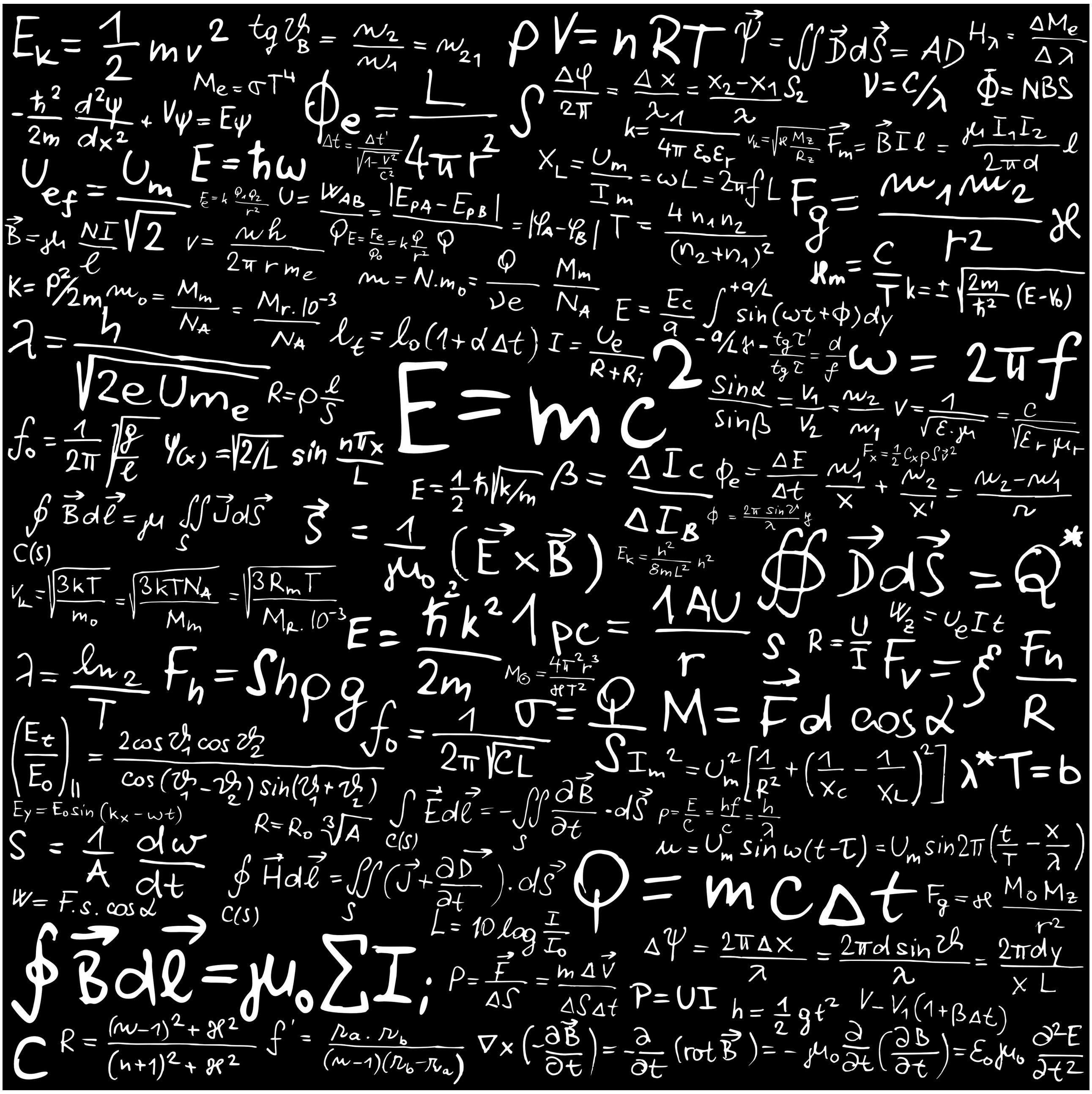
Petit rappel à l'intention des participants à l'Atelier de culture scientifique et technique sur la relativité générale



Différentielle d'une fonction de plusieurs variables et dérivées partielles

Une fonction f de n variables réelles $(x_1, x_2, \ldots, x_n)$ à valeur réelle est **différentiable** en un point $\mathbf{x}$ de coordonnées $(x_1, x_2, \ldots, x_n)$ si sa valeur au point $\mathbf{x}+\mathbf{h}$, avec $\mathbf{h}$ un vecteur de composantes $(h_1, h_2, \ldots, h_n)$, diffère en première approximation de sa valeur au point $\mathbf{x}$ de manière linéaire en $(h_1, h_2, \ldots, h_n)$ tel que :

$$f(\mathbf{x}+\mathbf{h}) = f(\mathbf{x}) + \mathbf{c}\cdot\mathbf{h} + r(\mathbf{x},\mathbf{h}), \text{ avec } \lim_{\|\mathbf{h}\|\to 0} \frac{r(\mathbf{x},\mathbf{h})}{\|\mathbf{h}\|} = 0$$

Cette variation linéaire $\mathbf{c}\cdot\mathbf{h}$, avec $\mathbf{c}$ un vecteur de n composantes réelles, est la **différentielle** de la fonction au point $\mathbf{x}$ notée df.

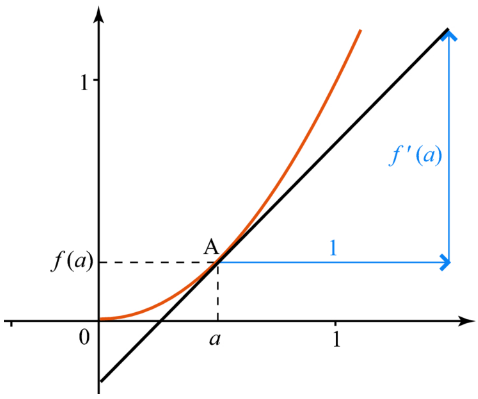
Une fonction f d'une variable réelle x à valeur réelle est différentiable en x si elle est **dérivable** en x. En effet, si :

$$f(x+h) = f(x) + c\,h + r(x,h), \text{ avec } \lim_{h\to 0}\frac{r(x,h)}{h} = 0$$

alors

$$f'(x) = \lim_{h\to 0}\frac{f(x+h)-f(x)}{h} = c + \lim_{h\to 0}\frac{r(x,h)}{h} = c$$

est la **dérivée** de la fonction f en x et la valeur de $f'(x)$ est égale à la pente de la tangente au graphe de la fonction passant par le point d'abscisse x et d'ordonnée $y = f(x)$.



Pour un accroissement h infinitésimal noté dx, la différentielle de la fonction est donnée par $df = c\,dx = f'(x)\,dx$ et la dérivée de la fonction en x s'écrit :

$$f'(x) = \left(\frac{df}{dx}\right)(x)$$

La **dérivée successive** d'ordre m de f est obtenue en dérivant successivement m fois la fonction, pour autant que cela soit possible, c'est-à-dire que les dérivées successives d'ordres inférieurs soient dérivables.

$$f^{(0)}(x) = f(x)$$

$$f^{(1)}(x) = \left(\frac{df^{(0)}}{dx}\right)(x) = \left(\frac{df}{dx}\right)(x) = f'(x)$$

$$f^{(2)}(x) = \left(\frac{df^{(1)}}{dx}\right)(x) = \frac{d}{dx}\left(\frac{df}{dx}\right)(x) = \left(\frac{d^2 f}{dx^2}\right)(x) = \left(\frac{df}{dx}\right)'(x) = f''(x)$$

$$\vdots$$

$$f^{(m)}(x) = \left(\frac{df^{(m-1)}}{dx}\right)(x)$$

En physique, lorsque la variable x correspond au temps, on utilise le point comme symbole de dérivation :

$$\dot{f}(t) = f'(t) = \left(\frac{df}{dt}\right)(t)$$

$$\ddot{f}(t) = \left(\frac{d\dot{f}}{dt}\right)(t) = \left(\frac{d^2 f}{dt^2}\right)(t)$$

Ainsi, lorsque la position d'un point matériel $\mathbf{x}(t)$ est représentée par ses trois composantes $(x(t), y(t), z(t))$, sa vitesse $\mathbf{v}(t)$ et son accélération $\mathbf{a}(t)$ sont données par :

$$\mathbf{v}(t) = \dot{\mathbf{x}}(t) = \begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \\ \dot{z}(t) \end{pmatrix} \text{ et } \mathbf{a}(t) = \ddot{\mathbf{x}}(t) = \begin{pmatrix} \ddot{x}(t) \\ \ddot{y}(t) \\ \ddot{z}(t) \end{pmatrix}$$

Pour une fonction de n variables, $df = \mathbf{grad}f(\mathbf{x}) \cdot \mathbf{h}$ où $\mathbf{grad}f(\mathbf{x})$ est le **gradient** de la fonction f au point $\mathbf{x}$. Le gradient de f est un vecteur dont les n composantes sont les n **dérivées partielles** de la fonction f au point $\mathbf{x}$ notées :

$$\left(\frac{\partial f}{\partial x_i}\right)(\mathbf{x}) = \lim_{h_i \to 0} \frac{f(x_1, x_2, \ldots, x_i + h_i, \ldots, x_n) - f(x_1, x_2, \ldots, x_i, \ldots, x_n)}{h_i} = \left(\frac{df_i}{dx}\right)(x_i),\ i = 1, \ldots, n$$

définies comme les dérivées des **fonctions partielles** $f_i(x_i)$ de f obtenues en ne faisant varier que la i^ème^ variable x_i et en figeant les valeurs des autres variables.

L'opération de dérivation partielle peut également être itérée successivement. Ainsi, les dérivées partielles de deuxième ordre s'écrivent :

$$\left(\frac{\partial^2 f}{\partial x_i^2}\right)(\mathbf{x}) = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_i}\right)(\mathbf{x}) \text{, ou encore } \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)(\mathbf{x}) = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)(\mathbf{x})$$

Pour alléger l'écriture, les physiciens réduisent parfois ces expressions de la manière suivante :

$$\left(\frac{\partial^2 f}{\partial x_i^2}\right)(\mathbf{x}) = \partial_{x_i}^2 f(\mathbf{x}) = \partial_i^2 f(\mathbf{x}) = \partial_{ii} f(\mathbf{x}) \text{, ou encore } \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)(\mathbf{x}) = \partial_{x_i x_j} f(\mathbf{x}) = \partial_{ij} f(\mathbf{x})$$

Pour des fonctions deux fois continûment différentiables, les dérivées partielles croisées ne dépendent pas de l'ordre de dérivation $\partial_{ij} f(\mathbf{x}) = \partial_{ji} f(\mathbf{x})$.

Le gradient de f peut s'écrire à l'aide d'un opérateur différentiel noté ∇ qui est un vecteur dont les composantes sont les opérateurs de dérivation partielle $(\partial_1, \partial_2, \ldots, \partial_n)$:

$$\mathbf{grad}f(\mathbf{x}) = \nabla f(\mathbf{x})$$

Pour un vecteur d'accroissement $\mathbf{h}$ infinitésimal de composantes $(dx_1, dx_2, \ldots, dx_n)$, la différentielle de la fonction est donnée par

$$df = \nabla f(\mathbf{x})\cdot \mathbf{dx} = \left(\frac{\partial f}{\partial x_1}\right)(\mathbf{x})\, dx_1 + \left(\frac{\partial f}{\partial x_2}\right)(\mathbf{x})\, dx_2 + \cdots + \left(\frac{\partial f}{\partial x_n}\right)(\mathbf{x})\, dx_n = \sum_{i=1}^{n} \partial_i f(\mathbf{x})\, dx_i$$

L'application de l'opérateur différentiel ∇ à un champ de vecteurs $\mathbf{v}(\mathbf{x})$ dont les composantes $(v_1, v_2, \ldots, v_n)$ sont des fonctions de n variables définit la **divergence** de $\mathbf{v}(\mathbf{x})$:

$$\mathrm{div}\mathbf{v}(\mathbf{x}) = \nabla\cdot \mathbf{v}(\mathbf{x}) = \left(\frac{\partial}{\partial x_1}\right) v_1(\mathbf{x}) + \left(\frac{\partial}{\partial x_2}\right) v_2(\mathbf{x}) + \cdots + \left(\frac{\partial}{\partial x_n}\right) v_n(\mathbf{x}) = \sum_{i=1}^{n} \partial_i v_i(\mathbf{x})$$

Enfin, itéré deux fois, l'opérateur différentiel ∇ permet de calculer le **Laplacien** d'une fonction noté Δ :

$$\Delta f(\mathbf{x}) = \nabla\cdot\nabla f(\mathbf{x}) = \left(\frac{\partial^2 f}{\partial x_1^2}\right)(\mathbf{x}) + \left(\frac{\partial^2 f}{\partial x_2^2}\right)(\mathbf{x}) + \cdots + \left(\frac{\partial^2 f}{\partial x_n^2}\right)(\mathbf{x}) = \sum_{i=1}^{n} \partial_i^2 f(\mathbf{x})$$

Matrices et tenseurs

Une **matrice** est un objet mathématique constitué d'éléments indexés par deux indices :

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{q1} & a_{q2} & \cdots & a_{qp} \end{pmatrix}$$

Par convention, le premier indice indique le numéro de la ligne et le second indice le numéro de la colonne où se trouve l'élément de matrice a_{ij}.

Lorsque le nombre de lignes est égal au nombre de colonnes, on parle de matrice carrée. L'addition de deux matrices carrées revient à sommer les éléments de matrice correspondants. Le produit de deux matrices est défini de la manière suivante :

$$\mathbf{A}\times\mathbf{B} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1q} \\ b_{21} & b_{22} & \cdots & b_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nq} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1q} \\ c_{21} & c_{22} & \cdots & c_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ c_{p1} & c_{p2} & \cdots & c_{pq} \end{pmatrix} = \mathbf{C}$$

où c_{ij} est donné par la combinaison linéaire des éléments de la i$^{\text{ème}}$ ligne de la matrice $\mathbf{A}$ et des éléments de la j$^{\text{ème}}$ colonne de la matrice $\mathbf{B}$:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

Le plus souvent, le produit de deux matrices n'est pas commutatif, $\mathbf{A}\times\mathbf{B}\neq\mathbf{B}\times\mathbf{A}$, sauf dans quelques cas particuliers, comme par exemple la multiplication d'une matrice par la matrice identité notée $\mathbf{I}$:

$$\mathbf{A}\times\mathbf{I}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}=\mathbf{I}\times\mathbf{A}=\mathbf{A}$$

La **transposée** d'une matrice $\mathbf{A}$ notée ${}^{\mathrm{T}}\mathbf{A}$ est obtenue en transposant les lignes par les colonnes de la matrice :

$${}^{\mathrm{T}}\mathbf{A}=\begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix}$$

Une matrice égale à sa transposée est dite **symétrique**. Une matrice $\mathbf{A}$ peut être inversée et sa **matrice inverse** notée $\mathbf{A}^{-1}$ si :

$$\mathbf{A}\times\mathbf{A}^{-1}=\mathbf{A}^{-1}\times\mathbf{A}=\mathbf{I}$$

La matrice inverse d'une matrice carrée à deux lignes et deux colonnes est donnée par :

$$\mathbf{A}^{-1}=\frac{1}{\det\mathbf{A}}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$$

où $\det\mathbf{A}$ est le **déterminant** de la matrice défini par :

$$\det\mathbf{A}=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=a_{11}a_{22}-a_{12}a_{21}$$

Une matrice est **régulière** si sont déterminant est non nul, sinon elle est dite **singulière**. Seules les matrices régulières sont inversibles et les matrices singulières ne peuvent pas être inversées. La notion de déterminant peut être étendue à toutes les matrices carrées et les mêmes règles s'appliquent.

Si les matrices ne sont que de simples tableaux de nombres indépendants de tout changement de base (référentiel), les tenseurs sont formés d'un ensemble de grandeurs indexées par un ou plusieurs indices qui, à l'instar des vecteurs (tenseurs à un indice ou tenseur d'ordre 1), sont affectés par les changements de base.

De manière plus abstraite, l'ensemble des formes linéaires d'un espace vectoriel constitue son **espace vectoriel dual**. Pour donner un exemple, une fonction f de n variables réelles à valeur réelle est une forme linéaire de l'espace vectoriel à n dimensions si elle est linéaire et homogène dans ses n variables, i.e. $f(\mathbf{x}+\mathbf{u})=f(\mathbf{x})+f(\mathbf{u})$ et $f(\lambda\mathbf{u})=\lambda f(\mathbf{u})$, avec $\mathbf{x}$, $\mathbf{u}$ des **vecteurs** de l'espace vectoriel et λ un **scalaire** (nombre réel). Les éléments de l'espace vectoriel dual sont appelés des **covecteurs**.

Si $(\mathbf{e}_1,\mathbf{e}_2,\ldots,\mathbf{e}_n)$ est une base de l'espace vectoriel, la décomposition d'un vecteur sur cette base s'écrit par convention de la manière suivante :

$$\mathbf{x} = x^1\mathbf{e}_1 + x^2\mathbf{e}_2 + \cdots + x^n\mathbf{e}_n = \sum_{i=1}^{n} x^i\mathbf{e}_i$$

Attention, dans cette notation, la numérotation de la composante du vecteur reportée en haut n'est pas un exposant ! On parle alors d'indice **contravariant** et de composante contravariante.

La base de l'espace vectoriel dual est numérotée avec des indices contravariants $(\mathbf{e}^1, \mathbf{e}^2, \ldots, \mathbf{e}^n)$ et la décomposition d'un covecteur sur cette base est donnée par des composantes dont la numérotation est reportée en bas. On parle alors d'indice **covariant** et de composante covariante :

$$\mathbf{u} = u_1\mathbf{e}^1 + u_2\mathbf{e}^2 + \cdots u_n\mathbf{e}^n = \sum_{i=1}^{n} u_i\mathbf{e}^i$$

Pour alléger l'écriture, on utilisera la règle de somme d'Einstein qui s'applique implicitement à toute répétition de deux indices covariant (noté en bas) et contravariant (noté en haut) de la manière suivante :

$$\mathbf{x} = x^i\mathbf{e}_i = \sum_{i=1}^{n} x^i\mathbf{e}_i \,,\ \mathbf{u} = u_i\mathbf{e}^i = \sum_{i=1}^{n} u_i\mathbf{e}^i$$

La fonction de n variables $f(\mathbf{x}) = f_1x^1 + f_2x^2 + \cdots + f_nx^n = f_ix^i$ est une forme linéaire appliquée au vecteur $\mathbf{x} = x^i\mathbf{e}_i$, $(f_1, f_2, \ldots, f_n)$ sont les composantes covariantes du covecteur $\mathbf{f} = f_i\mathbf{e}^i$ correspondant de l'espace vectoriel dual. En effet :

$$f(\mathbf{x}) = f(x^i\mathbf{e}_i) = x^i f(\mathbf{e}_i) = x^i f_j\mathbf{e}^j(\mathbf{e}_i) = x^i f_j\delta_i^j = x^i f_i$$

Les vecteurs de la base duale $(\mathbf{e}^1, \mathbf{e}^2, \ldots, \mathbf{e}^n)$ sont également des formes linéaires qui ont pour propriété :

$$\mathbf{e}^i(\mathbf{e}_j) = \delta_j^i \text{ où } \delta_j^i = \begin{cases} 1 \text{ si } i = j \\ 0 \text{ sinon} \end{cases} \quad \text{est le symbole de Kronecker}$$

Un tenseur permet de décrire une application multilinéaire $\mathbf{T}$ combinant des vecteurs (décrits par des composantes contravariantes) et des covecteurs (décrits par des composantes covariantes). Par exemple, pour une application combinant un covecteur et deux vecteurs, le tenseur T_{jk}^i est défini par:

$$T_{kj}^i = \mathbf{T}(\mathbf{e}^i, \mathbf{e}_j, \mathbf{e}_k)$$

Dans le formalisme tensoriel, les tenseurs A_k^i et B_j^l décrivent des formes bilinéaires combinant un covecteur et un vecteur. En utilisant la règle de somme d'Einstein, le produit de ces deux tenseurs :

$$C_j^i = A_k^i B_j^k$$

est simplement le produit matriciel !

Le produit scalaire de deux vecteurs $\mathbf{x} = x^i\mathbf{e}_i$ et $\mathbf{u} = u^j\mathbf{e}_j$ s'écrit :

$$\mathbf{x}\cdot\mathbf{u} = \left(x^i\mathbf{e}_i\right)\cdot\left(u^j\mathbf{e}_j\right) = x^i\left(\mathbf{e}_i\cdot\mathbf{e}_j\right)u^j = x^i\delta_{ij}u^j$$

En appliquant un changement de base $\tilde{\mathbf{e}}_l = P_l^k \mathbf{e}_k$ permettant d'exprimer les vecteurs $\mathbf{x}$ et $\mathbf{u}$ dans la nouvelle base $(\tilde{\mathbf{e}}_1, \tilde{\mathbf{e}}_2, \ldots, \tilde{\mathbf{e}}_n)$, le produit scalaire entre $\mathbf{x} = \tilde{x}^i \tilde{\mathbf{e}}_i$ et $\mathbf{u} = \tilde{u}^j \tilde{\mathbf{e}}_j$ devient :

$$\mathbf{x} \cdot \mathbf{u} = \left(\tilde{x}^i \tilde{\mathbf{e}}_i\right) \cdot \left(\tilde{u}^j \tilde{\mathbf{e}}_j\right) = \tilde{x}^i \left(\tilde{\mathbf{e}}_i \cdot \tilde{\mathbf{e}}_j\right) \tilde{u}^j = \tilde{x}^i P_i^k \left(\mathbf{e}_k \cdot \mathbf{e}_l\right) P_j^l \tilde{u}^j = \tilde{x}^i P_i^k \delta_{kl} P_j^l \tilde{u}^j = \tilde{x}^i g_{ij} \tilde{u}^j$$

où g_{ij} est le tenseur métrique dans la nouvelle base. Ce tenseur est symétrique et inversible par construction son inverse est noté par la même lettre g, mais avec des indices contravatiants (en haut) :

$$\left(g_{ij}\right)^{-1} = g^{ij} \text{ avec } g^{ik} g_{kj} = \delta_j^i = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

La relativité générale est développée dans un espace vectoriel à 4 dimensions : le temps $x^0 = ct$ et les trois dimensions spatiales (x^1, x^2, x^3). L'équation d'Einstein est alors donnée par :

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}$$

où $R_{\mu\nu}$ est le tenseur de Ricci, $g_{\mu\nu}$ le tenseur métrique, $T_{\mu\nu}$ le tenseur énergie-impulsion, R la courbure, Λ la constante cosmologique, G la constante de gravitation et c la vitesse de la lumière.

Dans le vide et loin de toute source gravitationnelle, le tenseur métrique est donné par la métrique de Minkowski utilisé dans le cadre de la relativité restreinte pour des composantes spatiales covariantes exprimées dans un référentiel Cartésien $(x^1, x^2, x^3) = (x, y, z)$:

$$g_{\mu\nu} = \eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Dans le vide au voisinage d'une masse sphérique statique (sans rotation) M (par exemple le Soleil), le tenseur métrique est donné par la métrique de Schwarzschild pour des composantes spatiales covariantes exprimées dans un référentiel en coordonnées sphériques $(x^1, x^2, x^3) = (r, \theta, \varphi)$:

$$g_{\mu\nu} = \begin{pmatrix} 1 - \dfrac{2GM}{rc^2} & 0 & 0 & 0 \\ 0 & -\dfrac{1}{1 - \dfrac{2GM}{rc^2}} & 0 & 0 \\ 0 & 0 & -r^2 & 0 \\ 0 & 0 & 0 & -r^2 \sin^2\theta \end{pmatrix}$$