Initiation à la recherche

Massonnet Louis
M1 MFA - Université de Paris
2020
Supervisor : Laurain Paul

Contents

1 Rappels de mathématiques **2**

1.1 Champs de vecteurs et dérivation . . . 2

1.1.1 Champs de vecteurs sur une variété . . . 2

1.1.2 La dérivée de Lie . . . 4

1.1.3 Connexions et dérivée covariante . . . 6

1.2 Tenseurs . . . 8

1.2.1 Définition et opérations sur les tenseurs . . . 8

1.2.2 Champs de tenseurs et dérivée covariante . . . 10

2 Physique **11**

2.1 L'Electromagnétisme . . . 11

2.1.1 Reformulation à la Yang-Mills . . . 11

2.1.2 Approche variationnelle . . . 13

2.2 Théorie de Yang-Mills . . . 14

2.3 Relativité générale et géométrie . . . 16

2.4 Théorie de Kaluza-Klein . . . 19

1 Rappels de mathématiques

1.1 Champs de vecteurs et dérivation

1.1.1 Champs de vecteurs sur une variété

Le but de ce chapitre est de définir la notion de champs de vecteurs sur une variété M, qui sera une variété réelle lisse et de dimension n dans toute la suite . Pour ce faire nous aurons d'abord besoin de définir les vecteurs et espaces tangents puis le fibré tangent. Ces deux notions nous permettrons de définir ce qu'est un champs de vecteurs sur M.

Il existe deux définitions (équivalentes) de vecteur tangent à une variété en un point. La première est donnée par une classe d'équivalence de courbes:

Définition : (vecteur tangent 1)
Soit M une variété différentielle. Soit $P \in M$ et soit (Ω, φ) une carte locale en P. Soient γ_1 et γ_2 des courbes sur M telles que $\gamma_1(0) = \gamma_2(0) = P$. On dit alors que γ_1 et γ_2 sont équivalentes, que l'on note $\gamma_1 \sim \gamma_2$, si $(\varphi \circ \gamma_1)'(0) = (\varphi \circ \gamma_2)'(0)$. Un vecteur tangent à M en P est alors une classe d'équivalence pour la relation $\sim$.

On peut également définir un vecteur tangent de la façon suivante :

Définition : (vecteur tangent 2)
Soit M une variété différentielle et $P \in M$. Soit (Ω, φ) une carte locale en P. Soit f une fonction à valeurs réelles définies sur un voisinage de P. On dit que f est plate en P si $(f \circ \varphi^{-1})'(\varphi(P)) = 0$. Cette définition a bien un sens car elle ne dépend pas de la carte locale en P. Un vecteur tangent à M en P est alors une application $X : f \mapsto X(f)$ définie sur l'ensemble des fonctions différentiables sur un voisinage de P telle que :
1) $\forall a, b \in \mathbb{R} : X(af + bg) = aX(f) + bX(g)$
2) $X(f) = 0$ si f est plate en P.

On peut voir que les conditions 1) et 2) impliquent qu'un vecteur tangent vérifie également :

$$X(fg) = f(P)X(f) + g(P)X(f)$$

On peut vérifier que les définitions 1 et 2 de vecteur tangent sont bien équivalentes. Cela nous permet de définir l'espace tangent à une variété en un point :

Définition : (espace tangent)
L'espace tangent à une variété M en un point P, noté $T_P M$, est l'ensemble des vecteurs tangents à M en P.

On peut munir $T_P M$ d'une structure d'espace vectoriel en posant :

$$(X + Y)(f) = X(f) + Y(f) \text{ et } (aX)(f) = a(X(f))$$

On peut désormais définir le fibré tangent :

<u>Définition :</u> (fibré tangent)
On définit le fibré tangent comme la réunion des espaces tangents :

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

On peut montrer que si M est de classe $\mathcal{C}^k$ alors TM peut être muni d'une structure de variété de classe $\mathcal{C}^{k-1}$.
On peut également remarquer que si on équipe TM de la projection :

$$\begin{array}{rccc} \pi : & TM & \longrightarrow & M \\ & X \in T_P M & \longmapsto & P \end{array}$$

cela fait de (TM, π) un fibré vectoriel de fibre $\mathbb{R}^n$. Nous sommes maintenant quasiment prêts pour définir ce qu'est un champs de vecteurs sur M. Il ne nous manque que la notion de section d'un fibré :

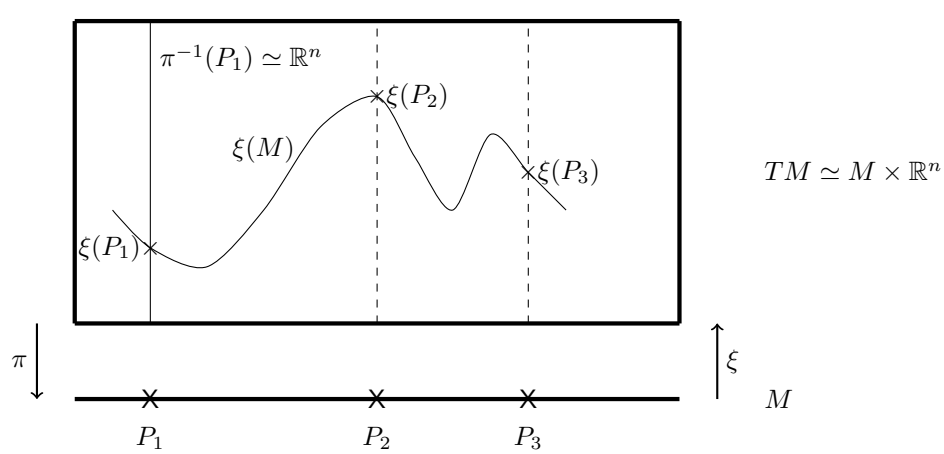
<u>Définition :</u> (section d'un fibré)
Une section d'un fibré (E, π) sur M est une application $\xi : M \longrightarrow E$ différentiable telle que $\pi \circ \xi = \mathrm{id}$

La définition d'un champs de vecteurs est alors simplement :

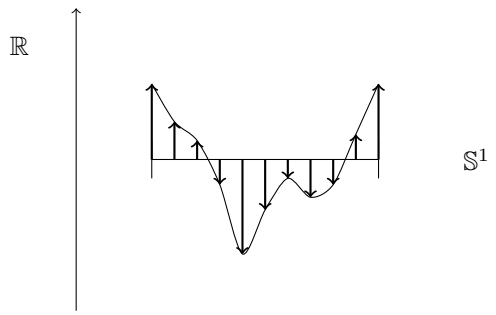
<u>Définition :</u> (champs de vecteurs)
Un champs de vecteurs est une section du fibré tangent.

Un champ de vecteurs est donc une application associant à chaque point d'une variété un vecteur de l'espace tangent en ce point (avec une condition de régularité).

Il est possible de se représenter simplement cette définition surtout dans le cas où le fibré tangent est trivial, c'est-à-dire lorsque que l'on peut l'exprimer comme $TM \simeq M \times \mathbb{R}^n$. Dans ce cas on peut résumer la situation par le schéma suivant :



On peut visualiser encore mieux la notion de vecteur sur un exemple plus concret : le cercle $\mathbb{S}^1$. Dans ce cas le fibré tangent est trivialisable et on peut donc le munir d'une carte telle que $T\mathbb{S}^1 \simeq \mathbb{S}^1 \times \mathbb{R}$. Un champs de vecteur sur $\mathbb{S}^1$ est alors une façon d'associer un réel à chaque point du cercle. Si on voit le cercle comme un segment sur lequel on a identifié les extrémités un champ de vecteur sur $\mathbb{S}^1$ ressemble donc au schéma suivant :



1.1.2 La dérivée de Lie

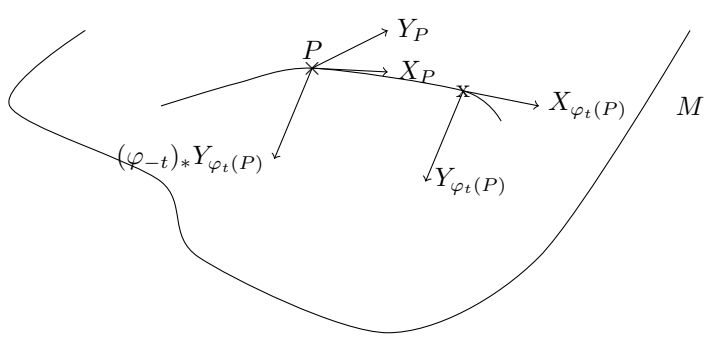
Dans ce chapitre on veut se donner un moyen de dériver un champs de vecteur par rapport à un autre. Pour ce faire on aimerait donner une définition qui ressemble à celle que l'on connaît pour les fonctions numériques : une limite de taux d'accroissement. Cependant, notre contexte est plus complexe que celui des fonctions numériques. En effet, on ne peut en principe pas comparer des vecteurs qui ne sont pas dans le même espace tangent. Il faut donc "transporter" les vecteurs pour pouvoir les comparer et cela peut se faire au moyen du flot.

<u>Définition :</u> (dérivée de Lie)
Soient X et Y des champs de vecteurs différentiables. Soit φ_t le flot de X. On définit alors la dérivée de Lie comme :

$$(\mathcal{L}_X Y)_P = \lim_{t\to 0} \tfrac{1}{t} \left[(\varphi_{-t})_* Y_{\varphi_t(P)} - Y_P \right]$$

où l'étoile en bas représente le pushforward, et l'indice P signifie que l'on prend la valeur en $P \in M$ du champs. On peut vérifier que la dérivée de Lie d'un champs de vecteur est bien un champs de vecteurs.
On peut même définir la dérivée de Lie d'une fonction différentiable :

$$\mathcal{L}_X(f) = \lim_{t\to 0} \tfrac{1}{t} (f \circ \varphi_t - f)$$



On remarque que dans cette définition la valeur de $\mathcal{L}_X Y$ dépend non seulement des variations de Y, ce qui est attendu pour une dérivation, mais également des variations de X. Cela nous poussera à définir une autre dérivation, dite covariante, qui ne dépendra que des variations de Y au prochain chapitre.

Il est possible de réexprimer la dérivée de Lie d'une autre manière, à l'aide du crochet de Lie.

Définition : (crochet de Lie)
Soient X et Y des champs de vecteurs. Le crochet de Lie de X et Y est le champs de vecteurs défini par :

$$[X, Y] = XY - YX$$

Pour rendre cette définition plus claire, nous pouvons exprimer $[X, Y]$ en coordonnées locales. Pour $x \in M$ alors $X(x) \in T_x M$. Si on note ∂_i une base de l'espace tangent en x alors on peut exprimer $X(x)$ en coordonnées locales comme :

$$X(x) = \sum_{i=1}^{n} X^i(x)\partial_i$$

les composantes du crochet de Lie en coordonnées locales sont alors :

$$[X, Y]^i = \sum_{j=1}^{n} \left(X^j(x)\frac{\partial Y^i}{\partial x_j}(x) - Y^j(x)\frac{\partial X^i}{\partial x_j}(x) \right)$$

Le crochet de Lie de deux champs de vecteurs vérifié les propriétés suivantes :

Propriétés :
Soient X, Y et Z des champs de vecteurs. Alors :
1) $[.,.]$ est antisymétrique, c'est à dire : $[X, Y] = -[Y, X]$
2) le crochet de Lie vérifie l'égalité de Jacobi : $[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0$

On a également la propriété suivante :

Proposition :
Pour deux champs de vecteurs X et Y on a :

$$\mathcal{L}_X Y = [X, Y]$$

Preuve :
Montrons que l'on a égalité en tout $P \in M$. Soit f une fonction différentiable sur un voisinage de P. Par définition :

$$\begin{aligned}
(\mathcal{L}_X Y)_P(f) &= \lim_{t \to 0} \frac{1}{t} \left[(\varphi_{-t})_* Y_{\varphi_t(P)}(f) - Y_P(f) \right] \\
&= \lim_{t \to 0} \frac{1}{t} \left[(\varphi_{-t})_* Y_{\varphi_t(P)}(f) - Y_{\varphi_t(P)}(f) + Y_{\varphi_t(P)}(f) - Y_P(f) \right] \\
&= \lim_{t \to 0} \frac{1}{t} \left[Y_{\varphi_t(P)}(f \circ \varphi_{-t}) - Y_{\varphi_t(P)}(f) + Y_{\varphi_t(P)}(f) - Y_P(f) \right]
\end{aligned}$$

On remarque alors que les deux derniers termes donnent :

$$\lim_{t \to 0} \frac{1}{t} \left[Y_{\varphi_t(P)}(f) - Y_P(f) \right] = X(Y(f))_P$$

alors que la différence des deux premiers s'exprime comme :

$$\lim_{t\to 0}\frac{1}{t}\left[Y_{\varphi_t(P)}(f\circ\varphi_{-t})-Y_{\varphi_t(P)}(f)\right]=-\lim_{t\to 0}\left[Y\left(\frac{f\circ\varphi_{-t}-f}{-t}\right)\right]_{\varphi_t(P)}=-Y(X(f))_P$$

ce qui montre bien le résultat voulu.

Nous avons donc défini une dérivation sur les champs de vecteurs sur une variété grâce à la dérivée de Lie, et nous avons même pu la réexprimer comme le crochet de Lie de deux champs. Nous avons remarqué que la dérivée de Lie a une propriété qui nous poussera à définir la dérivée covariante au prochain chapitre. Cependant, la dérivée de Lie peut s'avérer utile en physique. En effet prendre la dérivée de Lie d'un champs Y par rapport à un champs X revient à étudier les variations de Y le long du flot de X. Cette interprétation est particulièrement utile en physique, par exemple lorsque la variété M est munie d'une métrique g alors la nullité de la dérivée de Lie de g par rapport à un champs ξ est l'indication que la variété a une certaine symétrie. Dans le cadre de l'espace-temps de Schwarzschild en relativité générale le fait que l'espace temps soit stationnaire se traduit par :

$$\mathcal{L}_{\partial_t}g=0$$

tandis que l'invariance par rotation se traduit par :

$$\mathcal{L}_{\partial_\varphi}g=0$$

où g est la métrique de Scharzschild et ∂_t et ∂_φ sont des champs de vecteurs bien choisis. La dérivée de Lie est donc un moyen d'identifier des symétries.

1.1.3 Connexions et dérivée covariante

Dans ce chapitre et dans la suite on utilisera la convention de sommation d'Einstein qui consiste à sommer sur les indices répétés une fois en haut et une fois en bas. Ainsi une somme $\sum_{i=1}^{n} u_i v_i$ s'écrit simplement $u^i v_i$.

Nous avons vu lorsque que nous avions défini la dérivée de Lie $\mathcal{L}_X Y$ que cette quantité ne dépend pas seulement de $X_P \in T_P M$ mais de X. C'est pour cela que l'on définit une dérivée covariante. Dans ce chapitre on notera $\Gamma(M)$ l'ensemble des champs de vecteurs différentiables sur M.

<u>Définition :</u> (dérivée covariante)
Une dérivée covariante sur M (ou connexion) est une application billinéaire $D:\Gamma(M)\times\Gamma(M)\to\Gamma(M)$ vérifiant pour toute fonction différentiable f et tous champs de vecteurs X et Y:

1) $D(fX,Y)=fD(X,Y)$
2) D vérifie la formule de Leibniz, c'est-à-dire :

$$D(X,fY)=X(f)Y+fD(X,Y)$$

3) Si X est de classe $\mathcal{C}^r$ et Y de classe $\mathcal{C}^{r+1}$ alors $D(X,Y)$ est un champs de vecteurs de classe $\mathcal{C}^r$.
Dans la suite on notera systématiquement $D(X,Y):=D_XY$

La première condition impose que la valeur de $D_XY(P)$ pour un $P\in M$ ne dépend que de la valeur de X en P et pas de son comportement dans un voisinage de P. En effet, soient X et X' deux champs de vecteurs tels que $X(P)=X'(P)$. On peut exprimer X et X' en coordonnées locales au voisinage de P : $X=X^i\partial_i$ et $X'=X'^i\partial_i$.

Alors :

$$D_{X-X'}Y(P) = D_XY(P) - D_XY(P) = (X^i - X'^i)(P)D_iY(P) = 0$$

où on a utilisé la notation $D_i := D_{\partial_i}$. La dériviéee covariante ne dépend donc pas des variations de X autour de P contrairement à la dérivée de Lie.

Une fois encore, on peut écrire la dérivée covariante d'un champs de vecteur en coordonnées locales. Soit $P \in M$ et soient X et Y des champs de vecteurs sur M. On peut alors écrire X et Y en coordonnées locales au voisinage de P comme $X = X^i\partial_i$ et $Y = Y^i\partial_i$. La dérivée covariante s'écrit alors :

$$D_XY = D_{X^i\partial_i}Y^j\partial_j = X^iD_iY^j\partial_j$$

En utilisant la propriété 2) de la définition il vient :

$$D_XY = X^i(\partial_iY^j)\partial_j + X^iY^jD_i(\partial_j)$$

D'après la propriété 3) de la définition $D_i(\partial_j)$ est un champs de vecteurs. On peut donc lui aussi l'écrire en coordonnées locales dans la base $(\partial_k)_{k\in[\![1,n]\!]}$:

$$D_i(\partial_j) = \Gamma^k_{i\,j}\partial_k$$

Pour chaque connexion il existe donc n^3 fonctions $\Gamma^k_{i\,j}$ appelées symboles de Christoffel de la connexion. Les symboles de Christoffel sont donc moralement les "coefficients" de la connexion. A l'inverse, on peut montrer que se donner n^3 fonctions $\Gamma^k_{i\,j}$ permet de construire une connexion par la formule :

$$D_XY = \left(X^i\partial_iY^k + X^iY^j\Gamma^k_{i\,j}\right)\partial_k$$

Une connexion est donc totalement déterminée par ses symboles de Christoffel. Maintenant que l'on a défini ce qu'est une connexion nous sommes prêts à parler de quantités reliées à une connexion: sa courbure et sa torsion.

<u>Définition :</u>(torsion d'une connexion)
Soit D une connexion sur M. La torsion de D est l'application $T : \Gamma(M)\times\Gamma(M) \to \Gamma(M)$ définie par :

$$T(X,Y) = D_XY - D_YX - [X,Y]$$

On peut comme d'habitude exprimer $T(X,Y)$ en coordonnées locales.

$$\begin{aligned}
T(X,Y) &= D_XY - D_YX - [X,Y] \\
&= (XY^k + X^iY^j\Gamma^k_{i\,j})\partial_k - (YK^k + Y^iX^j\Gamma^k_{i\,j})\partial_k - [X,Y] \\
&= XY - YX + X^iY^j(\Gamma^k_{i\,j} - \Gamma^k_{j\,i})\partial_k - [X,Y] \\
&:= X^iY^jT^k_{i\,j}\partial_k
\end{aligned}$$

Si on considère la restriction de T à $T_PM \times T_PM$ pour un $P \in M$ définie par :

$$\begin{aligned}
T : T_PM \times T_PM &\to T_PM \\
(X_P, Y_P) &\mapsto T^i_{j\,k}(P)X^j_PY^k_P(\partial_i)_P
\end{aligned}$$

On voit que les $T^i_{j\,k}(P)$ sont les composantes d'un tenseur de type $(2,1)$ (en anticipant les notions de la prochaine section) puisqu'il transforme un couple de vecteurs en un vecteur.

On peut maintenant définir la courbure :

<u>Définition :</u> (courbure d'une connexion)
Soit D une connexion sur M. On définit alors sa courbure comme :

$$\begin{aligned} R : \Gamma(M) \times \Gamma(M) &\to \mathrm{Hom}(\Gamma(M), \Gamma(M)) \\ (X, Y) &\mapsto D_X \circ D_Y - D_Y \circ D_X - D_{[X,Y]} \end{aligned}$$

Cette fois le calcul pour exprimer $R(X,Y)(Z)$ en coordonnées locales est un peu plus lourd. Commencons par exprimer le premier terme :

$$\begin{aligned} D_X \circ D_Y(Z) &= D_X(D_Y(Z)) \\ &= \big(XD_Y(Z)^k + X^i D_Y(Z)^j \Gamma^k_{i,j}\big)\partial_k \\ &= X\big(YZ^k + Y^i Z^j \Gamma^k_{i\,j}\big)\partial_k + X^i\big(YZ^j + Y^m Z^l \Gamma^j_{m\,l}\big)\Gamma^k_{i\,j}\partial_k \end{aligned}$$

De même on obtient facilement :

$$D_Y \circ D_X(Z) = Y\big(XZ^k + X^i Z^j \Gamma^k_{i\,j}\big)\partial_k + Y^i\big(XZ^j + X^m Z^l \Gamma^j_{m\,l}\big)\Gamma^k_{i\,j}\partial_k$$

Et le troisième terme s'écrit :

$$D_{[X,Y]}(Z) = \big([X,Y]Z^k + [X,Y]^i Z^j \Gamma^k_{i\,j}\big)\partial_k$$

En injectant ces expressions dans la définition de la courbure on voit que certains termes se compensent pour ne laisser que :

$$R(X,Y)(Z) = \Big(XY^i Z^j \Gamma^k_{i\,j} - YX^i Z^j \Gamma^k_{i\,j} + X^i Y^m Z^l \Gamma^j_{m\,l}\Gamma^k_{i\,j} - Y^i X^m Z^l \Gamma^j_{m\,l}\Gamma^k_{i\,j}\Big)\partial_k$$

Finalement, en renommant les indices on arrive à l'expression suivante pour la courbure :

$$\begin{aligned} R(X,Y)(Z) &= X^m Y^i Z^l \big(\Gamma^j_{i\,l}\Gamma^k_{m\,j} - \Gamma^j_{m\,l}\Gamma^k_{i\,j} + \partial_m \Gamma^k_{i\,l} - \partial_i \Gamma^k_{m\,l}\big)\partial_k \\ &:= R^k_{mil} X^m Y^i Z^l \partial_k \end{aligned}$$

Les R^k_{mil} sont donc les composantes d'un tenseur de type $(3,1)$ puisqu'il transporme une famille de 3 vecteurs en 1 vecteur. Ce tenseur, appelé tenseur de Riemann, est central dans la théorie de la relativité générale. En effet, sa contraction $R^i_{j\,k\,i}$ que l'on appelle tenseur de Ricci intervient directement dans l'équation de champs d'Einstein et s'interprète comme une déformation de l'espace-temps.

1.2 Tenseurs

1.2.1 Définition et opérations sur les tenseurs

Dans ce chapitre on introduit le langage des tenseurs, qu'il est indispensable de connaitre pour pouvoir faire de la physique.
Dans ce chapitre E est un espace vectoriel de dimension finie n et on note E^* son dual. Comme E est de dimension finie on peut faire l'identification $E \simeq E^{**}$.

On peut alors définir un tenseur comme :

Définition : (tenseurs)
Soient p, q des entiers positifs. Un tenseur de type (p, q) (ou p fois contravariant et q fois covariant) est une forme linéaire agissant sur :

$$\underbrace{E \times ... \times E \times E}_{p} \times \underbrace{E^* \times ... \times E^*}_{q}$$

Par exemple, un élément de E^*, c'est-à-dire une forme linéaire sur E est un tenseur de type $(1, 0)$ tandis qu'un élément de E est un tenseur de type $(0, 1)$. Par suite, on peut définir le produit tensoriel de deux tenseurs :

Définition : (produit tensoriel)
Soit T_1 un tenseur de type (p_1, q_1) et T_2 un tenseur de type (p_2, q_2). On définit leur produit tensoriel $T_1 \otimes T_2$ comme un le tenseur de type $(p_1 + p_2, q_1 + q_2)$ tel que :

$$T_1 \otimes T_2(x, y) = T_1(x)T_2(y)$$

où x et y sont dans les bons espaces (c'est-à-dire x est dans l'espace sur lequel T_1 agit et y dans l'espace sur lequel T_2 agit).

On peut se donner une base $(e_1, ...e_n)$ de E et $(e^1, ..., e^n)$ une base de E^*. On peut alors se convaincre que l'ensemble des tenseurs de rang (p, q) est un espace vectoriel de dimension n^{p+q} dont une base est :

$$e^{i_1} \otimes ... \otimes e^{i_p} \otimes e_{j_1} \otimes ... \otimes e_{j_q}$$

Pour donner un exemple, un tenseur T de type $(1, 2)$ peut s'écrire sous la forme suivante :

$$T = \sum_{i,j,k \in [\![1,n]\!]} T_i^{jk} e^i \otimes e_j \otimes e_k = T_i^{jk} e^i \otimes e_j \otimes e_k$$

Une opération importante sur les tenseurs est la contraction. Le produit tensoriel permet à partir de deux tenseurs d'en créer un troisième de rang plus élevé, à l'inverse la contraction permet de diminuer le rang d'un tenseur. On la définit comme suit :

Définition : (contraction)
Soit T un tenseur de rang (p, q). La contraction de rang (k, l) (avec $k \in [\![1, p]\!]$ et $l \in [\![1, q]\!]$) de T est le tenseur de rang $(p - 1, q - 1)$ noté $C_l^k T$ où, dans une base, on a égalisé le $k-$ième indice en bas et le $l-$ième indice en haut et on a sommé sur les valeurs possibles.

Pour clarifier cette définition illustrons, la par un exemple. Soit T un tenseur de type $(2, 2)$. On peut l'écrire dans une certaine base comme :

$$T = T^{ij}{}_{kl} e_i \otimes e_j \otimes e^k \otimes e^l$$

Alors sa contraction $(1, 1)$ s'écrit :

$$C_1^1 T = T^{ij}{}_{il} e_j \otimes e^l$$

C'est cette contraction que l'on utilise en relativité générale pour construire le tenseur de Ricci utilisé dans l'équation d'Einstein à partir du tenseur de Riemann.

1.2.2 Champs de tenseurs et dérivée covariante

On sait que pour une variété M de dimension n on peut construire son fibré tangent TM que l'on peut munir d'une structure de variété de dimension $2n$ et que la projection $\pi : TM \to M$ le munit d'une structure de fibré vectoriel. On peut en fait généraliser cette construction. Soit $x \in M$ et soit T_xM l'espace tangent en x. On peut alors considérer l'espace des tenseurs de rang (p, q) sur T_xM, que l'on note $\overset{(p,q)}{\bigotimes} T_xM$. On peut alors définir :

<u>Définition :</u> (fibré des tenseurs)
On définit le fibré des tenseurs d'une variété M comme :

$$\overset{(p,q)}{\bigotimes} M = \bigcup_{x \in M} \overset{(p,q)}{\bigotimes} T_xM$$

On peut alors se convaincre qu'en procédant de la même manière que pour le fibré tangent on peut munir le fibré des tenseurs d'une structure de variété (de dimension $n + n^{p+q}$) et d'une structure de fibré vectoriel. On définit alors simplement un champs de tenseurs comme :

<u>Définition :</u> (champs de tenseurs)
Un champs de tenseurs est une section du fibré des tenseurs.

La notion de champs de tenseurs est également très importante en physique puisqu'elle est utilisée dans de nombreux domaines. On peut par exemple citer la relativité générale où la métrique est un champs de tenseurs de type $(0, 2)$ ou bien la physique des milieux continus où les contraintes appliquées sur un système sont décrites par le champs de tenseurs des contraintes.

On peut alors étendre la dérivée covariante, initialement définie pour les champs de vecteurs aux tenseurs :

<u>Définition :</u> (dérivée covariante pour les tenseurs)
Pour étendre la dérivée covariante aux tenseurs on demande naturellement que cette définition coïncide avec la définition initiale dans le cas d'un champs de vecteurs. On demande de plus les propriétés suivantes :
1) Soit f une fonction lisse sur M. Alors $D_X(f) = X(f)$
2) D_X ne change pas le type d'un tenseur
3) D_X commute avec la contraction
4) Si T_1 et T_2 sont deux champs de tenseurs alors :

$$D_X(T_1 \otimes T_2) = D_X(T_1) \otimes T_2 + T_1 \otimes D_X(T_2)$$

Cette définition nous permet donc de considérer les dérivées covariantes non seulement de champs de vecteurs, mais aussi de champs de formes ou de champs de tenseurs de rang quelconque.

2 Physique

2.1 L'Electromagnétisme

2.1.1 Reformulation à la Yang-Mills

Dans ce chapitre, on va exprimer la théorie de l'électromagnétisme grâce aux objets que l'on a défini précédemment. On se place dans $(\mathbb{R}^4, \eta)$ où η est la métrique définie par (en anticipant les définitions du prochain chapitre):

$$\eta = -\mathrm{d}x_0^2 + \sum_{k=1}^{n} \mathrm{d}x_k^2$$

Cet espace est simplement noté $\mathbb{R}^{3,1}$. Dans cette formulation de l'électromagnétisme, le champs électromagnétique est représenté par une 2-forme sur $\mathbb{R}^{3,1}$, notée F. Elle est donc de la forme :

$$F = \frac{1}{2} F_{ab} \mathrm{d}x^a \wedge \mathrm{d}x^b$$

avec la condition $F_{ab} = -F_{ba}$. On définit même :

$$F_{|x_0=t} = \begin{pmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & B_3 & -B2 \\ -E_2 & -B_3 & 0 & B_1 \\ -E_3 & B_2 & -B_1 & 0 \end{pmatrix}$$

où E et B sont respectivements les champs électriques et magnétiques. F est souvent appelé tenseur électromagnétique ou tenseur de Maxwell.
Les équations de Maxwell, connues depuis le 19ème siècle, s'écrivent habituellement :

$$\begin{cases} \mathrm{div}(B) = 0 \\ \dfrac{\partial B}{\partial t} + \mathrm{rot}(E) = 0 \\ \mathrm{div}(E) = \rho \\ -\dfrac{\partial E}{\partial t} + \mathrm{rot}(B) = J \end{cases}$$

où ρ et J sont respectivement la densité de charge et le courant électrique du milieu dans lequel on se place. Dans le vide ces quantités sont égales à 0. Notre but est de réexprimer ces équations dans le langage des formes.
On connait déjà pour des formes α et β les notions de produit extérieur, noté $\alpha \wedge \beta$ et de dérivée extérieure, notée par d. On peut alors définir le produit scalaire entre deux formes de même degré k à l'aide de la métrique :

$$< \alpha, \beta > := \eta^{i_1 j_1} ... \eta^{i_k j_k} \alpha_{i_1 ... i_k} \beta_{j_1 ... j_k}$$

On définit également l'opérateur adjoint de la différentielle, noté δ ou d^* par :

$$\int < \mathrm{d}\alpha, \beta > dv = \int < \alpha, \delta\beta > dv$$

On peut montrer que δ vérifie :

$$(\delta F)_a = -\eta^{bc} \partial_c F_{ab}$$

On veut désormais montrer que les équations de Maxwell se réduisent à :

$$\begin{cases} \mathrm{d}F = 0 \\ \delta F = j \end{cases}$$

où j est la 1-forme définie par : $j = \rho \mathrm{d}x^0 + J_i \mathrm{d}x^i$. Commencons par la deuxième équation. D'après la remarque précédente on a pour la composante 0 :

$$\begin{aligned}(\delta F)_0 &= -\eta^{ab}\partial_b F_{0a} \\ &= -\underbrace{\partial_t F_{00}}_{=0} + \underbrace{\partial_1 F_{01} + \partial_2 F_{02} + \partial_3 F_{03}}_{=\mathrm{div}(E)}\end{aligned}$$

et $j_0 = \rho$, donc la composante 0 de la deuxième équation nous donne bien la troisième équation de Maxwell, dite équation de Maxwell-Gauss.
Traitons les composantes 1,2 et 3 de la seconde équation. Prenons la composante 1, les deux autres étant similaires. On a alors :

$$\begin{aligned}(\delta_1 F)_1 &= -\eta^{ab}\partial_b F_{1a} \\ &= -\partial_t F_{10} + \partial_1 F_{11} + \partial_2 F_{12} + \partial_3 F_{13} \\ &= \partial_t E_1 + \partial_2 B_3 - \partial_3 B_2\end{aligned}$$

qui est bien la première composante de la quatrième équation de Maxwell, dite équation de Maxwell-Ampère. Le même traitement pour les deux autres composantes nous montrera que la deuxième équation est bien équivalente aux équations 2 et 3 de Maxwell.

Il reste désormais à traiter le cas de la première équation. On sait que F est de la forme suivante :

$$F = \frac{1}{2} F_{ab} \mathrm{d}x^a \wedge \mathrm{d}x^b$$

Sa différentielle est donc de la forme suivante :

$$\mathrm{d}F = \frac{1}{2}\partial_c(F_{ab})\mathrm{d}x^c \wedge \mathrm{d}x^a \wedge \mathrm{d}x^b$$

Commencons par exprimer F de la façon suivante :

$$F = E_1 \mathrm{d}t \wedge \mathrm{d}x^1 + E_2 \mathrm{d}t \wedge \mathrm{d}x^2 + E_3 \mathrm{d}t \wedge \mathrm{d}x^3 + B_3 \mathrm{d}x^1 \wedge \mathrm{d}x^2 - B_2 \mathrm{d}x^1 \wedge \mathrm{d}x^3 + B_1 \mathrm{d}x^2 \wedge \mathrm{d}x^3$$

rappelons également que les formes anticommutent : $\alpha \wedge \beta = -\beta \wedge \alpha$, donc $\alpha \wedge \alpha = 0$. Pour alléger les notations, on ne notera plus les wedges. Il nous faut donc calculer la différentielle $\mathrm{d}F$. Commencons par analyser les termes qui ne contiendront pas de $\mathrm{d}t$, qui ne peuvent être que les termes avec des B_i d'après le rappel ci dessus.
On a alors :

$$\begin{aligned}\text{termes sans } \mathrm{d}t &= \partial_3 B_3 \mathrm{d}x^3 \mathrm{d}x^1 \mathrm{d}x^2 - \partial_2 B_2 \mathrm{d}x^2 \mathrm{d}x^1 \mathrm{d}x^3 + \partial_1 B_1 \mathrm{d}x^1 \mathrm{d}x^2 \mathrm{d}x^3 \\ &= \partial_3 B_3 \mathrm{d}x^1 \mathrm{d}x^2 \mathrm{d}x^3 + \partial_2 B_2 \mathrm{d}x^1 \mathrm{d}x^2 \mathrm{d}x^3 + \partial_1 B_1 \mathrm{d}x^1 \mathrm{d}x^2 \mathrm{d}x^3 \\ &= \mathrm{div}(B) \mathrm{d}x^1 \mathrm{d}x^2 \mathrm{d}x^3\end{aligned}$$

Si $\mathrm{d}F = 0$ cela nous donne bien la première équation de Maxwell, dite équation de Maxwell-Thomson.
Penchons nous maintenant sur tous les autres termes. Ce sont donc ceux contenant $\mathrm{d}t$ et sont donc de la forme $\mathrm{d}t\mathrm{d}x^1\mathrm{d}x^2$ ou $\mathrm{d}t\mathrm{d}x^1\mathrm{d}x^3$ ou $\mathrm{d}t\mathrm{d}x^2\mathrm{d}x^3$. Analysons les termes de la forme $\mathrm{d}t\mathrm{d}x^1\mathrm{d}x^2$, les autres étant similaires.

On a alors :

$$\begin{aligned}\text{termes en } \mathrm{d}t\mathrm{d}x^1\mathrm{d}x^2 &= \partial_2 E_1 \mathrm{d}x^2\mathrm{d}t\mathrm{d}x^1 + \partial_1 E_2 \mathrm{d}x^1\mathrm{d}t\mathrm{d}x^2 + \partial_t B_3 \mathrm{d}t\mathrm{d}x^1\mathrm{d}x^2 \\ &= \partial_2 E_1 \mathrm{d}t\mathrm{d}x^1\mathrm{d}x^2 - \partial_1 E_2 \mathrm{d}t\mathrm{d}x^1\mathrm{d}x^2 + \partial_t B_3 \mathrm{d}t\mathrm{d}x^1\mathrm{d}x^2\end{aligned}$$

Si $\mathrm{d}F = 0$ cela nous donne alors la troisième composante de la deuxième équation de Maxwell, dite équation de Maxwell-Faraday. En faisant de même avec les termes de la forme $\mathrm{d}t\mathrm{d}x^1\mathrm{d}x^3$ et $\mathrm{d}t\mathrm{d}x^2\mathrm{d}x^3$ on retrouve l'équation complète. On a donc montré qu'en utilisant le formalisme des formes on peut ramener les 4 équations de Maxwell à un système de deux équations :

$$\begin{cases} \mathrm{d}F = 0 \\ \delta F = j \end{cases}$$

Comme $\mathbb{R}^4$ est simplement connexe, la relation $\mathrm{d}F = 0$ implique qu'il existe A tel que $F = \mathrm{d}A$. Un tel A est souvent appelé potentiel électromagnétique. Il est important de remarquer que A est défini à une différentielle près, puisque $\mathrm{d}A = \mathrm{d}(A + \mathrm{d}f)$ car $\mathrm{d}^2 = 0$.

2.1.2 Approche variationnelle

Les théories physiques sont souvent formulées en terme de ce que les physiciens appellent une action (une fonctionnelle pour les mathématiciens) dont les équations d'Euler-Lagrange fournissent les équations physiques. On peut donc formuler les équations de la partie précédente en terme d'une fonctionnelle. On définit la fonctionnelle la plus simple de F comme :

$$\begin{array}{rll} \mathcal{L} : & \left\{F \in \Omega^2(\mathbb{R}^{3,1}) \middle| \mathrm{d}F = 0\right\} & \to \mathbb{R} \\ & F & \mapsto \dfrac{1}{4}\displaystyle\int_{\mathbb{R}^{3,1}} < F, F > dv \end{array}$$

où $\Omega^2(M)$ désigne les fonctions de carré intégrable sur M.
Cette quantité $\mathcal{L}(F)$ est appelée action de Yang-Mills par les physiciens. On peut facilement calculer $< F, F >$ grâce à la définition du produit scalaire pour les formes et on trouve alors :

$$\mathcal{L}(F) = \frac{1}{4}\int_{\mathbb{R}^{3,1}} < F, F > dv = \frac{1}{2}\int_{\mathbb{R}^4} |B|^2 - |E|^2 dv$$

On peut montrer que les équations d'Euler-Lagrange pour cette fonctionnelle implique que :

$$\delta F = 0$$

ce qui donne bien les équations de Maxwell dans le vide. Cette formulation en terme d'une fonctionnelle permet par exemple de remarquer que la théorie de Kaluza-Klein (la relativité générale avec 5 dimensions, 4 d'espace dont une enroulée et une de temps) permet une unification de la relativité générale à 4 dimension que l'on connait et de l'électromagnétisme tel que nous l'avons écrit précédemment.

2.2 Théorie de Yang-Mills

Dans cette partie on énonce rapidement ce que sont les théories de Yang-Mills, théories absolument fondamentales dans la physique théoriques modernes. Cela nous permettra de remarquer que l'électromagnétisme peut-être vu comme un exemple de théorie de Yang-Mills.

On a défini dans les parties précédentes la dérivée covariante pour les fibrés vectoriels. Si on utilise la notation :

$$D_V(s) := Ds(V)$$

cela nous permet alors de parler de l'objet Ds pour un champs de vecteurs s. On peut alors exprimer cet objet en coordonnées locales, disont autour d'un point $x \in M$. Dans ces coordonnées on peut écrire $s = a^k \partial_k$. On a alors :

$$Ds = \mathrm{d}a^k \, \partial_k + a^k D\partial_k$$

et comme $D\partial_k$ est un champs de vecteurs on peut l'exprimer comme :

$$D\partial_k = A_k^j \partial_j$$

On remarque alors que les A_k^j doivent être des matrices à coefficients dans T_x^*M, c'est-à-dire des éléments de $\mathfrak{gl}(n, \mathbb{R}) \otimes T_x^*M$. On peut donc écrire que dans ces coordonnées locales on a :

$$D = \mathrm{d} + A$$

On a donc définit une connexion D sur un fibré vectoriel. On peut alors définir une connexion D^* sur le fibré dual par :

$$D^* = \delta + A^*$$

Si on note E un fibré vectoriel cela permet alors de construire une connexion sur $E \otimes E^*$. On note μ_i et μ^{*j} des bases de E et E^*. Pour une section s de $E \otimes E^*$ on a alors $s = s_j^i \mu_i \otimes \mu^{*j}$ et on a alors :

$$D(s_j^i \mu_i \otimes \mu^{*j}) = D(s) = \mathrm{d}s + [A, s]$$

On peut désormais définir la courbure d'une connexion comme :

<u>Définition :</u>(courbure d'une connexion)
Avec les notations précédentes on définit la courbure F d'une connexion D comme :

$$F = D^2$$

Exprimons la courbure plus simplement. Soit μ une section de E. Alors :

$$\begin{aligned}
F(\mu) &= (\mathrm{d}+A)^2\mu \\
&= (\mathrm{d}+A)(\mathrm{d}\mu + A\mu) \\
&= \mathrm{d}^2\mu + \mathrm{d}(A\mu) + A\mathrm{d}\mu + A\wedge A\mu \\
&= -A\mathrm{d}\mu + \mathrm{d}(A)\mu + A\mathrm{d}\mu + A\wedge A\mu \\
&= \mathrm{d}(A)\mu + A\wedge A\mu
\end{aligned}$$

On a donc simplement :

$$F = \mathrm{d}A + A\wedge A$$

Un calcul simple montre alors que :

$$DF = \mathrm{d}F + [A,F] = 0$$

cette identité est connue sous le nom d'identité de Bianchi.

On considère désormais un fibré principal P, c'est-à-dire un fibré au dessus de M dans lequel les fibres ont une structure de groupe de lie G, qui sera interprété comme le groupe d'invariance de notre théorie physique. On se donne une application :

$$A : TM \to \mathfrak{g}$$

où $\mathfrak{g}$ est l'algèbre de Lie de G. D'après ce qui précéde on sait que :

$$D = \mathrm{d} + A$$

est une connexion sur M. A est alors appelée 1-forme de connexion pour G. On peut alors définir la fonctionnelle de Yang-Mills pour une variété (Riemannienne ou Lorentzienne en anticipant les définitions de la prochaine section) avec une métrique g par :

$$\mathcal{L}_{YM}(D) = \int_M < F, F > dv_g$$

On a déja montré que pour toute connexion D on a $DF = 0$ et on peut montrer que l'équation d'Euler-Lagrange de cette fonctionnelle est :

$$D^*F = 0$$

ce qui conduit au système d'équations dit de Yang-Mills :

$$\begin{cases} DF = 0 \\ D^*F = 0 \end{cases}$$

Il est maintenant possible d'exprimer l'électromagnétisme tel qu'on l'a écrit à la partie précédente comme une théorie de Yang-Mills. On se place dans le cas où $M = \mathbb{R}^{3,1}$ et on veut que le groupe d'invariance de notre théorie physique soit $U(1)$. L'algèbre de Lie de $U(1)$ n'est autre que $i\mathbb{R} \simeq \mathbb{R}$. On se donne alors une application :

$$A : TM \to \mathbb{R}$$

qui vérifie les équations de Yang-Mills. Calculons la courbure dans notre cas. On sait que $F = \mathrm{d}A + A\wedge A$. Ici $A = A_i\mathrm{d}x^i$ avec $A_i \in \mathbb{R}$.

On a alors :

$$\begin{aligned}
A \wedge A &= A_i \mathrm{d}x^i \wedge A_j \mathrm{d}x^j \\
&= A_i A_j \mathrm{d}x^i \wedge \mathrm{d}x^j \\
&= \sum_{i>j} A_i A_j \mathrm{d}x^i \wedge A_j \mathrm{d}x^j + \sum_{j>i} A_i A_j \mathrm{d}x^i \wedge A_j \mathrm{d}x^j \\
&= \sum_{i>j} A_i A_j \mathrm{d}x^i \wedge A_j \mathrm{d}x^j - \sum_{j>i} A_j A_i \mathrm{d}x^j \wedge A_j \mathrm{d}x^i \\
&= \sum_{i>j} A_i A_j \mathrm{d}x^i \wedge A_j \mathrm{d}x^j - \sum_{i>j} A_i A_j \mathrm{d}x^i \wedge A_j \mathrm{d}x^j = 0
\end{aligned}$$

à l'avant dernière étape on utilise l'anti-commutativité du produit extérieur et à la dernière ligne on renomme i et j dans la seconde somme. Dans notre cas la courbure s'écrit alors $F = \mathrm{d}A$. On montre de même que l'on a dans ce cas $D = \mathrm{d}$ et $D^* = \delta$. Les équations de Yang-Mills pour le groupe d'invariance $U(1)$ sont donc :

$$\begin{cases} dF = 0 \\ \delta F = 0 \end{cases}$$

qui sont exactement les équations de Maxwell dans le vide telles qu'on les a écrites au chapitre précédent. C'est en ce sens que les physiciens appellent l'électromagnétisme "théorie de Jauge $U(1)$" (A est alors appelé le potentiel de jauge). Il est important de préciser que les simplifications que l'on a pu faire ici sont dues au fait que le groupe $U(1)$ est commutatif, ce qui n'est évidemment pas le cas pour toutes les théories de Yang-Mills. Par exemple le modèle standard de la physique des particules est obtenu comme la théorie de Yang-Mills correspondante au groupe d'invariance $G = SU(3) \times SU(2) \times U(1)$.

2.3 Relativité générale et géométrie

Dans ce chapitre on présente la théorie de la relativité générale qui est une théorie de la gravitation. Pour ce faire, nous aurons besoin de définir les objets en jeu dans cette théorie. L'espace temps est, en relativité générale, une variété Lorentzienne de dimension 4. Nous aurons donc besoin de définir les concepts de métrique puis de métrique Lorentzienne. Cela nous conduira jusqu'à la définition d'une connexion particulière sur les variétés Lorentzienne, dite connexion de Levi-Cevita, qui se comporte bien par rapport à la métrique.

Définition : (métrique)
On appelle métrique sur une variété différentielle M un champs de tenseur g de type $(0,2)$ tel que :
1) g est $C^\infty(M)$ bilinéaire. c'est à dire que pour toute fonction lisse f sur M on a :

$$g(fX + Y, Z) = fg(X,Z) + g(Y,Z) \text{ et } g(X, fY + Z) = fg(X,Y) + g(X,Z)$$

2) g est symétrique
3) g est définie, c'est-à-dire :

$$g(X,Y) = 0 \text{ pour tout } Y \Rightarrow X = 0$$

On peut exprimer une métrique en coordonnées locales de la façon suivante :

$$g = g_{ij} \mathrm{d}x^i \wedge \mathrm{d}x^j$$

On peut alors définir des cas particuliers importants de métriques, les métriques Riemannienne et Lorentzienne :

Définition : (métrique Riemannienne, Lorentzienne)
-Une métrique g sur une variété M de dimension m est dite Riemanienne si elle est de signature constante égale à $(m, 0)$ ou $(0, m)$ (c'est-à-dire que les valeurs propres des $g(x), x \in M$ ont toutes le même signe)
-Une métrique g sur une variété M de dimension m est dite Lorentzienne si elle est de signature constante égale à $(m-1, 1)$ ou $(1, m-1)$ (les valeurs propres des $g(x), x \in M$ sont à chaque fois toutes du même signe, sauf une).

Cela nous permet les définitions naturelles :

Définition : (variété Riemannienne, Lorentzienne)
-Une variété Riemannienne est une variété munie d'une métrique Riemannienne.
-Une variété Lorentzienne est une variété munie d'une métrique Lorentzienne.

Un exemple typique de variété Lorentzienne est donné par l'espace $\mathbb{R}^{3,1}$ du chapitre précédent où la métrique η est donné par :

$$\eta = -\mathrm{d}x_0^2 + \sum_{k=1}^{n} \mathrm{d}x_k^2$$

Cet espace $\mathbb{R}^{3,1}$ est appelé espace de Minkowski. On trouve souvent la notation $\eta = \mathrm{diag}(-1, 1, 1, 1)$ et la signature de la métrique est alors $(3, 1)$.

On peut également donner un exemple simple de variété Riemannienne : l'espace classique $\mathbb{R}^n$ avec sa métrique euclidienne standard. On y définit des coordonnées globales $(x_1, ..., x_n)$ et on définit la métrique g par $g(e_i, e_j) = \delta_i^j$ où les e_i forment la base canonique et δ_i^j est le symbole de kronecker.

Nous avons vu précédemment que l'on pouvait toujours définir une connexion sur une variété, mais que pour une variété quelconque donnée il n'existe pas une unique connexion. Cependant, dans le cadre Lorentzien on peut montrer qu'il n'existe qu'une seule connexion qui se comporte agréablement par rapport à la métrique. Plus précisément, on a le résultat suivant :

Théorème :
Soit (M, g) une variété Lorentzienne. Soient X, Y, Z des champs de vecteurs sur M. Alors il existe une unique connexion ∇ telle que :

$$Z.g(X, Y) = g(\nabla_Z X, y) + g(X, \nabla_Z Y)$$

et telle que la torsion de ∇ est nulle.

Comme on impose une condition entre la connexion et la métrique, on peut se douter qu'il existe une relation entre les coefficients de Christoffel de la connexion et la métrique. Commençons par remarquer que la condition de nullité de la torsion implique, par définition de la torsion, que les symboles de Christoffel vérifient :

$$\Gamma_{ij}^k = \Gamma_{ji}^k$$

Pour trouver cette relation, on peut utiliser notre condition dans le cas $Z = \partial_i, X = \partial_j$ et $Y = \partial_k$.

On a alors la relation suivante :

$$\begin{aligned}\partial_k g(\partial_i, \partial_j) := \partial_k g_{ij} &= g(\nabla_k \partial_i, \partial_j) + g(\partial_i, \nabla_k \partial_j) \\ &= g(\Gamma^l_{ki}\partial_l) + g(\partial_i, \Gamma^l_{kj}\partial_l) \\ &= \Gamma^l_{ki} g_{lj} + \Gamma^l_{kj} g_{il}\end{aligned}$$

En faisant une permutation des indices on remarque que :

$$\partial_i g_{jk} = \Gamma^l_{ij} g_{lk} + \Gamma^l_{ik} g_{jl} \text{ et } \partial_j g_{ki} = \Gamma^l_{jk} g_{li} + \Gamma^l_{kj} g_{kl}$$

En utilisant que la métrique est symétrique, donc $g_{ij} = g_{ji}$, que la connexion est sans torsion, et en sommant les deux premières égalités et en soustrayant la dernière, il vient :

$$\Gamma^l_{ik} = \frac{1}{2} g^{lj} \left(\partial_i g_{jk} + \partial_k g_{ij} - \partial_j g_{ki}\right)$$

où les g^{ij} sont définis par $g_{ik} g^{kj} = \delta^j_i$.

On connait maintenant les objets importants pour pouvoir parler de relativité générale.On peut énoncer les postulats de la relativité comme suit :
- L'espace temps est une variété Lorentzienne à 4 dimensions (M, g).
- La position d'une particule au repos vérifie l'équation des géodésiques :

$$x''^{\lambda}(t) + \Gamma^{\lambda}_{\mu\nu} x'^{\mu} x'^{\nu} = 0$$

- La métrique g vérifie l'équation d'Einstein.

Comme l'idée de base de la relativité générale est qu'il n'y a en fait pas de force de gravité mais que la chute des corps est due à une déformation de l'espace temps, on voit alors que l'équation d'Einstein doit être une équation portant sur la géométrie de l'espace temps, donc une équation sur la métrique. Comme pour les théories de Yang-Mills, on peut formuler la relativité générale grâce à une formlation variationnelle. Une action naturelle à poser est :

$$\mathcal{L}(g) = \int_M R \, dv_g$$

Cette action est appelée action d'Hilbert-Einstein. R est la courbure scalaire, obtenue par contraction du tenseur de Riemann défini précédemment et dv_g est l'élément de volume associé à la métrique : $dv_g = \sqrt{-\det(g)} dx$. On peut alors calculer l'équation d'Euler-Lagrange de cette action, qui donne l'équation de champs d'Einstein dans le vide:

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 0$$

2.4 Théorie de Kaluza-Klein

Les deux sections précédentes constituaient des introductions aux théories physiques importantes que sont l'électromagnétisme et la relativité générale. Dans cette partie on présente la théorie de Kaluza-Klein, qui est la première théorie à avoir tenté d'unifier la gravitation et l'électromagnétisme. L'idée de cette théorie est de réécrire la relativité générale non pas sur une variété Lorentzienne M mais sur un fibré principale $\overline{M}$ de fibre $\mathbb{S}^1$ sur M. On sait que M est équipée de sa métrique g et que l'on peut se donner une 1-forme de connexion A pour $U(1) \simeq \mathbb{S}^1$ sur M. On définit la projection $\pi : \overline{M} \to M$. On peut alors définir une métrique sur $\overline{M}$ par :

$$\overline{g} = \pi^* g + \pi^* A \otimes \pi^* A$$

où π^* est le pullback de π défini par $\pi^* g = g \circ \pi$. On définit alors la courbure pour la forme A par :

$$F = \mathrm{d}A$$

(il n'y a pas de terme en $A \wedge A$ car A est une 1-forme de connexion pour $U(1)$). On peut alors calculer la courbure scalaire pour cette métrique en utilisant les formules pour le tenseur de courbure et vérifier que la courbure scalaire $\overline{R}$ pour $\overline{M}$ vérifie :

$$\overline{R} = R - |F|^2 := R - < F, F >$$

l'action de Hilber-Einstein pour $\overline{M}$ s'écrit :

$$\mathcal{L}(\overline{g}) = \int\limits_{\overline{M}} \overline{R} \sqrt{-\mathrm{det}\overline{g}}\, dt\, dx^1\, dx^2\, dx^3\, d\xi$$

En utilisant ce qui précède et en intégrant sur $\mathbb{S}^1$ (en supposant que la mesure de $\mathbb{S}^1$ est 1 pour simplifier) on a alors :

$$\mathcal{L}(\overline{g}) = \int\limits_{M} (R - |F|^2) \sqrt{-\mathrm{det} g}\, dx$$

On remarque alors que cette fonctionnelle contient celle d'Hilbert-Einstein et celle de Yang-Mills pour l'électromagnetisme : on a bien un couplage entre la gravité et l'électromagnétisme.

References

[1] Paul Laurain. Introduction to the mathematics of general relativity. 2019.

[2] Theodore Frankel. The geometry of physics an introduction. 2001.

[3] Jürgen Jost. Geometry and physics. 2009.

[4] Thierry Aubin. A course in differential geometry. 2001.