{\partial y_1^2} = 0$$

Ce qui revient à dire que chacun des deux observateurs voit une onde sphérique centrée sur *lui-même* et s'écarter de lui à la vitesse  $c = 1$ .

Bon, revenons à nos oignons maintenant et voyons comment le choix de Lorentz définit de façon unique les coefficients de passage  $\alpha_{ij}$ . Nous avons déjà vu (eq.2.4) que pour que  $R_2$  soit un repère propre, nous devons avoir

$$\alpha_{10} = v\alpha_{00}$$

Remarquez qu'en passant, pour respecter les conventions en usage, nous avons noté  $v$  la vitesse du repère  $R_2$ , qui est bien sûr la même chose que la pente que nous notions  $p = -a_0/a_1$  pour la droite  $\Delta$ . L'utilisation des équation de passage (2.2,2.3) nous donne

$$\begin{aligned} ds^2 &= dx_0^2 - dx_1^2 \\ &= (\alpha_{00}^2 - \alpha_{10}^2)dy_0^2 + 2(\alpha_{00}\alpha_{01} - \alpha_{10}\alpha_{11})dy_0dy_1 + (\alpha_{01}^2 - \alpha_{11}^2)dy_1^2 \\ &= dy_0^2 - dy_1^2 \end{aligned}$$

où nous avons obtenu la dernière ligne en utilisant la conservation de la pseudo-distance. Comme ces relations sont vrai pour n'importe quelles valeurs de  $dy_0$  et  $dy_1$  tous les coefficients doivent être nuls. Un peu d'algèbre nous permet de déduire immédiatement (nous avons 4 équations pour les  $\alpha_{ij}$

---

5. Parler d'une onde sphérique avec une seule dimension spatiale peut paraître un abus de langage

## 2. La géométrie de l'espace temps.

et 4 inconnues)

$$\begin{aligned}\alpha_{00} &= \alpha_{11} = \pm 1/\sqrt{1-v^2} \\ \alpha_{01} &= v\alpha_{11} \quad ; \quad \alpha_{10} = v\alpha_{00}\end{aligned}$$

Nous utiliserons le signe plus pour conserver la même orientation de l'espace (plus sur ce point plus tard); Par ailleurs, pour alléger les notations, nous utiliserons

$$\gamma = 1/\sqrt{1-v^2}$$

qui est une convention de notation largement adoptée. En notation matricielle, nous pouvons donc réécrire les équations de passage ((2.2,2.3) par

$$\begin{pmatrix} x_0 \\ x_1 \end{pmatrix} = \gamma \begin{pmatrix} 1 & v \\ v & 1 \end{pmatrix} \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} \quad (2.7)$$

Notez que  $v$  est la pente de la droite  $\Delta$  dans le repère  $R_1$ . On peut facilement inverser ces relations et écrire les  $y$  en fonction des  $x$  (c'est cela qui est communément appelé transformation de Lorentz) pour obtenir

$$\begin{pmatrix} y_0 \\ y_1 \end{pmatrix} = \gamma \begin{pmatrix} 1 & -v \\ -v & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \quad (2.8)$$

Revenons un moment sur la signification du facteur  $\gamma$ . Nous voyons que nous pouvons l'écrire comme

$$\gamma = \frac{dx_0}{ds} \quad (2.9)$$

c'est à dire de combien varie la coordonnée temporelle quand on passe d'un point  $P$  à un point infiniment voisin  $P'$  séparé par la longueur d'arc  $ds$ . Par ailleurs, nous avons

$$\gamma v = \frac{dx_1}{ds}$$

ce qui donne à la matrice de passage une forme très symétrique, où les coordonnées sont traité sur un pied d'égalité.

**Remarque importante :** Vous remarquez que nous avons un problème si  $v > 1$ , nos expressions deviennent alors complexe. Ceci est l'hypothèse fondamental de la relativité : il existe une vitesse maximum de propagation des interactions, il ne peut pas y avoir interaction instantanée à distance. Dans tout ce qui suit, nous supposons  $v \leq 1$ . Cela implique également que dans tout ce qui suit,  $\gamma \geq 1$ . Remarquez cependant qu'il ne faut jamais ériger

## 2. La géométrie de l'espace temps.

un principe en tabou absolu et refuser de penser au delà ; si cela était le cas, Einstein n'aurait jamais inventé la relativité. Beaucoup ont cherché au delà de ce principe, qui par exemple est un sujet de recherche actif en mécanique quantique. On peut également imaginer un autre univers, inaccessible à nous, où tous ses objets bougent plus vite que la lumière ; cela reste compatible avec la formulation relativiste, tant que les interactions entre les deux mondes restent faible<sup>6</sup>. Les physiciens avaient cherché à détecter les particules de cet autre univers (les tachyons) dans les années 1970, sans succès.

**Généralisation.** Passons à l'espace physique à quatre dimension. Donnons nous un repère  $R_1$  muni des coordonnées  $x_0, x_1, x_2, x_3$  et donnons nous une droite  $\Delta : x_1 = vx_0, x_2 = x_3 = 0$ . Comme vous le remarquez, pour plus de simplicité, nous considérons le cas où la droite  $\Delta$  est dans le plan  $(x_0, x_1)$ . Pour revenir à la description classique, nous pouvons dire que la droite  $\Delta$  décrit un point matériel se déplaçant dans la direction  $x_1$  à vitesse constante  $v$ . Nous allons effectuer un changement de repère et passer aux coordonnées et passer au repère  $R_2$  muni des coordonnées  $y_i$  pour que l'équation de la droite dans ce repère soit  $y_1 = y_2 = y_3 = 0$ , c'est à dire le repère dans lequel le point matériel est au repos. En suivant exactement la même démarche que ci-dessus en utilisant la conservation de l'élément d'arc

$$ds^2 = dx_0^2 - (dx_1^2 + dx_2^2 + dx_3^2)$$

nous obtenons

$$\begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{pmatrix} = \gamma \begin{pmatrix} 1 & -v & 0 & 0 \\ -v & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

**Exercice 1. Généralisation** Considérons une droite  $\Delta$  quelconque  $x_i = v_i x_0$   $i = 1, 2, 3$ . Regroupons les trois coefficients  $x_i$  et  $v_i$  dans un vecteur  $\vec{x}$  (la position habituelle) et  $\vec{v}$  (la vitesse habituelle). Soit  $\vec{x}_{||}$  et  $\vec{x}_{\perp}$  les projections de  $\vec{x}$  parallèle et perpendiculaire au vecteur  $\vec{v} : \vec{x}_{||} = (\vec{x} \cdot \vec{v})\vec{v}/v^2$ ,  $\vec{x}_{\perp} = \vec{x} - \vec{x}_{||}$ . Montrer alors que les transformations de Lorentz s'écrivent :

$$\begin{aligned} y_0 &= \gamma (x_0 - \vec{v} \cdot \vec{x}) \\ y_1 &= \gamma (-vx_0 + (\vec{v} \cdot \vec{x})/v) \\ \vec{y}_{\perp} &= \vec{x}_{\perp} \end{aligned}$$

---

6. Voir l'article de Bilaniuk&Sudarshan, Physics Today **5** :43 (1969) *Particles beyond the light barrier*.

## 2. La géométrie de l'espace temps.

**Exercice.** Même chose qu'avant, mais cette fois avec la pseudo-distance

$$ds^2 = dx_0^2 + dx_1^2 - (dx_2^2 + dx_3^2)$$

**Exercice.** Montrer que nous pouvons écrire la matrice de passage de la transformation de Lorentz à 2d comme

$$\begin{pmatrix} \cosh \phi & -\sinh \phi \\ -\sinh \phi & \cosh \phi \end{pmatrix}$$

Cette forme est parfois utilisé pour souligner la ressemblance avec des rotations dans l'espace métrique habituel.

Bon, avec cette règle de calcul simple en main, visitons quelques conséquences de cette théorie et sa comparaison avec des faits expérimentaux.

## 2.4. Visite guidé.

### 2.4.1. vitesse de $R_1$ dans le repère $R_2$

Nous avons plus haut trouvé les règles de transformation de coordonnées entre un repère  $R_1$  et le repère propre d'une droite  $\Delta$  de pente  $v$  que nous avons appelé  $R_2$ . L'axe  $x_0$  du repère  $R_1$  est une droite  $\Delta'$  :  $x_1 = 0$ . Quelle est l'équation de cette droite ? Prenons un point  $P$  de coordonnée  $(x_0, 0)$ . En utilisant la transformation de Lorentz (2.8) (TL), nous trouvons les coordonnées de ce point dans le repère  $R_2$  :

$$P : (y_0 = x_0, y_1 = -vx_0)$$

L'équation de l'axe  $x_0$  dans le repère  $R_2$  s'écrit donc

$$y_1 = -vy_0 \tag{2.10}$$

On peut reformuler cela dans notre interprétation humaine du mouvement : au lieu de parler de la droite  $\Delta$ , nous pouvons imaginer un point matériel (vous dans le train) en train de parcourir l'espace à une dimension (les rails). Pour l'observateur de la gare, vous vous écartez de lui avec la vitesse  $v$ . Dans le repère  $R_2$  (le repère associé à vous dans le train), au lieu de parler de la droite  $\Delta'$ , vous pouvez imaginer la gare s'écarter de vous le long des rails. L'équation (2.10) veut simplement dire que si vous vous écartez de la gare à vitesse  $v$  dans le repère de la gare, la gare s'écarte de vous à la vitesse  $-v$  dans votre repère. Cela peut paraître trivial à dire, mais c'est bien de s'assurer que notre théorie conserve cette symétrie.

## 2. La géométrie de l'espace temps.

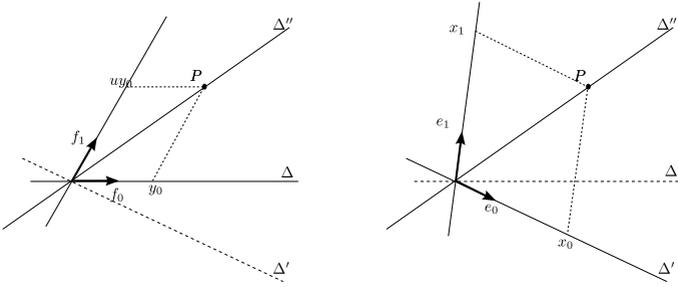


FIGURE 2.3. – Une droite quelconque dont l'équation est recherché dans deux repères différents.

### 2.4.2. Addition des vitesses.

Prenons maintenant notre repère  $R_2$ , et considérons une droite (passant à l'origine) de pente  $u$  :

$$\Delta'' : y_1 = uy_0$$

Quelle est l'équation de cette droite dans le repère  $R_1$  ( voir figure 2.3) ?  
 Considérons un point  $P$  sur cette droite, de coordonnées  $P : (y_0, uy_0)$  ; les coordonnées de cette droite, par TL inverse (2.7), sont données par

$$P : (x_0 = \gamma(1 + uv)y_0, x_1 = \gamma(v + u)y_0)$$

L'équation de la droite  $\Delta''$  est donc donnée par

$$x_1 = \frac{u + v}{1 + uv}x_0$$

et la pente de la droite  $\Delta''$  ( $dx_1/dx_0$ ) dans le repère  $R_1$  est

$$U = \frac{u + v}{1 + uv}$$

A nouveau, nous pouvons réinterpréter cela dans notre cadre du mouvement : par la fenêtre du train, vous projetez<sup>7</sup> un objet vers l'avant avec la vitesse  $u$  : cela veut dire que vous voyez l'objet s'écarter de vous avec la vitesse  $u$ . L'observateur de la gare verra lui l'objet s'écarter de lui avec la vitesse  $U$  donné ci-dessus.

7. Ceci est juste une expérience d'esprit, formellement interdit par la SNCF.

## 2. La géométrie de l'espace temps.

Remarquez deux cas limites intéressants. Si  $u$  et  $v$  sont très faible ( $u, v \ll 1$ ), alors

$$U \approx u + v$$

Si au contraire  $u = 1$ , nous avons

$$U = 1$$

ce qui veut dire que si un objet (un photon par exemple) s'écarte de vous avec la vitesse de la lumière, l'observateur de la gare mesurera également une vitesse de la lumière pour cet objet. Et cela bien sur quelque soit le signe de  $v$  ou de  $u$ . La théorie de M. Einstein explique bien l'expérience de Michelson-Morley.

**Généralisation.** Plaçons nous dans l'espace à quatre dimension. Soit dans le repère  $R_2$  une droite  $\Delta''$  donnée par les équations

$$y_i = u_i y_0 \quad i = 1, 2, 3$$

On peut bien sûr penser aux  $u_i$  comme les composantes de vitesse d'un point matériel dans le repère  $R_2$ . Cherchons l'équation de cette droite dans le repère  $R_1$ . Pour plus de simplicité, nous supposons que la droite porté par le vecteur  $\mathbf{f}_1$  est dans le plan des vecteurs  $\mathbf{e}_0, \mathbf{e}_1$ , ce qui veut dire que le repère  $R_2$  est en déplacement dans la direction  $x_1$  du repère  $R_1$  à vitesse constante  $v$ . La même démarche que ci-dessus nous montre que l'équation de la droite  $\Delta''$  dans le repère  $R_1$  s'écrit  $x_i = U_i x_0$  où

$$\begin{aligned} U_1 &= \frac{u_1 + v}{1 + v u_1} \\ U_2 &= \frac{u_2}{\gamma(1 + v u_1)} \\ U_3 &= \frac{u_3}{\gamma(1 + v u_1)} \end{aligned}$$

**Exercice.** Démontrez la relation ci-dessus.

### 2.4.3. Retard des horloges embarquées.

Reprenons l'expérience mentionné au début de ce manuscrit : vous prenez le train Grenoble-Valence. On choisit notre repère de la gare  $R_1$  pour que l'événement  $P_1$  "le train quitte la gare de grenoble" corresponde aux coordonnées ( $x_0 = 0, x_1 = 0$ ). Soit  $v$  la vitesse du train,  $L$  la distance Grenoble

## 2. La géométrie de l'espace temps.

valence. l'évènement  $P_2$  "le train arrive à la gare de Valence" est repéré par les coordonnées  $(x'_0 = L/v, x'_1 = L)$ . Dans votre repère  $R_2$  à vous assis dans le train, le point (l'évènement)  $P_1$  est repéré par la coordonnées  $(y_0 = 0, y_1 = 0)$  et le point  $P_2$  par les coordonnées  $(y'_0 = L/\gamma v, y'_1 = 0)$ . Le temps donné sur l'horloge de la gare valence est  $x'_0$  et celui donné par votre montre est

$$y'_0 = x'_0/\gamma$$

et comme  $\gamma > 1$ , vous voyez que votre montre a pris du retard.

**L'expérience de Rossi & Hall, 1941.** Le retard des horloges est très faibles aux vitesses habituelles pour les objets macroscopiques qui nous entourent. Pour un TGV roulant à 300km/h,  $v/c \approx 2 \times 10^{-7}$  et donc  $\gamma \approx 1 + 10^{-7}$ , à la limite de nos instruments de mesures. Pour mesurer cet effet, Rossi et Hall<sup>8</sup> ont effectué une expérience très astucieuse. Ils ont construit un appareil pour mesurer le flux des muons (une particule élémentaire); plus précisément, leurs appareils était capable de mesurer non seulement le nombre de muons arrivant sur le détecteur, mais également leur vitesse<sup>9</sup> et de distinguer ainsi les muons lent des muons rapides. La terre est bombardée par des rayons cosmiques, qui comprennent bien sûr des muons. Or, les muons ont un temps de vie finie et se désintègrent au bout de ce temps. De plus, les particules arrivant sur terre ont des vitesses non négligeable par rapport à  $c$ . En mesurant le flux de muons à haute altitude en montagne et à basse altitude, ils ont pu démontrer que le temps de désintégration mesuré en laboratoire dépendait de la vitesse des muons, parfaitement compatible avec la relativité restreinte. L'expérience est très jolie, puisqu'en plus, cela leurs permettait d'estimer le temps propre de désintégration des muons. Nous reprendrons en détail leurs calcul au problème 2.5.1.

### 2.4.4. Simultanéité.

Considérons un objet posé sur les rails, et donc fixe dans le repère  $R_1$ . Pour l'observateur de la gare, la coordonnée spatiale de cet objet reste constante  $x_1 = cte$ , autrement dit l'objet est caractérisé par une droite parallèle à l'axe

---

8. Bruno Rossi and David B. Hall, Phys. Rev. **59**, 223 (1941), *Variation of the Rate of Decay of Mesotrons with Momentum*. L'expérience a été raffinée et précisée quelques années plus tard au CERN : J. Bailey et al., Nucl. Phys. B150, 1 (1979), *Final report on the CERN muon storage ring including the anomalous magnetic moment and the electric dipole moment of the muon, and a direct test of relativistic time dilation*.

9. Leur appareil permettait en fait de distinguer seulement deux groupes de vitesses, qu'ils ont appelé rapide et lente, mais ils ne connaissaient pas la valeur de ces vitesses.

## 2. La géométrie de l'espace temps.

$x_0$ . Pour vous assis dans le train, l'objet est bien sûr pas immobile du tout, et à différents instants, vous mesureriez des positions  $y_1$  différents. Brefs, immobile pour les uns n'est pas immobile pour les autres.

L'espace est le temps étant symétrique, la même chose vaut pour des évènements *simultanés*. Supposons que l'observateur de la gare enregistre deux évènements (deux éclairs par exemples) au même instant à deux endroits différents<sup>10</sup>. Supposons que les coordonnées de ces évènements sont  $P_1 : (x_0, L_1)$  et  $P_2 : (x_0, L_2)$ . Dans le repère mobile  $R_2$  (vous dans le train), les coordonnées de ces points sont données par

$$\begin{aligned} P_1 & : (y_0 = \gamma(x_0 - vL_1), y_1 = \gamma(-vx_0 + L_1)) \\ P_2 & : (y'_0 = \gamma(x_0 - vL_2), y'_1 = \gamma(-vx_0 + L_2)) \end{aligned}$$

clairement,  $y_0 \neq y'_0$  : simultanée pour les uns n'est pas simultanée pour les autres. La différence de temps pour ces deux évènements dans le repère  $R_2$  est  $\Delta y_0 = \gamma v \Delta L$ . Ceci est probablement un des aspects de la théorie de la relativité qui a le plus troublé les gens au début du XXème siècle, qui ne comprenant pas cela, se sont inventé des "paradoxes".

### 2.4.5. Causalité.

Considérons maintenant deux évènements  $E_1(x_0, x_1)$  et  $E_2(x'_0, x'_1)$  se succédant dans le repère  $R_1$ , c'est à dire que  $x_0 < x_1$ . Est ce que dans le repère  $E_2$  ces deux évènements conservent l'ordre temporelle? La transformation de Lorentz nous donne immédiatement le résultat :

$$\begin{aligned} y_0 & = \gamma(x_0 - vx_1) \\ y'_0 & = \gamma(x'_0 - vx'_1) \end{aligned}$$

où encore

$$\Delta y_0 = \gamma(\Delta x_0 - v\Delta x_1)$$

Nous avons supposé que  $\Delta x_0 > 0$ , mais nous voyons que cela ne garantit pas que  $\Delta y_0 > 0$ , tout dépend de l'importance de la distance spatiale  $\Delta x_1$ . Considérons l'évènement "Eve tire sur Adam" et l'évènement "la balle pénètre Adam et le tue"<sup>11</sup>. Est-il possible que dans un autre repère, l'ordre des deux

---

10. Ceci n'a rien à voir avec le temps que les évènement mettent pour être repérés par l'observateur de la gare. On suppose que notre observateur a des collègues distribués par tout, y compris sur les lieux où les évènement arrivent ; ils enregistrent les coordonnées de l'évènement qu'ils observent sur place et communiquent le résultat à leur collègue de la gare. Il existe beaucoup de façon de réaliser expérimentalement ces "collègues" distribués.

11. exemple tiré des "aventures de monsieur Pomkins" par G. Gamov

## 2. La géométrie de l'espace temps.

évènement soit inversé? Cela remettrait en cause notre conception de la causalité. Nous voyons que pour respecter l'ordre des choses, il faut avoir

$$\Delta x_1 < \Delta x_0/v$$

Nous voulons avoir cela quelque soit le repère  $R_2$ ; comme  $v$  est au plus égale à 1 nous devons avoir

$$\Delta x_1 < \Delta x_0$$

C'est ce que l'on appelle le cône d'influence d'un évènement que nous avons rencontré plus haut : si nous acceptons que la causalité doit être absolument respecté, alors un évènement ne peut avoir de l'influence sur un autre que si la distance spatiale les séparant est plus faible que la distance temporelle. Nous retombons sur le principe qu'aucune interaction ne peut se propager plus vite qu'une vitesse maximum qui est celle de la lumière. Dans le cas mentionné plus haut, cela est certainement vrai tant qu'Eve ne tire pas avec des balles surluminales.

Remarquons à nouveau en passant que la causalité ne devrait pas être un tabou de notre pensée; elle est probablement un phénomène macroscopique. Dès les années 1940, les physiciens ont commencé à formuler des ondes ou des particules "remontant" le temps (l' équivalent de "la balle tiré par Eve provoqué par le meurtre d'Adam"). De nos jours, les particules qui remontent le temps sont appelées anti-particules et sont produites en masse dans des accélérateurs). Bien sûr, dans le cadre de ce manuscrit, nous nous tenons dans le cadre de la causalité.

### 2.4.6. équation d'onde.

Considérons une quantité  $V(P)$  dans l'espace à deux dimension pour rester simple. Nous supposons que cette quantité est distribué de façon très particulière : donnons nous un repère  $R_1$  muni des coordonnées  $x_i$ ; nous supposons que  $V$  est distribué de telle façon que

$$\frac{\partial^2 V}{\partial x_0^2} - \frac{\partial^2 V}{\partial x_1^2} = 0 \tag{2.11}$$

C'est à dire qu'en chaque point, les courbures du champ  $V$  dans les deux directions se compensent exactement. Donnons nous maintenant un repère  $R_2$  muni des coordonnées  $y_i$  reliées aux premiers par des transformation de Lorentz. Comment s'écrivent les équations du champ dans ce nouveau

## 2. La géométrie de l'espace temps.

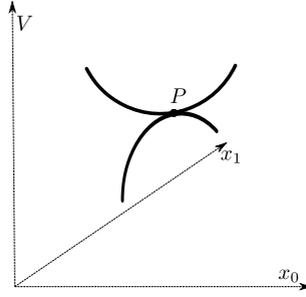


FIGURE 2.4. – Le champ  $V$  est de telle sorte qu'en chaque point de l'espace, sa courbure dans les deux directions se compensent.

repère ? Il nous suffit d'appliquer les règles de dérivations en chaîne :

$$\begin{aligned}\frac{\partial}{\partial x_0} &= \frac{\partial y_0}{\partial x_0} \frac{\partial}{\partial y_0} + \frac{\partial y_1}{\partial x_0} \frac{\partial}{\partial y_1} \\ &= \gamma \frac{\partial}{\partial y_0} - \gamma v \frac{\partial}{\partial y_1}\end{aligned}$$

et

$$\begin{aligned}\frac{\partial}{\partial x_1} &= \frac{\partial y_0}{\partial x_1} \frac{\partial}{\partial y_0} + \frac{\partial y_1}{\partial x_1} \frac{\partial}{\partial y_1} \\ &= -\gamma v \frac{\partial}{\partial y_0} + \gamma \frac{\partial}{\partial y_1}\end{aligned}$$

Ce qui nous donne, en l'appliquant deux fois :

$$\begin{aligned}\frac{\partial^2}{\partial x_0^2} &= \gamma^2 \frac{\partial^2}{\partial y_0^2} - 2\gamma v^2 \frac{\partial^2}{\partial y_0 \partial y_1} + \gamma^2 v^2 \frac{\partial^2}{\partial y_1^2} \\ \frac{\partial^2}{\partial x_1^2} &= \gamma^2 v^2 \frac{\partial^2}{\partial y_0^2} - 2\gamma v^2 \frac{\partial^2}{\partial y_0 \partial y_1} + \gamma^2 \frac{\partial^2}{\partial y_1^2}\end{aligned}$$

En soustrayant les deux lignes, puisque  $\gamma^2(1 - v^2) = 1$ , nous trouvons

$$\frac{\partial^2}{\partial x_0^2} - \frac{\partial^2}{\partial x_1^2} = \frac{\partial^2}{\partial y_0^2} - \frac{\partial^2}{\partial y_1^2}$$

ce qui veut dire que l'équation définissant le champ  $V$  s'écrit exactement de la même façon dans le repère  $R_1$  que dans le repère  $R_2$ . Souvenons nous

## 2. La géométrie de l'espace temps.

que le repère  $R_2$  est le repère en mouvement relatif à vitesse constante par rapport au repère  $R_1$ . L'équation (2.11) est l'équation de propagation de la lumière dans le vide (à vitesse  $c = 1$ ), si par  $V$  nous symbolisons n'importe quelle composante du champ électrique ou magnétique. Nous avons mentionné l'expérience de Michelson-Morley : à l'instant  $t_1 = 0$ ,  $x_1 = 0$ , on crée un flash lumineux, et l'observateur de la gare voit une onde lumineuse sphérique se propager à partir de la gare. L'observateur dans le train qui au moment  $t_2 = 0$ ,  $x_2 = 0$  se trouvait au niveau de la gare observe le même flash, et pour lui également, l'onde se propage à  $c = 1$ , elle est *sphérique et avec lui* toujours au centre.

C'est en cherchant les transformations qui laissent l'équation d'onde invariante que Lorentz ( et Larmor, et quelques autres) a abouti à ses fameuses transformations. Ces scientifiques ne saisissent pas encore la signification de ces transformations de coordonnées et il fallait attendre l'article de 1905 d'Einstein et de 1906 de Minkowski pour que l'espace et le temps s'intègrent dans la géométrie à quatre dimensions.

### **contraction des distances.**

**le cas  $v = 1$ .**

### **Effet Doppler.**

## **2.5. Problèmes.**

### **2.5.1. L'expérience de Rossi et Hall, 1941.**

### **2.5.2. L'expérience de Hafele & Keating, 1972.**

En 1972, des horloges atomiques ont été embarquées à bord des avions parcourant la terre vers l'est et ouest et revenant à leur point de départ<sup>12</sup>. Ces calculs sont cruciaux pour nos communications, puisqu'ils doivent être pris en compte dans le signal GPS.

---

12. J. C. Hafele & Richard E. Keating, Science **177**,166 - 168, *Around-the-World Atomic Clocks : Predicted Relativistic Time Gains*.

### **2.5.3. L'expérience de Trouton-Noble,1901**

The Mechanical Forces Acting on a Charged Electric Condenser Moving through Spac...Phil. Trans. R. Soc. Lond. A January 1, 1904 202 :165-181

# 3. Dynamique relativiste : formulation à l'ancienne.

## 3.1. Vitesse et impulsion à 4 dimensions.

La dynamique classique commence par la formule de Newton

$$\frac{d\mathbf{p}}{dt} = \mathbf{F}$$

De là, on construit le concept de l'énergie cinétique, potentielle, ... Bref, toute la mécanique que l'on connaît. C'est le même effort que l'on doit entreprendre ici, avec la contrainte de retrouver les formules de la mécanique classique à faible vitesse, et bien sûr une prédiction correcte des expériences à haute vitesse où la mécanique classique donne des résultats erronés<sup>1</sup>. La façon naturelle de faire cela simplement est à travers la formulation lagrangienne que nous verrons au prochain chapitre. Ici, nous suivrons essentiellement la démarche originale d'Einstein et de Minkowski (1905-1907).

Commençons par nous demander ce qu'est l'impulsion. On ne peut certes pas répondre  $\mathbf{p} = m\mathbf{v}$ , ceci est une relation que les gens ont trouvé entre impulsion et vitesse, dans le cadre des expériences. L'expérience mécanique la plus simple que l'on peut envisager est le choc élastique entre deux objets ponctuels : les physiciens se sont rendu compte que dans ces expériences, certaines quantités se conservaient quelque soit la vitesse des particules. Trois de ces quantités ont été regroupées dans un vecteur appelé *impulsion*, une autre a eu un nom propre toute seule, *énergie*. Quatre quantité conservées, quatre dimensions, cela devrait nous mettre la puce à l'oreille. Si l'on ne séparait pas artificiellement le temps et l'espace, peut être que les quatre quantités peuvent se grouper dans un seul vecteur à quatre dimension ?

Nous allons effectuer par la pensée des chocs élastiques entre particules, et chercher des quantités qui se conservent quel que soit le repère on l'observe

---

1. Les accélérateurs de particules comme l'ESRF à Grenoble, le CERN à Genève ou le SOLEIL à Paris sont conçu évidemment en utilisant la dynamique relativiste. De petits synchrotrons sont en conception pour utilisation en centre hospitalier pour la génération des rayons X.

### 3. Dynamique relativiste : formulation à l'ancienne.

les chocs. Nous avons maintenant bien établi les transformation de Lorentz, et il est évident que nos définitions classique de conservation d'impulsion et d'énergie deviennent repère-dépendant dans ces transformations. Commençons par formuler quelques concepts dans notre espace à 4 dimensions.

Nous avons vu que ce qui définit la transformation de Lorentz est la conservation de longueur d'arc  $ds$  définie par

$$ds^2 = \sum_{i=0}^3 \epsilon_i dx_i^2$$

où  $\epsilon_0 = 1$  et  $\epsilon_{i>0} = -1$ . Le fait qu'une des dimensions ait une signature différente des trois autres dans la définition de l'arc est le fait marquant de l'espace physique où nous ne trouvons et laisse sa trace à tous les vecteurs qui y sont définis. De façon général, la (pseudo-) norme<sup>2</sup> d'un vecteur  $\mathbf{u} = (u_0, u_1, u_2, u_3)$  est définie par

$$|\mathbf{u}|^2 = \sum_{i=0}^3 \epsilon_i u_i^2 \tag{3.1}$$

et nous pouvons même définir un (pseudo-) produit scalaire<sup>3</sup> entre deux vecteurs :

$$(\mathbf{u}, \mathbf{v}) = \sum_{i=0}^3 \epsilon_i u_i v_i \tag{3.2}$$

et définir l'orthogonalité de deux vecteurs par la nullité de leur (pseudo) produit scalaire.

Soit une courbe  $C$  dans l'espace. La courbe  $C$  est un objet géométrique ; par exemple, un cercle de centre  $0$  et de rayon  $R$  est le lieu de tout les points  $P$  tel que  $|\vec{OP}| = R$ . Pour nous simplifier la vie et manipuler la courbe par des nombre, nous nous donnons un repère  $R_1$  et nous donnons la courbe

---

2. Il n'y a aucune raison de voir cela comme une sorte de distance au sens de la géométrie habituelle. C'est seulement une forme quadratique qui se trouve conservée lors des changement de repaire, fait vérifié par des milliards d'expériences. Le nom de pseudo-norme est seulement utilisé pour souligner une certaine analogie avec la géométrie classique Riemannienne, sans vouloir pousser l'analogie plus loin.

3. un (vrai) produit scalaire a une propriété essentielle : si  $(\mathbf{u}, \mathbf{u}) = 0$ , alors  $\mathbf{u} = 0$ . Cette propriété n'est pas respecté par le produit présenté ici, puisque par exemple, le vecteur  $(1, 1, 0, 0)$  est orthogonal à lui même, sans être un vecteur nul. C'est pourquoi on ajoute le terme pseudo avant le terme.

### 3. Dynamique relativiste : formulation à l'ancienne.

comme une relation entre les coordonnées de ses points<sup>4</sup>. L'équation de la courbe dépend bien sûr du repère choisi. Nous pouvons décrire la courbe  $C$  de façon paramétrique en nous donnant quatre fonctions  $x_i = x_i(\zeta)$ .

Le vecteur tangent à la courbe est également un concept géométrique, mais ce vecteur n'est pas uniquement défini, puisque si  $\mathbf{u}$  est un vecteur tangent, alors  $\alpha\mathbf{u}$  où  $\alpha$  est un scalaire quelconque est également un vecteur tangent. Dans un repère donné, la tangente à la courbe  $C$  en un point  $P$  est donnée par le quadrivecteur  $(dx_i/d\zeta)$ ,  $i = 0, \dots, 3$ .

Nous avons le choix du paramètre pour décrire la courbe. Par exemple, il est courant d'utiliser une des coordonnées, souvent  $x_0$ , comme paramètre et donner les trois autres comme fonction de celui ci :  $x_i = x_i(x_0)$ ,  $i = 1, 2, 3$ . C'est que l'on fait couramment en mécanique classique où les positions  $x, y, z$  sont données en fonction du temps  $t$ . Dans ce cas, le vecteur tangent s'écrit

$$\mathbf{v} = \left(1, \frac{dx_1}{dx_0}, \frac{dx_2}{dx_0}, \frac{dx_3}{dx_0}\right) \quad (3.3)$$

et les trois dernières composantes sont ce que l'on appelle d'habitude vitesse.

Une façon plus astucieuse de choisir le paramètre  $\zeta$  est de prendre directement la longueur d'arc  $s$  entre deux points comme paramètre :  $x_i = x_i(s)$ ,  $i = 0, \dots, 3$ . Ce choix nous donne en fait le vecteur tangent unitaire ( qu'on appellera vecteur tangent propre ) à la courbe en un point :

$$\mathbf{v}_p = \left(\frac{dx_0}{ds}, \frac{dx_1}{ds}, \frac{dx_2}{ds}, \frac{dx_3}{ds}\right) \quad (3.4)$$

puisque'il est trivial de voir que  $|\mathbf{v}_p| = 1$ . Ce vecteur unique est défini de façon purement géométrique, l'expression ci-dessus est simplement son expression algébrique dans un repère donné. La signification du vecteur tangent propre est le suivant : Soit un point  $P$  et un point  $P'$  infiniment voisin sur la courbe  $C$ , séparé par la longueur d'arc  $ds$ . Alors le vecteur  $\mathbf{v}_p$  est défini par (Figure 3.1)

$$\mathbf{v}_p = \lim_{P' \rightarrow P} \frac{P\vec{P}'}{ds}$$

ce vecteur (contrairement à celui défini par l'équation 3.3) ne dépend absolument pas du choix des coordonnées : Si nous utilisons des coordonnées  $y_i$

---

4. Comme d'habitude, au lieu de la courbe  $C$ , vous pouvez imaginer un point matériel en train de "parcourir" une courbe à trois dimensions. Vous perdrez tout les sens de la géométrie, mais vous aurez plus de facilité de comparer cela à votre intuition. Le mieux est bien sûr d'avoir les deux images en tête.

### 3. Dynamique relativiste : formulation à l'ancienne.

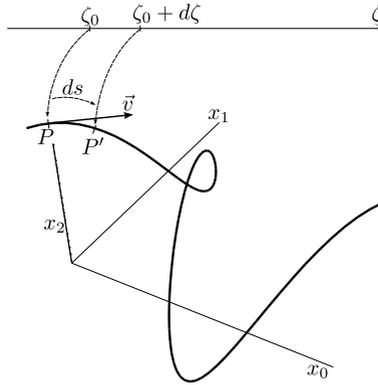


FIGURE 3.1. – Le vecteur tangent en un point  $P$  à une courbe. La courbe peut être paramétrée de façon arbitraire, la définition du vecteur en est indépendante. Comme il est difficile de tracer une courbe à 4 dimensions, nous nous sommes limités ici à 3.

d'un repère  $R_2$  pour repérer la courbe  $C$  et que nous construisons le vecteur tangent  $\mathbf{v}'_p = (dy_i/ds)$  et que par transformation de Lorentz nous recalculons ce vecteur dans le repère  $R_1$ , nous trouvons bien sûr l'expression (3.4).

Il serait instructif d'illustrer ces notions par un exemple tiré de la géométrie habituelle du plan. Une courbe  $C$  centrée sur l'origine peut être décrite par l'équation paramétrique  $x = R \cos \zeta, y = R \sin \zeta$  où  $\zeta \in [0, 2\pi]$ . Nous aurions pu paramétrer la courbe autrement, par exemple par  $x = R \cos \zeta^2, y = R \sin \zeta^2$ , où cette fois  $\zeta \in [0, \sqrt{2\pi}]$ , il s'agit bien sûr de la même courbe. Revenons à notre premier paramétrage, la longueur d'arc infinitésimale est alors donnée par  $ds = R d\zeta$  où encore  $\zeta = s/R$ . L'équation du cercle avec ce choix du paramètre est donnée par  $x = R \cos(s/R), y = R \sin(s/R)$  et le vecteur tangent par  $\mathbf{v} = (-\sin(s/R), \cos(s/R))$  qui est unitaire en géométrie Riemannienne. Une autre façon de donner l'équation du cercle est d'utiliser une des coordonnées (par exemple  $x$ ) comme paramètre et écrire  $y = \sqrt{R^2 - x^2}$ .

Le vecteur tangent propre va jouer un grand rôle par la suite. Nous avons mentionné aussi que les points matériels possèdent des qualités scalaires que l'on appelle masse et charge. La masse<sup>5</sup> est la quantité intrinsèque qui joue

---

5. Pourquoi deux quantités intrinsèques différentes, la masse *et* la charge ? Ne seraient-elles pas les divers reflets de la même chose ? Si on accélère une charge électrique, elle rayonne des ondes électromagnétiques, perd de l'énergie et ralentit. Ce rayonnement EM peut être vu comme une résistance à l'accélération, et donc une masse. On a quelques

### 3. Dynamique relativiste : formulation à l'ancienne.

un rôle déterminant dans les expériences de choc entres particules. Nous définissons le quadri-vecteur impulsion comme

$$\mathbf{p} = m\mathbf{v}_p$$

qui a nouveau est défini à l'aide des quantités intrinsèques indépendemment des coordonnées. Bien sûr, d'après notre construction,

$$|\mathbf{p}| = m$$

Souvent on note le 4-vecteur impulsion dans un repère donné comme<sup>6</sup>

$$\mathbf{p} = (E, \vec{p})$$

où la première composante est un scalaire appelé énergie et les trois autres sont regroupé dans le tri-vecteur  $\vec{p}$  appelé impulsion classique ou tri-impulsion. Nous avons bien sûr la relation trivial

$$E^2 - \vec{p} \cdot \vec{p} = m^2$$

qui se ramène, pour une particule immobile dans un repère à  $E = m$ , plus connu sous la forme

$$E = mc^2$$

quand le temps et l'espace ne sont pas mesuré dans les mêmes unités.

**relation entre vitesse propre et la vitesse habituelle.** Pour manipuler un vecteur avec des nombre, il faut exprimer ses composantes dans un repère donnée. Donnons nous un repère  $R_1$  où les coordonnées sont notées  $x_i$ . L'expression (3.4) donne alors l'expression de la vitesse propre quand le paramètre utilisé et la longueur d'arc. Dans un laboratoire, les expériences sont effectuées en mesurant des temps et des distances, il est donc beaucoup plus pratique d'utiliser le temps  $x_0$  comme paramètre. Si l'ont note les composantes de la vitesse habituelle par

$$v_i = \frac{dx_i}{dx_0} \quad i = 1, 2, 3$$

---

problèmes mathématiques de définir la masse électrique d'une charge ponctuelle (des divergences dans nos équations) ; le problème est souvent traité dans les chapitres avancés des livres d'électrodynamique.

6. Il existe toujours à ce niveau une grande confusion de notation, certain notant le vecteur par exemple  $(-E, \vec{p})$ . Cela dépend si l'on fait référence au composants covariant ou contra-variant. Ces concepts sont discutés en détail au chapitre dédié aux formes linéaires. En ce qui nous concerne, notre notation ici est cohérent avec le produit scalaire défini plus haut.

### 3. Dynamique relativiste : formulation à l'ancienne.

et en nous souvenant (eq 2.9) que  $\gamma = dx_0/ds$ , nous voyons que la vitesse propre s'écrit

$$\mathbf{v}_p = \gamma(1, v_1, v_2, v_3) = \gamma(1, \vec{v})$$

et le quadri-vecteur impulsion est

$$\mathbf{p} = \gamma(m, p_1, p_2, p_3) = \gamma(m, \vec{p})$$

où les  $p_i$  sont les composante du tri-vecteur impulsion habituel :  $p_i = mv_i$ .

Nous voyons donc que ce que nous définissions en mécanique classique comme les concepts séparés d'énergie cinétique et d'impulsion sont en fait les composante d'un même et unique vecteur quadri-dimensionnel que nous avons artificiellement séparé en singularisant la dimension temporelle.

Bien, nous avons passé notre temps à faire des constructions géométrique, faisons quelques expériences de chocs par la pensée pour vérifier la cohérence interne de ces expressions ; il nous faut ensuite les comparé à des faits expérimentaux avant des les accepter.

**Le concept de masse.** Certains auteurs préfèrent définir l'impulsion comme  $\mathbf{p} = m\mathbf{v}$  où  $\mathbf{v}$  est la vitesse classique (exprimée dans un repère donné). Pour rester cohérent, il font alors dépendre la masse du repère choisit,  $m = \gamma m_0$  où  $m_0$  est appelé masse propre. Les deux conventions sont équivalentes ; je trouve cependant cette formulation peu élégante et plus apte à la confusion. Dans tout ce manuscrit, la masse (comme la charge) est une quantité scalaire intrinsèque associée à une particule.

## 3.2. Chocs élastiques entre particules.

## 3.3. Le concept de Force.

Le concept de force est assez ambigu. Nous ne mesurons jamais une force directement, mais des déplacements dans le temps et dans l'espace. Voyons Comment les physiciens sont arrivés à formuler cette notion. La courbe (comme toujours, nous sommes à 4 dimensions) caractérisant une particule matérielle n'est pas toujours droite, mais représente des fléchissements et des courbures. Par de multiples expériences, les gens se sont rendu compte que ces courbures étudiées pour une particule matérielle étaient corrélées à la présence, à proximité, d'autres courbes d'autres particules matérielles.

Plusieurs éléments ont simplifiés la relation entre la courbure en un point d'une courbe et la présence d'autres courbe à proximité. Premièrement,

### 3. Dynamique relativiste : formulation à l'ancienne.

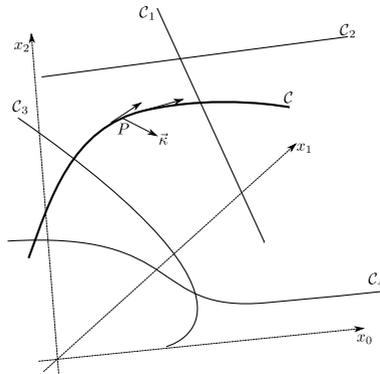


FIGURE 3.2. – Un point matériel est une courbe dans l'espace à quatre dimension muni de propriétés intrinsèques comme la masse et la charge. La courbure  $\vec{\kappa}$  (différence entre deux vecteurs tangents unitaires voisins) en un point  $P$  d'une courbe  $\mathcal{C}$  est due à la présence à proximité d'autres courbes  $\mathcal{C}_i$ . Evidemment, pour que l'approche soit cohérent, il en va de même pour tous les courbes. Nos calculs se simplifient si l'on réduit l'action des autres courbes à un "champ" de force occupant tout l'espace

### 3. Dynamique relativiste : formulation à l'ancienne.

comme nous l'avons dit, une particule matérielle possède une quantité intrinsèque propre, appelé masse. Toutes choses égales par ailleurs, la courbure en un point est inversement proportionnelle à cette quantité<sup>7</sup>. Ensuite, l'effet des autres lignes est linéaire : si la courbure vaut en un point vaut  $\vec{\kappa}_1$  en présence d'une courbe  $C_1$  et  $\vec{\kappa}_2$  en présence d'une courbe  $C_2$ , en présence des deux courbes, la courbure vaut  $\vec{\kappa}_1 + \vec{\kappa}_2$ .

Pour simplifier le calcul, les physiciens l'ont scindé en deux parties. D'abord, on peut calculer un champ de force  $\mathbf{F}(P)$  en chaque point de l'espace, résumant l'action de toutes les courbes présentes. Ensuite, il suffit d'égaliser la courbure locale d'une particule à la force définie en ce point. La force en un point n'est alors que la lecture de la courbure du point matériel en ce point. Cela simplifie beaucoup les choses, puisqu'une fois que nous avons calculé  $\mathbf{F}(P)$  due à la présence de certaines courbes, nous pouvons en déduire l'effet sur n'importe quelle particule sans reprendre le calcul depuis le début.

Cette vision a été formulée la première fois dans les années 1680 par Newton pour traiter la gravité et les lois de la mécanique en générale<sup>8</sup>. Elle a été tellement couronnée de succès qu'elle a servi de modèle pendant plus de deux siècles. Au fur et mesure cependant, le concept de champ a occupé le premier plan, surtout depuis les années 1920 et l'avènement de la mécanique quantique moderne : les physiciens se sont rendu compte qu'aux petites échelles, ils devaient décrire les particules elles mêmes par des champs<sup>9</sup> et non plus par des courbes. La physique de nos jours est formulée à travers la théorie de divers champs et de leurs interactions mutuelles. La physique classique peut cependant se formuler purement à travers "action à distance"<sup>10</sup>.

Revenons à nos moutons. La définition du (quadri-) vecteur de la courbure en un point  $P$  suit de près le modèle de notre définition de la vitesse

---

7. Ceci peut être quantifié parce que nous sommes capable de combiner les masses pour faire une masse plus grande. Cette capacité à combiner les masses nous fournit l'opération d'"addition" pour les masses.

8. Remarquez que ce concept a été très révolutionnaire pour l'époque. Le concept intuitif de la force était "ce que l'on exerce en poussant" : on n'exerçait des forces que par/sur des systèmes mécaniques en contact. Descartes (vers 1650) par exemple, pour expliquer le mouvement des planètes, imaginait l'espace rempli de tourbillon qui pouvait pousser/tirer les corps célestes. Beaucoup plus tard, le concept de "corps en contact" a perdu tout son sens quand les scientifiques se sont rendu compte que les forces en jeu n'étaient que la somme des forces intermoléculaires, également des actions à distances.

9. Un champ lui même peut être vu comme une surface, un volume,... plongé dans un espace à plus grande dimension.

10. Voir un très bel article de Wheeler et Feynman, *Rev. Modern Physics* **21** :425 (1949) : *Classical electrodynamics in terms of direct interparticle action.*

### 3. Dynamique relativiste : formulation à l'ancienne.

propre<sup>11</sup> :

$$\vec{\kappa} = \lim_{P \rightarrow P'} \frac{\mathbf{v}_p(P') - \mathbf{v}_p(P)}{ds} = \frac{d\mathbf{v}_p}{ds}$$

ce qui nous donne, en y intégrant la masse, la relation

$$\frac{d\mathbf{p}}{ds} = \mathbf{F} \quad (3.5)$$

Il est intéressant de noter un fait marquant qui découle directement du vecteur tangent unitaire. Nous savons que le vecteur vitesse propre est unitaire par définition et par conséquent  $|\mathbf{p}| = m$ . Le quadri-vecteur  $\mathbf{F}$  ne change donc que la direction de l'impulsion. Pour que cela soit possible, nous devons avoir obligatoirement l'orthogonalité (au sens de l'éq 3.2) entre la force et la vitesse propre  $v_p$  :

$$\sum_{i=0}^3 \epsilon_i F_i v_i = 0 \quad (3.6)$$

Cela est similaire, en mécanique classique 3d, à la dynamique d'une charge dans un champ magnétique, où la force classique est par définition de la force de Laplace, perpendiculaire à la vitesse classique et *ne travaille pas*.

Remarquez que la relation (3.6) est seulement la géométrisation quadridimensionnel du théorème de l'énergie cinétique en mécanique classique tridimensionnel : la variation de l'énergie cinétique égale le travail des forces. En mécanique classique,  $E_c = (1/2)m(\vec{v} \cdot \vec{v})$  et donc

$$\frac{dE_c}{dt} = m\vec{v} \cdot \frac{d\vec{v}}{dt}$$

Or,  $m d\vec{v}/dt = \vec{F}$  et nous avons donc  $dE_c/dt = \vec{F} \cdot \vec{v}$ . En mécanique classique, on sépare artificiellement l'espace et le temps, mais nous voyons que la relation géométrique (3.6) exprime exactement la même chose.

## 3.4. L'équation du mouvement dans un repère donné.

Nous avons vu l'équation (3.5) comme la définition du concept un peu vague de la force. Nous n'avons pas précisé la nature de la force (électromagnétique, gravitationnelle,...) ni son mode d'action sur la courbe. Nous

---

11. Le lecteur habitué à la géométrie différentielle (habituelle) a sûrement remarqué que c'est là la définition classique de la courbure ; seule la définition de la distance  $ds$  est différente.

### 3. Dynamique relativiste : formulation à l'ancienne.

verrons cela en détail quand nous introduirons la formulation lagrangienne au prochain chapitre. Pour l'instant, nous voyons que nous avons un peu de latitude à définir la force, et l'expression générale (3.5) se prête mal au calcul. Nous pouvons nous donner un repère  $R_1$ , paramétrer la courbe par la coordonnées  $x_0$  et définir la force de telle façon que

$$\frac{d\mathbf{p}}{dx_0} = \mathbf{F} \quad (3.7)$$

et utiliser ensuite les lois de transformations des 4-vecteurs pour exprimer cette relation dans un autre repère galiléen. Dans ces coordonnées,  $\mathbf{p} = m\gamma(1, \vec{v})$ ; il nous faut seulement effectuer proprement les dérivations. Définissons le 3-vecteur accélération de façon habituelle

$$\vec{a} = d\vec{v}/dx_0$$

Comme  $\gamma = (1 - \vec{v}\vec{v})^{-1/2}$ , nous avons

$$\frac{d\gamma}{dx_0} = \gamma^3 \vec{v}\vec{a}$$

Pour la composante spatiale, nous avons

$$\begin{aligned} \frac{d(\gamma\vec{v})}{dx_0} &= \frac{d\gamma}{dx_0} \vec{v} + \gamma\vec{a} \\ &= \gamma^3 (\vec{v}\vec{a})\vec{v} + \gamma\vec{a} \end{aligned}$$

Si nous écrivons également  $\mathbf{F} = (f_0, \vec{f})$ , nous avons l'équation de la dynamique relativiste pour la partie spatiale

$$m\gamma^3 (\vec{v}\vec{a})\vec{v} + m\gamma\vec{a} = \vec{f} \quad (3.8)$$

Pour les faibles vitesse  $v \ll 1$ , le premier terme de la somme est de l'ordre de  $v^2$  par rapport au second et peut-être négligé, ce qui nous donne l'équation de la dynamique classique  $\vec{f} = m\vec{a}$ . L'équation (3.8) paraît a priori plus compliqué que la dynamique classique, mais nous n'avons à prendre ces complications en compte que quand la vitesse devient important, comme par exemple les particules dans les accélérateurs. En général, la mécanique classique est une bonne approximation.

**Exemple 1. Dynamique dans un champ de force constant.** Prenons le cas le plus simple, un champ de force (spatialement) constant avec  $F_1 = f$ ,

### 3. Dynamique relativiste : formulation à l'ancienne.

$F_2 = F_3 = 0$ , et considérons une particule sans vitesse initiale dans le repère du laboratoire :  $v_1 = v_2 = v_3 = 0$ . L'équation du mouvement dans le repère du laboratoire s'écrit :

$$m \frac{d}{dx_0} \left( \frac{v_1}{\sqrt{1 - v_1^2}} \right) = f \quad (3.9)$$

qui s'intègre de façon triviale pour donner

$$\frac{v_1}{\sqrt{1 - v_1^2}} = (f/m)x_0$$

et donc

$$v_1 = \frac{(f/m)^2 x_0^2}{\sqrt{1 + (f/m)^2 x_0^2}}$$

et nous voyons que quand  $x_0 \rightarrow \infty$ ,  $v_1 \rightarrow 1$ . Pour les temps très long, la vitesse s'approche de la vitesse de la lumière sans jamais la dépasser. Nous avons déjà rencontré la barrière de la vitesse de la lumière au chapitre traitant de la cinétique pour ne pas violer la causalité; nous restons donc cohérent avec ce que nous avons énoncé : il n'est pas possible d'accélérer<sup>12</sup> un point matériel au delà de la vitesse de la lumière.

La barrière  $v = 1$  est apparu dans l'exemple particulier de champ de force constant, mais elle est bien sûr générale. Considérons un champs de force quelconque  $F_1 = f(x_i)$ , mais pour plus de simplicité gardons pour l'instant  $F_2 = F_3 = 0$ . Par raison de symétrie (voir la section suivante) nous savons que  $v_2$  et  $v_3$  resterons nulles si elles sont nulles à l'instant initial. L'équation (3.9) se réduit à, une fois que nous utilisons les règles de dérivation en chaîne

$$\frac{dv_1}{dx_0} = (f/m) \frac{1}{\gamma(1 + \gamma^2 v_1^2)}$$

Quand  $v \rightarrow 1$ ,  $\gamma \rightarrow \infty$  et par conséquent, quelque soit la valeur locale de la force,  $dv_1/dx_0 \rightarrow 0$  : plus on se rapproche de la barrière  $v = 1$ , plus l'accroissement de la vitesse chute et s'approche de zéro et il est donc impossible de dépasser  $v = 1$ .

**Exercice.** Généraliser l'argument ci-dessus au cas de champ de force quelconque dans toutes les directions.

---

12. classiquement en tout cas. On peut formuler des théories où l'on prend en compte des "tunneling" au delà de la barrière, mais cela dépasse de beaucoup le cadre de ce manuscrit.

**Un peu plus de détail.** L'équation (3.8) peut être mise en forme de façon légèrement plus parlante. Nous pouvons décomposer l'accélération dans la direction parallèle et perpendiculaire à la vitesse :

$$\begin{aligned}\vec{a}_{\parallel} &= (\vec{a} \cdot \vec{v}) \vec{v} / v^2 \\ \vec{a}_{\perp} &= \vec{a} - \vec{a}_{\parallel}\end{aligned}$$

et l'équation (3.8) s'écrit alors

$$\begin{aligned}\vec{f} &= m\gamma^3 v^2 \vec{a}_{\parallel} + m\gamma \vec{a} \\ &= m\gamma^3 \vec{a} - m\gamma^3 v^2 \vec{a}_{\perp}\end{aligned}$$

La relation ci-dessus met en évidence que contrairement à la mécanique classique, l'accélération n'est pas parallèle à la force. L'équation ci-dessus a été confirmée expérimentalement d'abord par Bucherer en 1909 lors de la mesure de la déflexion des électrons (rayons  $\beta$ ) en champ magnétique et électrique. En décomposant également la force dans les deux directions, et en massant encore un peu plus ces deux équations, on trouve :

$$\begin{aligned}\vec{f}_{\parallel} &= m\gamma^3 \vec{a}_{\parallel} \\ \vec{f}_{\perp} &= m\gamma \vec{a}_{\perp}\end{aligned}$$

Au début du siècle (XX<sup>ème</sup>), il était courant d'interpréter ce résultat comme si la particule avait une masse différente dans les deux directions.

### 3.5. Transformation des forces.

La définition (3.7) nous oblige à être très vigilant sur les transformations des 4-vecteurs entre deux repères en mouvement relatif<sup>13</sup>. D'ailleurs, si vous l'avez remarqué, pour passer de l'équation (3.5) à l'équation (3.7), on se place dans le repère où l'objet est au repos à l'instant initial. Il nous faut ensuite passer de ce repère au repère du laboratoire. En utilisant les lois des additions des vitesses et les transformations de Lorentz, on va donc trouver comment se

---

13. Ces calculs ont un caractère légèrement rébarbatif, surtout parce que nous ne savons pas encore ce que sont les forces et que nous utilisons une expression de la force qui dépend des coordonnées : nous écrivons  $F = dp/dx_0$  au lieu de  $dp/ds$ . Au prochain chapitre, nous verrons une formulation beaucoup plus rigoureuse, où les forces sont obtenues directement sous leurs formes invariants. Il est cependant utile de voir la présente approche, qui est, par sa formulation, proche de la mécanique Newtonienne classique, plus intuitif pour certains lecteurs.

### 3. Dynamique relativiste : formulation à l'ancienne.

transforment les impulsions et les forces entre deux repères. Pour simplifier et éviter les confusions, on considère le repère  $R_2$  en mouvement dans la direction  $x_0$  du repère  $R_1$  avec la vitesse  $\mathcal{V}$ , tandis que la tri-vitesse d'une particule dans les deux repères sera notés par  $\vec{v}$  et  $\vec{v}'$ ;  $\Gamma = (1 - \mathcal{V}^2)^{-1/2}$ . En étant très soigneux, nous trouvons

$$f_1 = f'_1 + \frac{\mathcal{V}v'_2}{1 + \mathcal{V}v'_1}f'_2 + \frac{\mathcal{V}v'_3}{1 + \mathcal{V}v'_1}f'_3 \quad (3.10)$$

$$f_2 = \frac{1}{\Gamma} \frac{1}{1 + \mathcal{V}v'_1} f'_2 \quad (3.11)$$

$$f_3 = \frac{1}{\Gamma} \frac{1}{1 + \mathcal{V}v'_1} f'_3 \quad (3.12)$$

Pour obtenir ces lois de transformations, nous partons de l'expression

$$\frac{d\mathbf{p}'}{dy_0} = \mathbf{F}'$$

qui donne la relation entre la variation de l'impulsion et la force, exprimé dans le repère  $R_2$ . Il nous faut maintenant tout exprimer dans le repère  $R_1$  pour aboutir à une relation du genre

$$\frac{d\mathbf{p}}{dx_0} = \mathbf{F}$$

Et ensuite identifier les divers termes.

Avant cela, notons une relation qui nous sera bien utile. En exprimant le quadri-vecteur impulsion dans un repère donné, nous l'écrivons  $(E, \vec{p})$  et nous savons que  $E^2 - \vec{p}\vec{p} = m^2$ . Si nous dérivons cette expression le long de la trajectoire par rapport aux coordonnées  $x_0$ , nous avons  $2E dE/dx_0 - 2\vec{p}.d\vec{p}/dx_0 = 0$ . Or,  $d\vec{p}/dx_0 = \vec{F}$ ,  $\vec{p} = \gamma m \vec{v}$  et  $E = \gamma m$ , ce qui nous donne

$$\frac{dE}{dx_0} = \vec{F} \cdot \vec{v}$$

Ensuite, notons que  $\mathbf{p}$  étant un quadri-vecteur que nous avons construit géométriquement (sans évoquer un système de coordonnées), nous pouvons utiliser les transformations de Lorentz pour relier ses coordonnées dans différents repères :

$$p'_1 = \Gamma(p_1 - \mathcal{V}E)$$

où nous avons simplement noté  $E$  à la place de  $p_0$ . Finalement, nous avons besoin de la transformation des dérivées entre les deux repères :

$$\frac{d}{dy_0} = \left( \frac{dx_0}{dy_0} \right) \frac{d}{dx_0}$$

### 3. Dynamique relativiste : formulation à l'ancienne.

Comme  $y_0 = \Gamma(x_0 - \mathcal{V}x_1)$ , nous obtenons<sup>14</sup>  $dy_0/dx_0 = \Gamma(1 - \mathcal{V}v_1)$  et par conséquent,

$$\frac{d}{dy_0} = \frac{1}{\Gamma(1 - \mathcal{V}v_1)} \frac{d}{dx_0}$$

Voilà, nous sommes bien équipées pour continuer. Commençons par regarder  $p'_1$  :

$$\begin{aligned} \frac{dp'_1}{dy_0} &= \frac{1}{\Gamma(1 - \mathcal{V}v_1)} \frac{d}{dx_0} (\Gamma(p_1 - \mathcal{V}E)) \\ &= \frac{1}{(1 - \mathcal{V}v_1)} \left( \frac{dp_1}{dx_0} - \mathcal{V} \frac{dE}{dx_0} \right) \\ &= \frac{1}{(1 - \mathcal{V}v_1)} \left( \frac{dp_1}{dx_0} - \mathcal{V} \vec{F} \vec{v} \right) \end{aligned}$$

Il nous suffit maintenant de noter que  $dp_1/dt = F_1$  pour obtenir

$$F'_1 = F_1 - \mathcal{V} \frac{F_2 v_2 + F_3 v_3}{1 - \mathcal{V}v_1}$$

Si nous avons inversé l'approche pour partir des  $p$  vers les  $p'$ , nous aurions obtenus les expressions (3.10-3.12). Nous laissons au lecteur le soin d'obtenir les deux autres relations.

**Remarques.** Il faut noter qu'il est difficile de séparer les forces des vitesses. Les vitesses entrent dans l'expression des forces, ce qui est moins habituel en mécanique Newtonienne classique ; là, très souvent la force dérive d'un potentiel  $V(x)$  et dépend seulement de la position, et non de la vitesse de la particule. Nous voyons maintenant que des considérations d'invariance relativiste nous rendent impossible ce genre de potentiel, et c'est pourquoi par exemple nous rencontrons des expressions du genre  $\vec{F} = q\vec{v} \times \vec{B}$ .

**Exercice.** Comment doivent se transformer les champs Électrique et Magnétique pour que la force de Laplace s'écrive

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B})$$

---

14. N'oublions pas que nous sommes en train d'effectuer les dérivations le long de la courbe d'univers de la particule, et que  $x_1$  dépend de  $x_0$ .

### 3.6. Application simple : le champ magnétique.

Le résultat de la section précédente nous donne directement l'origine du champ magnétique. Donnons nous un exemple très simple. Considérons un fil constitué de charges positives. Supposons, en admettant les calculs d'électrostatique, que le champ de force crée par ce fil est radial :  $\vec{F}_+^{(1)} = f(r)\vec{u}_r$ . Considérons maintenant un fil de charge négatif en mouvement par rapport aux charges positives. Dans le repère au repos par rapport aux charge négative, la force possède la même forme :  $\vec{F}_-^{(2)'} = -f(r)\vec{u}_r$ . Un courant dans un fil est l'union de ces deux fils : les ions positifs immobile par rapport au laboratoire et les électrons en mouvement. Pour calculer la force sur une particule chargée, il nous faut exprimer toutes les forces dans le même repère.

Pour exprimer le champ de force des charges négatives dans le repère  $R_1$  où les charges positives sont au repos, nous devons utiliser les expressions de transformation de force de la section précédente, et nous voyons que les deux forces ne se compensent pas :

$$(\vec{F}_+ + \vec{F}_-).\vec{u}_r = f(r) \left( 1 - \frac{1}{\Gamma} \frac{1}{1 + \mathcal{V}v'_1} \right)$$

et nous avons apparition des forces dans les autres directions également. Cette force supplémentaire, qui de plus dépend de la vitesse de la particule, à pris le nom de champ magnétique.

Quelques pages de calculs nous montrerai que la force de Laplace,  $\vec{F} = q(\vec{E} + \vec{v} \times \vec{B})$  est compatible avec la relativité : si nous calculons la force électrique et magnétique dans chaque repère en suivant les recettes d'électrostatique et de magnétostatique et que nous transformons cela dans un nouveau repère, nous aboutissons à à une force qui est égale à celle que nous aurions calculer directement dans ce repère<sup>15</sup>. Nous ne suivrons pas ce chemin, puisque nous obtiendrons cela automatiquement avec la formulation géométrique du champ au prochain chapitre.

---

15. Voir l'expérience de Trouton et Noble, 1901

## 4. Formulation lagrangienne de la physique et minimisations.

Nous supposons que le lecteur est familier avec le calcul variationnel et la formulation lagrangienne, nous allons nous contenter ici de rappeler les concepts et développer quelques exemples qui nous seront utiles<sup>1</sup>.

Nous savons trouver les extrémum locaux d'une fonction de plusieurs variables  $f(x_1, x_2, \dots, x_n)$ . Au point extremum  $P$  de coordonnées  $(x_i)$ , la fonction doit satisfaire à  $n$  équations

$$\frac{\partial f}{\partial x_i} = 0 ; i = 1, \dots, n$$

Cela veut dire qu'en ce point, la fonction est stationnaire, ou plate. Voyons cela de plus près : en un point quelconque, Si nous passons du point  $P(x_i)$  au point infiniment voisin  $P'(x_i + dx_i)$ , la variation causée dans la valeur de la fonction  $f$  est

$$df = \sum_i \frac{\partial f}{\partial x_i} dx_i + O(dx_i^2)$$

où le symbole  $O$  collecte tous les termes d'ordre supérieurs en  $dx_i$ . Le point  $P$  est un extremum si les variations linéaires en  $dx_i$  sont nulles (quelques soient les  $dx_i$ ); on dit alors que  $f$  en ce point est stationnaire.

Une fonction est une "machine" qui prend un nombre en entrée et produit un nombre en sortie, ce que l'on symbolise en mathématiques par le symbole  $f : \mathbb{R} \rightarrow \mathbb{R}$ . Ceci est une notation très judicieuse, nous disant d'un seul coup d'oeil quel genre d'objet est  $f$ , agissant sur quoi et produisant quoi. Nous pouvons imaginer des objets plus complexe qu'une fonction, par exemple une *fonctionnelle*, qui prend une "fonction" en entrée et produit un nombre en sortie :  $S : \mathbb{F} \rightarrow \mathbb{R}$ , où  $\mathbb{F}$  est l'espace des fonctions.

Les exemples concrets d'utilisation de fonctionnelle ont émergé très vite en physique, prédatant même de quelques années la formulation du calcul

---

1. Pour une vue approfondie, nous référons le lecteur au très beau livre de C. Lanczos, "variational principles of mechanics", Dover.

différentiel par Newton et Leibnitz dans les années 1680. L'exemple qui a marqué son temps est le principe de temps minimum de Fermat : étant donné un milieu d'indice  $n(x, y, z)$ , où en chaque point la lumière a une vitesse  $c/n$ , pour aller d'un point  $A$  à un point  $B$ , la lumière "choisit" la trajectoire qui minimise le temps de parcours. Le deuxième exemple<sup>2</sup>, apparu quelques années après la publication de *Principia* a été formulé par un des frères Bernouilli en 1684 comme un défi à la communauté scientifique : Supposons deux points reliés par une courbe  $y(x)$ , nous lâchons au point  $A$  sans vitesse initiale un objet matériel glissant sans frottement qui sous l'action de la gravité seule, arrive au bout d'un temps  $T$  au point  $B$ . Comment choisir la courbe  $y(x)$  pour que le temps  $T$  soit minimum ?

Nous voyons dans les exemples précédents qu'à chaque trajectoire correspond un nombre (le temps de parcours) et le travail de la nature ou du physicien est de trouver "la trajectoire" qui minimise ce temps. La solution générale à ce genre de problèmes a été donnée dans les années 1750 par Euler<sup>3</sup>, et son application à la mécanique énoncée par Lagrange dans les années 1780. Aujourd'hui, toute la physique fondamentale est formulée à travers cette formulation : les objets physiques (champs, trajectoire, ...) minimisent quelque chose qu'on appelle souvent une "action". Voyons cela de plus près.

Supposons une fonction  $f(t)$ , et étudions le problème où la fonctionnelle  $S[f(t)]$  qui nous intéresse donnée par une intégrale du genre

$$S = \int_a^b \mathcal{L}(f(t), f'(t), t) dt \quad (4.1)$$

Notons que l'intégrand, la fonction de trois variables  $\mathcal{L}$  est une fonction "normale", c'est à dire qu'à trois valeurs  $a, b, c$ , elle associe le nombre  $\mathcal{L}(a, b, c)$ . Cette fonction s'appelle le lagrangien. Si nous passons d'une fonction  $f(t)$  à une fonction infiniment voisin  $f(t) + \epsilon g(t)$ , la quantité  $S$  varie de  $\delta S$  :

$$\delta S = \epsilon S'[f(t), g(t)] + O(\epsilon^2) \quad (4.2)$$

---

2. fortement inspiré par le principe de Fermat

3. La solution d'Euler était une généralisation directe de la minimisation des fonctions de  $n$  variables. Il discrétisait une fonction sur un intervalle en  $N$  points, et considérait la valeur de la fonction en chaque point comme une variable indépendante; la fonctionnelle devenait alors une fonction de ces  $N$  variables et il suffisait ensuite de prendre la limite  $N \rightarrow \infty$ . Cette façon de voir est quelque peu abandonnée dans l'enseignement du calcul variationnel, mais elle est extrêmement fertile et est de retour depuis la popularisation des méthodes numériques de résolution d'équations différentielles comme les différences finies et les éléments finis. Nous référons le lecteur intéressé au livre (quelque peu austère) de Lanczos, "linear differential operators", Dover.

#### 4. Formulation lagrangienne de la physique et minimisations.

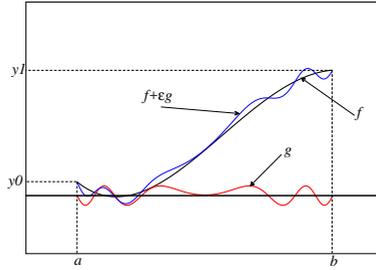


FIGURE 4.1. – Variation d’une fonctionnelle : on étudie la variation  $\delta S$  autour d’une fonction  $f(x)$ .

où  $S'$  est une autre fonctionnelle (ressemblant à une sorte de dérivée dans l’espace des fonctions) et  $O$  comme avant collecte les termes d’ordre supérieur. Comme avant, la fonction  $f$  est un extremum de  $S$ , si les variations linéaires au point  $f$  sont nulles, quelque soit la fonction  $g$ . En faite, pas tout à fait : nous ne considérons que les variations  $g$  telle que

$$g(a) = g(b) = 0$$

Avant de continuer, mettons nous en règle avec les conventions de notation : la fonction est souvent noté  $x(t)$ , ou simplement  $x$  ; sa dérivée  $x'(t)$  est notée  $\dot{x}$ . En conséquence, le lagrangien est notée  $\mathcal{L}(x, \dot{x}, t)$ . Les dérivées partielles du lagrangien par rapport à sa première et deuxième variable  $\partial_1$  et  $\partial_2$  sont notées

$$\frac{\partial \mathcal{L}}{\partial x}, \quad \frac{\partial \mathcal{L}}{\partial \dot{x}}$$

respectivement, ce qui au début constitue une source de confusion pour les étudiants. Si cela est le cas pour vous, remplacez tous les  $\dot{x}$  par un autre symbole, par exemple  $y$ , et une fois toute les dérivations effectuées, revenez aux notations initiales. Un développement limité du lagrangien et une intégration par partie nous donnent alors la variation  $\delta S$  à l’ordre 1 en  $\epsilon$  :

$$\delta S = \epsilon \int_a^b \left( \frac{\partial \mathcal{L}}{\partial x} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}} \right) g(t) dt$$

qui n’est nulle  $\forall g$  que si l’équation, appelé Euler-Lagrange, est vérifié par la fonction  $x$  :

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}} - \frac{\partial \mathcal{L}}{\partial x} = 0 \tag{4.3}$$

**Exemple fondamental : mécanique classique.** En mécanique classique le lagrangien est donnée par

$$\mathcal{L} = T - V$$

où  $T = (1/2)m\dot{x}^2$  est appelé l'énergie cinétique et  $V(x)$  est appelé énergie potentielle. Si nous posons  $F = -V'(x)$ , alors les équation d'Euler Lagrange nous donne l'équation de la dynamique Newtonienne :

$$m\ddot{x} = F$$

Beaucoup de personnes sont intriguées par le signe apparaissant dans le lagrangien :  $L = T - V$  et non  $L = T + V$ . Notons simplement que l'énergie cinétique implique des dérivations *par rapport au temps* tandis le potentiel implique essentiellement des coordonnées spatiales. Or, nous savons que dans l'espace-temps physique, l'élément d'arc vaut  $ds^2 = dt^2 - dx^2$  ; l'expression du lagrangien classique ne respecte bien sûr par l'invariance de l'élément d'arc, mais étant une forme approchée, elle conserve sa signature. Nous y viendrons en détail dans quelques sections.

**Généralisation.** On peut démontrer sans peine que si nous avons  $n$  fonctions  $x_i(t)$  d'une variable  $t$  et un lagrangien de la forme  $\mathcal{L} = \mathcal{L}(x_i, \dot{x}_i, t)$  alors les fonctions  $x_i$  obéissent à  $n$  équations

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}_i} - \frac{\partial \mathcal{L}}{\partial x_i} = 0 ; \quad i = 1, \dots, n \quad (4.4)$$

nous considérerons plus loin, au chapitre sur l'électrodynamique, le cas de  $m$  variable dépendant.

**Note sur "stationnaire" versus "minimum".** Dans le traitement variationnel, nous avons insisté sur le fait que la courbe  $x(t)$  doit rendre l'action  $S$  stationnaire, mais pas forcément minimum ou maximum<sup>4</sup>. Par ailleurs, Il peut arriver que nous trouvions des courbes  $x_2(t)$  loin de la solution  $x(t)$  donnée par l'équation d'Euler-Lagrange qui ont une action plus petite que celle de la courbe  $x(t)$ , mais qui ne rende pas l'action stationnaire. Pour comprendre cela, prenons l'exemple de la fonction  $y = (1/3)x^3 - x$ , qui possède un minimum local au point  $x = 1$  ; bien sûr,  $f(-4) < f(1)$ , mais au point  $x = -4$ , la fonction n'est pas stationnaire. Nous avons exactement la même chose pour les fonctionnelles.

---

4. Une très belle explication de cela a été donné par Feynman à travers le concept d'intégrale de chemin et la mécanique classique comme la limite de la mécanique quantique (Feynman & Hibbs, path integrals and quantum mechanics).

**Impulsion et forces généralisées.** La quantité

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{x}_i}$$

est appelé l'impulsion ou le moment conjuguée de la variable  $x_i$ . De même, la quantité

$$f_i = \frac{\partial \mathcal{L}}{\partial x_i}$$

est appelé force généralisée. De cette façon, nous pouvons écrire les équations d'Euler-Lagrange sous la forme de

$$\frac{dp_i}{dt} = f_i$$

Bien sûr, si on fait cela, ce n'est pas par raison esthétique, mais parce que la quantité  $p_i$  (contrairement à la quantité  $\dot{x}_i$ ) joue un très grand rôle. Supposons en effet que le Lagrangien ne contienne pas explicitement la variable  $x_i$ , alors l'équation d'Euler-Lagrange nous donne

$$\frac{d}{dt} p_i = 0$$

ou encore  $p_i = \text{Cte}$ , c'est à dire que l'impulsion  $p_i$  est une constante du mouvement et se conserve.

**Exemple** Considérons, en coordonnées polaire, une particule libre. Le lagrangien (= l'énergie cinétique, puisque  $V = 0$ ) s'écrit

$$\mathcal{L} = (m/2) \left( \dot{r}^2 + r^2 \dot{\theta}^2 \right)$$

Le lagrangien contient explicitement la variable  $r$ , mais pas la variable  $\theta$ . Donc

$$p_\theta = mr^2 \dot{\theta} = \text{Cte}$$

autrement dit, le moment cinétique se conserve.

Il existe une énoncée très précise à propos des lois de conservation et leur relation avec les symétries du système appelé théorème de Noether<sup>5</sup> Nous y viendront un peu plus tard.

---

5. Emmy Noether, 1918. Elle était une des personnes travaillant à l'époque phare de l'université de Goettingen, en compagnie des mathématiciens David Hilbert et Felix Klein et Herman Weyl, qui reformulaient la géométrie à travers la théorie des groupes et des invariants. Notons qu'en dehors des difficultés scientifiques, des femmes comme elle ou Kovalevskaya ou Skłodowska (Curie) devaient pénétrer un milieu mysogine exclusivement masculin.

Notons enfin qu'en prenant la transformée de Legendre du lagrangien par rapport à  $p$ , on aboutit à l'hamiltonien ; cela permet de transformer les équations différentielles de second ordre en (2 fois plus d') équations différentielles de premier ordre : les trajectoires s'obtiennent alors comme des flux dans l'espace des phases. Dans la suite de ce manuscrit, nous resterons avec les Lagrangiens.

## 4.1. Géométrie classique, formulation variationnelle et géodésiques.

Beaucoup de problème de géométrie sont formulés en terme variationnel : quelle est la courbe de longueur minimum entre deux points ? Quelle est la courbe fermée de périmètre fixé qui maximise l'aire englobé (problème isopérimétrique ou du berger) ? Quelle est la surface minimum qui relie deux courbes (problème de bulle de savon) ?

La première de ces questions joue un rôle fondamental en physique et s'appelle le problème de la détermination des géodésiques. Plaçons nous dans l'espace à deux dimensions muni de la notion habituelle de distance ; La courbe  $C$  de moindre longueur entre deux points  $P_1$  et  $P_2$  doit minimiser

$$L = \int_C ds$$

Le concept de distance minimum est purement géométrique et nous pouvons démontrer que la courbe de distance minimum est une droite par des méthode uniquement géométrique. Ceci dit, nous nous facilitons beaucoup la tâche si nous équipons notre espace de coordonnées cartésiennes  $(x, y)$  où l'élément d'arc d'écrit  $ds^2 = dx^2 + dy^2$ . Nous pouvons paramétrer notre courbe par une variable  $\zeta$  et chercher les deux fonctions  $x(\zeta)$  et  $y(\zeta)$ . Soit les deux extrémités  $P_1 = (0, 0)$  et le point  $P_2 = (a, b)$  correspondant au paramètre  $\zeta = 0$  et  $\zeta = 1$ . Alors le problème se ramène à minimiser la longueur

$$L = \int_0^1 \sqrt{\dot{x}^2 + \dot{y}^2} d\zeta$$

et les deux équations d'Euler-Lagrange se réduisent à

$$p_x = \frac{\dot{x}}{\sqrt{\dot{x}^2 + \dot{y}^2}} = K_1$$

$$p_y = \frac{\dot{y}}{\sqrt{\dot{x}^2 + \dot{y}^2}} = K_2$$

#### 4. Formulation lagrangienne de la physique et minimisations.

où  $K_1$  et  $K_2$  sont des constantes. En divisant la deuxième équation par la première, nous trouvons que  $\dot{y}/\dot{x} = K_3$ , ou encore

$$\frac{dy}{dx} = Cte.$$

qui est bien l'équation d'une droite. Notons également que nous avons une troisième loi de conservation

$$p_x^2 + p_y^2 = K_1^2 + K_2^2 = Cte$$

Ceci ne devrait pas être une surprise pour nous, puisqu'en regardant de plus près, nous voyons que le vecteur  $(p_x, p_y)$  est le vecteur tangent unitaire  $(dx/ds, dy/ds)$ .

Nous voyons d'ailleurs que nous aurions pu dès le début paramétrer la courbe par une des coordonnées, par exemple  $x$  et chercher directement la courbe  $y(x)$ . Nous aurions dû alors minimiser

$$L = \int_0^a \sqrt{1 + y'^2} dx$$

et trouver

$$\frac{y'}{\sqrt{1 + y'^2}} = K$$

ou encore

$$y' = K/\sqrt{1 - K^2} = cte$$

qui est bien sûr exactement le même résultat.

**Compliquons les choses.** Le problème est assez trivial jusque là, mais supposons que notre espace à deux dimensions n'est pas plate, comme par exemple en montagne. Dans ce cas, la distance la plus courte entre deux points n'est pas forcément la courbe  $y = ax + b$ , et parfois il vaut mieux de passer par un col un peu plus loin. La courbe de moindre distance entre deux points s'appelle un géodésique.

On peut regarder notre espace à deux dimensions comme une surface qui "baigne" dans un espace à trois dimensions et dont l'équation est donnée par  $z = f(x, y)$ . Nous pouvons comme avant, repérer les points de notre surface par deux coordonnées  $(x, y)$ ; par exemple, sur une sphère, nous avons besoin de deux angles  $(\theta, \phi)$  pour repérer un point. Quel est l'élément d'arc cette fois? Comme les deux points d'une surface baigne dans un espace à

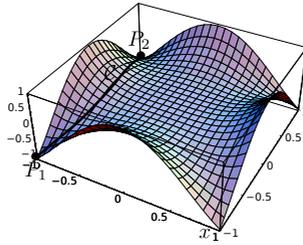


FIGURE 4.2. – Pour aller du point  $P_1$  au point  $P_2$ , la courbe de moindre longueur n'est pas  $x = 0$ , mais plutôt la courbe  $C$  montré en "gras".

trois dimensions,  $ds^2 = dx^2 + dy^2 + dz^2$ . Or,  $z = f(x, y)$  et donc  $dz = (\partial_x f)dx + (\partial_y f)dy$  et enfin

$$ds^2 = g_{11}dx^2 + (g_{12} + g_{21})dxdy + g_{22}dy^2$$

où  $g_{11} = (1 + (\partial_x f)^2)$ ,  $g_{12} = g_{21} = (\partial_x f)((\partial_y f))$ ,  $g_{22} = (1 + (\partial_y f)^2)$ ; notons que les coefficients  $g_{ij}$  sont des fonctions des coordonnées :  $g_{ij} = g_{ij}(x, y)$ . Pour trouver la géodésique, nous devons maintenant minimiser

$$L = \int_0^a \sqrt{g_{11} + (g_{12} + g_{21})y' + g_{22}y'^2} dx$$

**Exercice.** Obtenir l'équation d'Euler-Lagrange associé à ce problème.

Les coefficients  $g_{ij}$  sont les composantes de quelque chose qu'on appelle le tenseur métrique. Les habitants de la surface n'ont aucun élément pour savoir qu'ils baignent dans un espace de dimension supérieur ; par contre, ils peuvent mesurer les coefficients  $g_{ij}$  par des mesures purement locales, comme la somme des angles d'un (petit) triangle autour d'un point <sup>6</sup>. Il peuvent également penser que leur monde baigne dans un champ de force qui fait défléchir les droites, et appeler par exemple ce champ de force "la gravité". Dans ce cas, ils devraient même être capable, si ils cherchent les causes des  $g_{ii} \neq 1$  dans leur propre monde à deux dimensions, écrire des équations à dérivée partielle sur les  $g_{ij}$ , en supposant que ce tenseur lui même doit minimiser quelque chose à travers un lagragien.

---

6. Il nous faut un peu plus d'espace pour parler de ces mesures. Un très bon livre sur ce sujet est "la géométrie Riemannienne" par E. Cartan.

#### 4. *Formulation lagrangienne de la physique et minimisations.*

Nous poursuivrons ces idées dans l'espace-temps relativiste ; plus tard, au lieu de considérer l'espace courbe, nous étudierons une autre forme de couplage, plus simple, que nous appellerons champs électromagnétique. Mais il ne faut pas oublier que toutes les idées que nous exprimons correspondent à des objets géométriques et le choix des coordonnées ne doit pas occulter ce fait.

# 5. Les formes linéaires, les vecteurs, les tenseurs et les autres.

Nous allons nous donner un petit chapitre pour visiter les formes linéaires<sup>1</sup>. Il est possible que la première fois que vous voyez ce concept, vous vous demandiez pourquoi les mathématiciens ont inventé des objets aussi inutiles qui n'ont pas l'air d'apporter grand chose par rapport aux vecteurs. Si c'est le cas, un peu de patience. En fait, les formes linéaires sont le dual de notre espace vectoriel, une sorte de *l'autre côté du miroir*. La physique gagne énormément en beauté et cohérence à être formulé à l'aide de l'interaction entre ces objets et les vecteurs. De plus, cela ne demande pas beaucoup d'effort, puisque l'algèbre est déjà connu, c'est celui des vecteurs. Vous vous êtes demandé par exemple ce que signifie la trace d'une matrice et pourquoi elle est invariante? Vous pouvez évidemment dire que la trace est la somme des éléments diagonaux, mais cela ne vous donne pas la signification, seulement la façon de la calculer<sup>2</sup>.

## 5.1. Définitions.

Bon, commençons. Nous savons ce que sont les vecteurs. Donnons nous un espace vectoriel  $\mathcal{E}$ , c'est à dire un ensemble de vecteurs obéissant à certaines règles simples (voir annexe XXX). Nous avons besoin parfois d'associer des

---

1. Les formes linéaires sont apparu naturellement aux mathématiciens travaillant sur les ensembles de  $n$  équations à  $m$  inconnu : ce sont des vecteurs lignes au lieu des vecteurs colonnes. Ce concept se rencontrait partout, par exemple dans la théorie des équations différentielles linéaires et les conditions de solvabilité de celle ci. Le concept a pris une consistance géométrique propre et est devenu connu du grand public après l'apparition de l'article de Ricci et Levi-Civita en 1900 (*Méthodes de calcul différentiel absolu et leurs applications*), *Mathematische Annalen*, 54 :125-201, 1900) qui a marqué le début du calcul tensoriel tel qu'on le pratique aujourd'hui.

2. Par ailleurs, dès que vous utilisez des coordonnées non-cartésiennes, ou si votre espace n'est pas plat, vous ne pouvez pas échapper à utiliser les formes linéaires.

## 5. Les formes linéaires, les vecteurs, les tenseurs et les autres.

scalaires aux vecteurs. La norme d'un vecteur est par exemple un tel scalaire. La projection d'un vecteur sur un axe en est un autre exemple<sup>3</sup>. L'ensemble des fonctions qui associent un scalaire à un vecteur est bien sûr très vaste, mais celui qui nous intéresse particulièrement est l'ensemble des fonctions linéaires :

$$\begin{aligned} f : \mathcal{E} &\rightarrow \mathbb{R} \\ f(\lambda \mathbf{u} + \mu \mathbf{v}) &= \lambda f(\mathbf{u}) + \mu f(\mathbf{v}) \\ \lambda, \mu \in \mathbb{R} &, \quad \mathbf{u}, \mathbf{v} \in \mathcal{E} \end{aligned}$$

de telles fonctions sont appelées des formes linéaires<sup>4</sup>. Notez que la fonction “norme d'un vecteur” *n'est pas* une forme linéaire : la norme de la somme de deux vecteurs n'est pas la somme de leurs normes. Par contre, la composante (la projection sur un axe) d'un vecteur est bien une forme linéaire.

Bref, les formes linéaires agissent sur des vecteurs pour produire des scalaires. Il n'est pas difficile de voir que les formes linéaires constituent elles-mêmes un espace vectoriel : si  $\omega_1$  et  $\omega_2$  sont deux formes linéaires, nous pouvons donner un sens très précis à  $\lambda\omega_1 + \mu\omega_2$ . On peut aller un peu plus loin et démontrer que l'espace des formes linéaires définies sur  $\mathcal{E}$  a exactement la même dimension que  $\mathcal{E}$  ! L'espace de ces formes linéaires est appelé l'espace dual de  $\mathcal{E}$  et noté  $\mathcal{E}^*$ .

Bien, à ce stade, sans se noyer dans les abstractions, il faut juste retenir que  $\omega(\mathbf{u})$  où  $\mathbf{u}$  est un vecteur et  $\omega$  une forme linéaire, est un *scalaire*. En mécanique quantique, Dirac a introduit une notation très astucieuse pour distinguer les vecteurs et les formes linéaires : un vecteur est noté par  $|\psi\rangle$  (un bra), tandis que la forme linéaire agissant sur ce vecteur est noté par un  $\langle\phi|$  (un ket). L'action d'un bra sur un ket (un bracket donc) est noté  $\langle\phi|\psi\rangle$  et c'est bien sûr un scalaire.

Le lecteur est probablement plus familier avec la notation des vecteurs comme des “vecteurs colonnes” et les formes linéaires comme des “vecteurs lignes”. Pour distinguer les composantes de ces deux types d'objets, en géométrie nous utilisons des indices en haut et en bas pour distinguer les formes linéaires et les vecteurs. Nous allons voir que ces notations sont tout aussi astucieuses que celles de Dirac. Donnons nous une base dans l'espace  $\mathcal{E}$  comme par exemple les vecteurs  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ . Un vecteur quelconque  $\mathbf{u}$

---

3. Chaque composante d'un vecteur dans un repère donné est un scalaire, les composantes sont donc des fonctions agissant sur un vecteur et produisant un scalaire.

4. Le mot covecteur est également utilisé.

peut s'écrire comme une combinaison linéaire de ces vecteurs

$$\mathbf{v} = \sum_{i=1}^n x^i \mathbf{e}_i$$

Notez que nous avons noté les composants avec un indice en haut.

Einstein, lors du développement de la relativité générale (1910-1916), est devenu un très gros utilisateur du calcul tensoriel et s'est rendu compte qu'à chaque fois qu'il rencontrait une expression du genre  $x^i \mathbf{e}_i$ , il y avait un signe de sommation derrière. Il a donc laissé tomber le signe "somme" et poser la convention suivante : *un indice répété, une fois en haut, une fois en bas, veut dire la sommation sur cet indice.* C'est ce que l'on appelle la convention de sommation d'Einstein que nous suivront dorénavant. Ainsi, l'expression

$$\alpha_i \beta^k \Gamma_j^i$$

veut dire

$$\sum_{i=1}^n \alpha_i \beta^k \Gamma_j^i$$

Tandis que

$$\alpha_i \beta^k \Gamma_k^i$$

veut dire

$$\sum_{i=1}^n \sum_{k=1}^n \alpha_i \beta^k \Gamma_k^i$$

C'est juste une convention de notation qui nécessite un peu d'exercice mental au début mais qui nous simplifie la vie. De plus, nous remarquerons que les expressions bien formées ont toujours le même nombre d'indice haut que d'indice bas, un peu comme les parenthèses ouvertes et fermées, ce qui procure en plus une vérification syntaxique immédiat de ce que l'on écrit.

Revenons à nos oignons. Un vecteur quelconque s'écrit donc  $\mathbf{v} = x^i \mathbf{e}_i$ . L'espace des formes linéaires est également un espace vectoriel de même dimension que  $\mathcal{E}$ , nous pouvons donc nous donner une base dans cet espace  $\{\mathbf{e}^i\}$  et n'importe quelle forme linéaire  $\omega$  peut s'y écrire

$$\omega = y_j \mathbf{e}^j \quad y_j \in \mathbb{R}$$

Remarquez le changement des indices hauts et bas par rapport aux vecteurs : ceci est l'équivalent des bras et des kets de Dirac. Remarquez également que

le nom de l'indice répété n'a pas d'importance, puisque l'on somme dessus et que le résultat ne le contient donc pas.

Ayant en main les deux bases des deux espaces duaux, nous pouvons calculer l'action de n'importe quelle forme linéaire sur n'importe quel vecteur, en développant simplement les parenthèses<sup>5</sup> :

$$\omega(\mathbf{v}) = x^i y_j \mathbf{e}^j(\mathbf{e}_i) \quad (5.1)$$

Il nous suffit de connaître les  $n^2$  nombre  $a_i^j = \mathbf{e}^j(\mathbf{e}_i)$  pour être capable de calculer l'action de n'importe quelle forme sur n'importe quel vecteur. D'ailleurs, nous pouvons laisser tomber les parenthèse et écrire simplement  $\mathbf{e}^j \mathbf{e}_i$  au lieu de  $\mathbf{e}^j(\mathbf{e}_i)$ , cela n'a pas d'ambiguïté et nous économise deux parenthèses. Il ne vous pas échappé que les nombres  $a_i^j$  peuvent-être ordonnée dans un tableau à deux dimensions que nous appelons une matrice<sup>6</sup>.

Résumons : le calcul algébrique de  $\omega \mathbf{v}$  se fait en trois étapes : on décompose  $\omega$  sur une base de l'espace  $\mathcal{E}^*$  pour trouver les nombres  $y_j$  ; on décompose  $\mathbf{v}$  sur une base de l'espace  $\mathcal{E}$  pour trouver les nombres  $x^i$  ; nous avons déterminé une fois pour toute les nombres  $a_i^j = \mathbf{e}^j \mathbf{e}_i$  ; On combine tous ces nombres dans une expression telle que (5.1), et voilà. Remarquer que si nous avons choisi les bases intelligemment pour que

$$\mathbf{e}^j \mathbf{e}_i = \delta_i^j \quad (5.2)$$

la double sommation de l'expression (5.1) se ramène à une simple sommation

$$\omega \mathbf{v} = x^i y_i$$

c'est ce que vous faites quand vous multipliez un vecteur ligne (l'expression d'une une forme linéaire dans une base) par un vecteur colonne (l'expression d'un vecteur dans une base) : vous assumez que vous avez choisi une base telle que (5.2).

## 5.2. Association bi-univoque entre une forme et un vecteur à travers une expression quadratique.

Une forme linéaire est une machine (application) qui prend un vecteur en entrée et produit un scalaire en sortie, et cela de façon linéaire. Une forme

5. Faites l'exercice si vous n'êtes pas habitué à ces calculs.

6. En parlé chic, un tenseur mixte de rang 2

bilinéaire est une application qui prend deux vecteurs en entrée et produit un scalaire en sortie, et cela de façon linéaire dans chacun de ses variables :

$$\begin{aligned} B : E \times E &\rightarrow \mathbb{R} \\ B(\lambda \mathbf{u} + \mu \mathbf{v}, \mathbf{w}) &= \lambda B(\mathbf{u}, \mathbf{w}) + \mu B(\mathbf{v}, \mathbf{w}) \\ B(\mathbf{w}, \lambda \mathbf{u} + \mu \mathbf{v}) &= \lambda B(\mathbf{w}, \mathbf{u}) + \mu B(\mathbf{w}, \mathbf{v}) \end{aligned}$$

Si nous nous sommes données une base  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  dans l'espace  $E$ , alors nous avons juste besoin de  $n$  nombre  $b_{ij} = B(\mathbf{e}_i, \mathbf{e}_j)$  pour entièrement caractériser  $B$  et pouvoir calculer son effet sur n'importe quel couple de deux vecteurs  $\mathbf{u} = x^i \mathbf{e}_i$  et  $\mathbf{v} = y^j \mathbf{e}_j$  :

$$B(\mathbf{u}, \mathbf{v}) = b_{ij} x^i y^j$$

il suffit pour cela de simplement développer les parenthèses selon les règles de la bilinéarité<sup>7</sup>.

Donnons nous une forme bilinéaire  $B$ ; nous allons voir plus tard que se donner une forme bi-linéaire que l'on considère "invariant" revient à fixer la géométrie de l'espace. Considérons maintenant la forme linéaire

$$\omega = B(\mathbf{u}, \cdot)$$

Cette définition veut dire simplement que l'action de  $\omega$  sur un vecteur  $\mathbf{v}$  est :

$$\omega \mathbf{v} = B(\mathbf{u}, \mathbf{v})$$

Nous voyons que ceci est une façon d'associer à chaque vecteur  $\mathbf{u} \in E$ , une forme linéaire  $\omega \in E^*$  : chaque vecteur a ainsi une "image miroir" dans l'espace dual.

**Exemple fondamental.** Supposons que nous nous sommes donnée une base dans  $E$  tel que  $B(\mathbf{e}_i, \mathbf{e}_j) = \delta_{ij}$ ; ceci est une base orthonormée vis à vis de l'application  $B$ . Définissons la forme linéaire

$$\mathbf{e}^i = B(\mathbf{e}_i, \cdot)$$

Il en va sans dire que

$$\mathbf{e}^i \mathbf{e}_j = \delta_j^i$$

---

7. On appelle l'application linéaire ci-dessus un tenseur (0,2). Nous aurions pu imaginer d'autres type d'application bilinéaire, par exemple une prenant en entrée une forme linéaire et un vecteur (tenseur (1,1)), ou prenant en entrée 2 formes linéaires et 0 vecteur (tenseur (2,0)). Les composantes de ces tenseurs sont notées  $b_j^i$  et  $b^{ij}$ .

et il ne faut pas beaucoup d'effort pour démontrer que les  $\mathbf{e}^i$  forment une base de l'espace dual. Donnons nous maintenant un vecteur  $\mathbf{x} = x^i \mathbf{e}_i$ , la forme linéaire associée est  $\omega = B(\mathbf{x}, \cdot)$ . Dans la base des  $\mathbf{e}^i$ , nous avons

$$\omega = y_i \mathbf{e}^i$$

et il n'est pas difficile de démontrer que

$$y_i = x^i$$

C'est l'opération de transposition habituelle : pour obtenir un vecteur ligne à partir d'un vecteur colonne, juste "tourner" le vecteur.

De façon général, si notre forme bilinéaire s'écrit  $b_{ij}$  dans la base des  $\mathbf{e}_i$ , un vecteur  $x^i \mathbf{e}_i$  est associée à la forme  $y_i \mathbf{e}^i$  où

$$y_i = b_{ij} x^j$$

**Exemple fondamental 2.** Dans l'espace à 4 dimensions, supposons la forme bilinéaire  $B$  donnée, dans une base particulière, par

$$\begin{aligned} b_{11} &= 1; b_{22} = b_{33} = b_{44} = -1 \\ b_{i \neq j} &= 0 \end{aligned}$$

nous voyons alors qu'au vecteur  $(x^1, x^2, x^3, x^4)$ , nous associons la forme linéaire  $(x_1 = x^1, x_2 = -x^2, x_3 = -x^3, x_4 = -x^4)$ . Une liberté de langage souvent utilisé est de dire qu'au quadri-vecteur  $(x^1, \vec{x})$  nous associons la forme linéaire  $(x^1, -\vec{x})$ .

### 5.3. Covariant et contravariant.

Pour distinguer les deux types d'objets, vecteurs et formes, Ricci et Levi-Civita, dans leur article fondateur de 1900, ont utilisé les mots *contravariants* et *covariants*. Ces mots sont très répandus, et nous pourrions également les utiliser de temps à autres. Bien qu'au début, l'utilisation de ces mots étaient réservées aux composants des vecteurs et des formes, par abus de langage cela s'est étendu aux vecteurs et aux formes, et on peut ainsi entendre les mots "vecteur contravariant" pour dire vecteur et "vecteur covariant" pour dire forme. L'origine de ces mots est le suivant.

Prenons un vecteur  $\mathbf{x} = x^i \mathbf{e}_i$ . En notation matricielle, nous écrivons

$$\mathbf{x} = (\mathbf{e}_1, \dots, \mathbf{e}_n) \begin{pmatrix} x^1 \\ \dots \\ x^n \end{pmatrix}$$

c'est à dire que les composants  $x^i$  sont représentés par un vecteur colonne. Supposons que maintenant, nous changeons de base et utilisons la base

$$(\mathbf{e}'_1, \dots, \mathbf{e}'_n) = (\mathbf{e}_1, \dots, \mathbf{e}_n)P$$

où  $P$  est la matrice de passage. Les composantes du vecteur dans cette base seront les  $x'^i$  reliés aux anciens composants par

$$\begin{pmatrix} x'^1 \\ \dots \\ x'^n \end{pmatrix} = P^{-1} \begin{pmatrix} x^1 \\ \dots \\ x^n \end{pmatrix}$$

Ceci est obligatoire, puisque le vecteur  $\mathbf{x}$  ne doit pas dépendre de la base dans laquelle on l'a exprimé. Ce que nous voyons ici c'est que nous devons faire effectuer aux quantités  $x^i$  l'opération inverse de ce que nous avons fait subir aux vecteur  $\mathbf{e}_i$ , d'où le mot "contravariant" associé aux  $x^i$ . Nous laissons au lecteur l'explication du mot covariant.

## 5.4. Formes quadratiques invariantes et groupes de transformation.

Qu'est ce que la distance en géométrie classique? Elle est imbriquée dans l'idée que nous pouvons faire certaines transformations qui laissent invariant l'essence d'un objet. En l'occurrence, en géométrie classique, nous assumons que les rotations constituent un tel groupe de transformation : nous pouvons "tourner" ( et bien sûr translater) un triangle pour le superposer à un autre et voir si ils se couvrent parfaitement. Tourner un triangle ne change pas l'identité du triangle et laisse invariant son essence. Donnons nous maintenant des coordonnées cartésiens et deux points  $P_1 = (x^i)$  et  $P_2 = (y^i)$ . Nous pouvons vérifier sans peine que les opérations de rotation, que nous acceptons comme conservant l'essence des choses, laissent invariant la quantité  $d = (x^i - y^i)^2$ , et cette forme quadratique elle-même est associée à la forme bilinéaire de l'exemple fondamental cité plus haut. N'oublions pas que les vecteurs et les formes sont associées à travers cette forme bilinéaire.

Dans l'espace à quatre dimension dans lequel nous vivons, ce ne sont pas les rotations qui conservent l'essence des choses, mais les transformations de Lorentz que nous avons rencontrés au début de ce manuscrit. Ceci est un fait expérimental, et en rien plus étrange que les rotations; nous sommes simplement plus habitués aux rotations qu'aux groupes de Lorentz. Comme nous ne savons, les transformations de Lorentz laissent invariants la quantité

$ds^2 = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2$ ; cette forme quadratique est associée à la forme bilinéaire de l'exemple fondamental 2. Pour cette raison, quand on passe des vecteurs aux formes dans notre espace, nous devons changer le signe des composants spatiaux. La forme linéaire  $(E, -\vec{p})$  devient ainsi le vecteur  $^8 (E, \vec{p})$ . C'est une gymnastique à laquelle nous devons nous habituer pour la raison suivante : certaines quantités que nous obtiendrons seront des vecteurs, d'autres des formes ; si l'on veut écrire des égalités, comme par exemple des équations de mouvement, on doit transformer toutes les quantités en vecteur ou en forme, mais nous ne pouvons jamais évaluer une forme et un vecteur, se sont des objets appartenant à deux mondes différents.

Notons enfin que nous pouvons nous donner une forme quadratique et déduire le groupe de transformation linéaire qui la laisse invariant, ou nous donner un groupe de transformations et obtenir la forme quadratique qu'elles laissent invariant. Ces deux concepts sont strictement équivalents.<sup>9</sup>

## Groupes de Lorentz et de Poincaré.

### 5.5. La dérivation.

Soit une quantité distribuée dans l'espace ; nous équipant d'un système de coordonnées  $x^i$ , nous représentons cette quantité par la fonction scalaire  $f(x^1, \dots, x^n)$ . Nous pouvons former les dérivées partielles de cette quantité

$$p_i = \frac{\partial f}{\partial x^i}$$

Les éléments  $p_i$  doivent être considérés comme des éléments d'un vecteur ou d'une forme ? En langage plus usité, les  $p_i$  sont-ils contravariant ou covariant ? En notant l'indice en bas, nous avons bien sûr déjà répondu par co-

---

8. Si nous n'avions pas les contraintes de typographies, nous le noterions comme un vecteur colonne.

9. Le concept que la géométrie peut se ramener à la donnée d'une forme quadratique a été formulé pour la première fois par Riemann dans les années 1840 et a constitué la naissance des géométries non-euclidiennes. Lobachevsky l'avait précédé dans la formulation d'une forme particulière de géométrie non-euclidienne, mais avec des méthodes de la géométrie classique. Riemann cependant n'envisageait que les formes quadratiques définies positives, qui ne couvrent pas l'espace-temps relativiste. L'équivalence entre les groupes de transformations et les formes quadratiques invariantes a été formulé par Felix Klein vers 1870 (le programme Erlingen). Le groupe de transformation associé à l'espace physique qui nous entoure est le groupe de Lorentz, nom utilisé par Poincaré en 1905 quand il a formulé la relativité en langage de théorie des groupes. Le groupe le plus général qui comprend les translations et le groupe de Lorentz est appelé groupe de Poincaré.

variant et donc savons qu'ils forment les éléments d'une forme. Mais on peut établir cela sans ambiguïté en considérant la quantité  $df$ , qui par définition, s'écrit

$$df = \left( \frac{\partial f}{\partial x^i} \right) dx^i$$

La quantité  $df$  étant un scalaire<sup>10</sup>, les  $p_i$  doivent être covariants ( nous savons que les  $dx^i$  sont contravariant).

Prenons maintenant  $n$  quantités fonctions d'une seule variable, par exemple les coordonnées d'un point le long d'une courbe :  $x^i = x^i(t)$ . Nous pouvons former les dérivées

$$v^i = \frac{dx^i}{dt}$$

De quelle nature sont les quantités  $v^i$  ? La réponse est contenue dans la question précédente : considérons la variation de la quantité  $f$  le long de cette courbe. Nous avons

$$\begin{aligned} \frac{df}{dt} &= \left( \frac{\partial f}{\partial x^i} \right) \left( \frac{dx^i}{dt} \right) \\ &= p_i v^i \end{aligned}$$

Les  $v^i$  doivent être contravariants et forment un vecteur. Nous verrons plus tard que les impulsions s'obtiennent comme les  $p_i$  et les vitesses comme les  $v^i$ .

Dans beaucoup de livres, la covariance ou contravariance est déterminée en étudiant l'effet d'un changement de coordonnées ; nous le rappelons ici, sans insister. Faisons un changement de coordonnées (imaginez par exemple passer de cartésien à polaire) et appelons les nouvelles coordonnées  $x'^i$ . Les règles de dérivation nous donnent :

$$\begin{aligned} p'_i &= \frac{\partial f}{\partial x'^i} \\ &= \left( \frac{\partial x^j}{\partial x'^i} \right) \left( \frac{\partial f}{\partial x^j} \right) \end{aligned}$$

Pour trouver les composantes dans les nouvelles coordonnées, nous avons besoin des dérivées partielles des anciennes coordonnées par rapport aux nouvelles : anciennes  $\rightarrow$  nouvelles, donc covariant.

---

10. Dans la théorie des formes différentielles développée par E.Cartan dans les années 1920,  $df$  est appelé une *forme* différentielle et n'est donc par un scalaire. Nous avons dédié un chapitre sur la formulation de la relativité par les formes différentielles plus loin.

Prenons maintenant le cas des  $v^i$  :

$$\begin{aligned} v'^i &= \frac{dx'^i}{dt} \\ &= \left( \frac{\partial x'^i}{\partial x^j} \right) \frac{dx^j}{dt} \end{aligned}$$

Cette fois, nous avons utilisé les dérivées partielles des nouvelles coordonnées par rapport aux anciens : nouvelles  $\rightarrow$  ancienne, donc contravariant.

Vous commencez peut-être voir une des utilités des indices hauts et bas : ils suivent les mêmes règles que la combinaison des fractions rationnelles, genre :

$$\frac{\frac{4}{3}}{\frac{2}{8}}$$

**Exercice.** Soit le champs de vecteurs  $v^i(x^k)$ . Nous pouvons définir la dérivée de ce champs de vecteur par un tenseur dont les éléments sont

$$p_j^i = \frac{\partial v^i}{\partial x^j}$$

Est ce que notre notation est correcte ? Ou fallait-il noter  $p_{ij}$  ou peut-être  $p^{ij}$  ?

## 5.6. Retour sur la formulation lagrangienne des problèmes de géométries.

Indépendance par rapport au paramétrisation. Contre exemple : la mécanique classique.

Les fonctions homogènes.

La forme linéaire d'Euler-Lagrange  $\tilde{E}$ , nouvelle formulation des équations d'Euler-Lagrange : la trajectoire où  $\tilde{E} = 0$

## 5.7. Les tenseurs.

**Note.** Cette section n'est pas absolument nécessaire pour poursuivre la dynamique relativiste, en tout cas pas en première lecture. Le lecteur peut la contourner et passer directement à la suite, quitte à y revenir plus tard.

Cela dit, c'est une généralisation triviale de ce que nous venons de dire et ne demande pas beaucoup d'effort.

Après cet avertissement, revenons à nos moutons. En algèbre classique, nous pouvons, étant donné un nombre  $x$ , définir un nombre  $ax^2$  ou  $bx^3$  ou  $cx^n$ . Ces puissances nous sont très utile pour par exemple approximer une fonction proche d'un point :

$$f(x+h) = f(x) + f'(x)h + (1/2)f''(x)h^2 + \dots$$

Pouvons nous étendre ces concepts aux fonctions vectoriels, définir un vecteur au cube et le multiplier par une sorte de dérivée troisième? La réponse est oui. Considérons par exemple le nombre  $ax^2$ ; on peut définir une fonction de deux variables

$$f(x, y) = axy$$

et donner un sens à ce que veut dire  $ax^2$  en utilisant la fonction  $f$  :

$$ax^2 = f(x, x)$$

Cette façon un peu savante de redéfinir quelque chose qui nous est familier a l'avantage suivant : la fonction  $f$  est bilinéaire, c'est à dire linéaire dans chacune de ses variables :

$$f(x+y, z) = f(x, z) + f(y, z)$$

ou plus généralement,

$$f(\lambda x + \mu y, z) = \lambda f(x, z) + \mu f(y, z)$$

et bien sur la même chose pour l'autre variable<sup>11</sup>. Cela nous ouvre une voie royale pour définir la même chose pour les vecteurs :

**définition.** Un tenseur (covariant) de rang 2 est une application qui à deux vecteurs associe un scalaire (réelle ou complexe, mais ici les réelles nous suffisent) , et cela de façon bi-linéaire :

$$\begin{aligned} f & : E \times E \rightarrow \mathbb{R} \\ f(\lambda \mathbf{x} + \mu \mathbf{y}, \mathbf{z}) & = \lambda f(\mathbf{x}, \mathbf{z}) + \mu f(\mathbf{y}, \mathbf{z}) \\ \mathbf{x}, \mathbf{y}, \mathbf{z} \in E & , \quad \lambda, \mu \in \mathbb{R} \end{aligned}$$

---

11. Notez que nous pouvons nous payer le luxe de ne *pas* exiger de  $f$  d'être commutatif, nous n'en n'avons pas besoin.

et la même chose pour la deuxième variable.

Pour représenter un vecteur, nous nous donnons une base et présentons le vecteurs par un  $n$ -tuplet de  $\mathbb{R}^n$  (où  $n$  est la dimension de l'espace vectoriel), par exemple on note  $\mathbf{x} = (x^0, \dots, x^n)$ ; souvent on parle de la composante  $x^i$  au lieu de parler du vecteur lui-même. Nous allons faire la même chose avec les tenseurs de rang 2 : supposons que nous connaissons les  $n^2$  nombre

$$f_{ij} = f(e_i, e_j)$$

dans ce cas, nous connaissons l'effet de l'application du tenseur à n'importe quel deux vecteurs  $\mathbf{x} = x^i e_i$  et  $\mathbf{y} = y^j e_j$  : il suffit simplement d'appliquer la bilinéarité et de développer les parenthèse :

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}) &= f(x^i e_i, y^j e_j) \\ &= x^i f(e_i, y^j e_j) \\ &= x^i y^j f(e_i, e_j) \\ &= f_{ij} x^i y^j \end{aligned}$$

$f_{ij}$  est la représentation matricielle du tenseur.

Remarquer que nous pouvons définir une forme linéaire  $\tilde{\omega}$  par le tenseur  $f$  et le vecteur  $\mathbf{x}$  :  $\tilde{\omega} = f(\mathbf{x}, \cdot)$ . Ceci veut simplement dire que l'application de  $\omega$  à un vecteur  $\mathbf{y}$  donne le scalaire

$$\tilde{\omega}y = f(\mathbf{x}, \mathbf{y})$$

et cela ne vous prendra pas trop de temps de démontrer que dans la base conjuguée aux vecteurs  $e_i$ , les composantes de la forme s'écrivent :

$$\omega_j = f_{ij} x^i$$

Remarquez l'élégance de cette notation d'indice haut et bas, qui deviennent presque automatique.

Une autre façon de voir le tenseur (covariant) de rang 2 est de le voir comme le produit tensoriel<sup>12</sup> de deux formes linéaires. Donnons nous deux formes  $\tilde{\omega}$  et  $\tilde{\omega}'$  et définissons leurs produit tensoriel  $\tilde{\omega} \otimes \tilde{\omega}'$  par son action sur deux vecteurs :

$$\tilde{\omega} \otimes \tilde{\omega}'(\mathbf{x}, \mathbf{y}) = (\tilde{\omega}\mathbf{x})(\tilde{\omega}'\mathbf{y})$$

vous voyez que ce produit est un produit bilinéaire sur les vecteurs avec un résultat scalaire, c'est donc tout a fait un tenseur de rang deux comme nous

---

12. Au sens de la théorie des ensemble

5. *Les formes linéaires, les vecteurs, les tenseurs et les autres.*

l'avions défini. De plus, nous pouvons décomposer ces deux formes sur la base conjuguée des  $\tilde{e}^i$  et écrire ce tenseur comme

$$f_{ij}\tilde{e}^i \otimes \tilde{e}^j$$

**Généralisaton.** Un tenseur  $(r, s)$  est une application  $(r + s)$  linéaire sur  $r$  formes et  $s$  vecteurs.

## 6. Dynamique relativiste : formulation lagrangienne.

La formulation lagrangienne donne une puissance incroyable à la mécanique, en l'unifiant avec la géométrie. En mécanique non relativiste, la trajectoire d'une particule peut-être vue comme un géodésique dans un espace où la métrique est définie par la fonction potentiel. En relativité, ce lien devient encore plus élégants, puisque les champs d'interactions entre particules prennent eux-mêmes une signification géométrique. Nous allons maintenant visiter tout cela d'un peu plus près.

**Note importante :** A partir de maintenant, nous suivrons les conventions de notation de calcul tensoriel, faisant une distinction entre les vecteurs et les formes. Les coordonnées seront notées dorénavant par les indices en haut comme  $x^i$ . Les vecteurs seront noté en gras, comme par exemple  $\mathbf{p}$ , les formes par un tilde au dessus du symbole  $\tilde{p}$ . Ainsi, nous pouvons par exemple écrire  $\mathbf{p} = (p^0, p^1, p^2, p^3)$  et  $\tilde{p} = (p_0, p_1, p_2, p_3)$ . Il est de coutume de grouper les trois composante spatiale dans un tri-vecteur (les vecteurs de la dynamique classique), puisque la physique était formulée en termes de tri-vecteurs et que ces objets nous sont assez familier<sup>1</sup>. Les tri-vecteurs seront notés par une flèche au-dessus du symbole, par exemple  $\vec{p}$ . On peut donc voir par exemple une notation comme  $\mathbf{p} = (p^0, \vec{p})$ ; de plus, souvent en physique, la composante temporelle a un nom bien établi qu'on pourra utiliser, par exemple  $\mathbf{p} = (E, \vec{p})$  où  $\mathbf{A} = (V, \vec{A})$ .

**Note importante bis.** Comment passer des vecteurs aux formes et vis et versa en relativité restreinte, du à la forme simple du tenseur métrique. Ne pas prendre l'expression  $g_{ij}v^j$  pour plus compliqué qu'il ne l'est.

---

1. Ceci vient également du fait que la signature de ces composantes est identique.

## 6.1. Lagrangien relativiste.

Revenons maintenant à l'espace physique à quatre dimensions. Les particules, comme nous avons souvent insisté, sont des courbes dans cet espace. La géométrie de ces courbes, tant qu'il n'y a pas d'autres courbes à proximité, est très simple : ce sont des géodésiques qui minimisent la distance entre deux points<sup>2</sup>, c'est à dire que ce sont les solutions du problème variationnel suivant<sup>3</sup> :

$$S = m \int_A^B ds$$

La seule différence est que notre espace physique est équipée de l'élément d'arc

$$ds^2 = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2$$

D'ailleurs, pour nous simplifier ces notations, donnons nous le tenseur métrique

$$g = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

cela nous permet d'écrire

$$ds^2 = g_{ij} dx^i dx^j = dx_i dx^i$$

Comme au chapitre précédent, nous pouvons définir le vecteur vitesse propre le long d'une courbe

$$v_p^i = \frac{dx^i}{ds}$$

---

2. Le lagrangien relativiste a été formulé par Max Plank en 1906.

3. Comme nous l'avons dit, le signe du lagrangien n'a pas d'importance, puisque les trajectoires sont les extremum de l'action, pas des minimum (ou des maximum) ; ce qui compte sont les équations des trajectoires, qui sont les mêmes quelque soit le signe. Certains auteurs (dont Landau) posent  $S = -m \int ds$ , pour être cohérent avec la mécanique classique : en effet, en développant  $-m ds = -\sqrt{1-v^2} dt$  à l'ordre 1, on trouve bien  $(1/2)mv^2$ , c'est à dire l'énergie cinétique classique. Le choix  $\mathcal{L} = -m ds$  a cependant ses propres incohérences formels, puisque la signature de l'espace est  $(-1)$  en relativité et  $(+1)$  en mécanique classique, ce qui donnerait alors pour le *vecteur* impulsion en relativité  $-m\mathbf{v}_p$  et en mécanique classique  $+m\mathbf{v}$  (choix utilisé par C. Lanczos par exemple). Pour rester parfaitement cohérent entre classique et relativiste, on peut utiliser la signature  $-1$  pour l'espace 3d classique et poser  $\mathcal{L} = V - T$ . Cela imposera quelques gymnastiques avec les multiplications par  $(-1)$  partout en mécanique classique, l'effort qui n'en veut pas la peine vu le faible gain, limité au formel.

## 6. Dynamique relativiste : formulation lagrangienne.

Si nous utilisons un paramètre  $\zeta$  pour paramétrer la courbe, alors

$$v_p^i = \gamma \frac{dx^i}{d\zeta}$$

où  $\gamma = d\zeta/ds$ . Évidemment, si nous avons choisi une des coordonnées, par exemple  $x_0$ , comme paramètre, nous aurions

$$\mathbf{v}_p = \gamma(1, \vec{v})$$

où  $\vec{v}$  et le 3-vecteur vitesse classique.

De la même manière, nous pouvons déterminer les impulsions, en paramétrant la courbe par  $\zeta$  :

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{x}^i} \tag{6.1}$$

$$= m\gamma g_{ij} \dot{x}^j \tag{6.2}$$

Comme avant, nous aurions pu utiliser  $x_0$  comme paramètre. Dans ce cas, la 4-forme linéaire impulsion s'écrit

$$\tilde{p} = m\gamma(1, -\vec{v})$$

et donc, la quadri-vecteur impulsion comme

$$\mathbf{p} = m\gamma(1, \vec{v})$$

Le premier composant est appelé énergie, les trois autres sont groupés en un 3-impulsion.

En suivant ligne à ligne ce que nous avons dit dans le cadre de la géométrie Riemannienne, nous pouvons démontrer que les géodésiques sont ici aussi, des droites.

**Note sur les notations.** Il est important de se souvenir qu'en relativité restreinte, le métrique  $g_{ij}$  a une forme extrêmement simple et diagonale. Ainsi, quand nous écrivons

$$p_i = g_{ij} p^j$$

pour passer des vecteurs aux formes, le membre de droite est une somme dont un seul des éléments,  $j = i$  est non nul. Nous avons donc simplement  $p_0 = p^0$  et  $p_1 = -p^1, \dots$ . Ne prenez donc pas ces expressions pour plus compliqué qu'elles ne le sont.

## 6.2. Couplage d'une particule au champ.

Nous savons que les particules interagissent les unes avec les autres, que les déflexions de la courbe d'univers d'une particule sont dues à la présence des autres et que l'action des autres peut-être représentée par un champ qui remplit l'espace. Quelle est la nature de ce champs ? Considérons quelques possibilités simples :

1. Ce champ pourrait-être un scalaire  $f(P)$ , fonction du point  $P$ , auquel cas nous aurions une action de type

$$S = m \int_C (1 + f(P)) ds$$

venant pondérer la contribution des divers éléments infinitésimaux de la courbe  $C$  par le champ  $f$ .

2. Le champs pourrait-être vectoriel  $\mathbf{A}(P)$ , auquel cas l'élément d'arc sera pondéré par un pseudo-produit scalaire du genre  $(\mathbf{A}, ds)$ , ou encore, si on utilise les composants covariants de  $\mathbf{A}$ , par une action du type

$$S = m \int_A^B ds + q \int_A^B A_i dx^i$$

où  $q$  est un coefficient de couplage au champ.

3. Le champs pourrait-être tensoriel et l'action du type :

$$S = m \int_A^B ds + h \int_A^B g_{ij} dx^i dx^j$$

Nous pouvons continuer à l'infini les possibilités, ayant en tête que les champs eux même devraient pouvoir être décrits géométriquement pour respecter la cohérence interne. Nous ne connaissons pas d'interaction correspondant au premier cas. Le deuxième cas par contre décrit très bien une très grande quantité d'expériences et est appelé le champ électromagnétique. Nous ne poursuivrons pas ici le troisième cas, qui nous mènerai vers la relativité générale et la gravité.

## 6.3. Couplage au champ électromagnétique, formulation invariante.

Donnons nous un champs de vecteur  $\mathbf{A}$  dans l'espace (à 4 dimensions). Dans un système de coordonnées ce champ s'écrit  $(A^0, A^1, A^2, A^3)$ . Une certaine habitude use également la notation  $(V, \vec{A})$ , c'est à dire que la première

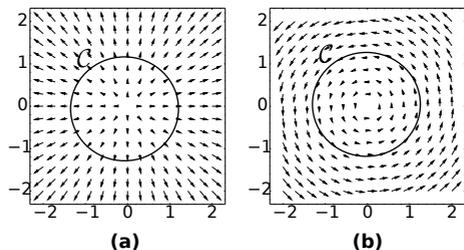


FIGURE 6.1. – Exemple de deux champs autour d’un point  $P$ , (a) sans torsion et (b) avec torsion. Pour mesurer la torsion, il faut intégrer le champ le long d’une courbe fermée : si le champ possède une torsion, les projections du champ sur la courbe s’ajoutent et l’intégrale est non-nulle. Le tenseur  $\tilde{F}$  mesure cette torsion.

composante prend le nom de “potentiel électrique” et les trois derniers sont groupés sous la dénomination “potentiel vecteur”. Pour la forme linéaire associée, nous avons  $\tilde{A} = (A_0, A_1, A_2, A_3)$  où  $A_0 = A^0$  et  $A_i = -A^i$ ,  $i > 0$ . On peut encore noter la forme par  $\tilde{A} = (V, -\vec{A})$ .

A partir d’un champ de forme linéaire, nous pouvons construire<sup>4</sup> un tenseur antisymétrique d’ordre 2<sup>5</sup> qui nous servira beaucoup par la suite :

$$F_{ij} = \frac{\partial A_j}{\partial x^i} - \frac{\partial A_i}{\partial x^j}$$

Ce tenseur mesure la “torsion” d’un champ (de vecteur ou de forme) en un point (Fig.6.1) : donnons nous un point  $P$  et un petit domaine délimité par la courbe  $\mathcal{C}$  dans le plan  $(x^i, x^j)$ , et intégrons la forme  $\tilde{A}$  le long de cette courbe, normalisée par l’air  $\mathcal{D}$  du domaine :

$$I = \frac{1}{\mathcal{D}} \int_{\mathcal{C}} A_i dx^i$$

Si le champ était sans “torsion”, l’intégrale serait nulle (Fig.6.1) ; sinon, on peut démontrer sans peine que l’intégrale vaut exactement le coefficient  $F_{ij}$  que nous avons calculé (quand  $\mathcal{D} \rightarrow 0$ ).

4. En théorie des formes différentielles, ceci est appelé la dérivée extérieure de la forme  $A_i dx^i$

5. qui est connu sous le nom de rotationnel à trois dimension

## 6. Dynamique relativiste : formulation lagrangienne.

Ce tenseur  $F$  s'appelle "le tenseur électromagnétique"; Normalement, ce tenseur devrait avoir 16 éléments, mais comme il est anti-symétrique, les coefficients diagonaux  $F_{ii}$  sont nulles, et comme  $F_{ij} = -F_{ji}$ , nous voyons qu'il a seulement 6 éléments indépendants. L'habitude est de nommer les trois composantes contenant l'indice 0 (les  $F_{0i}$ , donc associés à la composante temporelle) le tri-vecteurs  $\vec{E}$  champ électrique<sup>6</sup>; les trois autres sont appelé champ magnétique  $\vec{B}$ . Les éléments covariants  $F_{ij}$  peuvent être mis sous forme d'une matrice :

$$\tilde{F} = \begin{pmatrix} 0 & E^1 & E^2 & E^3 \\ -E^1 & 0 & -B^3 & B^2 \\ -E^2 & B^3 & 0 & -B^1 \\ -E^3 & -B^2 & B^1 & 0 \end{pmatrix}$$

Nous aurions également pu construire le tenseur contravariant  $F^{ij}$ . Nous verrons plus tard que de façon similaire à l'invariance d'élément d'arc  $ds^2 = dx_i dx^i$ , la quantité  $F_{ij} F^{ij} = 2(B^2 - E^2)$  joue le rôle de l'invariant de Lorentz pour le champs électromagnétique. Le métrique  $g_{ij}$  est très simple en relativité et nous pouvons facilement passer de covariant à contravariant : à chaque fois que l'on déplace un indice du haut vers le bas ou l'inverse, si il est associé au composant spatial, on change de signe. Par exemple,  $F_{10} = -E^1$ , donc  $F^{10} = E^1$

Le tenseur  $F_{ij}$  apparaît très naturellement dès que l'on écrit les équations d'Euler-Lagrange de la particule couplée au champs. Comme nous l'avons dit, l'action relativiste, en présence du champ s'écrit

$$S = m \int_C ds + q \int_C A_i dx^i$$

Comme dans la précédente section, pour manipuler cette intégrale nous paramétrons la courbe par un paramètre  $\zeta$ . L'action s'écrit

$$S = \int_{\zeta_0}^{\zeta_1} \left\{ m \sqrt{g_{ij} v^i v^j} + q A_j v^j \right\} d\zeta$$

où  $v^i = dx^i/d\zeta$  et l'indice  $i$  varie de 0 à 3. L'impulsion avec champ s'écrit

$$\begin{aligned} p_i^{ac} &= \frac{\partial \mathcal{L}}{\partial v^i} \\ &= m \gamma g_{ij} v^j + q A_i \end{aligned}$$

---

6. Insistons que nous mesurons le temps dans les mêmes unités que l'espace et donc  $c = 1$ ; si on mesurait le temps avec ses propres unités, le vecteur  $\vec{E}$  devrait être remplacé par le vecteur  $\vec{E}/c$ .

## 6. Dynamique relativiste : formulation lagrangienne.

Nous voyons ici qu'en présence du champ, nous devons ajouter un terme supplémentaire à l'impulsion qui *ne dépend pas* que de la particule, mais également du champs (l'action de toutes les autres particules de l'univers). Nous continuerons d'appeler l'impulsion sans champ la quantité  $p_i = m\gamma g_{ij}v^j$ , l'impulsion avec champ sera noté  $p_i^{ac} = p_i + qA_i$ .

Pour obtenir l'équation d'Euler-Lagrange, nous avons besoin de la dérivée des  $p_i^{ac}$  : pour un  $i$  fixé, nous avons<sup>7</sup> :

$$\begin{aligned} \frac{dp_i^{ac}}{d\zeta} &= \frac{dp_i}{d\zeta} + q \frac{dA_i}{d\zeta} \\ &= \frac{dp_i}{d\zeta} + q \frac{\partial A_i}{\partial x^j} v^j \end{aligned} \quad (6.3)$$

Insistons : l'expression ci-dessus contient une *somme* sur l'indice répété  $j$ .

Bien, nous avons traité le côté impulsion des équations d'Euler-Lagrange, nous devons maintenant calculer les forces généralisées. Tous les  $A_j$  dépendent à priori de la variable  $x^i$ , donc :

$$\begin{aligned} f_i &= \frac{\partial \mathcal{L}}{\partial x^i} \\ &= \frac{\partial A_j}{\partial x^i} v^j \end{aligned} \quad (6.4)$$

En regroupant les forces et les impulsions, nous trouvons donc

$$\frac{d}{d\zeta} (\gamma m g_{ij} v^j) = q \left( \frac{\partial A_j}{\partial x^i} - \frac{\partial A_i}{\partial x^j} \right) v^j$$

où encore

$$\frac{dp_i}{d\zeta} = q F_{ij} v^j \quad (6.5)$$

Nous allons dans un instant montrer l'équivalence de cette équation avec la forme plus connue  $d\vec{p}/dt = q(\vec{E} + \vec{v} \times \vec{B})$ , mais voyons d'abord deux conséquences importantes.

**Vitesse propre.** Nous voyons d'abord que la quadri-“force” appliquée à la particule est *proportionnelle* à sa quadri-vitesse, le coefficient de proportionnalité étant un tenseur qui pondère les diverses directions de l'espace. Cela devient encore plus manifeste si le paramètre choisi est l'élément d'arc  $s$ ,

---

7. Pour les lecteurs traumatisés par la différence entre les dérivées totales et partielles, nous avons inclus un appendix (cf AA) qu'il serait utile de lire avant de continuer.

## 6. Dynamique relativiste : formulation lagrangienne.

et que la vitesse dans l'équation (6.5) devient la vitesse propre : l'équation s'écrit alors simplement

$$\frac{dp_i}{ds} = \frac{q}{m} F_{ij} p^j$$

Ceci est simplement une équation différentielle de premier ordre ; le fait qu'elle soit tensorielle n'est pas vraiment une complication, au plus cela veut dire que nous avons un système de quatre équations linéaires de premier ordre. Par ailleurs, cette forme est remarquablement bien adaptée à la résolution numérique du mouvement de la particule.

**Invariance de jauge.** La deuxième chose à remarquer est que le coefficient de proportionnalité ( le tenseur  $\tilde{F}$  ) ne fait intervenir que la *torsion* du champ  $\tilde{A}$  : un champ sans torsion n'a aucun effet sur la particule. Plus exactement, quelque soit le champ  $\tilde{A}$ , nous pouvons l'écrire comme la somme de deux champs, l'une sans torsion et l'autre avec<sup>8</sup> :

$$\tilde{A} = \tilde{A}_{s,t} + \tilde{A}_t$$

Il n'y a que la partie avec torsion,  $\tilde{A}_t$  qui aura un effet sur la particule<sup>9</sup>. Nous voyons d'ailleurs que cela n'a pas de sens de parler de la partie sans torsion du champs, puisqu'elle n'est pas *mesurable* , la seule façon de mesurer le champ étant d'étudier son effet sur une particule<sup>10</sup> !

Donc, nous pouvons toujours ajouter au champ  $\tilde{A}$  un terme arbitraire  $\tilde{A}_{st}$  sans aucune conséquence pour les phénomènes physique que l'on observe. Toutes les expressions que l'on obtient doivent être invariant vis à vis de cette opération. On appelle cela *l'invariance de jauge*.

Dans certains problèmes, il peut être utile d'ajouter des formes particulières de  $\tilde{A}_{st}$  au vrai champ mesurable, pour mieux traiter certains conditions aux limites ; on parle ainsi par exemple de la jauge de Coulomb ou de la jauge de Lorentz<sup>11</sup>. De façon générale, considérons une fonction quelconque

---

8. Ceci est comme écrire une fonction comme la somme d'une fonction paire et impaire, ou encore une matrice comme la somme d'une matrice symétrique et anti-symétrique.

9. Il faut mettre un bémol en ce qui concerne la mécanique quantique. L'expérience d'Aharonov-Bohm tendrait à montrer qu'en mécanique quantique, la partie sans torsion joue également un rôle. La discussion de ce sujet cependant nous emmenerai trop loin des objectifs de ce manuscrit.

10. Les années 1980 ont vu la montée des théories *post-moderne* en physique, où les mesures et observations n'ont plus d'importance et il n'y a que la cohérence interne de la théorie qui compte. C'est ainsi que certains se sont mis a multiplier le nombre de dimension de l'espace (le nombre 11 est souvent cité ces derniers temps) et le nombre de monde possible ( $10^{500}$ ) pour essayer d'enrayer les divergences de leurs équations.

11. Voir le chapitre sur la dynamique du champ

## 6. Dynamique relativiste : formulation lagrangienne.

des coordonnées  $f(x^i)$ . Construisons la forme linéaire  $\tilde{\omega}$  dont les composants sont

$$\omega_i = \frac{\partial f}{\partial x^i}$$

il n'est alors pas difficile de démontrer que la torsion du champ  $\tilde{\omega}$  est nulle, simplement parce que

$$\frac{\partial^2 f}{\partial x^i \partial x^j} = \frac{\partial^2 f}{\partial x^j \partial x^i}$$

L'inverse de cela est également vrai : si la torsion d'un champ est nul, alors il dérive d'une fonction scalaire<sup>12</sup>. Donc, quelque soit la fonction  $f$ , nous pouvons toujours ajouter son gradient (son quadri-gradient, qui est bien sûr une forme) au champ  $\tilde{A}$ ,

$$\tilde{A} \rightarrow \tilde{A} + \text{grad} f$$

cela n'aura aucune conséquence. Cette constatation est souvent énoncé pour le vecteur  $\mathbf{A}$ , auquel cas, il faut prendre en compte la signature de l'espace :

$$A^i \rightarrow A^i + g^{ij} \frac{\partial f}{\partial x^j}$$

### 6.4. La force de Laplace.

Nous avons formulé les équations du mouvement sans privilégier une des coordonnées. Beaucoup de calcul sont fait cependant en privilégiant la coordonnées  $x_0$  et l'utilisant pour paramétrer la courbe. Nous perdons énormément en esthétique, mais à la fin de la journée, il faut gagner sa vie et calculer correctement les trajectoires dans le repère du laboratoire. Voyons cela de plus près.

Notons que si nous groupons les trois termes spatiaux ( $i > 0$ ) et que nous passons aux tri-vecteurs,

$$\vec{p}^{ac} = \gamma m \vec{v} + q \vec{A}$$

---

12. Cela s'appelle Lemme de Poincaré. En langage des formes différentielles, cela s'écrit  $d(d\omega) = 0$

# 7. Électrodynamique.

Au chapitre précédent, nous avons obtenu la courbe d'univers d'une particule étant donné un champ électromagnétique. Ce champ électromagnétique est lui-même le résultat de la présence des courbes d'univers de toutes les particules, nous devons donc donner une dérivation cohérente de son propre équation, c'est à dire écrire des équations locale qui une fois intégrées, nous donne la forme de  $\mathbf{A}$  dans tout l'espace due à la présence de toutes les particules. Avant de faire cela, nous avons besoin de généraliser la formulation variationnelle que nous avons vu précédemment aux fonctions de plusieurs variables.

## 7.1. Formulation variationnelle des champs.

Au chapitre 4, nous avons vu des lagrangiens où les quantités  $x^i$  (et leurs dérivées) étaient fonction d'une seule variable indépendante que nous avons appelé  $t$ . Une courbe dans l'espace à  $n$  dimensions par exemple peut-être paramétrée par *une* variable quelconque. Nous pouvons généraliser cela à plusieurs variable indépendante. Considérons par exemple une corde élastique, et désignons par  $u$  la hauteur de la corde à l'abscisse  $x$  et au temps  $t$ . Nous voyons que toute description de la fonction  $u$  nécessite *deux* variables indépendante. Nous pouvons également imaginer  $u$  comme une surface dans un espace à trois dimensions.

De façon générale, donnons nous une fonction scalaire  $f(x^1, x^2, \dots, x^n)$  et supposons que nous avons une action de la forme

$$S = \int_{\mathcal{D}} \mathcal{L}(f, \partial f / \partial x^1, \dots, \partial f / \partial x^n, x^1, \dots, x^n)$$

Quelle est alors l'équation d'Euler-Lagrange qui décrit la dynamique d'un tel champ ?

Les notations deviennent légèrement lourdes ici, pour les alléger, posons  $f_{,i} = \partial f / \partial x^i$ . Nous pouvons alors démontrer aussi facilement qu'avant que

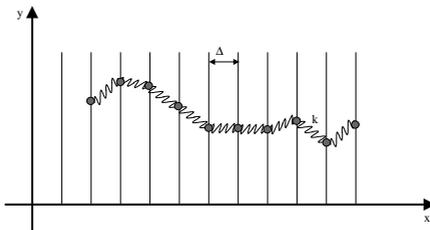


FIGURE 7.1. – Vue discrète d'une corde vibrante à un instant  $t_1$  donné, en fonction de l'abscisse  $x$ . La corde peut-être vue comme un ensemble de points reliés par des ressorts et contraints de se mouvoir seulement verticalement, approximation valide pour les faibles amplitudes.

l'équation d'Euler Lagrange s'écrit

$$\frac{\partial}{\partial x^i} \left( \frac{\partial \mathcal{L}}{\partial f_{,i}} \right) - \frac{\partial \mathcal{L}}{\partial f} = 0 \quad (7.1)$$

Cette fois, au lieu d'avoir  $\partial \mathcal{L} / \partial f'$ , nous avons une somme sur toutes les dérivations par rapport à toutes les variables indépendantes.

Faisons un exemple pour préciser les choses. Nous savons qu'en mécanique classique, le Lagrangien est  $L = T - V$ , l'énergie cinétique moins l'énergie potentielle. Reprenons notre corde élastique  $u(x, t)$ . Nous pouvons l'imaginer comme un ensemble de points reliés les uns aux autres, chaque point capable de bouger seulement verticalement. L'énergie cinétique de chaque point, d'abscisse  $x$ , de masse  $dm$  et de vitesse verticale  $\partial u / \partial t$  (figure 7.1)

$$T = (1/2)(dm) (\partial u / \partial t)^2$$

où  $dm = \rho dx$ . L'énergie potentielle de chaque point est proportionnel à l'écart à la hauteur de ces voisins (admettons cela pour l'instant)

$$V = (1/2)(dk)(\partial u / \partial x)^2$$

où la constante de ressort est  $dk = k dx$ . L'énergie cinétique et potentielle s'obtiennent en intégrant ces deux quantités sur toute la longueur de la corde  $L$  et l'action est l'intégrale de cette quantité sur l'intervalle de temps  $[0, T]$  :

$$\mathcal{S} = \frac{1}{2} \int_{[0, T] \times L} \left\{ \rho \left( \frac{\partial u}{\partial t} \right)^2 - k \left( \frac{\partial u}{\partial x} \right)^2 \right\} dx dt$$

## 7. Électrodynamique.

Le lagrangien est l'intégrand de cette intégrale. Si nous posons

$$u_{,t} = \partial u / \partial t ; u_{,x} = \partial u / \partial x$$

Nous pouvons écrire le Lagrangien comme

$$\mathcal{L} = (\rho u_{,t}^2 - k u_{,x}^2) / 2$$

et former les impulsions partielles ainsi que les forces :

$$\begin{aligned} \partial \mathcal{L} / \partial u_{,t} &= \rho u_{,t} \\ \partial \mathcal{L} / \partial u_{,x} &= -k u_{,x} \\ \partial \mathcal{L} / \partial u &= 0 \end{aligned}$$

Ce qui nous donne l'équation EL suivante :

$$\frac{\partial}{\partial t}(\rho u_{,t}) + \frac{\partial}{\partial x}(-k u_{,x}) - 0 = 0$$

ou encore, en nous souvenant de la définition des  $u_{,t}$  et  $u_{,x}$  :

$$\rho \frac{\partial^2 u}{\partial t^2} - k \frac{\partial^2 u}{\partial x^2} = 0$$

qui est l'équation bien connu de la propagation des ondes.

Un exemple plus général à plusieurs variables est le cas où  $f = f(x^1, \dots, x^n)$  et le Lagrangien s'écrit

$$\mathcal{L} = (1/2) \epsilon^{ij} \left( \frac{\partial f}{\partial x^i} \right) \left( \frac{\partial f}{\partial x^j} \right)$$

avec  $\epsilon^{ij} = \pm \delta^{ij}$ , ce qui nous donne l'équation d'euler-Lagrange

$$\sum \epsilon^{ii} \frac{\partial^2 f}{\partial (x^i)^2} = 0$$

Par exemple, au lagrangien

$$\mathcal{L} = \left( \frac{\partial u}{\partial t} \right)^2 - (\text{grad} u)^2$$

est associée l'équation d'onde

$$\frac{\partial^2 u}{\partial t^2} - \Delta u = 0$$

## 7. Électrodynamique.

Dans le cas le plus général, la fonction  $f$  lui même est vectoriel. La corde par exemple n'a aucune raison de ne vibrer que dans le plan vertical, nous pouvons l'imaginer ayant deux composantes  $u_y$  et  $u_z$ .

Dans le cas le plus général,  $f$  est une fonction vectorielle à  $m$  composantes  $(f^1, \dots, f^m)$  dépendant de  $n$  variables  $(x^1, \dots, x^n)$ , auquel cas nous obtenons une équation pour chaque composante :

$$\frac{\partial}{\partial x^i} \left( \frac{\partial \mathcal{L}}{\partial f^j_{,i}} \right) - \frac{\partial \mathcal{L}}{\partial f^j} = 0$$

Exemples : vibration d'une corde, champ élastique, équation d'élasticité.

Généralisation : Opérateur de Hodge, construction d'invariant associé au champ. [éviter le langage des formes différentielles pour l'instant].

Exemple 1 : fonction scalaire, gradient, forme du lagrangien, équation d'euler-lagrange avec Laplacien.

Exemple 2 : fonction vectorielle, forme du lagrangien, laplacien généralisée.

Exemple 3 : aperçu des champs tensoriels.

## 7.2. Formulation invariante du champs électromagnétique.

La forme de l'action du champs est une donnée expérimentale ; nous pouvons cependant nous donner quelques pistes de réflexions pour la trouver . Premièrement, nous cherchons une formulation purement géométrique du champ qui devra être indépendant des coordonnées. Nous avons vu que nous avons à notre disposition une quantité géométrique qui intervient dans les équation du mouvements : la torsion  $F$  du champs  $A$ . Cette torsion est cependant une quantité tensorielle, et le lagrangien doit être un scalaire. Nous devons donc construire un scalaire (avec une signification géométrique) à partir de la torsion. Nous pouvons construire plusieurs scalaire géométrique à partir du tenseur ; cependant, le plus naturel est de chercher quelque chose qui ressemble à l'invariant de Lorentz  $ds^2 = dx_i dx^i$  comme dans le cas des trajectoire. Le plus simple scalaire géométrique est alors

$$F_{ij} F^{ij}$$

**7.3. Les équations de Maxwell.**

**7.4. Interaction d'une particule avec son propre champ.**

## 8. Électrodynamique et l'action à distance.

Nous avons vu aux chapitres précédents que pour décrire l'interaction électromagnétique des particules, nous devons nous donner des particules *et* un champ, écrire un lagrangien en fonction de ces deux entités (et de leur couplage) et dériver ensuite les équation d'Euler-Lagrange pour donner la dynamique de chacun <sup>1</sup>.

Nous sommes maintenant tellement habitué au concept de champ que nous ne posons plus des questions sur sa réalité. D'un outil mathématiques pour simplifier les équations, le champ est devenu un objet réel et ayant sa propre existence. L'attitude positiviste en physique nous incite cependant à ne parler que des choses observables. Le champ est il observable ? La réponse n'est pas évidente : nous n'observons que des particules et le champ est détecté à travers son action sur les particules.

Peut on alors se passer complètement du champ électromagnétique et formuler une théorie qui ne prend en compte que les particules ? Pour un expérimentateur, la courbure de la ligne d'univers d'une particule est corrélée à la présence de la ligne d'univers d'autres particules à proximité. Pourquoi alors ne pas essayer de faire une théorie qui ne tient compte que de la géométrie des ces lignes, sans inventer un objet (champ) intermédiaire plus ou moins fictif ? Une telle théorie ressemblerai à la formulation originale de la gravité par exemple, où les corps agissent les uns sur les autres à travers l'espace, ce que l'on appelle action à distance.

La réponse à la question précédente est oui, à condition d'utiliser le concept de distance au sens de la géométrie Minkowkienne dans un espace à quatre dimension <sup>2</sup>. Dans l'espace de Minkowski, des points disjoints peuvent être

---

1. De façon séparé : soit on garde le champ fixe et on dérive la dynamique des particules, soit on grade les (lignes d'univers des) particules fixe et on dérive la dynamique du champ. Une approche variationnelle rigoureuse (faisant varier les deux entités) n'est pas possible et l'interaction d'une particule avec son propre champ pose des problème de cohérence de la formulation.

2. La théorie a été formulé par K. Shwarzschild en 1903 et développé par Trod (1922) et Fokker (1929). Wheeler et Feyman (*Classiscal Electrodynamics in terms of direct inter-*

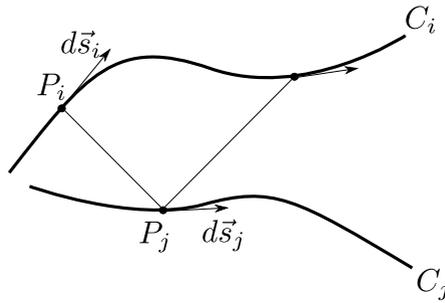


FIGURE 8.1. – L'action à “distance” entre deux lignes d'univers. Les points des deux courbes à distance nulle (connecté par un rayon lumineux) sont couplés et contribuent au lagrangien.

à distance nulle les uns des autres. Nous pouvons alors supposer que quand deux points appartenant à deux lignes d'univers sont à distance nulle<sup>3</sup>, ils exercent “une force” l'un sur l'autre et la géométrie des deux courbes est donnée par la somme de ces forces. Nous pouvons mettre cela sous forme variationnelle (Fig. 8.1) :

$$S = \sum_i m_i \int_{C_i} ds_i + \frac{1}{2} \sum_{i,j} \int \int_{C_i, C_j} q_i q_j \delta(\|P_i - P_j\|) ds_i ds_j \quad (8.1)$$

Le premier terme est l'action classique des particules isolées, où  $ds_i$  désigne l'élément d'arc le long de la courbe d'univers  $C_i$  de la particule  $i$ . Le deuxième terme décrit le couplage (deux par deux) des particules, où  $ds_i ds_j$  est le produit scalaire (Minkowskien) entre les deux vecteurs d'arc aux points  $P_i$  et  $P_j$  des deux courbes d'univers des deux particules. Seul les points à distance nulle (connecté par un rayon lumineux) contribuent au couplage, d'où la présence du delta de Dirac  $\delta(\|P_i - P_j\|)$  dans l'intégrale.

L'expression (8.1) a plusieurs traits intéressants. Premièrement, elle est formulée de façon entièrement géométrique et elle est donc invariante par transformation de Lorenz. Ensuite, nous pouvons remarquer qu'une particule  $i$ , le point  $P_i$  de sa courbe d'univers, interagit avec un point  $P_j$  d'une autre

---

*particle action*. Rev. Modern Phys., 21, 425. (1949)) l'ont utilisé pour tenter de résoudre le problème de self-interaction d'une particule chargée. Les développements récents peuvent être trouvés dans l'article de Hoyle et Narkila (*Cosmology and action at a distance electrodynamics*. Rev. Modern Phys., 67(1), 113. (1995).)

3. Les deux points sont en *contact*, au sens de Minkowski

## 8. Électrodynamique et l'action à distance.

particule dans son passé, mais également avec un point  $P'_j$  de la même particule dans *son futur*. Géométriquement parlant, il n'y a pas de passé et de futur, des concepts qui sont le résultat de notre choix de système de coordonnées. Remarquons que l'interaction d'une particule avec les ondes venant du futur n'est pas propre à la formulation "action à distance" de l'électrodynamique. Elle apparaît également dans la formulation classique avec champs (les solutions avancées) et elle est crucial si l'on veut rendre compte de la self interaction d'une particule avec son propre champ<sup>4</sup>.

Fixant toutes les lignes d'univers sauf celle de la particule 1, et en effectuant l'intégrale sur toutes les autres courbes, nous voyons que l'action pour la particule 1 s'écrit

$$S = m \int_C ds + q \int_C \mathbf{A} ds$$

où  $\mathbf{A}$  désigne l'intégrale sur toutes les autres courbes. Nous voyons que l'action pour une particule a exactement la bonne forme, pourvu que la quantité  $\mathbf{A}$  soit équivalent au champ électromagnétique. C'est ce que nous allons démontrer.

---

4. Voir n'importe quel *textbook* de l'électrodynamique qui traite de ce problème, comme par exemple le classique Jackson.

# 9. La formulation de la relativité à l'aide des formes différentielles.

Les tenseurs antisymétriques sont équivalents à des objets qu'on appelle les formes différentielles. Nous avons déjà rencontré les formes linéaires, qu'on appellera des 1-formes. Nous pouvons généraliser ces formes pour définir des  $n$ -formes. L'avantage est que les formes différentielles ont un contenu géométrique très fort, et les relations qui nous paraissent un peu compliquées en formulation tensorielle prennent un sens très simple quand formulées en terme de  $n$ -formes.

## 9.1. Les 1-formes.

Nous connaissons déjà les formes linéaires, que nous appellerons dorénavant les 1-formes : ce sont des applications linéaires agissant sur un vecteur et produisant un scalaire :  $\tilde{\omega}(\vec{u}) = a$ ,  $\vec{u} \in E$ ,  $\tilde{\omega} \in E^*$ ,  $a \in \mathbb{R}$ . Nous enlevons de plus les parenthèses et écrivons cela directement  $\tilde{\omega}\vec{u}$ . Par exemple, étant donné une fonction  $f(x)$ , nous pouvons associer à tout vecteur de l'espace  $\vec{t}$ , le scalaire  $(\text{grad}f)\vec{t}$  qui est la partie linéaire de la variation de la fonction  $f$  selon un déplacement  $\vec{t}$ .

Si nous nous sommes donné une base dans l'espace des vecteurs, nous disposons automatiquement d'une base dans l'espace des 1-formes différentielles. Supposons que nous nous donnons un système de coordonnées dans l'espace  $(x^1, \dots, x^2)$ . Cela veut dire que nous nous sommes données un ensemble de vecteurs  $(e_1, e_2, \dots, e_n)$  et un point d'origine  $O$ , tel que n'importe quel point de l'espace  $P$  peut être atteint<sup>1</sup> par le vecteur  $\vec{OP} = x^i e_i$ .

---

1. Nous voyons ici apparaître le concept de vecteur comme groupe de transformation de point. Nous pouvons imaginer toutes sortes de transformations  $T$  (mapping) des points de l'espace les uns dans les autres, qu'on peut écrire  $T(P) = P'$ . Parmi ces transformations, nous sélectionnons un sous-ensemble  $G$  tel que tous ses éléments commutent :  $g_1 g_2(P) = g_2 g_1(P) \forall g_1, g_2 \in G$ . De plus, nous exigeons de ces transformations de ne pas avoir de point

## 9. La formulation de la relativité à l'aide des formes différentielles.

A ces vecteurs  $e_i$ , nous associons des formes linéaires que nous appelons  $(dx^1, dx^2, \dots, dx^n)$ . Si  $\vec{t} = t^i e_i$ , alors  $dx^i(\vec{t}) = t^i$ . Une forme différentielle quelconque s'écrit alors, dans la base des  $dx^i$ ,  $\tilde{\omega} = a_i dx^i$ . Le lecteur rencontrant les formes différentielles (c'est leur nom) pour la première fois peut se sentir troublé : en tant que physicien, il avait l'habitude d'associer des "déplacements" infinitésimaux à ces symboles du genre  $dx$ , voilà que maintenant nous leurs donnons des significations torturées. A vrai dire, le changement est minime, nous avons chargé les vecteurs de gérer les déplacements et les formes différentielles de les "mesurer" en quelques sorte.

**Exemple.** Soit le système de coordonnées cartésien à 3 dimension, et le vecteur  $\vec{t} = (2, 3, 4)$ . Alors  $dx(\vec{t}) = 2$ ,  $dy(\vec{t}) = 3$ ,  $dz(\vec{t}) = 4$ .

**Exemple fondamental.** Etant donné une fonction  $f$ , la forme différentielle  $df$  et la forme qui au vecteurs  $\vec{t}$ , associe le scalaire

$$df(\vec{t}) = (p_i dx^i)(\vec{t})$$

où  $p_i = (\partial f / \partial x^i)$ . En gros,  $df(\vec{t})$  mesure le changement de  $f$  si l'on se déplace de  $\vec{t}$ .

**Champs de formes.** Si nous faisons dépendre la forme différentielle du point de l'espace, nous avons un champ de formes différentielles, exactement comme nous pouvons avoir des champs de vecteurs. Par exemple, on peut écrire en coordonnées cartésiennes,  $\omega = xdx + 3xydy + \sin(x)z^2 dz$ .

**Exemple.** Etant donné le champ de force de gravitation  $\vec{F} = -gdz$ , calculer le travail pour un déplacement  $\vec{t} = (2, 3, 4)$ . Le travail est l'application de la 1-forme force au vecteur déplacement, et nous avons ici  $W = -4g$ .

## 9.2. Intégration des formes différentielles.

Nous pouvons maintenant donner un sens géométrique à l'intégration, sans passer par la paramétrisation des courbes. Soit la courbe  $C$ , la 1-forme différentielle  $\omega$  (nous laissons tomber les  $\vec{\cdot}$ ) données en chaque point de l'espace.

---

invariant, c'est à dire qu'il n'existe pas un point particulier  $C$  tel que  $g(C) = C \forall g \in G$ . Les éléments de  $G$  sont appelés des vecteurs auxquels nous pouvons donner une structure d'espace vectoriel. L'espace muni de ces vecteurs est un espace affine. Nous ne continuons pas plus ces développements et leur généralisation aux espaces courbes (localement affines).

## 9. La formulation de la relativité à l'aide des formes différentielles.

Nous définissons l'intégrale de cette forme le long de la courbe comme suit : nous choisissons  $N + 1$  points  $P_0, \dots, P_N$  le long de la courbe, alors

$$\int_C \omega = \lim_{N \rightarrow \infty} \sum_{i=0}^N \omega(P_i \vec{P}_{i+1}) \quad (9.1)$$

Nous pouvons facilement vérifier que cette définition coïncide avec notre façon habituelle de paramétrer les courbes pour l'intégration.

**Exemple.** Intégrer la 1-forme  $\omega = ydx - xdy$  le long du demi-cercle de rayon 1 centré sur origine. On peut paramétrer la courbe par  $x = \cos t$ ,  $y = \sin t$ ; alors  $dx = -\sin t \cdot dt$  et  $dy = \cos t \cdot dt$ . Donc,

$$\int_C \omega = \int_0^\pi -(\sin^2 t + \cos^2 t) dt$$

### 9.3. Algèbre des formes différentielles.

Nous pouvons considérer une fonction  $f$  comme une 0-forme. Nous avons vu que nous pouvons construire à partir de cette 0-forme, une 1-forme dérivée  $df$  qui nous donne le changement de la 0-forme pour le déplacement  $\vec{t}$ . Nous pouvons continuer cette construction : quel est le changement induit dans la 1-forme  $\omega$  pour le déplacement  $\vec{t}$ ? Nous pouvons appeler cet objet  $d\omega$ . Si nous appelons  $E^*$  l'espace des 1-formes, nous voulons avoir  $d\omega(\vec{t}) \in E^*$ , et bien sûr nous exigeons de cet objet toutes les linéarités dont nous avons l'habitude.

$d\omega(\vec{t})$  est une 1-forme, donc nous pouvons l'appliquer à un vecteur  $\vec{s}$  pour obtenir un scalaire :  $d\omega(\vec{t})\vec{s} \in \mathbb{R}$ ; plus généralement, on peut appeler ce scalaire  $d\omega(\vec{t}, \vec{s})$ , c'est à dire que l'on peut considérer  $d\omega$  comme une 2-forme, prenant deux vecteurs en entrée et produisant un scalaire. On peut se demander d'ailleurs si il existe une relation entre  $d\omega(\vec{t}, \vec{s})$  et  $d\omega(\vec{s}, \vec{t})$ . Nous reviendrons à cette question plus un peu plus tard, généralisons les formes pour l'instant.

Nous appelons une  $k$ -forme, une fonction (application) linéaire de  $k$ -vecteurs à valeur réelle.  $\omega : E^k \rightarrow \mathbb{R}$ . De plus, nous exigeons de ces applications d'être anti-symétrique :  $\omega(\vec{t}_1, \vec{t}_2, \dots, \vec{t}_n) = -\omega(\vec{t}_2, \vec{t}_1, \dots, \vec{t}_n)$  ou n'importe quel autre permutation impaire de ses arguments. Il est trivial de démontrer que les  $k$ -formes ont une structure d'espace vectoriel, et si la dimension de l'espace est  $n$ , la dimension de l'espace  $E^{*k}$  des  $n$ -formes est  $n!/k!(n-k)!$

**Exemple.** Comment définir La 2-forme  $dxdy$  antisymétrique et linéaire ? Ces deux exigences nous laissent peu de choix, à part un signe que l'on peut fixer une fois pour toute à  $+1$  :  $dxdy(\vec{t}, \vec{s}) = dx(\vec{t})dy(\vec{s}) - dy(\vec{t})dx(\vec{s})$ . Nous voyons que si  $\vec{t} = \lambda\vec{s}$ , le résultat est nul, que cette définition est bien antisymétrique, et que  $dydx = -dxdy$ . De plus,  $dxdx = 0$ .

Les  $k$ -formes sont apparu naturellement lors du développement du calcul intégral des fonctions de plusieurs variables, et ce sont bien par exemple les  $dxdydz$  que nous voyons sous le signe de l'intégration. L'antisymétrie est nécessaire pour définir le sens de l'intégration et distinguer par exemple un flux entrant ou sortant d'une surface orientée. Le sens géométrique (comme application linéaire), indépendant du système de coordonnées, est apparu plus tard lors de travaux de E. Cartan. La définition (9.1) géométrique de l'intégration s'étend naturellement à l'intégration des  $k$ -formes sur un domaine  $D$  quelconque.

Nous avons mentionné la dimension des espaces  $E^{*k}$ . Par exemple, à 3 dimensions, l'espace des 1-formes et des 2-formes ont tout deux la dimension  $3!/2! = 3$ . C'est pour cela que nous (mal) représentons le rotationnel comme un vecteur. Notons que dans l'espace à  $n$  dimensions, les 0-formes et les  $n$ -formes sont tout deux de dimension 0, et qu'on les décrit à l'aide d'un seul scalaire, ou d'une seule fonction si leur valeur dépend du point de l'espace. De façon générale, l'espace des  $k$ -formes et des  $(n - k)$ -formes ont la même dimension. Nous viendrons un peu plus tard à la dualité entre ces deux espaces.

Notons enfin que sans trop insister, nous avons naturellement introduit le produit des formes : par exemple,  $dx$  et  $dy$  étant des 1-forme, nous avons utilisé une deux formes  $dxdy$ . De façon général, le produit d'une  $k$ -forme et d'une  $l$ -forme est une  $(k + l)$ -forme. on le note parfois comme  $\omega_1 \wedge \omega_2$ . Pour le calculer effectivement, il suffit de se donner une base et effectuer l'opération algébrique de multiplication, en respectant l'antisymétrie et la nullité des produit de type  $dxdx$ .

**Exemple fondamental : la  $k$ -forme flux.** Donnons nous, dans l'espace à  $n = 2$  dimension, un grand nombre de particules, qui comme nous le savons, sont caractérisées par leur courbe d'univers (Figure 9.1). Considérons un vecteur  $\vec{t}_1$  ayant pour origine le point  $P_1$ . combien de courbes traversent ce vecteur ? Nous appelons cette quantité scalaire le flux. Nous exigeons de cette quantité d'être algébrique, et donc changer de signe si l'on inverse l'orientation du vecteur  $\vec{t}_1$ . Si notre vecteur  $\vec{t}_1$  est petit (lire infinitésimal), le flux à travers ce vecteur est linéaire : le flux à travers le vecteur  $2\vec{t}_1$  est

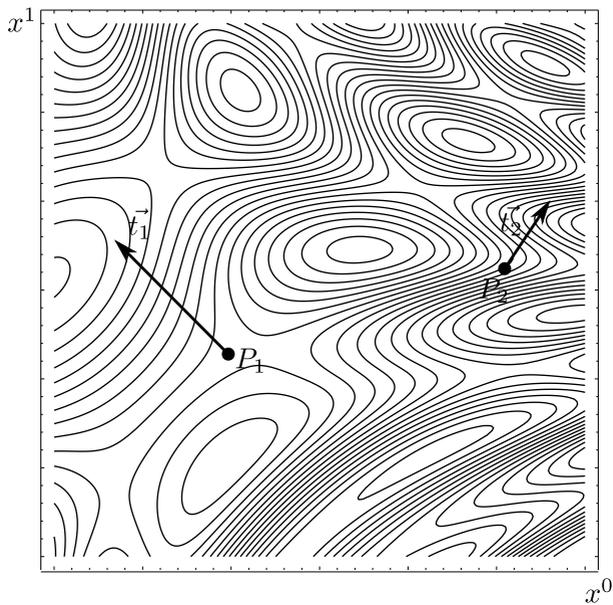


FIGURE 9.1. – Le flux de particules à travers une  $n - 1$  surface.

9. La formulation de la relativité à l'aide des formes différentielles.

deux fois plus grand que le flux à travers le vecteur  $\vec{t}$ . Nous voyons donc que le flux est une 1-forme appliquée à un vecteur. Appelons cette 1-forme le *courant*  $J$ . La mesure du flux dépend bien sûr du point de l'espace. Dans une base donnée, par exemple  $(x^0, x^1)$ , la forme pourrait s'écrire

$$J = j^0(x^0, x^1)dx^0 + j^1(x^0, x^1)dx^1$$

Si nous souhaitons connaître le flux à travers une courbe  $C$  non-infinitésimal, il suffit d'intégrer cette forme :

$$\Phi = \int_C J$$

Si la ligne  $C$  est orientée seulement dans la direction, disons  $x^1$ , le flux s'écrira

$$\Phi = \int_a^b j^1(x^0, x^1)dx^1$$

Ceci d'ailleurs est une façon de *mesurer* la 1-forme courant : en chaque point de l'espace, on mesure le flux à travers deux vecteurs selon les deux directions de la base, et nous avons accès ainsi à  $j^0$  et  $j^1$  en tout point.

Passons maintenant à l'espace à  $n = 3$  dimensions. Etant donné un point  $P$  est une 2-surface portée par deux vecteur  $(\vec{t}, \vec{s})$ , nous pouvons nous demander combien de ligne traversent cette surface? Cette fois, le flux est une 2-forme,  $\Phi = J(\vec{t}, \vec{s})$ . Bien sûr, nous pouvons continuer la généralisation, définir une 3-formes flux dans l'espace à  $n = 4$  dimensions

$$\omega = j^0(\mathbf{x})dx^1 dx^2 dx^3 + j^1(\mathbf{x})dx^0 dx^2 dx^3 + \dots$$

Si nous nous donnons une 3-surface  $D$ , (en géométrie habituelle, nous appelons cela un volume), le flux à travers elle est donnée par

$$\Phi = \int_D J$$

En particulier, si nous avons orienté la 3-surface de façon à être portée uniquement par les vecteurs  $\vec{e}_1, \vec{e}_2, \vec{e}_3$ , nous avons

$$\Phi = \int_D j^0(\mathbf{x})dx^1 dx^2 dx^3$$

Cette quantité est souvent appelée *la masse* contenue dans la 3-surface  $D$ , et  $j^0$  prend alors le nom de *densité*. Pour éviter les problèmes d'ordenancement des axes, le courant est parfois noté sous forme compacte par

$$J = \epsilon_{abcd}j^a dx^b dx^c dx^d$$

où  $\epsilon_{abcd}$  vaut  $\pm 1$  selon que la permutation  $(abcd)$  est paire ou impaire.

**Exemple 2.** Dans l'espace à quatre dimension, que vaut  $S = E \wedge H$  où  $E = E_x dx + E_y dy + E_z dz$  et  $H = H_x dxdt + H_y dydt + H_z dzdt$ ? En respectant notre algèbre, nous trouvons

$$S = (E_x H_y - E_y H_x) dx dy dt + \dots$$

Cette forme est plus connu comme le vecteur de Poynting.

## 9.4. La dérivation extérieure et l'intégration.

Soit l'espace à  $n$ -dimension muni des coordonnées  $x^i$ . Soit une  $k$ -forme différentielle  $\omega$ . Comment définir  $d\omega$ ? Nous généralisons le concept de dérivée des fonctions. Reprenons le cas d'une fonction (0-forme)  $f(\mathbf{x})$  et de sa différentielle (1-forme)  $df$  au point  $\mathbf{x}$ . Par définition,

$$df(\vec{t}) = \lim_{\epsilon \rightarrow 0} \frac{f(\mathbf{x} + \epsilon \vec{t}) - f(\mathbf{x})}{\epsilon}$$

Pour caractériser la différentielle  $df$  dans une base, il suffit de l'appliquer au  $n$  vecteurs  $\vec{e}_i$  de la base au point  $\mathbf{x}$ , noter ces valeurs (la tradition les nomme  $\partial f / \partial x_i$  ou plus simplement  $\partial_i f$ ) et finalement donner l'expression de  $df$  dans la base des  $dx^i$ .

$$df = (\partial_i f) dx^i$$

Nous pouvons très naturellement étendre cette définition aux  $k$ -formes.

**Exemple 1. Le rotationnel.** Prenons une 1-formes dans l'espace à trois dimension que nous écrivons

$$\omega = f_1(x, y, z) dx + f_2(x, y, z) dy + f_3(x, y, z) dz$$

Pour sa dérivée, nous avons

$$d\omega = d(f_1(x, y, z)) dx + d(f_2(x, y, z)) dy + d(f_3(x, y, z)) dz$$

En nous souvenant que  $dx dx = 0$  et que  $dx dy = -dy dx$ , nous pouvons écrire le résultat comme

$$d\omega = (\partial_x f_2 - \partial_y f_1) dx dy + (\partial_y f_3 - \partial_z f_2) dy dz + (\partial_z f_1 - \partial_x f_3) dz dx$$

Ceci n'est rien d'autre que ce que nous appelons le rotationnel.

**Exemple 1. La divergence.** Prenons une 2-forme dans l'espace à trois dimensions que nous écrivons

$$\omega = f_1(x, y, z)dx dy + f_2(x, y, z)dy dz + f_3(x, y, z)dz dx$$

un peu de gymnastique avec les permutations nous montre que

$$d\omega = (\partial_x f_1 + \partial_x f_2 + \partial_z f_3) dx dy dz$$

qui n'est autre que la divergence.

Nous voyons donc que les trois principaux opérateurs, grad, rot, div ne sont que des dérivées des  $k$ -formes.

**Lemme de Poincaré (exercice fondamental).** Il ne faut pas beaucoup d'effort, et on le laisse comme exercice au lecteur, de démontrer le lemme de Poincaré sur les dérivées secondes :

$$d(d\omega) = 0$$

**théorème de Stokes.** La généralisation et unification de concepts disparates à l'aide des formes différentielles va beaucoup plus loin. Nous savons (cela s'appelle le théorème fondamental de l'analyse) que si  $f(x) = F'(x)$ , alors

$$\int_a^b f(x)dx = F(b) - F(a)$$

C'est à dire que l'intégrale de la dérivée d'une 0-forme sur un domaine est égale au flux de la 0-forme à travers les frontières du domaine. Ce théorème se généralise très naturellement aux formes différentielles :

$$\int_D d\omega = \int_{\partial D} \omega$$

où  $\partial D$  désigne la frontière du domaine  $D$ . Cela nous donne par exemple pour les 1-formes

$$\int_{\Sigma} d\omega = \oint_C \omega$$

plus connu comme "la circulation d'un vecteur le long d'une courbe fermée est égale au flux du rotationnel à travers cette surface". Pour les 2-formes, nous obtenons une formulation du genre "le flux d'un champ de vecteur à travers une surface est égale à l'intégrale de sa divergence dans le volume". Nous voyons que toutes ces formulations sont des applications particulières du théorème de Stokes pour les formes.

## 9.5. Le déterminant et l'opérateur de Hodge.

### 9.5.1. Volume et déterminant.

Toutes les opérations que nous avons effectuées jusque là étaient purement "affine". Cependant, nous pouvons utiliser les  $k$ -formes pour effectuer des *mesures géométriques*, comme calculer la surface ou le volume engendré par deux ou trois vecteurs. Si l'on parle de surface et de volume, c'est que nous nous sommes donnée une géométrie, ou plus précisément une forme bilinéaire que nous considérons comme invariant et qui nous donne un produit scalaire. Une forme bi-linéaire  $g$  (non-dégénéré) que nous considérons comme "invariant géométrique" nous permet de définir un repère orthonormé, c'est à dire un ensemble de vecteurs  $\vec{e}_1, \dots, \vec{e}_n$  tel que

$$g(\vec{e}_i, \vec{e}_j) = \pm 1 \delta_{ij}$$

En nous donnant un repère orthonormé, nous pouvons nous définir les surfaces, volumes, ... unités générés par les vecteurs unités de la base. Par exemple, le *volume* engendré par  $n$  vecteurs unités dans un espace à  $n$  dimensions est :

$$Vol(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n) = 1$$

Nous voyons que le volume est une  $n$ -forme. Si maintenant, au lieu de mesurer le volume généré par les  $n$  vecteurs unitaires, nous mesurons le volume généré par  $n$  vecteurs quelconque  $\vec{v}_1, \dots, \vec{v}_n$ , nous pouvons, en utilisant les propriétés de linéarité des formes, calculer ce scalaire. Décomposons ces vecteurs dans notre base :

$$\vec{v}_i = a_i^j \vec{e}_j$$

alors, un peu de calcul nous montre que

$$Vol(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n) = \det(a_i^j) Vol(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$$

où *det* désigne le déterminant. D'ailleurs, ceci est la définition géométrique du déterminant : le volume généré par  $n$  vecteurs. Le déterminant s'introduit naturellement ici comme une forme alternées parce que les  $n$ -formes sont antisymétriques.

Comme nous nous sommes fixées un repère orthonormé, nous pouvons l'utiliser pour nous donner également une base dans l'espace des 1-forme comme d'habitude :

$$dx^i(\vec{e}_j) = \delta_j^i$$

## 9. La formulation de la relativité à l'aide des formes différentielles.

Et donc définir le  $n$ -forme volume directement en terme de ces 1-formes :

$$Vol = dx^1 dx^2 \dots dx^n$$

Noter que la  $n$ -forme  $Vol$  est un objet géométrique qui mesure les volumes, l'expression ci-dessus donne son expression dans une base particulière.

De façon générale, disposer de la forme bi-linéaire  $g$  pour les vecteurs nous permet également de définir un produit scalaire pour les formes, directement hérités des vecteurs. Pour les 1-formes, nous avons

$$\langle dx^i | dx^j \rangle = g(dx^i, dx^j) = g(\vec{e}_i, \vec{e}_j)$$

Le produit scalaire des  $k$ -formes s'hérite de celui des 1-forme, qui au vu de l'exigence d'anti-symétrie nous induit :

$$\langle u^1 \wedge u^2 \wedge \dots \wedge u^k | v^1 \wedge v^2 \wedge \dots \wedge v^k \rangle = \det (\langle u^i | v^j \rangle)$$

Par exemple, d'après cette définition,

$$\langle Vol | Vol \rangle = (-1)^q$$

où  $q$  est la signature de l'espace (le nombre de signe  $-$  dans la matrice diagonale de  $g$ )

### 9.5.2. Hodge et norme.

Il ne vous a pas échappé que les  $k$ -formes et les  $(n - k)$ -formes sont assez similaire : les deux occupent des espaces vectoriels de même dimension  $n!/k!(n - k)!$ . On pourrait donc associer de façon unique un élément de l'un à un élément de l'autre, et nous pouvons utiliser le métrique  $g$  pour effectuer cela. Voilà la procédure : au  $k$ -forme  $\tilde{\omega}_1$  nous associons le  $(n - k)$ -forme  $*\tilde{\omega}_1$  tel que

$$\tilde{\omega}_1 \wedge (*\tilde{\omega}_1) = \langle \tilde{\omega}_1 | \tilde{\omega}_1 \rangle Vol$$

(exercice : démontrer que cette association est unique). L'opérateur  $*$  est appelé l'opérateur de Hodge.

**Exemple 1.** Soit donné l'invariant de Lorentz dans l'espace à 4-dimensions. Nous avons

$$\begin{aligned} (*dx_0) &= dx_1 dx_2 dx_3 & ; & \quad (*dx_1) = -dx_0 dx_2 dx_3 \\ (*dx_2) &= dx_0 dx_1 dx_3 & ; & \quad (*dx_3) = -dx_0 dx_1 dx_2 \end{aligned}$$

**Exemple 2.** Nous pouvons faire la même chose pour les deux formes :

$$\begin{aligned} *(dx_0 dx_1) &= dx_2 dx_3 & ; & \quad * (dx_0 dx_2) = -dx_1 dx_3 \\ *(dx_0 dx_3) &= dx_1 dx_2 & ; & \quad * (dx_1 dx_2) = dx_0 dx_3 \\ *(dx_1 dx_3) &= -dx_0 dx_2 & ; & \quad * (dx_2 dx_3) = dx_0 dx_1 \end{aligned}$$

L'exercice est assez élémentaire, il suffit de calculer, une fois le produit extérieur formé, le nombre de permutation qui nous donne la  $n$ -forme  $Vol$ . L'opération n'est pas différent de "montée-descente" des indices pour les tenseurs.

Pour utiliser les  $k$ -formes dans l'espace physique, nous devons incorporer les invariants géométrique. Nous avons fait ce travail lors du traitement tensoriel, et nous avons vu comment associer des vecteurs et des formes linéaires en utilisant la pseudo métrique, et comment passer des covariants aux contravariants. C'est un travail similaire que nous devons effectuer ici.

## 9.6. Le lagrangien et les équations d'euler-lagrange.

Introduire

$$\mathcal{L} = A \wedge J + (1/2)F \wedge (*F)$$

où  $F = dA$ . en déduire les équations de Maxwell

$$\begin{aligned} dF &= 0 \\ d(*F) &= J \end{aligned}$$

La première relation est évidente puisque  $dF = d^2A$  (qu'on démontre en général en faisant joujou avec les élément du tenseur  $F$ ) et la deuxième est juste l'Euler-Lagrange de base. A la rigueur, écrire le lagrangien, pour être encore plus claire, sous la forme de

$$\mathcal{L} = A \wedge J + (1/2)dA \wedge (*dA)$$

Faire une variation  $A \rightarrow A + \alpha$ , effectuer une dérivation par partie pour dériver E-L dans le contexte des formes différentielle. Profiter pour établir la relation entre  $E, D$  et  $B, H$  comme les conjugués de Hodge ; Eventuellement, mentionner que ces expressions sont valide dans l'espace courbe, pour peu qu'on se donne les bonnes connexions affines.

# 10. Mécanique quantique relativiste : l'équation de Dirac.

10.1. Equation de Klein-Gordon.

10.2. Equation de Dirac.

10.3. Couplage au champ.

10.4. Electrodynamique quantique.

Un peu hors sujet peut-être.

# A. Les dérivées totales et partielles.

considérons une fonction d'une seule variable  $f(x)$ . Si au point  $x$ , nous effectuons un déplacement  $dx$ , nous pouvons approximer la variation de la fonction, à l'ordre 1, par

$$df = \frac{df}{dx} dx$$

où  $df/dx$  est la dérivée de la fonction au point  $x$ . Si nous avons plus qu'une variable par exemple 2, nous pouvons effectuer des déplacements dans plusieurs directions. Prenons le cas à deux variables  $f(x, y)$ . Nous pouvons alors approximer la variation de la fonction, proche de ce point, par

$$df = \left( \frac{\partial f}{\partial x} \right) dx + \left( \frac{\partial f}{\partial y} \right) dy \quad (\text{A.1})$$

La quantité  $\partial f / \partial x$  est le coefficient de proportionnalité entre la variation de la fonction  $f$  et le déplacement  $x$  si le déplacement dans la direction  $y$  est nul. Quand nous avons un déplacement quelconque  $(dx, dy)$  autour du point  $(x, y)$ , il suffit de combiner linéairement les deux variations, ce qui est fait dans l'expression (A.1).

Bien, maintenant répondons à la question suivante : de combien varie la fonction  $f$  autour du point  $(x, y)$  si je me déplace, non pas de façon quelconque, mais le long d'une courbe donnée  $\mathcal{C}$ ? Supposons que l'équation de la courbe soit donnée par  $y = y(x)$ . Dans ce cas, si j'effectue un déplacement  $dx$  dans la direction  $x$ , pour rester sur la courbe, je dois effectuer un déplacement  $dy = y'(x)dx$ . Nous pouvons donc calculer la variation de la fonction, quand la coordonnée  $x$  varie de  $dx$  :

$$df = \left( \frac{\partial f}{\partial x} \right) dx + \left( \frac{\partial f}{\partial y} \right) y'(x) dx$$

où encore

$$\frac{df}{dx} = \left( \frac{\partial f}{\partial x} \right) + \left( \frac{\partial f}{\partial y} \right) y'(x)$$

*A. Les dérivées totales et partielles.*

On peut encore dire cela autrement : en imposant la contrainte de rester sur la courbe  $\mathcal{C}$ , la fonction  $f$  devient une fonction d'une seule variable

$$f = f(x, y(x))$$

et donc nous pouvons à nouveau parler de sa dérivée totale  $df/dx$ .

## B. Un peu plus de rigueur : géométrie affine.

Nous avons utilisé les concepts de vecteurs, de parallèle, ... en géométrie affine, le lecteur les a beaucoup déjà vu au cours de sa vie. Mais il est temps de revenir sur ces notions et les définir très précisément.

Les vecteurs comme des groupes de transformation.

Parallèle comme une relation d'équivalence.

Les formes linéaires ;

Les formes quadratiques, distance et pseudo-distance.

## C. L'expérience de Michelson-Morley.