Coordonnées, variétés et espaces courbes

David Augier

Les notations en termes de coordonnées covariantes et contravariantes, c'est-à-dire indices en bas et en haut, interviennent de manière systématique dans l'écriture des lois physiques, notamment dans les théories incluant des considérations relativistes. En général, le principe de ces notations est souvent vu comme un détail, un artifice technique. Ces notes donnent quelques éclaircissements sur les concepts mis en jeu, d'un point de vue à la fois physique et mathématique : coordonnées covariantes et contravariantes, tenseur métrique, variétés, coordonnées curvilignes, géodésiques. Le niveau global du contenu est à partir de L2/deuxième année de classes préparatoires/L3. Le contenu mathématique concerne surtout des notions de base d'algèbre linéaire, bilinéaire et dualité. Avoir quelques idées sur la relativité permet de mieux aborder certaines remarques mais ce n'est pas indispensable. Le tout reste en évolution et pourra peut-être aboutir sur des applications de relativité générale, abordées de manière pragmatique sans rentrer dans les fondements de la théorie.

24 mai 2012

Table des matières

Chapitre I

Coordonnées covariantes et contravariantes

La détermination des coordonnées d'un vecteur dans un espace vectoriel (de dimension finie) se fait à partir du choix d'une base. Lorsque l'espace est muni d'un produit scalaire, il est possible d'utiliser un autre type de coordonnées. L'utilisation conjointe des deux types de coordonnées se trouve être très utile en relativité. Ce chapitre s'intéresse à l'établissement de ces coordonnées, au passage d'un type à l'autre, ainsi qu'à la généralisation des notions de vecteurs et formes linéaires, les tenseurs.

1 Coordonnées d'un vecteur

1.1 Coordonnées d'un vecteur dans une base

Soit E un espace vectoriel de dimension finie n. On appelle *base* de E une famille $(\vec{e}_1, \ldots, \vec{e}_n)$ telle que tout élément $\vec{v} \in E$ s'écrit de manière unique

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

Les (v^i) sont les *coordonnées* de $\vec{v}$ dans la base $(\vec{e}_i)$. Afin d'alléger les expressions, on omettra désormais le symbole Σ, les indices répétés en haut et en bas étant implicitement sommés (c'est la notation d'Einstein). Par convention, les éléments de la base $(\vec{e}_i)$ de E possèdent un indice en **bas** alors que les coordonnées (v^i) du vecteur $\vec{v}$ ont un indice en **haut** [1]. Un exemple est donné en dimension $n = 2$ sur la figure 1.

1. On reviendra sur cette convention très pratique plus tard.

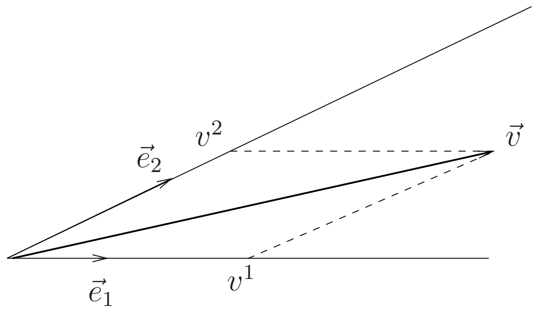

Figure 1 : coordonnées (v^i) d'un vecteur $\vec{v}$ en dimension 2

1.2 Utilisation d'un produit scalaire

Un autre jeu de coordonnées

Si l'espace E est muni d'un produit scalaire noté $(.,.)$ ou plus simplement ".", il existe une autre possibilité naturelle pour caractériser le vecteur $\vec{v}$: on pose

$$v_i = \vec{v}.\vec{e}_i \tag{1.1}$$

(bien noter cette fois que les indices sont en **bas** pour les (v_i)). Les (v_i) constituent un autre type de coordonnées pour le vecteur $\vec{v}$, c'est-à-dire que la donnée des (v_i) renseigne sans ambiguïté sur $\vec{v}$ (ce résultat sera établi dans les deux paragraphes suivants). La figure 2 montre clairement la différence entre les (v^i) et les (v_i). En revanche, si la base $(\vec{e}_i)$ est orthonormée, il n'existe aucune différence entre ces deux jeux de coordonnées.

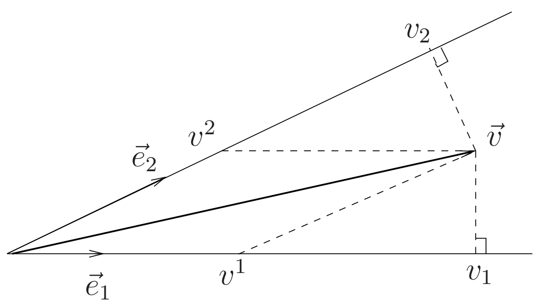

Figure 2 : coordonnées (v^i) et (v_i) d'un vecteur $\vec{v}$ en dimension 2, avec le produit scalaire usuel

Base associée : approche simple

On peut alors se demander quelle est la base correspondant aux coordonnées (v_i), c'est-à-dire une base $(\vec{\varepsilon}^i)$ telle que

$$\vec{v} = v_i \vec{\varepsilon}^i$$

Pour que $\vec{v}.\vec{e}_1$ soit égal à v_1, le plus simple serait que tous les $\vec{e}_i$ $(i \neq 1)$ vérifient $\vec{e}_i.\vec{\varepsilon}^1 = 0$. Ainsi, $\vec{\varepsilon}^1$ doit être orthogonal à tous les $\vec{e}_i$ $(i \neq 2)$ et normé de telle manière que $\vec{e}_1.\vec{\varepsilon}^1 = 1$. On procède de même avec tous les $\vec{\varepsilon}^i$. La situation assez simple est représentée en figure 3, mais ne permet pas de comprendre le fondement de la situation.

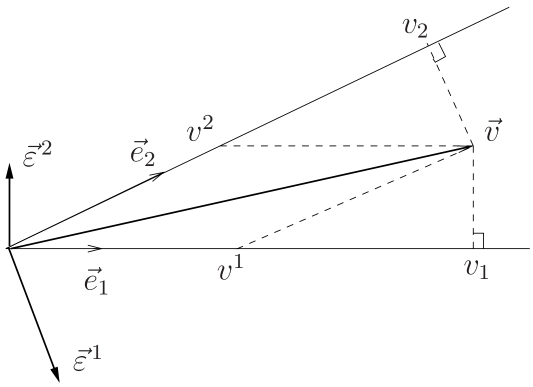

Figure 3 : coordonnées (v^i) et (v_i) d'un vecteur $\vec{v}$ en dimension 2, avec le produit scalaire usuel

Base associée : approche plus abstraite

Il faut alors se placer dans le dual E^* de E, c'est-à-dire l'espace des formes linéaires sur E. On rappelle rapidement qu'une forme linéaire φ sur un espace vectoriel E est une application linéaire de E sur son corps de base ($\mathbb{R}$ pour ce qui nous intéressera ici). On considère alors la base duale de $(\vec{e}_i)$, notée (ε^i), définie de manière unique par

$$\varepsilon^i(\vec{e}_j) = \delta^i_j$$

où δ^i_j appelé *symbole de Kronecker* vaut 1 si $i = j$ et 0 sinon. Il est assez facile de constater que $(\varepsilon^1, \ldots, \varepsilon^n)$ est une base de E^* [2]. Ainsi, si E est de dimension finie n, il en est de même pour E^*.
L'action de l'élément de la base duale ε^i sur le vecteur $\vec{v} = v^j\vec{e}_j$ est très simple :

$$\varepsilon^i(\vec{v}) = \varepsilon^i(v^j\vec{e}_j) = v^j\varepsilon^i(\vec{e}_j) = v^j\delta^i_j = v^i$$

ε^i agit sur le vecteur $\vec{v}$ en donnant sa coordonnée v^i dans la base antéduale $(\vec{e}_i)$. Dans la suite, la forme linéaire ε^i sera plutôt notée dx^i, et

$$\boxed{dx^i(\vec{v}) = v^i}$$

2. En effet, connaître une forme linéaire φ sur E est équivalent à connaître les images des vecteurs de la base $(\vec{e}_i)$. En considérant

$$\psi = \sum_{i=1}^{n} \varphi(\vec{e}_i)\varepsilon^i,$$

on vérifie que les formes linéaires φ et ψ coïncident sur la base $(\vec{e}_i)$, donc sur E par linéarité. Elles sont donc égales.

À un niveau fondamental, l'existence d'un produit scalaire $(,)$ sur E permet d'associer canoniquement (c'est-à-dire sans se référer à une base) à tout vecteur $\vec{v}$ une forme linéaire $\tilde{v}$ (et réciproquement) par

$$\tilde{v} = (\vec{v}, .) \quad \text{ou plus explicitement} \quad \tilde{v}(\vec{x}) = (\vec{v}, \vec{x}) \text{ pour tout } \vec{x} \in E$$

Cette correspondance est un isomorphisme[3] : il est totalement équivalent de parler d'un vecteur ou de sa forme linéaire associée.
Du coup, plutôt que de considérer les coordonnées (v^i) du vecteur $\vec{v}$, il est équivalent de s'intéresser aux coordonnées de la forme linéaire $\tilde{v}$. En fait, ces coordonnées sont justement les (v_i) dans la base duale. Montrons ce résultat. Comme (ε^i) est une base de E^*, $\tilde{v}$ s'écrit forcément

$$\tilde{v} = \alpha_i \varepsilon^i$$

Le tout est de montrer que $\alpha_i = v_i = \vec{v}.\vec{e}_i$ (relation (1.1)). D'une part,

$$\tilde{v}(\vec{e}_i) = \alpha_j \varepsilon^j(\vec{e}_i) = \alpha_j \delta_i^j = \alpha_i$$

D'autre part,

$$\tilde{v}(\vec{e}_i) = (\vec{v}, \vec{e}_i) = v_i$$

ce qui établit le résultat.

Au final, pour décrire un vecteur $\vec{v}$, on peut donner ses coordonnées (v^i) dans une base $(\vec{e}_i)$ ou de manière équivalente les coordonnées (v_i) de $\tilde{v}$ dans la base duale (ε^i).

1.3 Tenseur métrique

Descente des indices

Il s'agit essentiellement de rendre plus systématique le passage entre les coordonnées (v^i) et (v_i) d'un vecteur $\vec{v}$. Par définition (relation (1.1)),

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

Au final,

$$\boxed{v_i = g_{ij} v^j} \tag{1.2}$$

en notant

$$\boxed{g_{ij} = (\vec{e}_i, \vec{e}_j) = g_{ji}}$$

g est nommé *tenseur métrique*. D'un point de vue mathématique, g représente la matrice associée au produit scalaire. Pour la suite, on notera particulièrement son usage pratique en terme d'ascenseur à indices. Notamment, dans la relation (1.2), il y a une sommation sur l'indice j : c'est la notation d'Einstein que l'on a déjà évoquée. L'indice i est en bas, à droite et à gauche de l'égalité. Il faut tout le temps être attentif et vérifier cette homogénéité de hauteur entre chaque terme d'une égalité.

3. En effet, l'application associée $\tilde{}$: $E \to E^*$ est linéaire, de noyau réduit à $\{0\}$.

Tenseur métrique

Revenons un peu sur la nature du tenseur métrique. g prend en argument deux vecteurs et donne en résultat un nombre. Il s'agit d'une forme bilinéaire ou de manière équivalente d'un élément de $E^* \otimes E^*$ (le $\otimes$ désigne un produit tensoriel [4]). Formellement,

$$g = g_{ij}\varepsilon^i \otimes \varepsilon^j = g_{ij}dx^i \otimes dx^j = g_{ij}dx^i dx^j \tag{1.3}$$

(le $\otimes$ est souvent omis). En effet,

$$g(\vec{u}, \vec{v}) = g_{ij}dx^i dx^j(\vec{u}, \vec{v}) = g_{ij}dx^i(\vec{u})dx^j(\vec{v}) = g_{ij}v^i w^j = (\vec{u}, \vec{v})$$

Le sens physique donné à la relation (1.3) est de déterminer le carré de la "longueur" entre deux points dont les coordonnées diffèrent de dx^i. Cette longueur sera plutôt notée ds^2 et

$$\boxed{ds^2 = g_{ij}dx^i dx^j} \tag{1.4}$$

Par exemple, dans l'espace $\mathbb{R}^n$ muni d'une base orthonormée $g_{ij} = \delta_{ij} = \begin{pmatrix} 1 & & (0) \\ & \ddots & \\ (0) & & 1 \end{pmatrix}$.

Montée des indices

On a vu précédemment que descendre les indices (c'est-à-dire passer des coordonnées du vecteur à celles de la forme linéaire associée) se faisait grâce au tenseur métrique g. L'opération inverse est réalisé par l'inverse de g : on pose

$$\boxed{g^{ij} = \text{élément } (ij) \text{ de } g^{-1}} \qquad \text{d'où} \qquad \boxed{g^{ij}g_{jk} = \delta^i_k}$$

Il faudra être attentif au fait que g avec les indices en haut correspond en fait à l'inverse du tenseur métrique. L'égalité de droite n'est rien d'autre que l'écriture en coordonnées de $g^{-1}g = Id$ (rappelons que l'élément (ik) du produit matriciel AB s'écrit : $(AB)_{ik} = \sum_j A_{ij}B_{jk}$.

g^{-1} va alors naturellement servir à faire monter les indices. On repart de la relation (1.2) :

$$v_i = g_{ij}v^j$$

4. Pour faire très simple, le produit tensoriel de deux espaces E et F, noté $E \otimes F$, est l'espace vectoriel dont une base est donnée par les $\vec{e}_i \otimes \vec{f}_j$, où $(\vec{e}_i)$ est une base de E et $(\vec{f}_j)$ une base de F. On peut en donner un exemple physique. Si E repère les états d'une particule et F ceux d'une autre particule, $E \otimes F$ repère les états des deux particules, c'est-à-dire qu'il faut se donner à la fois un vecteur de E et un de F, ou plus exactement une combinaison linéaire de ce genre de termes pour décrire l'état des deux particules. Mathématiquement, $E \otimes F$ est introduit de manière à pouvoir considérer une application bilinéaire sur $E \times F$ comme une application linéaire sur $E \otimes F$.

On multiplie de chaque côté par g^{ki} :

$$g^{ki}v_i = g^{ki}g_{ij}v^j = \delta^k_j v^j = v^k$$

d'où en changeant le nom des indices

$$v^i = g^{ij}v_j \tag{1.5}$$

Différentes écritures du produit scalaire

Le produit scalaire de deux vecteurs peut alors aisément être calculé dans la base $(\vec{e}_i)$. Si $\vec{v} = v^i\vec{e}_i$ et $\vec{w} = w^j\vec{e}_j$,

$$(\vec{v}, \vec{w}) = (v^i\vec{e}_i, w^j\vec{e}_j) = v^i w^j(\vec{e}_i, \vec{e}_j) = g_{ij}v^i w^j = g(\vec{v}, \vec{w})$$

On utilisera beaucoup par la suite ce genre de calcul, dit *en coordonnées*, qui se prête aisément à des calculs systématiques.

Une autre possibilité pour exprimer le produit scalaire de deux vecteurs est de considérer l'action de $\tilde{v}$ sur $\vec{w}$, ou réciproquement de $\tilde{w}$ sur $\vec{v}$:

$$(\vec{v}, \vec{w}) = \tilde{v}(\vec{w}) = \tilde{w}(\vec{v})$$

Comme $\tilde{v} = v_i\varepsilon^i$ et $\tilde{w} = w_i\varepsilon^i$, on déduit

$$(\vec{v}, \vec{w}) = v_i\varepsilon^i(w^j\vec{e}_j) = v_i w^j\delta^i_j = v_i w^i = v^i w_i$$

On retiendra que le produit scalaire de deux vecteurs s'écrit indifféremment

$$\vec{v}.\vec{w} = g_{ij}v^i w^j = v_i w^i = v^i w_i \tag{1.6}$$

Le tout est de bien faire attention aux différentes hauteurs d'indice.

Cas de la relativité restreinte

La relation caractéristique de la géométrie euclidienne (en coordonnées cartésiennes orthonormées)

$$ds^2 = dx^2 + dy^2 + dz^2$$

est remplacée en relativité restreinte par [5]

$$ds^2 = dt^2 - dx^2 - dy^2 - dz^2 = \eta_{\mu\nu}dx^\mu dx^\nu$$

5. Conformément aux usages de relativité, on pose $c = 1$. Il ne faudra évidemment pas oublier de rétablir l'homogénéité pour les applications numériques...

C'est ce que l'on appelle la *métrique minkowskienne*. La différence notable avec une métrique euclidienne est la présence de signes différents sur la diagonale [6] (pour une métrique euclidienne, il est toujours possible d'écrire g comme l'identité en se plaçant dans une base orthonormée). g ne correspond plus à un produit scalaire mais à une forme bilinéaire symétrique définie (le caractère positif a disparu). En fait, cela ne change rien à tout ce qui précède ! Le point important lors de l'utilisation du produit scalaire était l'identification canonique entre E et E^*. Cette identification peut tout aussi bien se faire à l'aide d'une forme bilinéaire définie. Le tout est que le déterminant de g doit être non nul. Par abus de langage, on continue à considérer g comme un produit scalaire dans le cas minkowskien.

Les vecteurs appartenant à l'espace minkowskien de dimension 4 sont appelés *quadrivecteurs*. On va en donner deux exemples élémentaires :

- quadrivecteur position d'un événement : c'est la donnée des quatre coordonnées d'un point. En coordonnées covariantes [7], on obtient naturellement

$$\underset{\sim}{x} = (x^0, x^1, x^2, x^3) = (t, \vec{r}) = (t, x, y, z)$$

  Le quadrivecteur $\underset{\sim}{x}$ [8] peut aussi s'exprimer sous forme de coordonnées covariantes $x_0 = g_{0\mu} x^\mu = x^0 = t$ et $x_i = \eta_{i\mu} x^\mu = -x^i$ pour $i = 1, 2, 3$. Au final,

$$\underset{\sim}{x} = (x_0, x_1, x_2, x_3) = (t, -\vec{r}) = (t, -x, -y, -z)$$

- quadrivecteur *énergie-impulsion* d'une particule : sous forme de coordonnées contravariantes [9], le quadrivecteur s'écrit

$$\underset{\sim}{p} = (p^0, p^1, p^2, p^3) = (E, \vec{p}) = (E, p_x, p_y, p_z)$$

6. Il est à noter qu'une autre convention possible est d'écrire

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2$$

(aucune convention ne s'impose particulièrement. . .).

7. On a bien ici les coordonnées contravariantes car $\underset{\sim}{x} = x\vec{e}_x + y\vec{e}_y + z\vec{e}_z + t\vec{e}_t$, où $\vec{e}_t$ est un vecteur unitaire représentant le temps dans un espace-temps à quatre dimensions.

8. La petite vague sous le x indique la nature quadrivectorielle de la quantité.

9. Le quadrivecteur $\underset{\sim}{p}$ se définit en fait comme $\underset{\sim}{p} = m\frac{d\underset{\sim}{x}}{d\tau}$, où $d\tau$ est l'intervalle de temps propre de la particule entre deux événements, et $d\underset{\sim}{x}$ la variation du quadrivecteur position correspondante. Vu les coordonnées contravariantes de la position, on en déduit celles de l'énergie-impulsion. Ainsi, $p^0 = m\frac{dt}{d\tau} = \gamma m$, et $p^1 = m\frac{dx}{d\tau} = \gamma m\frac{dx}{dt} = p_x$.

où E est l'énergie de la particule et $\vec{p} = \gamma m \vec{v}$ son impulsion. En considérant les coordonnées covariantes, on obtient de même que précédemment

$$\underset{\sim}{p} = (p_0, p_1, p_2, p_3) = (E, -\vec{p}) = (E, -p_x, -p_y, -p_z)$$

La "norme" de $\underset{\sim}{p}$ s'écrit alors

$$\underset{\sim}{p}^2 = \underset{\sim}{p}.\underset{\sim}{p} = \eta_{\mu\nu} p^\mu p^\nu = p_\mu p^\mu = E^2 - \vec{p}^2$$

En anticipant légèrement ce qui suit (voir page 16), $\underset{\sim}{p}.\underset{\sim}{p}$ ne dépend pas du référentiel inertiel choisi (cela revient à dire que le produit scalaire de deux vecteurs est indépendant de la base choisie), donc a la valeur obtenue dans le référentiel dans lequel la particule est au repos. Alors, $E' = mc^2 = m$ (énergie au repos), et $\vec{p}' = \vec{0}$, d'où on déduit

$$\underset{\sim}{p}.\underset{\sim}{p} = E^2 - \vec{p}^2 = m^2$$

1.4 Changement de base, coordonnées covariantes et contravariantes

Imaginons qu'on choisisse une nouvelle base de E, notée $(\vec{e}_{1'}, \ldots, \vec{e}_{n'})$ [10]. Ce choix est tout aussi pertinent que celui de la base $(\vec{e}_1, \ldots, \vec{e}_n)$. Le vecteur $\vec{v}$ a une existence intrinsèque, c'est-à-dire indépendante du choix de la base. Par léger abus de langage, on le dit invariant. En revanche, les coordonnées dépendent du choix de la base. Plus précisément, les deux types de coordonnées de $\vec{v}$ présentent des comportements différents. Les coordonnées v_i vérifiant

$$v_i = \underbrace{\vec{v}}_{\text{invariant}} .\vec{e}_i$$

vont varier "en même temps" que les vecteurs de la base. Les (v_i) sont appelées *coordonnées covariantes*. En revanche, les (v^i) sont telles que

$$\underbrace{\vec{v}}_{\text{invariant}} = v^i \vec{e}_i$$

et "compensent" la variation des vecteurs de la base. Les (v^i) sont appelées *coordonnées contravariantes*.

Précisons les expressions des changements de coordonnées. Les nouvelles coordonnées contravariantes s'écrivent en fonction des anciennes

$$X' = \Lambda X \quad \text{où} \quad \Lambda = P^{-1}$$

10. La nouvelle base s'écrit formellement $(\vec{e}'_1, \ldots, \vec{e}'_n)$ mais la convention choisie ici de faire porter les primes par les indices permet d'écrire les différentes relations de changement de base de manière naturelle. On passera indifféremment d'une notation à l'autre par la suite.

en notant P la matrice de changement de base de $(\vec{e}_1, \ldots, \vec{e}_n)$ à $(\vec{e}_{1'}, \ldots, \vec{e}_{n'})$, X le vecteur colonne des anciennes coordonnées et X' celui des nouvelles. Avec les indices, cela donne

$$\boxed{v^{i'} = \Lambda^{i'}{}_j v^j} \tag{1.7}$$

Ensuite, comme

$$\vec{v} = v^i \vec{e}_i = v^{j'} \vec{e}_{j'} = \Lambda^{j'}{}_k v^k \vec{e}_{j'} = v^k \Lambda^{j'}{}_k \vec{e}_{j'}$$

on déduit

$$\vec{e}_i = \Lambda^{j'}{}_i \vec{e}_{j'} \tag{1.8}$$

On peut alors déduire la loi de changement de base pour les coordonnées covariantes :

$$v_i = \vec{v}.\vec{e}_i = \vec{v}.(\Lambda^{j'}{}_i \vec{e}_{j'}) = \Lambda^{j'}{}_i \vec{v}.\vec{e}_{j'}$$

d'où

$$\boxed{v_i = \Lambda^{j'}{}_i v_{j'}} \tag{1.9}$$

On a obtenu des lois de changement de base pour les coordonnées covariantes et contravariantes. Passons au changement de base inverse. Cela nécessite l'utilisation de l'inverse de la matrice Λ, en posant pour harmoniser les notations

$$\Lambda_{i'}{}^j = \left(\Lambda^{-1}\right)^j{}_{i'} \quad \text{avec} \quad \Lambda_{i'}{}^j \Lambda^{i'}{}_k = \delta^j_k \tag{1.10}$$

Il faut bien faire attention à la position des indices : si l'indice haut est à gauche, il s'agit de la matrice Λ et de la matrice Λ^{-1} dans le cas inverse. On compose alors la relation (1.7) par Λ^{-1} :

$$\Lambda_{i'}{}^k v^{i'} = \Lambda_{i'}{}^k \Lambda^{i'}{}_j v^j = \delta^k_j v^j = v^k$$

d'où

$$\boxed{v^i = \Lambda_{j'}{}^i v^{j'}} \tag{1.11}$$

De même que précédemment, on déduit le changement de base des vecteurs :

$$\vec{e}_{i'} = \Lambda_{i'}{}^j \vec{e}_j \tag{1.12}$$

et celui des coordonnées covariantes

$$\boxed{v_{i'} = \Lambda_{i'}{}^j v_j} \tag{1.13}$$

En confrontant les expressions (1.7), (1.8), (1.9), (1.11), (1.12) et (1.13), on comprend bien la dénomination coordonnées covariantes/contravariantes.

Cas de la relativité restreinte

Ainsi, pour une transformation de Lorentz standard avec une vitesse $\vec{V} = V\vec{u}_x$, en posant $\gamma = 1/\sqrt{1-V^2}$,

$$\begin{cases} t' = \gamma\,(t - Vx) \\ x' = \gamma(x - Vt) \\ y' = y \\ z' = z \end{cases} \qquad \text{d'où} \quad \Lambda^{i'}{}_j = \begin{pmatrix} \gamma & -\gamma V & 0 & 0 \\ -\gamma V & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(comme précédemment, t est considéré comme la coordonnée zéro). Un changement de référentiel inertiel, c'est-à-dire d'observateur intertiel du point de vue physique) s'interprète alors mathématiquement comme un simple changement de base.

2 Tenseurs

2.1 Définition et exemples

On va chercher à généraliser la notion de vecteur/forme linéaire. Rappelons ainsi qu'une forme linéaire est un objet prenant en argument un vecteur et renvoyant de manière linéaire un nombre (par la suite, on nommera les nombres scalaires). Elle est naturellement représentée par des coordonnées covariantes. Symétriquement, un vecteur peut être vu comme un objet prenant en argument une forme linéaire et renvoyant linéairement un scalaire [11]. On peut chercher à généraliser ce genre de comportement. Un endomorphisme (application linéaire de E dans E) peut être vu comme un objet prenant en argument un vecteur et renvoyant de manière linéaire un vecteur. Ainsi, l'endomorphisme M agit sur le vecteur $\vec{v}$ en donnant le vecteur $\vec{w}$ par

$$w^i = M^i{}_j v^j$$

Il s'agit de la notation matricielle usuelle dans la base $(\vec{e}_i)$. On écrit alors l'expression complète :

$$M = M^i{}_j \vec{e}_i \otimes \varepsilon^j$$

Ainsi, un tenseur p fois contravariant et q fois covariant (dit de type (p,q)) prend en argument q vecteurs et fournit de manière linéaire un objet pouvant agir linéairement sur p formes linéaires [12]. Par convention, un tenseur de type $(0,0)$ est un *scalaire*. Un vecteur est un tenseur de type $(1,0)$, une forme linéaire de type $(0,1)$, un endomorphisme de type $(1,1)$, le tenseur métrique de type $(0,2)$… Dans les bases $(\vec{e}_i)$ et (ε^i), un tenseur T de type (p,q) s'écrit

$$T = T^{i_1\ldots i_p}{}_{j_1\ldots j_q} \vec{e}_{i_1} \otimes \ldots \otimes \vec{e}_{i_p} \otimes \varepsilon^{j_1} \otimes \varepsilon^{j_q} \tag{1.14}$$

L'expression en coordonnées donnée par $T^{i_1\ldots i_p}{}_{j_1\ldots j_q}$ d'un tenseur T permet d'effectuer de manière systématique différentes opérations :

11. Par la relation $\vec{v}(\tilde{w}) = \tilde{w}(\vec{v})$.

12. Il est équivalent de dire de ce tenseur prend en argument p formes linéaires et fournit un objet pouvant agir sur q vecteurs.

- lors d'un changement de base, les coordonnées du tenseur T de type (p, q) se transforment selon les lois (1.7) et (1.13) vues précédemment (une matrice Λ par coordonnée) :

$$T^{i'_1 \ldots i'_p}{}_{j'_1 \ldots j'_q} = \Lambda^{i'_1}{}_{i_1} \ldots \Lambda^{i'_p}{}_{i_p} \Lambda_{j'_1}{}^{j_1} \ldots \Lambda_{j'_q}{}^{j_q} T^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} \tag{1.15}$$

- l'action du tenseur T sur les formes linéaires $\tilde{w}_1, \ldots, \tilde{w}_p$ et les vecteurs $\vec{v}_1, \ldots, \vec{v}_q$ s'effectue simplement :

$$T(\tilde{w}_1, \ldots, \tilde{w}_p, \vec{v}_1, \ldots, \vec{v}_q) = T^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} w_{1,i_1} \ldots w_{p,i_p} v_1^{j_1} \ldots v_q^{j_q}$$

2.2 Champs de tenseurs

Jusqu'à maintenant, on n'a considéré que des tenseurs uniformes dans l'espace, c'est-à-dire qui ne dépendent pas du point considéré de l'espace. Par exemple, on s'est intéressé aux coordonnées d'un vecteur $\vec{v}$ donné. Un *champ de tenseurs* correspond à un tenseur prenant des valeurs différentes en chaque point de l'espace. Par exemple, un champ de vecteurs est donné par la vitesse $\vec{v}$ en chaque point d'un écoulement.

Lors d'un changement de base, les coordonnées contravariantes d'un point de l'espace passent de $(x^1, \ldots, x^n)$ à $(x'^1, \ldots, x'^n)$. Ainsi, au vecteur $\vec{v}(x^1, \ldots, x^n)$ correspond le vecteur $\vec{v}\,'(x'^1, \ldots, x'^n)$, avec l'identité

$$\vec{v}(x^1, \ldots, x^n) = \vec{v}\,'(x'^1, \ldots, x'^n)$$

qui exprime le fait que la vitesse de l'écoulement est au final la même quelle que soit la base choisie. Les fonctions exprimant $\vec{v}$ en fonction des x^i et $\vec{v}\,'$ en fonction des x'^i sont alors *a priori* différentes [13].

Ainsi, lors d'un changement de base, l'expression (1.15) se généralise à

$$T^{i'_1 \ldots i'_p}{}_{j'_1 \ldots j'_q}(x') = \Lambda^{i'_1}{}_{i_1} \ldots \Lambda^{i'_p}{}_{i_p} \Lambda_{j'_1}{}^{j_1} \ldots \Lambda_{j'_q}{}^{j_q} T^{i_1 \ldots i_p}{}_{j_1 \ldots j_q}(x) \tag{1.16}$$

où x (x') représente les coordonnées dans la première (seconde) base.

13. Rien ne vaut un exemple. Dans un espace de dimension 1 muni d'un vecteur normé $\vec{e}_x$, on considère le champ de vecteurs $\vec{v} = x^2 \vec{e}_x$. On choisit une autre base $\vec{e}_{x'} = 2\vec{e}_x$. Alors le point de coordonnée contravariante x dans la base $\vec{e}_x$ a pour coordonnée contravariante $x' = x/2$ dans la base $\vec{e}_{x'}$ car $x\vec{e}_x = \frac{x}{2}(2\vec{e}_x) = x'\vec{e}_{x'}$. On en déduit la fonction $\vec{v}\,'$:

$$\vec{v}\,'(x') = \vec{v}(x) = x^2 \vec{e}_x = \frac{x^2}{2}(2\vec{e}_x) = 2x'^2 \vec{e}_{x'}$$

Les fonctions $\vec{v}(x)$ et $\vec{v}\,'(x')$ sont alors différentes malgré le fait qu'elles représentent le même vecteur au même point.

2.3 Montées et descentes d'indices d'un tenseur

De la même manière que le tenseur métrique g permet de transformer un vecteur en forme linéaire, il permet aussi de transformer un tenseur de type (p, q) en tenseur de type (p', q') avec $p' + q' = p + q$. En d'autres termes, il est possible de faire monter ou descendre des indices. Par exemple, le tenseur T de type $(2, 0)$ peut être transformé en un tenseur de type $(1, 1)$ que l'on notera de même T :

$$T^{i}{}_{j} = g_{jk}T^{ik}$$

Le tenseur métrique (son inverse) fait descendre (monter) l'indice voulu. Attention en revanche à bien écrire les indices toujours dans le même ordre : ici, le second indice est descendu. Descendre le premier indice mène au résultat différent

$$T_{i}{}^{j} = g_{ik}T^{kj}$$

Suite à l'identification entre l'espace E et son dual E^{*}, les différents coordonnées du tenseur T (avec un nombre différent d'indice en haut/bas) correspondent effectivement au même tenseur T mais représenté de manière différente.

2.4 Comment engendrer de nouveau tenseurs ?

À partir d'un jeu de tenseurs, il est possible de créer de nouveau tenseurs.

La première possibilité est le produit tensoriel $T \otimes U$ des tenseurs T de type (p, q) et U de type (r, s). $T \otimes U$ est un tenseur de type $(p+r, q+s)$ dont les coordonnées sont simplement les

$$(T \otimes U)^{i_1 \ldots i_p}{}_{j_1 \ldots j_q}{}^{k_1 \ldots k_r}{}_{l_1 \ldots l_s} = T^{i_1 \ldots i_p}{}_{j_1 \ldots j_q} U^{k_1 \ldots k_r}{}_{l_1 \ldots l_s}$$

La seconde possibilité est de faire agir un tenseur T sur un tenseur U, c'est ce qu'on appelle une *contraction*. Par exemple, si T est de type $(2, 0)$ et U de type $(1, 1)$, on utilise le fait que U prend en argument un vecteur et que T donne en "sortie" des vecteurs. On peut ainsi former les nouveaux tenseurs V ou W de type $(2, 0)$ dont les coordonnées sont

$$V^{jk} = T^{ij}U^{k}{}_{i} \quad \text{ou} \quad W^{ik} = T^{ij}U^{k}{}_{j}$$

Le principe est le même avec des tenseurs de types différents. On peut recommencer le processus autant de fois que désiré.

Le produit scalaire entre deux vecteurs apparaît donc comme un cas particulier de contraction entre un tenseur de type $(1, 0)$ et un tenseur de type $(0, 1)$ donnant au final un scalaire (tenseur de type $(0, 0)$) :

$$g(\vec{v}, \vec{w}) = g_{ij}v^{i}w^{j} = v_{i}w^{i}$$

2.5 Cas particulier : les scalaires

D'un point de vue tensoriel, les *scalaires* sont des tenseurs de type $(0, 0)$. Ils ne sont donc pas affectés par un changement de base vu la loi de transformation (1.15). Vérifions le explicitement sur un exemple.

Le produit scalaire de deux vecteurs (expression (1.6)) :

$$(\vec{v}, \vec{w}) = g_{ij} v^i w^j$$

qui est la contraction du tenseur métrique $(0, 2)$ avec deux vecteurs (tenseurs de type $(1, 0)$) est un exemple de scalaire. Évaluée dans une autre base, cette quantité serait égale à $g_{i'j'} v^{i'} w^{j'}$. Les deux quantités ne semblent pas forcément égales. À l'aide des lois de transformation, on récrit

$$g_{i'j'} v^{i'} w^{j'} = \Lambda_{i'}{}^k \Lambda_{j'}{}^l g_{kl} \, \Lambda^{i'}{}_m v^m \Lambda^{j'}{}_n w^n$$

On retrouve ici la matrice Λ et son inverse (expression (1.10)), d'où

$$g_{i'j'} v^{i'} w^{j'} = (\Lambda_{i'}{}^k \Lambda^{i'}{}_m)(\Lambda_{j'}{}^l \Lambda^{j'}{}_n) g_{kl} v^m w^n = \delta^k_m \delta^l_n g_{kl} v^m w^n = g_{kl} v^k w^l$$

Il est donc bien équivalent d'évaluer un produit scalaire en coordonnées dans une base ou une autre.

Cas de la relativité restreinte, dérivée covariante

L'intérêt en relativité restreinte est évident. Le changement de base correspond dans ce cas à un changement de référentiel inertiel, c'est-à-dire d'observateur. Ainsi, tout observateur inertiel mesurera la même valeur, si la quantité observée se comporte comme un scalaire. Ce sera donc le cas pour le produit scalaire de deux quadrivecteurs. Par exemple, le scalaire $\underset{\sim}{p}.\underset{\sim}{p}$ correspondant au carré du quadrivecteur énergie-impulsion d'une particule devra prendre la même valeur pour tous les observateurs, particulièrement pour l'observateur possédant un mouvement tangent à la particule. Pour ce dernier, la particule a une vitesse nulle, et donc $\underset{\sim}{p} = (E, \vec{0})$, avec $E = mc^2 = m$ d'après la formule d'Einstein. On en déduit que pour tous les observateurs, $\underset{\sim}{p}.\underset{\sim}{p} = m^2$.

Continuons un peu la discussion dans le cas d'un scalaire dépendant du point d'observation $S(x^0, \ldots, x^3)$. Un autre observateur mesurera la même quantité en chaque point, mais correspondant aux coordonnées $(x'^0, \ldots, x'^3)$. Vu les coordonnées différentes, ce n'est plus la même fonction des coordonnées, on la notera donc S'. Ainsi,

$$S(x^0, \ldots, x^3) = S'(x'^0, \ldots, x'^3)$$

Entre deux points proches séparés de δx^μ ou $\delta x'^\mu$, les deux observateurs mesureront la même variation du scalaire

$$\delta S = \frac{\partial S}{\partial x^\mu} \delta x^\mu = \delta S' = \frac{\partial S'}{\partial x'^\nu} \delta x'^\nu = \frac{\partial S'}{\partial x^{\nu'}} \delta x^{\nu'}$$

δx^μ est un quadrivecteur, donc

$$\delta x^\mu = \Lambda_{\nu'}{}^\mu \delta x^{\nu'}$$

On déduit

$$\frac{\partial S'}{\partial x^{\nu'}} = \Lambda_{\nu'}{}^\mu \frac{\partial S}{\partial x^\mu}$$

On reconnaît là la loi de transformation d'un tenseur 1 fois covariant. On réécrit cela sous la forme symbolique plus compacte

$$\partial_{x^{\mu'}} = \Lambda_{\mu'}{}^{\nu} \partial_{x^{\nu}}$$

On est alors amené à poser

$$\boxed{(\partial_t, \partial_x, \partial_y, \partial_z) = \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)}$$

C'est ce qu'on appelle la *dérivée covariante*. De manière plus générale, elle peut être vue comme un tenseur de type $(1, 0)$ et par contraction, elle transforme un tenseur de type (p, q) en tenseur de type $(p, q + 1)$. De même, on définit la dérivée contravariante

$$\boxed{(\partial^t, \partial^x, \partial^y, \partial^z) = \left(\frac{\partial}{\partial t}, -\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z}\right)}$$

vu la forme de la métrique minkowskienne. La dérivée contravariante est un tenseur de type $(1, 0)$ agissant par contraction sur un tenseur de type (p, q) pour en faire un tenseur de type $(p + 1, q)$.

De tout cela, on peut déduire d'invariance par changement de référentiel inertiel de quelques lois physiques :

- équation de d'Alembert : soit S un (champ) scalaire satisfaisant à l'équation d'onde

  $$\frac{\partial^2 S}{\partial t^2} - \frac{\partial^2 S}{\partial x^2} - \frac{\partial^2 S}{\partial y^2} - \frac{\partial^2 S}{\partial z^2} = 0$$

  On peut se demander la forme prise par cette équation dans un autre référentiel inertiel. on va se servir de ce que l'on a vu précédemment. $\partial^\mu S$ est un tenseur de type $(1, 0)$ et $\partial_\mu \partial^\mu S$ est à nouveau un scalaire. Explicitons ce terme

  $$\partial_\mu \partial^\mu S = \left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}\right) S$$

  Cela correspond exactement à l'expression du d'Alembertien. Vu que cette quantité scalaire est nulle dans le référentiel inertiel de départ

  $$\boxed{\partial_\mu \partial^\mu S = 0}$$

  on en déduit que dans un autre référentiel inertiel, l'équation prend exactement la même forme $\partial_{\mu'} \partial^{\mu'} S' = 0$. On dit que l'équation de d'Alembert est invariante (ou covariante) relativiste ;

- équation de conservation : imaginons qu'on puisse définir un quadrivecteur charge-courant $\underset{\sim}{j} = (\rho, \vec{j})$ en coordonnées contravariantes. Alors sa contraction avec la dérivée covariante donne le scalaire

$$\partial_\mu j^\mu = \partial_t \rho + \partial_x j^x + \partial_y j^y + \partial_z j^z$$

Dans un référentiel inertiel de référence, on reconnaît l'équation de conservation de la charge, et ainsi

$$\partial_\mu j^\mu = 0$$

On sait alors que par changement de référentiel inertiel, cette équation garde exactement la même forme : $\partial'_\mu j'^\mu = 0$. Là encore, l'équation de conservation garde la même forme dans tous les référentiels inertiels, on dit qu'elle est invariante (ou covariante) relativiste.

Les deux équations précédentes (d'Alembert et conservation) jouent donc des rôles importants en physique relativiste. Similairement, on désire obtenir en relativité des équations dites covariantes, c'est-à-dire prenant des formes similaires dans des référentiels inertiels différents. Les concepts de tenseurs, coordonnées covariantes et contravariantes permettent d'exprimer ces lois sous des formes simples [14].

2.6 Rappels finaux sur la gymnastique d'indices

Les points essentiels sont de veiller à :

- la sommation implicite sur des indices bas/hauts (convention d'Einstein), de plus ces indices ne doivent donc apparaître que deux fois. Par exemple, des expressions telles que

$$u^i v^i \quad \text{ou} \quad u^{ij} v_i w^i{}_k$$

n'ont aucun sens ;

- les hauteurs des indices dans les égalités doivent être les mêmes de chaque côté. Ainsi,

$$U^{ij} = T_i V^j$$

n'est pas correct.

Attention, le symbole de Kronecker δ^i_j valant 0 si $i \neq j$ et 1 si $i = j$ n'est pas un tenseur. Cette propriété ne peut en effet pas être conservée par descente/montée d'indice. Si c'était le cas, on aurait :

$$\delta_{ij} = g_{ik} \delta^k_j = g_{ij}$$

ce qui n'est pas le cas (en dehors du cas d'une base orthonormée).

14. En pratique, il suffit d'avoir des tenseurs de même type de chaque côté des équations.

Chapitre II

Variétés

Si la relativité restreinte peut être décrite en termes de tenseurs sur $\mathbb{R}^4$ avec la métrique $ds^2 = dt^2 - dx^2 - dy^2 - dz^2$, l'un des fondements de la relativité générale est que les masses vont avoir un effet sur la structure de l'espace-temps, c'est-à-dire sur la métrique. On désire alors étendre les notions précédentes de coordonnées puis de vecteurs et tenseurs à des situations plus générales que les espaces vectoriels.

1 Introduction

Une *variété* M est un espace topologique qui localement ressemble à $\mathbb{R}^n$, n étant appelé *dimension* de la variété : pour tout $P \in M$, il existe un ouvert U_P de M contenant P, un ouvert O_P de $\mathbb{R}^n$ et un homéomorphisme $\varphi_P : U_P \to O_P$ appelé *carte locale* en P.

Une illustration graphique est donnée en figure 1 dans le cas $n = 2$.

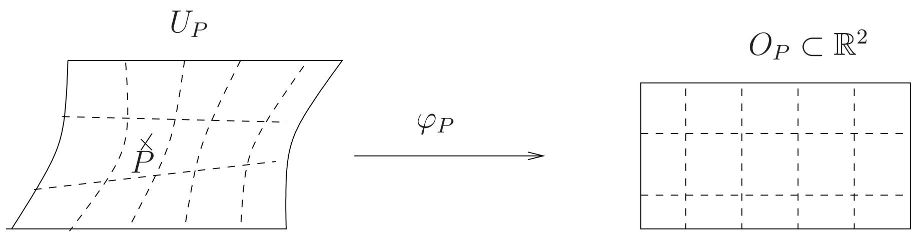

Figure 1 : exemple de variété de dimension 2, les pointillés de U_P et de O_P sont images les uns des autres par φ_P

Quelques remarques peuvent alors être apportées :

- la définition d'une variété M [1] sous-entend qu'une topologie existe sur la variété afin de pouvoir définir les applications continues φ_P. C'est pour cela que M est préalablement supposé être un espace topologique, c'est-à-dire un ensemble muni d'un ensem-

1. Le M désignant la variété provient de *manifold*, variété en anglais.

ble de parties de M qui en constituent les ouverts [2]. Notamment, la variété doit être connexe afin d'avoir unicité de sa dimension n ;

- on rappelle qu'un homéomorphisme est une bijection continue dont l'inverse est également continu. L'existence d'un homéomorphisme entre O_P et U_P assure ici que ces deux ouverts ont les mêmes propriétés topologiques. Typiquement, cela signifie qu'on peut continûment déformer l'un en l'autre ;
- on peut citer les exemples suivants :
  - $\mathbb{R}^n$: une carte est alors donnée par l'identité ;
  - la sphère de dimension n $S^n \subset \mathbb{R}^{n+1}$. Par exemple, S^1 est le cercle du plan (figure 2). Une carte autour d'un point peut être définie par une donnée angulaire [3]. Similairement, S^2 la sphère de l'espace de dimension 3 est aussi une variété. Un point peut être repéré par deux angles : sa latitude et sa longitude, et localement cette correspondance est homéomorphe (ce qui revient à dire que la Terre est localement plate) ;

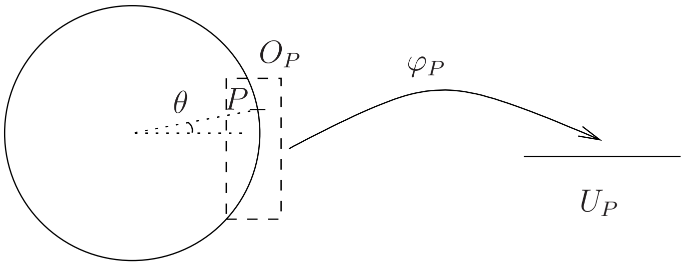

Figure 2 : variété $S^1 \subset \mathbb{R}^2$ dont un ouvert O_P est homéomorphe à un ouvert U_P de $\mathbb{R}$

  - le tore (c'est-à-dire une chambre à air) $S^1 \times S^1$;
  - en revanche, les cas de branchements comme sur la figure 3 ne conviennent pas.

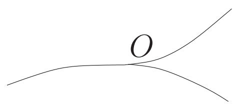

Figure 3 : cas d'un espace topologique qui ne soit pas une variété : courbe avec branchement

Les différents exemples habituels de variétés donnés ci-dessus présentent la particularité commune d'être plongés dans un espace de dimension supérieure. Cela n'est en aucun cas une nécessité et par exemple, la variété S^n existe intrinsèquement, sans besoin de se référer à $\mathbb{R}^{n+1}$;

- l'ensemble des différentes cartes φ_P décrivant la variété M se nomme un *atlas*. Tout point P de M appartient au moins à un ouvert définissant une carte.

---

2. Il y a quelques conditions : M et l'ensemble vide doivent être des ouverts, une union quelconque d'ouverts est un ouvert de même qu'une intersection finie.

3. Dans le cas de S^1, une seule carte ne peut suffire à couvrir la totalité de la variété : S^1 ne peut pas être homéomorphe à une partie de $\mathbb{R}$.

Le but de tout ce qui précède est de réaliser du calcul différentiel sur des fonctions définies sur une variété M. Avant d'aborder cet aspect se pose le problème des zones de M recouvertes par au moins deux cartes. Ces cartes doivent se recouvrir régulièrement, c'est-à-dire que si $P \in M$ est à la fois dans les ouverts O_a et O_b, alors l'application $\varphi_a \, o \, \varphi_b^{-1}$: $\varphi_b(U_a \cap U_b) \to \varphi_a(U_a \cap U_b)$ soit un $\mathcal{C}^\infty$-difféomorphisme[4] au sens usuel des fonctions d'un ouvert de $\mathbb{R}^n$ dans un autre ouvert de $\mathbb{R}^n$ (figure 4). Explicitement, cela signifie qu'on

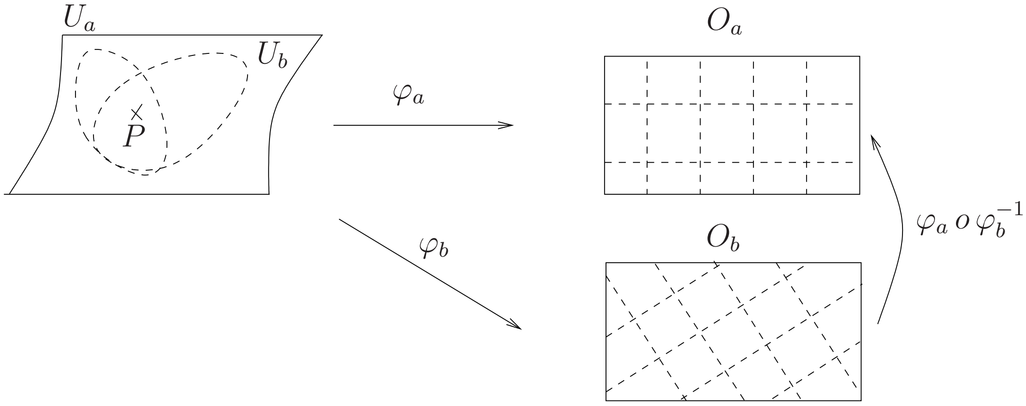

Figure 4 : cas d'un point repérable sur plusieurs cartes

peut disposer de plusieurs jeux de coordonnées, par exemple $\varphi_a(Q) = (x^1(Q), \ldots, x^n(Q))$ et $\varphi_b(Q) = (x'^1(Q), \ldots, x'^n(Q))$ fournis par deux cartes. Dans ce cas, les nouvelles coordonnées peuvent s'exprimer en fonction des anciennes, et la condition de difféomorphisme est équivalente à dire que les fonctions $x'^i(x^1, \ldots, x^n)$ doivent être $\mathcal{C}^\infty$ et le déterminant du jacobien $\det \left(\dfrac{\partial x'^i}{\partial x^j} \right)$ ne peut pas s'annuler.

À l'aide de ce qui précède, il est alors possible de faire du calcul différentiel sur la variété M. En effet, au lieu de considérer la variété, il est possible en utilisant des cartes de se ramener localement à $\mathbb{R}^n$. Plus spécifiquement, considérons :

- une fonction $f : M \to \mathbb{R}$. On dira que f est $\mathcal{C}^\infty$ si c'est le cas dans toute carte de M, c'est-à-dire si $f \, o \, \varphi_P^{-1} : O_P \to \mathbb{R}$ est $\mathcal{C}^\infty$ pour tout P de M. On notera $\mathcal{C}^\infty(M)$ l'ensemble de ces fonctions ;
- une courbe tracée sur M. C'est une application $\gamma : J \to M$, où J est un intervalle de $\mathbb{R}$. γ sera considérée comme $\mathcal{C}^\infty$ si elle l'est dans les différentes cartes.

Dans la suite, on supposera systématiquement que les variétés vérifient ces conditions.

2 Vecteurs d'une variété, espace tangent

Les vecteurs sont les éléments d'un espace vectoriel. Définir des vecteurs à partir d'une variété n'est pas immédiat. L'exemple d'une sphère montre la difficulté. On va en fait définir

4. Un $\mathcal{C}^\infty$- difféomorphisme est une bijection de classe $\mathcal{C}^\infty$ telle que son inverse le soit aussi.

en tout point d'une variété un plan tangent, de la même manière qu'il est possible de considérer un plan tangent à une sphère.

2.1 Dérivées des fonctions définies sur la variété

Dans $\mathbb{R}^n$, un vecteur X est repéré par ses composantes contravariantes $(X^1, \ldots, X^n)$ dans une base quelconque. On peut alors associer à X une dérivation sous la forme :

$$Xf = X(f) = \sum_{i=1}^{n} X^i \frac{\partial f}{\partial x^i}$$

où f est une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$. Même s'il est impossible de faire rigoureusement la même chose dans le cas d'une variété, on peut s'en inspirer.

Soit M une variété, et $x \in M$. Une *dérivation* en P est une application $X_P : \mathcal{C}^\infty(M) \to \mathbb{R}$ telle que

$$\forall \alpha, \beta \in \mathbb{R},\ \forall f, g \in \mathcal{C}^\infty(M), \quad \begin{array}{l} X_P(\alpha f + \beta g) = \alpha X_P(f) + \beta X_P(g) \\ X_P(fg) = f(P)X_P(g) + X_P(f)g(P) \end{array}$$

Dans cette définition, on reconnaît sans surprise la linéarité de la dérivée, ainsi que la dérivation d'un produit de fonction. Il est à noter qu'il n'est pas nécessaire que les fonctions soient définies et $\mathcal{C}^\infty$ sur M entier, un voisinage de P est suffisant.
L'ensemble des dérivations en x se nomme espace vectoriel[5] tangent en P à M et est noté $T_P M$. Ensuite, l'*espace tangent* à la variété M est

$$TM = \bigcup_{P \in M} T_P M$$

On définit finalement un *champ de vecteurs* noté X sur une variété comme une application de $X : M \to TM$ qui à chaque point P de la variété associe une dérivation X_P en P.

Grâce aux cartes locales, il est possible de donner une base de $T_P M$. On considère la carte locale $\varphi : U \to O$ autour de P, et on note $\varphi(Q) = (x^1(Q), \ldots, x^n(Q))$ pour $Q \in U$. Alors, pour $f \in \mathcal{C}^\infty(M)$, $f \, o \, \varphi^{-1}$ est une fonction d'un ouvert de $\mathbb{R}^n$ dans $\mathbb{R}$, et une base des dérivations est constituée des dérivées partielles selon les différentes directions :

$$(\partial_{1,P}, \ldots, \partial_{n,P}) \text{ est une base de } T_P M, \quad \text{où} \quad \partial_{i,P} f = \frac{\partial (f \, o \, \varphi^{-1})}{\partial u^i}(\varphi(P))\,^6$$

Ainsi, dans cette base, une dérivation X_P s'écrit

$$X_P = X_P^i \partial_{i,P}$$

5. Il est clair par linéarité que l'on obtient bien un espace vectoriel.
6. Les notations deviennent rapidement très lourdes, on a donc rapidement tendance à simplifier…

Les (X_P^i) forment alors les coordonnées contravariantes de X_P. En effet, dans une autre carte φ' telle que $\varphi'(Q) = (x'^1(Q), \ldots, x'^n(Q))$,

$$X_P = X_P^j \partial_{j,P} = X_P'^i \partial'_{i,P} = X_P^{i'} \partial_{i',P}$$

(la dernière égalité sert simplement à systématiser les notations en faisant porter le prime par les indices). Or, pour une fonction g définie sur un ouvert de $\mathbb{R}^n$, la dérivation de fonctions composées donne

$$\frac{\partial g}{\partial x'^i}(x^1, \ldots, x^n) = \sum_j \frac{\partial g}{\partial x^j}(x^1, \ldots, x^n) \frac{\partial x^j}{\partial x'^i}$$

soit de manière plus compacte en reprenant les notations du chapitre précédent :

$$\partial'_{i,P} = \partial_{i',P} = \Lambda_{i'}{}^j \partial_{j,P} \quad \text{avec} \quad \Lambda_{i'}{}^j = \frac{\partial x^j}{\partial x'^i}$$

et

$$X^{i'} = \Lambda^{i'}{}_j X^j, \quad \text{avec} \quad \Lambda^{i'}{}_j = \frac{\partial x'^i}{\partial x^j}$$

Le parallèle avec les expressions (1.12) et (1.7) est immédiat : les X_P^i se comportent bien comme des coordonnées contravariantes de vecteurs, et les $\partial_{i,P}$ comme des vecteurs. Le but est atteint : en chaque point de la variété, il a été possible d'"attacher" un espace vectoriel.

Néanmoins, la grande différence avec une variété qui soit identifiable à un espace vectoriel plat (c'est-à-dire muni d'une métrique uniforme) est que les matrices de changement de base Λ dépendent du point considéré sur la variété. En conséquence, il n'est plus possible d'obtenir une base identique $(\vec{e}_1, \ldots, \vec{e}_n)$ pour tous les espaces T_PM. En général, cela n'a d'ailleurs aucun sens : les differents espaces tangents T_PM n'ont aucun lien naturel entre eux, et comparer deux vecteurs de T_PM et T_QM (avec $P \neq Q$) n'a aucun sens [7].

2.2 Tangentes aux courbes tracées sur la variété

Il est aussi possible de donner un aspect plus géométrique aux vecteurs d'une variété. Soit P un point de la variété M. On considère alors les courbes γ de classe $\mathcal{C}^\infty$ tracées sur la variété à partir d'un intervalle de $\mathbb{R}$ de la forme $]-a, a[$ telles que $\gamma(0) = P$. Alors, l'espace tangent T_PM est aussi défini par

$$T_PM = \{\dot\gamma(0) \,:\, \gamma \text{ tracée sur } M \text{ telle que } \gamma(0) = P\}$$

Cette définition doit être comprise en terme de cartes locales. Dans la carte φ, $\gamma(t) = (\gamma^1(t), \ldots, \gamma^n(t))$ [8]. Ainsi, $\dot\gamma(0) = (\dot\gamma^1(0), \ldots, \dot\gamma^n(0))$. La situation est représentée en figure 5 [9].

7. Une conséquence de tout cela est que la dérivation covariante telle qu'elle a été définie au chapitre précédent ne possède plus un caractère tensoriel. Pour rétablir ce caractère, il faut définir un parallélisme entre vecteurs d'espaces tangents différents au moyen d'une connexion (chapitre xx).

8. Il y a là un abus de notation, il faudrait plutôt écrire $\varphi \, o \, \gamma(t) = (\gamma^1(t), \ldots, \gamma^n(t))$, et $\frac{d}{dt}(\varphi \, o \, \gamma)(0) = (\dot\gamma^1(0), \ldots, \dot\gamma^n(0))$.

9. Là encore, l'espace tangent T_PM existe indépendamment d'un espace de dimension supérieure dans lequel la variété M serait plongé, $\mathbb{R}^3$ dans le cas de la figure 5.

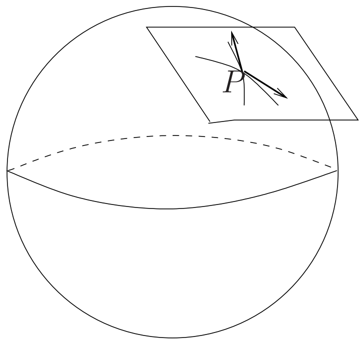

Figure 5 : espace $T_P M$ tangent à S^2 en x

On peut donner rapidement quelques arguments liant les deux approches. À partir d'une courbe γ tracée sur M vérifiant les conditions précédentes, on peut obtenir une dérivation $\dot{\gamma}(0) = X_{\gamma,P}$ par

$$\dot{\gamma}(0)(f) = X_{\gamma,P}(f) = \frac{d(f\, o\, \gamma)}{dt}(0)$$

où $f \in \mathcal{C}^\infty(M)$. Réciproquement, à partir d'un vecteur X_P, il est possible de tracer une courbe sur M possédant X_P comme dérivée à l'origine. En effet, si $X_P = (X_P^1, \ldots, X_P^n)$ dans la carte φ (on supposera que $\varphi(P) = (0, \ldots, 0)$), la courbe $\gamma(t) = \varphi^{-1}(X_P^1 t, \ldots, X_P^n t)$ convient bien.

3 Formes linéaires sur une variété, espace cotangent

À partir de l'espace $T_P M$ tangent en P à M, il est facile de définir $T_P^* M = (T_P M)^*$ appelé *espace cotangent* en P à M. C'est l'espace des formes linéaires sur $T_P M$.

Soit la carte locale φ définie par $\varphi(Q) = (x^1(Q), \ldots, x^n(Q))$. Comme on l'a vu précédemment, une base de $T_Q M$ est formée des dérivées partielles selon les directions en question

$$\left(\left(\frac{\partial}{\partial x^1} \right)_P, \ldots, \left(\frac{\partial}{\partial x^n} \right)_P \right) = (\partial_{1,P}, \ldots, \partial_{n,P})$$

Alors, la base duale associée est notée $(dx_P^1, \ldots, dx_P^n)$ et vérifie

$$dx^i_P(\partial_{j,P}) = \delta^i_j$$

En conséquence, si $\omega_P \in T_P^* M$, alors ω s'écrit

$$\omega_P = \omega_{P,j}\, dx_P^j$$

Enfin, l'espace cotangent à M est défini par

$$T^* M = \bigcup_{P \in M} T_P^* M$$

De même que pour les vecteurs, on définit un champ de formes linéaires sur M par une application $\omega : M \to T^*M$ associant à chaque point de la variété une forme linéaire $\omega(P) = \omega_P$ sur T_PM.

4 Tenseurs sur une variété

À partir du moment où on sait définir des champs de vecteurs et des champs de formes linéaires sur une variété, il est facile de définir des *champs de tenseurs*. Il suffit de procéder de la même manière qu'au chapitre précédent. On rappelle rapidement le principe. Soit P un point de la variété. Une forme linéaire en P agit linéairement sur un vecteur en P en donnant un scalaire. De manière symétrique, un vecteur en P agit sur une forme linéaire en P de manière à donner en résultat un scalaire. Un tenseur T_P qui est p fois contravariant et q fois covariant agit sur q vecteurs et donne en sortie un objet pouvant agir sur p formes linéaires (on omet de préciser que cela concerne des quantités définies au point P de la variété). Dans la carte locale φ, on écrit alors, similairement au cas d'un espace vectoriel (expression (1.14)) :

$$T_P = T_P^{i_1 \dots i_q}{}_{j_1 \dots j_q}\, \partial_{i_1,P} \otimes \dots \otimes \partial_{i_q,P} \otimes dx_P^{j_1} \otimes \dots \otimes dx_P^{j_q}$$

Ensuite, le champ de tenseur T est obtenu en considérant $T_P^{i_1 \dots i_q}{}_{j_1 \dots j_q}$ comme une fonction de P, point repérant la position sur la variété.

L'exemple le plus utilisé par la suite sera le tenseur métrique, deux fois covariant, défini en chaque point de la variété

$$g(P) = g_{ij}(P) dx_P^i \otimes dx_P^j = g_{ij} dx^i dx^j$$

(la dernière égalité n'étant qu'une simplification d'écriture). La donnée du tenseur métrique permet de définir des équivalents de distance sur la variété.

5 Quelques exemples

5.1 Plan euclidien

Le plan euclidien $E = \mathbb{R}^2$ peut évidemment être vu comme une variété, munie de la carte triviale. On considère une base orthonormée $(\vec{e}_x, \vec{e}_y)$. En tout point P de coordonnées (contravariantes) (x, y), on obtient l'espace tangent $T_PE \simeq E$ de base $(\partial_{P,x}, \partial_{P,y}) \simeq (\vec{e}_x, \vec{e}_y)$. En fait, on identifie alors naturellement la base de l'espace vectoriel de départ (qui n'existe pas dans le cas général d'une variété) et celle de chaque espace tangent. Le tenseur métrique est alors $g_{ij} = \delta_{ij}$.

En changeant la carte, c'est-à-dire les coordonnées, la différence sera plus claire. On va s'intéresser aux coordonnées polaires, valides en dehors de l'origine O. Le point P est désormais repéré par (r, θ) (figure 6) [10]

10. En fait, une seule carte ne suffit pas à définir les coordonnées polaires. En partant d'un point sur $(O, \vec{e}_x)$

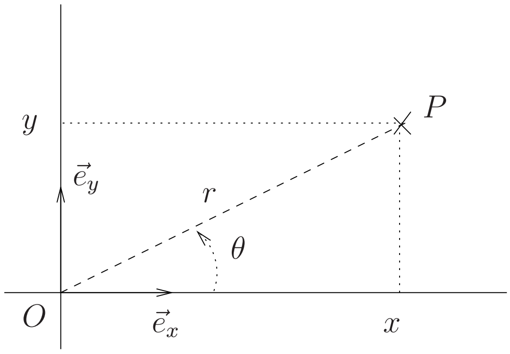

Figure 6 : coordonnées polaires dans le plan

Une base de l'espace vectoriel tangent $T_P E$ est alors $\left(\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}\right) = (\partial_r, \partial_\theta)$ [11]. On va alors exprimer le tenseur métrique g'_{ij} dans ces nouvelles coordonnées. Soit une fonction $f(x, y)$ s'écrivant en coordonnées polaires : $\tilde{f}(r, \theta) = f(x, y)$. Par dérivation composée, on obtient

$$\frac{\partial \tilde{f}}{\partial r} = \frac{\partial x}{\partial r}\frac{\partial f}{\partial x} + \frac{\partial y}{\partial r}\frac{\partial f}{\partial y}$$

Or, $x = r\cos\theta$ et $y = r\sin\theta$, d'où

$$\frac{\partial \tilde{f}}{\partial r} = \cos\theta\frac{\partial f}{\partial x} + \sin\theta\frac{\partial f}{\partial y} \quad \text{soit} \quad \partial_r = \cos\theta\,\partial_x + \sin\theta\,\partial_y$$

Similairement,

$$\partial_\theta = -r\sin\theta\,\partial_x + r\cos\theta\,\partial_y$$

En se rappelant que (∂_x, ∂_y) est une base orthonormée, on obtient

$$g'_{rr} = \partial_r.\partial_r = (\cos\theta\,\partial_x + \sin\theta\,\partial_y).(\cos\theta\,\partial_x + \sin\theta\,\partial_y) = \cos^2\theta + \sin^2\theta = 1$$

En réitérant l'opération, on aboutit à

$$g'_{r\theta} = g'_{\theta r} = 0 \quad \text{et} \quad g'_{\theta\theta} = r^2$$

soit [12]

$$\boxed{g' = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix} \qquad \text{c'est-à-dire} \qquad g' = dr^2 + r^2\,d\theta^2} \tag{2.1}$$

donc à $\theta = 0$, puis en faisant un tour autour de O, la coordonnée θ prend la valeur 2π en revenant sur le point original. Les coordonnées polaires sont donc définies localement (en dehors de O).

11. Cette base est locale, c'est-à-dire varie d'un point à l'autre de la variété (figure 7) et le point P de la variété auquel elle se rattache devrait donc être indiqué. Pour alléger les notations, cela est sous-entendu.

12. Un calcul "direct" de la métrique est en fait légitime et bien plus aisé. Comme $x = r\cos\theta$ et $y = r\sin\theta$,

$$g = dx^2 + dy^2 = (dr\,\cos\theta - r\sin\theta\,d\theta)^2 + (dr\,\sin\theta + r\cos\theta\,d\theta)^2 = dr^2 + r^2\,d\theta^2$$

La base ($\hat{r} = \partial_r, \hat{\theta} = \partial_\theta$) n'est donc pas orthonormée, seulement orthogonale (∂_r est normé, mais ∂_θ est de norme r). Il est utile de faire une précision à ce point. Sur une variété, il est possible de choisir une métrique g quelconque (il s'agit seulement d'une tenseur deux fois covariant symétrique). Ici, on s'est donné la nature euclidienne du plan, et donc la métrique en base cartésienne orthonormée. On a ensuite effectué un changement de coordonnées en polaires, puis cherché le changement de métrique correspondant. Il aurait aussi été possible de calculer g' par changement de coordonnées tensorielles :

$$g'_{ij} = g_{i'j'} = \Lambda_{i'}{}^k \Lambda_{j'}{}^l g_{kl} \quad \text{avec} \quad \Lambda_{i'}{}^k = \frac{\partial x^k}{\partial x^{i'}}$$

ce qui donne ici

$$\begin{aligned} \Lambda_r{}^x &= \frac{\partial x}{\partial r} = \cos\theta & \Lambda_r{}^y &= \frac{\partial y}{\partial r} = \sin\theta \\ \Lambda_\theta{}^x &= \frac{\partial x}{\partial \theta} = -r\sin\theta & \Lambda_\theta{}^y &= \frac{\partial y}{\partial \theta} = r\cos\theta \end{aligned}$$

Au final, par exemple,

$$\begin{aligned} g'_{\theta\theta} &= \Lambda_\theta{}^x \Lambda_\theta{}^x g_{xx} + \Lambda_\theta{}^x \Lambda_\theta{}^y g_{xy} + \Lambda_\theta{}^y \Lambda_\theta{}^x g_{yx} + \Lambda_\theta{}^y \Lambda_\theta{}^y g_{yy} \\ &= (-r\sin\theta)^2\, 1 + 0 + 0 + (r\cos\theta)^2\, 1 = r^2 \end{aligned}$$

et de même pour les autres composantes. On retrouve bien le résultat (2.1).

Ainsi, g et g' sont identiques à un changement de carte près, et les mesures de distances le long de chemins données par la métrique seront les mêmes, que l'on utilise les coordonnées cartésiennes ou les coordonnées polaires.

Comme on l'a vu précédemment, dans la carte $(x^1, \ldots, x^n)$, le vecteur ∂_{x^1} peut être interprété comme tangent à la courbe $\gamma(t) = (t, \mathrm{cte}, \ldots, \mathrm{cte})$ et passant par le point désiré. Cela est illustré sur la figure 7.

5.2 Espace euclien de dimension 3

La première possibilité est d'utiliser une base orthonormée $(\vec{e}_x, \vec{e}_y, \vec{e}_z)$ afin d'obtenir des coordonnées contravariantes cartésiennes (x, y, z). Le tenseur métrique est alors évidemment de la forme $g_{ij} = \delta_{ij}$. Ensuite, le passage aux coordonnées sphériques (r, θ, φ) (figure 8) se fait comme précédemment avec les coordonnées polaires, menant au tenseur métrique

$$g' = dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\varphi^2) \tag{2.2}$$

Sphère S^2

C'est une variété de dimension 2, qui peut-être vue comme une partie de $\mathbb{R}^3$. Une carte possible est constituée des deux dernières coordonnées sphériques (θ, φ) (colatitude, longitude), ce qui mène à la métrique

$$g = d\theta^2 + \sin^2\theta\, d\varphi^2 \tag{2.3}$$

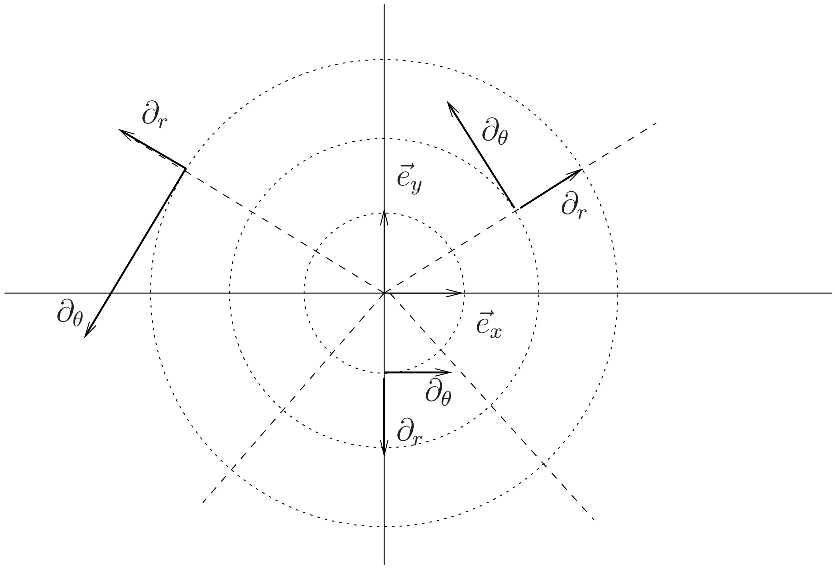

Figure 7 : courbes $r = \mathrm{cte}$ (pointillés), $\theta = \mathrm{cte}$ (tirets), et vecteurs tangents ∂_r et ∂_θ

(qualitativement, on part de la métrique en sphériques de $\mathbb{R}^3$ (expression (2.2)) en supprimant dr vu que la coordonnée radiale est désormais fixée, à $r = 1$ pour simplifier).

5.3 Espace-temps minkowskien

C'est une variété M de dimension 4 munie de la carte $(x^0, x^1, x^2, x^3) = (t, x, y, z))$ et de la métrique minkowskienne (dite plate) que l'on rappelle (voir page 10)

$$g = dt^2 - dx^2 - dy^2 - dz^2 = \eta_{\mu\nu} dx^\mu dx^\nu$$

Les coordonnées covariantes sont donc données par $(x_0, x_1, x_2, x_3) = (t, -x, -y, -z)$.

5.4 Espace-temps de Schwarzschild

C'est une variété M de dimension 4, munie de la carte (t, r, θ, φ) et de la métrique

$$g = \left(1 - \frac{2m}{r}\right) dt^2 - \frac{dr^2}{1 - \dfrac{2m}{r}} - r^2 (d\theta^2 + \sin^2\theta \, d\varphi^2)$$

On reconnaît une métrique de type minkowskienne $(+ - --)$, présentant des liens certains avec les coordonnées sphériques de l'expression (2.2). Ceux-ci seront commentés plus tard en détail. On observe simplement qu'un déplacement radial de dr ne correspond pas à la même distance suivant les valeurs de r à cause du dénominateur en $1 - \frac{2m}{r}$: l'espace est dit courbe.

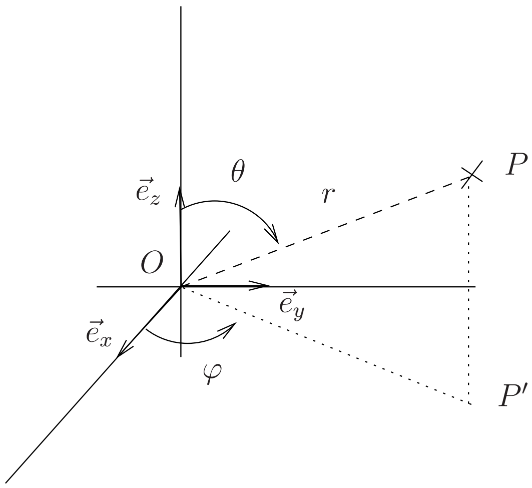

Figure 8 : coordonnées sphériques (r, θ, φ) pour un point P, P' est le projeté orthogonal de P dans le plan $(O, \vec{e}_x, \vec{e}_y)$

Chapitre III

Géodésiques

Étant donnée une variété M munie d'une métrique g, on peut calculer la distance [1] entre deux point P et Q par

$$d(P,Q) = \int_P^Q |ds| = \int_P^P \sqrt{|g_{\mu\nu} dx^\mu dx^\nu|}$$

Étant donnés deux points P et Q de la variéte, on cherche alors à trouver les chemins entre P et Q rendant extrémale, ou de manière plus précise stationnaire par rapport aux variations, la distance entre P et Q.

1 Équations des géodésiques

2 Applications

2.1 Plan euclidien

2.2 Sphère S^2

2.3 Espace-temps de Schwarzschild

1. La quantité ainsi évaluée n'a pas forcément les propriétés d'une distance au sens mathématique du terme, elle peut par exemple être nulle entre deux points disctincts. En revanche, dans le cas d'un espace euclidien, on retrouve bien la distance usuelle issue de la norme 2. Par analogie, la dénomination de distance est restée.