

Transport parallèle et connexions en Géométrie et en Physique

Jean-Pierre Bourguignon

Centre de Mathématiques U.R.A. 169 du C.N.R.S. Ecole Polytechnique F-91128
PALAISEAU Cedex (France)

Le thème de cet exposé¹, "*Transport parallèle et connexions en Géométrie et en Physique*", est vraiment au cœur de la problématique de ce colloque consacré à "Un siècle de Géométrie" car c'est entre 1830 et 1930 que s'est élaborée la notion de connexion, cette Longue Marche alliant étroitement Physique et Mathématiques. Presque tous les grands géomètres de cette période ont contribué à cette transformation. Je n'ai donc aucun effort à faire pour qu'apparaissent dans mon exposé des noms de grands mathématiciens comme Carl Friedrich Gauss, Bernhard Riemann, et Elie Cartan, qui complètent le titre de ce colloque. Pour faire cette présentation, deux options s'ouvraient à moi :

- m'en tenir à une démarche historique, prendre le point de vue (qui était presque obligatoire il y a quelques années seulement) de présenter les notions de la façon la plus récente et moderne, en ne faisant que quelques allusions à l'Histoire.

J'ai choisi la première solution. Peut-être est-ce pour ne pas faire comme on m'a appris ? Une raison moins passionnelle tient à ma volonté de mettre en évidence les diverses formes sous lesquelles est apparue la notion de connexion avant de s'imposer comme centrale². Il y a, je crois, peu d'exemples de concepts mathématiques qui ont, comme celui de connexion, mis un siècle et demi à s'imposer. Cet exposé va en effet montrer comment cette notion disparaît dès qu'elle apparaît, comment, peu après son introduction, elle se transforme en un puissant outil de calcul, puis se subordonne à un autre concept, pour enfin triompher en existant vraiment pour elle-même. Pour terminer ma démonstration, j'ai d'ailleurs demandé aux organisateurs de me permettre de présenter des résultats postérieurs à 1930 car il a fallu attendre les années 1970 pour que les connexions soient vraiment étudiées pour elles-mêmes, et la reconnaissance par les physiciens de leur pertinence dans la modélisation des interactions fondamentales est pour beaucoup dans ce phénomène. Cette épopée permet de corriger une idée fausse selon laquelle l'apparition des concepts devrait suivre une marche inexorable. Cette quête a aussi conforté ma conviction qu'il est nécessaire d'apporter une attention plus grande aux conditions d'élaboration des concepts pour en faciliter l'accès, y compris dans le cadre

¹ Recherche soutenue par le programme européen C.E.E. G.A.D.G.E.T. SC1-0105-C.

² Je tiens à appeler l'attention des lecteurs sur l'article [10] qui contient une excellente discussion de beaucoup des points abordés dans cet article.

de l'enseignement. Avant d'entrer dans le vif du sujet, je tiens à exprimer ma gratitude aux organisateurs du colloque pour la confiance qu'ils m'ont témoignée en me proposant de faire ce retour aux sources. Je dois à l'hospitalité d'Ohio State University d'avoir pu trouver le temps nécessaire à l'étude de l'ensemble des articles historiques que je présente ; que tous les collègues qui ont rendu ce séjour possible et fructueux y trouvent une expression de ma gratitude. La préparation matérielle de ce texte a été facilitée par l'aide technique que j'ai reçue d'Eugène Durenard et de Madame Dezetter ; qu'ils en soient remerciés.

I. Avant la Naissance des Connexions

Pour commencer, remontons un peu dans l'Histoire, ... jusqu'à Isaac Newton.

Il est bien connu que Newton n'a pas seulement eu des contributions fondamentales à la Mécanique et à la Physique, mais aussi aux Mathématiques. Pour formuler les équations fondamentales de la Mécanique sous leur forme la plus standard, à savoir exprimer que la résultante des forces appliquées à un système mécanique, à une particule par exemple, est proportionnelle à son accélération, Newton a dû développer un calcul d'un nouveau type : le Calcul Différentiel dont le concept fondamental est celui de *dérivée d'une fonction*.

La notion d'*accélération* n'est pas des plus concrètes. Celle de *vitesse instantanée* suppose déjà une démarche assez abstraite, faisant intervenir l'ensemble des positions du mobile au cours d'un intervalle de temps entourant un instant donné pour n'en retenir qu'une information *infinitésimale*, la dérivée de la fonction position à cet instant. Pour définir l'accélération, il est nécessaire d'aller un cran plus loin encore et de considérer la dérivée *seconde* : étant données les vitesses instantanées d'une particule le long d'une trajectoire (qu'il est naturel d'attacher à la position d'un mobile x de masse d'inertie m à l'instant t où elle est calculée), il faut transporter parallèlement les vitesses instantanées prises en des points voisins et calculer le vecteur vitesse instantanée de cette nouvelle courbe *abstraite*. Ceci conduit à la forme classique des équations de la Mécanique

$$F = m \frac{d^2x}{dt^2} .$$

Ces développements remontent à trois siècles, et sont le sous-bassement des *Philosophiae naturalis Principia Mathematica* de Newton parus en 1687.

Le pas suivant, qui est loin d'être négligeable, a été franchi par Joseph Louis de Lagrange dans [16]. Ce dernier a généralisé le Calcul des Variations d'Euler au cas de systèmes de coordonnées arbitraires. Lagrange a parfaitement compris une notion assez subtile³, celle de *quantité d'accélération*, et c'est là une de ses contributions importantes à la Mécanique. Ainsi il a pu énoncer ce que l'on appelle aujourd'hui *les équations d'Euler-Lagrange* qui étendent les équations de Newton aux systèmes conservatifs repérés dans des systèmes de coordonnées quelconques : pour un système régi par un lagrangien L dépendant des positions par les coordonnées q^i et des composantes $\dot{q}^i$ du vecteur vitesse instantanée, la quantité d'accélération $m d^2x/dt^2$ est remplacée par l'expression de Lagrange ayant pour composantes

³ ... même aujourd'hui, elle nécessite qu'on y réfléchisse un peu.

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^i} \right) - \frac{\partial L}{\partial q^i}.$$

Cette expression doit être égale aux forces appliquées au système. Dans le cadre de notre discussion, le point important est que, pour définir cette notion, on a recours à une réflexion géométrique qui est rendue nécessaire par la nature subtile de la dérivation seconde lorsqu'on travaille dans un système de coordonnées quelconques. En effet, en prenant un point de vue moderne, i.e. en se mettant sur une variété différentielle (ce qui est le cadre naturel pour modéliser des systèmes mécaniques ou physiques un peu généraux), les vecteurs tangents en deux points distincts de l'espace de configuration (que sont les vecteurs vitesses instantanées en deux instants distincts) sont des quantités a priori incommensurables. On rencontre donc une difficulté pour définir l'accélération qui suppose qu'on compare les vecteurs-vitesses en des points distincts. Lagrange parvient à la surmonter en utilisant une donnée supplémentaire, le *lagrangien* L qui est une fonction dépendant des positions et des vitesses : dans le cas d'une particule dans un potentiel V , on a $L = E_c - V$ où E_c désigne l'énergie cinétique de la particule qui est une fonction quadratique de la vitesse. Ceci est fait de façon un peu implicite dans [16], mais Lagrange ne semble pas intéressé⁴ à développer les concepts nécessaires pour analyser complètement le processus par lequel cette équation se forme, ce qui aurait pu le mettre sur la voie de la notion de connexion. Continuons à explorer la période qui précède immédiatement celle retenue pour la conférence. L'étape suivante a un rapport explicite avec la Géométrie, ce qui n'était pas le cas pour les travaux de Lagrange. Je voudrais brièvement présenter les "*Disquisitiones Generales circa superficies curvas*" (cf. [9]) qui furent publiées par Gauss en 1827, mais dont les résultats essentiels étaient connus de lui bien avant. Dans cet article de Géométrie, Gauss utilise les courbes géodésiques qui sont des plus courts chemins pour établir des propriétés géométriques alors que chez Lagrange les géodésiques interviennent essentiellement pour leur intérêt mécanique. Pour développer la théorie des surfaces, Gauss utilise la propriété fondamentale suivante : *il existe des paramétrages de toute surface par des systèmes de coordonnées qui se déduisent des vecteurs-vitesses dans le plan tangent en un point* (ces coordonnées sont depuis appelées *coordonnées gaussiennes*). Gauss met notamment en évidence l'influence de la courbure sur la géométrie d'une surface ; il montre en particulier comment l'expression de la métrique dans ces coordonnées géodésiques fait seulement intervenir la racine carrée de la courbure et comment la courbure est déterminée complètement par la métrique et n'est pas simplement une donnée *extrinsèque*, i.e. dépendant de la forme de la surface dans l'espace ambiant. C'est son fameux *Theorema Egregium*. Grâce à cette étude, il comprend notamment l'influence de la courbure sur les propriétés angulaires des triangles (cette motivation était directement liée à son travail en Géodésie). C'est la première utilisation un peu systématique des géodésiques pour faire une Géométrie générale, mais Gauss, lui non plus, n'éprouve pas le besoin d'introduire un procédé de comparaison entre espaces tangents.

⁴ Ces termes exacts sont les suivants : "*Les méthodes que j'expose ne demandent ni constructions, ni raisonnements géométriques ou mécaniques, mais seulement des opérations algébriques, assujetties à une marche régulière et uniforme*".

II. Une Première Occasion Manquée : La Naissance de la Géométrie de Riemann

La généralisation en dimension supérieure de ce que Gauss a fait pour les surfaces est due à Riemann. C'est le contenu de sa dissertation inaugurale (cf [25]) qui était un exercice imposé. Ce document fondamental, à la présentation duquel Gauss a assisté en 1854, n'a été publié que beaucoup plus tard en 1868, après la mort de Riemann.

Si, dans ce texte (qui n'est pas un article scientifique au sens ordinaire du terme), Riemann développe un grand programme et donne des formules fondamentales pour le développement de la Géométrie appelée aujourd'hui *riemannienne*, il ne fait qu'y esquisser les démonstrations sans énoncer de résultats de façon systématique et linéaire. C'est dans ce document que Riemann introduit les *étendues à n dimensions* (qu'on appelle aujourd'hui *variétés différentielles de dimension n*) et considère des expressions différentielles quadratiques, les *métriques riemanniennes*, qui permettent d'y mesurer la longueur des courbes⁵. Il se sert alors, comme le faisait Gauss, des plus courts chemins pour construire un système de coordonnées géodésiques : dans un voisinage d'un point donné a , un point q est repéré par le vecteur-vitesse de la géodésique issue de a et atteignant le point q au bout du temps 1. Mais ces courbes n'apparaissent pas comme définies par des équations (dans le texte de Riemann, il n'y a en fait aucune équation). Il se préoccupe par contre de mesurer en quoi les espaces généralisés qu'il a introduits dévient de l'espace ordinaire (qu'il propose d'ailleurs d'appeler *plat*). Dans les coordonnées géodésiques l'expression de la métrique n'est plus donnée par des constantes, mais par un développement commençant par l'élément de longueur au point considéré. Le terme suivant contient une quantité plus complexe dont Riemann identifie la dépendance par rapport aux directions et qui se réduit pour chaque sous-surface à la courbure introduite par Gauss : c'est la *courbure de Riemann*. Des développements plus explicites se trouvent dans [24]⁶. En notations modernes, la formule de Riemann est donnée par

$$g(q) = \sum_{i,j=1}^n g_{ij}(a) dx^i dx^j - \frac{1}{6} \sum_{i,j,k,l=1}^n R_{ikjl} (q^k dx^i - q^i dx^k)(q^l dx^j - q^j dx^l) + \dots$$

où g désigne le produit scalaire généralisé et où les q^i sont les coordonnées du point q dans le système de coordonnées géodésiques centré en a . Dans cette formule apparaît clairement que, *dans les coordonnées géodésiques, la métrique riemannienne est osculatrice à la métrique euclidienne*. Au premier ordre, l'écart ne dépend que des éléments de surface engendrés par le vecteur directeur de la géodésique le long de laquelle on se déplace et par le vecteur à mesurer. (Rappelons qu'à cette époque-là le calcul extérieur n'existait pas.) A ce point, nous n'avons toujours ni les connexions, ni le transport parallèle !

⁵ Noter qu'il ne propose cette étude que comme un début, et envisage explicitement de considérer ensuite le cas où la longueur serait donnée par la racine quatrième d'une expression de degré 4 des éléments différentiels.

⁶ A cause de son état de santé, il n'a pu développer complètement son argumentation ce qui semble lui avoir coûté le prix qu'il brigait.

III. Les $\left\{ \begin{smallmatrix} j & k \\ i & \end{smallmatrix} \right\}$ d'E.B. Christoffel

Le mathématicien qui a vraiment introduit la notion de connexion est Edwin Bruno Christoffel : dans l'article [7] paru une année après la publication du texte de Riemann⁷, il s'est intéressé aux *conditions que doivent vérifier des expressions différentielles de degré 2 pour pouvoir se transformer l'une dans l'autre par un changement de variables*.

Le point de vue de Christoffel est strictement celui d'un mathématicien intéressé par la *théorie des invariants*. Il résout un problème de *classification*. Il part donc de deux expressions $\sum_{i,j=1}^n g_{ij}(x^k) dx^i dx^j$ et $\sum_{\alpha,\beta=1}^n g_{\alpha\beta}(y^\gamma) dy^\alpha dy^\beta$, dans lesquelles les quantités (g_{ij}) et $(g_{\alpha\beta})$ sont des matrices de fonctions inversibles sur leur domaine de définition, et il cherche une transformation f qui échange ses deux expressions (ce sont donc les expressions de la même métrique riemannienne dans deux cartes reliées par le difféomorphisme local inconnu f). Il est à noter que dans [6] il considère le problème plus général de l'équivalence d'expressions homogènes de degré n ce qui souligne que la Géométrie riemannienne (i.e. définie par un produit scalaire à coefficients variables) n'avait pas encore établi la suprématie que nous lui connaissons maintenant.

Dans [7], Christoffel introduit les symboles $\left\{ \begin{smallmatrix} j & k \\ i & \end{smallmatrix} \right\}$ de la façon suivante. Si les variables y et x vérifient la relation $y = f(x)$, alors $dy^\alpha = \sum_{i=1}^n (\partial f^\alpha / \partial x^i) dx^i$ où $(\partial f^\alpha / \partial x^i)$ désigne la *matrice jacobienne* de f . L'idée fondamentale est de partir de la relation différentielle

$$\sum_{i,j=1}^n g_{ij}(x^k) dx^i dx^j = \sum_{\alpha,\beta=1}^n g_{\alpha\beta}(y^\gamma) dy^\alpha dy^\beta$$

et de la dériver une fois en supposant les coordonnées (x^k) et (y^γ) , et leurs différentielles, reliées par la transformation f . Ce faisant, on fait donc apparaître des dérivées partielles secondes de la transformation f que l'on cherche, mais aussi des dérivées partielles premières des coefficients de la métrique.

Christoffel remarque alors que, pour isoler une dérivée partielle seconde de la transformation f , il est commode d'introduire les coefficients⁸ notés aujourd'hui

$$\left\{ \begin{smallmatrix} i & \\ j & k \end{smallmatrix} \right\} = \frac{1}{2} \sum_{l=1}^n g^{il} \left(\frac{\partial g_{jl}}{\partial x^k} + \frac{\partial g_{kl}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^l} \right)$$

où (g^{ij}) désigne la matrice inverse de la matrice (g_{ij}) . Un fait technique tout à fait important pour plus tard est la non-dégénérescence de la métrique, car dans l'écriture des symboles de Christoffel apparaît l'inverse de la matrice de la métrique. Cela lui donne l'équation fondamentale

$$\frac{\partial^2 f^\alpha}{\partial x^i \partial x^j} = \sum_{l=1}^n \left\{ \begin{smallmatrix} l & \\ i & j \end{smallmatrix} \right\} \frac{\partial f^\alpha}{\partial x^l} - \sum_{\beta,\gamma=1}^n \left\{ \begin{smallmatrix} \alpha & \\ \beta & \gamma \end{smallmatrix} \right\} \frac{\partial f^\beta}{\partial x^i} \frac{\partial f^\gamma}{\partial x^j}.$$

⁷ Christoffel fait explicitement référence dans son texte aux travaux de Riemann (en mentionnant d'ailleurs Dedekind). Je ne sais par contre pas si des notes de Riemann avaient circulé parmi les mathématiciens de l'époque.

⁸ L'automatisme maintenant habituel de calcul sur la position des indices est perturbé dans l'article de Christoffel. En effet il ne met pas les indices en positions opposées dans l'expression de la métrique, ni ne se préoccupe de la variance de la matrice inverse des coefficients de la métrique.

A aucun moment, Christoffel ne s'intéresse à la relation qui pourrait exister entre ces quantités et des courbes particulières dans l'espace. En particulier, aucune remarque n'est faite pour dire que ces coefficients sont naturellement présents dans l'équation des géodésiques $t \mapsto (\gamma^i(t))$ que l'on obtient en appliquant les équations de Lagrange à l'énergie E_c que définit la métrique riemannienne g , équation qui s'écrit

$$\frac{d^2 \gamma^i}{dt^2} + \sum_{j,k=1}^n \left\{ \begin{matrix} i \\ j \ k \end{matrix} \right\} \frac{d\gamma^j}{dt} \frac{d\gamma^k}{dt} = 0 .$$

Ces symboles jouent un rôle important dans la résolution du problème étudié par Christoffel et dont il donne une solution presque complète. Comme condition nécessaire à l'équivalence des deux métriques, il obtient notamment que l'évaluation dans les deux cartes des composantes de la courbure de Riemann, qui se calculent à partir des symboles de Christoffel par la formule

$$R_{ijk}{}^l = \frac{\partial \left\{ \begin{matrix} l \\ j \ k \end{matrix} \right\}}{\partial x^i} - \frac{\partial \left\{ \begin{matrix} l \\ i \ k \end{matrix} \right\}}{\partial x^j} + \sum_{m=1}^n \left(\left\{ \begin{matrix} m \\ j \ k \end{matrix} \right\} \left\{ \begin{matrix} l \\ i \ m \end{matrix} \right\} - \left\{ \begin{matrix} m \\ i \ k \end{matrix} \right\} \left\{ \begin{matrix} l \\ j \ m \end{matrix} \right\} \right) ,$$

doivent se correspondre par la matrice jacobienne qui est déterminée par les deux expressions métriques, i.e. qu'elles satisfont l'égalité

$$\sum_{l=1}^n R_{ijk}{}^l \frac{\partial f^l}{\partial x^i} = \sum_{\alpha,\beta,\gamma=1}^n \frac{\partial f^\alpha}{\partial x^i} \frac{\partial f^\beta}{\partial x^j} \frac{\partial f^\gamma}{\partial x^k} R_{\alpha\beta\gamma}{}^\delta .$$

Christoffel établit un certain nombre d'identités pour ces quantités à quatre indices (antisymétries, et ce qu'on appelle aujourd'hui la première identité de Bianchi). Mais il va aussi plus loin, et montre que le même type d'identité doit être satisfait pour certaines expressions qui sont des *dérivées modifiées de la courbure*. Ces expressions qui seront identifiées ultérieurement comme les dérivées covariantes de la courbure sont un peu compliquées. Christoffel voit cependant qu'il y a un procédé général de formation d'objets nouveaux puisqu'il applique ce procédé de dérivation aux composantes de la métrique et mentionne que les expressions obtenues sont identiquement nulles. Christoffel trouve ainsi une famille infinie de conditions pour que les deux expressions différentielles soient équivalentes, à savoir que les composantes de la courbure et de ses dérivées covariantes soient mises en correspondance par la matrice jacobienne. Sa démonstration du caractère suffisant (cf. [8]) qui repose sur l'étude des conditions de compatibilité du système différentiel n'est pas tout à fait correcte. Il remarque qu'un nombre fini de ces conditions suffit pour assurer l'équivalence dans le cas générique. En présence d'un groupe d'automorphismes non trivial (i.e. lorsque la métrique a des isométries), il a conscience que la situation est plus complexe. En fait la solution finale de ce problème sera donnée dans le cas analytique par Elie Cartan (cf. [3]) et, dans le cas général, par Shlomo Sternberg dans les années 50. C'est donc l'aspect "*outil de calcul*" qui a poussé Christoffel à introduire les symboles qui portent son nom. Bien que mis en œuvre, le concept proprement dit de connexion (dans sa version *dérivation covariante*) n'est donc pas étudié par Christoffel.

IV. G. Ricci et le Calcul Différentiel Absolu

C'est Gregorio Ricci-Curbastro qui a vraiment reconnu que les formules introduites par Christoffel permettaient de créer un outil d'une grande importance théorique. Il a présenté cela dans l'article [22] publié en 1888 (voir aussi [21]), qui est resté relativement inconnu. Grâce à ce nouvel outil, il est en effet possible de sortir des difficultés créées par le caractère non intrinsèque des opérations de dérivation partielle effectuées dans des systèmes de coordonnées locales.

L'article [23], qui est généralement considéré comme la référence sur ce point, s'intitule "*Méthodes de calcul différentiel absolu et leurs applications*" et est paru aux Mathematische Annalen en 1900. Ricci l'a écrit en collaboration avec un des ses élèves, Tullio Levi-Civita. Il est important de souligner la modernité de ce texte qui est *l'acte de naissance et le manuel fondamental du Calcul Tensoriel*. Une de leurs motivations en écrivant cet article était de donner une démonstration de l'invariance par changement de coordonnées de l'opérateur de Laplace-Beltrami (qui généralise le laplacien ordinaire aux variétés riemanniennes).

Le point de vue qui y est adopté est tout à fait intrinsèque, et les notions de *covariance* et de *contravariance* clairement définies. Les auteurs s'y préoccupent abondamment de ce qu'une équation établie dans un système de coordonnées soit encore valable dans un autre système de coordonnées. C'est d'ailleurs de cette analyse que les auteurs tirent le nom d'*absolu* qu'ils donnent à cette extension du Calcul Différentiel. La dérivée covariante d'un champ de vecteurs X est un champ d'endomorphismes dont les composantes $(X^i_{;j})$ se déduisent des composantes (X^i) de X dans un système de coordonnées x^i par la formule

$$X^i_{;j} = \frac{\partial X^i}{\partial x^j} + \sum_{k=1}^n \left\{ \begin{matrix} i \\ j \ k \end{matrix} \right\} X^k .$$

Le lien avec les géodésiques associées à cette métrique est naturellement fait dans [23], et il est montré que leur quantité d'accélération (i.e. la dérivation covariante du vecteur-vitesse de la géodésique le long de lui-même) est nulle.

L'article est bâti comme un cours dans lequel les préoccupations physiques sont pratiquement à l'égal des préoccupations mathématiques. Passons en revue les titres des différents chapitres : I. Algorithme du Calcul différentiel ; II. La Géométrie Intrinsèque comme instrument de calcul ; III. Applications analytiques ; IV. Applications géométriques ; V. Applications mécaniques ; VI. Applications physiques.

C'est donc sous la forme de dérivation covariante que le concept de connexion s'est formalisé pour la première fois. Cette nouvelle dérivation se déduit de la métrique riemannienne (qu'il était alors de coutume d'appeler la *forme fondamentale* par extension du cas des surfaces). Ainsi développé, ce nouveau Calcul allait pouvoir devenir un outil de base pour la Physique Mathématique.

V. L'Aventure Mathématique de la Relativité Générale

La naissance de la théorie de la Relativité Générale a constitué une occasion extraordinaire pour développer le Calcul Différentiel Absolu, et pousser à l'approfondissement de la notion de connexion. La création de cette théorie, qui a indiscutablement été l'œuvre d'Albert Einstein, s'est échelonnée sur plusieurs années. Le point de départ est bien entendu la Théorie de la Relativité Restreinte. Dans ce cadre, l'espace-temps est un espace vectoriel lorentzien dans lequel les rayons lumineux sont des génératrices des cônes de lumière et où les particules libres se meuvent suivant des lignes droites du type temps. Je donne deux références pour la Relativité Générale. La première, [12], est, à tort à mon avis, considérée comme une tentative avortée. Il est intéressant de noter qu'elle a deux parties : l'une, écrite par Einstein, porte sur les aspects physiques, et l'autre traitant des aspects mathématiques a, elle, été écrite par son ami Marcel Grossmann. Einstein y justifie le saut épistémologique fondamental de la Relativité Générale consistant à *concevoir les effets de la Gravitation comme des modifications de la Géométrie de l'espace-temps*. Il y propose donc que le tenseur métrique (lorentzien puisque nous sommes dans l'espace-temps) serve de potentiel de gravitation. Ce sont les inhomogénéités du champ de gravitation qui accélèrent les particules. C'est seulement dans la deuxième référence, [13], qu'il donne les équations de champ correctes (après une course effrénée avec David Hilbert!). Le coût mathématique de cette démarche est considérable puisqu'il lui faut utiliser les théories mathématiques les plus avancées de son temps. Une des sources de la Relativité Générale est la réflexion menée par Einstein sur *l'identification qu'il convient de faire sur la base des expériences physiques entre masse d'inertie et masse gravitationnelle*. Quand on adopte le point de vue de l'inertie, on est naturellement amené à étudier les trajectoires des particules tombant librement, et à utiliser ces courbes pour définir une Géométrie, par exemple en les considérant comme des géodésiques pour une dérivation covariante. A ce point, nous sommes donc confrontés à deux familles d'objets mathématiques : *une connexion* définie par l'intermédiaire de ses géodésiques, d'autre part *une métrique lorentzienne* correspondant à la distribution de masse gravitationnelle. Dans le prolongement de la Relativité Restreinte, Einstein les relie alors comme suit : *les vecteurs-vitesses des photons, particules de masse nulle tombant librement, doivent être partout tangents au cône de lumière*. Dès lors, la connexion définie par ces géodésiques est nécessairement celle que Christoffel et Ricci et Levi Civita avaient déterminée à partir des coefficients de la métrique⁹. Ainsi, *alors qu'apparaissait l'occasion d'introduire de façon autonome une dérivation covariante, celle-ci est immédiatement subordonnée à la métrique lorentzienne*. Noter que le caractère non tensoriel des symboles de Christoffel prend ici une signification physique : le fait que, le long d'une géodésique, il existe un système de coordonnées où ces symboles s'annulent correspond au fait que le passage à des systèmes de référence accélérés crée des effets inertiels et gravitationnels. Du point de vue des connexions, l'aventure de la Théorie de la Relativité Générale a été une nouvelle occasion perdue : *la notion de connexion n'a pas encore acquis un statut pour elle-même*.

⁹ Noter que, dans la définition des symboles de Christoffel, seule la non-dégénérescence de la métrique intervient. Il est donc possible d'étendre aux métriques lorentziennes les considérations introduites jusque là seulement pour les métriques riemanniennes.

VI. La Naissance du Transport Parallèle

C'est dans [18] écrit par Levi Civita en 1917 qu'apparaît la notion de *transport parallèle* ce qui constitue une étape très importante dans l'autonomisation de la notion de connexion.

C'est là que le processus géométrique par lequel on peut définir l'accélération dans un espace général est vraiment analysé de façon propre. Levi-Civita envisage la résolvante de l'équation différentielle du premier ordre définissant les champs de vecteurs, dits *parallèles*, dont la dérivée covariante le long d'une courbe est nulle. Grâce à cette résolvante, on peut en effet comparer les vecteurs tangents en deux points distincts et par là-même définir le vecteur accélération d'une courbe comme dans un espace vectoriel : en transportant parallèlement les vecteurs-vitesses le long de la courbe jusqu'à un point de référence, on définit une courbe dans l'espace vectoriel tangent en ce point qui est un espace fixe. Les *géodésiques* sont donc les courbes dont le vecteur-vitesse est transporté parallèlement sur lui-même.

Levi Civita fait une analyse assez poussée de ce que peut être une notion de parallélisme. Cela le conduit à travailler dans un espace vectoriel ambiant pour pouvoir le comparer au parallélisme ordinaire. Il interprète notamment le *transport parallèle le long d'une courbe tracée sur une sous-variété d'un espace euclidien munie de la métrique induite comme le roulement sans glissement le long de cette courbe*. C'est Hermann Weyl qui va vraiment se libérer de l'idée de transport parallèle conçu en termes géométriques traditionnels. Dans [28], [29] et dans les versions successives de [30] (livre fondamental pour les relations entre Mathématiques et Physique), Weyl définit une notion générale de *connexion affine* et développe une tentative d'unification de la Gravitation et de l'Electromagnétisme à laquelle il donne le nom de *Weltgeometrie*. Il interprétait le champ scalaire nécessaire pour représenter le potentiel électromagnétique comme le facteur conforme d'une métrique (il s'agit donc d'un potentiel réel). Cette tentative était suggérée par les propriétés d'invariance conforme de l'Electromagnétisme. Mais cette théorie n'a pas pu être validée physiquement.

Pourtant, sur le plan mathématique, ces recherches ont permis à Weyl de bien dégager la notion de transport parallèle associée à une dérivation covariante générale. Elle a aussi été chez Weyl une puissante motivation pour *développer la Géométrie Conforme*. Mais, peut-être encore plus important pour l'avenir, il pressent à cette occasion l'existence d'équations qui admettent une invariance plus générale que l'invariance par un groupe de transformations de dimension finie. C'est la première apparition d'une *Théorie de Jauge* ("Eichinvarianz" dans le texte). Dans [31], après l'avènement de la Mécanique Quantique, Weyl reprend ces idées d'un point de vue physique en donnant une nouvelle version de sa Théorie de Jauge. Il prend cette fois le facteur de phase imaginaire pur et il décrit très précisément comment ce changement est lié à la considération du groupe multiplicatif des nombres complexes de module 1 (dont l'axe imaginaire est l'algèbre de Lie), alors que la construction précédente était reliée au groupe multiplicatif des nombres réels. Cette approche a été par la suite considérée comme une alternative à la présentation traditionnelle des équations de Maxwell (voir aussi [19]). Dans une des sections suivantes, nous verrons combien, convenablement généralisé au cas de groupes de Lie compacts non nécessairement abéliens, ce point de vue s'est révélé fécond.

VII. L'Apport d'Elie Cartan

Elie Cartan a évidemment joué un très grand rôle dans l'émergence des concepts géométriques modernes, et les connexions sont au cœur de son œuvre. On sait maintenant que l'aboutissement sera la *Théorie des Espaces Fibrés*, et la notion de connexion dans ces espaces. Il ne fait cependant pas de doute qu'on trouve tout cela en germe dans les travaux d'Elie Cartan. A propos des connexions, on pourrait citer beaucoup d'articles d'Elie Cartan ; je n'en cite que deux, [1] et [2]. Dans le cadre de la Théorie de la Relativité Générale, Elie Cartan a fait beaucoup de tentatives pour libérer la métrique et la connexion du lien fort qui existait entre elles. Ses efforts n'ont pas été couronnés de succès comme le montre par exemple l'échec de la Théorie du Parallélisme Absolu. On peut mentionner l'étude fine de la notion de *torsion* d'une connexion qu'il a menée à ce propos.

On doit à Elie Cartan une autre notion importante attachée aux connexions : celle de *groupe d'holonomie*. Il s'agit du sous-groupe du groupe linéaire de l'espace tangent en un point qui est engendré par les transformations résultant du transport parallèle d'un vecteur le long de tout lacet issu de ce point. Cette notion introduit en fait une extension tout à fait intéressante de la notion de groupe d'invariance, puisqu'au lieu de considérer des propriétés conservées par des transformations, on s'intéresse ici à des notions infinitésimales qui restent bien définies même lorsqu'elles sont transportées parallèlement le long de n'importe quel chemin. Cette approche est l'acte de naissance de certaines Géométries appelées génériquement du nom de *Géométries de Chemins*. On peut noter que certaines Géométries spéciales (comme les Géométries Kählériennes ou Quaternionniennes) peuvent être détectées par le biais de leur groupe d'holonomie. Les articles d'Elie Cartan sur les *espaces (localement) symétriques* ont joué un rôle considérable dans le développement de la Géométrie moderne car ces espaces se sont révélés être des modèles extrêmement précieux à cause à la fois de leur grande diversité et de la possibilité de les décrire complètement de façon algébrique. Cette classe d'espaces peut être distinguée de beaucoup de façons équivalentes, à première vue difficiles à rapprocher : l'une d'entre elles consiste à demander que le *transport parallèle laisse invariant le tenseur de courbure de Riemann* comme l'indique le titre de [2].

VIII. Les Connexions dans la Théorie Moderne des Espaces Fibrés

Le pas suivant nous fait dépasser 1930. Il a été franchi avec la théorie d'Ehresmann (cf. [11]) qui a donné finalement un contenu réellement géométrique à la notion de connexion. Pour simplifier, nous allons nous restreindre à la famille des espaces fibrés *principaux*. On se place sur un espace P sur lequel un groupe de Lie G agit sans point fixe de telle sorte que toutes les orbites sont difféomorphes. Il est dès lors naturel d'introduire $M = P/G$, l'espace des orbites, sur lequel P se projette naturellement ce qui donne naissance à la fibration $P \rightarrow M$ (i.e. la contre-image d'un ouvert suffisamment petit de M est le produit de cet ouvert par la *fibre-type*, ici un espace difféomorphe à G). Le groupe G s'appelle dans ce contexte le *groupe structural*. Dans ce cadre une *connexion* n'est alors rien d'autre que la donnée en tout point p de P d'un sous-espace de l'espace tangent

en p à P transverse à l'espace tangent à la fibre (appelé *espace horizontal* puisqu'on représente le plus souvent l'espace tangent à la fibre comme vertical). Comme le groupe G agit sur P , il est naturel de demander en plus que cette donnée soit équivariante sous l'action du groupe, i.e. qu'au point $p.g$ (pour $g \in G$) le sous-espace horizontal soit l'image du sous-espace horizontal au point p par l'action de g . Avec cette définition, une connexion apparaît comme un objet purement géométrique, et l'existence d'une telle représentation est importante : ainsi nous ne sommes pas contraints de travailler avec des fibrés de repères ou des fibrés qui leur sont associés comme le fibré des vecteurs tangents. Toutes les notions attachées à celle de connexion peuvent s'interpréter dans ce cadre (cf. [19]). Ainsi le transport parallèle est le résultat du relèvement horizontal d'une courbe (ce qui se traduit, pour une géodésique, par le fait que le relèvement de son vecteur-vitesse est à tout instant horizontal). Le groupe d'holonomie prend aussi une signification géométrique intéressante comme sous-groupe opérant dans la fibre qui permet une *réduction* du groupe structural du fibré. De même la courbure apparaît dans ce contexte comme l'*obstruction à l'intégrabilité de la distribution horizontale définissant la connexion*, d'où une interprétation purement géométrique de la platitude.

Il est à noter que les préoccupations physiques étaient totalement absentes chez Ehresmann. Ce point vaut la peine d'être souligné car il montre que *des cheminements différents par des physiciens et des mathématiciens peuvent aboutir à des objets très similaires*. Certains physiciens mathématiciens avaient proposé très tôt des constructions de cette nature. Ainsi une théorie, développée en 1919 par Theodor Kaluza dans [14], puis reprise en 1926 par Oscar Klein dans [15], visait à unifier l'Electromagnétisme et la Gravitation en considérant des métriques spéciales dans un espace à cinq dimensions. La construction de ces métriques faisait déjà appel à un supplémentaire de l'espace tangent à l'espace élargi. Les dérivations covariantes peuvent être aussi vues comme des *opérateurs différentiels du premier ordre à symbole injectif*. Ce point de vue, développé notamment par Jean-Louis Koszul, s'est révélé important dans l'étude des structures complexes sur les espaces totaux d'espaces fibrés et dans l'introduction de fibrés de coefficients qui sont indispensables dans les diverses formes du Théorème de l'Indice. Cela a d'ailleurs amené Daniel Quillen à donner une généralisation algébrique de la notion de connexion sous le nom de *superconnexion*. Ces travaux ont été amplifiés par Jean-Michel Bismut. Dans la section suivante, nous développons encore une autre facette des connexions.

IX. Le Triomphe des Théories de Jauge

Il y a aujourd'hui plusieurs raisons indépendantes pour lesquelles les connexions sont reconnues comme des objets fondamentaux à la fois par les mathématiciens et par les physiciens. La publication en 1954 de [33] écrit par Chen Ning Yang et Robert Mills a représenté un saut qualitatif dans la géométrisation de la Physique. La Théorie de Yang-Mills n'a pas eu un succès immédiat, et surtout ce n'est que plus tard qu'elle a été identifiée comme une théorie s'appuyant fondamentalement sur des connexions sur un espace fibré¹⁰. L'objectif initial de Yang et Mills était de développer un modèle classique des interactions fortes (celles qui sont responsables de la cohésion des noyaux) bien que

¹⁰ Il semble que plusieurs physiciens mathématiciens comme Andrzej Trautmann aient compris cela bien plus tôt, mais ils sont restés isolés.

ces interactions soient fondamentalement de nature quantique. Ils proposèrent donc à cet effet un lagrangien en espérant que les extrémales de l'action qu'il définit permettraient de donner les premiers termes d'un développement perturbatif dont on pourrait tirer des informations quantiques. *Le lagrangien proposé par Yang et Mills n'était rien d'autre que la norme L^2 de la courbure d'une connexion* sur des fibrés ayant pour base l'espace-temps et pour groupe structural le groupe non abélien SU_2 . A cette époque, ce groupe était considéré comme le groupe de symétrie associé au nombre quantique de spin isotopique ; depuis lors, d'autres nombres quantiques ont été mis en évidence pour les interactions fortes ce qui a conduit à prendre SU_3 comme groupe de symétrie des interactions fortes. Cet article proposant de prendre l'espace des connexions comme espace des paramètres pour une Théorie des Champs est remarquable de clairvoyance. Ce modèle a eu finalement plus de succès quand il a été repris par Steven Weinberg et Abdus Salam (cf. [26] et [27]) pour expliquer les *interactions faibles* (responsables de la radioactivité β) au point de permettre l'unification de ces interactions avec l'Electromagnétisme sous le nom d'interaction *électrofaible*. Une étape importante pour le succès de ces théories a été, en 1971, la preuve par Gerhart 't Hooft du fait que *ces théories pouvaient être renormalisées*, i.e. qu'on peut définir un procédé cohérent pour éliminer les termes infinis apparaissant dans certaines intégrales qui permettent d'estimer les perturbations quantiques. Aujourd'hui, le succès de ces approches, dites *de jauge*, est presque total. Elles servent de cadre de pensée à plusieurs branches de la Physique Théorique. Pour la Théorie des Particules Élémentaires, la conception suivante est maintenant bien installée. On distingue

- les *champs représentant des particules de matière*, qui, au niveau classique, ont des fonctions d'ondes identifiées à des *sections de certains fibrés* dont le groupe structural reflète les interactions dans lesquelles cette particule est engagée ;
- des *champs d'interaction* correspondant à des *particules virtuelles d'échange* qui sont toujours des bosons (comme les photons, ou les vecteurs de l'interaction faible que sont les bosons Z^0 , W^+ et W^- dits *intermédiaires*), dont la version classique correspond à des *connexions* sur les fibrés mentionnés précédemment. On voit combien il était fondamental de disposer de la notion de connexion sur des fibrés de groupe structural un groupe de Lie n'opérant pas sur l'espace-temps. *C'est donc dans le cadre de théories physiques que, pour la première fois, l'espace de toutes les connexions est apparu comme un objet digne d'intérêt.* Sur une variété de dimension 4 orientée, l'espace des minima de la fonctionnelle de Yang-Mills (convenablement normalisé en le quotientant par le groupe de jauge) est un objet que les physiciens souhaitaient vivement comprendre vers la fin des années 70 en liaison avec le calcul de perturbations quantiques. De nombreux mathématiciens (parmi lesquels Michael F. Atiyah, Vladimir Drinfeld, Nigel Hitchin, Yuri Manin, Isadore M. Singer, Clifford Taubes, Karen Uhlenbeck) ont contribué à mieux comprendre cet espace de modules. En 1982, Simon Donaldson a complètement renversé la problématique en montrant comment *cet espace de connexions minimisantes contient des informations subtiles sur la structure différentielle de la variété de base.* Ce fut le point de départ d'une révolution de la topologie différentielle des variétés de dimension 4, ... et d'un intérêt accru d'autres mathématiciens pour les connexions. Simon Donaldson (et d'autres) ont étendu la portée de cette approche à l'étude de divers *espaces de modules*, comme ceux des *fibrés stables* qui intéressent beaucoup les géomètres

algébristes, ou les *espaces de représentations* de certains groupes en les identifiant à des espaces de connexions holomorphes ou plates.

X. L'Apport des Connexions dans l'Etude de Propriétés Globales

Dans les dernières décennies, d'autres démarches totalement indépendantes ont aussi amené les mathématiciens à étudier les connexions pour elles-mêmes. Faisons un petit saut en arrière, et regardons comment les connexions ont participé à une des évolutions fondamentales de la Géométrie dans ce siècle, à savoir *le passage du local au global*. Une des étapes décisives de cette évolution a été la description de certaines classes de cohomologie dites *caractéristiques* en terme de courbure grâce à la fameuse *théorie de Chern-Weil*. Le point de départ en fut la découverte en 1944 par Shiing-Shen Chern¹¹ (cf. [4]) d'une nouvelle preuve du théorème de Gauss-Bonnet exprimant la caractéristique d'Euler-Poincaré par une intégrale de courbure. La construction de base en était une *transgression* de la forme de courbure d'une surface, i.e. l'opération consistant à la remonter en la différentielle extérieure de la forme de connexion sur le fibré en cercles tangents. Une nouvelle fois, les connexions sont introduites pour disparaître dans le résultat final ! En 1974, S.S. Chern et James Simons (cf. [5]) ont étendu la théorie de Chern-Weil aux classes caractéristiques secondaires avec, cette fois, dépendance en la connexion. La quantité introduite par Chern et Simons en dimension 3 a d'ailleurs trouvé sa place dans certaines Théories des Champs contemporains sous le nom de *lagrangien de Wess-Zumino*. L'opération de transgression est aussi fondamentale dans la cohomologie B.R.S. qui permet la description de certaines *anomalies* rencontrées en quantifiant certaines théories. L'étude de ces objets a encore été stimulé récemment par le point de vue développé en 1989 par Edward Witten dans [32] sur le nouvel invariant des nœuds que constitue le polynôme de Jones. Son détour utilise une Théorie des Champs pour le lagrangien de Wess-Zumino et pour chaque connexion la trace du transport parallèle le long du nœud. Ces deux exemples montrent l'actualité des recherches menées autour des connexions, et la place prise par l'espace de toutes les connexions dans la Géométrie Moderne. Bien que le cheminement historique ait été tortueux, nous pouvons dire aujourd'hui que *la notion de connexion occupe une place centrale en Géométrie et en Physique*.

¹¹ A ce propos, notons que la première personne qui a exprimé un nombre caractéristique d'une variété de dimension 4 par une intégrale de courbure est en fait un physicien, Cornelius Lanczos, qui, étudiant dans [17] les équations de champ engendrées par les lagrangiens quadratiques en la courbure, a constaté qu'un lagrangien particulier n'engendrait aucune équation de champ : c'était celui qui exprimait la caractéristique d'Euler-Poincaré en dimension 4.

Bibliographie

- [1] E. Cartan, *Sur les variétés à connexion affine et la théorie de la relativité générale*, Ann. Sci. Ec. Norm. Paris **40** (1923), 325-412.
- [2] E. Cartan, *Sur les espaces de Riemann dans lequel le transport par parallélisme conserve la courbure*, Rend. Acc. Lincei **3** (1926), 544-547.
- [3] E. Cartan, *Leçons sur la Géométrie des espaces de Riemann*, Deuxième édition, Gauthier-Villars, Paris (1946).
- [4] S.S. Chern, *A simple intrinsic proof of the Gauss-Bonnet formula for closed Riemannian manifolds*, Ann. Math. **45** (1944), 747-752.
- [5] S.S. Chern, J. Simons, *Characteristic forms and geometrical invariants*, Ann. Math. **99** (1974), 48-69.
- [6] E.B. Christoffel, *Über die Transformation ganzer homogener Differentialausdrücke*, Monatsber. Kgl. Preuss. Akad. Wiss. Berlin (1969), 1-6.
- [7] E.B. Christoffel, *Über die Transformation der homogenen Differentialausdrücke zweiten Grades*, J. reine angew. Math. **70** (1869), 46-70.
- [8] E.B. Christoffel, *Über ein die Transformation der homogenen Differentialausdrücke zweiten Grades betreffendes Theorem*, J. reine angew. Math. **70** (1869), 241-245.
- [9] P. Dombrowski, *150 years after Gauss' Disquisitiones Generales circa Superficies Curvas*, Astérisque **62** (1979).
- [10] J. Ehlers, *Christoffel's work on the equivalence problem for Riemannian spaces and its importance for modern field theories of Physics*, in *E.B. Christoffel's Festschrift*, Aachen (1980), 526-542.
- [11] C. Ehresmann, *Les connexions infinitésimales dans un espace fibré différentiable*, in *Colloque de Topologie*, Bruxelles (1950), 29-55.
- [12] A. Einstein, *Die Grundlagen der allgemeinen Relativitätstheorie*, Ann. Phys. **49** (1916), 769-822.
- [13] A. Einstein, M. Grossmann, *Entwurf einer allgemeinerten Relativitätstheorie und einer Theorie der Gravitation, I. Physikalischer Teil*, Z. Mat. Phys. **62** (1913), 225-244 ; *idem*, *II. Mathematischer Teil*, ibidem, 245-261.
- [14] T. Kaluza, *Zum Unitätsproblem der Physik*, Sitzungber. Preuss. Akad. Wiss. Phys., Math. Klasse (1921), 966-972.
- [15] O. Klein, *Quantentheorie und fünf-dimensionale Relativitätstheorie*, Z. Phys. **37** (1926), 895-906.
- [16] J.L. Lagrange, *Mechanique Analytique*, Paris (1788).
- [17] C. Lanczos, *A remarkable property of the Riemann-Christoffel tensor in four dimensions*, Ann. Math. **39** (1938), 842-850.
- [18] T. Levi-Civita, *Nozione di parallelismo in una varietà qualunque et conseguente specificazione geometrica della curvatura riemanniana*, Rend. Circ. Mat. Palermo **42** (1917), 173-204.
- [19] A. Lichnerowicz, *Théorie globale des connexions et des groupes d'holonomie*, Cremonese, Roma (1955).
- [20] F. London, *Quantenmechanische Deutung der Theorie von Weyl*, Z. Phys. **42** (1927), 375-389.
- [21] G. Ricci, *Sui parametri e gli invarianti delle forme quadratiche differenziali*, Ann. Mat. Pura Appl. **XIV** (1886).
- [22] G. Ricci, *Delle derivazioni covarianti e controvarianti*, Stud. ed. Univ. Padova, Padova (1888).
- [23] G. Ricci, T. Levi-Civita, *Méthodes de calcul différentiel absolu et leurs applications*, Math. Ann. **54** (1900), 125-201.

- [24] B. Riemann, *Commentatio mathematica, qua respondere tentatur quaestioni ab Ill^{ma} Academia Parisiensi propositae* (1861), in *Riemann's Gesamm. Math. Werke* XXII, 2. Aufl. (1892), 391-423.
- [25] B. Riemann, *Über die Hypothesen, welche der Geometrie zu Grunde liegen*, in *Riemann's Gesamm. Math. Werke* XIII, Abhandl. Kgl. Gesell. Wiss. Göttingen XIII (1868), 272-287.
- [26] A. Salam, *Gauge unification of fundamental forces*, Rev. Modern Phys. **92** (1980), 525-536.
- [27] S. Weinberg, *Conceptual foundations of the unified theory of weak and electromagnetic interactions*, Rev. Modern Phys. **92** (1980), 515-524.
- [28] H. Weyl, *Gravitation und Elektrizität*, Sitz. Kgl. Preuss. Akad. Wiss. (1918), 465.
- [29] H. Weyl, *Reine Infinitesimalgeometrie*, Math. Z. **2** (1918), 384-411.
- [30] H. Weyl, *Raum, Materie, Zeit*, Springer, Berlin (1918).
- [31] H. Weyl, *Elektron und Gravitation*, Z. Phys. **56** (1929), 330-352.
- [32] E. Witten, *Quantum Field Theory and the Jones polynomial*, Commun. Math. Phys. **121** (1989), 351-399.
- [33] C.N. Yang, R.L. Mills, *Conservation of isotopic spin and isotopic gauge invariance*, Phys. Rev. **96** (1954), 191-196.