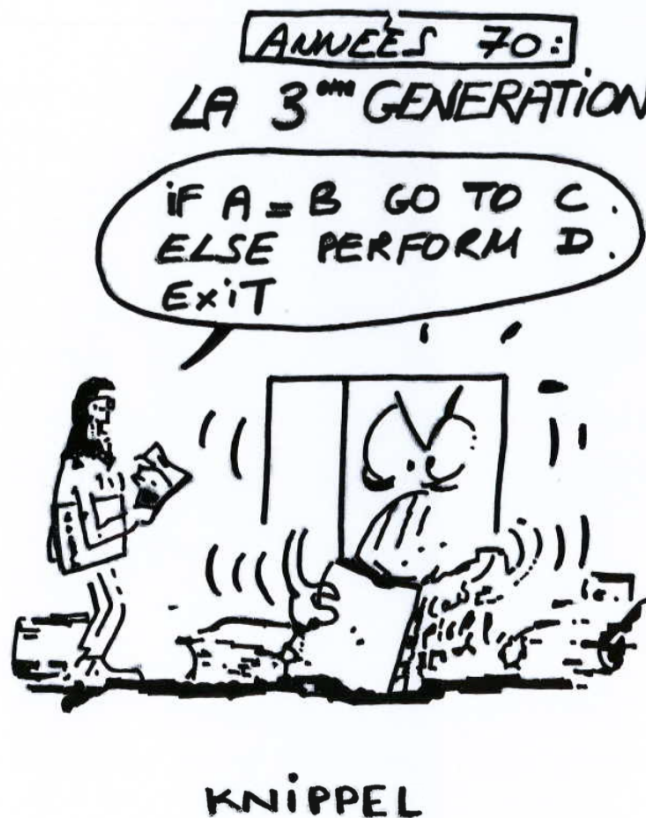


BIAA

Bulletin d'Informatique Approfondie & Applications
Computation - Information
Volumes 2015

100 - 101 - 102



B.I.A.A.

BULLETIN D'INFORMATIQUE APPROFONDIE ET APPLICATIONS
Revue fondée par Edmond Bianco
Publication trimestrielle de l'Université d'Aix-Marseille
ISSN 0291-5413

Le bulletin d'informatique approfondie et applications est une revue pluridisciplinaire destinée à éclairer les connaissances fondamentales informatiques. Les fondements sont un domaine vaste allant de la structure intérieure de l'ordinateur, où se matérialise la machine universelle, à l'algorithme qui devient programme, pour aboutir à la notion de système. Nous contribuons ainsi à ce que les autres disciplines plus anciennes (sciences humaines et de la société, sciences de la matière et de l'énergie, sciences mathématiques, sciences de la nature, sciences de la terre, sciences de l'univers, sciences de la vie, etc.) n'aient pas tendance à considérer l'informatique comme un simple outil définitivement figé. Il importe de continuer à maîtriser les développements fondamentaux de l'informatique pour que nos disciplines puissent en tirer un meilleur parti.

Notre publication est ouverte à l'ensemble de la communauté scientifique. Le périodique est diffusé vers les bibliothèques universitaires de France et vers quelques bibliothèques des cinq continents.

DIRECTEUR DE LA PUBLICATION

Jean - Michel Knippel

RESPONSABLE DE L'ÉDITION

Eric Olivier

SERVEUR DE PUBLICATION

Christian Blanvillain

SECRETARIAT

Kalassoumi Adjilani
Université d'Aix-Marseille
Site St Charles, Case 33
3 place Victor Hugo
F -13331 Marseille Cedex 3
Téléphone : +33 (0) 413 550 252
Télécopie : +33 (0) 491 509 110

DÉPOSITAIRE

Université d'Aix-Marseille
Bibliothèque Universitaire
Site St Charles
3 place Victor Hugo
F -13331 Marseille Cedex 3
Téléphone : +33 (0) 413 550 579
Télécopie : +33 (0) 491 957 557

IMPRIMEUR

Université d'Aix-Marseille
Service Reprographie
Site St Charles
3 place Victor Hugo
F -13331 Marseille Cedex 3
Téléphone : +33 (0) 413 550 626
Télécopie : +33 (0) 413 550 650

COMITÉ SCIENTIFIQUE

Pr. Patrick Abellard (Université du Sud, Toulon)
Françoise Adreit (Université de Toulouse I)
France Chappaz (Université de Provence)
Georges Chappaz (Université d'Aix-Marseille)
M'hamed Charifi (Consultant autonome)
Jean - Paul Coste (Université de Provence)
Pr. Roger Cusin (Université de la Méditerranée)
Jean - Claude Fumanal (Université Paul Cézanne)
Alain de Gantès (Université d'Aix-Marseille)
Jean Gonella (Université d'Aix-Marseille)
Pr. Bernard Goossens (Université de Perpignan)
Sami Hilala (Université d'Aix-Marseille)
Patrick Isoardi (Université d'Avignon et Pays de Vaucluse)
Robert Jacquier (Université Paul Cézanne)
Jean - Michel Knippel (Université d'Aix-Marseille)
Jean - Philippe Lehmann (Université d'Avignon et Pays de Vaucluse)
Pr. Agathe Merceron (Technische Fachhochschule, Berlin)
Nadia Mesli (Université d'Aix-Marseille)
Eric Olivier (Université d'Aix-Marseille)
Patrick Sanchez (C.N.R.S., Marseille)
Rolland Stutzmann (I.U.T. de Strasbourg Sud)
Pr. André Tricot (E.S.P.E., Toulouse)

CORRESPONDANT(E)S

Pr. Mohamed Tayeb Laskri (Université Badji Mokhtar, Afrique)
Sylvie Monjal (Cégep de Sainte Foy, anciennement Académie de Québec, Amériques)
Moussa HadjAli (Université Virtuelle de Syrie, Asie)
José Rouillard (Université des Sciences et Technologies de Lille, Europe)
Kalina Yacef (Université de Sydney, Océanie)

Table Des Matières

| | |
|--|----|
| BULLETIN n° 100 (2015) | 7 |
| ÉDITORIAL : Le cent et le sang (par Jean - Michel KNIPPEL) | 9 |
| Informatique, vérité, théorème de Gödel (par Edmond BIANCO) | 13 |
| 1. Informatique, vérité, théorème de Gödel | 13 |
| 2. Démonstration de Gödel | 18 |
| 3. Raisonnement | 19 |
| 4. Quelques réflexions | 19 |
| Références | 20 |
| 100 ans de Relativité Générale (I/III): covariance (par Eric OLIVIER) | 23 |
| 1. Introduction | 23 |
| 2. Changements de bases dans un espace vectoriel | 24 |
| 3. Dualité et dualité euclidienne | 25 |
| 4. Calcul tensoriel | 28 |
| 5. Variétés différentielles | 29 |
| 6. Crochet de Lie de deux champs de vecteurs | 32 |
| 7. Champs de tenseurs sur une variété | 35 |
| 8. Dérivations covariantes et connexions | 38 |
| 9. Variétés pseudo-riemanniennes et connexion de Levi-Civita | 40 |
| 10. Torsion et g -compatibilité d'une connexion | 44 |
| 11. Théorème d'existence et d'unicité de Koszul | 46 |
| 12. Transport parallèle | 48 |
| 13. Géodésiques inertielles : Théorème de Levi-Civita | 50 |
| 14. Dérivation covariante des champs de tenseurs | 52 |
| 15. Opérateurs différentiels | 54 |
| 16. Notes | 60 |
| 17. Appendices : Connexion de Levi-Civita d'une nappe de \mathbb{R}^3 | 60 |
| Références | 63 |
| La cosmologie newtonnienne (la minute d'Anselme Lanturlu) (par Jean-Pierre PETIT) | 65 |
| 1. Le grumeau pulsant | 65 |
| 2. Mythe de la caverne | 67 |
| Deux ou trois choses que je sais de Sakharov (la minute d'Anselme Lanturlu) (par Jean-Pierre PETIT) | 69 |
| VOUZZAVEDIBISAR : Injustice à haut débit (par RISS de CHARLIE HEBDO) | 73 |
| BULLETIN n° 101 (2015) | 75 |
| ÉDITORIAL : Faut-il laisser tomber l'allemand, le chinois ou l'espagnol et faire code informatique en seconde langue ? (par Jean-Pierre ARCHAMBAULT et Rémi BOULLE) | 77 |
| Contradiction ou indécidabilité : il faut choisir ! (par Patrick ISOARDI) | 83 |
| 1. Formaliser les mathématiques: le rêve de Hilbert | 83 |
| 2. Le Théorème d'incomplétude de Gödel | 89 |
| 3. Bibliographie | 94 |

| | |
|---|-----|
| Références | 95 |
| 100 ans de Relativité Générale (II/III): courbure (par Eric OLIVIER) | 97 |
| 1. Introduction | 97 |
| 2. Structures algébriques du tenseur de Riemann | 98 |
| 3. Tenseurs de torsion et de courbure d'une connexion | 101 |
| 4. Identités de Ricci | 103 |
| 5. Les symétries du tenseur de courbure | 105 |
| 6. Exponentielle de Riemann et coordonnées normales | 108 |
| 7. Le tenseur Ricci et l'équation d'Einstein | 110 |
| 8. Interprétation géodésique du tenseur de Riemann | 112 |
| 9. Courbure de Gauss | 117 |
| 10. Notes | 121 |
| Références | 122 |
| VOUZZAVEDIBISAR : Der Dirigent - Le directeur (par Annagret TAESCHNER) | 125 |
| BULLETIN n° 102 (2015) | 127 |
| ÉDITORIAL : Charles CORGE et INFORMATIQUE QUANTIQUE (par Jean - Michel KNIPPEL) | 129 |
| The Imitation Game : et si vous préféreriez l'original ? (par Jean LASSÈGUE) | 131 |
| 100 ans de Relativité Générale (III/III): déflexion de la lumière (par Eric OLIVIER) | 139 |
| 1. Introduction | 139 |
| 2. Première constante du mouvement newtonien : l'énergie | 141 |
| 3. Deuxième constante du mouvement newtonien: le moment cinétique | 142 |
| 4. Troisième constante du mouvement : le vecteur d'excentricité | 144 |
| 5. Déflexion newtonienne I | 146 |
| 6. Déflexion newtonienne II | 148 |
| 7. Déflexion newtonienne III | 149 |
| 8. La métrique de Schwarzschild | 151 |
| 9. Déflexion de la lumière pour la Relativité Générale | 156 |
| 10. Notes | 161 |
| Références | 162 |
| 11. Postscriptum : Textes choisis | 164 |
| VOUZZAVEDIBISAR : Videac. Télétruc. Janvier 1982 (par Jean-Pierre PETIT) | 169 |

BIAA

Bulletin d'Informatique Approfondie & Applications
Computation - Information
Volumes 2015

100 - 101 - 102



Couverture : dessin de Michel Avezard (Zevard) interprété par la MMIAGe (1985-1987)

ÉDITORIAL : Le cent et le sang

Jean - Michel KNIPPEL

Résumé. – Il s'en est passé des événements dans ce premier trimestre 2015. Le 7 janvier 2015, attentat contre Charlie Hebdo. Douze personnes trouvent la mort et onze personnes blessées. Le sang a coulé. Quant au bulletin, il pensait fêter dans la joie son numéro cent, le 100, mais il y a eu un numéro zéro, le 0. Nous en sommes au numéro 101, en réalité. Quoi d'autre le 7 janvier 2015 ?

Accueillis provisoirement par le bateau "Je Sers", une partie des 80 demandeurs d'asile et réfugiés tibétains ont été pris en charge par la ville de Conflans-Sainte-Honorine, le 7 janvier 2015 ... en attendant mieux. Ils sont en France depuis plus de trois ans pour certains. Plus d'une centaine de Tibétains, réfugiés et demandeurs d'asile, ont pour abris de fortune des tentes installées dans les niches du pont de Conflans-Sainte-Honorine qui enjambe la Seine. D'autres, quatre-vingts plus précisément, sont accueillis depuis le début de l'hiver par le bateau-chapelle "Je Sers", situé non loin de là. Mais les trois péniches reliées les unes aux autres ne peuvent tous les prendre en charge, et leur situation est devenue urgente en ces jours de grand froid.

Solution provisoire ?

« Nous accueillons ces femmes, ces hommes et ces enfants tibétains, entre autres. Nous leur fournissons un endroit pour dormir, de quoi manger et nous les accompagnons dans leurs démarches administratives, confie Hugues Fresnau, en charge de l'accueil des Tibétains sur le bateau. Mais notre capacité d'accueil n'est pas infinie, et même si quelques-uns sont pris en charge par les urgences sociales, il faut rapidement trouver une solution pour que ces gens ne soient pas livrés à eux-mêmes en plein hiver. »¹

Depuis l'attentat contre "Charlie Hebdo", le 7 janvier dernier, Tom Lafon achète chaque semaine le journal satirique. Il a 10 ans... Tom n'était pas né le 11 septembre 2001. Tom avait 10 ans le 7 janvier 2015 et il n'a pas tout de suite compris ce qu'il venait de se passer. Alors, il a fait comme beaucoup de monde, les adultes surtout, et il a regardé en boucle les images diffusées à la télévision. Il voulait comprendre pourquoi une telle horreur. Passionné d'histoire du haut de ses 10 ans, le monde l'intéresse. La Seconde Guerre mondiale, les Tours jumelles, la tuerie à Charlie Hebdo... il veut savoir ? : « On ne tue pas des gens juste parce qu'ils défendent la liberté d'expression. C'est injuste ? ! »

Alors, quand le numéro 1.178 de Charlie Hebdo est sorti, il s'est précipité à la librairie d'Issoudun pour l'acheter : « J'avais demandé à mes parents s'ils voulaient bien me donner 3 euros pour l'acheter et ils l'ont fait tout de suite, sans me poser de question. » Mais quand il est arrivé, il n'y en avait déjà plus. « J'ai dû revenir trois fois mais j'ai fini

1. <http://www.gazettevaldoise.fr/2015/01/07/les-tibetains-loggeront-salle-eugene-lecorre/>

par l'avoir. Je suis rentré à la maison et je l'ai lu dans ma chambre. Ce sont les dessins que j'aime surtout. Ils sont trop drôles. »

Matthieu Ricard montre que la violence n'a cessé de baisser depuis des siècles. Par exemple, le nombre d'homicides en Angleterre pour 100.000 habitants était de 110 en 1350, de 10 au XVI^e siècle et désormais de un. « En Europe, on a entre cinquante et cent fois moins de chances d'être tué », assure le moine bouddhiste, qui rappelle que « c'était un spectacle d'aller voir les pendus au Moyen-Âge. A tel point qu'un village français a loué un condamné à mort à une autre ville parce que les gens se plaignaient qu'il n'y avait pas eu d'exécution publique depuis longtemps. »²

L'aventure du Bulletin d'Informatique Approfondie et Applications a commencé en 1981 avec un numéro 0 en mars 1981 autour d'Edmond Bianco au Département de Mathématiques et d'Informatique de la Faculté des Sciences de Luminy à Marseille. L'équipe du Laboratoire d'Informatique Théorique de Luminy composait le socle de la publication. Son comité de rédaction s'articulait autour d'Edmond Bianco avec Gilbert Cousin, Francis Donnat, Patrick Isoardi, Jean-Philippe Lehmann, Joseph Roller et Rolland Stuzmann. Les séminaires du mercredi après midi permettaient de débattre des futures articles et thèmes de publication : construction de compilateurs, théorie des systèmes, conduite de processus, etc.

Le temps a passé et nous sommes au numéro 100. Dans l'Avant - Propos du numéro 0, la Rédaction écrivait « Le plus simple, nous semble-t-il, pour présenter un nouveau bulletin est encore d'en réaliser un premier numéro. N'en déplaise aux jeteurs de sort, nous essayerons de faire en sorte qu'il y en ait un second. Au moins. Pourquoi vouloir à tout pris essayer d'expliquer, essayer d'introduire : ce bulletin existe et s'il doit vivre il vivra, sinon tant pis. »

Nous faisons comme Charlie Hebdo, nous continuons « ... C'EST REPARTI! » figure dans la couverture du numéro 1179 daté du 25 février 2015 de Charlie Hebdo. Pour nous, ... C'EST PARTI! POUR LES NUMEROS 100 ET 101 ET 102 de la collection 2015 du bulletin. Notre aventure « comme la nôtre ne réclame qu'optimisme, beaucoup d'optimisme ». ³

La théorie de la relativité a 100 ans en 2015 et voilà Eric Olivier parti sur les traces d'Albert Einstein et Jean-Pierre Petit pour s'approcher des frontières de l'informatique quantique. Certes, « que la théorie d'Einstein remplace avantageusement celle de Newton, qu'elle constitue un cadre de pensée infiniment plus fécond, personne n'en doute. »⁴ Le plus important problème auquel la relativité est aujourd'hui confrontée est celui de la conciliation avec la physique quantique. « Hélas, la théorie quantique n'est pas compatible avec la relativité générale. Tenter de conjointre ces deux fondements de notre physique est une question considérée comme essentielle depuis près d'un siècle. Nous n'avons toujours pas aujourd'hui de réponse tout à fait satisfaisante. »⁵

2. <http://www.franceinfo.fr/emission/tout-et-son-contraire/2014-ete/matthieu-ricard-les-sages-ne-sont-plus-des-modeles-07-22-2014-05-45>

3. Bulletin numéro 0. Avant-Propos

4. Charles Corge. *L'informatique quantique qu'est-ce et pourquoi faire ?* Editions Ellipses. 2012

5. Aurélien Barrau. Professeur à l'Université Joseph Fourier de Grenoble

A mes yeux, il est naturel de rassembler des réflexions sur des travaux d'Einstein et de Gödel, car le dénominateur commun est le doute. Alan Turing écrivait « Einstein met en doute les axiomes d'Euclide quand ils s'appliquent à des corps solides ... »⁶

Nous ne sommes pas certain de pouvoir contacter directement Albert Einstein et Kurt Gödel pour relire les articles que nous publions de nouveau dans ce numéro, textes d'Edmond Bianco de 1991 et de Jean-Pierre Petit sortis en 1990. Edmond Bianco a rejoint Albert Einstein, Kurt Gödel et André Lentin, entre autres, où ? Quant à Patrick Isoardi, nous pouvons le joindre directement en pays de Drôme et Jean-Pierre Petit en pays de Pertuis.

Alors Optimisme, bon sang !

6. Andrew Hodges. *Alan Turing ou l'énigme de l'intelligence* Editions Payot. 1988

Informatique, vérité, théorème de Gödel

Edmond BIANCO¹

Résumé. – Si, comme pourraient le prétendre les médias, nous pénétrons dans l'ère de l'intelligence artificielle, il est bon, loin de vouloir briser les élans d'imagination, d'observer quelque peu les limites du domaine dans lequel on patauge. Il est prudent de voir clairement ce qu'on ne peut pas faire. A l'imagination l'initiative de trouver ce qu'on peut faire. En évitant l'écueil qui consiste à ramener la pensée informatique à une liste ouverte de recettes, il devient plus que nécessaire de montrer que l'informatique a sa méthodologie propre, et de mettre en évidence tout ce qui est commun dans les théories de Turing, Post, Von Neumann, Gödel et autres. Or, si abstraite que soit la démonstration de Gödel, elle utilise simplement une technique de base de l'informatique. Cela, les contemporains de Gödel ne pouvaient le voir.

1. Informatique, vérité, théorème de Gödel

En écrivant ces mots, je n'ai nulle prétention à redécouvrir le Théorème de Gödel, je préfère renvoyer à l'excellent ouvrage de J.- Y. Girard [GNNG89] cité en référence. Mon souci est autre. J'ai eu la chance de voir naître ce phénomène qu'on a nommé sur le tard "Informatique", qui a, je pense avec juste raison, englobé de nombreux phénomènes qui se sont petit à petit distingués les uns des autres tout en conservant, cependant, comme point commun, l'objet qu'on a rebaptisé toujours sur le tard : ordinateur. Mais par un étrange métaphénomène c'est l'une de ces branches du savoir, certes économiquement la plus visible, la programmation, qui masque pratiquement toutes les autres. Un peu comme si l'on prétendait que la "comptabilité" c'est les mathématiques.

Dans le théorème de Gödel, plusieurs domaines importants de la réflexion peuvent être dégagés, ou même engagés. Il me semble qu'il en est un autre pour lequel il est intéressant d'insister.

Quand on a commencé à se poser des questions sur la signification des choses et des moyens d'expression, c'est devenu un jeu amusant que de construire des paradoxes. Le paradoxe du menteur en est un exemple célèbre : Epiménide le Crétois dit de lui-même qu'il est un menteur. Le mot "menteur" possède de toute évidence plusieurs sens distincts dont un autre exemple apparaît dans le problème du voyageur. Les pérégrinations de ce dernier l'amènent à une fourche, et là il ne sait plus où aller. Heureusement se trouvent au carrefour un couple d'individus capables de le renseigner, mais l'un d'eux, on ne sait pas lequel, est un menteur, quelle est alors la question à poser. La solution brutale, implique que le menteur, fieffé menteur, dit systématiquement des contre-vérités. Peut-on alors continuer à le qualifier de menteur, à partir du moment où il est démasqué, un simple décodage le rend autant vecteur de vérité que son comparse. Le véritable mensonge est plus subtil, il consiste en fait, à utiliser essentiellement la vérité, et n'en travestir que juste ce qu'il faut pour atteindre un but hypothétique, tout en restant vraisemblable.

1. Note de la rédaction : article paru dans le numéro 30 du bulletin de décembre 1991.

L'antinomie de Russel met en évidence une sorte de manœuvre linguistique un peu voisine. Je définis un ensemble normal comme ne se contenant pas lui-même. Il est facile de trouver des exemples : l'ensemble de tous les clous n'est pas un clou. Cela implique immédiatement que je me mette à imaginer des ensembles, alors non normaux, qui se contiennent eux-mêmes. Cette notion est moins évidente, il devient nécessaire d'illustrer la propriété. Et force est bien de chercher une définition assez vague pour que ça marche. Exemple. L'ensemble des choses pensables, est une chose pensable. Donc l'ensemble des choses pensables est un élément de lui-même. Devant le vague de ce que représente une chose pensable, on ressent le besoin de trouver une idée plus précise. L'ensemble des objets rouges est un objet rouge. L'ensemble des masses de sable est une masse de sable. On perçoit qu'on frôle à chaque instant le jeu de mots. "Masse" passe beaucoup mieux que "objet", tout simplement parce que le mot objet sous-entend presque nettement une unicité dans la constitution. Il est alors réjouissant de se préoccuper de savoir, si l'ensemble de tous les ensembles normaux est lui aussi normal. S'il l'est, il tombe sous le sens qu'étant normal, nouvelle définition vague, il est aussi un élément de l'ensemble. Mais ceci est précisément la définition d'un ensemble non normal. Donc s'il est normal alors, conclusion, il est non normal.

Le paradoxe marche d'autant mieux que le mot pivot possède un sens très large, il est visiblement d'ordre sémantique, des termes comme volume, surface font très bien l'affaire.

En toute généralité, un raisonnement sera d'autant plus rigoureux qu'il portera sur des notions plus précisément définies, et surtout très soigneusement délimitées. C'est précisément, à la limite, le cas des mathématiques, dont l'efficacité consiste à vider les mots de toute substance afin de n'en conserver que le squelette. Le théorème de Gödel intervient alors pour montrer l'importance du hiatus qui existe entre ce dont on parle, et ce qui permet d'en parler. Le langage et le méta-langage.

Le même hiatus existe dans la société, mais dans un domaine d'application légèrement différent. Prenons le cas de l'enseignement des mathématiques. Nombre d'élèves ont du mal à se faire à l'idée d'utiliser des mots vides de sens, pour eux le mot conserve toujours un peu de sa richesse, et l'application du raisonnement mathématique devient alors difficile voire impossible. L'enseignant, oubliant un peu les difficultés qu'il a vécues dans son passé, oublie en même temps d'insister lourdement sur l'importance du côté conventionnel. De telle sorte que bien trop souvent des élèves doués d'une intelligence fine se trouvent rejetés par l'aridité de la manipulation de concepts vides. Tout le monde ne peut être passionné par l'emballage, juste pour l'intérêt de l'emballage. Et cela explique également qu'on puisse revenir sur le tard aux mathématiques, quand on a pu surmonter l'écœurement d'un enseignement exsangue. Prenons un exemple. Les mots "vrai" et "faux". Ils sont définis sans ambiguïté en logique. Quelque chose qui est vrai est définitivement vrai. Point. Quand j'apprends les mathématiques, je me rends compte qu'il y a deux univers, l'univers mathématique, où tout est rigoureusement "vrai", bref un univers pur, cristallin. Et surtout ceux qui peuvent y accéder sont considérés comme les plus intelligents. Et puis l'autre univers où rien ne peut jamais être aussi systématique, toutes les vérités sont non seulement fluctuantes, mais encore jamais nettement

établies. Telles théories politiques semblent permettre de résoudre de graves conflits et on s'aperçoit à l'usage qu'elles ont tendance à les aggraver. L'élève qui, pour diverses raisons commence dans la vie avec une étiquette de cancre, se trouve dans la situation difficile suivante : s'il admet ce fait, c'est visiblement définitif, puisque toute vérité est éternelle, alors pourquoi essayer de sortir du puits ? S'il ne l'admet pas, il va se heurter à une sorte de barrière de potentiel, à son premier petit succès il va recevoir des quolibets indirects du style : comment avez-vous pu rater cet exercice alors que même lui a pu le comprendre. La vérité essentielle demeure : crétin je suis, crétin je reste. Devenir fort en math implique de rendre fausse une vérité définitive, ce qui revient à remettre en question la mathématique elle-même. Essayons alors d'imaginer la frustration du cancre qui perçoit nettement qu'il ne pourra jamais participer à l'élaboration des bombes à billes, à la dispersion de la dioxine ou à la conception d'un surrégénérateur.

Le raisonnement formel implique deux idées de base, d'abord que les vérités établies se conservent pour tous les systèmes d'objets qui sont munis des propriétés exigées au départ. Ensuite, il faut être sûr que le raisonnement, dans le domaine choisi, n'aboutira jamais à une contradiction. Par exemple que sur tel type de figure on réussisse à montrer que deux droites à la fois se coupent et ne se coupent pas.

Il y a ainsi de nombreux chapitres des mathématiques pour lesquels l'expérience des mathématiciens, et la qualité des résultats obtenus permettent seuls, jusqu'à présent, d'en affirmer la cohérence.

Mais à partir de là, l'idée a germé selon laquelle le processus du raisonnement mathématique est un processus mécanique, en foi de quoi il devient judicieux d'en concevoir le mécanisme. Première étape l'axiomatisation. On se donne un ensemble fini de propriétés de base, et un jeu de règles de déduction. De quelle nature est alors le mécanisme ? Eh bien il suffit de se donner les moyens de représenter les axiomes et de se donner les moyens d'en déduire tous les théorèmes possibles et imaginables, par application mécanique des règles de déduction, elles-mêmes dûment représentées pour qu'une machine puisse faire tout le travail.

Comment se présente alors le problème ? Nous nous donnons de la sorte des signes et des jeux de signes auxquels nous attribuons une signification arbitraire, mais telle qu'on puisse montrer que toutes les "bonnes constructions", ainsi définies par notre système mécanique comportent un sens cohérent. Hélas Gödel a montré que cette "méta-assertion" est fausse. Tant pis pour la mécanique, tant mieux pour l'esprit humain.

Le principe de la démonstration passe par la définition d'un "codage" dont on est certain qu'il permet de représenter de manière bien distincte des objets formellement distincts. Ensuite on se place par exemple dans le cadre de l'arithmétique de Péano, et on montre qu'on peut construire une formule, parfaitement concevable, mais dont il est visible qu'elle déclare d'elle-même qu'elle est fausse.

On se donne le moyen mécanique de construire et de reconnaître toutes les formules rigoureusement vraies du système. Ce sont des considérations méta-linguistiques qui permettent alors de les interpréter. Il doit être parfaitement clair que seules les vérités telles qu'elles apparaissent dans le méta-langage sont importantes puisque c'est comme

ça que le système est utile. Il doit être un support écrit non ambigu dans lequel il suffit de lire une formule construite, déclarée vraie, pour lire une vérité utilisable.

Ainsi se plaçant dans un système connu et cohérent on prouve qu'on peut, en respectant les règles de ce système, construire au moins une formule reconnue vraie, qui, si on l'en croit en l'interprétant, signifie qu'elle-même est fausse.

Cela ne prouve qu'une seule chose, c'est que dès lors qu'on a construit un système formel, donc constructible mécaniquement, on sépare l'univers de la réflexion en deux parties : celle qui est contenue dans le système et qui est capable de distinguer ce qui est intérieurement vrai de ce qui ne l'est pas, et celle qui permet d'observer le système de l'extérieur et d'en déduire certaines propriétés que le mécanisme ne peut pas mettre en évidence, ce sont les méta-propriétés. Ce que montre Gödel, c'est que, vouloir inclure les méta-propriétés dans le système lui-même, le rend inconsistant. Il suffit de montrer qu'il est possible, en respectant les règles du système, d'y inclure une propriété contradictoire.

Rappelons les axiomes de Péano :

- 1) Zéro est un nombre.
- 2) Le successeur immédiat d'un nombre est un nombre.
- 3) Zéro n'est pas le successeur immédiat d'un nombre.
- 4) Il n'existe pas deux nombres distincts qui possèdent le même successeur.
- 5) Toute propriété qui appartient à Zéro et à tout successeur d'un nombre qui a cette propriété, appartient à tous les nombres.

On va choisir un ensemble suffisant de signes et de variables et un codage biunivoque : d'abord dix signes :

0 s = $\exists \cup \supset \neg ()$, ceci n'est pas limitatif mais doit suffire, puis des variables à substitution numérique :

x y z ...

des variables à substitution propositionnelle :

p q r ...

des variables prédicatives :

P Q R ...

Toutes ces variables en quantité suffisante et dénombrable. On code les dix signes avec les entiers 1, 2, 3, ... 10.

0 s = $\exists \cup \supset \neg ()$,
1 2 3 4 5 6 7 8 9 10

pour les variables numériques on choisit les nombres premiers plus grands que 10 :

x y z ...
11 13 17 19 ...

pour les variables de proposition, les carrés de ces nombres premiers :

p q r ...
 $11^2 13^2 17^2 19^2$...

pour les variables prédicatives, les cubes de ces mêmes nombres :

P Q R ...
 $11^3 13^3 17^3 19^3$...

On peut de la sorte coder toutes les formules dont on a besoin, par exemple :

$$(\exists x)(x = s 0)$$

qui se lit : il existe un x tel que ce x soit le successeur immédiat de zéro, formule tirée des axiomes 1 et 2 de Péano. Je peux également construire la formule :

$$(\exists x)(x = s y)$$

qui est l'expression de l'axiome 2. On peut coder ces formules puisque chacun de leurs signes porte un code :

$$\begin{array}{cccccccc} (& \exists & x &) & (& x & = & s & 0 &) \\ 8 & 4 & 11 & 9 & 8 & 11 & 3 & 2 & 1 & 9 \end{array}$$

On cherche alors à coder une telle formule de manière unique et telle que toute autre formule qui en différerait si peu que ce soit porte un code distinct. On peut procéder ainsi : ces entiers servent d'exposant aux premiers successifs et on fait le produit de l'ensemble :

$$2^8 \times 3^4 \times 5^{11} \times 7^9 \times 11^8 \times 13^{11} \times 17^3 \times 19^2 \times 23^1 \times 29^9 = n1$$

Il est parfaitement clair que si je transforme si peu que ce soit ma formule, son code devient très différent. Imaginons qu'on permute = et s. Alors dans le nombre précédent le facteur :

$$17^3 \times 19^2$$

devient :

$$17^2 \times 19^3$$

Si on se réfère au théorème de l'arithmétique qui dit que tout nombre entier se décompose en une suite unique de facteurs premiers, on constate que l'on obtient pour ces deux formules deux codes distincts. Réciproquement partant d'un tel code, la décomposition en facteurs premiers permet de retrouver facilement la formule. Tout nombre qui est l'image d'une formule du système est désigné par "nombre de Gödel" de la formule. Je dirai indistinctement "code de" ou "nombre de Gödel de". De la même manière, la deuxième formule portera le nombre de Gödel :

$$(\exists x)(x = s y)$$

$$2^8 \times 3^4 \times 5^{11} \times 7^9 \times 11^8 \times 13^{11} \times 17^3 \times 19^2 \times \mathbf{23^{13}} \times 29^9 = n2$$

Ce code diffère du précédent par le facteur marqué en gras. Désignons par n1 et n2 ces deux codes. On peut être amenés à se donner un code pour la suite des deux formules, ou, plus généralement, pour toute suite quelconque de formules. Cela permettrait d'obtenir également un nombre de Gödel unique pour chacune des suites possibles, qu'on appelle plutôt des démonstrations. Ainsi si je considère que la suite des deux règles précédentes constitue une démonstration, j'en construis le code de la manière unique, en prenant le nombre de Gödel de chaque formule et en le portant en exposant de la suite des nombres premiers à partir de 2. De sorte que la démonstration :

$$(\exists x)(x = s y)$$

$$(\exists x)(x = s 0)$$

va porter le nombre de Gödel :

$$2^{n1} \times 3^{n2} = k$$

visiblement, le raisonnement déjà appliqué marche encore, avec ce procédé deux démonstrations différentes porteront obligatoirement des codes différents. C'est ce qu'on peut constater en intervertissant simplement l'ordre des deux formules.

2. Démonstration de Gödel

On se donne alors deux moyens pour la démonstration. D'abord je constate que je peux coder les deux formules suivantes :

$$F1(x) (\text{dém} (k, n1))$$

$$F2(x) (\neg \text{dém} (k, n1))$$

qui se lisent, la première : « pour tout x, la suite de formules de code k est une démonstration de la formule de code n1. »

La seconde : « pour tout x, la suite de formules de code k n'est pas une démonstration de la formule de code n1. »

J'ai là deux énoncés méta-linguistique que je peux parfaitement coder dans le système mécanique, ce sont des formules qui peuvent être construites automatiquement pour tout n1 possible et pour tout k correspondant.

Il est à ce moment-là évident que si je prends au hasard une suite quelconque de formules valables cette suite ne sera généralement pas la démonstration de la dernière formule, et dans ce cas F2 est une formule "vraie".

Je me donne alors un autre moyen de calculer. Si j'observe la formule :

$$(\exists x)(x = s y)$$

je constate que ce qu'elle exprime, c'est qu'il existe toujours un entier successeur immédiat d'un entier quelconque, donc je peux en déduire autant de formules parfaitement vraies dans notre système, au moyen du remplacement de y par n'importe quel entier. Ainsi si je veux dire que 4 est successeur de 3², j'ai la formule vraie :

$$(\exists x)(x = sss 0)$$

avec les notations choisies 3 s'écrit évidemment sss 0. Mais tout autre moyen de faire apparaître un entier quelconque conserve à la formule sa valeur de vérité. En particulier si je remplace le code de y par le code de l'entier code de n2. Auquel cas je donne un nom à cette opération qui consiste à remplacer le code de y par le code du nombre de Gödel de la formule. La nouvelle formule porte à ce moment là un autre nombre de Gödel qu'on peut facilement calculer.

$$2^8 \times 3^4 \times 5^{11} \times 7^9 \times 11^8 \times 13^{11} \times 17^3 \times 19^2 \times \mathbf{23^{13}} \times 29^9 = n2$$

C'est le terme $\mathbf{23^{13}}$ qu'il faut remplacer par une suite :

$$23^2 \times 29^2 \times \dots \times p^2$$

si p est le n1ème premier à partir de 23. La nouvelle formule portera donc un code qu'on désignera par N2. Cette opération qui n'est en fait qu'une simple substitution, sera désignée par la notation :

$$\text{subst} (n2, \mathbf{23^{13}}, n2)$$

qui se lit : « Substitution dans la formule de nombre de Gödel n2, du code de l'entier n2 au terme $\mathbf{23^{13}}$. »

Cette opération donne une formule de code N2. Je reprends alors la formule :

$$F2(k) (\neg \text{dém} (k, n2))$$

Cette formule est en fait une démonstration dans laquelle je remplace la dernière formule de code n2, par la même formule sur laquelle j'ai appliqué la substitution. Cela donne :

2. Note de la rédaction : correction introduite en gras souligné

$F2'(k) (\neg \text{dém}(k, \text{subst}(k, 23^{13}, k)))$

La formule $F2$ porte par définition le code k , j'applique la substitution sur la formule complète $F2$, et cela aura pour effet de remplacer l'élément 23^{13} par le code, obtenu comme on l'a vu, de l'entier k . L'application de ce calcul donne comme code un nombre de Gödel facile à calculer qu'on désignera par K .

On se demande alors quel est le code ou nombre de Gödel de $F2'$. La nature de l'opération

$\text{subst}(k, 23^{13}, k)$

montre de toute évidence que K est le nombre de Gödel de $F2'$.

3. Raisonnement

Par construction on a vu qu'à chaque formule correspond un seul nombre de Gödel, de plus chaque fois qu'on rencontre un nombre de Gödel, on peut reconstituer la formule dont il est le code. Il est alors clair que si je sais qu'une déduction provenant d'une suite de formules vraies amène à une formule également vraie, si un code qui correspond à la déduction entraîne un code qui correspond à la conclusion, on peut méta-déduire que ce code exprime une vérité dans le système.

Or, si j'interprète la formule $F2'$ voici ce qu'elle dit :

Hypothèse : « Quel que soit k la formule $F2'$ n'est pas une démonstration de la dernière formule. »

Conclusion : « Or, cette dernière formule porte comme nombre de Gödel précisément le nombre de Gödel de $F2'$. Qui affirme donc qu'elle n'est pas une démonstration d'elle-même, et ceci quel que soit k . Il n'existe donc pas de démonstration de $F2'$.

Et c'est $F2'$ qui le dit. »

4. Quelques réflexions

Pour construire un mécanisme, quel qu'il puisse être, il faut avoir à l'avance un ensemble d'idées. Pour fixer ces idées il devient nécessaire de trouver un moyen graphique de les exprimer. D'où l'invention du symbole. Mais le symbole n'est pas la pensée, car cela voudrait dire que les peuples qui n'ont pas d'écriture ne pensent pas. Mais il existe toutefois de bonnes raisons d'être inquiet quand on voit ce que pensent les peuples qui en ont une.

Quels que soient les signes employés, ou leurs combinaisons, ce n'est que par ce qu'on leur accole de signification arbitraire, qu'ils peuvent être intéressants. Le symbole en lui-même n'a aucun pouvoir, aucune valeur. Il n'en est pas de même de son pouvoir évocateur, qui implique l'existence d'une imagination observante. Quand on se livre à la construction d'un système formel, que fait-on en réalité ? On essaye désespérément de supprimer toutes les ambiguïtés que l'on considère comme gênantes. Il paraît ainsi naturel de se priver d'une source riche, sinon importante d'expression.

Le système tronqué de la sorte ne devra vraisemblablement plus pouvoir exprimer un certain nombre de vérités utiles mais d'un niveau qui dépasse les possibilités laissées au système. Or, c'est bien le phénomène que cette démonstration met en évidence. Les lois de construction des images de vérités permettent de construire des images que le système

est parfaitement incapable d'interpréter. La force de la pensée de Gödel réside dans le fait qu'il a créé un raisonnement de type informatique. D'abord il fabrique un code dont la propriété essentielle est la biunivocité. L'algorithme de construction de son code, basé sur un méta-théorème de l'arithmétique, fonctionne dans les deux sens. Partant du code il peut refabriquer la formule.

On ne fait rien d'autre quand on utilise un "langage machine" qui présente toujours deux formes, l'une plus commode pour le programmeur, comporte un maximum d'objets symboliques tels que des identificateurs. Mais c'est un code finalement de forme binaire, pour l'autre, donc enregistrable en mémoire et qui sera soumis à la machine universelle.

Une démonstration n'est en fait, qu'une construction de formules et une vérification qu'une formule obtenue est une bonne formule. Informatiquement on peut imaginer une sorte de machine universelle qui, à la donnée des axiomes de Péano dûment codés, et à la lecture d'un jeu adéquat de règles de déduction également codées construirait toutes les formules possibles.

C'est ainsi que partant des deux objets :

$$(\exists x)(x = s 0)$$

$$(\exists x)(x = s y)$$

par une construction combinatoire, il est clair que la machine en question tombera obligatoirement sur une formule du type F2'. Ce qu'il faut entendre par là c'est qu'elle va construire tous les nombres de Gödel possibles jusqu'à un certain ordre et on peut montrer facilement qu'il finira par tomber sur le code d'une version de F2'. Mais ce que l'automate ne possède pas c'est le moyen de relever ce que cette formule a de remarquable. Pour lui, c'est une formule vraie, donc un théorème. Il n'a aucun moyen d'en saisir le méta-sens. C'est le contre-exemple qui montre qu'il existe dans le système, des formules qu'il considère comme "vraies" mais dont l'interprétation "normale" n'a aucune signification. Il ne s'agit là, en fait que d'une "méta-propriété" supplémentaire, et pas tellement évidente.

Note de la rédaction : [GNNG89] est la seule référence explicitement citée par Edmond Bianco. Les autres références, qui se rapportent au sujet ont été insérées, font partie de sa bibliothèque : [Pea89], [Göd31],[Göd62], [CN04], [Wan90], [And15], [Del02], [Hof85], [Göd56], [You05], [Göd65].

RÉFÉRENCES

- [And15] D. Andler. *Kurt Gödel, 1906-1978*. Encyclopaedia Universalis [en ligne], 2015.
- [CN04] P. Cassou-Noguès. *Gödel*. Paris, Les Belles Lettres, 2004.
- [Del02] J.-P. Delahaye. *L'intelligence et le calcul, de Gödel aux ordinateurs quantiques*. Belin. Pour la Science, 2002.
- [GNNG89] K. Gödel, E. Nagel, J. Newman, and J.-Y. Girard. *Le Théorème de Gödel*. Seuil, 1989.
- [Göd31] K. Gödel. Über formal unentscheidbare Sätze der principia mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, 38 :173–198, 1931.
- [Göd56] K. Gödel. *Gödel's Lost Letter and P=NP : a personal view of the theory of computation*. Richard J. Lipton, 1956.
- [Göd62] K. Gödel. *On Formally Undecidable Propositions*. New York, Basic Books, 1962.

- [Göd65] Kurt Gödel. On undecidable propositions of formal mathematical systems, (1934) lecture notes taken by Kleene and Rosser at the Institute for Advanced Study : reprinted in M. Davis (ed.), 1965.
- [Hof85] D. Hofstadter. *Gödel Escher Bach. Les brins d'une guirlande éternelle*. Interéditions, 1985.
- [Pea89] G. Peano. *Arithmetices principia, nova methodo exposita*. Turin, Bocca, 1889.
- [Wan90] H. Wang. *Gödel*. Armand Colin, 1990.
- [You05] P. Yourgrau. *Einstein/Gödel. Quand deux génies refont le monde*. Dunod, 2005.

100 ans de Relativité Générale (I/III) : covariance

Eric OLIVIER^{1,2}

Résumé. – Einstein publie l'article fondateur de la Relativité Générale en 1916 – il y a 100 ans. C'est l'aboutissement de 10 ans de travail qui lui permettent d'intégrer la gravitation à l'espace-temps de la Relativité Restreinte (1905). La manière dont les idées scientifiques d'Einstein se développent est très liée à sa personnalité : en étudiant la Relativité Générale, on est frappé par la manière dont s'articulent des idées d'apparence naïve à la technicité mathématique. En trois notes, nous partons des fondements mathématiques de la Relativité Générale, afin de comprendre l'effet de la gravitation sur les rayons lumineux. La première note est consacrée à la covariance, c'est-à-dire au calcul tensoriel : cela nous permettra d'introduire les symboles de Christoffel, la connexion de Levi-Civita et la caractérisation des géodésiques inertielles par le transport parallèle (Théorème de Levi-Civita). La deuxième note [Oli15a] traite de la courbure riemannienne : c'est le point crucial de la géométrie riemannienne qui donne tout son sens à l'équation d'Einstein de la Relativité Générale. Enfin dans la troisième note [Oli15b], nous abordons la déflexion des rayons lumineux, d'abord du point de vue newtonien, puis en donnant la résolution des équations d'Einstein par Schwarzschild : en application, nous donnons le calcul relativiste complet de la déflexion des rayons lumineux rasant la surface solaire.

1. Introduction

Lorsqu'il commence à réfléchir à la manière d'intégrer la gravitation à l'espace-temps de la Relativité Restreinte, Einstein bute sur des difficultés mathématiques très sérieuses. C'est un ami proche (Marcel Grossmann) qui l'oriente vers les tous nouveaux concepts de la géométrie différentielle et du calcul tensoriel. Plus précisément, c'est la notion de *courbure* introduite par Gauss pour les surfaces, puis développée par Riemann pour des objets géométriques plus généraux (les variétés pseudo-riemanniennes), qui semble le bon concept pour généraliser la Relativité Restreinte. Einstein réalise que la *force centrifuge ressentie* sur un manège tournant, doit être de nature gravitationnelle : sa Relativité Restreinte ne prend pas en compte un référentiel qui serait en rotation par rapport à un référentiel lorentzien, et c'est là que la gravitation se dévoile ! Il fera aussi souvent allusion à *l'expérience de pensée* où un *observateur* se trouve en chute libre dans une cabine d'ascenseur, constatant – par la pensée – l'annulation de la gravitation terrestre. Là encore, le référentiel lié à la cabine ne peut être considéré comme lorentzien, du fait de son accélération uniforme par rapport au référentiel terrestre. Ces réflexions semblent indiquer que la gravitation est liée au référentiel utilisé, ou plus concrètement à la manière dont on mesure les positions dans l'espace-temps. La grande idée de la Relativité Générale est que les lois de la physique (à commencer par celles de la gravitation) doivent être exprimées indépendamment du référentiel : en d'autres termes – et contrairement aux Relativités Galiléenne et Restreinte – aucune sous-classe de référentiels ne peut être privilégiée au sein de l'ensemble des référentiels possibles. Selon la terminologie d'Einstein, la Relativité Générale se comprend en termes de lois *covariantes* (terminologie introduite par Ricci en 1888 [Ric88]) : afin de pouvoir énoncer sa nouvelle théorie de la gravitation,

1. GDAC-I2M UMR 7373 CNRS Université d'Aix-Marseille

2. eric.olivier@univ-amu.fr

Einstein va devoir assimiler les derniers progrès de la géométrie de son temps – essentiellement le calcul tensoriel – en suivant les traces de Gauss, Riemann, Christoffel, Ricci et Levi-Civita.

Cette première note est consacrée aux concepts mathématiques nécessaires à la formulation de la Relativité Générale : en particulier, il nous faudra introduire le formalisme du calcul différentiel et tensoriel sur les variétés, puis introduire les structures pseudo-riemanniennes, afin de pouvoir définir rigoureusement la notion de géodésiques inertielles. Pour Einstein, ces géodésiques doivent décrire les mouvements des corpuscules uniquement soumis à la gravitation : pour comprendre cela, il nous faudra partir du formalisme lagrangien afin de voir comment les équations des géodésiques inertielles se traduisent en termes de transport parallèle (Théorème de Levi-Civita). Nous aborderons ces questions par les bases de calcul différentiel absolue de Ricci et Levi-Civita en regardant particulièrement l'articulation qui existe entre la notion de connexion et le transport parallèle introduit par Levi-Civita.

Remerciement : Je tiens à remercier chaleureusement Christian Faivre pour m'avoir mis au défi d'un exposé "le plus simple possible" permettant de comprendre pourquoi la gravitation courbe les rayons lumineux. Entre autres, ce projet a été l'occasion de moult discussions passionnées qui ont animé la salle du Dugommier. D'une certaine manière la Relativité Générale donne un élément de réponse important à cette question, puisqu'elle permet de prévoir avec une grande précision, les mesures des déflexions lumineuses dues à la gravitation. Aller plus loin, en essayant de voir comment s'articulent l'électrodynamique (quantique) et la Relativité Générale, permettrait une bonne (meilleure) description des rayons lumineux : mais cette tâche se trouve être très difficile, car un photon n'est pas un objet de la Relativité Générale...

2. Changements de bases dans un espace vectoriel

Soit $\mathcal{E} = (\varepsilon_1, \dots, \varepsilon_n)$ une base d'un espace vectoriel E . Tout vecteur $v \in E$ se décompose de manière unique sous forme d'une combinaison linéaire, soit $v = v^i \varepsilon_i$ (notations sommatoires d'Einstein³) : ici, les coordonnées v^i (indice en haut) sont appelées \mathcal{E} -composantes contravariantes. Pour l'instant, la « *contravariance*⁴ » proprement dite traduit simplement le fait que les indices des composantes sont en position supérieure (nous verrons plus loin que la « *covariance* » concerne les indices en position inférieure). Ces notations ont aussi l'avantage de permettre l'utilisation des notations sommatoires d'Einstein ; dans la suite, nous donnerons un sens précis au formalisme de la covariance/contravariance, avec l'introduction systématique des tenseurs. Le changement de coordonnées de \mathcal{E} vers une autre base $\hat{\mathcal{E}} = (\hat{\varepsilon}_1, \dots, \hat{\varepsilon}_n)$, est effectué grâce à la matrice de passage⁵

3. La notation d'Einstein signifie une sommation sur tous les indices i, j, k, \dots se trouvant à la fois en position supérieure et inférieure dans une expression multiplicative : ainsi, par exemple, la décomposition d'un vecteur v de coordonnée (v^1, \dots, v^n) dans une base $\mathcal{E} = (\varepsilon_1, \dots, \varepsilon_n)$ s'écrit-elle $v = v^i \varepsilon_i = \sum_{i=1}^n v^i \varepsilon_i$.

4. La terminologie *covariance/contravariance* apparaît dans l'article de Ricci [Ric88] en 1888 (en italien). Dans leur article [RTLC00] de synthèse en français Ricci et Levi-Civita parlent de *contrevariance*.

5. Une matrice est ici considérée comme un tableau de nombres réels (les coefficients), dont la position est repérée par les indices de ligne/colonne. L'ensemble $\mathcal{M}_n(\mathbb{K})$ des matrices $n \times n$ à coefficients dans $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} est munie d'une structure naturelle de \mathbb{K} -espace vectoriel ; on peut aussi voir $\mathcal{M}_n(\mathbb{K})$ comme une algèbre non commutative en définissant le produit des matrices : si $A = (a^i_j)$ et $B = (b^i_j)$, alors le produit

$P = (p^i_j)$ – i étant l'indice de ligne et j l'indice de colonne – de sorte que

$$\hat{\varepsilon}_j = p^i_j \varepsilon_i$$

(retenir qu'on exprime les vecteurs de la *nouvelle* base en fonction des vecteurs de la base *initiale*, ce qui semble raisonnable). Par suite, pour un vecteur $v = v^i \varepsilon_i = \hat{v}^j \hat{\varepsilon}_j$ arbitraire, les relations entre les \mathcal{E} -composantes v^i et les $\hat{\mathcal{E}}$ -composantes \hat{v}^i s'obtiennent en écrivant $\hat{v}^j \hat{\varepsilon}_j = \hat{v}^j p^i_j \varepsilon_i = v^i \varepsilon_i$: cela se traduit par le système d'équations :

$$v^i = \hat{v}^j p^i_j$$

(où les *anciennes* composantes sont fonctions des *nouvelles*). Notons $P^{-1} = (q^i_j)$ l'inverse de la matrice P (i.e. $p^i_k q^k_j = \delta^i_j$ – symbole de Kronecker). Soit $A = (a^i_j)$ la matrice d'un endomorphisme (linéaire) ψ de E dans la base \mathcal{E} (i.e. $\psi(\varepsilon_j) = a^i_j \varepsilon_i$) ; alors,

$$\psi(v) = a^i_j v^j \varepsilon_i = a^i_j (\hat{v}^r p^j_r) (q^s_i \hat{\varepsilon}_s) = (q^s_i a^i_j p^j_r) \hat{v}^r \hat{\varepsilon}_s$$

de sorte que la matrice de ψ dans la base $\hat{\mathcal{E}}$ s'écrit

$$(1) \quad \hat{A} = P^{-1} A P$$

Soit $\xi : E \times E \rightarrow \mathbb{R}$ une forme bilinéaire ; par définition, les $b_{ij} := \xi(\varepsilon_i, \varepsilon_j)$ sont les \mathcal{E} -composantes de ξ , de sorte que $\xi(v, w) = v^i w^j b_{ij}$ dès que $v = v^i \varepsilon_i$ et $w = w^j \varepsilon_j$. Pour obtenir les $\hat{\mathcal{E}}$ -composantes $\hat{b}_{ij} = \xi(\hat{\varepsilon}_i, \hat{\varepsilon}_j)$ de ξ , notons $v = \hat{v}^i \hat{\varepsilon}_i$ et $w = \hat{w}^j \hat{\varepsilon}_j$; alors il vient :

$$(2) \quad \xi(v, w) = v^i w^j b_{ij} = (\hat{v}^r p^i_r) (\hat{w}^s p^j_s) b_{ij} = \hat{v}^r \hat{w}^s (p^i_r b_{ij} p^j_s)$$

et finalement

$$(3) \quad \hat{b}_{ij} = p^i_r b_{ij} p^j_s$$

Par convention appelons $B = (b_{ij})$ (resp. $\hat{B} = (\hat{b}_{ij})$) la matrice de ξ dans la base \mathcal{E} (resp. $\hat{\mathcal{E}}$) – i étant l'indice de ligne et j l'indice de colonne. Avec P^* désignant la transposée de P , nous pouvons traduire (3) en disant que la matrice de ξ dans la base $\hat{\mathcal{E}}$ est

$$(4) \quad \hat{B} = P^* B P$$

3. Dualité et dualité euclidienne

Le dual E^* d'un \mathbb{R} -espace vectoriel E de dimension finie n , est le \mathbb{R} -espace vectoriel des formes linéaires sur E (encore appelées covecteurs), i.e. des applications linéaires $\alpha : E \rightarrow \mathbb{R}$. L'espace E^* est de dimension n et toute base $\mathcal{E} = (\varepsilon_1, \dots, \varepsilon_n)$ de E est associée à une base duale $\mathcal{E}' = (\varepsilon^1, \dots, \varepsilon^n)$ qui est la base de E^* formées des covecteurs ε^i définis par le système équations (c.f. [BG68, § 2.7 p. 75]) :

$$(5) \quad (\forall j) \quad \varepsilon^i(\varepsilon_j) = \delta^i_j$$

Maintenant, munissons E d'une structure euclidienne : cela signifie que nous fixons une forme bilinéaire $(u, v) \mapsto g(u, v)$ qui est à la fois symétrique (i.e. $g(u, v) = g(v, u)$), positive (i.e. $g(u, u) \geq 0$) et définie (i.e. $g(u, u) = 0 \iff u = 0$). Par définition, g est le produit scalaire de l'espace euclidien E . Nous utiliserons aussi la notation spéciale $\langle u|v \rangle$ pour désigner la valeur de $g(u, v)$; $u \mapsto \|u\| := \langle u|u \rangle^{1/2}$ est une norme sur E qui fait de $(E, \|\cdot\|)$ un espace de Banach. Soit $\mathcal{E} = (\varepsilon_1, \dots, \varepsilon_n)$ une base (quelconque) de E . La symétrie de g

AB est la matrice $(a^i_k b^k_j)$. Par définition, la transposée de $A = (a^i_j)$ est la matrice $A^* = (a^i_j)$ t.q. $\alpha^i_j = a^j_i$: lorsque $\mathbb{K} = \mathbb{C}$, la transposition $A \mapsto A^*$ fait de $\mathcal{M}_n(\mathbb{R})$ une C^* -algèbre.

(en tant que produit scalaire) se traduit par la symétrie de ses \mathcal{E} -composantes $g_{ij} = g(\varepsilon_i, \varepsilon_j)$ en ce sens que $g_{ij} = g_{ji}$. D'autre part, les g_{ij} permettent aussi de définir les \mathcal{E} -composantes covariantes $v_i \in \mathbb{R}$ (indice en bas) d'un vecteur $v = v^i \varepsilon_i$ en posant

$$(6) \quad v_i := g(\varepsilon_i, v) = g(\varepsilon_i, v^j \varepsilon_j) = g_{ij} v^j$$

(opération généralisée avec l'abaissement/relèvement des indices tensoriels). Enfin, notons que la structure euclidienne sur E détermine une classe spéciale de bases de E : ainsi, $\mathcal{E} = (\varepsilon_1, \dots, \varepsilon_n)$ sera dite **orthonormée** si et seulement si $g_{ij} = g(\varepsilon_i, \varepsilon_j) = \delta_{ij}$ (symbole de Kronecker, avec deux indices covariants).

Munir E d'un produit scalaire g permet de voir la dualité entre E et E^* comme une isométrie involutive : nous parlerons dans ce cas de **g -dualité**. Pour voir cela, notons que tout vecteur v est associé au covecteur $v^* = g(v, \cdot)$ et réciproquement que tout covecteur α est associé au vecteur α^* t.q. $\alpha = g(\alpha^*, \cdot)$. Étant déterminée par le produit scalaire g , l'opération $v \mapsto v^*$ (resp. $\alpha \mapsto \alpha^*$) de E sur E^* (resp. de E^* sur E) est appelée **g -transposition** : c'est une involution en ce sens que $(v^*)^* = v$ et $(\alpha^*)^* = \alpha$, pour tout $v \in E$ et tout $\alpha \in E^*$.

Remarque 3.1. (1) : Dans les notations de Dirac, un vecteur v se note comme un « ket » soit $v \equiv |v\rangle$ et le covecteur associé comme un « bra » soit $v^* \equiv \langle v|$. Si w est un deuxième vecteur, on retrouve l'expression du produit scalaire $v^*(w) = g(v, w) = \langle v|w\rangle$.

(2) : Attention : étant donné $\mathcal{E} = (\varepsilon_1, \dots, \varepsilon_n)$ une base (quelconque) de E , le vecteur ε^i de la base duale \mathcal{E}' n'est pas en général égal à ε_i^* . En fait l'égalité $\varepsilon^i = \varepsilon_i^*$ n'a lieu – pour tout i – que dans le cas où \mathcal{E} est une base orthogonale.

L'espace E^* est naturellement muni d'un produit scalaire $(\alpha, \beta) \mapsto g^*(\alpha, \beta)$, où par définition, $g^*(\alpha, \beta) = g(\alpha^*, \beta^*)$: l'application $v \mapsto v^*$ est alors une isométrie de (E, g) sur (E^*, g^*) dont l'isométrie réciproque est l'application $\alpha \mapsto \alpha^*$. Par définition les \mathcal{E} -composantes de g^* sont les

$$(7) \quad g^{ij} := g^*(\varepsilon^i, \varepsilon^j)$$

Rappelons que la base $\mathcal{E}' = (\varepsilon^1, \dots, \varepsilon^n)$ de E^* se déduit de la base $\mathcal{E} = (\varepsilon_1, \dots, \varepsilon_n)$ de E par les relations de dualité en (5) qui sont indépendantes du produit scalaire sur E . Les \mathcal{E}' -composantes covariantes⁶ α_i d'un covecteur α sont les coordonnées de α dans la base duale \mathcal{E}' , ce qui donne $\alpha = \alpha_i \varepsilon^i$. De même que pour les composantes contravariantes des vecteurs en (6), les indices covariants des composantes d'un covecteur α peuvent être relevés en posant :

$$(8) \quad \alpha^i := g^*(\varepsilon^i, \alpha) = g^*(\varepsilon^i, \varepsilon^j) \alpha_j = g^{ij} \alpha_j$$

Les définitions que nous venons de poser peuvent paraître abstraites (et elles le sont !) : elles prennent cependant un sens avec la proposition suivante (voir aussi la Figure 1).

Proposition 3.2. Notons g_{ij} (resp. g^{ij}) les composantes covariantes (resp. contravariantes) du produit scalaire g d'un espace euclidien E rapporté à une base $\mathcal{E} = (\varepsilon_1, \dots, \varepsilon_n)$ et soit $\mathcal{E}' = (\varepsilon^1, \dots, \varepsilon^n)$ la base duale associée à \mathcal{E} . Si $v_i = g(\varepsilon_i, v)$ (resp. $\alpha^i = g^*(\varepsilon^i, \alpha)$) sont les composantes

6. Covariantes, car les indices sont en position inférieure : c.f. infra pour la signification tensorielle.

covariantes (resp. contravariantes) de $v = v^i \varepsilon_i$ (resp. de $\alpha = \alpha_i \varepsilon^i$), alors :

- (i) : $v^* = v_i \varepsilon^i$ et $\alpha^* = \alpha^i \varepsilon_i$
- (ii) : $v = v_i \varepsilon^{i*}$ et $\alpha = \alpha^i \varepsilon_i^*$
- (iii) : $\varepsilon_j^* = g_{ij} \varepsilon^i$ et $\varepsilon^{j*} = g^{ij} \varepsilon_i$
- (iv) : $g_{ij} g^{jk} = \delta_i^k$ et $g(\varepsilon^{i*}, \varepsilon_j) = \delta^i_j$

Preuve. (i) – (ii) : D’après (6) nous avons $v^*(\varepsilon_j) = g(v, \varepsilon_j) = v^i g(\varepsilon_i, \varepsilon_j) = v^i g_{ij} = v_j$. Le fait que $v^* = v_i \varepsilon^i$ se déduit de la définition des ε^i en (5) : en effet, pour tout j :

$$(v_i \varepsilon^i)(\varepsilon_j) = v_i \varepsilon^i(\varepsilon_j) = v_i \delta^i_j = v_j = v^*(\varepsilon_j)$$

La preuve de $\alpha^* = \alpha^i \varepsilon_i$ est analogue et (ii) découle de (i) par g -transposition.

(iii) : Comme $\varepsilon_j = \delta^i_j \varepsilon_i$, en appliquant (i) il vient : $\varepsilon_j^* = (g_{ik} \delta^k_j) \varepsilon^i = g_{ij} \varepsilon^i$. La preuve de $\varepsilon^{j*} = g^{ij} \varepsilon_i$ est analogue. (En général $\varepsilon^i \neq \varepsilon_i^*$, cette égalité n’ayant lieu (pour tout i) que lorsque la base $\mathcal{E} = (\varepsilon_1, \dots, \varepsilon_n)$ est orthonormée, de sorte que dans ce cas $g(\varepsilon_i, \varepsilon_j) = \delta_{ij}$.)

(iv) : La g -dualité entre E et E^* est involutive en ce sens que pour tout $v \in E$ nous avons $(v^*)^* = v$: par suite, si les v^i sont les composantes contravariantes de v alors

$$v^i \varepsilon_i = v = (v^*)^* = (v^k g_{kj} \varepsilon^j)^* = (v^k g_{kj}) g^{ji} \varepsilon_i = v^k (g_{kj} g^{ji}) \varepsilon_i$$

Par suite, l’identité $v^i = v^k (g_{kj} g^{ji})$ étant satisfaite pour tout i et tout $v \in E$, nous déduisons que $g_{kj} g^{ji} = \delta^i_k$. Pour la deuxième identité de (iv), nous avons

$$g(\varepsilon^{i*}, \varepsilon_j) = g(g^{ik} \varepsilon_k, \varepsilon_j) = g^{ik} g_{kj} = \delta^i_j$$

□

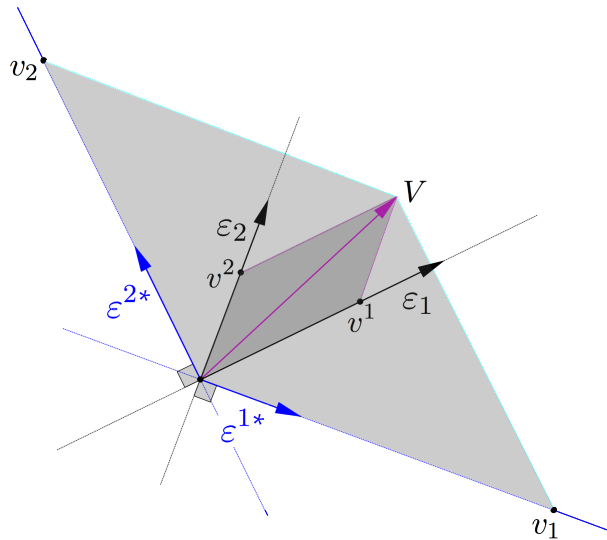


FIGURE 1. Système de coordonnées covariant/contravariant relatif à une base (quelconque) $(\varepsilon_1, \varepsilon_2)$ de \mathbb{R}^2 muni du produit scalaire canonique, $g(\cdot, \cdot) = \langle \cdot | \cdot \rangle$. Tout vecteur v se représente de deux manières soient $v = v^i \varepsilon_i$ et $v = v_i \varepsilon^{i*}$. La deuxième identité de la partie (iv) de la Proposition 3.2 affirme que $\langle \varepsilon^{i*} | \varepsilon_j \rangle = \delta^i_j$: cela signifie que ε^{i*} est un vecteur (de E) qui est orthogonal à chacun des ε_j dès que $j \neq i$, avec $\langle \varepsilon^{i*} | \varepsilon_i \rangle = 1$.

4. Calcul tensoriel

Soient E et F deux \mathbb{R} -espaces vectoriels de dimensions respectives p et q (finies). Le produit tensoriel $u \otimes v$ de $u \in E$ et $v \in F$ est la forme linéaire sur $E^* \times F^*$ t.q.

$$u \otimes v(\xi, \zeta) = u(\xi)v(\zeta)$$

Par définition le produit tensoriel de E et F est l'espace vectoriel noté $E \otimes F$ qui est engendré par les $u \otimes v$, pour (u, v) décrivant $E \times F$. Si $\mathcal{E} = (\varepsilon_1, \dots, \varepsilon_p)$ et $\mathcal{F} = (\eta_1, \dots, \eta_r)$ sont respectivement des bases de E et F , alors pour tout $u = u^i \varepsilon_i \in E$ et tout $v = v^j \eta_j \in F$,

$$u \otimes v = u^i v^j \varepsilon_i \otimes \eta_j$$

On peut alors montrer que les $\varepsilon_i \otimes \eta_j$ pour $1 \leq i \leq p$ et $1 \leq j \leq r$ forment une base de $E \otimes F$ notée $\mathcal{E} \otimes \mathcal{F}$, de sorte que $E \otimes F$ est de dimension pq . Si $E = F$ alors $u \otimes v$ et $v \otimes u$ sont deux éléments de $E \otimes E =: \otimes^2 E$: il est important de remarquer (et immédiat à vérifier) que – mis à part le cas où u et v sont proportionnels – nous avons

$$u \otimes v \neq v \otimes u$$

Si H est un troisième espace vectoriel de dimension r alors pour tout $(u, v, w) \in E \times F \times H$, nous posons $u \otimes v \otimes w := (u \otimes v) \otimes w$, ce qui permet de définir $E \otimes F \otimes H$. Il est immédiat de vérifier que le produit tensoriel est associatif, i.e. que $(u \otimes v) \otimes w = u \otimes (v \otimes w)$; ainsi, le produit tensoriel d'un nombre quelconque (fini) d'espaces vectoriels (de dimensions finies) se définit par une induction associative non commutative.

Un cas important est celui du produit tensoriel de copies de E et de son dual E^* : un élément de $(\otimes^p E) \otimes (\otimes^q E^*) =: \otimes^{p,q} E$ est appelé un tenseur de valence (on dit aussi de type) (p, q) sur E . Nous avons vu que toute base $\mathcal{E} = (\varepsilon_1, \dots, \varepsilon_n)$ de E est associée à la base duale $\mathcal{E}' = (\varepsilon^1, \dots, \varepsilon^n)$ de E^* par les relations en (5) : dans ce cas, les composantes d'un tenseur T de valence (p, q) sur E sont les réels $T^{i_1 \dots i_p}_{j_1 \dots j_q}$ t.q :

$$T = T^{i_1 \dots i_p}_{j_1 \dots j_q} \varepsilon_{i_1} \otimes \dots \otimes \varepsilon_{i_p} \otimes \varepsilon^{j_1} \otimes \dots \otimes \varepsilon^{j_p}$$

Remarque 4.1. Soit T un tenseur de valence $(1, 1)$ sur E et T^i_j ses composantes relativement à la base $\mathcal{E} = (\varepsilon_1, \dots, \varepsilon_n)$; alors les T^i_j sont aussi les coefficients de la matrice (dans la base \mathcal{E}) de l'endomorphisme ψ de E qui à tout vecteur $v = v^i \varepsilon_i$ associe $T^i_j v^j \varepsilon_i$, soit encore que

$$\psi(\varepsilon_j) = T^i_j \varepsilon_i$$

La trace d'une matrice $A = (a^i_j)$ est le réel $\text{tr}(A) = a^i_i$. Cette définition naïve, nous permet d'introduire la contraction d'indices pour les tenseurs : plus précisément, la contraction $[^1_1] : \otimes^{p,q} E \rightarrow \otimes^{p-1,q-1} E$ est définie par deux conditions soient, pour tout $v_1, \dots, v_p \in E$ (resp. $\alpha_1, \dots, \alpha_q \in E^*$) :

$$\text{(CO1)} : [^1_1](v_1 \otimes v_2 \otimes \dots \otimes v_p \otimes \alpha_1 \otimes \alpha_2 \otimes \dots \otimes \alpha_q) = \alpha_1(v_1)(v_2 \otimes \dots \otimes v_p \otimes \alpha_2 \otimes \dots \otimes \alpha_q)$$

et pour tout $T, S \in \otimes^{p,q} E$

$$\text{(CO2)} : [^1_1](T + S) = [^1_1](T) + [^1_1](S)$$

Si E^* est muni de la base duale de la base choisie pour E et si $T^{i_1 i_2 \dots i_p}_{j_1 j_2 \dots j_q}$ sont les composantes d'un tenseur $T \in \otimes^{p,q} E$, alors la contraction $[^1_1]$ de T s'écrit en composantes :

$$[^1_1](T)^{i_2 \dots i_p}_{j_2 \dots j_q} = T^{i_1 i_2 \dots i_p}_{i j_2 \dots j_q}$$

De même, pour tout $1 \leq r \leq p$ et tout $1 \leq s \leq q$ il est possible de définir la contraction $[^r_s] : \otimes^{p,q} E \rightarrow \otimes^{p-1,q-1} E$: il y a donc pq contractions possibles définies sur $\otimes^{p,q} E$. Enfin, rappelons que pour E rapporté à une base $\mathcal{E} = (\varepsilon_1, \dots, \varepsilon_n)$, un tenseur $T \in \otimes^{1,1} E$ de composante T^i_j est associé à un endomorphisme ψ de E , où T^i_j soient les coefficients de la matrice de ψ dans la base \mathcal{E} (c.f. Remarque 4.1) : dans ce cas, l'unique contraction de T s'identifie à un tenseur de $\otimes^{0,0} E$, c'est-à-dire à un scalaire : ce scalaire correspond à la trace de ψ .

5. Variétés différentielles

5.1. Cartes et atlas. Un espace topologique M séparé et à base dénombrable⁷ est une variété topologique de dimension n (≥ 1) s'il existe un atlas \mathfrak{A} , i.e. une famille de couples $(\mathcal{U}, \mathbf{x})$ où chaque $\mathbf{x} : \mathcal{U} \rightarrow \mathbb{R}^n$ est un homéomorphisme d'un ouvert \mathcal{U} de M sur un ouvert de \mathbb{R}^n , t.q. l'ensemble des ouverts \mathcal{U} forme un recouvrement ouvert (localement fini) de M et vérifiant la condition de compatibilité suivante : pour tout $(\mathcal{U}, \mathbf{x})$ et $(\mathcal{V}, \mathbf{y})$ dans \mathfrak{A} t.q. $\mathcal{U} \cap \mathcal{V}$ ne soit pas vide, l'application de *changement de coordonnées* $\mathbf{x} \circ \mathbf{y}^{-1}$ de $\mathbf{y}(\mathcal{U} \cap \mathcal{V})$ sur $\mathbf{x}(\mathcal{U} \cap \mathcal{V})$ est un homéomorphisme. Dans la suite nous supposerons que M est une *variété différentielle lisse et orientable*, ce qui signifie que tous les homéomorphismes $\mathbf{x} \circ \mathbf{y}^{-1}$ définis ci-dessus sont des difféomorphismes C^∞ préservant l'orientation de \mathbb{R}^n . Nous dirons qu'un couple $(\mathcal{U}, \mathbf{x})$ formé d'un ouvert de M et d'une application $\mathbf{x} : \mathcal{U} \rightarrow \mathbb{R}^n$ est un *système de coordonnées sur M* lorsque l'ensemble $\mathfrak{A} \cup \{(\mathcal{U}, \mathbf{x})\}$ est lui-même un atlas de M ; par définition⁸ la *carte associée à \mathbf{x}* est l'application X définie sur $\mathbf{x}(\mathcal{U})$, à valeur dans M et telle que $X \circ \mathbf{x}$ est l'identité de \mathcal{U} . Enfin, une *partie S de M est une sous-variété différentielle de codimension d* , si S est une variété différentielle de dimension $n - d$ et si tout point X de S appartient au domaine d'un système de coordonnées $(\mathcal{U}, \mathbf{x})$ de M , de sorte que $\mathbf{x}(\mathcal{U} \cap S) = \mathbf{x}(\mathcal{U}) \cap (\mathbb{R}^{n-d} \times \{0\}^d)$. Le *théorème du plongement de Whitney (1936)*, nous permet de supposer, sans perte de généralités, que M est une *sous-variété de (ou encore plongée dans) l'espace euclidien \mathbb{R}^m* pour un certain $m \geq n$.

5.2. Vecteurs tangents et dérivations directionnelles. Soit M une variété (lisse et orientable) de dimension n plongée dans l'espace euclidien⁹ \mathbb{R}^m . Nous notons $C^\infty(M)$ l'espace vectoriel des applications $f : M \rightarrow \mathbb{R}$ de classe C^∞ sur M (i.e. pour tout système de coordonnées $(\mathcal{U}, \mathbf{x})$ de carte associée X , l'application $f \circ X : \mathbf{x}(\mathcal{U}) \rightarrow \mathbb{R}$ est de classe C^∞ sur $\mathbf{x}(\mathcal{U})$ au sens usuel). L'application (lisse) $\gamma : [a; b] \rightarrow \mathbb{R}^m$ est un *chemin de M* si son support est inclus dans M et si de plus elle est injective sur¹⁰ $]a; b[$; c'est un *germe de chemin en X* si de plus $a < 0 < b$ avec $\gamma(0) = X$. Par définition l'espace tangent en X est l'ensemble $T_X M$ formé des dérivées en $t = 0$ (au sens du calcul différentiel dans \mathbb{R}^m) des germes de chemin de M en X . Par la structure de variété différentielle de M , chaque $T_X M$ est un sous-espace vectoriel de \mathbb{R}^m de dimension n ; si maintenant, $t \mapsto \gamma_1(t)$ et $t \mapsto \gamma_2(t)$ sont deux germes de chemin en X , dont les dérivées en $t = 0$ sont toutes les

7. Un espace topologique est dit séparé lorsqu'il possède un sous ensemble dénombrable et dense ; il est à base dénombrable si chacun de ses points possède une base de voisinage dénombrable.

8. Convention non universelle utilisée par Adrien Douady.

9. Le produit scalaire canonique de \mathbb{R}^m sera toujours noté $\langle \cdot | \cdot \rangle$.

10. Le support $\gamma([a; b])$ de γ est une sous variété de M qui est lisse dès que $\gamma(a) \neq \gamma(b)$ (dans le cas où $\gamma(a) = \gamma(b)$ il faut regarder de plus près).

deux égales au vecteur v alors, pour tout $f \in \mathcal{C}^\infty(M)$, chaque application $f \circ \gamma_i$ est \mathcal{C}^∞ sur un voisinage de 0 et les dérivées de $f \circ \gamma_1(t)$ et $f \circ \gamma_1(t)$ prennent une valeur commune en $t = 0$ notée $v(f)$: cette valeur s'interprète comme la **dérivée directionnelle de f suivant le vecteur tangent v** . Il est facile de vérifier que l'application $v(\cdot) : \mathcal{C}^\infty(M) \rightarrow \mathbb{R}$ est une **dérivation en X** , en ce sens qu'elle vérifie la **règle de dérivation de Leibniz en X** soit, pour tout $f, g \in \mathcal{C}^\infty(M)$:

$$(9) \quad v(fg) = v(f)g(X) + f(X)v(g)$$

Le **fibré tangent TM** est l'union disjointe $\coprod_X T_X M$ des espaces tangents : c'est un sous-ensemble de la grassmannienne¹¹ $\text{Gr}_n(\mathbb{R}^m)$. Soit V une section du fibré tangent, i.e. une application $V : X \mapsto V|_X$ de M dans TM t.q. $V|_X \in T_X M$; pour $f : M \rightarrow \mathbb{R}$, la valeur de l'application $V(f) : M \rightarrow \mathbb{R}$ prise en X est définie comme la dérivation directionnelle $V|_X(f)$ de f suivant $V|_X$: alors la section V est dite \mathcal{C}^∞ si $V(f) \in \mathcal{C}^\infty(M)$ dès que $f \in \mathcal{C}^\infty(M)$: c'est par définition un **champ de vecteurs sur M** . Nous notons $\mathfrak{T}_0^1(TM)$ l'**ensemble des champs de vecteurs sur M** (la notation sera généralisée avec l'introduction des champs de tenseurs : c.f. § 4) : il est facile de vérifier que $\mathfrak{T}_0^1(TM)$ forme un espace vectoriel. De plus, pour tout $V \in \mathfrak{T}_0^1(TM)$, l'application $V(\cdot) : f \mapsto V(f)$ est un endomorphisme de $\mathcal{C}^\infty(M)$: on déduit alors de (9) que l'application $V(\cdot) : \mathcal{C}^\infty(M) \rightarrow \mathbb{R}$ est une **dérivation**, en ce sens qu'elle vérifie la **formule de dérivation de Leibniz** soit, pour tout $f, g \in \mathcal{C}^\infty(M)$:

$$(10) \quad V(fg) = V(f)g + fV(g)$$

Les dérivations sur M – i.e. les endomorphismes de $\mathcal{C}^\infty(M)$ satisfaisant l'identité de Leibniz – forment un espace vectoriel. L'application $V \mapsto V(\cdot)$, qui a un champ de vecteurs associe la dérivation correspondante, est une application linéaire injective (il est facile de vérifier que $V(f) = 0$ pour tout f entraîne que V est le champ nul). Dans la suite nous identifierons toujours un champ de vecteurs V avec la dérivation $V(\cdot)$: ainsi, pour tout $X \in M$ et $f \in \mathcal{C}^\infty(M)$, la notation $V|_X$ (resp. $V|_X(f)$) désigne un élément de $T_X M$ et donc un vecteur de \mathbb{R}^m (resp. un réel qui représente la dérivée directionnelle de f suivant $V|_X$) alors que $V(f) : X \mapsto V|_X(f)$ est un élément de $\mathcal{C}^\infty(M)$. Pour légitimer complètement cette identification il faut montrer que la correspondance $V \mapsto V(\cdot)$ entre champs de vecteurs et dérivations, est un isomorphisme d'espace vectoriel (le point difficile concernant la surjectivité de l'application $V \mapsto V(\cdot)$ est levée par le Lemme d'Hadamard¹²).

5.3. Base holonôme d'un système de coordonnées. Soit (e_1, \dots, e_m) la base canonique de \mathbb{R}^m . Étant donné $(\mathcal{U}, \mathbf{x})$ un système de coordonnées de M et $X \in \mathcal{U}$ nous notons $\partial_i|_X$ le vecteur de $T_X M$ associé au germe de chemin $\gamma_i : t \mapsto \mathbf{X}(\mathbf{x}(X) + te_i)$, où $\mathbf{X} : \mathbf{x}(\mathcal{U}) \rightarrow \mathbb{R}^m$ est la carte associée à \mathbf{x} . L'application $\partial_i : X \mapsto \partial_i|_X$ est un champ de vecteurs défini

11. La grassmannienne $\text{Gr}_n(\mathbb{R}^m)$ est ensemble des sous-espaces vectoriels de dimension n de \mathbb{R}^m .

12. Soit \mathcal{U} un ouvert de \mathbb{R}^n étoilé en $X = (x_1, \dots, x_n)$. Alors, pour $f \in \mathcal{C}^\infty(\mathcal{U})$ et $Y = (y_1, \dots, y_n) \in \mathcal{U}$,

$$f(Y) - f(X) = \int_0^1 \frac{d}{dt} f(X + t(Y - X)) dt = \sum_{i=1}^n (y_i - x_i) \int_0^1 \frac{\partial f}{\partial x_i}(X + t(Y - X)) dt$$

Si $f_i(Y) := \int_0^1 \frac{\partial f}{\partial x_i}(X + t(Y - X)) dt$, alors (**théorème de dérivation sous le signe intégral**) les applications $Y \mapsto f_i(Y)$ sont localement \mathcal{C}^∞ en $X = 0$ et nous pouvons écrire : $f(Y) = f(X) + \sum_{i=1}^n (y_i - x_i) f_i(Y)$

sur¹³ \mathcal{U} : c'est le i -ème champ holonôme associée à $(\mathcal{U}, \mathbf{x})$. Le fait que $(\partial_1|_X, \dots, \partial_n|_X)$ soit une base de $T_X M$ pour tout $X \in \mathcal{U}$, entraîne que $(\partial_1, \dots, \partial_n)$ est une base du $\mathcal{C}^\infty(\mathcal{U})$ -module des champs de vecteurs sur \mathcal{U} . En d'autres termes, tout champ $V \in \mathfrak{T}_0^1(TM)$ est associé à n -applications $V^i : \mathcal{U} \rightarrow \mathbb{R}$ de $\mathcal{C}^\infty(\mathcal{U})$ et t.q.¹⁴ $V = V^i \partial_i$ (identité valable sur \mathcal{U}). On dira que $(\partial_1, \dots, \partial_n)$ est la base holonôme des \mathcal{U} -champs de vecteurs pour le système de coordonnées $(\mathcal{U}, \mathbf{x})$. Supposons que (x^1, \dots, x^n) soient les composantes cartésiennes de \mathbf{x} , en ce sens que $\mathbf{x}(X) = x^i(X) \mathbf{e}_i$, pour tout $X \in \mathcal{U}$. Alors $(x^1(X), \dots, x^n(X))$ sont appelées les coordonnées locales de X pour le système de coordonnées $(\mathcal{U}, \mathbf{x})$: en vue des formules de changement de bases holonômes il sera avantageux de noter

$$\partial_i = \frac{\partial}{\partial x^i}$$

de sorte que $\partial_i(f) = \partial f / \partial x^i$, pour tout $f \in \mathcal{C}^\infty(M)$. De plus, si $\hat{f} := f \circ \mathbf{X}$, alors il y a une identification entre ∂_i et la dérivation partielle (au sens usuel du calcul différentiel sur $\mathbf{x}(\mathcal{U})$) et qui se traduit par l'identité $\partial_i|_X(f) = \partial \hat{f} / \partial x^i$, dès que $\mathbf{x}(X) = (x^1, \dots, x^n)$.

5.4. Courbes intégrales d'un champ de vecteurs. Soit V un champ de vecteurs sur une variété M de dimension n , supposée plongée dans \mathbb{R}^m . Alors pour tout $X \in M$, il existe une unique courbe intégrale maximale $t \mapsto \phi_V^t(X) \in M$ i.e. un chemin sur M définie sur un intervalle ouvert I_X contenant 0, maximal pour l'inclusion et t.q

$$(11) \quad \phi_V^0(X) = X \quad \text{et} \quad \frac{d}{dt} \phi_V^t(X) = V|_{\phi_V^t(X)}$$

la dernière identité ayant lieu pour tout $t \in I_X$. (En particulier le flot (maximal) $t \mapsto \phi_V^t(X)$ d'un champ de vecteur V en X est un germe de chemin en X .) Considérons maintenant que $Y := \phi_V^s(X)$ pour $s \in I_X$ donné. Alors (exercice) $t \mapsto \phi_V^{t+s}(X)$ est la solution maximale du problème de Cauchy en Y avec $I_Y = I_X - s$. Par suite (unicité du problème de Cauchy) il vient $\phi_V^{t+s}(X) = \phi_V^t(Y)$ pour tout $t \in I_Y$, soit encore :

$$(12) \quad \phi_V^{t+s}(X) = \phi_V^t(\phi_V^s(X))$$

L'identité (12) est la propriété (locale) de cocycle des courbes intégrales de V . Le champ de vecteurs V est dit complet si et seulement si $I_X = \mathbb{R}$, pour tout $X \in M$. Du fait que $I_{\phi_V^s(X)} = (I_X - s)$ dès que $s \in I_X$, la complétude de V est équivalente à l'existence d'un $\varepsilon > 0$ t.q. $I_X \supset]-\varepsilon; \varepsilon[$, pour tout $X \in M$. Notons que si V est complet, (12) assure que l'ensemble $\{\phi_V^t(\cdot) ; t \in \mathbb{R}\}$ réalise un groupe de difféomorphismes (de M) à un paramètre.

Le théorème d'existence et d'unicité d'une solution maximale du problème de Cauchy (sur un ouvert d'un espace vectoriel euclidien) s'adapte au cas des courbes intégrales d'un champ de vecteurs défini sur une variété M plongée dans \mathbb{R}^m . Pour voir cela on peut résoudre le problème en coordonnées locales, ou encore supposer que le champ de vecteur est (au moins localement) la restriction à M d'un champ de vecteurs défini sur un voisinage ouvert de M .

13. Ce n'est pas un champ de vecteurs sur M , puisque la définition de ∂_i dépend du système de coordonnées $(\mathcal{U}, \mathbf{x})$ et que donc $\partial_i|_X$ n'est définie que pour $X \in \mathcal{U}$.

14. Nous utilisons ici les notations d'Einstein, de sorte que $V^i \partial_i = \sum_i V^i \partial_i$.

6. Crochet de Lie de deux champs de vecteurs

6.1. Crochet de lie comme dérivation. Soit \mathbb{K} un corps commutatif, E un \mathbb{K} -espace vectoriel et $[\cdot, \cdot] : E \times E \rightarrow E$ une application bilinéaire : l'algèbre $(E, [\cdot, \cdot])$ est appelée une algèbre de Lie lorsque l'application bilinéaire $[\cdot, \cdot]$ est un **crochet de Lie**, c'est-à-dire qu'elle est antisymétrique (i.e. $[u, v] = -[v, u]$) et satisfait l'**identité de Jacobi** en ce sens que

$$(13) \quad [u, [v, w]] + [w, [u, v]] + [v, [w, u]] = 0$$

(Lorsqu'il est non nul, le crochet de Lie n'est ni commutatif ni associatif¹⁵.) Le modèle des algèbres de Lie est l'espace euclidien \mathbb{R}^3 où le crochet de Lie est le produit vectoriel $[u, v] = u \times v$. Nous allons voir que l'espace vectoriel des champs de vecteurs sur une variété M est naturellement munie d'un crochet de Lie. Soient en effet U et V deux champs de vecteurs sur M et tout pour $f \in \mathcal{C}^\infty(M)$ posons

$$[U, V](f) := U(V(f)) - V(U(f))$$

Il est clair que $f \mapsto [U, V](f)$ est un endomorphisme linéaire de $\mathcal{C}^\infty(M)$. Pour montrer que $[U, V]$ est un champ de vecteurs **il reste donc à vérifier la règle de Leibniz** : si g est une autre fonction de $\mathcal{C}^\infty(M)$, alors

$$\begin{aligned} [U, V](fg) &= U[V(fg)] - V[U(fg)] \\ &= \left(U[V(f)g + fV(g)] \right) - \left(V[U(f)g + fU(g)] \right) \\ &= \left(U[V(f)]g + V(f)U(g) + U(f)V(g) + fU[V(g)] \right) - \\ &\quad \left(V[U(f)]g + U(f)V(g) + V(f)U(g) + fV[U(g)] \right) \\ &= \left(U[V(f)] - V[U(f)] \right)g + f \left(U[V(g)] - V[U(g)] \right) \end{aligned}$$

soit encore $[U, V](fg) = [U, V](f)g + f[U, V](g)$. D'autre part, il est aussi facile de vérifier que $[\cdot, \cdot] : \mathfrak{X}_0^1(TM) \times \mathfrak{X}_0^1(TM) \rightarrow \mathfrak{X}_0^1(TM)$ est une application bilinéaire qui satisfait l'identité de Jacobi et réalise donc un crochet de Lie sur $\mathfrak{X}_0^1(TM)$.

Proposition 6.1. $(\mathfrak{X}_0^1(TM), [\cdot, \cdot])$ est une algèbre de Lie ; lorsque le corchet de Lie $[U, V]$ des deux champs de vecteurs U et V is identiquement nul, on dit que U et V commutent.

Le crochet de Lie $[U, V]$ de deux champs de vecteurs est un champ de vecteurs qui mesure le degré de non-commutativité des deux champs en question : le sens de cette affirmation vient essentiellement du **Lemme de Schwarz**. Soit \mathcal{U} un ouvert de \mathbb{R}^n – muni de sa structure euclidienne et rapporté à la base canonique (e_1, \dots, e_n) ; rappelons que si $f = f(x^1, \dots, x^n) : \mathcal{U} \rightarrow \mathbb{R}$ est de classe C^2 , alors l'**identité de commutation entre dérivées partielles** affirme que pour tout $1 \leq i, j \leq n$

$$(*) \quad \frac{\partial}{\partial x^i} \left(\frac{\partial f}{\partial x^j} \right) = \frac{\partial}{\partial x^j} \left(\frac{\partial f}{\partial x^i} \right).$$

Il est possible d'interpréter cette identité en terme de crochet de Lie en écrivant

$$(*) \iff [\partial_i, \partial_j] = 0$$

15. Suivant certaines définitions, et à cause de la non associativité du crochet de Lie, une algèbre de Lie peut ne pas être considérée comme une algèbre au sens strict.

Ici, nous notons $(\partial_1, \dots, \partial_n)$ la base holonôme du système de coordonnées triviale de l'ouvert \mathcal{U} (i.e. l'identité de \mathcal{U}) de sorte que $\partial_i|_X = \mathbf{e}_i$, pour tout $X \in \mathcal{U}$. En général le crochet de Lie $[V, U]$ est non nul : on peut voir cela avec des champs de vecteurs sur l'ouvert \mathcal{U} : dans ce cas, nous profitons du fait [l'espace vectoriel \$\mathfrak{T}_0^1\(T\mathcal{U}\)\$ des champs de vecteurs sur \$\mathcal{U}\$ est un \$C^\infty\(\mathcal{U}\)\$ -module de dimension \$n\$](#) dont une base est la base holonôme $(\partial_1, \dots, \partial_n)$. Si $U = U^i \partial_i$ et $V = V^j \partial_j$ sont deux champs de vecteurs sur \mathcal{U} alors d'une part

$$\begin{aligned} U(V) &= (U^i \partial_i)(V^j \partial_j) = U^i \partial_i(V^j \partial_j) \\ &= U^i (\partial_i(V^j) \partial_j + V^j \partial_{ij}) = U^i \partial_i(V^j) \partial_j + U^i V^j \partial_{ij} \end{aligned}$$

Comme d'autre part $V(U) = V^i \partial_i(U^j) \partial_j + V^i U^j \partial_{ij}$, nous déduisons du [Lemme de Schwarz](#) que la i -ème composante du crochet $[U, V]$ est $U^i \partial_i(V^j) - V^i \partial_i(U^j)$.

Proposition 6.2. *Soient U, V deux champs de vecteurs dont les composantes (en coordonnées) sont respectivement U^i et V^i : alors les composantes de $[U, V]$ sont*

$$(14) \quad [U, V] = \left(U^i \partial_i(V^j) - V^i \partial_i(U^j) \right) \partial_j$$

6.2. Interprétation géométrique du crochet de Lie. Supposons que la variété M coïncide avec un ouvert de \mathbb{R}^m et que $\mathbf{x} = (x^1, \dots, x^m)$ est le système de coordonnées cartésiennes (i.e. $\partial_i = \partial/\partial x^i = \mathbf{e}_i$). Étant donné un point X de M fixé et $U = U^i \mathbf{e}_i$ un champ de vecteurs sur M , la courbe intégrale $t \mapsto \phi_U^t(X)$ de U issue de X est déterminée comme l'unique solution en $\varphi = (\varphi^1, \dots, \varphi^m)$ du problème de Cauchy :

$$(15) \quad (\forall k) \quad \frac{d\varphi^k}{dt} = U^k(\varphi(t)) \quad \text{avec} \quad \varphi(0) = X$$

Le développement de Taylor de w en 0 au second ordre s'écrit

$$(16) \quad (\forall k) \quad \varphi^k(t) = w(0) + t \frac{d\varphi^k}{dt}(0) + \frac{t^2}{2} \frac{d^2\varphi^k}{dt^2}(0) + \mathcal{O}(t^3)$$

Or ([méthode d'Euler](#)) d'après (15) nous avons la condition initiale $y(0) = X$, de sorte que pour $t = 0$ l'équation différentielle entraîne que

$$\frac{d\varphi^k}{dt}(0) = U^k(X) \quad \text{et} \quad \frac{d^2\varphi^k}{dt^2}(0) = \frac{\partial U^k}{\partial x^i}(X) \frac{d\varphi^i}{dt} = U^i(X) \frac{\partial U^k}{\partial x^i}(X)$$

Soit alors ε (suffisamment petit) et notons $H = \phi_U^\varepsilon(X)$: alors, d'après (16)

$$(17) \quad H^k = X^k + \varepsilon U^k(X) + \frac{\varepsilon^2}{2} U^i(X) \frac{\partial U^k}{\partial x^i}(X) + \mathcal{O}(\varepsilon^3)$$

Si maintenant $t \mapsto \phi_V^t(H)$ est la courbe intégrale d'un deuxième champ de vecteurs $V = V^i \mathbf{e}_i$ (issue du point H), alors en notant $I := \phi_V^\varepsilon(H)$, il vient (au second ordre en ε) :

$$I^k = H^k + \varepsilon V^k(H) + \frac{\varepsilon^2}{2} V^i(H) \frac{\partial V^k}{\partial x^i}(H) + \mathcal{O}(\varepsilon^3)$$

Pour nous ramener en X , nous utilisons (17) pour la valeur de H^k mais nous avons aussi,

$$\varepsilon V^k(H) = \varepsilon V^k(X + \varepsilon U^i(X) \mathbf{e}_i) + \mathcal{O}(\varepsilon^3) = \varepsilon V^k(X) + \varepsilon^2 U^i(X) \frac{\partial V^k}{\partial x^i}(X) + \mathcal{O}(\varepsilon^3)$$

et enfin

$$\frac{\varepsilon^2}{2} V^i(H) \frac{\partial V^k}{\partial x^i}(H) + \mathcal{O}(\varepsilon^3) = \frac{\varepsilon^2}{2} V^i(X) \frac{\partial V^k}{\partial x^i}(X) + \mathcal{O}(\varepsilon^3)$$

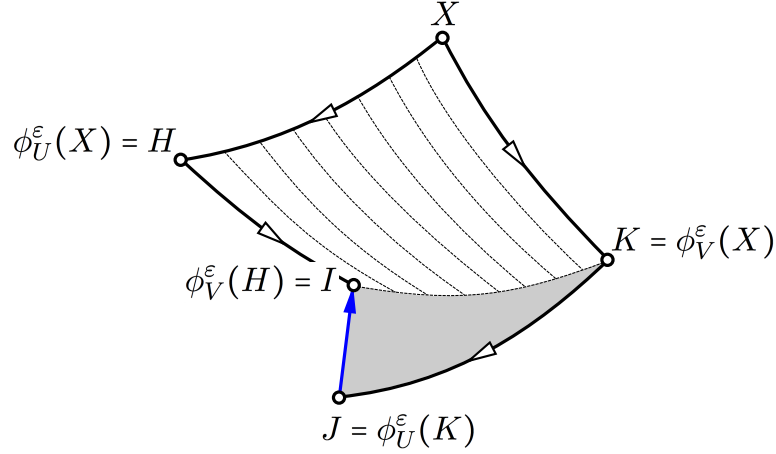


FIGURE 2. Soient $t \mapsto \phi_U^t(X)$ et $t \mapsto \phi_V^t(X)$ les courbes intégrales respectives de deux champs de vecteurs U et V issues d'un point X quelconque de M . Si $H := \phi_U^\varepsilon(X)$ et $I := \phi_V^\varepsilon(H)$ (resp. $K := \phi_V^\varepsilon(X)$ et $J := \phi_U^\varepsilon(K)$), alors l'approximation $I \approx J$ a lieu au premier ordre en ε , mais n'a pas lieu au second ordre en ε (du moins, si U et V ne commutent pas). Dans tous les cas le défaut géométrique de commutation des champs U et V se traduit par un troisième champ, i.e. le crochet de Lie $[U, V]$ de sorte qu'en X (et pour ε suffisamment petit) :

$$I - J = \phi_V^\varepsilon \circ \phi_U^\varepsilon(X) - \phi_U^\varepsilon \circ \phi_V^\varepsilon(X) = \varepsilon^2 [U, V]|_X + \mathcal{O}(\varepsilon^3)$$

En regroupant ces développements nous obtenons (à partir de X et au second ordre en ε) :

$$\begin{aligned} I^k &= X^k + \varepsilon U^k(X) + \frac{\varepsilon^2}{2} U^i(X) \frac{\partial U^k}{\partial x^i}(X) \\ &\quad + \varepsilon \left(V^k(X) + \varepsilon U^i(X) \frac{\partial V^k}{\partial x^i}(X) \right) + \frac{\varepsilon^2}{2} V^i(X) \frac{\partial V^k}{\partial x^i}(X) + \mathcal{O}(\varepsilon^3) \end{aligned}$$

soit encore

$$\begin{aligned} (18) \quad I^k &= X^k + \varepsilon \left(U^k(X) + V^k(X) \right) \\ &\quad + \varepsilon^2 U^i(X) \frac{\partial V^k}{\partial x^i}(X) \\ &\quad + \frac{\varepsilon^2}{2} \left(U^i(X) \frac{\partial U^k}{\partial x^i}(X) + V^i(X) \frac{\partial V^k}{\partial x^i}(X) \right) + \mathcal{O}(\varepsilon^3) \end{aligned}$$

En notant $K := \phi_V^\varepsilon(X)$ et $J = \phi_U^\varepsilon(K)$, nous obtenons de manière symétrique

$$\begin{aligned} (19) \quad J^k &= X^k + \varepsilon \left(V^k(X) + U^k(X) \right) \\ &\quad + \varepsilon^2 V^i(X) \frac{\partial U^k}{\partial x^i}(X) \\ &\quad + \frac{\varepsilon^2}{2} \left(V^i(X) \frac{\partial V^k}{\partial x^i}(X) + U^i(X) \frac{\partial U^k}{\partial x^i}(X) \right) + \mathcal{O}(\varepsilon^3) \end{aligned}$$

et par suite

$$I^k - J^k = \varepsilon^2 \left(U^i(X) \frac{\partial U^k}{\partial x^i}(X) - V^i(X) \frac{\partial V^k}{\partial x^i}(X) \right) + \mathcal{O}(\varepsilon^3)$$

En reconnaissant les composantes du crochet de Lie de U et V (c.f. (6.2) in Proposition 14) et compte tenu des définitions des points I et J (de \mathbb{R}^m), nous pouvons finalement écrire :

$$(20) \quad \phi_V^\varepsilon \circ \phi_U^\varepsilon(X) - \phi_U^\varepsilon \circ \phi_V^\varepsilon(X) = \varepsilon^2[U, V]|_X + \mathcal{O}(\varepsilon^3)$$

Pour étendre (20) au cas des variétés, supposons maintenant que M est une variété de dimension n plongée dans \mathbb{R}^m et que U et V sont deux champs de vecteurs définies M . Alors, pour tout $X \in M$, il existe un voisinage de X (dans \mathbb{R}^m) pour lequel les champs de vecteurs U et V sont les restrictions à M de champs de vecteurs définis sur un ouvert de \mathbb{R}^m : cela nous permet d'énoncer la proposition suivante.

Proposition 6.3. *Soient M une variété de dimension n plongée dans \mathbb{R}^m et V et U deux champs de vecteurs sur M ; si $t \mapsto \phi_U^t(X)$ et $t \mapsto \phi_V^t(X)$ désignent respectivement les courbes intégrales de U et V issues d'un point X quelconque de M , alors pour tout ε suffisamment petit :*

$$(21) \quad \phi_V^\varepsilon \circ \phi_U^\varepsilon(X) - \phi_U^\varepsilon \circ \phi_V^\varepsilon(X) = \varepsilon^2[U, V]|_X + \mathcal{O}(\varepsilon^3)$$

7. Champs de tenseurs sur une variété

7.1. Changements de bases holonômes et co-holonômes. Soit M une variété de dimension n (plongée dans \mathbb{R}^m). Soit $(\partial_1, \dots, \partial_n)$ la base holonôme des coordonnées $\mathbf{x} = (x^1, \dots, x^n)$. Rappelons que nous écrivons aussi $\partial_i = \partial/\partial x^i$, cette notation ayant l'avantage de faire apparaître les composantes cartésiennes x^i du système de coordonnées en question ; elle est aussi très utilisée pour traiter la question des changements de coordonnées, un point essentiel du calcul tensoriel. En effet, si $\hat{\mathbf{x}} = (\hat{x}^1, \dots, \hat{x}^n)$ est un deuxième système de coordonnées locales dont le domaine intersecte celui de \mathbf{x} , alors nous avons (dans l'intersection des domaines de \mathbf{x} et $\hat{\mathbf{x}}$) :

$$(22) \quad \frac{\partial}{\partial \hat{x}^\alpha} = \frac{\partial x^a}{\partial \hat{x}^\alpha} \frac{\partial}{\partial x^a}$$

Le changement de base holonômes est une traduction de l'expression du jacobien du difféomorphisme de changement de système de coordonnées qui donne l'identité

$$\frac{\partial x^a}{\partial \hat{x}^\beta} \frac{\partial \hat{x}^\beta}{\partial x^c} = \delta^a_c$$

D'autre part, si T_X^*M désigne le dual de $T_X M$ (i.e. l'espace des formes linéaires sur $T_X M$), la base co-holonôme $(dx^1|_X, \dots, dx^n|_X)$ des coordonnées locales \mathbf{x} au point X , est la base de T_X^*M associée par dualité à la base holonôme correspondante, avec :

$$dx^i|_X \left(\frac{\partial}{\partial x^j} \Big|_X \right) = \delta^i_j$$

On vérifie alors que si $(d\hat{x}^1, \dots, d\hat{x}^n)$ est la base co-holonôme des coordonnées locales $\hat{\mathbf{x}} = (\hat{x}^1, \dots, \hat{x}^n)$ alors les formules de changement de bases co-holonôme s'écrivent ¹⁶

$$(23) \quad d\hat{x}^\alpha = \frac{\partial \hat{x}^\alpha}{\partial x^a} dx^a$$

16. La formule de changement co-holonôme est formellement consistante, mais il est instructif de la vérifier rigoureusement : ainsi, avec le changement de base $d\hat{x}^\alpha = \theta^\alpha_a dx^a$, il vient :

$$\delta^\beta_\alpha = d\hat{x}^\alpha \left(\frac{\partial}{\partial \hat{x}^\beta} \right) = \theta^\alpha_b dx^b \left(\frac{\partial x^b}{\partial \hat{x}^\beta} \frac{\partial}{\partial x^b} \right) = \theta^\alpha_b \frac{\partial x^b}{\partial \hat{x}^\beta} dx^b \left(\frac{\partial}{\partial x^b} \right) = \theta^\alpha_b \frac{\partial x^b}{\partial \hat{x}^\beta}$$

ce qui signifie que (θ^α_b) est la matrice inverse de $(\partial x^b / \partial \hat{x}^\beta)$, c'est à dire que $\theta^\alpha_b = \partial \hat{x}^\beta / \partial x^b$.

7.2. Champ de tenseurs. Par définition $\mathfrak{T}_q^p(TM)$ est l'espace vectoriel des champs de tenseurs de valence (p, q) sur M : plus précisément, un élément F de $\mathfrak{T}_q^p(TM)$ est une section C^∞ du fibré vectoriel $\Pi_X \otimes^{p,q} T_X M$ qui s'exprime dans chaque système de coordonnées $\mathbf{x} = (x^1, \dots, x^n)$ en écrivant

$$F = F^{i_1 \dots i_p}_{j_1 \dots j_q} \frac{\partial}{\partial x^{i_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{i_p}} \otimes dx^{j_1} \otimes \dots \otimes dx^{j_p}$$

et où les $F^{i_1 \dots i_p}_{j_1 \dots j_q}$ sont les \mathbf{x} -composantes de F . Notons que $\mathfrak{T}_0^1(TM)$ correspond à l'espace des champs de vecteurs sur M (ce qui explique notre choix de notation initial du § 5.2) et que $\mathfrak{T}_1^0(TM)$ est l'espace des champs de covecteurs sur M , i.e. des 1-formes. Par convention nous noterons $\mathfrak{T}_0^0(TM)$ l'espace $C^\infty(M)$ des champs scalaires sur M .

En pratique, il y a deux conditions qui caractérisent les champs de tenseurs. **Il y a d'abord une condition de régularité** : le tenseur $F|_X$ s'identifie à une forme multilinéaire t.q. pour tout $\mathbf{V} \in \times_1^p \mathfrak{T}_0^1(TM)$ et $\boldsymbol{\xi} \in \times_1^q \mathfrak{T}_1^0(TM)$ fixés, l'application

$$X \mapsto F|_X(\mathbf{V}|_X, \boldsymbol{\xi}|_X)$$

est de classe C^∞ sur M . **La deuxième condition** concerne la compatibilité¹⁷ de l'expression du champ de tenseur relativement au coordonnées locales. C'est une généralisation de la formule des changements de bases holonômes/co-holonômes en (22) et (23) : si $\hat{\mathbf{x}} = (\hat{x}^1, \dots, \hat{x}^n)$ est un deuxième système de coordonnées locales alors

$$\begin{aligned} F &= F^{i_1 \dots i_p}_{j_1 \dots j_q} \frac{\partial}{\partial x^{i_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{i_p}} \otimes dx^{j_1} \otimes \dots \otimes dx^{j_p} \\ &= \left(\frac{\partial \hat{x}^{r_1}}{\partial x^{i_1}} \dots \frac{\partial \hat{x}^{r_p}}{\partial x^{i_p}} \frac{\partial x^{j_1}}{\partial \hat{x}^{s_1}} \dots \frac{\partial x^{j_q}}{\partial \hat{x}^{s_q}} F^{i_1 \dots i_p}_{j_1 \dots j_q} \right) \frac{\partial}{\partial \hat{x}^{r_1}} \otimes \dots \otimes \frac{\partial}{\partial \hat{x}^{r_p}} \otimes d\hat{x}^{s_1} \otimes \dots \otimes d\hat{x}^{s_p} \end{aligned}$$

Par suite, si $\hat{F}^{r_1 \dots r_p}_{s_1 \dots s_q}$ sont les $\hat{\mathbf{x}}$ -composantes de F , alors :

$$(24) \quad \hat{F}^{r_1 \dots r_p}_{s_1 \dots s_q} = \frac{\partial \hat{x}^{r_1}}{\partial x^{i_1}} \dots \frac{\partial \hat{x}^{r_p}}{\partial x^{i_p}} \frac{\partial x^{j_1}}{\partial \hat{x}^{s_1}} \dots \frac{\partial x^{j_q}}{\partial \hat{x}^{s_q}} F^{i_1 \dots i_p}_{j_1 \dots j_q}$$

Regardons deux exemples importants. Si V est un champ de vecteurs qui s'écrit $V = V^i \partial / \partial x^i$ dans la base holonôme d'un système de coordonnées \mathbf{x} , alors les composantes V^i de V dans ce système de coordonnées sont les composantes (contravariantes) de V en tant que tenseur. Si $V = \hat{V}^j \partial / \partial \hat{x}^j$ dans les coordonnées d'un autre système de coordonnées $\hat{\mathbf{x}}$, alors le changement de bases holonômes (22) donne

$$V = V^i \frac{\partial}{\partial x^i} = V^i \frac{\partial \hat{x}^j}{\partial x^i} \frac{\partial}{\partial \hat{x}^j} \quad \text{soit} \quad \hat{V}^j = V^i \frac{\partial \hat{x}^j}{\partial x^i}$$

Notons encore que l'espace des p -formes différentielles s'identifie à $\mathfrak{T}_p^0(TM)$. Ainsi, en partant (par exemple) de l'expression d'une 2-forme ξ dans les coordonnées locales de \mathbf{x} , nous obtenons son expression dans les coordonnées de $\hat{\mathbf{x}}$, en écrivant

$$\xi = \xi_{ab} dx^a \otimes dx^b = \xi_{ab} \left(\frac{\partial x^a}{\partial \hat{x}^\alpha} d\hat{x}^\alpha \right) \otimes \left(\frac{\partial x^b}{\partial \hat{x}^\beta} d\hat{x}^\beta \right) = \left(\frac{\partial x^a}{\partial \hat{x}^\alpha} \frac{\partial x^b}{\partial \hat{x}^\beta} \xi_{ab} \right) d\hat{x}^\alpha \otimes d\hat{x}^\beta$$

17. On parle aussi (Einstein, Weinberg,...) de propriété de covariance (avec une ambiguïté relativement à la dualité covariant/contravariant des tenseurs) : ainsi un opérateur différentiel qui est covariant est un champ de tenseur.

Si nous notons $\hat{\xi}_{\alpha\beta}$ les composantes de ξ dans les coordonnées locales de \hat{x} alors :

$$(25) \quad \hat{\xi}_{\alpha\beta} = \frac{\partial x^a}{\partial \hat{x}^\alpha} \frac{\partial x^b}{\partial \hat{x}^\beta} \xi_{ab}$$

Le lemme suivant est très important en pratique.

Proposition 7.1. Soient M une variété de dimension n et F une forme multilinéaire définie sur

$$\left(\times_1^p \mathfrak{T}_1^0(TM) \right) \times \left(\times_1^q \mathfrak{T}_0^1(TM) \right)$$

Alors F est un champ de tenseurs (p, q) si et seulement si F est $C^\infty(M)$ -multilinéaire (i.e. $C^\infty(M)$ -linéaire suivant chacune des $p + q$ entrées de F).

Preuve. Par définition, si F est un champ de tenseurs (p, q) alors F est $C^\infty(M)$ -multilinéaire. Pour la réciproque, considérons – pour simplifier – le cas $p = q = 1$ et notons F^a_b les composantes de F dans un système de coordonnées x : il s’agit de vérifier la formule de compatibilité tensorielle lors du passage du système de coordonnées x à un autre système \hat{x} (dont le domaine intersecte celui de x). Pour voir cela, soit $(\xi, V) \in \mathfrak{T}_1^0(TM) \times \mathfrak{T}_0^1(TM)$: alors, en utilisant la $C^\infty(M)$ -bilinearité de F , la définition de F^b_a et les formules de changement de base holonôme et co-holonôme dans (22) et (23) appliquées à V et ξ respectivement, il vient avec $\xi = \xi_b dx^b = \hat{\xi}_\beta d\hat{x}^\beta$ et $V = V^a \partial/\partial x^a = \hat{V}^\alpha \partial/\partial \hat{x}^\alpha$:

$$\begin{aligned} F(\xi, V) &= F\left(\xi_b dx^b, V^a \frac{\partial}{\partial x^a}\right) = \xi_b V^a F\left(dx^b, \frac{\partial}{\partial x^a}\right) \\ &= \xi_b V^a F^b_a \\ &= \left(\frac{\partial \hat{x}^\beta}{\partial x^b} \hat{\xi}^\beta\right) \left(\frac{\partial x^a}{\partial \hat{x}^\alpha} \hat{V}^\alpha\right) F^b_a = \hat{\xi}^\beta \hat{V}^\alpha \left(\frac{\partial \hat{x}^\beta}{\partial x^b} \frac{\partial x^a}{\partial \hat{x}^\alpha} F^b_a\right) \end{aligned}$$

Ainsi, par définition des coefficients \hat{F}^β_α de F dans les coordonnées de \hat{x} , nous obtenons la propriété de compatibilité des tenseurs $(1, 1)$, soit

$$(26) \quad \hat{F}^\beta_\alpha = \frac{\partial \hat{x}^\beta}{\partial x^b} \frac{\partial x^a}{\partial \hat{x}^\alpha} F^b_a$$

□

Corollaire 7.2. Soient M une variété de dimension n et $F : \times_1^p \mathfrak{T}_0^1(TM) \rightarrow \mathfrak{T}_0^1(TM)$ une application multilinéaire et pour tout système $x = (x^1, \dots, x^n) : \mathcal{U} \rightarrow \mathbb{R}^n$ de coordonnées locales, soient $F^j_{i_1 \dots i_p} : \mathcal{U} \rightarrow \mathbb{R}$ les applications t.q.

$$(27) \quad F\left(\frac{\partial}{\partial x^{i_1}}, \dots, \frac{\partial}{\partial x^{i_p}}\right) = F^j_{i_1 \dots i_p} \frac{\partial}{\partial x^j}$$

Alors les $F^j_{i_1 \dots i_p}$ sont les composantes d’un champ de tenseurs $(1, p)$ si et seulement si F est $C^\infty(M)$ -multilinéaire (i.e. $C^\infty(M)$ -linéaire sur les p entrées de F).

Remarque 7.3. Nous verrons (c.f. Définition 8.1 infra) qu’une connexion sur une variété M est une application bilinéaire $D : \mathfrak{T}_0^1(TM) \times \mathfrak{T}_0^1(TM) \rightarrow \mathfrak{T}_0^1(TM)$ qui est $C^\infty(M)$ -linéaire suivant la première entrée, mais vérifiant une règle de Leibnitz sur la seconde entrée : plus précisément, si U et V sont deux champs de vecteurs sur M et si $f \in C^\infty(M)$, alors

$$D(U, fV) = U(f)V + fD_U(V)$$

Si g est une structure riemannienne sur M (c.f. § 9), alors l'application (c.f. Formule de Koszul en (48))

$$(U, V, W) \mapsto g(\mathbf{D}(U, V), W)$$

est une forme trilinéaire sur $\mathfrak{T}_0^1(TM)$, mais ce n'est pas un tenseur $(0, 3)$.

8. Dérivations covariantes et connexions

La variété M est de dimension n , plongée isométriquement dans \mathbb{R}^m et nous considérons que $\mathbf{x} = (x^1, \dots, x^n)$ est un système de coordonnées locales donné. Si f est un champ scalaire sur M , alors les dérivées partielles $\partial f / \partial x^\mu$ sont les composantes d'un tenseur $(0, 1)$ (nous verrons ci-après qu'il s'agit du gradient covariant de f) : en effet, si $\hat{\mathbf{x}} = (\hat{x}^1, \dots, \hat{x}^n)$ est une autre système de coordonnées locales, alors la formule de la dérivée des fonctions composées donne directement la compatibilité tensorielle, i.e.

$$\frac{\partial f}{\partial \hat{x}^\nu} = \frac{\partial f}{\partial x^\mu} \frac{\partial x^\mu}{\partial \hat{x}^\nu}$$

Maintenant, si V est un champ de vecteurs de \mathbf{x} -composantes V^μ alors les $\partial V^\nu / \partial x^\mu$ ne sont pas les composantes d'un tenseur $(1, 1)$: en effet, si $V = \hat{V}^\mu$ sont les composantes de V dans une autre système $\hat{\mathbf{x}}$ de coordonnées locales, alors

$$V^\mu \frac{\partial}{\partial x^\mu} = V = \hat{V}^\alpha \frac{\partial}{\partial \hat{x}^\alpha} = \hat{V}^\alpha \frac{\partial x^\mu}{\partial \hat{x}^\alpha} \frac{\partial}{\partial x^\mu} \quad \text{i.e.} \quad \hat{V}^\alpha \frac{\partial x^\mu}{\partial \hat{x}^\alpha} = V^\mu$$

Par suite

$$\frac{\partial V^\mu}{\partial x^\nu} \frac{\partial x^\nu}{\partial \hat{x}^\beta} = \frac{\partial V^\mu}{\partial \hat{x}^\beta} = \frac{\partial}{\partial \hat{x}^\beta} \left(\hat{V}^\alpha \frac{\partial x^\mu}{\partial \hat{x}^\alpha} \right) = \frac{\partial \hat{V}^\alpha}{\partial \hat{x}^\beta} \frac{\partial x^\mu}{\partial \hat{x}^\alpha} + \hat{V}^\alpha \frac{\partial^2 x^\mu}{\partial \hat{x}^\beta \partial \hat{x}^\alpha}$$

et finalement

$$(28) \quad \frac{\partial V^\mu}{\partial x^\nu} = \frac{\partial \hat{V}^\alpha}{\partial \hat{x}^\beta} \frac{\partial x^\mu}{\partial \hat{x}^\alpha} \frac{\partial \hat{x}^\beta}{\partial x^\nu} + \hat{V}^\alpha \frac{\partial^2 x^\mu}{\partial \hat{x}^\beta \partial \hat{x}^\alpha} \frac{\partial \hat{x}^\beta}{\partial x^\nu}$$

Le deuxième terme de cette dernière somme n'étant pas nul, les $\partial V^\mu / \partial x^\nu$ ne sont pas les composantes d'un tenseur $(1, 1)$. Grâce à la notion de **connexion**, nous allons voir comment, pour tout système de coordonnées locales \mathbf{x} , **chaque dérivation partielle $\partial / \partial x^\nu$** peut être *corrigée*¹⁸ en une **dérivation partielle covariante**¹⁹ soit D_ν , de sorte que les $D_\nu(V^\mu)$ constituent les composantes d'un tenseur.

Définition 8.1. Soit M une variété différentielle de dimension n ; une **connexion (affine)** sur M est une application **\mathbb{R} -bilinéaire** $(U, V) \mapsto \mathbf{D}_U(V)$ de $\mathfrak{T}_0^1(TM) \times \mathfrak{T}_0^1(TM)$ dans $\mathfrak{T}_0^1(TM)$ t.q. pour tout $f \in C^\infty(M)$ et U, V deux champs de vecteurs quelconques :

$$(29) \quad \mathbf{D}_{fU}(V) = f \mathbf{D}_U(V)$$

$$(30) \quad \mathbf{D}_U(fV) = U(f)V + f \mathbf{D}_U(V)$$

i.e. $U \mapsto \mathbf{D}_U(V)$ est $C^\infty(M)$ -linéaire et $V \mapsto \mathbf{D}_U(V)$ vérifie la règle de Leibniz. Les **coefficients de la connexion** sont définis pour chaque système de coordonnées locales comme les fonctions $\Gamma^\sigma_{\mu\nu}$ définies (localement) de sorte que $\mathbf{D}_{\partial_\mu}(\partial_\nu) = \Gamma^\epsilon_{\mu\nu} \partial_\epsilon$.

18. Nous verrons qu'il n'y a pas une manière unique de *corriger* les dérivations partielles en dérivations covariantes : cela est à relier au fait (c.f. § 11) que les connexions forment un sous-espace affine (strict) de l'espace des applications linéaires de $\mathfrak{T}_0^1(TM) \times \mathfrak{T}_0^1(TM)$ dans $\mathfrak{T}_0^1(TM)$.

19. Une dérivation partielle covariante n'est pas une dérivation au sens strict, du fait qu'elle n'est pas associée à un champ de vecteurs (même local comme c'est le cas pour les dérivation partielles).

Soit D une connexion sur M ; alors pour V fixé, les coefficients de D nous permettent d'écrire (en coordonnées)

$$\begin{aligned} D_U(V) = U^\mu D_{\partial_\mu}(V^\nu \partial_\nu) &= U^\mu \left(\frac{\partial V^\nu}{\partial x_\mu} \partial_\nu + V^\nu D_{\partial_\mu}(\partial_\nu) \right) \\ &= U^\mu \left(\frac{\partial V^\nu}{\partial x_\mu} \partial_\nu + V^\nu \Gamma^\epsilon_{\mu\nu} \partial_\epsilon \right) \\ &= U^\mu \left(\frac{\partial V^\nu}{\partial x_\mu} \partial_\nu + V^\sigma \Gamma^\nu_{\mu\sigma} \partial_\nu \right) = U^\mu \left(\frac{\partial V^\nu}{\partial x_\mu} + V^\sigma \Gamma^\nu_{\mu\sigma} \right) \partial_\nu \end{aligned}$$

Par définition les dérivées partielles covariantes des x -composantes de V sont les

$$(31) \quad D_\nu(V^\mu) := \frac{\partial V^\mu}{\partial x^\nu} + V^\sigma \Gamma^\mu_{\sigma\nu}$$

Par le Corollaire 7.2, la $C^\infty(M)$ -linéarité des applications partielles $U \mapsto D_U(V)$ nous permet d'énoncer la propriété essentielle des dérivations covariantes d'une connexion.

Proposition 8.2. Soit D une connexion sur M ; si V est un champ de vecteur de x -composantes V^μ alors les dérivées partielles covariantes $D_\mu(V^\nu)$ sont les composantes d'un tenseur $(1, 1)$ t.q.

$$D_{\partial_\mu}(V) = D_\mu(V^\nu) \partial_\nu$$

Admettons provisoirement l'existence d'une connexion D et soient $\Gamma^\mu_{\sigma\nu}$ (resp. $\hat{\Gamma}^\mu_{\sigma\nu}$) ses coefficients dans un système de coordonnées $\mathbf{x} = (x^1, \dots, x^n)$ (resp. $\hat{\mathbf{x}} = (\hat{x}^1, \dots, \hat{x}^n)$). Pour V un champ de vecteurs fixé, nous utilisons la compatibilité tensorielle des dérivées covariantes $D_\nu(V^\mu)$ définie en (31), de sorte que par (28)

$$D_\nu(V^\mu) = \frac{\partial \hat{V}^\alpha}{\partial \hat{x}^\beta} \frac{\partial x^\mu}{\partial \hat{x}^\alpha} \frac{\partial \hat{x}^\beta}{\partial x^\nu} + \hat{V}^\alpha \frac{\partial^2 x^\mu}{\partial \hat{x}^\beta \partial \hat{x}^\alpha} \frac{\partial \hat{x}^\beta}{\partial x^\nu} + V^\sigma \Gamma^\mu_{\sigma\nu} = \left(\frac{\partial \hat{V}^\alpha}{\partial \hat{x}^\beta} + \hat{V}^\sigma \hat{\Gamma}^\alpha_{\sigma\beta} \right) \frac{\partial x^\mu}{\partial \hat{x}^\alpha} \frac{\partial \hat{x}^\beta}{\partial x^\nu}$$

soit, après simplification, et du fait que $V^\sigma = \hat{V}^\alpha (\partial x^\sigma / \partial \hat{x}^\alpha)$,

$$\hat{V}^\alpha \frac{\partial^2 x^\mu}{\partial \hat{x}^\beta \partial \hat{x}^\alpha} \frac{\partial \hat{x}^\beta}{\partial x^\nu} + \hat{V}^\alpha \frac{\partial x^\sigma}{\partial \hat{x}^\alpha} \Gamma^\mu_{\sigma\nu} = \hat{V}^\sigma \hat{\Gamma}^\alpha_{\sigma\beta} \frac{\partial x^\mu}{\partial \hat{x}^\alpha} \frac{\partial \hat{x}^\beta}{\partial x^\nu}$$

soit encore

$$\hat{V}^\alpha \frac{\partial^2 x^\mu}{\partial \hat{x}^\beta \partial \hat{x}^\alpha} \frac{\partial \hat{x}^\alpha}{\partial x^\mu} + \hat{V}^\alpha \frac{\partial x^\sigma}{\partial \hat{x}^\alpha} \frac{\partial \hat{x}^\alpha}{\partial x^\mu} \frac{\partial x^\nu}{\partial \hat{x}^\beta} \Gamma^\mu_{\sigma\nu} = \hat{V}^\sigma \hat{\Gamma}^\alpha_{\sigma\beta}$$

Le champ de vecteur V étant arbitraire, nous obtenons le système d'identités

$$(32) \quad \hat{\Gamma}^\alpha_{\sigma\beta} = \frac{\partial^2 x^\mu}{\partial \hat{x}^\beta \partial \hat{x}^\alpha} \frac{\partial \hat{x}^\alpha}{\partial x^\mu} + \frac{\partial x^\sigma}{\partial \hat{x}^\alpha} \frac{\partial \hat{x}^\alpha}{\partial x^\mu} \frac{\partial x^\nu}{\partial \hat{x}^\beta} \Gamma^\mu_{\sigma\nu}$$

Ces identités montrent que les coefficients de connexion $\Gamma^\mu_{\sigma\nu}$ ne sont pas les composantes d'un tenseur $(1, 2)$ (comme la notation pourrait le laisser suggérer).

Proposition 8.3. Supposons que chaque système de coordonnées $\mathbf{x} : \mathcal{U} \rightarrow \mathbb{R}$ sur M est associée à une famille de fonctions $C^\sigma_{\mu\nu}(\cdot|\mathbf{x}) : \mathcal{U} \rightarrow \mathbb{R}$ supposées lisses ; si la condition de compatibilité

$$(33) \quad C^\alpha_{\sigma\beta}(\cdot|\hat{\mathbf{x}}) = \frac{\partial^2 x^\mu}{\partial \hat{x}^\beta \partial \hat{x}^\alpha} \frac{\partial \hat{x}^\alpha}{\partial x^\mu} + \frac{\partial x^\sigma}{\partial \hat{x}^\alpha} \frac{\partial \hat{x}^\alpha}{\partial x^\mu} \frac{\partial x^\nu}{\partial \hat{x}^\beta} C^\mu_{\sigma\nu}(\cdot|\mathbf{x})$$

est satisfaite dès que les domaines de \mathbf{x} et de $\hat{\mathbf{x}}$ s'intersectent, alors les $C^\sigma_{\mu\nu}(\cdot|\mathbf{x})$ sont les coefficients d'une connexion D sur M dans le système de coordonnées \mathbf{x} .

Preuve. Soient \mathbf{x} et $\hat{\mathbf{x}}$ deux systèmes de coordonnées dont les domaines s'intersectent ; dans la suite nous noterons simplement $C^\sigma{}_{\mu\nu} := C^\sigma{}_{\mu\nu}(\cdot|\mathbf{x})$ et $\hat{C}^\alpha{}_{\sigma\beta} := C^\alpha{}_{\sigma\beta}(\cdot|\hat{\mathbf{x}})$. Étant donné $V = V^\nu \partial/\partial \mathbf{x}^\nu = \hat{V}^\beta \partial/\partial \hat{\mathbf{x}}^\beta$ un champ de vecteur, nous voulons démontrer la formule de compatibilité tensorielle, soit :

$$(34) \quad \left(\frac{\partial V^\nu}{\partial \mathbf{x}^\mu} + V^\sigma C^\nu{}_{\sigma\mu} \right) \frac{\partial \mathbf{x}^\mu}{\partial \hat{\mathbf{x}}^\alpha} \frac{\partial \hat{\mathbf{x}}^\beta}{\partial \mathbf{x}^\nu} = \frac{\partial \hat{V}^\beta}{\partial \hat{\mathbf{x}}^\alpha} + \hat{V}^\gamma \hat{C}^\beta{}_{\gamma\alpha}$$

Comme $V^\nu = \hat{V}^\beta (\partial \mathbf{x}^\nu / \partial \hat{\mathbf{x}}^\beta)$ nous avons

$$(35) \quad \frac{\partial V^\nu}{\partial \mathbf{x}^\mu} = \frac{\partial V^\nu}{\partial \hat{\mathbf{x}}^\alpha} \frac{\partial \hat{\mathbf{x}}^\alpha}{\partial \mathbf{x}^\mu} = \frac{\partial}{\partial \hat{\mathbf{x}}^\alpha} \left(\hat{V}^\beta \frac{\partial \mathbf{x}^\nu}{\partial \hat{\mathbf{x}}^\beta} \right) \frac{\partial \hat{\mathbf{x}}^\alpha}{\partial \mathbf{x}^\mu} = \frac{\partial \hat{V}^\beta}{\partial \hat{\mathbf{x}}^\alpha} \frac{\partial \mathbf{x}^\nu}{\partial \hat{\mathbf{x}}^\beta} \frac{\partial \hat{\mathbf{x}}^\alpha}{\partial \mathbf{x}^\mu} + \hat{V}^\beta \frac{\partial^2 \mathbf{x}^\nu}{\partial \hat{\mathbf{x}}^\alpha \hat{\mathbf{x}}^\beta} \frac{\partial \hat{\mathbf{x}}^\alpha}{\partial \mathbf{x}^\mu}$$

Nous pouvons maintenant transformer le membre de droite de (34) :

$$\begin{aligned} \left(\frac{\partial V^\nu}{\partial \mathbf{x}^\mu} + V^\sigma C^\nu{}_{\sigma\mu} \right) \frac{\partial \mathbf{x}^\mu}{\partial \hat{\mathbf{x}}^\alpha} \frac{\partial \hat{\mathbf{x}}^\beta}{\partial \mathbf{x}^\nu} &= \left(\frac{\partial \hat{V}^\beta}{\partial \hat{\mathbf{x}}^\alpha} \frac{\partial \mathbf{x}^\nu}{\partial \hat{\mathbf{x}}^\beta} \frac{\partial \hat{\mathbf{x}}^\alpha}{\partial \mathbf{x}^\mu} + \hat{V}^\beta \frac{\partial^2 \mathbf{x}^\nu}{\partial \hat{\mathbf{x}}^\alpha \hat{\mathbf{x}}^\beta} \frac{\partial \hat{\mathbf{x}}^\alpha}{\partial \mathbf{x}^\mu} + \hat{V}^\gamma \frac{\partial \mathbf{x}^\sigma}{\partial \hat{\mathbf{x}}^\gamma} \frac{\partial \mathbf{x}^\mu}{\partial \hat{\mathbf{x}}^\sigma} C^\nu{}_{\sigma\mu} \right) \frac{\partial \mathbf{x}^\mu}{\partial \hat{\mathbf{x}}^\alpha} \frac{\partial \hat{\mathbf{x}}^\beta}{\partial \mathbf{x}^\nu} \\ &= \frac{\partial \hat{V}^\beta}{\partial \hat{\mathbf{x}}^\alpha} + \hat{V}^\beta \frac{\partial^2 \mathbf{x}^\nu}{\partial \hat{\mathbf{x}}^\alpha \hat{\mathbf{x}}^\beta} \frac{\partial \hat{\mathbf{x}}^\alpha}{\partial \mathbf{x}^\nu} + \hat{V}^\gamma \frac{\partial \mathbf{x}^\sigma}{\partial \hat{\mathbf{x}}^\gamma} \frac{\partial \mathbf{x}^\mu}{\partial \hat{\mathbf{x}}^\sigma} C^\nu{}_{\sigma\mu} \\ &= \frac{\partial \hat{V}^\beta}{\partial \hat{\mathbf{x}}^\alpha} + \hat{V}^\gamma \frac{\partial^2 \mathbf{x}^\nu}{\partial \hat{\mathbf{x}}^\alpha \hat{\mathbf{x}}^\gamma} \frac{\partial \hat{\mathbf{x}}^\alpha}{\partial \mathbf{x}^\nu} + \hat{V}^\gamma \frac{\partial \mathbf{x}^\sigma}{\partial \hat{\mathbf{x}}^\gamma} \frac{\partial \mathbf{x}^\mu}{\partial \hat{\mathbf{x}}^\sigma} C^\nu{}_{\sigma\mu} \\ &= \frac{\partial \hat{V}^\beta}{\partial \hat{\mathbf{x}}^\alpha} + \hat{V}^\gamma \left(\frac{\partial^2 \mathbf{x}^\nu}{\partial \hat{\mathbf{x}}^\alpha \hat{\mathbf{x}}^\gamma} \frac{\partial \hat{\mathbf{x}}^\alpha}{\partial \mathbf{x}^\nu} + \frac{\partial \mathbf{x}^\sigma}{\partial \hat{\mathbf{x}}^\gamma} \frac{\partial \mathbf{x}^\mu}{\partial \hat{\mathbf{x}}^\sigma} C^\nu{}_{\sigma\mu} \right) \\ &= \frac{\partial \hat{V}^\beta}{\partial \hat{\mathbf{x}}^\alpha} + \hat{V}^\gamma \hat{C}^\beta{}_{\gamma\alpha} \end{aligned}$$

où nous avons utilisé la formule (33) de compatibilité. Ainsi les

$$D_\mu(V^\nu) := \frac{\partial V^\nu}{\partial \mathbf{x}^\mu} + V^\sigma C^\nu{}_{\sigma\mu}$$

sont les composantes d'un tenseur (1, 1). La conclusion vient facilement en vérifiant que $(U, V) \mapsto \mathbf{D}_U(V) = U^\mu D_\mu(V^\nu) \partial_\nu$ est une connexion t.q. $\mathbf{D}_{\partial_\mu}(\partial_\nu) = C^\sigma{}_{\mu\nu} \partial_\sigma$. □

9. Variétés pseudo-riemanniennes et connexion de Levi-Civita

Sans aborder le problème de l'existence, nous avons vu comment s'articulent les notions de dérivations covariantes et de connexions sur une variété M – chacune d'elle pouvant se déduire de l'autre et réciproquement. Cependant, mis à part l'idée initiale (un peu vague) consistant à chercher une version covariante des dérivées partielles des composantes des champs de vecteurs, il est difficile de se faire une idée précise de l'intérêt de ces notions : ceci viendra avec la caractérisation des géodésiques (pour le lagrangien inertiel) par le transport parallèle (c.f. § 12). Dans cette section, nous définissons la notion de variété pseudo-riemannienne : la métrique nous permet alors de définir successivement les [symboles de Christoffel](#) puis la [connexion de Levi-Civita](#) comme la connexion dont les coefficients coïncident avec les symboles de Christoffel. Nous verrons au § 11 en quel sens la connexion de Levi-Civita d'une variété pseudo-riemannienne est caractérisée par deux conditions, à savoir l'[absence de torsion](#) et la [compatibilité avec la métrique](#).

9.1. Structures pseudo-riemanniennes. Soit M une variété de dimension n (qu'on suppose toujours plongée dans \mathbb{R}^m) et g un tenseur de valence $(0, 2)$: alors g est une métrique pseudo-riemannienne sur M – on dit aussi que (M, g) est une variété pseudo-riemannienne – si (1) : g est symétrique, en ce sens que $g(U, V) = g(V, U)$, pour tout couple (U, V) de champs de vecteurs sur M et (2) : g est non dégénéré, ce qu'on peut traduire en disant que pour tout $X \in M$ et tout système de coordonnées x en X , le déterminant de la matrice²⁰ $G := (g_{ij}(X))$ est non nul. Par le **Théorème Spectral**, il existe une matrice orthogonale P (dépendant de X) et une matrice diagonale $\text{diag}(\lambda_1, \dots, \lambda_n)$ (à coefficients diagonaux dépendant aussi de X) t.q. $D = P^{-1}GP$. La signature de g (en X) est le couple d'entiers (p, q) t.q. $p + q = n$ et où p (resp. q) est le nombre de λ_i qui sont négatifs (resp. positifs) : on montre que la signature de g en X ne dépend pas de X et c'est par définition la signature de g . La métrique g est dite riemannienne si sa signature est égale à $(0, n)$: cela signifie que la restriction $g|_X$ à $T_X(M)$ est un produit scalaire euclidien ; dans ce cas, il n'y a pas de perte de généralité à supposer²¹ que (M, g) est isométriquement plongée dans \mathbb{R}^m (muni de sa structure euclidienne canonique), c'est-à-dire que $g|_X$ coïncide avec le produit scalaire induit par la structure euclidienne de \mathbb{R}^m sur $T_X M$. Un cas important pour la relativité générale est celui où (M, g) est de dimension 4 et de signature²² $(1, 3)$ ou $(3, 1)$: on parle alors de variété lorentzienne.

9.2. Abaissement/relèvement des indices tensoriels. De même qu'un produit scalaire euclidien sur un espace vectoriel E , permet de définir le relèvement/abaissement des indices covariants/contravariants d'un tenseur (c.f. indications en § 3), une structure pseudo-riemannienne g sur M autorise le relèvement/abaissement des indices covariants/contravariants des composantes des champs de tenseurs. Nous ne présentons pas formellement l'utilisation de cette technique, préférant plutôt donner un exemple d'application. Supposons donc que (M, g) est une variété pseudo-riemannienne et soit F un tenseur de valence $(1, p)$; alors, pour tout $X \in M$, il existe une application p -linéaire $\tilde{F}|_X$ de $\times_1^p T_X M$ et à valeurs dans $T_X M$ telle que pour tout i_1, \dots, i_p

$$\tilde{F}|_X \left(\frac{\partial}{\partial x^{i_1}}, \dots, \frac{\partial}{\partial x^{i_p}} \right) = F^j{}_{i_1 \dots i_p}(X) \frac{\partial}{\partial x^j} \Big|_X$$

où les $F^j{}_{i_1 \dots i_p}$ sont les x -composantes de F . Alors, en posant

$$(36) \quad F_{i_1 \dots i_p} := g \left(\frac{\partial}{\partial x^i}, \tilde{F} \left(\frac{\partial}{\partial x^{i_1}}, \dots, \frac{\partial}{\partial x^{i_p}} \right) \right) = g \left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right) F^j{}_{i_1 \dots i_p} = g_{ij} F^j{}_{i_1 \dots i_p}$$

(abaissement de l'indice contravariant) nous définissons les x -composantes d'un champ de tenseurs $(0, p + 1)$. Maintenant (c.f. (7) et Proposition 3.2), si nous notons g^* le $(2, 0)$ tenseur t.q. (pour tout X) $(g^*)|_X = (g|_X)^*$, alors les x -composantes de g^* sont les g^{ij} t.q. $g^{ik} g_{ki} \equiv \delta^i_j$; alors nous pouvons démontrer de manière analogue à (36) que (relèvement du premier indice covariant) :

$$F^i{}_{i_1 \dots i_p} = g^{ij} F_{j i_1 \dots i_p}$$

20. Matrice, en tant que simple tableau $n \times n$ de nombres réels, i.e. un élément de $\mathcal{M}_n(\mathbb{R})$.

21. D'après le Théorème de Plongement de Nash (1954-56) [Nas54, Nas56].

22. On trouve aussi la notation $(- + +)$ pour la signature $(1, 3)$ et $(- - +)$ pour la signature $(3, 1)$.

9.3. Symboles de Christoffel. Soit (M, g) est une variété riemannienne de dimension n plongée isométriquement dans \mathbb{R}^m (nous traiterons ultérieurement le cas pseudo-riemannien). Étant donné $(\mathcal{U}, \mathbf{x})$ un système de coordonnées, nous notons $\mathbf{X}^\nu(x^1, \dots, x^n)$ (pour $\nu = 1, \dots, m$ et $x = (x^1, \dots, x^n) \in \mathbf{x}(\mathcal{U})$) les m composantes cartésiennes de la carte \mathbf{X} associée à \mathbf{x} (et considérée comme une application de $\mathbf{x}(\mathcal{U})$ à valeurs dans l'espace ambiant \mathbb{R}^m). Alors par définition de $(\partial_1, \dots, \partial_n)$ (base holonôme de \mathbf{x}) et de $(\mathbf{e}_1, \dots, \mathbf{e}_m)$ (base canonique de \mathbb{R}^m), nous avons pour $X = \mathbf{X}(x^1, \dots, x^n)$

$$\partial_\mu|_X = \frac{\partial \mathbf{X}}{\partial x^\mu} := \sum_{\nu=1}^m \frac{\partial \mathbf{X}^\nu}{\partial x^\mu} \mathbf{e}_\nu$$

De même que \mathbf{X} , les différentielles partielles $\partial \mathbf{X} / \partial x^\mu$ sont des applications de $\mathbf{x}(\mathcal{U})$ dans \mathbb{R}^m : cela nous permet de définir les différentielles partielles secondes

$$\frac{\partial^2 \mathbf{X}}{\partial x^\mu \partial x^\nu} := \frac{\partial}{\partial x^\mu} \left(\frac{\partial \mathbf{X}}{\partial x^\nu} \right)$$

Soit $V = V^\mu \partial_\mu$ un champ de vecteurs et posons $\hat{V} := V \circ \mathbf{X}$ (resp. $\hat{V}^\mu = V^\mu \circ \mathbf{X}$) : ainsi, \hat{V} est une application définie sur un ouvert de \mathbb{R}^n (i.e. $\mathbf{x}(\mathcal{U})$), à valeur dans \mathbb{R}^m (l'espace ambiant de M) qui de plus vérifie $\hat{V}(x) = V|_X \in T_X M$ dès que $x = \mathbf{x}(X)$. Du fait que $\partial_\nu|_X = \partial \mathbf{X} / \partial x^\nu$, il vient $\hat{V}(x) = \hat{V}^\nu(x) \partial \mathbf{X} / \partial x^\nu$ et en appliquant la différentielle (usuelle) du produit nous obtenons

$$(37) \quad \frac{\partial \hat{V}}{\partial x^\mu} = \frac{\partial \hat{V}^\nu}{\partial x^\mu} \frac{\partial \mathbf{X}}{\partial x^\nu} + \hat{V}^\nu(x) \frac{\partial^2 \mathbf{X}}{\partial x^\mu \partial x^\nu}$$

Le premier terme est dans $T_X M$, mais pas nécessairement le second. Par définition, [Les symboles de Christoffel dans les coordonnées locales](#) d'un système de coordonnées $\mathbf{x} : \mathcal{U} \rightarrow \mathbb{R}^n$ sont les applications $\mathbf{C}^\lambda_{\mu\nu} : \mathcal{U} \rightarrow \mathbb{R}$ t.q. nous ayons

$$(38) \quad (\forall \sigma) \quad \left\langle \frac{\partial^2 \mathbf{X}}{\partial x^\mu \partial x^\nu} - \mathbf{C}^\lambda_{\mu\nu}(\mathbf{X}(x)) \frac{\partial \mathbf{X}}{\partial x^\lambda} \middle| \frac{\partial \mathbf{X}}{\partial x^\sigma} \right\rangle = 0$$

D'une part, nous tirons de (38) que

$$(39) \quad \left\langle \frac{\partial^2 \mathbf{X}}{\partial x^\mu \partial x^\nu} \middle| \frac{\partial \mathbf{X}}{\partial x^\sigma} \right\rangle = \left\langle \frac{\partial \mathbf{X}}{\partial x^\lambda} \middle| \frac{\partial \mathbf{X}}{\partial x^\sigma} \right\rangle \mathbf{C}^\lambda_{\mu\nu}(\mathbf{X}(x)) = G_{\lambda\sigma} \mathbf{C}^\lambda_{\mu\nu}(\mathbf{X}(x))$$

où $G_{ij} = g_{ij} \circ \mathbf{X}$. Mais d'autre part, du fait que $G_{\alpha\beta} = \left\langle \frac{\partial \mathbf{X}}{\partial x^\alpha} \middle| \frac{\partial \mathbf{X}}{\partial x^\beta} \right\rangle$ il vient aussi

$$\begin{aligned} \frac{\partial G_{\nu\lambda}}{\partial x^\mu} &= \left\langle \frac{\partial^2 \mathbf{X}}{\partial x^\mu \partial x^\nu} \middle| \frac{\partial \mathbf{X}}{\partial x^\lambda} \right\rangle + \left\langle \frac{\partial \mathbf{X}}{\partial x^\nu} \middle| \frac{\partial^2 \mathbf{X}}{\partial x^\mu \partial x^\lambda} \right\rangle \\ \frac{\partial G_{\lambda\mu}}{\partial x^\nu} &= \left\langle \frac{\partial^2 \mathbf{X}}{\partial x^\nu \partial x^\lambda} \middle| \frac{\partial \mathbf{X}}{\partial x^\mu} \right\rangle + \left\langle \frac{\partial \mathbf{X}}{\partial x^\lambda} \middle| \frac{\partial^2 \mathbf{X}}{\partial x^\nu \partial x^\mu} \right\rangle \\ \frac{\partial G_{\mu\nu}}{\partial x^\lambda} &= \left\langle \frac{\partial^2 \mathbf{X}}{\partial x^\lambda \partial x^\mu} \middle| \frac{\partial \mathbf{X}}{\partial x^\nu} \right\rangle + \left\langle \frac{\partial \mathbf{X}}{\partial x^\mu} \middle| \frac{\partial^2 \mathbf{X}}{\partial x^\lambda \partial x^\nu} \right\rangle \end{aligned}$$

soit encore avec (39)

$$\frac{\partial G_{\nu\lambda}}{\partial x^\mu} + \frac{\partial G_{\lambda\mu}}{\partial x^\nu} - \frac{\partial G_{\mu\nu}}{\partial x^\lambda} = 2 \left\langle \frac{\partial^2 \mathbf{X}}{\partial x^\mu \partial x^\nu} \middle| \frac{\partial \mathbf{X}}{\partial x^\lambda} \right\rangle = 2G_{\lambda\sigma} \mathbf{C}^\lambda_{\mu\nu}(\mathbf{X}(x))$$

et finalement (en terme des fonctions définies dans le domaine du système de coordonnées considéré)

$$(40) \quad \mathbf{C}^\lambda{}_{\mu\nu} = \frac{1}{2}g^{\lambda\sigma}(\partial_\mu(g_{\nu\sigma}) + \partial_\nu(g_{\sigma\mu}) - \partial_\sigma(g_{\mu\nu}))$$

Nous transformons (37) en écrivant

$$\begin{aligned} \frac{\partial \hat{V}}{\partial x^\mu} &= \frac{\partial \hat{V}^\nu}{\partial x^\mu} \frac{\partial \mathbf{X}(x)}{\partial x^\nu} + \hat{V}^\nu(x) \left(\mathbf{C}^\lambda{}_{\mu\nu}(\mathbf{X}(x)) \frac{\partial \mathbf{X}}{\partial x^\lambda} + \frac{\partial^2 \mathbf{X}}{\partial x^\mu \partial x^\nu} - \mathbf{C}^\lambda{}_{\mu\nu}(\mathbf{X}(x)) \frac{\partial \mathbf{X}}{\partial x^\lambda} \right) \\ &= \left(\frac{\partial \hat{V}^\nu}{\partial x^\mu} + \hat{V}^\nu(x) \mathbf{C}^\lambda{}_{\mu\nu}(\mathbf{X}(x)) \right) \frac{\partial \mathbf{X}}{\partial x^\lambda} + \hat{V}^\nu(x) \left(\frac{\partial^2 \mathbf{X}}{\partial x^\mu \partial x^\nu} - \mathbf{C}^\lambda{}_{\mu\nu}(\mathbf{X}(x)) \frac{\partial \mathbf{X}}{\partial x^\lambda} \right) \end{aligned}$$

qui est la décomposition de $\partial \hat{V} / \partial x^\mu$ suivant $T_X M$ et son orthogonal. Cette présentation (valable pour une variété riemannienne de dimension n et plongée isométriquement dans \mathbb{R}^m), nous permet de poser la définition suivante (dans le cas pseudo-riemannien) :

Définition 9.1. Soit (M, g) une variété pseudo-riemannienne de dimension n ; alors,

(i) : en coordonnées, les *symboles de Christoffel de première espèce* sont les $\mathbf{C}_{\lambda\mu\nu} : \mathcal{U} \rightarrow \mathbb{R} \text{ t. } q$

$$\mathbf{C}_{\lambda\mu\nu} = \frac{1}{2}(\partial_\mu(g_{\sigma\nu}) + \partial_\nu(g_{\mu\sigma}) - \partial_\sigma(g_{\mu\nu}))$$

(ii) : en coordonnées, les *symboles de Christoffel de deuxième espèce* sont les $\mathbf{C}^\lambda{}_{\mu\nu} : \mathcal{U} \rightarrow \mathbb{R} \text{ t. } q$

$$\mathbf{C}^\lambda{}_{\mu\nu} = g^{\lambda\sigma} \mathbf{C}_{\sigma\mu\nu} = \frac{1}{2}g^{\lambda\sigma}(\partial_\mu(g_{\sigma\nu}) + \partial_\nu(g_{\mu\sigma}) - \partial_\sigma(g_{\mu\nu}))$$

Par des calculs directs utilisant la symétrie du tenseur métrique nous pouvons démontrer (exercice) les propriétés suivantes du symboles de Christoffel. Nous verrons (c.f. Proposition 10.1) que ces propriétés s'interprètent en correspondance avec des propriétés spécifiques portant sur les connexions (c.f. carte mentale in Fig. 3).

Proposition 9.2. Les symboles de Christoffel vérifient

$$(41) \quad \mathbf{C}^\sigma{}_{\mu\nu} = \mathbf{C}^\sigma{}_{\nu\mu}$$

$$(42) \quad \mathbf{C}_{\sigma\mu\nu} = \mathbf{C}_{\sigma\nu\mu}$$

$$(43) \quad \partial_\sigma(g_{\mu\nu}) = \mathbf{C}_{\mu\nu\sigma} + \mathbf{C}_{\nu\mu\sigma}$$

$$(44) \quad \partial_\sigma(g_{\mu\nu}) = \mathbf{C}_{\mu\sigma\nu} + \mathbf{C}_{\nu\sigma\mu}$$

Théorème 9.3. Soit (M, g) une variété pseudo-riemannienne de dimension n ; alors l'application $(U, V) \mapsto \nabla_U(V)$ s'exprimant en coordonnées locales par

$$\nabla_U(V) = U^\mu \nabla_\mu(V^\nu) \partial_\nu \quad \text{où} \quad \nabla_\mu V^\nu := \partial_\mu V^\nu + V^\epsilon \mathbf{C}^\nu{}_{\epsilon\mu}$$

définit une connexion sur M dont les coefficients coïncident avec les symboles de Christoffel, i.e.

$$\nabla_{\partial_\mu}(\partial_\nu) = \mathbf{C}^\sigma{}_{\mu\nu} \partial_\sigma$$

Par définition \mathbf{D} est la *connexion de Levi-Civita*.

Par la Proposition 8.3, le Théorème 9.3 est une corollaire de la proposition suivante.

Proposition 9.4. Si $\mathbf{C}^c{}_{ab}$ (resp. $\hat{\mathbf{C}}^\gamma{}_{\alpha\beta}$) sont les symboles de Christoffel en \mathbf{x} (resp. $\hat{\mathbf{x}}$), alors

$$(45) \quad \hat{\mathbf{C}}^\gamma{}_{\alpha\beta} = \frac{\partial \hat{\mathbf{x}}^\gamma}{\partial \mathbf{x}^c} \frac{\partial \mathbf{x}^a}{\partial \hat{\mathbf{x}}^\alpha} \frac{\partial \mathbf{x}^b}{\partial \hat{\mathbf{x}}^\beta} \mathbf{C}^c{}_{ab} + \frac{\partial \hat{\mathbf{x}}^\gamma}{\partial \mathbf{x}^c} \frac{\partial^2 \mathbf{x}^c}{\partial \hat{\mathbf{x}}^\alpha \partial \hat{\mathbf{x}}^\beta}$$

Preuve. Par définition, nous avons respectivement

$$\mathbb{C}^c_{ab} = \frac{1}{2}g^{cd} \left(\frac{\partial g_{db}}{\partial x_a} + \frac{\partial g_{da}}{\partial x_b} - \frac{\partial g_{ab}}{\partial x_d} \right) \quad \text{et} \quad \hat{\mathbb{C}}^c_{\alpha\beta} = \frac{1}{2}\hat{g}^{\gamma\delta} \left(\frac{\partial \hat{g}_{\beta\delta}}{\partial \hat{x}_\alpha} + \frac{\partial \hat{g}_{\alpha\delta}}{\partial \hat{x}_\beta} - \frac{\partial \hat{g}_{\alpha\beta}}{\partial \hat{x}_\delta} \right)$$

(où le " $\hat{}$ " marque les symboles dans les coordonnées locales de \hat{x}). Le changement de variables sur le tenseur métrique donne

$$\hat{g}_{\alpha\beta} = \frac{\partial x^a}{\partial \hat{x}^\alpha} \frac{\partial x^b}{\partial \hat{x}^\beta} g_{ab} \quad \text{et} \quad \hat{g}^{\alpha\beta} = \frac{\partial \hat{x}^\alpha}{\partial x^a} \frac{\partial \hat{x}^\beta}{\partial x^b} g^{ab}$$

Par suite

$$\begin{aligned} \frac{1}{2}\hat{g}^{\gamma\delta} \left(\frac{\partial \hat{g}_{\beta\delta}}{\partial \hat{x}_\alpha} \right) &= \frac{1}{2} \left(\frac{\partial \hat{x}^\gamma}{\partial x^c} \frac{\partial \hat{x}^\delta}{\partial x^d} g^{cd} \right) \frac{\partial}{\partial \hat{x}_\alpha} \left(\frac{\partial x^b}{\partial \hat{x}^\beta} \frac{\partial x^d}{\partial \hat{x}^\delta} g_{bd} \right) \\ &= \frac{1}{2} \left(\frac{\partial \hat{x}^\gamma}{\partial x^c} \frac{\partial \hat{x}^\delta}{\partial x^d} g^{cd} \right) \left(\frac{\partial}{\partial \hat{x}_\alpha} \left(\frac{\partial x^b}{\partial \hat{x}^\beta} \frac{\partial x^d}{\partial \hat{x}^\delta} \right) g_{bd} + \frac{\partial x^b}{\partial \hat{x}^\beta} \frac{\partial x^d}{\partial \hat{x}^\delta} \frac{\partial x^a}{\partial \hat{x}^\alpha} \frac{\partial g_{bd}}{\partial x_a} \right) \\ &= \frac{1}{2} \left(\frac{\partial \hat{x}^\gamma}{\partial x^c} \frac{\partial \hat{x}^\delta}{\partial x^d} \right) \frac{\partial}{\partial \hat{x}_\alpha} \left(\frac{\partial x^b}{\partial \hat{x}^\beta} \frac{\partial x^d}{\partial \hat{x}^\delta} \right) \delta_b^c + \frac{\partial \hat{x}^\gamma}{\partial x^c} \frac{\partial x^b}{\partial \hat{x}^\beta} \frac{\partial x^a}{\partial \hat{x}^\alpha} \left(\frac{1}{2} g^{cd} \left(\frac{\partial g_{bd}}{\partial x_a} \right) \right) \\ &= \frac{1}{2} \left(\frac{\partial \hat{x}^\beta}{\partial x^b} \frac{\partial \hat{x}^\delta}{\partial x^d} \right) \left(\frac{\partial^2 x^b}{\partial \hat{x}_\alpha \partial \hat{x}^\beta} \frac{\partial x^d}{\partial \hat{x}^\delta} + \frac{\partial x^b}{\partial \hat{x}^\beta} \frac{\partial^2 x^d}{\partial \hat{x}_\alpha \partial \hat{x}^\delta} \right) + \frac{\partial \hat{x}^\gamma}{\partial x^c} \frac{\partial x^b}{\partial \hat{x}^\beta} \frac{\partial x^a}{\partial \hat{x}^\alpha} \left(\frac{1}{2} g^{cd} \left(\frac{\partial g_{bd}}{\partial x_a} \right) \right) \\ &= \frac{1}{2} \left(\frac{\partial \hat{x}^\beta}{\partial x^b} \frac{\partial^2 x^b}{\partial \hat{x}_\alpha \partial \hat{x}^\beta} + \frac{\partial \hat{x}^\delta}{\partial x^d} \frac{\partial^2 x^d}{\partial \hat{x}_\alpha \partial \hat{x}^\delta} \right) + \frac{\partial \hat{x}^\gamma}{\partial x^c} \frac{\partial x^b}{\partial \hat{x}^\beta} \frac{\partial x^a}{\partial \hat{x}^\alpha} \left(\frac{1}{2} g^{cd} \left(\frac{\partial g_{bd}}{\partial x_a} \right) \right) \end{aligned}$$

et par un calcul analogue

$$\begin{aligned} \frac{1}{2}\hat{g}^{\gamma\delta} \left(\frac{\partial \hat{g}_{\alpha\delta}}{\partial \hat{x}_\beta} \right) &= \frac{1}{2} \left(\frac{\partial \hat{x}^\beta}{\partial x^b} \frac{\partial^2 x^b}{\partial \hat{x}_\alpha \partial \hat{x}^\beta} + \frac{\partial \hat{x}^\delta}{\partial x^d} \frac{\partial^2 x^d}{\partial \hat{x}_\alpha \partial \hat{x}^\delta} \right) + \frac{\partial \hat{x}^\gamma}{\partial x^c} \frac{\partial x^b}{\partial \hat{x}^\beta} \frac{\partial x^a}{\partial \hat{x}^\alpha} \left(\frac{1}{2} g^{cd} \left(\frac{\partial g_{bd}}{\partial x_a} \right) \right) \\ \frac{1}{2}\hat{g}^{\gamma\delta} \left(\frac{\partial \hat{g}_{\alpha\beta}}{\partial \hat{x}_\delta} \right) &= \frac{1}{2} \left(\frac{\partial \hat{x}^\beta}{\partial x^b} \frac{\partial^2 x^b}{\partial \hat{x}_\alpha \partial \hat{x}^\beta} + \frac{\partial \hat{x}^\delta}{\partial x^d} \frac{\partial^2 x^d}{\partial \hat{x}_\alpha \partial \hat{x}^\delta} \right) + \frac{\partial \hat{x}^\gamma}{\partial x^c} \frac{\partial x^b}{\partial \hat{x}^\beta} \frac{\partial x^a}{\partial \hat{x}^\alpha} \left(\frac{1}{2} g^{cd} \left(\frac{\partial g_{bd}}{\partial x_a} \right) \right) \end{aligned}$$

Nous obtenons (45) en additionnant les trois équations. □

10. Torsion et g -compatibilité d'une connexion

Soit $(U, V) \mapsto D_U(V)$ une connexion (affine); alors D est dite **sans torsion** lorsque pour tous champs de vecteurs U, V :

$$(46) \quad D_U(V) - D_V(U) = [U, V]$$

De plus, lorsque g est une métrique pseudo-riemannienne sur M , alors D est dite **métrique ou encore g -compatible**, si pour tout triplet (U, V, W) de champs de vecteurs sur M , l'identité suivante est satisfaite, soit:

$$(47) \quad U(g(V, W)) = g(D_U(V), W) + g(V, D_U(W))$$

La propriété de g -compatibilité en (47) possède une belle symétrie formelle: plus concrètement, nous verrons (c.f. § 12) qu'elle est la condition assurant la propriété d'isométrie du transport parallèle. À ce stade, un premier indice de l'importance de la g -compatibilité est donné par le Théorème d'existence et d'unicité de Koszul (c.f. Théorème 11.2).

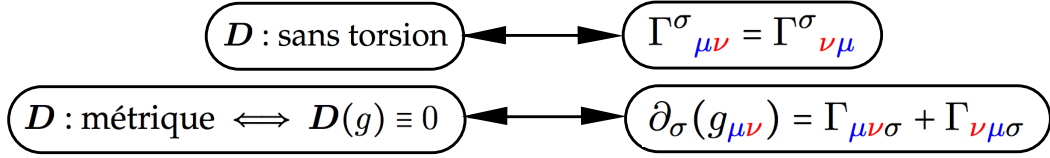


FIGURE 3. Traduction de l'absence de torsion et de la g -compatibilité d'une connexion D sur les coefficients $\Gamma^\sigma_{\mu\nu}$ et $\Gamma_{\mu\nu\sigma}$.

Proposition 10.1. Soit M une variété de dimension n munie d'une connexion D de coefficient $\Gamma^\sigma_{\mu\nu}$; alors (i) : D est sans torsion si et seulement si les coefficients $\Gamma^\sigma_{\mu\nu}$ sont symétriques, i.e. :

$$\Gamma^\sigma_{\mu\nu} = \Gamma^\sigma_{\nu\mu}$$

(ii) : si de plus g est une structure riemannienne sur M , alors D est métrique si et seulement si

$$\partial_\sigma(g_{\mu\nu}) = \Gamma_{\mu\nu\sigma} + \Gamma_{\nu\mu\sigma}$$

(où par définition nous avons posé $\Gamma_{\mu\nu\sigma} := g_{\mu\epsilon}\Gamma^\epsilon_{\nu\sigma}$).

Preuve. (i) : Étant donné U et V deux champs de vecteurs sur M , il vient :

$$\begin{aligned} U^\mu(D_\mu(V^\nu)) - V^\mu(D_\mu(U^\nu)) &= U^\mu(\partial_\mu(V^\nu) + V^\epsilon\Gamma^\nu_{\epsilon\mu}) - V^\mu(\partial_\mu(U^\nu) + U^\epsilon\Gamma^\nu_{\epsilon\mu}) \\ &= U^\mu\partial_\mu(V^\nu) - V^\mu\partial_\mu(U^\nu) + U^\mu V^\epsilon\Gamma^\nu_{\epsilon\mu} - V^\mu U^\epsilon\Gamma^\nu_{\epsilon\mu} \\ &= [U, V]^\nu + U^\mu V^\epsilon(\Gamma^\nu_{\epsilon\mu} - \Gamma^\nu_{\mu\epsilon}) \end{aligned}$$

(ii) : Étant donné U, V et W trois champs de vecteurs sur M , il vient :

$$\begin{aligned} &g(\mathbf{D}_U(V), W) + g(V, \mathbf{D}_U(W)) \\ &= U^\sigma(\partial_\sigma(V^\mu) + V^\epsilon\Gamma^\mu_{\epsilon\sigma})g_{\mu\nu}W^\nu + U^\sigma(\partial_\sigma(W^\mu) + W^\epsilon\Gamma^\mu_{\epsilon\sigma})g_{\mu\nu}V^\nu \\ &= U^\sigma(\partial_\sigma(V^\mu)g_{\mu\nu}W^\nu + V^\epsilon\Gamma^\mu_{\epsilon\sigma}g_{\mu\nu}W^\nu + V^\nu\Gamma^\mu_{\epsilon\sigma}g_{\mu\nu}W^\epsilon + V^\nu g_{\mu\nu}\partial_\sigma(W^\mu)) \\ &= U^\sigma(\partial_\sigma(V^\mu)g_{\mu\nu}W^\nu + V^\mu\Gamma^\epsilon_{\mu\sigma}g_{\epsilon\nu}W^\nu + V^\epsilon\Gamma^\mu_{\nu\sigma}g_{\mu\epsilon}W^\nu + V^\nu g_{\mu\nu}\partial_\sigma(W^\mu)) \\ &= U^\sigma(\partial_\sigma(V^\mu)g_{\mu\nu}W^\nu + V^\mu\Gamma^\epsilon_{\mu\sigma}g_{\epsilon\nu}W^\nu + V^\mu\Gamma^\epsilon_{\nu\sigma}g_{\epsilon\mu}W^\nu + V^\nu g_{\mu\nu}\partial_\sigma(W^\mu)) \\ &= U^\sigma(\partial_\sigma(V^\mu)g_{\mu\nu}W^\nu + V^\mu(g_{\nu\epsilon}\Gamma^\epsilon_{\mu\sigma} + g_{\mu\epsilon}\Gamma^\epsilon_{\nu\sigma})W^\nu + V^\nu g_{\mu\nu}\partial_\sigma(W^\mu)) \\ &= U^\sigma(\partial_\sigma(V^\mu)g_{\mu\nu}W^\nu + V^\mu\partial_\sigma(g_{\mu\nu})W^\nu + V^\nu g_{\mu\nu}\partial_\sigma(W^\mu)) + V^\mu W^\nu(\Gamma_{\nu\mu\sigma} + \Gamma_{\mu\nu\sigma} - \partial_\sigma(g_{\mu\nu})) \\ &= U^\sigma\partial_\sigma(V^\mu g_{\mu\nu}W^\nu) + V^\mu W^\nu(\Gamma_{\nu\mu\sigma} + \Gamma_{\mu\nu\sigma} - \partial_\sigma(g_{\mu\nu})) \end{aligned}$$

Cependant, nous avons aussi

$$U(g(V, W)) = U(V^\mu g_{\mu\nu}W^\nu) = U^\sigma\partial_\sigma(V^\mu g_{\mu\nu}W^\nu)$$

ce qui donne finalement

$$U(g(V, W)) - g(\mathbf{D}_U(V), W) + g(V, \mathbf{D}_U(W)) = V^\mu W^\nu(\Gamma_{\nu\mu\sigma} + \Gamma_{\mu\nu\sigma} - \partial_\sigma(g_{\mu\nu}))$$

□

11. Théorème d'existence et d'unicité de Koszul

Il existe une présentation algébrique très simple de la connexion de Levi-Civita due à Koszul [Kos50]. Le point de départ est le fait que **l'ensemble des connexions sur M (s'il n'est pas vide!) est une sous-espace affine de l'espace vectoriel des applications \mathbb{R} -bilinéaire sur $\mathfrak{T}_0^1(TM)$** . En effet, soit $B(TM) := \coprod_{X \in M} B(T_X M)$, où $B(T_X M)$ l'espace des applications \mathbb{R} -bilinéaires de $T_X M$. Une **forme de connexion** sur M est par définition une section $\Theta : X \mapsto \Theta|_X$ de $B(TM)$ qui est C^∞ : **l'ensemble des formes de connexion est un espace vectoriel** identifié à l'espace des applications $C^\infty(M)$ -bilinéaires sur $\mathfrak{T}_0^1(TM)$: ainsi, si Θ est une forme de connexion alors, pour tout $f \in C^\infty(M)$,

$$\Theta(fV, W) = f\Theta(V, W) = \Theta(V, fW)$$

et par suite (règle de Leibniz) **une forme de connexion n'est pas une connexion**.

Proposition 11.1. *L'ensemble des connexions sur M est un espace affine dont la base vectorielle est l'espace vectoriel des formes de connexion sur M .*

Preuve. Soient $D^{(1)}$ et $D^{(2)}$ deux connexions affines : pour $V, W \in \mathfrak{T}_0^1(TM)$ et $f \in C^\infty(M)$ on a d'une part $D_{fV}^{(1)}(W) - D_{fV}^{(2)}(W) = f(D_V^{(1)}(W) - D_V^{(2)}(W))$ et que d'autre part,

$$(D_V^{(1)} - D_V^{(2)})(fW) = (V(f)W + fD_V^{(1)}W) - (V(f)W + fD_V^{(2)}W) = f(D_V^{(1)}W - D_V^{(2)}W)$$

Cela montre que la différence de deux connexions est une forme de connexion. \square

Théorème 11.2 (Koszul). *Toute variété pseudo-riemannienne (M, g) est associée à une unique connexion ∇ – appelée **connexion de Levi-Civita** – et caractérisée par le fait que (1) : ∇ est sans torsion et (2) : g est parallèle (i.e. ∇ est métrique). De plus, la connexion de Levi-Civita est entièrement déterminée par la **formule de Koszul** affirmant que pour tout $V, W \in \mathfrak{T}_0^1(TM)$*

$$(48) \quad 2g(\nabla_U(V), W) = U(g(V, W)) + V(g(W, U)) - W(g(U, V)) + g([U, V], W) + g([W, U], V) - g([V, W], U)$$

Preuve. Montrons d'abord l'**unicité** de la connexion de Levi-Civita. Comme ∇ est g -compatible, nous avons pour tous champs de vecteurs U, V et W ,

$$\begin{aligned} U(g(V, W)) &= g(\nabla_U(V), W) + g(V, \nabla_U(W)) \\ V(g(W, U)) &= g(\nabla_V(W), U) + g(W, \nabla_V(U)) \\ W(g(U, V)) &= g(\nabla_W U, V) + g(U, \nabla_W(V)) \end{aligned}$$

En retranchant la troisième identité à la somme des deux premières, il vient :

$$(49) \quad U(g(V, W)) + V(g(W, U)) - W(g(U, V)) = g(\nabla_U(V) + \nabla_V(U), W) + g(\nabla_U(W) - \nabla_W(U), V) + g(\nabla_V(W) - \nabla_W(V), U)$$

D'autre part, la torsion de ∇ étant nulle l'identité (49) se simplifie :

$$2g(\nabla_U(V), W) = U(g(V, W)) + V(g(W, U)) - W(g(U, V)) + g([U, V], W) + g([W, U], V) - g([V, W], U)$$

où on reconnaît la formule de Koszul en (48). Par non dégénérescence de g , la connexion ∇ est uniquement déterminée de sorte que si U et V sont deux champs de vecteurs sur M , le champ $\nabla_U(V)$ est défini comme l'unique champ de vecteurs sur M vérifiant (48). Pour l'existence, montrons que la formule de Koszul définissant $\nabla_U(V)$ fait de ∇ une connexion g -compatible et sans torsion. D'abord, pour toutes fonctions f :

$$\begin{aligned} 2g(\nabla_U(fV), W) &= U(g(fV, W)) + (fV)(g(W, U)) - W(g(U, fV)) + \\ &\quad g([U, fV], W) + g([W, U], fV) - g([fV, W], U) \\ &= U(f) \cdot g(V, W) - W(f) \cdot g(U, V) + U(f) \cdot g(V, W) + \\ &\quad W(f) \cdot g(U, V) + f \cdot 2g(\nabla_U(V), W) \\ &= 2g(U(f) \cdot V + f \nabla_U(V), W) \\ 2g(\nabla_{fU}V, W) &= (fU)(g(V, W)) + V(g(W, fU)) - W(g(fU, V)) + \\ &\quad g([fU, V], W) + g([W, fU], V) - g([V, W], fU) \\ &= V(f) \cdot g(W, U) - W(f) \cdot g(U, V) - V(f) \cdot g(U, W) + \\ &\quad W(f) \cdot g(U, V) + f \cdot 2g(\nabla_U(V), W) \\ &= 2g(f \nabla_U(V), W) \end{aligned}$$

D'autre part, ∇ est bien sans torsion :

$$\begin{aligned} &2g(\nabla_U(V) - \nabla_V(U), W) \\ &= U(g(V, W)) + V(g(W, U)) - W(g(U, V)) + g([U, V], W) + g([W, U], V) - g([V, W], U) \\ &\quad - V(g(U, W)) - U(g(W, V)) + W(g(V, U)) - g([V, U], W) - g([W, V], U) + g([U, W], V) \\ &= 2g([U, V], W) \end{aligned}$$

Enfin, ∇ est bien g -compatible :

$$\begin{aligned} &2g(\nabla_U(V), W) + 2g(V, \nabla_U(W)) \\ &= U(g(V, W)) + V(g(W, U)) - W(g(U, V)) + g([U, V], W) + g([W, U], V) - g([V, W], U) \\ &\quad + U(g(W, V)) + W(g(V, U)) - V(g(U, W)) + g([U, W], V) + g([V, U], W) - g([W, V], U) \\ &= 2U(g(V, W)) \end{aligned}$$

□

Remarque 11.3. Nous avons initialement défini ∇ comme la connexion dont les coefficients Γ^k_{ij} coïncident avec les symboles de Christoffel \mathbf{C}^k_{ij} (c.f. Définition 9.1 et Théorème 9.3). Nous pouvons cependant retrouver directement ce résultat par la formule de Koszul donnée en (48) : en

effet, si $(\partial_1, \dots, \partial_n)$ est une base holonôme et si $(U, V, W) = (\partial_i, \partial_j, \partial_k)$ alors,

$$2g(\nabla_{\partial_i}(\partial_j), \partial_k) = \partial_i(g(\partial_j, \partial_k)) + \partial_j(g(\partial_k, \partial_i)) - \partial_k(g(\partial_i, \partial_j)) + g([\partial_i, \partial_j], \partial_k) + g([\partial_k, \partial_i], \partial_j) - g([\partial_j, \partial_k], \partial_i)$$

et comme $[\partial_i, \partial_j] = 0$, nous retrouvons

$$g_{i\epsilon} \Gamma^{\epsilon}_{jk} = \frac{1}{2} \partial_i(g(\partial_j, \partial_k)) + \partial_j(g(\partial_k, \partial_i)) - \partial_k(g(\partial_i, \partial_j)) = \frac{1}{2} (\partial_i(g_{jk}) + \partial_j(g_{ki}) - \partial_k(g_{ij}))$$

12. Transport parallèle

Soit M une variété de dimension n plongée dans l'espace \mathbb{R}^m . Étant donné $\gamma : [a; b] \rightarrow M$ un chemin sur M , nous dirons que l'application $V : t \mapsto V(t)$ de $[a; b]$ dans \mathbb{R}^m est un champ de vecteur le long de γ si elle est C^∞ et si $V(t) \in T_{\gamma(t)}M$, pour tout $a \leq t \leq b$. Toute connexion D sur M est associée à un transport parallèle le long des chemins sur M . Avant de définir cette notion, nous commençons par un cas particulier, pour lequel $\gamma(t)$ coïncide avec la position $\phi_U^t(X)$ de la courbe intégrale d'un champ de vecteur U au temps t et de position initiale X et où $V(t)$ coïncide avec la valeur $V|_{\gamma(t)}$ prise par un deuxième champ de vecteur V sur M . En posant

$$(50) \quad \frac{DV(t)}{dt} := D_U(V)|_{\gamma(t)}$$

nous définissons un nouveau champ de vecteur le long de γ et qui – en un sens – est dérivé de $V(t)$. Pour aller un peu plus loin, remarquons qu'en se plaçant dans n'importe quel système $\mathbf{x} = (x^1, \dots, x^n)$ de coordonnées locales, les propriétés de la connexion D (dont nous notons $\Gamma^\sigma_{\mu\nu}$ les coefficients), nous donnent :

$$D_U(V) = (U^\mu \partial_\mu(V^\sigma) + U^\mu V^\nu \Gamma^\sigma_{\mu\nu}) \partial_\sigma$$

Maintenant nous notons $(\gamma^1(t), \dots, \gamma^n(t))$ les composantes cartésiennes de $\mathbf{x}(\gamma(t))$; alors du fait que $\gamma(t) = \phi_U^t(X)$, nous avons $U^\mu(\gamma(t)) = \dot{\gamma}^\mu(t)$. Par suite, en notant (abusivement) $V(t) := V(\gamma(t))$, i.e. $V^\nu(t) := V^\nu(\gamma(t))$, il vient²³ :

$$D_U(V)|_{\gamma(t)} = (\dot{\gamma}^\mu(t) \partial_\mu(V^\sigma) + \dot{\gamma}^\mu(t) V^\nu(t) \Gamma^\sigma_{\mu\nu}(\gamma(t))) \partial_\sigma$$

soit encore, d'après la notation en (50)

$$\frac{DV(t)}{dt} = (\dot{V}^\sigma(t) + \dot{\gamma}^\mu(t) V^\nu(t) \Gamma^\sigma_{\mu\nu}(\gamma(t))) \partial_\sigma$$

Cette introduction nous permet de poser la définition suivante :

Définition 12.1. Soit D une connexion sur M ; alors, tout champ de vecteurs $t \mapsto V(t)$ le long d'un chemin $\gamma : [a; b] \rightarrow M$ est associé à un champ de vecteurs $t \mapsto DV(t)/dt$ le long de γ t.q pour tout système $\mathbf{x} = (x^1, \dots, x^n)$ de coordonnées locales (dont le domaine contient $\gamma(t)$)

$$\frac{DV(t)}{dt} = (\dot{V}^\sigma(t) + \dot{\gamma}^\mu(t) V^\nu(t) \Gamma^\sigma_{\mu\nu}(\gamma(t))) \partial_\sigma|_{\gamma(t)}$$

23. Pour tout fonction $t \mapsto \alpha(t)$ de la variable réelle et à valeurs dans \mathbb{R}^m , nous utilisons la notation de Newton $\dot{\alpha}(t) := d\alpha/dt$ (indépendamment du nom de la variable réelle).

et où $(\gamma^1(t), \dots, \gamma^n(t))$ sont les coordonnées cartésiennes de $\alpha(\gamma(t))$. Par définition $t \mapsto V(t)$ est transporté parallèlement le long de γ ssi $DV(t)/dt = 0$, pour tout $a \leq t \leq b$; en coordonnées :

$$\frac{DV(t)}{dt} = 0 \iff (\forall \sigma) \quad \dot{V}^\sigma(t) + \dot{\gamma}^\mu(t)V^\nu(t)\Gamma^\sigma_{\mu\nu}(\gamma(t)) = 0$$

L'ensemble $\mathfrak{F}_0^1(T_\gamma M)$ des champs de vecteurs le long d'un chemin $t \mapsto \gamma(t)$ sur M est un espace vectoriel : il est facile de vérifier que la dérivation covariante $D/dt : \mathfrak{F}_0^1(T_\gamma M) \rightarrow \mathfrak{F}_0^1(T_\gamma M)$ le long de γ associée à la connexion \mathbf{D} est une application \mathbb{R} -linéaire qui vérifie la règle de Leibnitz, i.e. pour toute application lisse $t \mapsto \alpha(t) \in \mathbb{R}$ et pour tout champ de vecteur V_t le long de γ :

$$\frac{D(f(t)V(t))}{dt} = \dot{\alpha}(t)V(t) + \alpha(t)\frac{DV(t)}{dt}$$

Le champ du vecteur vitesse $t \mapsto d\gamma/dt$ est un champ de vecteur le long de γ alors que l'accélération $t \mapsto d^2\gamma/dt^2$ (calculée dans l'espace ambiant \mathbb{R}^m) n'en est pas nécessairement un. D'autre part, le champ du vecteur vitesse $t \mapsto d\gamma/dt$ le long de γ est transporté parallèlement par γ si et seulement si

$$(51) \quad (\forall \sigma) \quad \ddot{\gamma}^\sigma(t) + \dot{\gamma}^\mu(t)\dot{\gamma}^\nu(t)\Gamma^\sigma_{\mu\nu}(\gamma(t)) = 0$$

Nous verrons (c.f. infra § 13) comment le système d'équations différentielles (51) permet de décrire les géodésiques sur M , lorsque M est munie d'une métrique pseudo-riemannienne et associée à la connexion de Levi-Civita. La proposition suivante est une conséquence directe du Théorème de Cauchy-Lipchitz.

Proposition 12.2. Soit $\gamma :]a; b[\rightarrow M$ un germe de chemin en un point $X \in M$ (i.e. $a < 0 < b$ et $\gamma(0) = X$) et de support est dans le domaine d'un système de coordonnées ; alors, pour tout $a < t < b$ il existe une application linéaire $\Pi_\gamma^t : T_X(0) \rightarrow T_{\gamma(t)}M$ (appelé *propagateur*) t.q. pour tout vecteur $v \in T_X M$ le champ de vecteur $t \mapsto \Pi_\gamma^t(v)$ est transporté parallèlement le long de γ , en ce sens que c'est la solution maximale du problème de Cauchy

$$(52) \quad (\forall \sigma) \quad \dot{V}^\sigma(t) + \dot{\gamma}^\mu(t)V^\nu(t)\Gamma^\sigma_{\mu\nu}(\gamma(t)) = 0 \quad \text{et} \quad V^\sigma(0) = v^\sigma$$

Rappelons que lorsque M est munie d'une métrique pseudo-riemannienne g , la connexion \mathbf{D} est dite g -compatible (ou encore métrique) si pour tout $U, V, W \in \mathfrak{F}_0^1(TM)$

$$(53) \quad U(g(V, W)) = g(\mathbf{D}_U(V), W) + g(V, \mathbf{D}_U(W))$$

La proposition suivante montre que dans le cas où \mathbf{D} est une connexion métrique (et en particulier pour la connexion de Levi-Civita), chaque propagateur Π_γ^t , le long d'un germe de chemin γ en un point X de M est une isométrie de $T_X M$ sur $T_{\gamma(t)}M$.

Proposition 12.3. Soit (M, g) une variété pseudo-riemannienne et soit D/dt la dérivation covariante associée à une connexion \mathbf{D} supposée métrique : alors, étant données $t \mapsto V_t$ et $t \mapsto W_t$ deux champs de vecteurs quelconques le long d'un chemin γ sur M :

$$(54) \quad \frac{d}{dt}g(V_t, W_t) = g\left(\frac{DV_t}{dt}, W_t\right) + g\left(V_t, \frac{DW_t}{dt}\right)$$

En particulier si V et W sont transportés parallèlement par γ , alors

$$\frac{d}{dt}g(V_t, W_t) = 0$$

Preuve. Soit D une connexion métrique sur (M, g) et soit $t \mapsto V_t$ et $t \mapsto W_t$ deux champs de vecteurs le long d'un chemin γ . Pour démontrer (54), il suffit de considérer d'une part que les champs de vecteurs $t \mapsto V_t$ et $t \mapsto W_t$ le long de γ sont des restrictions de deux champs de vecteurs sur M – soient V et W – de sorte que $V_t = V|_{\gamma(t)}$ et $W_t = W|_{\gamma(t)}$ et d'autre part, que γ est une courbe intégrale d'un troisième champ de vecteur U , dont le support est contenu dans le domaine d'une système de coordonnées locales (dans ce cas on a donc $\dot{\gamma}(t) = U|_{\gamma(t)}$). Alors, d'après la définition de la compatibilité métrique de D ,

$$(55) \quad U|_X(g(V, W)) = g(D_U(V)|_X, W|_X) + g(V|_X, D_U(W)|_X)$$

Or, pour $X = \gamma(t)$ il vient $D_U(V)|_{\gamma(t)} = DV/dt$ et $D_U(W)|_{\gamma(t)} = DW/dt$, et (55) devient :

$$U|_{\gamma(t)}(g(V, W)) = g\left(\frac{DV_t}{dt}, W_t\right) + g\left(V_t, \frac{DW_t}{dt}\right)$$

Par définition de U relativement au chemin γ , nous avons $U|_{\gamma(t)}(f) = d/dt(f \circ \gamma)$, pour tout champ scalaire f sur M : nous obtenons (54) en prenant $f(X) := g(U|_X, V|_X)$. \square

13. Géodésiques inertielles : Théorème de Levi-Civita

Dans ce paragraphe, nous supposons que la variété pseudo-riemannienne (M, g) (de dimension n) est munie de la connexion de Levi-Civita $(U, V) \mapsto \nabla_U(V)$. Rappelons que les coefficients de ∇ coïncident avec les symboles de Christoffel i.e.

$$\nabla_{\partial_\mu}(\partial_\nu) = \mathbf{C}^\sigma_{\mu\nu} \partial_\sigma \quad \text{où} \quad \mathbf{C}^\sigma_{\mu\nu} = \frac{1}{2} g^{\sigma\epsilon} (\partial_\mu(g_{\epsilon\nu}) + \partial_\nu(g_{\mu\epsilon}) - \partial_\epsilon(g_{\mu\nu}))$$

Ainsi, ∇ est une connexion qui est sans torsion et g -compatible (et c'est la seule connexion ayant ces deux propriétés). Enfin, U et V étant deux champs de vecteurs sur M , le champ de vecteurs $\nabla_U(V)$ s'écrit (en coordonnées)

$$(56) \quad \nabla_U(V) = U^\mu \nabla_\mu(V^\nu) \partial_\nu \quad \text{avec} \quad \nabla_\mu(V^\nu) := \partial_\mu(V^\nu) + V^\epsilon \mathbf{C}^\nu_{\epsilon\mu}$$

où on reconnaît les composantes des dérivées partielles covariantes $\nabla_\mu(V^\nu)$ associée à V .

Nous allons maintenant voir comment les géodésiques sur M définies par un principe de moindre action, peuvent être caractérisées comme solutions d'une équation différentielle du second ordre s'interprétant en terme de transport parallèle. Pour cela, considérons d'abord que \mathcal{V} est un ouvert de \mathbb{R}^n et soit

$$(q^1, \dots, q^n, \dot{q}^1, \dots, \dot{q}^n) = (q, \dot{q}) \mapsto L(q, \dot{q})$$

un lagrangien²⁴ défini sur $\mathcal{V} \times \mathbb{R}^n$. Le chemin $\mathbf{q} = (q^1, \dots, q^n) : [a; b] \rightarrow \mathcal{V}$ sera appelé une L -géodésique s'il satisfait le système des équations d'Euler-Lagrange i.e.

$$(\forall \sigma) \quad \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^\sigma}(\mathbf{q}(t), \dot{\mathbf{q}}(t)) = \frac{\partial L}{\partial q^\sigma}(\mathbf{q}(t), \dot{\mathbf{q}}(t))$$

24. Ici on se restreint à un lagrangien stationnaire en ce sens qu'il ne dépend pas d'un paramètre temporel; on exige seulement que les dérivées partielles $\partial L / \partial q^\mu$ et $\partial L / \partial \dot{q}^\mu$ existent et soient continues.

Théorème 13.1 (Levi-Civita). Soit (M, g) une variété pseudo-riemannienne de dimension n munie de la connexion de Levi-Civita et soit $\mathbf{x} : \mathcal{U} \rightarrow \mathbb{R}^n$ une carte. Nous considérons le **lagrangien-inertiel** défini pour tout $(q, \dot{q}) \in \mathbf{x}(\mathcal{U}) \times \mathbb{R}^n$ en posant

$$L(q, \dot{q}) = \frac{1}{2} \dot{q}^\mu \dot{q}^\nu G_{\mu\nu}(q)$$

où $G_{\mu\nu}(q) = g_{\mu\nu}(X)$ dès que $q = \mathbf{x}(X)$. Si $\gamma^1, \dots, \gamma^n$ sont les \mathbf{x} -composantes d'un germe de chemin $\gamma : [a; b] \rightarrow \mathcal{U}$ en un point $X = \gamma(0) \in \mathcal{U}$ – en ce sens que $\mathbf{x}(\gamma(t)) = (\gamma^1(t), \dots, \gamma^n(t))$ – alors $t \mapsto (\gamma^1(t), \dots, \gamma^n(t))$ est une L -géodésique (dans $\mathbf{x}(\mathcal{U})$) ssi le champ du vecteur vitesse de $t \mapsto \gamma(t)$ est transporté parallèlement suivant γ , ce qui se traduit par le système différentiel

$$(\forall k) \quad \ddot{\gamma}^\sigma(t) + \dot{\gamma}^\mu(t) \dot{\gamma}^\nu(t) \mathbf{C}^\sigma_{\mu\nu}(\gamma(t)) = 0$$

On dit que γ est une géodésique issue de X . De plus (Théorème de Cauchy-Lipschitz) γ est déterminées par le conditions initiales $\gamma(0) = X$ et $\dot{\gamma}(0) = \dot{\gamma}^\mu(0) \partial_\mu|_X =: v \in T_X M$ et on écrira

$$(57) \quad \gamma(t) = \exp_X(tv)$$

Preuve. Nous commençons par calculer les dérivées partielles de L . D'une part,

$$\frac{\partial L}{\partial q^\sigma} = \frac{\partial}{\partial q^\sigma} \left(\frac{1}{2} \dot{q}^\mu \dot{q}^\nu G_{\mu\nu}(q) \right) = \frac{1}{2} \dot{q}^\mu \dot{q}^\nu \frac{\partial G_{\mu\nu}}{\partial q^\sigma}$$

et d'autre part (en utilisant la symétrie $G_{\mu\sigma} = G_{\sigma\mu}$)

$$\frac{\partial L}{\partial \dot{q}^\sigma} = \frac{\partial}{\partial \dot{q}^\sigma} \left(\frac{1}{2} \dot{q}^\mu \dot{q}^\nu G_{\mu\nu}(q) \right) = \dot{q}^\mu G_{\mu\sigma}(q)$$

Les équations d'Euler-Lagrange pour $t \mapsto (\gamma^1(t), \dots, \gamma^n(t))$ ($\in \mathbf{x}(\mathcal{U})$) s'écrivent alors :

$$(\forall \sigma) \quad \ddot{\gamma}^\mu(t) g_{\mu\sigma}(\gamma(t)) + \dot{\gamma}^\mu(t) \dot{\gamma}^\nu(t) \frac{\partial g_{\mu\sigma}}{\partial \mathbf{x}^\nu}(\gamma(t)) = \frac{1}{2} \dot{\gamma}^\mu(t) \dot{\gamma}^\nu(t) \frac{\partial g_{\mu\nu}}{\partial \mathbf{x}^\sigma}(\gamma(t))$$

Or, il est évident par la convention sommatoire d'Einstein que

$$\dot{\gamma}^\mu(t) \dot{\gamma}^\nu(t) \frac{\partial g_{\mu\sigma}}{\partial \mathbf{x}^\nu}(\gamma(t)) = \dot{\gamma}^\nu(t) \dot{\gamma}^\mu(t) \frac{\partial g_{\nu\sigma}}{\partial \mathbf{x}^\mu}(\gamma(t))$$

d'où l'identité

$$\dot{\gamma}^\mu(t) \dot{\gamma}^\nu(t) \frac{\partial g_{\mu\sigma}}{\partial \mathbf{x}^\nu}(\gamma(t)) = \frac{1}{2} \dot{\gamma}^\mu(t) \dot{\gamma}^\nu(t) \left(\frac{\partial g_{\mu\sigma}}{\partial \mathbf{x}^\nu}(\gamma(t)) + \frac{\partial g_{\nu\sigma}}{\partial \mathbf{x}^\mu}(\gamma(t)) \right)$$

et nous obtenons une forme équivalente des équations d'Euler-Lagrange, soient :

$$(\forall \sigma) \quad \ddot{\gamma}^\mu(t) g_{\sigma\nu}(\gamma(t)) + \frac{1}{2} \dot{\gamma}^\mu(t) \dot{\gamma}^\nu(t) \left(\frac{\partial g_{\mu\sigma}}{\partial \mathbf{x}^\nu}(\gamma(t)) + \frac{\partial g_{\nu\sigma}}{\partial \mathbf{x}^\mu}(\gamma(t)) - \frac{\partial g_{\mu\nu}}{\partial \mathbf{x}^\sigma}(\gamma(t)) \right) = 0$$

soient encore, par définition des symboles de Christoffel de première espèce :

$$(\forall \sigma) \quad \ddot{\gamma}^\mu(t) g_{\sigma\nu}(\gamma(t)) + \dot{\gamma}^\nu(t) \dot{\gamma}^\mu(t) \mathbf{C}^\sigma_{\mu\nu}(\gamma(t)) = 0$$

Comme $g^{\sigma\sigma} g_{\mu\sigma} = g^{\sigma\sigma} g_{\sigma\mu} = \delta^\sigma_\mu$ et $g^{\sigma\sigma} \mathbf{C}^\sigma_{\mu\nu} = \mathbf{C}^\sigma_{\mu\nu}$, nous retrouvons :

$$(\forall \sigma) \quad \ddot{\gamma}^\sigma(t) + \dot{\gamma}^\mu(t) \dot{\gamma}^\nu(t) \mathbf{C}^\sigma_{\mu\nu}(\gamma(t)) = 0$$

□

Remarque 13.2. Supposons (M, g) riemannienne, isométriquement plongée dans \mathbb{R}^m euclidien et munie de la connexion de Levi-Civita. Dire que $t \mapsto \gamma(t)$ est une géodésique sur M signifie que le vecteur vitesse $t \mapsto d\gamma/dt$ est transporté parallèlement le long de $t \mapsto \gamma(t)$. Par suite, la connexion de Levi-Civita étant en particulier métrique, nous en déduisons (Proposition 12.3) que la norme de $\dot{\gamma}(t)$ (i.e. la norme dans l'espace tangent aussi que dans l'espace ambiant) reste constante le long de γ et donc $0 = d/dt \langle \dot{\gamma}(t) | \dot{\gamma}(t) \rangle = 2 \langle \ddot{\gamma}(t) | \dot{\gamma}(t) \rangle$. Cela signifie que l'accélération de γ (calculée dans \mathbb{R}^m euclidien) est normale au vecteur vitesse.

14. Dérivation covariante des champs de tenseurs

Dans ce paragraphe, nous supposons que (M, g) est pseudo-riemannienne et munie d'une connexion D quelconque. En composante, la dérivée covariante d'un $(1, 0)$ -tenseur $V = V^\mu \partial_\mu$ est un tenseur de valence $(1, 1)$ noté $D(V)$ et dont les composantes sont

$$D(V)^\mu{}_\nu = D_\nu(V^\mu) = \partial_\nu(V^\mu) + V^\epsilon \Gamma^\mu{}_{\epsilon\nu}$$

La dérivaison covariante se généralise en une transformation D de l'ensemble $\mathfrak{T}(TM) := \coprod_{p,q} \mathfrak{T}_q^p(TM)$ de tous les tenseurs sur M déterminée par les axiomes (DC0 : 5) dans la Proposition 14.1 ci-dessous. Tous ces axiomes sont raisonnablement naturels, excepté (peut-être) l'axiome (DC3) qui fait intervenir la contraction des indices des champs de tenseurs. Rappelons (c.f. § 4) que lorsque E est un espace euclidien de dimension finie, la contraction d'indice $[^r_s]$ est définie sur le produit tensoriel $\otimes^{p,q} E$ (pour $p, q \geq 1$, $1 \leq r \leq p$ et $1 \leq s \leq q$) et à valeur dans $\otimes^{p-1, q-1} E$. En utilisant la métrique g sur M , cette application s'étend simplement aux champs de tenseurs en une application

$$[^r_s] : \mathfrak{T}_q^p(TM) \rightarrow \mathfrak{T}_{q-1}^{p-1}(TM)$$

où avec un abus de notations $[^r_s](F)|_X = [^r_s](F|_X)$. Nous verrons que (DC3) permet la définition inductive de D , de sorte que la dérivaison covariante d'une forme différentielle ξ s'écrit en composantes

$$D(\xi)_{\mu\nu} = D_\nu(\xi_\mu) = \partial_\nu(\xi_\mu) - \xi_\epsilon \Gamma^\epsilon{}_{\mu\nu}$$

Proposition 14.1. Étant donnée (M, g) est une variété pseudo-Riemannienne munie d'une connexion D quelconque, la dérivaison covariante des tenseurs (relativement à D) est l'unique application $D : \mathfrak{T}(TM) \rightarrow \mathfrak{T}(TM)$ qui vérifie les six axiomes suivantes :

(DC0) : pour tout $f \in \mathfrak{T}_0^0(TM) = \mathcal{C}^\infty(M)$, $D(f)$ est le tenseur $(0, 1)$ (appelé gradient covariant de f) dont les composantes sont

$$D(f)_\mu = D_\mu(f) = \partial_\mu(f)$$

(DC1) : pour tout $V \in \mathfrak{T}_0^1(TM)$ de composantes V^μ , la dérivée covariante $D(V)$ est le tenseur $(1, 1)$ de composantes

$$D(V)^\mu{}_\nu = D_\nu(V^\mu) = \partial_\nu(V^\mu) + V^\sigma \Gamma^\mu{}_{\sigma\nu}$$

(DC2) : la dérivée covariante d'un tenseur F de valence (p, q) est un tenseur $D(F)$ de valence $(p, q + 1)$ dont on écrit les composantes sous la forme (conventionnelle)

$$D(F)^{\mu_1 \dots \mu_p}{}_{\nu_1 \dots \nu_q \epsilon} =: D_\epsilon(F^{\mu_1 \dots \mu_p}{}_{\nu_1 \dots \nu_q}) = F^{\mu_1 \dots \mu_p}{}_{\nu_1 \dots \nu_q \epsilon}$$

(DC3) : la dérivation covariante D commute avec la contraction d'indice $[^r_s]$ en ce sens que

$$D([^r_s](F)) = [^r_s](D(F)) \quad (\text{lorsque cela a un sens})$$

(DC4) : pour $A, B \in \mathfrak{T}_q^p(TM)$

$$D(A + B) = D(A) + D(B)$$

(DC5) : D satisfait la règle de Leibnitz généralisée : pour $A \in \mathfrak{T}_q^p(TM)$ et $B \in \mathfrak{T}_s^r(TM)$

$$D(A \otimes B) = D(A) \otimes B + A \otimes D(B)$$

Par définition l'application

$$F^{\mu_1 \dots \mu_p}_{\nu_1 \dots \nu_q} \mapsto D_\epsilon(F^{\mu_1 \dots \mu_p}_{\nu_1 \dots \nu_q}) = F^{\mu_1 \dots \mu_p}_{\nu_1 \dots \nu_q; \epsilon}$$

est la dérivée partielle covariante D_ϵ étendue aux composantes des tenseurs.

Au lieu de donner une démonstration de la Proposition 14.1 (i.e. la vérification de la consistance des axiomes) nous préférons donner un aperçu de l'articulation de ces axiomes en effectuant le calcul de la dérivée covariante dans le cas (important) d'une forme différentielle ξ . Soit alors V un champ de vecteurs arbitrairement donné ; alors, le tenseur $\xi \otimes V$ de composante $\xi_\nu V^\mu$ se contracte d'une seule manière possible sous forme d'un tenseur $(0, 0)$ i.e. le champ scalaire $X \mapsto \xi|_X(V)$, que nous notons $\xi(V)$, et dont les composantes sont les $V^\nu \xi_\nu$. D'une part, d'après (DC0)

$$(58) \quad D_\sigma(\xi(V)) = \xi(V)^\nu_{;\sigma} = \partial_\sigma(\xi(V)^\nu_\nu) = \partial_\sigma(V^\nu \xi_\nu) = \partial_\sigma(V^\nu) \xi_\nu + V^\nu \partial_\sigma(\xi_\nu)$$

mais d'autre part, d'après (DC4 – 5)

$$D_\sigma(\xi \otimes V)^\mu_\nu = (\xi \otimes V)^\mu_{;\sigma} = \xi_{\nu;\sigma} V^\mu + \xi_\nu V^\mu_{;\sigma}$$

et du fait que $\xi(V) = [^1_1](\xi \otimes V)$ nous pouvons utiliser (DC3) et écrire

$$D_\sigma(\xi(V)) = [^1_1](\xi \otimes V)^\mu_{;\sigma} = \xi_{\nu;\sigma} V^\nu + \xi_\nu V^\nu_{;\sigma}$$

Finalement en exprimant $D_\sigma(V^\nu) = V^\nu_{;\sigma}$ par (DC1) il vient avec (58)

$$\partial_\sigma(V^\nu) \xi_\nu + V^\nu \partial_\sigma(\xi_\nu) = \xi_{\nu;\sigma} V^\nu + \xi_\nu (\partial_\sigma(V^\nu) + V^\epsilon \Gamma^\nu_{\epsilon\sigma})$$

soit encore,

$$\xi_{\nu;\sigma} V^\nu = V^\nu \partial_\sigma(\xi_\nu) - \xi_\nu V^\epsilon \Gamma^\nu_{\epsilon\sigma} = \partial_\sigma(\xi_\nu) V^\nu - \xi_\epsilon V^\nu \Gamma^\epsilon_{\nu\sigma}$$

Cette dernière identité étant valable pour tout V , nous en concluons que

$$D_\sigma(\xi_\nu) = \partial_\sigma(\xi_\nu) - \xi_\epsilon \Gamma^\epsilon_{\nu\sigma}$$

Plus généralement, les axiomes de la dérivation covariante des tenseurs permettent de déduire (de manière analogue aux calculs donnant l'expression de $D_\sigma(\xi_\nu)$) la formule de dérivation covariante des tenseurs en composantes, i.e. en donnant l'expression des [dérivations partielles covariantes](#) étendues aux composantes de valence quelconques.

Proposition 14.2. Soient $F^{\mu_1 \dots \mu_p}_{\nu_1 \dots \nu_q}$ les composantes d'un tenseur $F \in \mathfrak{T}_q^p(TM)$; alors la dérivée covariante DF est un tenseur de $\mathfrak{T}_{q+1}^p(TM)$ dont les composantes sont²⁵ :

$$\begin{aligned} D(F)^{\mu_1 \dots \mu_p}_{\nu_1 \dots \nu_q \epsilon} &= D_\epsilon(F^{\mu_1 \dots \mu_p}_{\nu_1 \dots \nu_q}) \\ &= \partial_\epsilon(F^{\mu_1 \dots \mu_p}_{\nu_1 \dots \nu_q}) + \sum_{r=1}^p F^{\mu_1 \dots (\sigma \leftarrow \mu_r) \dots \mu_p}_{\nu_1 \dots \nu_q} \Gamma^{\mu_r}_{\sigma \epsilon} \\ &\quad - \sum_{s=1}^q F^{\mu_1 \dots \mu_p}_{\nu_1 \dots (\sigma \leftarrow \nu_s) \dots \nu_q} \Gamma^{\sigma}_{\nu_s \epsilon} \end{aligned}$$

et par définition, pour tout $U = U^i \partial_i$, nous posons :

$$D_U(F) := U^\epsilon D_\epsilon(F^{\mu_1 \dots \mu_p}_{\nu_1 \dots \nu_q}) \frac{\partial}{\partial x^{\mu_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{\mu_p}} \otimes dx^{\nu_1} \otimes \dots \otimes dx^{\nu_q}$$

La proposition suivante est une conséquence de la formule des dérivées covariantes des tenseurs en composantes, pour une connexion quelconque (c.f. Proposition 14.2).

Théorème 14.3. Si D est une connexion sur une variété pseudo-riemannienne (M, g) alors,

$$(i) \iff (ii) \iff (iii)$$

où nous avons les propositions suivantes, soient (i) : D est métrique ; (ii) : $D(g) \equiv 0$ et (iii) : les dérivations partielles covariantes de D commutent avec l'abaissement/relèvement d'indices.

Preuve. (i) \iff (ii) : Étant donnée D une connexion arbitraire de coefficients $\Gamma^\sigma_{\mu\nu}$, la formule de la dérivée covariante des tenseurs en composantes (c.f. Proposition 14.2) appliquée au tenseur métrique nous permet d'écrire :

$$D_\lambda(g_{\mu\nu}) = \partial_\lambda(g_{\mu\nu}) - g_{\nu\epsilon} \Gamma^\epsilon_{\lambda\mu} - g_{\mu\epsilon} \Gamma^\epsilon_{\lambda\nu} = \partial_\lambda(g_{\mu\nu}) - (\Gamma_{\nu\lambda\mu} + \Gamma_{\mu\lambda\nu})$$

L'équivalence annoncée vient avec la partie (ii) de la Proposition 10.1.

(ii) \iff (iii) : Nous considérons le cas de l'abaissement d'indice, le cas du relèvement étant analogue (le fait que $D(g) = 0$ entraîne aussi que $D_\sigma(g^{\mu\nu}) = 0$). Si nous supposons que $D(g) \equiv 0$, alors pour tout tenseur $F \in \mathfrak{T}_q^p(TM)$ de x -composantes $F^{\mu_1 \dots \mu_p}_{\nu_1 \dots \nu_q}$ le calcul de la dérivée covariante du produit donne

$$\begin{aligned} D_\sigma(g_{\tau\mu_i} F^{\mu_1 \dots \mu_p}_{\nu_1 \dots \nu_q}) &= D_\sigma(g_{\tau\mu_i}) F^{\mu_1 \dots \mu_p}_{\nu_1 \dots \nu_q} + g_{\tau\mu_i} D_\sigma(F^{\mu_1 \dots \mu_p}_{\nu_1 \dots \nu_q}) \\ &= g_{\tau\mu_i} D_\sigma(F^{\mu_1 \dots \mu_p}_{\nu_1 \dots \nu_q}) \end{aligned}$$

la conclusion venant du fait que $D_\sigma(g_{\tau\mu_i}) = 0$. Réciproquement, si nous supposons la commutation des dérivées partielles covariantes D_σ avec l'abaissement des indices, alors nous pouvons écrire : $D_\sigma(g_{\mu\nu}) = g_{\mu\nu} D_\sigma(1) = 0$. □

15. Opérateurs différentiels

15.1. Gradient polaire classique. Une difficulté de l'analyse vectorielle classique porte sur l'expression des opérateurs différentiels (gradient, divergence, laplacien, ...) lors d'un changement de système de coordonnées. Considérons par exemple que \mathbb{R}^2 est rapporté

25. Nous notons $(\sigma \leftarrow k)$ le remplacement d'un indice k par l'indice ϵ .

à la base canonique $(\mathbf{e}_1, \mathbf{e}_2)$ et soit $\mathcal{U} =]0; +\infty[\times \mathbb{R}$. L'expression cartésienne du gradient (vecteur) $\text{grad}(f)$ d'un champ scalaire $(x, y) \mapsto f(x, y)$ défini sur \mathcal{U} s'écrit :

$$(59) \quad \text{grad}(f) = \frac{\partial f}{\partial x} \mathbf{e}_1 + \frac{\partial f}{\partial y} \mathbf{e}_2$$

Notons $\hat{f}(\rho, \theta) = f(\rho \cos(\theta), \rho \sin(\theta))$ l'expression de f dans les coordonnées polaires et $(\mathbf{e}_\rho, \mathbf{e}_\theta)$ la base *mobile* polaire au point $(\rho \cos(\theta), \rho \sin(\theta))$ (i.e. $\mathbf{e}_\rho := \cos(\theta)\mathbf{e}_1 + \sin(\theta)\mathbf{e}_2$ et $\mathbf{e}_\theta := -\sin(\theta)\mathbf{e}_1 + \cos(\theta)\mathbf{e}_2$) : un rapide examen montre que

$$\frac{\partial \hat{f}}{\partial \rho} \mathbf{e}_\rho + \frac{\partial \hat{f}}{\partial \theta} \mathbf{e}_\theta$$

ne correspond pas au gradient en (59). Pour obtenir l'expression du gradient polaire, il nous faut partir d'une définition précise. Nous utilisons l'*heuristique* du calcul infinitésimal de Leibnitz : si $X \in \mathcal{U}$ et si dX est un vecteur déplacement infinitésimal,

$$f(X + dX) - f(X) = \langle \text{grad}(f) | dX \rangle$$

En coordonnées cartésiennes, $X = (x, y)$ et $dX = dx\mathbf{e}_1 + dy\mathbf{e}_2$: nous retrouvons l'expression du gradient de f en (59), par l'approximation différentielle de f en (x, y) qui s'écrit

$$f(x + dx, y + dy) - f(x, y) = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$$

Pour trouver l'expression polaire du gradient de f , nous considérons le déplacement polaire infinitésimal, soit $dX = d\rho \mathbf{e}_\rho + \rho d\theta \mathbf{e}_\theta$: ainsi nous pouvons écrire

$$\begin{aligned} \langle \text{grad}(f) | dX \rangle &= \hat{f}(\rho + d\rho, \theta + d\theta) - \hat{f}(\rho, \theta) \\ &= \frac{\partial \hat{f}}{\partial \rho} d\rho + \frac{\partial \hat{f}}{\partial \theta} d\theta \\ &= \left\langle \frac{\partial \hat{f}}{\partial \rho} \mathbf{e}_\rho + \frac{1}{\rho} \frac{\partial \hat{f}}{\partial \theta} \mathbf{e}_\theta \middle| d\rho \mathbf{e}_\rho + \rho d\theta \mathbf{e}_\theta \right\rangle = \left\langle \frac{\partial \hat{f}}{\partial \rho} \mathbf{e}_\rho + \frac{1}{\rho} \frac{\partial \hat{f}}{\partial \theta} \mathbf{e}_\theta \middle| dX \right\rangle \end{aligned}$$

et finalement, l'accroissement dX étant quelconque (sic)

$$(60) \quad \text{grad}(f) = \frac{\partial \hat{f}}{\partial \rho} \mathbf{e}_\rho + \frac{1}{\rho} \frac{\partial \hat{f}}{\partial \theta} \mathbf{e}_\theta$$

15.2. Gradient pseudo-riemannien. Retrouvons la définition du gradient vecteur pour un champ scalaire $f : M \rightarrow \mathbb{R}$ défini sur une variété pseudo-riemannienne (M, g) munie de la connexion de Levi-Civita ∇ . Ici la situation est simplifiée par l'axiome (DC0) de la dérivation covariante des tenseurs : dans un système $\mathbf{x} = (x^1, \dots, x^n)$ de coordonnées locales, les dérivées partielles $\partial_i(f) = \partial f / \partial x^i$ coïncident avec les dérivées partielles covariantes $\nabla_i(f)$ et sont les composantes d'un tenseur $(0, 1)$, i.e. le gradient covariant²⁶ $\nabla(f)$. Le gradient vecteur $\text{grad}(f)$ (ou gradient contravariant) s'obtient par le relèvement de l'indice covariant de $\nabla_i(f)$: les composantes de $\text{grad}(f)$ sont donc les $g^{ij} \partial_j(f)$, soit :

$$(61) \quad \text{grad}(f) = g^{ij} \partial_j(f) \partial_i$$

26. Attention aux confusions possibles avec les notations d'analyse vectorielle classique !

Revenons à l'ouvert de $\mathcal{U} =]0; +\infty[\times \mathbb{R}$ de \mathbb{R}^2 : dans l'analyse vectorielle classique \mathcal{U} est considéré (implicitement comme une variété riemannienne munie de la métrique euclidienne canonique. Dans le cas considéré ici, les *coordonnées polaires* sont déterminées par le système de coordonnées $\mathcal{U} \ni (x, y) \mapsto (\rho(x, y), \theta(x, y))$ où

$$\rho(x, y) := (x^2 + y^2)^{1/2} \quad \text{et} \quad \theta(x, y) := \text{Arcsin}(y/x)$$

Alors la base holonôme polaire s'exprime en fonction de la base mobile polaire $(\mathbf{e}_\rho, \mathbf{e}_\theta)$ avec $\partial_\rho = \mathbf{e}_\rho$ et $\partial_\theta = \rho \mathbf{e}_\theta$: les coefficients polaires de la métrique (plate) sont alors

$$\begin{pmatrix} \langle \partial_\rho | \partial_\rho \rangle & \langle \partial_\rho | \partial_\theta \rangle \\ \langle \partial_\theta | \partial_\rho \rangle & \langle \partial_\theta | \partial_\theta \rangle \end{pmatrix} = \begin{pmatrix} g_{\rho\rho} & g_{\rho\theta} \\ g_{\rho\theta} & g_{\theta\theta} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \rho^2 \end{pmatrix} \quad \text{i.e.} \quad \begin{pmatrix} g^{\rho\rho} & g^{\rho\theta} \\ g^{\rho\theta} & g^{\theta\theta} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1/\rho^2 \end{pmatrix}$$

Si $f(x, y)$ est la valeur prise par l'application $f : \mathcal{U} \rightarrow \mathbb{R}$ au point de coordonnées cartésiennes (x, y) alors la valeur correspondante en coordonnées polaires est $\hat{f}(\rho, \theta) = f(\rho \cos(\theta), \rho \sin(\theta))$: avec la définition tensorielle en (61) nous retrouvons

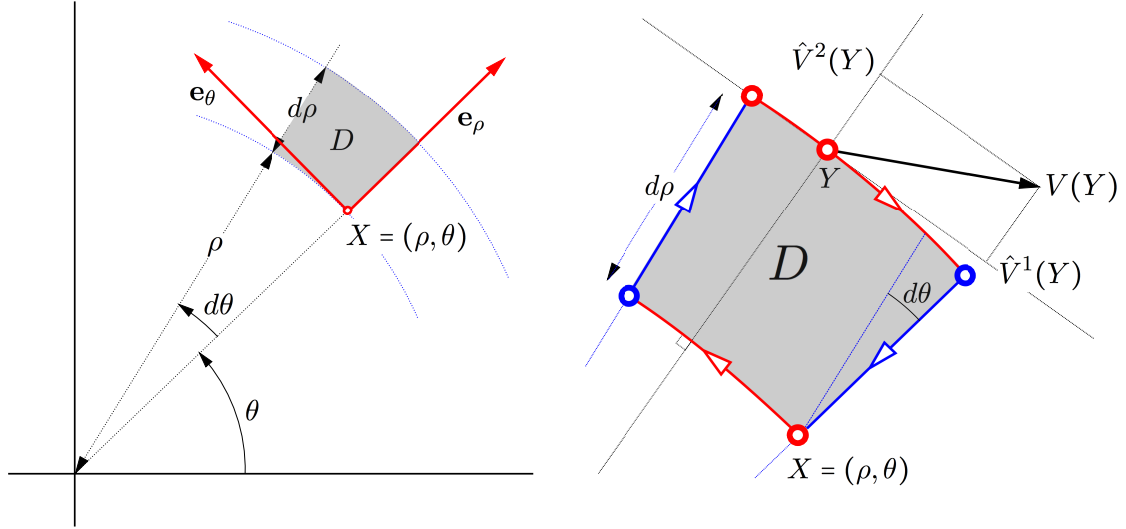
$$\text{grad}(f) = \frac{\partial \hat{f}}{\partial \rho} \mathbf{e}_\rho + \frac{1}{\rho} \frac{\partial \hat{f}}{\partial \theta} \mathbf{e}_\theta$$

15.3. Divergence polaire classique. Le calcul de la divergence d'un champ de vecteur est un outils essentiel de l'analyse vectorielle. La divergence est – entre autre – liée à l'équation des ondes ainsi qu'à l'équation de la chaleur ; elle permet aussi de caractériser l'écoule des fluide incompressibles. Soit $X \mapsto V(X)$ un champ de vecteur défini pour tout point $X \in \mathcal{U} =]0; +\infty[\times \mathbb{R}$. Écrivons (V^1, V^2) les composantes cartésiennes de V ; si (ρ, θ) (resp. (x, y)) sont les coordonnées polaires (resp. cartésiennes) de X , alors, $V^i(x, y) = V^i \circ \Phi(\rho, \theta)$, où $\Phi(\rho, \theta) = (\rho \cos(\theta), \rho \sin(\theta))$. Par suite, en partant de la définition de la divergence de V en coordonnées cartésiennes, nous obtenons successivement :

$$\begin{aligned} \text{div}(V) &= \left\langle \text{grad}(V^1) \middle| \mathbf{e}_1 \right\rangle + \left\langle \text{grad}(V^2) \middle| \mathbf{e}_2 \right\rangle \\ &= \left\langle \left(\frac{\partial V^1 \circ \Phi}{\partial \rho} \right) \mathbf{e}_\rho + \frac{1}{\rho} \left(\frac{\partial V^1 \circ \Phi}{\partial \theta} \right) \mathbf{e}_\theta \middle| \cos(\theta) \mathbf{e}_\rho - \sin(\theta) \mathbf{e}_\theta \right\rangle + \\ &\quad \left\langle \left(\frac{\partial V^2 \circ \Phi}{\partial \rho} \right) \mathbf{e}_\rho + \frac{1}{\rho} \left(\frac{\partial V^2 \circ \Phi}{\partial \theta} \right) \mathbf{e}_\theta \middle| \sin(\theta) \mathbf{e}_\rho - \cos(\theta) \mathbf{e}_\theta \right\rangle \\ &= \frac{\partial V^1 \circ \Phi}{\partial \rho} \cos(\theta) - \frac{1}{\rho} \frac{\partial V^1 \circ \Phi}{\partial \theta} \sin(\theta) + \frac{\partial V^2 \circ \Phi}{\partial \rho} \sin(\theta) + \frac{1}{\rho} \frac{\partial V^1 \circ \Phi}{\partial \theta} \cos(\theta) \end{aligned}$$

Maintenant, si (\hat{V}^1, \hat{V}^2) sont les composantes polaires de V , alors

$$V^1 \circ \Phi = \cos(\theta) \hat{V}^1 - \sin(\theta) \hat{V}^2 \quad \text{et} \quad V^2 \circ \Phi = \sin(\theta) \hat{V}^1 + \cos(\theta) \hat{V}^2$$

FIGURE 4. *Domaine polaire infinitésimal.*

de sorte que

$$\begin{aligned}
 \operatorname{div}(V) &= \frac{\partial}{\partial \rho} \left(\cos(\theta) \hat{V}^1 - \sin(\theta) \hat{V}^2 \right) \cos(\theta) - \frac{1}{\rho} \frac{\partial}{\partial \theta} \left(\cos(\theta) \hat{V}^1 - \sin(\theta) \hat{V}^2 \right) \sin(\theta) + \\
 &\quad \frac{\partial}{\partial \rho} \left(\sin(\theta) \hat{V}^1 + \cos(\theta) \hat{V}^2 \right) \sin(\theta) + \frac{1}{\rho} \frac{\partial}{\partial \theta} \left(\sin(\theta) \hat{V}^1 + \cos(\theta) \hat{V}^2 \right) \cos(\theta) \\
 &= \cos^2(\theta) \frac{\partial \hat{V}^1}{\partial \rho} + \frac{1}{\rho} \left(\sin(\theta) \hat{V}^1 - \cos(\theta) \frac{\partial \hat{V}^1}{\partial \theta} + \cos(\theta) \hat{V}^2 + \sin(\theta) \frac{\partial \hat{V}^2}{\partial \theta} \right) \sin(\theta) + \\
 &\quad \sin^2(\theta) \frac{\partial \hat{V}^1}{\partial \rho} + \frac{1}{\rho} \left(\cos(\theta) \hat{V}^1 + \sin(\theta) \frac{\partial \hat{V}^1}{\partial \theta} - \sin(\theta) \hat{V}^2 + \cos(\theta) \frac{\partial \hat{V}^2}{\partial \theta} \right) \cos(\theta) \\
 &= \frac{\partial \hat{V}^1}{\partial \rho} + \frac{1}{\rho} \hat{V}^1 + \frac{1}{\rho} \frac{\partial \hat{V}^2}{\partial \theta} \cos(\theta)
 \end{aligned}$$

et finalement

$$\operatorname{div}(V) = \frac{1}{\rho} \frac{\partial(\rho \hat{V}^1)}{\partial \rho} + \frac{1}{\rho} \frac{\partial \hat{V}^2}{\partial \theta}$$

Un raisonnement analogue en polaire ne donne pas le même résultat, puisque

$$\begin{aligned}
 \langle \operatorname{grad}(\hat{V}^1) | \mathbf{e}_\rho \rangle + \langle \operatorname{grad}(\hat{V}^2) | \mathbf{e}_\theta \rangle &= \left\langle \frac{\partial \hat{V}^1}{\partial \rho} \mathbf{e}_\rho + \frac{1}{\rho} \frac{\partial \hat{V}^1}{\partial \theta} \mathbf{e}_\theta \middle| \mathbf{e}_\rho \right\rangle + \left\langle \frac{\partial \hat{V}^2}{\partial \rho} \mathbf{e}_\rho + \frac{1}{\rho} \frac{\partial \hat{V}^2}{\partial \theta} \mathbf{e}_\theta \middle| \mathbf{e}_\theta \right\rangle \\
 &= \frac{\partial \hat{V}^1}{\partial \rho} + \frac{1}{\rho} \frac{\partial \hat{V}^2}{\partial \theta}.
 \end{aligned}$$

Cela est dû au fait que la base $(\mathbf{e}_\rho, \mathbf{e}_\theta)$ dépend des coordonnées locales polaires (ρ, θ) . D'une manière purement géométrique (et dans l'approximation différentielle) la divergence de V en (ρ, θ) est constante dans le domaine polaire infinitésimal en (ρ, θ) (d'aire $\rho d\rho d\theta$). Par suite, en intégrant le flux sortant de V à travers le bord de ce domaine (forme

locale du **théorème du flux divergence** : voir Figure 4), il vient :

$$\begin{aligned} \operatorname{div}(V) &= \frac{1}{\rho d\rho d\theta} \left(\hat{V}^1(\rho + d\rho, \theta)(\rho + d\rho)d\theta - \hat{V}^1(\rho, \theta)\rho d\theta + \hat{V}^2(\rho, \theta + d\theta)d\rho - \hat{V}^2(\rho, \theta)d\rho \right) \\ &= \frac{1}{\rho d\rho d\theta} \left(\left(\hat{V}^1 + \frac{\partial \hat{V}^1}{\partial \rho} d\rho \right) (\rho d\theta + d\rho d\theta) - \hat{V}^1 \rho d\theta + \left(\hat{V}^2 + \frac{\partial \hat{V}^2}{\partial \theta} d\theta \right) d\rho - \hat{V}^2 d\rho \right) \\ &= \frac{1}{\rho} \left(\hat{V}^1 + \frac{\partial \hat{V}^1}{\partial \rho} \right) + \frac{1}{\rho} \frac{\partial \hat{V}^2}{\partial \theta} + \frac{d\rho}{\rho} \frac{\partial \hat{V}^1}{\partial \rho} \end{aligned}$$

et nous retrouvons

$$\operatorname{div}(V) = \frac{1}{\rho} \left(\hat{V}^1 + \frac{\partial \hat{V}^1}{\partial \rho} \right) + \frac{1}{\rho} \frac{\partial \hat{V}^2}{\partial \theta} = \frac{1}{\rho} \frac{\partial(\rho \hat{V}^1)}{\partial \rho} + \frac{1}{\rho} \frac{\partial \hat{V}^2}{\partial \theta}$$

Le calcul de la divergence d'un champ de vecteur dans le cadre pseudo-riemannien ne peut se traiter élémentairement (comme nous l'avons fait au § 15.2 pour le gradient). Plus généralement, la définition des opérateurs différentiels sur une variété dépend essentiellement de la dérivation covariante.

15.4. Divergence et laplacien pseudo-riemanniens. Lorsque (M, g) est pseudo-riemannienne et munie de la connexion de Levi-Civita ∇ , la divergence peut être définie pour les champs de tenseurs. Étant donné $F \in \mathfrak{T}_q^{p+1}(TM)$ de \mathbf{x} -composantes $F^{i_0 \dots i_p}_{j_1 \dots j_q}$, la **divergence de F** est le tenseur $\operatorname{div}(F) \in \mathfrak{T}_q^p(TM)$ de \mathbf{x} -composantes

$$\operatorname{div}(F)^{i_1 \dots i_p}_{j_1 \dots j_q} = \nabla(F)^{i_1 \dots i_p \epsilon}_{j_1 \dots j_q \epsilon} = \nabla_{\epsilon}(F^{i_1 \dots i_p \epsilon}_{j_1 \dots j_q}) = F^{i_1 \dots i_p \epsilon}_{j_1 \dots j_q \epsilon}$$

Proposition 15.1. *La divergence d'un champ de vecteur V est un champ scalaire, soit*

$$(62) \quad \operatorname{div}(V) = \nabla_i V^i = \partial_i(V^i) + V^j \mathbf{C}^i_{ij} = \frac{1}{\sqrt{|g|}} \partial_i (V^i \sqrt{|g|})$$

où g désigne le déterminant de la matrice (g_{ij}) .

La preuve de la Proposition 15.1 est basée sur la contraction de l'exposant contravariant des symboles de Christoffel²⁷.

Lemme 15.2. *Si $g := \det(g_{ij})$ alors la contraction des symboles de Christoffel s'écrit :*

$$(63) \quad \mathbf{C}^i_{ik} = \frac{1}{2} g^{ij} \partial_k(g_{ij}) = \frac{1}{2g} \partial_k(g) = \frac{1}{\sqrt{|g|}} \partial_k \sqrt{|g|} = \partial_k (\log \sqrt{|g|})$$

Preuve. Soit \mathbf{x} un système de coordonnées fixé. Partant de l'expression des symboles de Christoffel et profitant de la symétrie du tenseur métrique il vient :

$$\begin{aligned} \mathbf{C}^i_{ik} &= \frac{1}{2} (g^{ij} \partial_i(g_{jk}) + g^{ij} \partial_k(g_{ij}) - g^{ij} \partial_j(g_{ik})) \\ (i \leftrightarrow j) &= \frac{1}{2} (g^{ij} \partial_i(g_{jk}) + g^{ij} \partial_k(g_{ij}) - g^{ji} \partial_i(g_{jk})) = \frac{1}{2} g^{ij} \partial_k(g_{ij}) \end{aligned}$$

Soient X et Y deux points du domaine de \mathbf{x} : pour simplifier nous supposons que $\mathbf{x}(X) = 0$ et que $\mathbf{x}(Y) = (y^1, \dots, y^n)$. Nous allons utiliser le fait que

$$\det(\delta^i_j + h^i_j) \approx 1 + h^i_i + \max_{ij} \{|h^i_j|\} \varepsilon(h) \quad (\text{avec } \varepsilon(h) \rightarrow 0 \text{ and } h \rightarrow 0)$$

27. Il n'y a qu'une contraction possible du fait de la symétrie $\mathbf{C}^k_{ij} = \mathbf{C}^k_{ji}$.

Alors, nous pouvons écrire

$$\begin{aligned}
 \mathbf{g}(Y) = \det(g_{ij}(Y)) &= \det\left(g_{ij}(X) + y^k \partial_k|_X(g_{ij}) + \dots\right) \\
 &= \det\left(g_{i\epsilon}\left(\delta^\epsilon_j + g^{\epsilon i} y^k \partial_k|_X(g_{ij})\right) + \dots\right) \\
 &= \mathbf{g}(X) \det\left(\delta^\epsilon_j + g^{\epsilon i} y^k \partial_k|_X(g_{ij}) + \dots\right) \\
 &= \mathbf{g}(X) \left(1 + g^{\epsilon i} y^k \partial_k|_X(g_{i\epsilon}) + \dots\right) = \mathbf{g}(X) + y^k \left(\mathbf{g}(X) g^{ji} \partial_k|_X(g_{ij})\right) + \dots
 \end{aligned}$$

ce qui entraîne que $\partial_k|_X(\mathbf{g}) = \mathbf{g}(X) g^{ij} \partial_k|_X(g_{ij})$, soit encore que

$$\frac{\partial_k(\mathbf{g})}{2\mathbf{g}} = \frac{1}{2} g^{ij} \partial_k(g_{ij})$$

□

Preuve de la Proposition 15.1. L'expression du symbole de Christoffel contracté \mathbf{C}^i_{ij} (c.f. Lemma 15.2) permet d'écrire la divergence sous la forme :

$$\operatorname{div}(V) = \partial_i(V^i) + \frac{V^i}{\sqrt{|\mathbf{g}|}} \partial_i(\sqrt{|\mathbf{g}|})$$

Or la dérivation partielle usuelle (des fonctions numériques) donne $1/\phi \partial_i(\psi\phi) = \partial_i(\psi) + (\psi/\phi)\partial_i(\phi)$ de sorte que pour $\phi = \sqrt{|\mathbf{g}|}$ et $\psi = V^i$ il vient :

$$\frac{1}{\sqrt{|\mathbf{g}|}} \partial_i(V^i \sqrt{|\mathbf{g}|}) = \partial_i(V^i) + \frac{V^i}{\sqrt{|\mathbf{g}|}} \partial_i(\sqrt{|\mathbf{g}|})$$

□

Pour terminer ce paragraphe, notons que le laplacien $\Delta(f)$ d'un champ scalaire $f : M \rightarrow \mathbb{R}$ est défini comme la divergence du gradient vecteur de f i.e.

$$\Delta(f) = \operatorname{div}(\operatorname{grad}(f)) = \nabla_i(g^{ij} \partial_j(f))$$

Proposition 15.3. En coordonnées locales, l'expression du laplacien de $f : M \rightarrow \mathbb{R}$ est

$$\Delta(f) = g^{ij} \left(\partial_{ij}(f) - \partial_k(f) \mathbf{C}^k_{ij} \right) = \frac{1}{\sqrt{|\mathbf{g}|}} \partial_i \left(\sqrt{|\mathbf{g}|} g^{ij} \partial_j(f) \right)$$

(Le laplacien ainsi défini en pseudo-riemannien est l'*opérateur de Laplace-Beltrami*.)

Preuve. D'une part, nous pouvons utiliser le fait que $\nabla_i g^{ij} = 0$ pour écrire

$$\Delta(f) = \nabla_i(g^{ij} \partial_j(f)) = g^{ij} \nabla_i(\partial_j(f)) = g^{ij} \left(\partial_{ij}(f) - \partial_k(f) \mathbf{C}^k_{ij} \right)$$

D'autre part, en utilisant l'expression de la divergence en (62),

$$\Delta(f) = \operatorname{div}(\operatorname{grad}(f)) = \frac{1}{\sqrt{|\mathbf{g}|}} \partial_i \left(\sqrt{|\mathbf{g}|} g^{ij} \partial_j(f) \right)$$

□

16. Notes

Le livre de Bishop et Goldberg [BG68] donne une présentation complète du calcul tensoriel. On pourra aussi consulter, avec un grand intérêt, le travail de Bourguignon [Bou92] sur l'évolution des notions de transport parallèle et de connexion entre 1830 et 1930. De ce papier, nous tirons quelques points de repères : la notion de connexion apparaît chez Christoffel [Chr69], le calcul différentiel absolu étant introduit par Ricci en 1888 [Ric88]. Le papier de référence sur le calcul différentiel absolu est cependant celui que publient Ricci et son élève Levi-Civita en 1900 [RTLC00] : il peut être considéré comme « *l'acte de naissance et le manuel fondamental du Calcul Tensoriel* ». Un premier cycle de développement des idées géométriques initiées par Gauss et Riemann s'achève avec le papier de Levi-Civita de 1917 [LC17] et l'introduction du transport parallèle : ainsi, (Théorème de Levi-Civita), un chemin satisfait l'équation des géodésiques inertielles si et seulement si il transporte parallèlement son vecteur vitesse. Il faut préciser ici que la forme du lagrangien inertiel impose que le transport parallèle du Théorème de Levi-Civita, soit nécessairement effectué relativement à la connexion de Levi-Civita. La présentation de Koszul [Kos50], place l'espace affine des connexions comme paradigme : la connexion de Levi-Civita est alors caractérisée par le Théorème d'Unicité : ce point de vue offre une multitude d'avantages, mais on y perd le lien avec le formalisme lagrangien.

17. Appendices : Connexion de Levi-Civita d'une nappe de \mathbb{R}^3

17.1. Une connexion. Soit \mathcal{O} un ouvert de \mathbb{R}^m : comme d'habitude, \mathbb{R}^m est supposé munie de sa structure euclidienne usuelle et de sa base canonique $(\mathbf{e}_1, \dots, \mathbf{e}_m)$. Étant donnés deux champs de vecteurs V et W définis sur \mathcal{O} et de composantes cartésiennes respectives V^i et W^i , on voudrait définir le champ de vecteurs $D_V(W)$, de sorte qu'en tout $X = (x^1, \dots, x^n)$ de \mathcal{O} , le vecteur $D_V(W)|_X$ soit la dérivée en $t = 0$ du champ $t \mapsto W|_{\phi_V^t(X)}$ le long de la courbe intégrale $t \mapsto \phi_V^t(X)$ de V issue de X . Plus précisément, par approximations différentielles

$$\begin{aligned} \frac{1}{t} \left(W|_{\phi_V^t(X)} - W|_{\phi_V^0(X)} \right) &= \frac{1}{t} \left(W^k(\phi_V^t(X)) - W^k(X) \right) \mathbf{e}_k \\ &\approx \frac{1}{t} \left(W^k(X + tV^\epsilon(X)\mathbf{e}_\epsilon) - W^k(X) \right) \mathbf{e}_k \approx V^\epsilon(X) \frac{\partial W^k}{\partial x^\epsilon} \mathbf{e}_k \end{aligned}$$

Par définition, nous posons

$$(64) \quad D_V(W) = V^\epsilon(X) \frac{\partial W^k}{\partial x^\epsilon} \mathbf{e}_k$$

L'application $(V, W) \mapsto D_V(W)$ est bilinéaire et de plus, pour tout $f \in \mathcal{C}^\infty(\mathcal{U})$,

$$(65) \quad D_{fV}(W) = fD_V(W) \quad \text{et} \quad D_V(fW) = V(f)W + fD_V(W)$$

En d'autres termes, $(V, W) \mapsto D_V(W)$ est une connexion sur \mathcal{O} . Pour tout $X \in \mathcal{O}$, soit $\text{Sp}_X(V, W)$ le sous-espace vectoriel de \mathbb{R}^m engendré par $V|_X$ et $W|_X$ et notons $D_V^{\text{tan}}(W)|_X$ la projection orthogonale de $D_V(W)|_X$ sur $\text{Sp}_X(V, W)$. Les propriétés de D en (65) se décomposent suivant la somme orthogonale $\text{Sp}_X(V, W) \boxplus \text{Sp}_X(V, W)^\perp$ (dans \mathbb{R}^m) en tout point X de M , nous obtenons

$$(66) \quad D_{fV}^{\text{tan}}(W) = fD_V^{\text{tan}}(W) \quad \text{et} \quad D_V^{\text{tan}}(fW) = V(f)W + fD_V^{\text{tan}}(W)$$

ce qui signifie que $(V, W) \mapsto D_V^{\text{tan}}(W)$ est aussi une connexion de Koszul sur \mathcal{O}

Supposons maintenant que (M, g) est une variété riemannienne de dimension n plongée isométriquement dans \mathbb{R}^m et soit V et W deux champs de vecteurs sur M . Alors V et W peuvent respectivement être considérés comme les restrictions locales de deux champs de vecteurs \hat{V} et \hat{W} définis sur un voisinage ouvert – disons \mathcal{O} – de M dans \mathbb{R}^m (\hat{V} et \hat{W} n'étant pas en général uniques). Alors, pour tout $X \in M$ le vecteur $D_V^{\text{tan}}(W)|_X := D_{\hat{V}}^{\text{tan}}(\hat{W})|_X$ appartient à $T_X M$: de plus, la valeur $D_V^{\text{tan}}(W)|_X$ ne dépend pas des extensions \hat{V} et \hat{W} et $X \mapsto D_V^{\text{tan}}(W)|_X$ définit un champ de vecteur sur M . Il découle alors de (66) que l'application bilinéaire $(V, W) \mapsto D_V^{\text{tan}}(W)$ est une connexion de Koszul sur M . Il n'est pas difficile de montrer que cette connexion coïncide avec la connexion de Levi-Civita de (M, g) , en ce sens qu'elle est sans torsion et g -compatible (nous n'utiliserons pas cette dernière remarque dans la suite de cette appendice).

17.2. Le cas d'une nappe. Notons (x^1, x^2, x^3) les coordonnées cartésiennes de \mathbb{R}^3 et soit $(x^1, x^2) \mapsto f(x^1, x^2)$ une application numérique définie sur un ouvert \mathcal{V} de \mathbb{R}^2 . Nous nous intéressons à la surface $\Sigma := \{(x^1, x^2, f(x^1, x^2)) ; (x^1, x^2) \in \mathcal{V}\}$ munie de sa structure riemannienne g obtenue par le plongement isométrique dans \mathbb{R}^3 . Nous notons $(\Sigma, \mathbf{x} = (x^1, x^2))$ le système de coordonnées trivial t.q. $\mathbf{x}(x^1, x^2, f(x^1, x^2)) = (x^1, x^2)$, les vecteurs holonômes ∂_1 et ∂_2 de ce système de coordonnées s'expriment simplement dans la base canonique de \mathbb{R}^3 avec :

$$(67) \quad \partial_1 = \mathbf{e}_1 + \left(\frac{\partial f}{\partial x^1} \right) \mathbf{e}_3 \quad \text{et} \quad \partial_2 = \mathbf{e}_2 + \left(\frac{\partial f}{\partial x^2} \right) \mathbf{e}_3$$

où $\mathbf{f} = f \circ \mathbf{x} = f(x^1, x^2)$. Nous appliquons la construction décrite au paragraphe précédent pour définir une connexion. Pour cela notons que tout champ de vecteur V défini sur M est associé à un champ de vecteur \hat{V} sur $\mathcal{V} \times \mathbb{R}$ dont la valeur prise en (x^1, x^2, x^3) coïncide avec celle de V prise en $(x^1, x^2, f(x^1, x^2))$. Maintenant, étant donnés $\hat{V} = \hat{V}^\nu \mathbf{e}_\nu$ et $\hat{W} = \hat{W}^\mu \mathbf{e}_\mu$ deux champs de vecteurs sur $\mathcal{V} \times \mathbb{R}$, nous savons la définition en (64) que

$$D_{\hat{V}}(\hat{W}) = \hat{V}^\nu \left(\frac{\partial \hat{W}^\mu}{\partial x^\nu} \right) \mathbf{e}_\mu$$

Dans le cas où $\hat{W} = \hat{\partial}_i$ nous avons $(\hat{W}^1, \hat{W}^2, \hat{W}^3) = (\delta^1_j, \delta^2_j, \partial f / \partial x^j)$ et

$$\begin{pmatrix} \frac{\partial \hat{W}^1}{\partial x^1} & \frac{\partial \hat{W}^1}{\partial x^2} & \frac{\partial \hat{W}^1}{\partial x^3} \\ \frac{\partial \hat{W}^2}{\partial x^1} & \frac{\partial \hat{W}^2}{\partial x^2} & \frac{\partial \hat{W}^2}{\partial x^3} \\ \frac{\partial \hat{W}^3}{\partial x^1} & \frac{\partial \hat{W}^3}{\partial x^2} & \frac{\partial \hat{W}^3}{\partial x^3} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \frac{\partial^2 f}{\partial x^1 \partial x^j} & \frac{\partial^2 f}{\partial x^2 \partial x^j} & 0 \end{pmatrix}$$

Il découle alors de (64) que

$$D_{\hat{\partial}_i}(\hat{\partial}_j) = \frac{\partial^2 f}{\partial x^i \partial x^j} \mathbf{e}_3$$

Nous prenons $\hat{\partial}_1 \times \hat{\partial}_2 / \|\hat{\partial}_1 \times \hat{\partial}_2\|$ comme vecteur normal au plan vectoriel engendré (ponctuellement) par $\hat{\partial}_1$ et $\hat{\partial}_2$, où plus précisément nous avons

$$(68) \quad \hat{\partial}_1 \times \hat{\partial}_2 = -\frac{\partial f}{\partial x^1} \mathbf{e}_1 - \frac{\partial f}{\partial x^2} \mathbf{e}_2 + \mathbf{e}_3 \quad \text{avec} \quad \|\hat{\partial}_1 \times \hat{\partial}_2\| = \sqrt{1 + \left(\frac{\partial f}{\partial x^1}\right)^2 + \left(\frac{\partial f}{\partial x^2}\right)^2}$$

Alors, par définition,

$$\begin{aligned} \mathbf{D}_{\hat{\partial}_i}^{\text{tan}}(\hat{\partial}_j) &= \mathbf{D}_{\hat{\partial}_i}(\hat{\partial}_j) - \left\langle \mathbf{D}_{\hat{\partial}_i}(\hat{\partial}_j) \left| \frac{\hat{\partial}_1 \times \hat{\partial}_2}{\|\hat{\partial}_1 \times \hat{\partial}_2\|} \right. \right\rangle \frac{\hat{\partial}_1 \times \hat{\partial}_2}{\|\hat{\partial}_1 \times \hat{\partial}_2\|} \\ &= \frac{\partial^2 f}{\partial x^i \partial x^j} \left[\mathbf{e}_3 + \frac{1}{\|\hat{\partial}_1 \times \hat{\partial}_2\|^2} \left(\frac{\partial f}{\partial x^1} \mathbf{e}_1 + \frac{\partial f}{\partial x^2} \mathbf{e}_2 - \mathbf{e}_3 \right) \right] \\ &= \frac{\partial^2 f}{\partial x^i \partial x^j} \left[\mathbf{e}_3 + \frac{1}{\|\hat{\partial}_1 \times \hat{\partial}_2\|^2} \left(\frac{\partial f}{\partial x^1} \left(\mathbf{e}_1 + \frac{\partial f}{\partial x^1} \mathbf{e}_3 - \frac{\partial f}{\partial x^1} \mathbf{e}_3 \right) \right. \right. \\ &\quad \left. \left. + \frac{\partial f}{\partial x^2} \left(\mathbf{e}_2 + \frac{\partial f}{\partial x^2} \mathbf{e}_3 - \frac{\partial f}{\partial x^2} \mathbf{e}_3 \right) - \mathbf{e}_3 \right) \right] \\ &= \frac{1}{\|\hat{\partial}_1 \times \hat{\partial}_2\|^2} \frac{\partial^2 f}{\partial x^i \partial x^j} \left[\|\hat{\partial}_1 \times \hat{\partial}_2\|^2 \mathbf{e}_3 + \frac{\partial f}{\partial x^1} \hat{\partial}_1 - \left(\frac{\partial f}{\partial x^1} \right)^2 \mathbf{e}_3 + \right. \\ &\quad \left. + \frac{\partial f}{\partial x^2} \hat{\partial}_2 - \left(\frac{\partial f}{\partial x^2} \right)^2 \mathbf{e}_3 - \mathbf{e}_3 \right] \\ &= \frac{1}{\|\hat{\partial}_1 \times \hat{\partial}_2\|^2} \frac{\partial^2 f}{\partial x^i \partial x^j} \left(\frac{\partial f}{\partial x^1} \hat{\partial}_1 + \frac{\partial f}{\partial x^2} \hat{\partial}_2 \right) \end{aligned}$$

Nous savons que la restriction de \mathbf{D}^{tan} à Σ – toujours notée (abusivement) \mathbf{D}^{tan} – définit une connexion sur Σ : alors, dans les coordonnées locales de \mathbf{x} , il vient $\mathbf{D}_{\hat{\partial}_i}^{\text{tan}}(\hat{\partial}_j) = \Gamma^1_{ij} \hat{\partial}_1 + \Gamma^2_{ij} \hat{\partial}_2$ où les coefficients de \mathbf{D}^{tan} s'écrivent (avec $\mathbf{f} = f(\mathbf{x}^1, \mathbf{x}^2)$)

$$(69) \quad \Gamma^k_{ij} = \frac{1}{\|\hat{\partial}_1 \times \hat{\partial}_2\|^2} \frac{\partial^2 \mathbf{f}}{\partial \mathbf{x}^i \partial \mathbf{x}^j} \frac{\partial \mathbf{f}}{\partial \mathbf{x}^k}$$

Proposition 17.1. Lorsque Σ est munie de la métrique riemannienne induite par la structure euclidienne de \mathbb{R}^3 , alors $(V, W) \mapsto \mathbf{D}_V^{\text{tan}}(W)$ coïncide avec la connexion de Levi-Civita \mathbf{D} .

Preuve. Nous allons montrer que les coefficients Γ^k_{ij} de \mathbf{D}^{tan} dans les coordonnées locales \mathbf{x} et donnés en (69), coïncident avec les symboles de Christoffel \mathbf{C}^k_{ij} . De l'expression en (67) des vecteurs holonômes ∂_1 et ∂_2 dans la base canonique de \mathbb{R}^3 , nous déduisons

$$\begin{pmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{pmatrix} = \begin{pmatrix} 1 + \left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}^1}\right)^2 & \frac{\partial \mathbf{f}}{\partial \mathbf{x}^1} \frac{\partial \mathbf{f}}{\partial \mathbf{x}^2} \\ \frac{\partial \mathbf{f}}{\partial \mathbf{x}^1} \frac{\partial \mathbf{f}}{\partial \mathbf{x}^2} & 1 + \left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}^2}\right)^2 \end{pmatrix}$$

ce qui donne (grâce à l'expression en (68) de $\|\hat{\partial}_1 \times \hat{\partial}_2\|$)

$$\begin{pmatrix} g^{11} & g^{12} \\ g^{21} & g^{22} \end{pmatrix} = \frac{1}{\|\hat{\partial}_1 \times \hat{\partial}_2\|^2} \begin{pmatrix} 1 + \left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}^2}\right)^2 & -\frac{\partial \mathbf{f}}{\partial \mathbf{x}^1} \frac{\partial \mathbf{f}}{\partial \mathbf{x}^2} \\ -\frac{\partial \mathbf{f}}{\partial \mathbf{x}^1} \frac{\partial \mathbf{f}}{\partial \mathbf{x}^2} & 1 + \left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}^1}\right)^2 \end{pmatrix}$$

Par suite, en utilisant l'expression des symboles de Christoffel en fonction des composantes de la connexion, nous pouvons écrire

$$\begin{aligned}
2\|\partial_1 \times \partial_2\|^2 \mathbb{C}^1_{ij} &= \|\partial_1 \times \partial_2\|^2 \left(g^{11} \left(\frac{\partial g_{1j}}{\partial x^i} + \frac{\partial g_{i1}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^1} \right) + g^{12} \left(\frac{\partial g_{2j}}{\partial x^i} + \frac{\partial g_{i2}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^2} \right) \right) \\
&= \left(1 + \left(\frac{\partial f}{\partial x^2} \right)^2 \right) \left(\frac{\partial^2 f}{\partial x^i \partial x^1} \frac{\partial f}{\partial x^j} + \frac{\partial f}{\partial x^1} \frac{\partial^2 f}{\partial x^i \partial x^j} + \frac{\partial^2 f}{\partial x^j \partial x^i} \frac{\partial f}{\partial x^1} + \frac{\partial f}{\partial x^i} \frac{\partial^2 f}{\partial x^j \partial x^1} \right. \\
&\quad \left. - \frac{\partial^2 f}{\partial x^1 \partial x^i} \frac{\partial f}{\partial x^j} - \frac{\partial f}{\partial x^i} \frac{\partial^2 f}{\partial x^1 \partial x^j} \right) \\
&\quad - \frac{\partial f}{\partial x^1} \frac{\partial f}{\partial x^2} \left(\frac{\partial^2 f}{\partial x^i \partial x^2} \frac{\partial f}{\partial x^j} + \frac{\partial f}{\partial x^2} \frac{\partial^2 f}{\partial x^i \partial x^j} + \frac{\partial^2 f}{\partial x^j \partial x^i} \frac{\partial f}{\partial x^2} + \frac{\partial f}{\partial x^i} \frac{\partial^2 f}{\partial x^j \partial x^2} \right. \\
&\quad \left. - \frac{\partial^2 f}{\partial x^2 \partial x^i} \frac{\partial f}{\partial x^j} - \frac{\partial f}{\partial x^i} \frac{\partial^2 f}{\partial x^2 \partial x^j} \right) \\
&= 2 \left(1 + \left(\frac{\partial f}{\partial x^2} \right)^2 \right) \frac{\partial f}{\partial x^1} \frac{\partial^2 f}{\partial x^i \partial x^j} - 2 \left(\frac{\partial f}{\partial x^2} \right)^2 \frac{\partial f}{\partial x^1} \frac{\partial^2 f}{\partial x^j \partial x^i} = 2 \frac{\partial f}{\partial x^1} \frac{\partial^2 f}{\partial x^i \partial x^j}
\end{aligned}$$

c'est-à-dire (et de même pour \mathbb{C}^2_{ij})

$$\mathbb{C}^1_{ij} = \frac{1}{\|\partial_1 \times \partial_2\|^2} \frac{\partial^2 f}{\partial x^j \partial x^i} \frac{\partial f}{\partial x^1} = \Gamma^1_{ij}$$

□

Remarque 17.2. Il est aussi possible de vérifier que D^{tan} coïncide avec la connexion de Levi-Civita, grâce au Théorème d'unicité de Koszul et en vérifiant directement par la Proposition 10.1 que D^{tan} est sans torsion et g -compatible. L'absence de torsion est facile à vérifier, puisqu'on voit directement sur (69) que $\Gamma^k_{ij} = \Gamma^k_{ji}$. Pour vérifier la g -compatibilité, il est possible de prouver directement (par des calculs plus fastidieux) que $\partial_k(g_{ij}) = \Gamma_{ijk} + \Gamma_{jik}$.

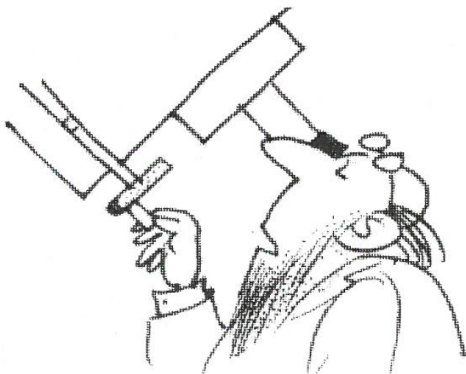
RÉFÉRENCES

- [BG68] R.L. Bishop and S.I. Goldberg. *Tensor Analysis on Manifolds*. Dover Books on Mathematics. Dover Publications, 1968.
- [Bou92] J.-P. Bourguignon. Transport parallèle et connexions en géométrie et en physique, in "1830-1930 : a century of geometry". Springer-Verlag, *Lect. Notes in Physics*, 402 :150–164, 1992.
- [Chr69] E.B. Christoffel. Ueber die Transformation der homogenen Differentialausdrücke zweiten Grades. *J. reine angew. Math.*, 70 :46–70, 1869.
- [Kos50] J. L. Koszul. Homologie et cohomologie des algèbres de Lie. *Bulletin de la Société Mathématique de France*, 78 :65–127, 1950.
- [LC17] T. Levi-Civita. Nozione di parallelismo in una varietà qualunque e conseguente specificazione geometrica della curvatura riemanniana. *Rend. Circ. Mat. Palermo (1884-1940)*, 42 :173–204, 1917.
- [Nas54] J. Nash. C^1 -isometric imbeddings. *Annals of Mathematics*, 60 :383–396, 1954.
- [Nas56] J. Nash. The imbedding problem for Riemannian manifolds. *Annals of Mathematics*, 63 :20–63, 1956.
- [Oli15a] E. Olivier. 100 ans de Relativité Générale II/III : courbure. *Bull. Inf. App. & App.*, 101 :97–123, 2015.
- [Oli15b] E. Olivier. 100 ans de Relativité Générale III/III : déflexion de la lumière. *Bull. Inf. App. & App.*, 102 :139–167, 2015.
- [Ric88] G. Ricci. Delle derivazioni covarianti e contravarianti e del loro uso nella analisi applicata. *Stud. ed. Univ. Padova, Padova*, 3, 12 :23 p., 1888.
- [RTL00] G. Ricci and Tullio T. Levi-Civita. Méthodes de calcul différentiel absolu et leurs applications. *Mathematische Annalen (Springer)*, 54 :125–201, 1900.

La cosmologie newtonnienne (*la minute d'Anselme Lanturlu*)

Jean-Pierre PETIT^{1,2}

Résumé. – Note de Jean-Pierre Petit, portant sur la modélisation de l'Univers par les méthodes newtonniennes (initialement publiée dans le volume no 2 de Quadrature de Janvier-Février 1990 p. 26-28)



Historiquement, et cela on le verra dans les articles suivants, Einstein commença en 1917 par créer un modèle d'univers *stationnaire*, qui n'évoluait pas dans le temps. Tout simplement parce que personne n'aurait imaginé une seconde qu'il pu en être autrement. C'est le russe Friedman qui sauta ce pas quelques années plus tard et la cosmologie contemporaine est basée sur ses travaux. Rageur, Einstein avait alors lâché : « Ah, si j'avais pu imaginer que l'univers était instationnaire, j'aurais trouvé ce modèle avant Friedman. » Si, comme disent les lacedémoniens... Si on avait pensé à cela, on aurait pu s'engager dans cette voie avant même que la relativité restreinte n'apparût.

Mais ceci fut redécouvert qu'en 1934, par Milne et Mac Crea, à la surprise générale, d'ailleurs. L'équation de Friedman, clef des cosmologies, avait été découverte à l'issue de calculs affreusement complexes, nécessitant des outils mathématiques comme les *tenseurs*. Or voici que Milne et Mac Crea montraient, sur

moins d'une page de calcul, que tout ceci pouvait découler du simple formalisme de la mécanique des fluides la plus classique, sans transformation de Lorentz, avec une vitesse de la lumière infinie. Ceci provient du fait que la machine cosmologique est en quelque sorte « parasitée » par la relativité restreinte.



FIGURE 1

1. Le grumeau pulsant

Nous allons donner une version simplifiée de la théorie de Milne et Mac Crea. Imaginons que nous soyons au 19^{ème} siècle. La relativité restreinte n'existe pas. Beaucoup de gents sont encore prêts à croire à la thèse des actions à distance, c'est-à-dire au fait que la gravitation agit instantanément d'un point à l'autre du cosmos. Considérons alors un milieu fait de poussières. Celles-ci vont s'attirer mutuellement. Isolons par la pensée dans ce milieu une sphère de rayon R en supposant que les particules situées en son centre resteront immobiles. Il existe alors un théorème que le mathématicien retrouvera facilement, qui montre qu'une particule située à la surface de cette sphère de rayon R subit

1. Directeur de Recherches au CNRS
2. jppetit1937@yahoo.fr

l'attraction gravitationnelle de celles qui sont situées à l'intérieur de ce volume, exactement comme si toute la masse était concentrée en son centre. Ceci découle du fait que la force gravitationnelle est en $1/r^2$ « Figure 1 ».

La force qui s'exerce sur cette particule, équivalent à une force de pesanteur, de direction radiale par raison de symétrie, est :

$$F = -\frac{GMm}{R^2}$$

M est la masse du grumeau de poussière, m celle de la particule, G la constante de gravité et $R(t)$ le rayon du grumeau sphérique. Appliquons maintenant la relation : la force égale masse multipliée par accélération. L'accélération sera la dérivée seconde de R , notée R'' . Et on aura :

$$mR'' = -\frac{GMm}{R^2} \quad \text{ou} \quad R^2 R'' + GM = 0$$

avec $GM = \text{constante}$. Ceci est une équation différentielle liant R et R'' . C'est exactement le résultat trouvé par Friedman, bien que le chemin qu'il ait suivi ait été sensiblement différent. Une telle équation possède trois solutions. Friedman commença par trouver la solution *cycloïdale* :

$$R : \sqrt[3]{GM}(1 - \cos \theta)$$

avec $t = \theta - \sin \theta$. Vous pouvez vérifier en dérivant³. Cette solution, dite paramétrique puisque R et t se calculent à partir d'un paramètre θ , à l'allure ci-dessous (Figure 2). Dans cette solution $R(t)$ passe par la valeur 0. Le grumeau de gaz pulse et R passe par une valeur maximale. Elle évoque une univers *pulsant*.

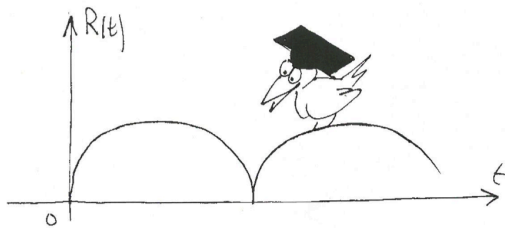


FIGURE 2. Solution cycloïdale

Einstein fut très vexé de voir un illustre inconnu lui ravir au poteau cette découverte, car il considérait la Relativité Générale, dont il avait jeté les bases, comme sa propriété... Cette solution bâtie par Friedman était en effet beaucoup plus élégante que la sienne (qui ne tenait pas debout, en vérité).

3. Rappelons qu'alors $\frac{dR}{dt} = \frac{dR}{d\theta} \frac{d\theta}{dt}$ et que : $R'' = \frac{d^2R}{dt^2} = \frac{d}{d\theta} \left(\frac{dR}{dt} \right) \frac{d\theta}{dt}$.

Ne faisant pas preuve d'un grand esprit sportif, il se détourna de la Relativité Générale, pendant une dizaine d'années, sans faire grand cas de la solution trouvée par l'outsider, qui ne reçut aucun prix pour son invention (et en tout cas, pas le prix Nobel, qu'il eût mérité cent fois).

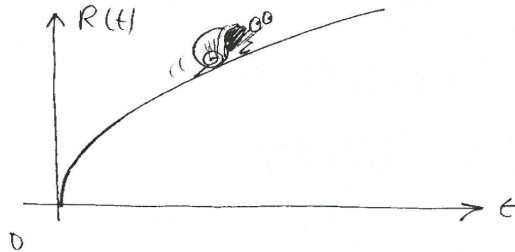


FIGURE 3. Loi parabolique d'Einstein - de Sitter

Einstein attendit tranquillement que Friedman meure et publia sous son nom et sous celui de son collègue de Sitter une autre solution de l'équation sous forme d'une fonction puissance. Retrouvons la rapidement : Soit à chercher une solution de la forme $R = at^m$. On a : $R' = amt^{m-1}$ et

$$R'' = am(m-1)t^{m-2}$$

d'où

$$a^3 m(m-1)t^{3m-2} + GM = 0$$

Comme le produit GM ne dépend pas du temps, il faut que $m = 2/3$, ce qui donnera

$$R = \sqrt[3]{\frac{9GM}{2}} t^{2/3}$$

C'est la solution *parabolique*, dite d'Einstein-de Sitter (voir Figure 3).

Il apparut par ailleurs qu'il existait une troisième solution, dite *hyperbolique* où apparaissent cette fois des lignes hyperboliques :

$$R = \sqrt[3]{GM}(\text{ch}(\theta) - 1)$$

avec $t = \text{sh}(\theta) - \theta$. Tout ceci étant très facile à programmer et à tracer à l'aide d'une calculatrice scientifique programmable. (Figure 4.)

2. Mythe de la caverne

Pourquoi est-il possible de retrouver un des résultats majeur de la Relativité Générale à l'aide de calculs aussi simple ? En opérant plus correctement nous serions parti des équations de conservation de la mécanique des fluides, ainsi de ce qu'on appelle l'équation de Poisson (qui traduit simplement, mathématiquement, le fait que la force gravitationnelle soit en $1/r^2$). Mais, même sous cette forme, le calcul tient quand même sur une page. L'explication est la suivante. Lorsqu'on procède au tir d'un véhicule spatial en direction de la lune, en toute rigueur on devrait tenir compte des corrections relativistes liées à la vitesse du mobile et aux effets de courbure dus à la présence de la terre et de la lune. Mais ces effets sont si infimes qu'on utilise carrément les bêtes équations newtonniennes. L'espace est courbe, mais, sur d'aussi « faibles » distances et avec des concentrations de masses aussi « infimes », on l'assimile à un espace euclidien, plat (c'est comme confondre la sphère terrestre et son plan tangent, pour des opérations d'arpentage à distances modestes).

La vitesse du train spatial est également si petite devant c qu'on décide de négliger les corrections de la relativité restreinte. Autrement dit, dans les équations rigoureuses, relativistes, tout ce passe comme si on faisait tendre c vers l'infini, ce qui redonnerait les équations newtonniennes, les équations « de tous les jours ».

Ces équations relativistes quelles sont-elles ? Et bien nous ne les connaissons pas, tout simplement. Nous pensons qu'elles existent, mais nous n'avons pas réussi à les construire. Tout ce qu'on peut dire, c'est que sur de courtes distances et à des vitesses V faibles devant c , ces équations doivent dégénérer selon les équations de la mécanique des fluides classique. À la manière de Platon, nous devons considérer ces équations « usuelles » comme les « ombres » d'équations plus complexes, que nous ignorons. Le philosophe grec avait jadis aussi suggéré que tout ce qui était tangible, perceptible, n'était l'ombre d'une réalité plus complexe.

La prochaine fois, nous tenterons de vous expliquer comment s'y est pris Einstein pour jeter les bases des la Relativité Générale.

J.P.P.

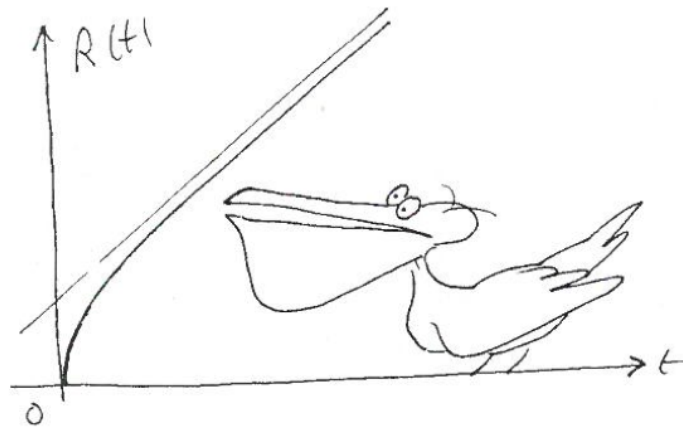


FIGURE 4. La solution hyperbolique

Deux ou trois choses que je sais de Sakharov (*la minute d'Anselme Lanturlu*)

Jean-Pierre PETIT^{1,2}

Résumé. – Extrait d'une note de Jean-Pierre Petit à propos de l'oeuvre scientifique d'Andrei Sakharov publié dans Quadrature n°3 Mars-Avril 1990.

Andrei Sakharov est prix Nobel de la Paix. On le connaît surtout comme le grand refusnik soviétique, le courageux militant pour les droits de l'homme. Certains savent aussi qu'il a contribué fortement à la mise au point de la bombe H soviétique.

On a dit il y a quelques années que le Kremlin l'avait tenu reclus pendant de longues années parce qu'il aurait été détenteur de secrets scientifiques et militaires qui auraient fait de lui un véritable danger en cas de passage à l'ouest.

On connaît très bien l'oeuvre scientifique de Sakharov depuis la publication en 1982 de l'ensemble de ses travaux, traduits en anglais. Les éditions Anthropos (15 rue Lacépède, Paris 75005) traduisent par la suite cet ouvrage en français en 1984. Si vous êtes un physicien théoricien chevronné, précipitez vous sur ce livre. Mais si vous êtes néophyte ces pages ne vous diront pas grand chose, car il s'agit de la reproduction intégrale des articles scientifiques originaux, et non d'un texte de vulgarisation accessible au grand public.

Sakharov est né le 21 mai 1921 à Moscou. Son père était un professeur de physique connu, auteur de nombreux ouvrages de vulgarisation. Avant la guerre la vie des Sakharov correspondait à celle de bourgeois cultivés relativement aisés (vis à vis des normes soviétiques). Andreï commença ses études supérieures en 1938 et les acheva en pleine guerre, en 1942. Il fut alors directement affecté à une grande usine d'armement, sur la Volga, où il travailla jusqu'en 1945. Ceci lui donna une expérience d'homme de terrain, d'ingénieur, qui vint compléter une solide formation de physicien théoricien. En 1945 il devint l'assistant d'Igor Tamm, futur prix Nobel et en 1948

l'équipe s'attela à la réalisation d'une arme thermonucléaire, avec le succès que l'on sait.

En 50 Tamm et Sakharov jetèrent les premières bases de la filière de la fusion contrôlée et inventèrent le Tokamak, ultérieurement développé par l'académicien Artsimovitch. En 1952 Sakharov travailla activement dans le domaine de la MHD (en français magnétohydrodynamique, en russe MGD ou magnétogaz dynamique). Un système à magnétostriction produira alors un champ record de deux mille cinq cent teslas en 1964. Toutes ces recherches constituèrent la base de travaux développés par E.Velikhov, actuel vice-président de l'Académie des Sciences d'URSS, et sur lequel se fonde l'arsenal des armes spatiales soviétiques.

Sakharov fut élu à l'Académie des Sciences en 1953. En 1967 sa carrière scientifique connut un changement profond. En parallèle avec son engagement de militant de la paix, il se réorienta vers la cosmologie et publia des travaux encore peu connus sur un modèle de structure gémeilaire de l'univers.

Comme il est difficile de résumer un tel personnage en un aussi court article, nous nous axerons sur la trame scientifique. Rappelons seulement que sa prise de position pour la limitation et l'interdiction des armes nucléaires date du début des années 1950 et entraîna dès 1961 un conflit violent avec Nikita Kroutchev. Sakharov contribua activement à la signature, en 1962, du traité d'interdiction des essais nucléaires dans l'atmosphère, dans l'espace et sous la mer. Il reçut le prix Nobel de la Paix en 1975 et fut exilé à Gorki en 1980, après privation de toutes ses distinctions honorifiques.

[...]

1967 marque un virage à angle droit dans la trajectoire de l'académicien. En même temps

1. Directeur de Recherches au CNRS
2. jppetit1937@yahoo.fr

qu'il s'engageait totalement dans sa lutte pour les droits de l'homme il décida d'abandonner toute recherche à vocation militaire. En jurant à ses collègues qu'il ne trahira aucun secret, il bifurqua brutalement vers la cosmologie. Là encore sa contribution s'avèra totalement originale.

La question point de départ était que se passe-t-il à $t < 0$?

Dans un papier de 1967 Sakharov fournit une réponse pour le moins originale. Si on suit, au long d'une « expérience de pensée », une ligne d'univers, à la traversée de la singularité nommée Big Bang la flèche du temps subit une inversion (notons que le physicien anglais non moins célèbre Hawking suggéra dans un papier de 1987 une inversion similaire de cette flèche du temps, mais cette fois lorsque l'univers semble lassé de s'étendre et se recontracte de manière... rétrochrone).

Sakharov compléta cette vision en suggérant que les phénomènes situés de l'autre côté du Big Bang, c'est à dire sur son « versant rétrochrone » puissent n'être que des images « CPT » de ce qui se passe sur notre versant « diachrone ».

Que veut dire CPT ? Il faut entendre C pour charge, P pour « parité » et T pour temps. La parité correspond à l'orientation droite gauche. Lorsqu'on inverse la charge, le temps et qu'on réalise une symétrie en miroir (c'est-à-dire qu'on remplace les objets par leurs « images énantiomorphes ») on change la matière en antimatière (et vice versa). Sakharov apporta donc une réponse à l'éternelle question « mais où est passée l'antimatière ? ». Elle serait ainsi « de l'autre côté de ce miroir spatio temporel » qu'est le Big Bang³.

On sait que notre versant d'univers présente une « violation du principe de parité ». Lorsqu'on considère deux réactions nucléaires qui sont en miroir, spatialement parlant, les résultats sont les mêmes, mais les durées de ces phénomènes diffèrent de façon mesurable. C'est une des découvertes scientifiques majeures de l'après guerre. Sakharov suggéra que cette violation du principe puisse être inversée dans le versant gemellaire d'univers, symétrique du nôtre. Sa vision cosmologique complète implique une synthèse des constituants matériels de l'univers, des baryons (bary, en

grec, veut dire lourd) à partir de quarks, tandis que les anti-baryons résulteraient de la fusion d'antiquarks. Au cours du temps la violation du principe de parité aurait permis dans notre versant d'univers une production en excès de matière, au détriment de l'antimatière, dans la proportion de un pour un milliard. C'est ce qui aurait évité la totale annihilation au premier centième de seconde. Notre versant d'univers contiendrait ainsi un excès de matière et un excès d'antiquarks. Le scénario serait évidemment inverse dans cet autre versant de l'univers qui contiendrait un excès d'antimatière et un excès de quarks (en fait l'image en miroir de notre situation spatio-temporelle). Notons que cette asymétrie baryonique va, selon Sakharov, avec une durée de vie du proton finie, évaluée par Sakharov à 1050 ans, idée qu'il a été un des tous premiers à lancer.

Comme le note Susskind, de Stanford University, commentateur des travaux de Sakharov dans l'ouvrage paru aux éditions Anthropos :

« Avant les audacieuses hypothèses de Sakharov, la seule réponse était que Dieu avait créé l'univers avec plus de matière que d'antimatière, un point c'est tout ».

La place manque évidemment dans ce court article pour déployer tout l'éventail des directions de recherche suivies par Sakharov, véritable « généraliste de la physique ». Nous terminerons en reproduisant la dernière page de son discours de réception de prix Nobel, qui date de 1975 :

« Il y a des milliers d'années les tribus humaines souffraient de grandes privations dans la lutte pour l'existence. Il était alors important, non seulement de savoir manier une matraque, mais de posséder la capacité de penser intelligemment, de tenir compte du savoir et de l'expérience engrangés par la tribu et de développer les liens qui établiraient les bases d'une coopération avec d'autres tribus. Aujourd'hui la race humaine doit affronter une épreuve analogue. Plusieurs civilisations pourraient exister dans l'espace infini, parmi lesquelles des sociétés qui pourraient être plus sages et plus "performantes" que la nôtre. Je soutiens l'hypothèse cosmologique selon laquelle le développement de l'Univers se répète un nombre infini de fois, suivant des caractéristiques essentielles. D'autres civilisations, y compris certaines plus "performantes", sont inscrites un nombre infini de fois sur les pages "suivantes" »

3. Voir par exemple les premières minutes de l'univers, de Steven Weinberg, au Seuil, ou la traduction que j'en ai faite en bande dessinée : mon album Big Bang, toujours dans la série des Aventures d'Anselme Lanturlu.

ou "précédentes" du Livre de l'Univers. Néanmoins nous ne devons pas minimiser nos efforts sacrés en ce monde, où comme de faibles lueurs dans l'obscurité, nous avons surgi pour un instant du néant de l'inconscience obscure à l'existence matérielle. Nous devons respecter les exigences de la raison et créer une vie qui soit digne de nous-mêmes et des buts que nous percevons à peine. »

J.P.P.

P.S. : Le lecteur curieux pourra trouver une présentation imagée de ce thème dans ma bande dessinée LE TOPOLOGICON, paru aux éditions Belin, 8 rue Férou, Paris 75006. Le scientifique plus averti pourra se référer aux deux articles que j'ai publiés en 1977 aux comptes rendus de l'Académie des Sciences de Paris (en ignorant ce qui avait été fait par Sakharov) et intitulés "Univers Enantiomorphes à temps propres opposés" et "Univers en interaction avec leur image dans le miroir du temps".

INJUSTICE À HAUT DÉBIT

BILAN DU SOMMET, DE L'e-GB : BIENTÔT, PEUT-ÊTRE, LE HAUT DÉBIT SERA RÉSERVÉ AUX SITES QUI ONT DU FRIC ET LE BAS DÉBIT AUX SITES GAGNE-PETIT

JE COMMANDE DES PÂTES AUX TRUFFES... ELLES ARRIVENT 20 MINUTES APRÈS !



JE COMMANDE UNE PIZZA ... ELLE ARRIVE 6 MOIS APRÈS



JE COMMANDE UNE PUTE DE LUXE ...ELLE ARRIVE 35 MINUTES APRÈS



JE COMMANDE UN GOSSE SUR UN SITE PÉDOPHILE... ...IL ARRIVE 20 ANS APRÈS



BIAA

Bulletin d'Informatique Approfondie & Applications
Computation - Information
Volumes 2015

100 - 101 - 102



Couverture : dessin de Michel Avezard (Zevard) interprété par la MMIAGe (1985-1987)

ÉDITORIAL : Faut-il laisser tomber l'allemand, le chinois ou l'espagnol et faire code informatique en seconde langue ?

Jean-Pierre ARCHAMBAULT et Rémi BOULLE¹

Résumé. – La question de l'apprentissage de langages informatiques à l'école fait son chemin. C'est le sens du projet Code.org qui mobilise des stars du web et des nouvelles technologies pour inciter à l'enseignement de lignes de code aux têtes blondes outre-Atlantique. En France, le patron de Free, Xavier Niel, tenait des propos similaires lors de sa conférence annonçant son projet d'école informatique gratuite. *Propos recueillis par Benjamin Weil (Journal en ligne Atlantico - 3 juin 2013).*

Atlantico : Un récent rapport de l'Académie des sciences² préconise l'enseignement de l'informatique en tant que matière à part entière dans l'Éducation nationale. Faut-il aller plus loin et enseigner les langages informatiques dans les écoles et établissements du secondaire ? Quels peuvent être les avantages pour les élèves d'apprendre à coder ?

Jean-Pierre Archambault : Les langages informatiques et la programmation sont l'un des domaines de la science informatique, avec l'algorithmique, la théorie de l'information et les machines (architecture, ordinateurs, réseaux...). A ce titre, ils ont nécessairement leur place dans un enseignement de l'informatique.

De plus, ils permettent de comprendre ce qu'est l'informatique, de percevoir sa « nature profonde », de s'en imprégner. Pour s'approprier des notions (fichier, protocole de communication, « verrou mortel »...), rien de tel que d'écrire des « petits » programmes.

Mais les vertus pédagogiques de l'activité de programmation et des langages valent également pour l'apprentissage des autres disciplines. Encore faut-il que les élèves sachent programmer ! La programmation est un élément de cursus informatique apprécié des élèves, car elle les place dans une situation active et créative, dans laquelle ils peuvent eux-mêmes fabriquer un objet. **On constate en effet avec l'ordinateur une transposition des comportements classiques que l'on observe dans le domaine de la fabrication des objets matériels.** À la manière d'un artisan qui prolonge ses efforts tant que son ouvrage n'est pas effectivement terminé et qu'il fonctionne, un lycéen, qui par ailleurs se contentera d'avoir résolu neuf questions sur dix de son problème de mathématiques (ce qui

1. Jean-Pierre Archambault est président de l'association Enseignement Public & Informatique (EPI) dont l'un des objectifs est de promouvoir l'enseignement de l'informatique en France. (<http://www.epi.asso.fr>)

Rémi Boulle est vice-président de l'April (Association de promotion et de défense du logiciel libre) en charge des questions d'éducation.

2. L'enseignement de l'informatique en France - Il est urgent de ne plus attendre. Rapport de l'Académie des sciences - Mai 2013. (http://www.academie-sciences.fr/pdf/rapport/rads_0513.pdf)

n'est déjà pas si mal!), s'acharnera jusqu'à ce que « tourne » le programme de résolution de l'équation du second degré que son professeur lui a demandé d'écrire, pour qu'il cerne mieux les notions d'inconnue, de coefficient et de paramètre. **Ces potentialités pédagogiques de la programmation, qui favorisent l'activité intellectuelle, sont parfois paradoxalement et curieusement oubliées par des pédagogues avertis** (qui, par ailleurs apprécient les vertus de l'ordinateur et d'internet, outil pédagogique).

Enfin, la programmation est une excellente école de la rigueur, de la logique. Vraiment, pourquoi s'en priver ?

Rémi Boule : L'école doit former le futur citoyen à la société du XXI^e siècle donc oui, il est urgent d'agir. Nous faisons fausse route depuis trop longtemps. **Après plus de 12 ans d'échec le B2i (brevet informatique et internet) a malheureusement conditionné le mode de représentation de l'informatique à l'école de nombreux décideurs : l'informatique ne serait qu'un simple outil.** C'est depuis le début une mauvaise réponse à un vrai problème : faire acquérir à tous des connaissances en informatique.

Il faut vraiment être éloigné de la pratique ou suivre les intérêts d'entreprises privées pour ne pas le voir (voir à ce sujet le tour de France du numérique³ organisé par le café pédagogique soutenu par Microsoft qui fait la tournée des CRDP).

Au-delà des langages informatiques et dans l'idée de former le futur citoyen, il faut aussi un enseignement qui permette aux élèves de **comprendre les enjeux de l'informatique dans notre société** : neutralité des réseaux, licences libres, fonctionnement de l'Internet, droit d'auteur, problèmes liés aux DRM, dangers des brevets logiciels, libération des données publiques, Wikipédia, travail collaboratif. . .

Il y a un prérequis fondamental à un enseignement de l'informatique au service de tous : se débarrasser complètement des logiciels propriétaires et des formats propriétaires et fermés. Ce sont des freins à tout cela, ils parasitent le système éducatif et brident les innovations.

N'est-ce pas trop ambitieux de vouloir aborder d'autres langages, alors même que le niveau général des élèves en français (orthographe, grammaire) tendrait à s'effondrer ? Y a-t-il des résistances doctrinales face à l'enseignement du code informatique ?

Jean-Pierre Archambault : Une réponse pragmatique : dans les années 1980, des instituteurs enseignaient la programmation Logo à leurs élèves avec la fameuse tortue. Et cela marchait très bien. **A mon sens, aborder d'autres langages est plutôt de nature à aider l'enseignement du français. Un peu comme la connaissance d'une langue étrangère facilite l'apprentissage d'une nouvelle langue et renforce la maîtrise que l'on a de sa langue maternelle.**

Il y a des résistances de nature doctrinale, et d'autres, à l'enseignement de l'informatique en général et du code informatique en particulier. Ce qui est nouveau a toujours du

3. <http://www.april.org/le-service-public-deducation-sceren-cndp-est-il-le-nouveau-showroom-de-microsoft-avec-la-complicite-du-cafe-pedagogique>

mal à émerger. Déjà, Confucius mettait en garde : *«Lorsque tu fais quelque chose, sache que tu auras contre toi ceux qui voulaient faire la même chose, ceux qui voulaient faire le contraire et l'immense majorité de ceux qui ne voulaient rien faire.»*

Rappelons l'essentiel. L'informatique et le numérique sont présents dans la vie quotidienne de tout un chacun. L'informatisation est la forme contemporaine de l'industrialisation. On ne compte plus les débats de société suscités par l'informatique, la science qui est au cœur du numérique. Il s'agit donc, avec un enseignement de l'informatique au collège puis au lycée dans une discipline scolaire en tant que telle, après une sensibilisation à l'école primaire, de donner à tous les élèves les moyens de comprendre et de maîtriser le monde dans lequel ils vivent, d'encourager leur créativité, de former « l'homme, le travailleur et le citoyen », à savoir les missions de l'École.

Il faut donner à tous les élèves la culture générale de leur époque.

Des premiers pas ont été faits : création à la rentrée 2012 de l'enseignement de spécialité optionnel « Informatique et sciences du numérique » (ISN), extension en Terminale ES et L à la rentrée 2014 sous forme d'une option, enseignement d'informatique à la rentrée 2013 pour les élèves des classes préparatoires aux grandes écoles scientifiques. Ils en appellent beaucoup d'autres.

Rémi Boule : C'est simplement devenu nécessaire. Sans forcément rajouter une discipline de plus, on peut tout à fait introduire un enseignement de l'informatique dans le cadre des cours de technologie et ce dès le collège. **Il est temps de former des élèves qui auront la capacité de devenir des citoyens informatiquement éclairés ou des créateurs et plus uniquement des consommateurs de prisons numériques que l'on pilote avec son doigt.**

Des résistances ? Oui, elles sont importantes. J'en vois essentiellement deux. D'une part, **les mêmes personnes qui ont fait perdre plus de 10 ans au système éducatif avec le B2i et parlent encore, en 2013, de « révolution du numérique »** ne veulent toujours pas entendre parler d'un enseignement de l'informatique. Elles sont prises en otage de leurs propres constructions intellectuelles erronées et font un lobbying forcené à longueur de colloques et de salons « entre soi ».

D'autre part, **l'opportunisme et la naïveté de certains élus des départements ou des régions.** Ce sont eux qui équipent collèges et lycées en matériel informatique. L'achat massif de matériel informatique est un moyen commode et opportun pour montrer à leurs électeurs leur attachement à l'école. Par conséquent on déploie des milliers de tablettes sans discernement ni consultation. Naïvement ils pensent que cela répond à une demande des personnels enseignants.

Ces deux obstacles s'entretiennent en plus mutuellement. Les premiers, vu leur pauvre représentation de l'informatique n'ont besoin que de « minitels tactiles 2.0 » pour ce qu'ils appellent « innover pédagogiquement » et les autres les leur fournissent trop contents de

croire satisfaire des besoins et œuvrer pour l'intérêt général. Où est la plus-value pédagogique ? Dans ce schéma-là, l'informatique en tant qu'objet d'enseignement n'a malheureusement pas sa place et c'est un problème pour la société.

Alors qu'en France on équipe massivement des élèves avec des prisons numériques de type iPad, en Angleterre, on leur distribue des Raspberry Pi ou des Arduino. C'est environ 20 fois moins cher et permet de réelles pratiques pédagogiques créatrices de connaissances en informatique. Comme l'a dénoncé l'association Pullco : « en choisissant de distribuer des iPad aux collégiens, le Conseil Général de la Corrèze a fait le choix politique de suivre un effet de mode, en privilégiant le paraître au bon sens. »

Que les politiques se réunissent avec les professeurs pour décider ensemble ce que sera l'école du XXI^{ème} siècle. Pour nous il est clair que ce n'est pas celle où le « numérique », mis à toutes les sauces, est un simple outil. C'est celle qui utilise des logiciels libres, des formats ouverts et où on apprend l'informatique au sens large. Espérons qu'au moins cette priorité soit enfin inscrite dans la loi⁴.

Justement, comment est abordée la question de l'enseignement de l'informatique par les décideurs politiques ?

Rémi Boule : Malheureusement, les récents débats parlementaires sur le projet de loi de refondation de l'école montre que nous régressons : l'amendement 118 qui visait à introduire un enseignement de l'informatique⁵ a été rejeté au motif que les programmes étaient déjà trop chargés. Certes mais en 2013 à l'heure où l'informatique occupe une place centrale dans la société, c'est une position que n'est plus tenable, il aurait suffi de l'enseigner dans le cadre des cours de technologie.

Sauf coup de théâtre à l'assemblée, l'école risque fort de continuer à enseigner aux élèves comment utiliser des outils créés par d'autres et qu'ils ne seront que difficilement en mesure de créer eux-mêmes ! A-t-on besoin de cela dans le contexte économique actuel ?

Quels langages informatiques peuvent être enseignés en priorité aux élèves selon les niveaux et quelles sont leurs spécificités ?

Jean-Pierre Archambault : Nous avons parlé de Logo à l'école primaire. Un manuel scolaire destiné aux lycéens d'ISN a été édité l'année dernière par les éditions Eyrolles⁶. Cet ouvrage a été publié avec le concours de l'EPI, la SIF et l'INRIA. Les programmes sont écrits en Java. Une version de ce manuel va être publiée pour la rentrée 2013 avec les programmes écrits en Python. **L'important est que le langage choisi contienne les « grandes » instructions : la déclaration, l'affectation, la séquence, le test et la boucle ; les types de base ; les fonctions.** Il existe de nombreux langages. Il faut en choisir un bien adapté à des objectifs pédagogiques, en évitant soigneusement la « guerre de religion »

4. <http://www.april.org/la-commission-des-affaires-culturelles-et-de-leducation-de-lassemblee-nationale-confirme-la-priorite>

5. http://www.senat.fr/seances/s201305/s20130523/s20130523019.html#Niv3_art_Article_26

6. <http://www.editions-eyrolles.com/Livre/9782212135435/informatique-et-sciences-du-numerique>

qui, comme l'expérience le prouve, n'est jamais bien loin. Et l'important c'est aussi ce qu'il y a en amont de la programmation : la compréhension du problème à résoudre, la conception du bon algorithme, son expression en langage naturel.

Rémi Boule : Il n'y a pas seulement les langages. Il faut comprendre comment fonctionne un réseau, ce qu'est un format de fichier, utiliser un traitement de texte de façon structurée, le web, quels sont les enjeux citoyens de la neutralité des réseaux, comment fonctionne une machine, qu'est-ce qu'un algorithme... Comme dit, un prérequis est l'utilisation massive de logiciels et ressources libres qui ne peut que libérer les pratiques pédagogiques et l'innovation.

Je signale à ce propos une initiative très intéressante. L'IUT de Blagnac propose en partenariat avec « Upstream University »⁷ à ses étudiants de participer à un projet de développement de logiciel libre (OpenStack de mémoire). Leur parcours n'est validé que si leur code est « comité Upstream » c'est à dire si leur code informatique est intégré dans celui du projet. Pour cela les étudiants doivent bien sûr savoir coder mais aussi avec des notions sur les licences libres, le travail collaboratif et beaucoup d'autres connaissances connexes qu'ils devraient posséder au sortir du Lycée... Nous avons là ce qui est attendu après le baccalauréat, à nous de faire en sorte en amont que cela soit possible et accessible à qui le souhaite dès maintenant.

7. <http://upstream-university.org/>

UN COMMENTAIRE

Nos articles sont ouverts aux commentaires sur une période de 7 jours. Face à certains abus et dérives, nous vous rappelons que cet espace a vocation à partager vos avis sur nos contenus et à débattre mais en aucun cas à proférer des propos calomnieux, violents ou injurieux. Nous vous rappelons également que nous modérons ces commentaires et que nous pouvons être amenés à bloquer les comptes qui contreviendraient de façon récurrente à nos conditions d'utilisation. (Atlantico.fr)

Par autodidacte.info - 03/06/2013 - 17 :22

Ne pas en faire une nécessité mais proposer en option

C'est vrai que de très nombreux pays industrialisés ne jurent que par l'économie numérique, et que la France est à la traîne. De là à imposer le code à l'école, bof. D'autant plus que, comme le disent certains dans leurs commentaires, apprendre à lire n'enseigne pas à réfléchir... S'il s'agit de « pondre du code » (expression du milieu informatique dont je fais partie) sans comprendre le pourquoi du comment, cela ne donnera qu'une armée de perroquets. Et quel code, d'abord ? Attention à l'enfermement et aux lobbys de grosses entreprises qui vont forcément chercher à mettre en avant leur technologie. « Code is poetry » dit l'Autre (le code c'est de la poésie). Et c'est vrai qu'il est bon d'en apprendre plusieurs (Lisp, Fortran, C, Python...) pour se faire son idée du langage en tant que concept. Tout comme il est bon de comprendre le concept intellectuel de l'Allemand, de l'Anglais, etc... Car lire « dans le texte » n'a pas son pareil : on s'approprie l'Autre au travers de sa propre histoire. L'éducation à la programmation pourrait se faire de façon ludique et créative, mais en gardant en tête un impératif : l'apprentissage de la réflexion par soi-même, pas de la consommation !

Contradiction ou indécidabilité : il faut choisir !

Patrick ISOARDI¹

Résumé. – À propos du théorème d'incomplétude de Gödel.

1. Formaliser les mathématiques : le rêve de Hilbert

“Je suis un menteur” .

L'individu qui considère que dans le monde qui l'entoure tout ne peut être que vrai ou faux, ne peut pas statuer sur la vérité de cette phrase. En effet :

- s'il considère que la phrase “Je suis un menteur” est vraie, alors celui qui l'énonce est un menteur car c'est ce qu'affirme la phrase. Les menteurs ne prêchant que le faux, cette phrase est fautive ; il y a contradiction avec l'hypothèse de départ.

- s'il considère que la phrase “Je suis un menteur” est fautive, celui qui l'énonce fait partie des menteurs car ce qu'il dit est faux. Mais disant qu'il est un menteur, c'est donc vrai. Cela contredit l'hypothèse sur les menteurs.

La phrase du menteur semble ne pouvoir être ni vraie ni fautive. C'est un paradoxe.

Le langage naturel autorise ce genre de discours, mais le langage rigoureux des Mathématiques ne devrait permettre d'exprimer que des vérités non contradictoires. Certains mathématiciens et logiciens du début du XX siècle, notamment Russel, Whitehead ou Hilbert ont durement travaillé pour ne plus employer le langage naturel dans les mathématiques. Ils voulaient les formaliser en proposant un système de notation rigoureux permettant d'écrire sans ambiguïté les idées élémentaires et des méthodes de raisonnement qui à partir de ces idées élémentaires démontreraient exactement et complètement tous les théorèmes de Mathématiques que le système ainsi formalisé peut produire.

Il s'agissait là d'un objectif similaire à celui de Leibniz qui quelques temps plus tôt voulait réduire le raisonnement à du calcul et les paradoxes à des erreurs de syntaxe : “ les chimères ne pourront être écrites en ces caractères ! ” C'est en fait cet objectif qui s'est révélé chimérique lorsqu'en 1931 Gödel démontra que si un système formalisé contient les opérations arithmétiques élémentaires et qu'on le veut non contradictoire alors il existe des vérités qu'il ne peut pas démontrer.

Ainsi le paradoxe ne réside pas dans le langage utilisé pour écrire la phrase du menteur, mais dans l'incapacité de déterminer par le raisonnement si elle est vraie ou fautive.

L'étude qui suit précise l'environnement dans lequel Gödel a travaillé. Elle devrait permettre d'appréhender les idées essentielles de la démonstration de son théorème d'incomplétude.

1. Ndr : d'après le numéro 73 du bulletin de 2006.

1.1. Système formel. C'est le cadre dans lequel Hilbert espérait que les mathématiques pourraient être produites, en partant des idées élémentaires sur lesquelles on appliquerait de façon mécanique des règles définissant les processus mentaux du raisonnement.

Un système formel se définit par la construction suivante :

- un ensemble fini de symboles ; c'est le vocabulaire avec lequel on peut écrire des mots et des phrases,
- des règles de formation avec lesquels on écrit des *formules* qui sont des phrases syntaxiquement correctes ; c'est la grammaire du système,
- un sous ensemble de formules : les *axiomes*,
- on définit une liste finie de relations entre les formules : les *règles d'inférence*.

Les deux premiers points définissent un langage et les deux derniers la notion de démonstration.

Une *démonstration* est une suite finie de formules dont chacune est soit un axiome soit une conséquence immédiate de l'application d'une règle d'inférence sur la formule précédente.

Un *théorème* du système est une formule dont il existe une démonstration.

A partir des axiomes le système ainsi formalisé produira mécaniquement et complètement tous les théorèmes².

1.2. L'indécidabilité : un grain de sable dans les systèmes mécaniques. Programmons un ordinateur pour qu'il applique successivement toutes les règles d'inférences à tous les axiomes et aux théorèmes générés. Notons au préalable que ce processus peut être infini. Imaginons à présent que nous voulions savoir si une certaine formule F est un théorème. Quatre cas peuvent se présenter à l'ordinateur :

1- Il liste F et pas non F \Rightarrow La formule F est un théorème.

2- Il liste non F et pas F \Rightarrow La formule non F est un théorème.

3- Il liste F et non F \Rightarrow Le système est *contradictoire* ou inconsistant, c'est à dire qu'il démontre à la fois une chose et son contraire. Ce que Hilbert voulait éviter pour les mathématiques.

4- Il ne liste ni F ni non F :

- Si l'ordinateur s'arrête, on dit que la formule F est *indécidable* dans le sens où il est impossible de démontrer formellement F ou nonF par une démonstration déductive des axiomes. F indécidable \Leftrightarrow F et nonF ne sont pas démontrables. On substitue ici la propriété : " est démontrable " à celle de " est un théorème " ; cela est dû à l'apparition de l'indécidabilité que l'on n'imaginait pas avec un " bon " système formel qui devrait toujours se trouver dans un des deux premiers cas.

2. Un tel système formel est déductif et mécanique au sens où il n'est pas nécessaire de faire appel à l'intelligence. On ne sait pas forcément de quoi on parle ni si ce qu'on dit est vrai. Ici la question n'est pas de savoir si une formule F est vraie mais si elle est bien la conséquence de l'application de règles d'inférences sur les axiomes ; dans ce cas F est un théorème et la suite des formules obtenues par l'application précédente est sa démonstration. Lorsqu'il n'y a pas de vérité intuitive le système formel est souvent qualifié de syntaxique ou encore typographique.

Bien que F soit indécidable, si elle est vraie (au sens où F détermine une propriété que l'on peut vérifier par ailleurs), nous dirons que le système est *incomplet*. Un système formel complet permettrait de démontrer tout ce qui est vrai. C'est ce que Hilbert recherchait.

- Si l'ordinateur ne s'arrête pas, on ne peut pas dire que F est indécidable mais que l'ordinateur n'a pas encore trouvé une démonstration pour F ou nonF, son travail n'étant pas encore terminé. Par exemple il n'est toujours pas démontré que tout nombre pair est la somme de 2 nombres premiers ; pourtant cela semble vrai aussi loin qu'on est allé avec un ordinateur. C'est une conjecture, celle de Goldbach : une formule (qui semble) vraie qui n'est pas (encore) démontrée.

Après cette analyse quelques questions se posent, notamment :

- Comment faire la distinction entre une conjecture et une formule qui n'est pas démontrable ?

- Comment peut-on savoir qu'une formule est vraie si on ne peut pas la démontrer ?

1.3. Système mécanique et vérité. La question précédente revient en fait sur la comparaison entre l'homme et la machine. Plus précisément : si un humain est capable de savoir que la phrase qu'il donne à un ordinateur est vraie, pourquoi l'ordinateur ne serait-il pas capable de découvrir la vérité ?

Voici la phrase : "L'ordinateur ne répondra jamais vrai à cette phrase". Que fait l'ordinateur ?

- S'il répond vrai, il affirme que "L'ordinateur ne répondra jamais vrai à cette phrase". Or ce n'est pas le cas, puisqu'il vient justement de répondre vrai à la phrase. Si l'ordinateur ne se contredit pas, il ne peut donc pas répondre vrai.

- S'il répond faux, il affirme que "L'ordinateur ne répondra jamais vrai à cette phrase" est une affirmation fautive. Or l'affirmation n'est pas fautive puisque l'ordinateur vient justement de ne pas répondre vrai. Si l'ordinateur ne se contredit pas, il ne peut donc pas répondre faux. L'ordinateur ne peut répondre ni vrai ni faux à la question. Elle est pour lui indécidable.

Et nous, pouvons nous répondre à la question?... La phrase dit : "L'ordinateur ne répondra jamais vrai à cette phrase". Nous venons de voir qu'en effet, l'ordinateur ne peut pas répondre vrai. Cette phrase est donc une vérité que l'on vient de vérifier " par ailleurs " en dehors de l'ordinateur et que celui-ci ne pourra jamais découvrir³. Ce n'est pas une conjecture, mais bien une certitude ; une vérité qu'un système mécanique ne pourra jamais démontrer. Gödel l'exprimera par une formule arithmétique dans un système formel, rendant celui-ci incomplet.

1.4. Mécanique, intelligence. Cette différence entre " l'homme qui sait " et " la machine qui ne peut pas trouver la vérité " a été mise en évidence dans un système formel⁴ par

3. Observons toutefois que l'indécidabilité de l'ordinateur est là justement parce que la phrase soulève un paradoxe. Si un système mécanique arrivait à donner une valeur de vérité à un paradoxe, cela serait inquiétant ! En cette circonstance, l'ordinateur réagit exactement de la manière attendue : quand on lui demande de démontrer un paradoxe, il n'y arrive pas, et c'est heureux !

4. Symboles : M, I, U. Axiomes : MI. Règles : (1) xI → xIU (2) Mx → Mxx (3) xIII → xU (4) xUU → x. La lettre "x" ne sert qu'à remplacer n'importe quelle chaîne.

Douglas Hofstadter dans son livre : "Gödel Escher Bach" (page 38) où il pose la question : La formule MU est-elle un théorème du système ?

Jouons un peu à l'intérieur de ce système formel. Par exemple l'application de la règle 1 sur l'axiome MI produit MIU alors que les règles 3 et 4 ne peuvent s'appliquer sur cet axiome. La démonstration produite par l'application sur l'axiome MI de la suite des règles 223 génère MI MII MIII MIU.

Après quelques amusements de cette nature, la question initiale revient : ce mode mécanique de dérivation produira-t-il MU un jour ?

Cette énigme n'est pourtant qu'un problème d'arithmétique. Si on se place à l'extérieur du système l'observation des règles et de l'axiome MI montre que le nombre de I ne peut pas être un multiple de 3 et la règle 3 ne produira donc jamais xU. MU n'est donc pas un théorème du système formel, mais celui-ci ne pourra le démontrer alors qu'un simple raisonnement arithmétique donne la solution.

Ainsi les deux exemples précédents montrent simplement qu'un système mécanique ne peut pas toujours étudier de l'intérieur avec ses propres règles, la vérité d'une phrase ou la théorèmeté d'une formule. En revanche, on peut le faire de l'extérieur par des raisonnements métamathématiques.

Une nouvelle question se pose : si on inclut le raisonnement (ici l'arithmétique) à l'intérieur du système formel, peut-on alors déterminer la nature théorématique d'une formule ? Gödel montrera que la réponse reste négative ; le raisonnement ne permet pas toujours de déterminer la vérité.

1.5. Vérité et démonstration. Essayons d'être plus précis sur la relation entre vérité et démontrabilité ou théorèmeté.

La phrase : "x est pair" exprime la propriété arithmétique de parité que le sujet x peut avoir. On ne sait pas si cette propriété est vraie ou fausse, mais lorsqu'on remplace x par 5 par exemple, on obtient la phrase : "5 est pair" qui est fausse. Selon le même procédé, "14 est pair" est une phrase vraie. Ainsi, lorsqu'on remplace x par un entier, on pourra diviser celui-ci par 2 et comparer le reste de cette division à 0 pour démontrer ou réfuter l'affirmation. La suite des opérations de division et de comparaison constitue un procédé mécanique de décision (algorithme) qui donne toujours au bout d'un temps fini, un résultat vrai (si le reste est 0) ou faux (si le reste est 1). Toutefois on observera que le résultat n'est pas un théorème, mais une valeur de vérité.

On retrouve les deux manières classiques de résoudre les problèmes :

- la méthode algorithmique remise en vigueur avec l'essor de l'informatique, qui prescrit une suite d'opérations arithmétiques à effectuer, terminée par un test qui répond au bout d'un temps fini par vrai ou faux lorsque la propriété est décidable.

- la méthode déductive des systèmes formels où il s'agit de combiner un ensemble d'axiomes au moyen d'inférences logiques pour démontrer des théorèmes (qui sont alors considérés comme des vérités).

Quelle relation unit les 2 méthodes ?

Si $P(x)$ signifie "x à la propriété P" et que $\forall n$ il existe un algorithme qui détermine au bout d'un temps fini la valeur de vérité de la formule $P(n) \Rightarrow$ la propriété $P(x)$ est dite décidable. De plus :

si $P(x)$ est décidable et que $P(n)$ est vraie $\Rightarrow P(n)$ est démontrable dans un système formel. (1)

Dans le cas où $P(n)$ est faux $\Rightarrow P(n)$ est réfutable. Mais la réfutation n'est pas naturelle dans les systèmes formels.

Sautons dans un nouveau niveau d'abstraction où le problème de l'infini va devenir de plus en plus pressant mais dont la réflexion nous aidera par la suite.

Considérons la formule existentielle : " $\exists x P(x)$ " et la formule universelle : " $\forall x P(x)$ " dans lesquelles $P(x)$ est une propriété décidable. Nous remarquerons que $P(x)$ représente une infinité de formules alors que les 2 autres sont chacune d'elles une simple formule contenant une variable et qui sera vraie ou fausse.

Par exemple si $P(x)$: " x est pair" alors " $\exists x P(x)$ " est vraie car $P(2)$ est vraie et " $\forall x P(x)$ " est fausse car $P(1)$ est faux. Plus généralement :

- si on est sûr de trouver un n tel que $P(n)$ est vrai alors " $\exists x P(x)$ " est vrai et un système formel trouvera dans la suite $P(0), P(1), P(2), \dots$ le théorème $P(n)$. Ce qui démontre la formule " $\exists x P(x)$ ".

" $\exists x P(x)$ " vrai \Rightarrow " $\exists x P(x)$ " démontrable. (2)

Mais si on ne sait pas que ce n existe, tant qu'on ne l'a pas trouvé $P(0), P(1), P(2), \dots$ sont faux et " $\exists x P(x)$ " l'est aussi et peut le rester indéfiniment, sans pouvoir le démontrer voire l'affirmer car peut-être qu'un jour on trouvera ce n .

- de la même façon, si on est sûr de trouver un n tel que $P(n)$ est fausse alors " $\forall x P(x)$ " est fausse. Tant qu'on n'a pas trouvé ce n , $P(0), P(1), \dots$ sont vrais ; " $\forall x P(x)$ " reste vraie sans pouvoir le démontrer car la recherche peut être infinie.

Ainsi : " $\forall x P(x)$ " est réfutable mais si elle est vraie, l'étude cas par cas ne peut pas le démontrer de manière mécanique. En effet si " $\forall x P(x)$ " est vraie alors $P(0), P(1), P(2), \dots$ sont des théorèmes mais leur obtention est sans fin. Ce serait le cas de la conjecture de Goldbach qui pourrait être vraie, mais qui en même temps ne peut être démontrée à partir des axiomes de l'arithmétique.

Il est difficile d'énoncer d'autres relations entre vrai et démontrable, mais celles-ci ont permis à Gödel d'écrire dans un système formel non contradictoire contenant les opérations arithmétiques élémentaires une formule vraie qui ne sera jamais démontrable rendant ce système incomplet.

1.6. Métalangage et arithmétique. Revenons à la théorémété ou démontrabilité de la formule F . Comme nous l'avons vu précédemment c'est une propriété universelle manifestement extérieure au système formel. L'étude des propriétés de ce type se fait habituellement au moyen d'un métalangage différent de celui du système formel étudié. Généralement, il faut créer un second système formel qui traite du premier. L'inquiétude est de continuer cette construction à l'infini. Par contre si la phrase " F est démontrable" est elle-même une formule du système formel initial, celui-ci contient des formules ayant

pour sujet d'autres formules du système ; il contient donc son métalangage et il peut s'introspecter notamment au sujet de la nature théorématique de ses formules. Existe-t-il un tel système formel capable de s'introspecter ?

Vérifier que F est démontrable dans un système formel revient à chercher s'il existe au moins une démonstration dont la suite de formules constituée par l'application de règles d'inférences sur un axiome ou un théorème se termine par la formule F. Tant qu'on ne l'a pas trouvé, il faut continuer car rien ne nous dit qu'il n'existe pas et si tel est le cas dans ce mode mécanique de recherche sans fin, nous ne saurons jamais que F n'est pas démontrable. Par contre si on se donne la formule F et une suite finie D de formules, on peut toujours facilement vérifier en un nombre fini d'étapes si D est une démonstration de F. Appelons DEM(F,D) la phrase : "D est une démonstration de F. Comme pour la propriété de parité, quand on connaît F et D, un algorithme peut fournir en un temps fini le résultat vrai, faux. La décision se fait par des décalages sur les formules de D et des comparaisons pour vérifier que les règles d'inférences utilisées sont bien celles du système formel. Décalages et comparaisons sont des opérations arithmétiques.

Gödel montre que l'arithmétique est un métalangage pour l'étude des systèmes formels. Cela évite la régression infinie des métalangages où les symboles utilisés par le langage et le métalangage ne sont pas les mêmes.

1.7. Arithmétique formelle. L'arithmétique s'intéresse aux opérations (+, x) sur les entiers naturels et à leurs propriétés. Pour que l'arithmétique soit un métalangage pour l'étude des systèmes formels, il faut formaliser les opérations arithmétiques. Ces opérations sont des fonctions qui combinent des valeurs de N et donnent un résultat dans N ; par exemple $z = f(x,y)$. Ici z n'est pas une valeur de vérité mais un entier, or une formule définit une relation entre deux ou plusieurs objets et cette relation se voit attribuée une valeur de vérité. Comment passer de la fonction arithmétique f à la relation ? En remplaçant x,y,z par un triplet d'entiers k,m,n on pourra toujours vérifier si l'égalité $n = f(k,m)$ est vraie ou fautive. Ainsi :

toute fonction arithmétique f à 2 variables peut être *exprimée* dans un système formel par une formule $F(x,y,z)$ qui définit une propriété décidable du nombre $z = f(x,y)$. (3)

Dans le même souci de formalisation, un nombre dont les entiers font partie est un concept dans lequel nous devons distinguer sa valeur n qui intervient dans les calculs arithmétiques de son écriture formelle, celle qui interviendra dans les formules que nous noterons n⁵. Il existe plusieurs écritures formelles d'un même nombre : 10 en décimal, X en romain, 3130 en ASCII, sssssssss0 généralement utilisée dans un système formel⁶.

Ainsi formalisée, l'arithmétique peut être incluse dans un système formel.

5. si le gras n'apparaît pas, le lecteur saura faire la différence entre l'écriture qui apparaît dans les formules et la valeur utilisée dans les calculs.

6. L'ensemble des entiers naturels N est un ensemble ordonné infini, qui possède cette propriété remarquable que ses éléments peuvent être obtenus en se donnant le premier 0 et la fonction s (successeur) qui à tout entier associe le suivant : $s(0)=1$, $s(1)=ss(0)=2$, $s(2)=sss(0)=3$,... Plus généralement un entier quelconque n s'écrira s...(n fois s)...s0

2. Le Théorème d'incomplétude de Gödel

Si un système formel contenant les opérations arithmétiques élémentaires n'est pas contradictoire alors il est incomplet.

Disons au préalable que l'astuce de Gödel consiste à faire en sorte que tout "raisonnement" à l'intérieur du système formel se ramène à des calculs arithmétiques. C'est pour cette raison que le système évoqué doit nécessairement formaliser les opérations arithmétiques de base.

Pour appréhender l'ensemble de la démarche, énumérons les étapes de la démonstration :

1 - Gödel montre comment construire dans son système une formule arithmétique G qui exprime la phrase métamathématique : " G n'est pas démontrable ".

2 - Il montre alors que G est démontrable si sa négation non G est démontrable. Si le système produit G , il produit non G et réciproquement \Rightarrow il est contradictoire.

3 - Or si comme il se doit l'arithmétique n'est pas contradictoire, par négation de l'implication précédente G et non G ne sont pas des théorèmes $\Rightarrow G$ est indécidable.

4 - Gödel montre que pourtant G est vraie ; vraie au sens où G dit qu'un entier possède une certaine propriété arithmétique que l'on peut définir et vérifier par ailleurs.

5 - Puisque la formule G est vraie mais indécidable, le système est incomplet.

A travers l'arithmétique c'est bien évidemment toutes les Mathématiques qu'il devenait impossible de formaliser sans contradiction. Nous voilà maintenant en mesure d'approfondir les grandes lignes de la démonstration du théorème d'incomplétude.

2.1. La numérotation de Gödel. Le fait d'assigner des nombres à des formules réduit la logique du raisonnement à des calculs arithmétiques.

- La première étape assigne à chaque symbole du système formel un nombre non nul différent : n_1, n_2, \dots, n_k .

- Puis Gödel affecte à chaque formule un nombre qui est le produit des premiers nombres premiers élevés à la puissance du nombre représentant les symboles qui y figurent : $2^{n_1} \times 3^{n_2} \times \dots \times p^{n_k}$

où $2, 3, \dots, p$ est la suite ordonnée des nombres premiers.

- Selon un procédé identique, Gödel affecte un nombre à chaque suite de formules (démonstration) : $2^{m_1} \times 3^{m_2} \times \dots \times q^{m_s}$ où $2, 3, \dots, q$ est la suite ordonnée des nombres premiers et m_1, m_2, \dots, m_s sont les nombres de Gödel des formules de la suite.

Du fait de l'unicité de la décomposition d'un nombre en produit de facteurs premiers, cette méthode établit un codage bi-univoque entre les expressions du système formel et un sous-ensemble de \mathbb{N} . Le codage est ainsi possible dans tout système contenant l'arithmétique et c'est justement parce qu'il est possible qu'on rencontrera là aussi l'indécidabilité.

Inversement un nombre étant donné, on peut déterminer si c'est un nombre de Gödel⁷ et si le nombre est bien de Gödel, on peut retrouver exactement l'expression qu'il code :

- un symbole,

7. Ce n'est pas toujours le cas. $8800 = 2^5 \times 5^2 \times 11^1$ n'est pas un nombre de Gödel car la suite des nombres premiers n'est pas respectée ; il manque les nombres premiers 3 et 7.

- une formule,
- une démonstration

Gödel plonge ainsi le système formel initial dans un codage numérique. L'important est de comprendre que tout est nombre, dans cette approche gödelienne des mathématiques. Les nombres peuvent être considérés comme les atomes au sens où tout le reste peut être construit à partir d'eux. Pour exemple appliquons la numérotation de Gödel au système formel du paragraphe § 1.4.

A chaque symbole du système, on fait correspondre un nombre : M → 1 I → 2 U → 3

Ainsi, pour une formule on a un nombre de Gödel :

l'axiome MI → $2^1 \times 3^2 = 18$, la formule MU → $2^1 \times 3^3 = 54, \dots$

La succession des règles 223 sur l'axiome génère les formules MI, MII, MIII, MIU dont les nombres de Gödel respectifs sont :

$$18, \quad 2^1 \times 3^2 \times 5^2 = 450, \quad 2^1 \times 3^2 \times 5^2 \times 7^2 \times 11^2 = 2668050, \quad 2^1 \times 3^2 \times 5^3 = 2250$$

Le nombre de Gödel de cette démonstration est donc :

$$2^{18} \times 3^{450} \times 5^{2668050} \times 7^{2250}$$

Inversement, ce nombre étant donné, la décomposition en produit de nombres premiers donnera la suite unique des nombres de Gödel des formules qui, décomposés eux-même en produit de nombres premiers, donneront par décodage des symboles l'écriture typographique de ces formules. De même que l'ensemble de règles typographiques engendre des théorèmes, le codage de Gödel engendre des nombres entiers ; le passage de la typographie au nombre est biunivoque.

La question initiale : "MU est-elle un théorème" ? devient : " 54 est-il un nombre que peut produire le système formel" ?

2.2. L'arithmétique : langage universel pour l'étude des systèmes formels. L'étape suivante de Gödel est une utilisation ingénieuse de l'arithmétique pour vérifier des propriétés métamathématiques. Puisque toutes les formules ou suites de formules du système formel se trouvent codées par un nombre (de Gödel), il sera possible d'exprimer une propriété d'une formule du système par une simple formule arithmétique sur les codes ; les relations métamathématiques se réduisent ainsi à des opérations arithmétiques. Par exemple : la formule " MIII contient la formule MII" est une relation métamathématique entre deux formules du système. Cela revient à dire que 450 (nombre de Gödel de MII) est un diviseur de 2668050 (nombre de Gödel de MIII). Un simple calcul arithmétique permettra de le démontrer ou de le réfuter. Certains énoncés métamathématiques concernant des propriétés des formules du système formel pourront être exprimés et démontrés à l'intérieur du système lui-même.

Plus généralement, considérons une propriété arithmétique $P(x)$: "x à la propriété P". Le sujet x est destiné à être remplacé par un entier quelconque n pour vérifier dans quels cas $P(n)$ est vraie et affirmer en conséquence que $P(n)$ est un théorème.

Peut-on remplacer x par F ? Notre système formel ne traite que de propriétés arithmétiques c'est à dire des propriétés sur les entiers. Le sujet x ne peut être qu'un entier ;

par exemple $P(4)$: "4 est pair"⁸. La formule F est une suite de symboles et non pas un entier. En associant le nombre de Gödel f à la formule F , $P(f)$: "f est pair"⁹ devient décidable dans le système. Elle exprime : "F est pair". Même si cela n'a pas grand sens, en arithmétique la formule $P(f)$ est correcte.

Pour procéder à cette substitution de façon automatique dans le corps d'une démonstration par exemple, Gödel définit une fonction $\text{sub}(p,f)$ de la façon suivante :

- si p n'est pas un nombre de Gödel $\Rightarrow \text{sub}(p,f)=0$,

- si p est le nombre de Gödel d'une formule ou d'une propriété P de sujet x , on substitue f aux x . On obtient la formule $P(f)$ dont on calcule le nombre de Gödel z et on pose $\text{sub}(p,f)=z$.

Aucun doute $\text{sub}(p,f)$ est calculable et arithmétique. Si on considère un triplet d'entiers k,m,n on peut calculer $\text{sub}(k,m)$ pour savoir s'il est égal ou non à n . Selon l'encadré (3) on peut exprimer dans le système la substitution par une propriété décidable notée $\text{SUB}(x,y,z)$ telle que :

$\text{SUB}(p,f,z)$ construit la formule $P(f)$ de nombre de Gödel $z=\text{sub}(p,f)$. (4)

Si P est une propriété arithmétique alors la substitution transforme une phrase métamathématique en une formule arithmétique sur les entiers.

2.3. Autoréférence. Si on veut reconduire le paradoxe du menteur G dit : "G est menteur" dans le système formel, il faut que la phrase G détermine une propriété P et se l'applique à elle-même. En effet, si G dit : "F est menteur" :

- si G dit vrai \Rightarrow ce que dit F est faux,

- si G dit faux \Rightarrow ce que dit F est vrai.

Il n'y a pas de paradoxe. L'autoréférence semble nécessaire au paradoxe. Appliquons-la à la propriété arithmétique de parité $P(x)$: "x est pair" de nombre de Gödel p . Cela donne " $P(x)$ est pair" ou encore "x est pair" est pair" qui n'a pas grand sens. Par contre, comme nous l'avons vu précédemment " p est pair" en a un car p est un entier ; c'est vrai ou faux. Ainsi on peut exprimer dans l'arithmétique formelle $P(p)$: " p est pair" mais disons tout de suite que cela ne crée pas l'autoréférence recherchée à cause justement de la rigueur du calcul arithmétique. En effet, substituer p à x calcule pour la formule " p est pair" un nouveau nombre de Gödel qui ne peut pas être p . Plus précisément pour que $P(x)$ se réfère à elle-même, il faut remplacer les x de $P(x)$ par p . On obtient par substitution $\text{SUB}(p,p,z)$ la formule $P(p)$: " p est pair" de nombre de Gödel $z = \text{sub}(p,p)$. Pour avoir l'autoréférence il faudrait que $z = p$ c'est à dire $p = \text{sub}(p,p)$. C'est impossible car $P(x)$ ne se réduit pas au seul sujet x . On ne peut donc pas exprimer une autoréférence directe.

Oublions un temps l'autoréférence. Dans la phrase du menteur, remplaçons G par sa citation (la phrase entre-guillemets). On obtient G : "'G est menteur" est menteur" ou encore G : "Z est menteur" si on appelle Z la citation de G . Attention, cette nouvelle

8. 4 est l'écriture formelle du nombre " quatre ". Cette écriture pourrait être ssss0 selon la note 6. La formule $P(4)$: " 4 est pair " s'écrira alors dans le système formel " $\exists e \text{ss}0.e=\text{ssss}0$ "

9. f est l'écriture formelle du nombre de Gödel f de la formule F

phrase Z : "G est menteur" n'a pas le même sens que la phrase initiale. Z est ici une phrase et non pas une personne comme G l'est dans la phrase du menteur ; nous y reviendrons.

Pour l'heure on peut considérer que :

- G donne la propriété " est menteur " à une phrase Z,
- Z donne la même propriété à la personne G.

Ainsi dans un premier temps G ne parle pas d'elle-même mais d'une phrase Z. Comme il s'avère que Z cite G \Rightarrow G parlera d'elle-même. De plus si on ne fait pas de différence entre personne et phrase, on arrive indirectement à l'autoréférence à partir d'une seule phrase génératrice $G(x)$ qui attribue la propriété " est menteur " à un sujet x . Quelle est cette phrase génératrice $G(x)$? Comment obtient-on la phrase autoreproductrice G ?

G : " "G est menteur" est menteur " est obtenue en substituant G au sujet x de $G(x)$ qui construit Z que l'on qualifie de menteur.

Dans un système formel si k est le nombre de Gödel de $G(x)$, cette dernière phrase G serait obtenue par la formule programme : "SUB(k, k, z) et P(z)". En effet SUB(k, k, z) substitue k au sujet x de $G(x)$ et construit la formule $Z=G(k)$ de nombre de Gödel $z = \text{sub}(k, k)$. P(z) affirme que Z à la propriété P. Ce programme dit donc : "G(k) à la propriété P".

Pour qu'il y ait autoréférence, il faut que " SUB(k, k, z) et P(z) " = G(k). Ainsi la formule génératrice serait $G(x)$: "SUB(x, x, y) et P(y)". Vérifions que l'on atteint bien l'autoréférence.

En substituant k à x dans $G(x)$ on obtient à la formule $G(k)$: "SUB(k, k, y) et P(y)" qui à l'image de G construit la formule Y de nombre de Gödel $y = \text{sub}(k, k)$ et lui donne la propriété P. $G(k)$: "Y à la propriété P".

En remplaçant x par k dans $G(k)$, selon la définition de l'encadré (4), $G(k)$ a pour nombre de Gödel $\text{sub}(k, k)$. Or $\text{sub}(k, k)$ est unique et nous venons de voir que $y = \text{sub}(k, k)$ donc $Y=G(k)$. Ainsi :

$$G(k) : \text{"SUB}(k, k, y) \text{ et P}(y)\text{"} \Leftrightarrow G(k) : \text{"G}(k) \text{ à la propriété P}\text{"} ; k \text{ nombre de Gödel de } G(x). \text{ (5)}$$

$G(k)$ est une formule autoreproductrice, amenant l'autoréférence. Elle pourrait exprimer la phrase G précédente si la propriété " est menteur " était une propriété arithmétique ; ce n'est pas le cas. Mais Gödel trouvera le moyen de projeter dans l'arithmétique une propriété analogue à celle-ci : la propriété métamathématique : "n'est pas démontrable".

2.4. Paradoxe. Revenons à nouveau à la phrase du menteur. Nous constatons que le caractère paradoxal n'est pas seulement dû à l'autoréférence il dépend aussi du choix de la propriété. Si G dit : "G est Crétois", ce n'est pas paradoxal, c'est vrai ou faux suivant la nationalité de G. Pour reproduire le paradoxe du menteur il faut une phrase qui énonce qu'elle ment, ou encore qu'elle est fausse.

"G est menteur" est un énoncé sur un objet du langage. Il attribue une propriété P à l'objet G ; c'est la personne G qui ment.

Si on remplace G par sa citation comme précédemment, on a un énoncé sur un énoncé du langage. On est dans le métalangage : ““G est menteur” est menteur”. Ici c’est la citation qui ment. C’est peut-être du mauvais français, mais cela a un sens ; c’est faux de dire “G est menteur” ou encore G dit : ““ G est menteur” est fausse”. Donc G ne ment pas. On est effectivement en plein paradoxe !

On a attribué la même propriété de menteur à la personne G et à sa citation Z : “G est menteur”. Ainsi par substitution, le double effet amène la contradiction. C’est une autoréférence contradictoire.

S’il y a supercherie d’identité entre ces deux niveaux de lecture, il n’y en a pas lorsque tout est nombre. En attribuant un entier à chaque formule, Gödel a exprimé dans le langage de l’arithmétique une formule autoréférente paradoxale syntaxiquement correcte.

2.5. La formule autoréférente de Gödel. Revenons à la question métamathématique qui nous préoccupe depuis le début : la formule F est-elle un théorème du système formel ? F est un théorème ou encore F est démontrable s’il existe une démonstration D qui aboutisse à F. Cela peut être exprimé dans le système par la formule à une variable : “ $\exists d$ DEM(f,d)”. f est l’écriture formelle du nombre de Gödel de F, d est l’écriture formelle du nombre de Gödel de la démonstration D. Comme nous l’avons vu précédemment, au paragraphe § 1.6, DEM(f,d) est une propriété arithmétique décidable¹⁰ ; il suffit simplement de savoir si d et d+f sont premiers entre eux.

La formule autoréférente choisie par Gödel dit d’elle même qu’elle n’est pas démontrable ; G : “G n’est pas démontrable”. Pour obtenir son écriture formelle, partons de la formule autoréférente de l’encadré (5) ; G(k) : “G(k) à la propriété P”. Remplaçons :

- G(k) par G de nombre de Gödel g

- P est la propriété métamathématique : “n’est pas démontrable”, exprimée dans le système par la propriété arithmétique P(x) : “non($\exists d$ DEM(x,d))” On obtient en projection de la formule métamathématique G : “G n’est pas démontrable” dans un système formel, la formule arithmétique G : “non($\exists d$ DEM(g,d))”.

On vient de mélanger des propriétés purement arithmétiques et des formules contenant une référence numérique (le nombre de Gödel g) à la notation utilisée pour formuler ces propriétés arithmétiques. Plus précisément, la propriété “n’est pas démontrable” ne fait pas partie de l’arithmétique pure (celle qui ne fait référence à aucun système de notation). C’est parce que ces notations sont codées par les nombres de Gödel qu’on peut leur appliquer de l’arithmétique. Le code permet de projeter des énoncés métamathématiques en formules arithmétiques à l’intérieur du calcul lui-même. Dans la construction précédente, g est associé à une formule arithmétique du système. La question posée est de savoir si un nombre g possède une certaine propriété arithmétique P(x). Donc la formule G se trouvent bien à l’intérieur de l’arithmétique. Il n’y a pas supercherie !

Codage des formules et projection dans l’arithmétique constituent la clé du raisonnement élaboré par Gödel.

10. Il faut bien faire la distinction avec la formule précédente : “ $\exists d$ DEM(f,d)” qui est arithmétique mais ne sera démontrable que si elle est vraie selon l’encadré (2), car si on ne sait pas que le d existe la recherche peut être sans fin.

2.6. Une contradiction dans l'arithmétique. Gödel montre alors que la formule arithmétique G est démontrable \Leftrightarrow non G est démontrable.

1 – si G est démontrable \Rightarrow il existe une démonstration de G et donc un nombre d tel que $DEM(\mathbf{g}, \mathbf{d})$ est vrai. Donc " $\exists \mathbf{d} DEM(\mathbf{g}, \mathbf{d})$ " vrai \Rightarrow la formule " $\exists \mathbf{d} DEM(\mathbf{g}, \mathbf{d})$ " est démontrable selon l'encadré (2) \Rightarrow non[non($\exists \mathbf{d} DEM(\mathbf{g}, \mathbf{d})$)] est démontrable \Rightarrow non G est démontrable. Ainsi, si G est démontrable, sa négation l'est aussi ; le système formel est contradictoire.

2 – si non G est démontrable \Rightarrow " $\exists \mathbf{d} DEM(\mathbf{g}, \mathbf{d})$ " est démontrable. Cela signifie comme nous l'avons vu dans l'analyse des formules existentielles c.f. § 1.6 qu'en remplaçant d par $0, 1, 2, \dots$ on trouvera un nombre d tel que d et $d+g$ soient premiers entre eux, c'est-à-dire une démonstration de G . Mais si on le pense, on ne peut pas le démontrer car la recherche de la démonstration (le nombre d) peut être sans fin. On est dans une situation inconfortable où non G démontrable \Rightarrow G sera démontrable, mais on ne sait pas quand... La contradiction existe, sans pouvoir la démontrer. C'est ce que Gödel a appelé la *w*-inconsistance.

Ainsi Gödel ne démontre pas comme on le dit souvent l'inconsistance absolue de l'arithmétique. Il dit simplement qu'un système formel contenant de l'arithmétique n'est pas parfait car on peut écrire des contradictions ou bien si comme il se doit l'arithmétique n'est pas contradictoire il existe des problèmes qui ne peuvent être résolus. En effet, par négation de l'implication précédente, G et non G ne sont pas démontrables \Rightarrow G est indécidable.

2.7. Incomplétude ou contradiction. Il n'y a rien de remarquable à construire dans l'arithmétique une formule indécidable, sauf si c'est une vérité. Philosophiquement, c'est beaucoup plus frustrant d'avoir des vérités indémonstrables que des problèmes non résolus. Et pourtant la formule arithmétique G est vraie ! Cette vérité est établie grâce à un raisonnement métamathématique et une projection dans le formalisme de l'arithmétique et non pas en la déduisant formellement des axiomes du système. Si on admet que l'arithmétique n'est pas contradictoire, on vient de voir que le système ne peut pas démontrer G . Donc la formule " G n'est pas démontrable" est bien vraie ; c'est le raisonnement métamathématique établi dans l'exemple de l'ordinateur, au début de l'exposé dans le paragraphe § 1.3. Cette phrase métamathématique est exprimée à l'intérieur de l'arithmétique et la raison d'être de la projection étant que les formules métamathématiques vraies sont projetées dans le système par des formules arithmétiques vraies : G est donc vraie.

Au final Gödel a construit dans un système formalisant l'arithmétique non contradictoire une formule vraie et indécidable rendant ce système formel incomplet.

En allant un peu plus loin, Gödel montrera que même en posant comme axiome que G est vraie, on peut encore trouver une formule vraie qui n'est pas démontrable, signifiant qu'il est inutile d'ajouter des axiomes supplémentaires, le système reste incomplet. Ainsi la réalité restera toujours plus importante que l'ensemble des connaissances formalisables et on peut prétendre des choses vraies sans avoir à les justifier.

3. Bibliographie

Cette étude s'appuie essentiellement sur les deux livres suivants :

1 - *Le théorème de Gödel* [GNNG89] de E. Nagel, J.-R. Newman, K. Gödel, J.-Y. Girard – Editions du Seuil. 1989

Ce livre comprend :

* traduit de l'anglais l'ouvrage de Ernest Nagel & James R. Newman de 1958, "la démonstration de Gödel",

* traduit de l'allemand l'article de Gödel de 1931, "Sur les propositions formellement indécidables des Principia Mathematica et des systèmes apparentés",

* un commentaire de Jean-Yves Girard, "le champ du signe ou la faillite du réductionnisme".

2 - *Gödel Escher Bach* [Hof85] : les brins d'une guirlande éternelle de D.Hofstadter – Interéditions. 1985

* Cette étude est aussi inspirée par de nombreux articles que l'on trouve sur Internet avec les mots clés suivants : Théories et systèmes formels - Propositions indécidables - Arithmétique et complétude - Logique et paradoxes - Autoréférence - Le programme de Hilbert - ...

RÉFÉRENCES

[GNNG89] K. Gödel, E. Nagel, J. Newman, and J.-Y. Girard. *Le Théorème de Gödel*. Seuil, 1989.

[Hof85] D. Hofstadter. *Gödel Escher Bach. Les brins d'une guirlande éternelle*. Interéditions, 1985.

100 ans de Relativité Générale (II/III) : courbure

Eric OLIVIER^{1 2}

Résumé. – Riemann généralise la notion de courbure d'une surface introduite par Gauss pour des objets géométriques différentiels de dimension quelconque (le concept de variété riemannienne n'existe pas encore!). Pour cela il introduit ce qui sera par la suite appelé le tenseur de courbure de Riemann (ou de Riemann-Christoffel). Dans le sillage de Gauss et de Riemann, les travaux de Christoffel, Ricci et Levi-Civita, ont été déterminants dans la réalisation du projet d'Einstein qui l'a mené de la Relativité Restreinte à la Relativité Générale et où la courbure de l'espace-temps remplace – en un sens – la *force* de gravitation newtonnienne.

1. Introduction

Étant donnée (M, g) une variété pseudo-riemannienne de dimension n , le *tenseur de Riemann*³ est un tenseur de valence $(1, 3)$ qui généralise, en dimension quelconque la notion de courbure introduite par Gauss pour les surfaces. En coordonnées, les composantes $R^\tau{}_{\sigma\mu\nu}$ du tenseur de Riemann possèdent (au signe prêt) une unique contraction non nulle de l'indice contravariant : ici, nous retrouvons notre sujet, puisque le *tenseur de Ricci* de composantes $R_{\mu\nu} := R^\epsilon{}_{\mu\epsilon\nu}$ joue un rôle central dans la Relativité Générale. En particulier (ce sera le point de départ de [Oli15b, § 9] sur la déflexion des rayons lumineux) la métrique de Schwarzschild est déterminée par une résolution de l'équation d'Einstein dans le vide i.e.

$$R_{\mu\nu} = 0$$

pour un espace-temps lorentzien à symétrie (spatiale) sphérique. Dans cette note, nous commençons par introduire les tenseurs de torsion et de courbure d'un point de vue général (définition de Koszul), lorsque (M, g) est munie d'une connexion affine à priori quelconque. Cela nous permet de raffiner les propriétés de symétrie du tenseur de courbure en fonction des propriétés spécifiques de la connexion. La connexion de Levi-Civita est déterminée par son absence de torsion et sa compatibilité métrique (Théorème d'unicité de Koszul) : elle est essentiellement liée à la Relativité Générale du fait qu'elle détermine les géodésiques du lagrangien inertiel (Théorème de Levi-Civita : c.f. [Oli15a, Théorème 14.1]). Le tenseur de Riemann proprement dit correspond au tenseur de courbure de la connexion de Levi-Civita : conséquence de cette spécificité, celui-ci vérifie (en particulier) un ensemble de quatre identités de symétrie (deux propriétés d'antisymétrie, la première identité de Bianchi et la bisymétrie) que nous regarderons en détail. Il est essentiel pour l'équation d'Einstein de la Relativité Générale (c.f. § 7), que le tenseur de Ricci soit symétrique en ce sens que

$$R_{\mu\nu} = R_{\nu\mu}$$

1. GDAC-I2M UMR 7373 CNRS Université d'Aix-Marseille
2. eric.olivier@univ-amu.fr
3. Einstein parle du tenseur de *Riemann-Christoffel*.

cette symétrie découlant elle-même de la symétries du tenseur métrique et de la bisymétrie du tenseur de Riemann. (Parmi l'ensemble des tenseurs de courbure, seul le tenseur de Riemann vérifie la bisymétrie : voir la carte mentale de la Figure 2.)

Un point important abordé porte sur l'interprétations géométrique du tenseur de Riemann. Une rapide présentation du concept de transport parallèle nous amènera au Théorème de Levi-Civita – déjà mentionné – et caractérisant les géodésiques inertielles par un système différentiel déduit du transport parallèle. Nous serons alors capable de retrouver l'expression des composantes $R^T_{\sigma\mu\nu}$ par le transport parallèle d'un champ de vecteur le long d'un circuit infinitésimal de géodésiques. Enfin, nous terminerons cette note par les origines historiques du tenseur de Riemann, en revisitant la courbure gaussienne dans les termes de la théorie moderne des tenseurs.

Remerciement : Je tiens à remercier chaleureusement John Hubbard : nos nombreuses discussions sur le tenseur de courbure (entre autres), m'ont grandement éclairé/guidé pour la rédaction de cette deuxième note sur la Relativité Générale. Un regret : faute de place, je n'ai pu intégrer beaucoup de ses remarques, essentiellement sur les idées originales de Gauss et Riemann, et qui sont difficiles à trouver dans la littérature : j'espère qu'elles pourront être publiées par ailleurs, sous une forme ou sous une autre.

2. Structures algébriques du tenseur de Riemann

Dans ce paragraphe, $\mathcal{S}(\mathbb{R}^d)$ et $\mathcal{S}_B(\mathbb{R}^d)$ désignent deux sous-espaces de l'espace vectoriel des formes quadrilinéaires définies sur \mathbb{R}^d ($d \geq 2$). Plus précisément, étant donnée $\varphi : \times_1^4 \mathbb{R}^d \rightarrow \mathbb{R}$ une forme quadrilinéaire, nous considérons les propositions suivantes :

- (S1) : $\varphi(u, v, w, z) = -\varphi(v, u, w, z)$ (Première antisymétrie)
- (S2) : $\varphi(u, v, w, z) = -\varphi(u, v, z, w)$ (Deuxième antisymétrie)
- (S3) : $\varphi(u, v, w, z) + \varphi(w, u, v, z) + \varphi(v, w, u, z) = 0$ (Identité de Bianchi)
- (S4) : $\varphi(u, v, w, z) = \varphi(w, z, u, v)$ (Bisymétrie)

Alors, l'espace vectoriel $\mathcal{S}(\mathbb{R}^d)$ (resp. $\mathcal{S}_B(\mathbb{R}^d)$) est formé des formes quadrilinéaires φ définies sur E et satisfaisant les propriétés (S1)-(S4)⁴ (resp. (S1)-(S2)-(S3)).

Proposition 2.1. $\mathcal{S}_B(\mathbb{R}^d)$ est un sous-espace vectoriel de $\mathcal{S}(\mathbb{R}^d)$, en ce sens que

$$\left((S1) \ \& \ (S2) \ \& \ (S3) \right) \implies (S4)$$

Preuve (Milnor). Soit O le point de \mathbb{R}^3 de coordonnées $(0, 0, 0)$; pour tout point A de \mathbb{R}^3 , nous notons A' l'image de A dans la symétrie centrale de centre O (avec la convention que $A'' = A$). Nous considérons alors que A, B, C, A', B', C' sont les sommets d'un octaèdre (régulier) de \mathbb{R}^3 de centre O et que A, B, C sont les trois sommets d'une même face de l'octaèdre. Étant donné $\varphi \in \mathcal{S}_B(\mathbb{R}^d)$ et $(u, v, w, z) \in \times_1^4 \mathbb{R}^d$, chaque sommet – disons H – de l'octaèdre est associé au réel $\rho(H)$ correspondant à la valeur de φ sur une permutation (a, b, c, d) du quadruplet (u, v, w, z) et déterminée par les deux conditions :

- (C1) : $(\rho(A), \rho(B), \rho(C)) = (\varphi(u, v, w, z), \varphi(w, u, v, z), \varphi(v, w, u, z))$
- (C2) : $\rho(H) = \varphi(a, b, c, d) \iff \rho(H') = \varphi(c, d, a, b)$

4. Il est évident que (S1) et (S4) entraîne (S2).

(voir Fig. 1). Par définition de la fonction ρ , l'identité de Bianchi (S3) entraîne $\rho(A) + \rho(B) + \rho(C) = 0$. Mais d'autre part, en utilisant (S1) et (S2), nous pouvons aussi écrire

$$\begin{aligned} \rho(A) + \rho(B') + \rho(C') &= \varphi(u, v, w, z) + \varphi(v, z, w, u) + \varphi(u, z, v, w) \\ &= -\varphi(u, v, z, w) - \varphi(v, z, u, w) - \varphi(z, u, v, w) \end{aligned}$$

soit encore $\rho(A) + \rho(B') + \rho(C') = 0$, en utilisant (S3) une nouvelle fois. Finalement

$$(1) \quad 2\rho(A) + \rho(B) + \rho(C) + \rho(B') + \rho(C') = 0$$

et comme de même

$$(2) \quad 2\rho(A') + \rho(B) + \rho(C) + \rho(B') + \rho(C') = 0$$

nous obtenons $\rho(A) = \rho(A')$ en retranchant (1) et (2).

□

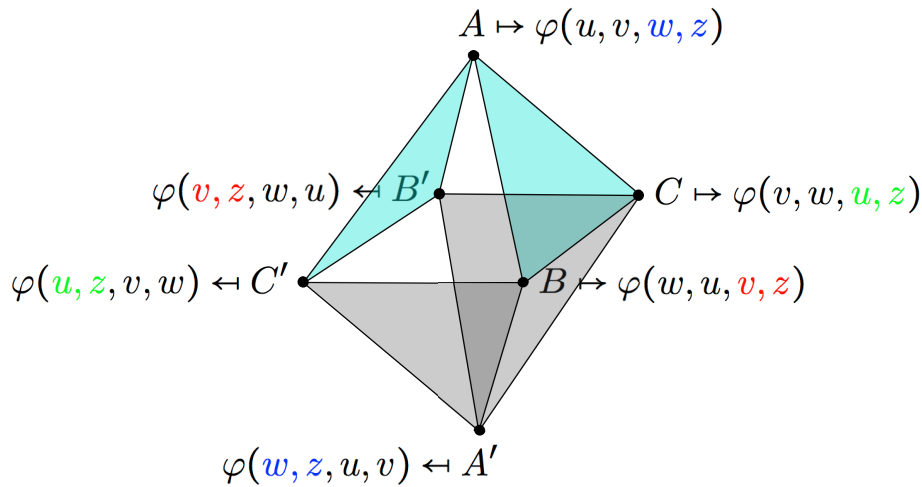


FIGURE 1. Diagramme de Milnor.

Soit (e_1, \dots, e_d) la base canonique de \mathbb{R}^d ($d \geq 2$) et notons $\varphi_{\sigma\tau\mu\nu} := \varphi(e_\sigma, e_\tau, e_\mu, e_\nu)$. Pour commencer, l'égalité $\mathcal{S}_B(\mathbb{R}^d) = \mathcal{S}(\mathbb{R}^d)$ est satisfaite lorsque d est égal à 2 ou à 3 ; en effet, pour tout ⁵ $1 \leq \mu, \nu, \sigma \leq 3$ nous avons :

$$\begin{aligned} 3\varphi_{[\mu\nu]\sigma} &= 3\varphi_{[\nu\mu]\sigma} = 3\varphi_{[\nu\mu\mu]\sigma} = \varphi_{\nu\mu\mu\sigma} + \varphi_{\mu\mu\nu\sigma} + \varphi_{\mu\nu\mu\sigma} = 0 \\ 3\varphi_{[\mu\nu\sigma]\mu} &= 3\varphi_{[\nu\sigma\mu]\mu} = 3\varphi_{[\sigma\mu\nu]\mu} = \varphi_{\sigma\mu\nu\mu} + \varphi_{\mu\nu\sigma\mu} + \varphi_{\nu\sigma\mu\mu} = 0 \end{aligned}$$

d'où $\varphi_{[\mu\nu\sigma]\tau} = 0$ pour tout $1 \leq \mu, \nu, \sigma, \tau \leq 3$: l'égalité $\mathcal{S}_B(\mathbb{R}^d) = \mathcal{S}(\mathbb{R}^d)$, est donc valable pour $d = 2, 3$. Il se trouve que l'inclusion $\mathcal{S}_B(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d)$ est stricte pour $d > 3$. Par exemple pour $d = 4$, notons $(e_1^*, e_2^*, e_3^*, e_4^*)$ la base duale de la base canonique de \mathbb{R}^4 (i.e. $e_\mu^*(e_\nu) = \delta_{\mu\nu}$), de sorte que $e_\mu^* \wedge e_\nu^*(u, v)$ (produit extérieur) désigne le mineur d'indice (μ, ν) du couple de vecteurs (u, v) . Alors, il est facile de vérifier que

$$\varphi := (e_1^* \wedge e_2^*) \otimes (e_3^* \wedge e_4^*) + (e_3^* \wedge e_4^*) \otimes (e_1^* \wedge e_2^*)$$

est bien dans $\mathcal{S}(\mathbb{R}^4)$: or $\varphi \notin \mathcal{S}_B(\mathbb{R}^4)$, puisque $\varphi_{[123]4} = \varphi_{1234} = 1$.

5. Soit $(x_1, \dots, x_k) \mapsto f(x_1, \dots, x_k)$ soit une fonction numérique de k indéterminées x_1, \dots, x_n : alors $f([x_1, \dots, x_k]) = 1/k \sum_\sigma f(x_{\sigma(1)}, \dots, x_{\sigma(k)})$, où σ est une permutation circulaire de $\{1, \dots, k\}$.

Remarque 2.2. (1) : Soit $\varphi \in \mathcal{S}(\mathbb{R}^2)$: alors pour $(u, v, w, z) \in \times_1^4 \mathbb{R}^2$,

$$\begin{aligned}\varphi(u, v, w, z) &= u^1 v^2 w^1 z^2 \varphi_{1212} + u^2 v^1 w^1 z^2 \varphi_{2112} + u^1 v^2 w^2 z^1 \varphi_{1221} + u^2 v^1 w^2 z^1 \varphi_{2121} \\ &= (u^1 v^2 w^1 z^2 - u^2 v^1 w^1 z^2 - u^1 v^2 w^2 z^1 + u^2 v^1 w^2 z^1) \varphi_{1212} \\ &= (u^1 v^2 - u^2 v^1)(w^1 z^2 - w^2 z^1) \varphi_{1212}\end{aligned}$$

Nous en déduisons que l'espace vectoriel $\mathcal{S}(\mathbb{R}^2)$ est de dimension 1 et que tout $\varphi \in \mathcal{S}(\mathbb{R}^2)$ s'écrit

$$(3) \quad \varphi = (\mathbf{e}_1^* \wedge \mathbf{e}_2^*) \otimes (\mathbf{e}_1^* \wedge \mathbf{e}_2^*) \varphi_{1212}$$

où $\mathbf{e}_1^* \wedge \mathbf{e}_2^*$ est le déterminant sur \mathbb{R}^2 . Dans le cas où $\varphi \in \mathcal{S}(\mathbb{R}^d)$ pour $d \geq 2$, nous utilisons – de même que précédemment – la quadrilinéarité combinée aux symétries (S1) et (S4) pour obtenir :

$$\begin{aligned}\varphi &= \sum_{1 \leq \mu < \nu \leq d} (\mathbf{e}_\mu^* \wedge \mathbf{e}_\nu^*) \otimes (\mathbf{e}_\mu^* \wedge \mathbf{e}_\nu^*) \varphi_{\mu\nu\mu\nu} + \\ &\quad \sum_{1 \leq \mu < \nu \leq \sigma < \tau \leq d} \left((\mathbf{e}_\mu^* \wedge \mathbf{e}_\nu^*) \otimes (\mathbf{e}_\sigma^* \wedge \mathbf{e}_\tau^*) + (\mathbf{e}_\sigma^* \wedge \mathbf{e}_\tau^*) \otimes (\mathbf{e}_\mu^* \wedge \mathbf{e}_\nu^*) \right) \varphi_{\mu\nu\sigma\tau}\end{aligned}$$

ce qui nous permet d'obtenir la dimension de $\mathcal{S}(\mathbb{R}^d)$, soit :

$$\dim(\mathcal{S}(\mathbb{R}^d)) = \frac{1}{2} \left(\frac{d(d-1)}{2} \right) \left(\frac{d(d-1)}{2} + 1 \right)$$

On pourra trouver dans [Wei72, Chap. 6, § 7], la formule

$$\dim(\mathcal{S}_B(\mathbb{R}^d)) = \frac{1}{12} d^2 (d^2 - 1)$$

(2) : Soit $(u, v, w) \mapsto \det(u, v, w) = \mathbf{e}_1^* \wedge \mathbf{e}_2^* \wedge \mathbf{e}_3^*(u, v, w)$ le déterminant sur \mathbb{R}^3 . Alors, pour $u, v \in \mathbb{R}^3$ fixés, l'application $\det(u, v, \cdot)$ est une forme linéaire sur \mathbb{R}^3 : par définition, le produit vectoriel de u et v est le vecteur noté $u \times v$ t.q. $\det(u, v, w) = \langle u \times v | w \rangle$ (identité du produit mixte). Notons que si $\text{Vec}(u, v)$ est l'endomorphisme linéaire de \mathbb{R}^3 t.q. $\text{Vec}(u, v)w = (u \times v) \times w$, alors :

- (i) : $\text{Vec}(u, v) = -\text{Vec}(v, u)$ (Antisymétrie du produit vectoriel)
- (ii) : $\text{Vec}(u, v)w + \text{Vec}(w, u)v + \text{Vec}(v, w)u = 0$ (Identité de Jacobi)

Considérons maintenant l'application $(u, v, w, z) \mapsto \mathcal{V}(u, v, w, z)$, t.q.

$$\mathcal{V}(u, v, w, z) := \langle \text{Vec}(u, v)w | z \rangle = \det(\text{Vec}(u, v), w, z)$$

D'après (i) et du fait que le déterminant est une forme trilinéaire alternée, nous avons

$$\mathcal{V}(u, v, w, z) = -\mathcal{V}(v, u, w, z) \quad \text{et} \quad \mathcal{V}(u, v, w, z) = -\mathcal{V}(u, v, z, w)$$

Or, en prenant le produit scalaire par un vecteur z sur l'identité de Jacobi en (ii), nous obtenons :

$$\mathcal{V}(u, v, w, z) + \mathcal{V}(w, u, v, z) + \mathcal{V}(v, w, u, z) = 0$$

ce qui signifie que $\mathcal{V} \in \mathcal{S}_B(\mathbb{R}^3)$. Nous savons qu'il y a égalité entre $\mathcal{S}_B(\mathbb{R}^3)$ et $\mathcal{S}(\mathbb{R}^3)$: par suite \mathcal{V} satisfait la bisymétrie $\mathcal{V}(u, v, z, w) = \mathcal{V}(z, w, u, v)$. Notons que (dans le cas particulier du produit vectoriel) l'identité de Bianchi ainsi que la bisymétrie de \mathcal{V} peuvent être déduites (indépendamment et directement) de l'identité Gibbs, i.e. $\text{Vec}(u, v)w = \langle u | w \rangle v - \langle v | w \rangle u$.

3. Tenseurs de torsion et de courbure d'une connexion

Dans ce paragraphe, nous donnons la présentation formelle de Koszul des tenseurs de torsion et courbure d'une connexion (c.f. [Kos50]). Étant donnée (M, g) une variété pseudo-riemannienne de dimension n , rappelons (c.f. [Oli15a]) qu'un (champ de) tenseurs de type (p, q) sur M est une section C^∞ du fibré vectoriel

$$\coprod_X \otimes^{p,q} T_X M = \coprod_X \left(\otimes_1^p T_X M \right) \otimes \left(\otimes_1^q T_X^* M \right)$$

où les p -premiers produits tensoriels portent sur p -copies de l'espace $T_X M$ tangent à M en X et les q -derniers produits tensoriels sont sur q -copies de l'espace cotangent en X (i.e. le dual $T_X^* M$) : nous notons $\mathfrak{T}_q^p(TM)$ l'espace vectoriel des (champs de) tenseurs de type (p, q) sur M . Si les $\partial_i = \partial/\partial x^i$ sont les vecteurs de la base holonôme (base locale des champ de vecteurs) d'un système $\mathbf{x} = (x^1, \dots, x^n)$ de coordonnées locales, alors les dx^i sont les co-vecteurs de la base duale (base locale des formes différentielles), de sorte que $dx^\mu(\partial_\nu) = \partial_\nu(dx^\mu) = \delta^\mu_\nu$ (symbole de Kronecker). Ainsi, la décomposition (en coordonnées) d'un tenseur $F \in \mathfrak{T}_q^p(TM)$ s'écrit (avec les notations sommatoires d'Einstein)

$$F = F^{\mu_1 \dots \mu_p}_{\nu_1 \dots \nu_q} \frac{\partial}{\partial x^{\mu_1}} \otimes \dots \otimes \frac{\partial}{\partial x^{\mu_p}} \otimes dx^{\nu_1} \otimes \dots \otimes dx^{\nu_q}$$

où les applications $X \mapsto F^{\mu_1 \dots \mu_p}_{\nu_1 \dots \nu_q}(X)$ définies sur le domaine de \mathbf{x} sont appelées les \mathbf{x} -composantes du tenseur F . L'espace $\mathfrak{T}_0^1(TM)$ (resp. $\mathfrak{T}_1^0(TM)$) s'identifie à l'espace des champs de vecteurs (resp. des formes différentielles) sur M ; par convention $\mathfrak{T}_0^0(TM)$ s'identifie à l'espace $C^\infty(M)$ des champs scalaires sur M . Nous supposons que (M, g) est munie d'une connexion (ou plus précisément une connexion affine) arbitraire \mathbf{D} : rappelons (c.f. [Oli15a, Définition 8.1]) que \mathbf{D} est une application \mathbb{R} -bilinéaire $(U, V) \mapsto \mathbf{D}_U(V)$ sur $\mathfrak{T}_0^1(TM)$ pour laquelle (en coordonnées) chaque application partielle $U \mapsto \mathbf{D}_U(V)$ est $C^\infty(M)$ -linéaire alors que chaque application $V \mapsto \mathbf{D}_U(V)$ satisfait la règle de Leibnitz i.e. pour tout champ de vecteur V et tout champ scalaire f

$$\mathbf{D}_U(fV) = U(f)V + f\mathbf{D}_U(V)$$

Les coefficients de \mathbf{D} dans un système de coordonnées locales \mathbf{x} sont les fonctions numériques $\Gamma^\sigma_{\mu\nu}$ définies uniquement dans le domaine de \mathbf{x} et telles que

$$\mathbf{D}_{\partial_\mu}(\partial_\nu) = \Gamma^\epsilon_{\mu\nu} \partial_\epsilon$$

Enfin, par définition des dérivations partielles covariantes D_1, \dots, D_n associées à \mathbf{D} :

$$(4) \quad \mathbf{D}_U(V) = U^\mu D_\mu(V^\sigma) \partial_\sigma \quad \text{où} \quad D_\mu(V^\sigma) = \partial_\mu(V^\sigma) + V^\epsilon \Gamma^\sigma_{\mu\epsilon}$$

Bien que d'un abord relativement simple, la notion de connexion demeure à priori mystérieuse. En fait il s'agit d'une formalisation algébrique de la notion géométrique – bien plus intuitive – de transport parallèle (c.f. § 12 et [Bou92] sur l'histoire des connexions et du transport parallèle). Cependant, un intérêt immédiat de cette notion, est de permettre des définitions assez claires des tenseurs de torsion et de courbure.

Proposition 3.1 (Tenseur de torsion). *Soit (M, g) une variété pseudo-riemannienne de dimension n muni d'une connexion \mathbf{D} . L'application bilinéaire $(U, V) \mapsto T(U, V)$ de $\mathfrak{T}_0^1(TM) \times \mathfrak{T}_0^1(TM) \times$ et à valeurs dans $\mathfrak{T}_0^1(TM)$ est définie en posant*

$$(5) \quad T(U, V) = \mathbf{D}_U(V) - \mathbf{D}_V(U) - [U, V].$$

alors (i) : l'application T est $C^\infty(M)$ -linéaire suivant ses deux entrées et (ii) : dans la base holonôme $(\partial_1, \dots, \partial_N)$ d'un système de coordonnées locales,

$$T(U, V) = U^\mu V^\nu T^\kappa_{\mu\nu} \partial_\kappa \quad \text{où} \quad T^\kappa_{\mu\nu} = \Gamma^\kappa_{\mu\nu} - \Gamma^\kappa_{\nu\mu}$$

(iii) : les $T^\kappa_{\mu\nu}$ sont les composantes d'un tenseur (1, 2) appelé *tenseur de torsion* de D .

Preuve. Les propositions (i) et (ii) se vérifient par calculs directs, la proposition (iii) étant une conséquence de (i) et (ii)⁶. □

Proposition 3.2 (Tenseur de courbure). Soit (M, g) une variété pseudo-riemannienne de dimension n muni d'une connexion D . L'application tri-linéaire $(U, V, W) \mapsto R(U, V, W)$ de $\mathfrak{T}_0^1(TM) \times \mathfrak{T}_0^1(TM) \times \mathfrak{T}_0^1(TM)$ et à valeurs dans $\mathfrak{T}_0^1(TM)$ est définie en posant

$$(6) \quad R(U, V, W) = D_U(D_V(W)) - D_V(D_U(W)) - D_{[U, V]}(W).$$

alors (i) : l'application R est $C^\infty(M)$ -linéaire suivant ses trois entrées et (ii) : dans la base holonôme $(\partial_1, \dots, \partial_N)$ d'un système de coordonnées locales

$$R(U, V, W) = U^\mu V^\nu W^\sigma R^\kappa_{\sigma\mu\nu} \partial_\kappa$$

(attention à l'ordre des indices) où les quantités

$$(7) \quad R^\kappa_{\sigma\mu\nu} = \partial_\mu(\Gamma^\kappa_{\nu\sigma}) - \partial_\nu(\Gamma^\kappa_{\mu\sigma}) + \Gamma^\kappa_{\mu\epsilon} \Gamma^\epsilon_{\nu\sigma} - \Gamma^\kappa_{\nu\epsilon} \Gamma^\epsilon_{\mu\sigma}$$

sont les composantes d'un tenseur (1, 3) appelé *tenseur de courbure de D* .; (iii) : pour $U, V \in \mathfrak{T}_0^1(TM)$ fixés, l'application $W \mapsto R(U, V, W)$ est un endomorphisme linéaire de $\mathfrak{T}_0^1(TM)$, notée (abusivement⁷) $R(U, V)$ et (donc) associée à un tenseur (1, 1) dont les composantes sont

$$R(U, V)^\kappa_\sigma = U^\mu V^\nu R^\kappa_{\sigma\mu\nu}$$

(iv) : La forme quadrilinéaire $U, V, W, Z \in \mathfrak{T}_0^1(TM)$ t.q.

$$(8) \quad \mathcal{R}(U, V, W, Z) := g(R(U, V, W), Z)$$

est un tenseur de valence (0, 4) appelé *tenseur de courbure covariant de D* , dont les composantes sont les $R_{\kappa\sigma\nu\mu} = g_{\kappa\epsilon} R^\epsilon_{\sigma\nu\mu}$: ainsi, pour tout $U, V, W, Z \in \mathfrak{T}_0^1(TM)$:

$$\mathcal{R}(U, V, W, Z) = U^\mu V^\nu W^\sigma Z^\kappa R_{\kappa\sigma\mu\nu}$$

Preuve. (i) : Une connexion est $C^\infty(M)$ -linéaire sur la première entrée et satisfait la règle de Leibniz sur la deuxième entrée. Par suite, pour tout⁸ $f \in C^\infty(M)$:

$$\begin{aligned} D_{fU}(D_U(W)) &= f D_U(D_U(W)) \\ D_V(D_{fU}(W)) &= D_V(f D_U(W)) = V(f) D_U(W) + f D_V(D_U(W)) \\ D_{[fU, V]}(W) &= f D_{[U, V]}(W) - V(f) D_U(W) \end{aligned}$$

6. On pourra aussi remarquer (exercice) que (a) : si D et \hat{D} sont deux connexions quelconques de coefficients respectifs $\Gamma^\sigma_{\mu\nu}$ et $\hat{\Gamma}^\sigma_{\mu\nu}$ alors les symboles $\Gamma^\sigma_{\mu\nu} - \hat{\Gamma}^\sigma_{\mu\nu}$ sont les composantes d'un tenseur de valence (1, 2) et (b) : si les $\Gamma^\sigma_{\mu\nu}$ sont les coefficients d'une connexion D , alors les $\Gamma^\sigma_{\nu\mu}$ (permutation des indices μ et ν) sont aussi les coefficients d'une connexion \hat{D} ; de (a) et (b) nous déduisons que le $T^\sigma_{\mu\nu} := \Gamma^\sigma_{\mu\nu} - \Gamma^\sigma_{\nu\mu}$ sont les composantes d'un tenseur (1, 2). On retrouve ainsi le tenseur de torsion associé à une connexion D .

7. Remarquons que si U et V commutent (et en particulier pour les champs holonômes) alors :

$$R(U, V)W = D_U(D_V(W)) - D_V(D_U(W)) = D_{[U, V]}(W)$$

8. Pour la troisième identité, nous utilisons le fait que $[fU, V] = f[U, V] - V(f)U$.

d'où $R(fU, V, W) = fR(U, V, W)$ (et de même $R(U, fV, W) = fR(U, V, W)$). Il nous reste à démontrer que $R(U, V, fW) = fR(U, V, W)$: pour cela, remarquons que

$$D_U(D_V(fW)) = D_U(V(f)W + fD_V(W)) = D_U(V(f)W) + D_U(fD_V(W))$$

soit encore

$$(9) \quad D_U(D_V(fW)) = U(V(f))W + V(f)D_U(W) + U(f)D_V(W) + fD_U(D_V(W))$$

et de même (en permutant les rôles des champs de vecteurs U et V)

$$(10) \quad D_V(D_U(fW)) = V(U(f))W + U(f)D_V(W) + V(f)D_U(W) + fD_V(D_U(W))$$

Par soustraction des identités (9) et (10) nous obtenons

$$(11) \quad D_U(D_V(fW)) - D_V(D_U(fW)) = [U, V](f)W + f(D_U(D_V(W)) - D_V(D_U(W)))$$

Enfin, l'identité $D_{[U, V]}(fW) = [U, V](f)W + fD_{[U, V]}(W)$ permet d'exprimer $[U, V](f)W$; en substituant le résultat obtenu dans (11) nous tirons $R(U, V, fW) = fR(U, V, W)$.

(ii) : Si $(\partial_1, \dots, \partial_n)$ est la base holonôme d'une carte, alors $[\partial_\mu, \partial_\nu] = 0$ et

$$\begin{aligned} R(\partial_\mu, \partial_\nu, \partial_\sigma) &= D_{\partial_\mu}(D_{\partial_\nu}(\partial_\sigma)) - D_{\partial_\nu}(D_{\partial_\mu}(\partial_\sigma)) - D_{[\partial_\mu, \partial_\nu]}(\partial_\sigma) \\ &= D_{\partial_\mu}(\Gamma^\kappa_{\nu\sigma}\partial_\kappa) - D_{\partial_\nu}(\Gamma^\kappa_{\mu\sigma}\partial_\kappa) \\ &= \left(\partial_\mu(\Gamma^\kappa_{\nu\sigma})\partial_\kappa + \Gamma^\kappa_{\nu\sigma}D_{\partial_\mu}(\partial_\kappa)\right) - \left(\partial_\nu(\Gamma^\kappa_{\mu\sigma})\partial_\kappa + \Gamma^\kappa_{\mu\sigma}D_{\partial_\nu}(\partial_\kappa)\right) \\ &= \left(\partial_\mu(\Gamma^\kappa_{\nu\sigma})\partial_\kappa + \Gamma^\kappa_{\nu\sigma}\Gamma^\epsilon_{\mu\kappa}\partial_\epsilon\right) - \left(\partial_\nu(\Gamma^\kappa_{\mu\sigma})\partial_\kappa + \Gamma^\kappa_{\mu\sigma}\Gamma^\epsilon_{\nu\kappa}\partial_\epsilon\right) \\ &= \left(\partial_\mu(\Gamma^\kappa_{\nu\sigma})\partial_\kappa + \Gamma^\epsilon_{\nu\sigma}\Gamma^\kappa_{\mu\epsilon}\partial_\kappa\right) - \left(\partial_\nu(\Gamma^\kappa_{\mu\sigma})\partial_\kappa + \Gamma^\epsilon_{\mu\sigma}\Gamma^\kappa_{\nu\epsilon}\partial_\kappa\right) \quad (\epsilon \leftrightarrow \kappa) \\ &= \left(\partial_\mu(\Gamma^\kappa_{\nu\sigma}) - \partial_\nu(\Gamma^\kappa_{\mu\sigma}) + \Gamma^\epsilon_{\nu\sigma}\Gamma^\kappa_{\mu\epsilon} - \Gamma^\epsilon_{\mu\sigma}\Gamma^\kappa_{\nu\epsilon}\right)\partial_\kappa \end{aligned}$$

soit encore (en réordonnant les produits de symboles de Christoffel)

$$R(\partial_\mu, \partial_\nu, \partial_\sigma) = R^\kappa_{\sigma\mu\nu}\partial_\kappa \quad \text{avec} \quad R^\kappa_{\sigma\mu\nu} := \partial_\mu(\Gamma^\kappa_{\nu\sigma}) - \partial_\nu(\Gamma^\kappa_{\mu\sigma}) + \Gamma^\kappa_{\mu\epsilon}\Gamma^\epsilon_{\nu\sigma} - \Gamma^\kappa_{\nu\epsilon}\Gamma^\epsilon_{\mu\sigma}$$

Le fait que les $R^\kappa_{\lambda\mu\nu}$ sont les composantes d'un tenseur (1, 3) découle directement de la partie (i) affirmant que $(U, V, W) \mapsto R(U, V, W)$ est $\mathcal{C}^\infty(M)$ -linéaire en ses trois entrées.

(iii) : est une conséquence directe de (i).

(iv) : Pour montrer que la forme quadrilinéaire définie en(8) est une tenseur (0, 4), il suffit de remarquer que pour $U, V, W, Z \in \mathfrak{X}_0^1(TM)$

$$\begin{aligned} \mathcal{R}(U, V, W, Z) &= g(U^\mu V^\nu W^\sigma R^\epsilon_{\sigma\mu\nu}\partial_\epsilon, Z^\kappa\partial_\kappa) \\ &= U^\mu V^\nu W^\sigma Z^\kappa g(\partial_\epsilon, \partial_\kappa) R^\epsilon_{\sigma\mu\nu} = U^\mu V^\nu W^\sigma Z^\kappa g_{\kappa\epsilon} R^\epsilon_{\sigma\mu\nu} \end{aligned}$$

□

4. Identités de Ricci

Revenons maintenant au cas où la variété pseudo-riemannienne (M, g) est munie d'une connexion D dont nous notons (en coordonnées) $\Gamma^\sigma_{\mu\nu}$ les coefficients. Étant donné V un champ de vecteur de composantes V^μ , la $\mathcal{C}^\infty(M)$ -linéarité de l'application partielle $U \mapsto D_U(V)$ entraîne que les $D_\nu(V^\nu)$ définis en (4) sont les composantes d'un tenseur

(1, 1) noté $D(V)$. L'application linéaire $D : \mathfrak{T}_0^1(TM) \rightarrow \mathfrak{T}_1^1(TM)$ est par définition la dérivation covariante associée à la connexion \mathbf{D} ⁹, chaque système de coordonnées locales correspondant à un système (D_1, \dots, D_n) de **dérivations partielles covariantes**. Intuitivement, D_μ s'interprète comme une *correction* de la dérivation partielle ∂_μ , cette correction lui donnant des propriétés tensorielles adaptées à la structure de variété différentielle. Il faut ici pointer une différence entre les dérivations partielles et les dérivations partielles covariantes : contrairement au lemme de Schwarz pour les dérivées partielles, la composition des dérivations partielles covariantes ne commutent pas : nous allons voir (c.f. Proposition 4.2) comment le *crochet de Lie* $[D_\mu, D_\nu] = D_{\mu\nu} - D_{\nu\mu}$ est liée au tenseur de courbure de \mathbf{D} . Le point de départ est l'extension de la dérivation covariante D à l'espace complet $\coprod_{p,q} \mathfrak{T}_q^p(TM)$ des champs tenseurs (c.f. [Oli15a, Proposition 14.1]) : si $A^{\tau_1 \dots \tau_p}_{\sigma_1 \dots \sigma_q}$ sont les composantes d'un tenseur $A \in \mathfrak{T}_q^p(TM)$ alors la dérivée covariante $D(A)$ est le tenseur de $\mathfrak{T}_{q+1}^p(TM)$ dont les composantes sont :

$$(12) \quad D_\mu(A^{\tau_1 \dots \tau_p}_{\sigma_1 \dots \sigma_q}) := \partial_\mu(A^{\tau_1 \dots \tau_p}_{\sigma_1 \dots \sigma_q}) + \sum_{r=1}^p A^{\tau_1 \dots (\epsilon \leftarrow \tau_r) \dots \tau_p}_{\sigma_1 \dots \sigma_q} \Gamma^{\tau_r}_{\epsilon \mu} - \sum_{s=1}^q A^{\tau_1 \dots \tau_p}_{\sigma_1 \dots (\epsilon \leftarrow \sigma_s) \dots \sigma_q} \Gamma^{\epsilon}_{\sigma_s \mu}$$

Enfin, pour tout champ de vecteur U , de \mathbf{x} -composante U^k nous pouvons définir :

$$D_U(A) := U^\epsilon (D_\epsilon(A^{\tau_1 \dots \tau_p}_{\sigma_1 \dots \sigma_q})) \partial_{\tau_1} \otimes \dots \otimes \partial_{\tau_p} \otimes d\mathbf{x}^{\sigma_1} \otimes \dots \otimes d\mathbf{x}^{\sigma_q}$$

De même que pour les dérivations partielles nous notons $D_{\tau_1 \dots \tau_p}$ la composition des p dérivations partielles covariantes $D_{\tau_1}, \dots, D_{\tau_p}$. Rappelons aussi (lemme de Schwarz) que le crochet de Lie des dérivées partielles s'annule, ce qu'on note $[\partial_\mu, \partial_\nu] \equiv 0$: nous utilisons aussi la notation en *crochet le lie* pour les dérivées partielles covariante en posant :

$$[D_\mu, D_\nu](A^{\tau_1 \dots \tau_p}_{\sigma_1 \dots \sigma_q}) = D_{\mu\nu}(A^{\tau_1 \dots \tau_p}_{\sigma_1 \dots \sigma_q}) - D_{\nu\mu}(A^{\tau_1 \dots \tau_p}_{\sigma_1 \dots \sigma_q})$$

Le « *crochet de Lie* » des dérivées partielles covariantes est lié à la courbure ainsi qu'à la torsion de la connexion.

Théorème 4.1. *Soit (M, g) une variété pseudo-riemannienne et \mathbf{D} une connexion ; alors pour tout champ de vecteur V de composantes V^μ et toute forme différentielle ξ de composantes ξ_μ :*

$$[D_\mu, D_\nu](V^\sigma) := V^\epsilon R^\sigma_{\epsilon\mu\nu} - D_\epsilon(V^\sigma) T^\epsilon_{\mu\nu} \quad \text{et} \quad [D_\mu, D_\nu](\xi_\sigma) = -\xi_\epsilon R^\epsilon_{\sigma\mu\nu} - D_\eta(\xi_\sigma) T^\eta_{\mu\nu}$$

où les $R^\sigma_{\epsilon\mu\nu}$ (resp. $T^\epsilon_{\mu\nu}$) sont les composantes du tenseur de courbure (resp. de torsion) de \mathbf{D} .

Preuve. Soit $(\partial_1, \dots, \partial_n)$ la base holonôme d'un système de coordonnées \mathbf{x} ; pour $V = V^\mu \partial_\mu$ (et de même pour une forme différentielle $\xi = \xi_\mu d\mathbf{x}^\mu$), nous utilisons (12) pour

9. Chaque connexion $(U, V) \mapsto D_U(V)$ sur M est automatiquement associée à une dérivation covariante $D : \mathfrak{T}_0^1(TM) \rightarrow \mathfrak{T}_1^1(TM)$ ainsi qu'à un système de dérivation covariante (D_1, \dots, D_n) défini pour chaque système de coordonnées locales.

écrire

$$\begin{aligned}
& D_\mu(D_\nu(V^\sigma)) \\
&= \partial_\mu(D_\nu(V^\sigma)) - D_\eta(V^\sigma)\Gamma^\eta_{\mu\nu} + D_\nu(V^\epsilon)\Gamma^\sigma_{\mu\epsilon} \\
&= \partial_{\mu\nu}(V^\sigma) + \partial_\mu(V^\epsilon)\Gamma^\sigma_{\nu\epsilon} + V^\epsilon\partial_\mu(\Gamma^\sigma_{\nu\epsilon}) - D_\eta(V^\sigma)\Gamma^\eta_{\mu\nu} + \partial_\nu(V^\epsilon)\Gamma^\sigma_{\mu\epsilon} + V^\eta\Gamma^\epsilon_{\nu\eta}\Gamma^\sigma_{\mu\epsilon} \\
&= \left(\partial_{\mu\nu}(V^\sigma) + \partial_\mu(V^\epsilon)\Gamma^\sigma_{\nu\epsilon} + \partial_\nu(V^\epsilon)\Gamma^\sigma_{\mu\epsilon}\right) + \left(V^\epsilon\partial_\mu(\Gamma^\sigma_{\nu\epsilon}) + V^\eta\Gamma^\epsilon_{\nu\eta}\Gamma^\sigma_{\mu\epsilon}\right) - D_\eta(V^\sigma)\Gamma^\eta_{\mu\nu} \\
&= \left(\partial_{\mu\nu}(V^\sigma) + \partial_\mu(V^\epsilon)\Gamma^\sigma_{\nu\epsilon} + \partial_\nu(V^\epsilon)\Gamma^\sigma_{\mu\epsilon}\right) + \left(V^\epsilon\partial_\mu(\Gamma^\sigma_{\nu\epsilon}) + V^\epsilon\Gamma^\eta_{\nu\epsilon}\Gamma^\sigma_{\mu\eta}\right) - D_\eta(V^\sigma)\Gamma^\eta_{\mu\nu}
\end{aligned}$$

Finalement nous obtenons

$$[D_\mu, D_\nu](V^\sigma) = V^\epsilon \left(\partial_\mu(\Gamma^\sigma_{\nu\epsilon}) - \partial_\nu(\Gamma^\sigma_{\mu\epsilon}) + \Gamma^\sigma_{\mu\eta}\Gamma^\eta_{\nu\epsilon} - \Gamma^\sigma_{\nu\eta}\Gamma^\eta_{\mu\epsilon} \right) - D_\eta(V^\sigma) \left(\Gamma^\eta_{\mu\nu} - \Gamma^\eta_{\nu\mu} \right) \quad \square$$

Corollaire 4.2 (Identités de Ricci). *Soit (M, g) une variété pseudo-riemannienne et D une connexion sans torsion ; alors pour tout champ de vecteur V et toute forme différentielle¹⁰ ξ :*

$$[D_\mu, D_\nu](V^\sigma) := V^\epsilon R^\sigma_{\epsilon\mu\nu} \quad \text{et} \quad [D_\mu, D_\nu](\xi_\sigma) = -\xi_\epsilon R^\epsilon_{\sigma\mu\nu}$$

Preuve. Les identités de Ricci découlent directement du Théorème 4.1 du fait que la connexion D est supposée sans torsion (i.e. $T^\sigma_{\mu\nu} = \Gamma^\sigma_{\mu\nu} - \Gamma^\sigma_{\nu\mu} = 0$). □

Les identités de Ricci s'appliquent plus généralement sous la forme suivante :

Théorème 4.3 (Identités de Ricci généralisées). *Soit (M, g) une variété pseudo-riemannienne et D une connexion sans torsion : si $F^{\sigma_1 \dots \sigma_p}_{\tau_1 \dots \tau_q}$ sont les composantes d'un tenseur (p, q) alors :*

$$\begin{aligned}
[D_\mu, D_\nu](F^{\sigma_1 \dots \sigma_p}_{\tau_1 \dots \tau_q}) &= \sum_{i=1}^p F^{\sigma_1 \dots (\epsilon \leftarrow \sigma_i) \dots \sigma_p}_{\tau_1 \dots \tau_q} R^{\sigma_i}_{\epsilon\mu\nu} \\
&= - \sum_{i=1}^q F^{\sigma_1 \dots \sigma_p}_{\tau_1 \dots (\epsilon \leftarrow \tau_i) \dots \tau_q} R^\epsilon_{\tau_i\mu\nu}
\end{aligned}$$

5. Les symétries du tenseur de courbure

Soit (M, g) une variété pseudo-riemannienne de dimension n et D une connexion affine (à priori quelconque) de coefficients $\Gamma^\sigma_{\mu\nu}$ et associée au système de dérivations partielles covariantes (D_1, \dots, D_n) . Le tenseur de courbure associé à la connexion D est le tenseur $(1, 3)$ (abusivement) identifié à l'application $(U, V, W) \mapsto R(U, V, W) = R(U, V)W$ définie par (6) ; ses composantes $R^\tau_{\sigma\mu\nu}$ données par (7) s'expriment en fonction des coefficients de la connexion et de leur dérivées premières. De plus pour U, V fixés, $R(U, V)$ s'identifie à un tenseur $(1, 1)$ de composantes $R(U, V)^\tau_\sigma = U^\mu V^\nu R^\tau_{\sigma\mu\nu}$. La **première antisymétrie** du tenseur de courbure est une conséquence directe de la définition du tenseur

10. Nous savons aussi (c.f. [Oli15a, Théorème 14.2]) que la g -compatibilité d'une connexion D est équivalent à la commutation des composantes de la métrique avec les dérivées partielles covariantes : ainsi, en écrivant $\xi^\sigma = g^{\sigma\epsilon}\xi_\epsilon$, il vient $[D_\mu, D_\nu](\xi^\sigma) = [D_\mu, D_\nu](g^{\sigma\epsilon}\xi_\epsilon) = g^{\sigma\epsilon}[D_\mu, D_\nu](\xi_\epsilon)$ et par suite

$$[D_\mu, D_\nu](\xi_\sigma) = g_{\sigma\epsilon}(\xi^\eta R^\epsilon_{\eta\mu\nu}) = \xi^\eta (g_{\sigma\epsilon} R^\epsilon_{\eta\mu\nu}) (g^{\eta\tau} \xi_\tau) R_{\sigma\eta\mu\nu} = \xi_\tau (g^{\tau\eta} R_{\sigma\eta\mu\nu})$$

Or (c.f. (iii)-Proposition 5.1 infra), la compatibilité métrique de D entraîne aussi l'antisymétrie $R_{\sigma\eta\mu\nu} = -R_{\eta\sigma\mu\nu}$: par suite $g^{\tau\eta} R_{\sigma\eta\mu\nu} = -g^{\tau\eta} R_{\eta\sigma\mu\nu} = -R^\tau_{\sigma\mu\nu}$ et finalement : $[D_\mu, D_\nu](\xi_\sigma) := -\xi_\tau R^\tau_{\sigma\mu\nu}$.

lui même (6) et du fait que $[V, U] = -[U, V]$: ainsi, nous obtenons $R(U, V) = -R(V, U)$, ce qui se traduit en composantes par l'identité :

$$R^\sigma{}_{\tau\mu\nu} = -R^\sigma{}_{\tau\nu\mu}$$

Maintenant, nous **supposons que la connexion D est sans torsion**. Étant donnés U, V, W trois champs de vecteurs, nous composons l'équation $D_V(W) - D_W(V) - [V, W] = 0$ par D_U : par permutation circulaire de U, V, W , nous obtenons trois équations, soient :

$$(13) \quad D_U(D_V(W)) - D_U(D_W(V)) - D_U([V, W]) = 0$$

$$(14) \quad D_W(D_U(V)) - D_W(D_V(U)) - D_W([U, V]) = 0$$

$$(15) \quad D_V(D_W(U)) - D_V(D_U(W)) - D_V([W, U]) = 0$$

Or, l'absence de torsion de D , nous permet aussi d'écrire successivement

$$D_U([V, W]) = D_{[V, W]}(U) + [U, [V, W]]$$

$$D_W([U, V]) = D_{[U, V]}(W) + [W, [U, V]]$$

$$D_V([W, U]) = D_{[W, U]}(V) + [V, [W, U]]$$

Par suite en additionnant membre à membre les équations (13) (14) et (15) nous obtenons

$$(16) \quad R(U, V)W + R(W, U)V + R(V, W)U = [U, [V, W]] + [W, [U, V]] + [V, [W, U]]$$

L'identité de Jacobi du crochet de Lie sur les champs annule membre de gauche de (16) : nous obtenons alors la **première identité de Bianchi** :

$$R(U, V)W + R(W, U)V + R(V, W)U = 0$$

En composantes, nous pouvons écrire

$$R^\sigma{}_{\tau\mu\nu} + R^\sigma{}_{\nu\tau\mu} + R^\sigma{}_{\mu\nu\tau} = 0$$

La **deuxième antisymétrie du tenseur de courbure** n'est vérifiée que lorsque **la connexion est métrique et sans torsion**, i.e. lorsque D coïncide avec la connexion de Levi-Civita ∇ . En effet, par application du Théorème 4.3 au tenseur métrique, il vient :

$$(17) \quad \nabla_\alpha(\nabla_\beta(g_{\mu\nu})) - \nabla_\beta(\nabla_\alpha(g_{\mu\nu})) = g_{\sigma\nu}R^\sigma{}_{\mu\alpha\beta} + g_{\mu\sigma}R^\sigma{}_{\nu\alpha\beta} = R_{\nu\mu\alpha\beta} + R_{\mu\nu\alpha\beta}$$

Or la dérivée covariante $\nabla(g)$ (de composantes $\nabla_\sigma(g_{\mu\nu})$) du tenseur métrique (pour la connexion de Levi-Civita) est identiquement nulle et par suite (17) devient

$$R_{\mu\nu\alpha\beta} = -R_{\nu\mu\alpha\beta}$$

Notons enfin que lorsque $D = \nabla$ le tenseur de courbure (i.e. le tenseur de Riemann) satisfait à la fois la première antisymétrie, l'identité de Bianchi et la deuxième antisymétrie : par suite (c.f. Proposition 2.1) le tenseur de Riemann covariant satisfait aussi **la bisymétrie**, c'est-à-dire qu'en composantes :

$$R_{\mu\nu\alpha\beta} = R_{\alpha\beta\mu\nu}$$

Proposition 5.1. *Soit (M, g) une variété pseudo-riemannienne de dimension n et soit $(U, V, W) \mapsto R(U, V)W$ le tenseur de courbure associé à une connexion D à priori arbitraire. Alors (i) : La première antisymétrie i.e.*

$$R^\tau{}_{\sigma[\mu\nu]} = 0$$

est toujours valide ; (ii) : la première identité de Bianchi, soit

$$R^{\tau}{}_{[\sigma\mu\nu]} = 0$$

est satisfaite dès que la connexion D est sans torsion ; (iii) : lorsque la connexion D est métrique et sans torsion (i.e. D coïncide avec la connexion de Levi-Civita) le tenseur de courbure satisfait aussi la deuxième antisymétrie, ainsi que la bisymétrie, soient respectivement :

$$R_{[\tau\sigma]\mu\nu} = 0 \quad \text{et} \quad R_{\tau\sigma\mu\nu} = R_{\mu\nu\tau\sigma}$$

Nous avons vu (Proposition 5.1) que la première identité de Bianchi est satisfaite par le tenseur de courbure d'une connexion D , du moment que la connexion en question est sans torsion. Nous terminons ce paragraphe en établissant la **seconde identité de Bianchi** (ou encore **identité différentielle de Bianchi**) : elle s'exprime à l'aide de la dérivation covariante des tenseurs.

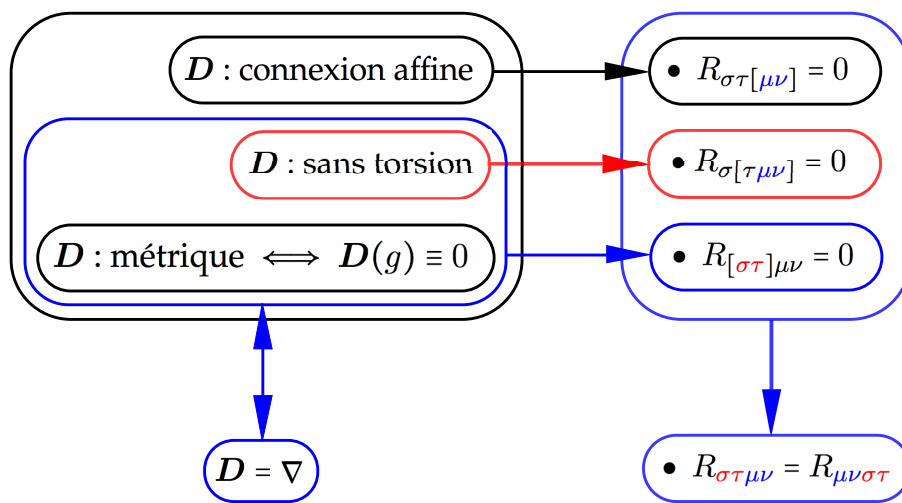


FIGURE 2. Les propriétés de symétrie du tenseur de courbure d'une connexion D relativement aux propriétés de D . Le tenseur de Riemann proprement dit correspond au cas où D est métrique et sans torsion, ce qui signifie (Théorème de Koszul) que D coïncide avec la connexion de Levi-Civita notée ∇ .

Proposition 5.2. Soit M une variété et $(U, V, W) \mapsto R(U, V)W$ le tenseur de courbure associé à une connexion D , supposée sans torsion. La **seconde identité de Bianchi** affirme que pour tout triplet (U, V, W) de champ de vecteurs sur M ,

$$D_U(R(V, W)) + D_V(R(W, U)) + D_W(R(U, V)) = 0$$

ce qui en composante s'écrit

$$D_\epsilon(R^\tau{}_{\sigma\mu\nu}) + D_\mu(R^\tau{}_{\sigma\nu\epsilon}) + D_\nu(R^\tau{}_{\sigma\epsilon\mu}) = 0$$

Preuve (d'après [Wal10]). En identifiant les dérivées partielles covariantes des deux membres de l'identité de Ricci $[D_\mu, D_\nu](\xi_\sigma) = -\xi_\tau R^\tau{}_{\sigma\mu\nu}$ (c.f. Corollaire 4.2), nous obtenons par la règle de Leibniz appliquée au second membre :

$$D_\epsilon[D_\mu, D_\nu](\xi_\sigma) = -D_\epsilon(\xi_\tau)R^\tau{}_{\sigma\mu\nu} - \xi_\tau D_\epsilon(R^\tau{}_{\sigma\mu\nu})$$

Il vient alors successivement

$$\begin{aligned}\xi_\tau D_\epsilon(R^\tau_{\sigma\mu\nu}) &= -D_\epsilon[D_\mu, D_\nu](\xi_\sigma) - D_\epsilon(\xi_\tau)R^\tau_{\sigma\mu\nu} \\ &= -D_\epsilon[D_\mu, D_\nu](\xi_\sigma) + \left(-D_\epsilon(\xi_\tau)R^\tau_{\sigma\mu\nu} - D_\tau(\xi_\sigma)R^\tau_{\epsilon\mu\nu}\right) + D_\tau(\xi_\sigma)R^\tau_{\epsilon\mu\nu} \\ &= -D_\epsilon[D_\mu, D_\nu](\xi_\sigma) + [D_\mu, D_\nu](D_\epsilon(\xi_\sigma)) + D_\tau(\xi_\sigma)R^\tau_{\epsilon\mu\nu}\end{aligned}$$

où nous avons utilisé l'identité de Ricci généralisée (c.f. Théorème 4.3) avec le tenseur (0, 2) de composante $D_\epsilon(\xi_\tau)$. En développant les crochets des dérivées secondes $[D_\mu, D_\nu]$ et en effectuant les permutations sur les indices ϵ, μ, ν , il vient

$$\begin{aligned}\xi_\tau D_\epsilon(R^\tau_{\sigma\mu\nu}) &= -\left(D_{\epsilon\mu\nu}(\xi_\sigma) - D_{\epsilon\nu\mu}(\xi_\sigma)\right) + \left(D_{\mu\nu\epsilon}(\xi_\sigma) - D_{\nu\mu\epsilon}(\xi_\sigma)\right) + D_\tau(\xi_\sigma)R^\tau_{\epsilon\mu\nu} \\ \xi_\tau D_\mu(R^\tau_{\sigma\nu\epsilon}) &= -\left(D_{\mu\nu\epsilon}(\xi_\sigma) - D_{\mu\epsilon\nu}(\xi_\sigma)\right) + \left(D_{\nu\epsilon\mu}(\xi_\sigma) - D_{\epsilon\nu\mu}(\xi_\sigma)\right) + D_\tau(\xi_\sigma)R^\tau_{\mu\nu\epsilon} \\ \xi_\tau D_\nu(R^\tau_{\sigma\epsilon\mu}) &= -\left(D_{\nu\epsilon\mu}(\xi_\sigma) - D_{\nu\mu\epsilon}(\xi_\sigma)\right) + \left(D_{\epsilon\mu\nu}(\xi_\sigma) - D_{\mu\epsilon\nu}(\xi_\sigma)\right) + D_\tau(\xi_\sigma)R^\tau_{\nu\epsilon\mu}\end{aligned}$$

soit encore, en additionnant membres à membres ces trois équations :

$$\xi_\tau \left(D_\epsilon(R^\tau_{\sigma\mu\nu}) + D_\mu(R^\tau_{\sigma\nu\epsilon}) + D_\nu(R^\tau_{\sigma\epsilon\mu}) \right) = D_\tau(\xi_\sigma) \left(R^\tau_{\epsilon\mu\nu} + R^\tau_{\mu\nu\epsilon} + R^\tau_{\nu\epsilon\mu} \right)$$

La conclusion vient de la première identité de Bianchi (c.f. (ii)-Proposition 5.1) et du fait que ξ est une forme différentielle quelconque. □

Remarque 5.3. *Il existe une notation classique donnant une expression condensée des dérivées covariantes des tenseurs : si $F^{\mu_1 \dots \mu_p}_{\nu_1 \dots \nu_q}$ sont les composantes d'un tenseur (p, q) , alors :*

$$F^{\mu_1 \dots \mu_p}_{\nu_1 \dots \nu_q; \epsilon} := D_\epsilon \left(F^{\mu_1 \dots \mu_p}_{\nu_1 \dots \nu_q} \right)$$

Ainsi, (lorsque D est sans torsion) nous pouvons écrire la seconde identité de Bianchi

$$R^\tau_{\sigma[\mu\nu; \epsilon]} = 0$$

6. Exponentielle de Riemann et coordonnées normales

Dans ce paragraphe, nous considérons que (M, g) est une variété (pseudo)-riemannienne de dimension n . D'après le Théorème de Levi-Civita (c.f. [Oli15a, Théorème 13.1]), nous savons qu'un chemin $t \mapsto \gamma(t)$ dans M est une géodésique inertielle si et seulement si γ transporte parallèlement son propre vecteur vitesse relativement à la connexion de Levi-Civita : cela se traduit – dans un système x de coordonnées locales – par le fait que les composantes cartésiennes γ^μ de γ (i.e. $\gamma = X(\gamma^1, \dots, \gamma^n)$, avec X la carte associée à x) satisfont le système différentiel

$$(18) \quad \ddot{\gamma}^\sigma(t) + \dot{\gamma}^\mu(t)\dot{\gamma}^\nu(t)\mathbb{C}^\sigma_{\mu\nu}(\gamma(t)) = 0$$

où les $\mathbb{C}^\sigma_{\mu\nu}$ désignent les symboles de Christoffel pour les coordonnées locales x . La géodésique inertielle γ est dite issue d'un point X lorsque $\gamma(0) = X$; une telle géodésique est entièrement déterminée par le vecteur vitesse initial, soit $v = \dot{\gamma}(0)$: pour préciser les paramètres de la géodésique, nous notons $\gamma(t) = \exp_X(tv)$, c'est-à-dire que $(\gamma(0), \dot{\gamma}(0)) = (X, v)$. Si \mathcal{V} est le sous-ensemble de $T_X M$ des v t.q. $\exp(tv)$ soit défini pour tout $0 \leq t \leq 1$, alors l'application $v \mapsto \exp_X(v) =: \exp_X(1v)$ définie pour tout $v \in \mathcal{V}$ est [par définition l'exponentielle de Riemann](#) : nous supposons par la suite que (M, g) est

géodésiquement complète, ce qui signifie que \exp_X est définie sur tout $T_X M$, pour tout $X \in M$. En général, l'exponentielle de Riemann n'est ni injective, ni surjective (penser par exemple à la 2-sphère pour l'injectivité, ou encore à une variété non connexe pour la surjectivité). Cependant (et nous l'admettrons) il existe toujours un voisinage ouvert de 0 dans $T_X M$ pour lequel \exp_X réalise un difféomorphisme sur son image \mathcal{U} (ouvert de M). Supposons que X appartienne au domaine d'un système de coordonnées $\mathbf{x} : \mathcal{U} \rightarrow \mathbb{R}^n$ et soit \mathbf{X} la carte associée à \mathbf{x} : alors \mathbf{x} est dit **normal (ou encore inertielle) en X** , si la base holonôme $(\partial_1|_X, \dots, \partial_n|_X)$ de $T_X M$ est orthonormée et si de plus il existe un voisinage de 0 dans \mathbb{R}^n t.q. pour tout (y^1, \dots, y^n) dans ce voisinage

$$(19) \quad \exp_X(y^i \partial_i|_X) = Y \iff \mathbf{X}(y^1, \dots, y^n) = Y \iff (y^1, \dots, y^n) = \mathbf{x}(Y)$$

Nous admettrons l'existence de systèmes de coordonnées normal en chaque point X de M , ce qui revient à admettre l'existence du voisinage de 0 dans $T_X M$ sur lequel \exp_X réalise un difféomorphisme.

Proposition 6.1. *Soit (M, g) une variété pseudo-riemannienne de dimension n , de signature*

$$(\epsilon_1, \dots, \epsilon_n) = (+1, \dots, +1, -1, \dots, -1)$$

Si \mathbf{x} un système de coordonnées normal en X , alors :

$$\mathbf{x}(X) = (0, \dots, 0) \quad g_{\mu\nu}(X) = \epsilon_\mu \delta_{\mu\nu} \quad \mathbf{C}^k_{\mu\nu}(X) = 0 \quad \partial_\sigma|_X(g_{\mu\nu}) = 0 \quad \partial_\sigma|_X(g^{\mu\nu}) = 0$$

Preuve. Soit X un point du domaine d'un système (quelconque) de coordonnées \mathbf{x} et $t \mapsto \gamma(t) \in \mathcal{U}$ une géodésique inertielle issue de X (i.e. $\gamma(0) = X$). Si $\gamma^1(t), \dots, \gamma^n(t)$ sont les composantes de $\gamma(t)$ – en ce sens que $\mathbf{x}(\gamma(t)) = (\gamma^1(t), \dots, \gamma^n(t))$ – alors $\gamma(t) = \exp_X(tv)$ où $v = \dot{\gamma}^i(0)\partial_i|_X$. Alors le chemin $t \mapsto (\gamma^1(t), \dots, \gamma^n(t))$ (dans \mathbb{R}^n) est un germe de chemin en $\mathbf{x}(X)$ solution du système différentiel :

$$(20) \quad (\forall \sigma) \quad \ddot{\gamma}^\sigma(t) + \dot{\gamma}^\mu(t)\dot{\gamma}^\nu(t)\mathbf{C}^\sigma_{\mu\nu}(\gamma(t)) = 0$$

Soit \mathbf{x} un système de coordonnées normal en X : alors, d'une part, $(\partial_1|_X, \dots, \partial_n|_X)$ est (définition) une base orthonormée de $T_X M$; d'autre part, l'identité $\exp_X(0v) = X$ implique d'après (19) que $\mathbf{x}(X) = (0, \dots, 0)$. Enfin, (19) entraîne aussi que

$$\exp_X(tv^i \partial_i|_X) = \gamma(t) \iff (tv^1, \dots, tv^n) = \mathbf{x}(\gamma(t))$$

en d'autres termes $tv^i = \gamma^i(t)$, pour tout t et tout $1 \leq i \leq n$ et par suite $(\dot{\gamma}^i(t), \ddot{\gamma}^i(t)) \equiv (v^i, 0)$. D'après (20) il vient $v^\mu v^\nu \mathbf{C}^\sigma_{\mu\nu}(\gamma(t)) = 0$. La géodésique considérée en X étant arbitraire, nous obtenons que $v^\mu v^\nu \mathbf{C}^\sigma_{\mu\nu}(X) = 0$, pour tout $(v^1, \dots, v^n) \in \mathbb{R}^n$: cela entraîne $\mathbf{C}^\sigma_{\mu\nu}(X) = 0$, pour tout triplet d'indices (σ, μ, ν) . Or, $\mathbf{C}_{\sigma\mu\nu} = g_{\sigma\epsilon} \mathbf{C}^\epsilon_{\mu\nu}$ et comme $\mathbf{C}_{\sigma\mu\nu} + \mathbf{C}_{\nu\mu\sigma} = \partial_\mu(g_{\nu\sigma})$, nous obtenons d'abord $\partial_\mu|_X(g_{\nu\sigma}) = 0$; l'identité $\partial_\mu|_X(g^{\nu\sigma}) = 0$, s'en déduit du fait que $g^{\nu\sigma} g_{\sigma\mu} = \delta^\nu_\mu$. □

Remarque 6.2. *Pour la connexion de Levi-Civita, les composantes $R_{\kappa\sigma\mu\nu} := g_{\kappa\epsilon} R^\epsilon_{\sigma\mu\nu}$ du tenseur de courbure déduites de (7) par un calcul direct, s'écrivent :*

$$(21) \quad R_{\kappa\sigma\mu\nu} = \frac{1}{2} \left(\partial_{\sigma\nu}(g_{\kappa\mu}) - \partial_{\kappa\nu}(g_{\sigma\mu}) + \partial_{\kappa\mu}(g_{\sigma\nu}) - \partial_{\sigma\mu}(g_{\kappa\nu}) \right) + g_{\epsilon\eta} \left(\mathbf{C}^\epsilon_{\kappa\mu} \mathbf{C}^\eta_{\sigma\nu} - \mathbf{C}^\epsilon_{\kappa\nu} \mathbf{C}^\eta_{\sigma\mu} \right)$$

Il est alors possible de retrouver rapidement les identités de symétrie du tenseur de Riemann (covariant par exemple) en se plaçant dans un système des coordonnées normal en un point X donnée : dans ce cas, nous obtenons par la Proposition 6.1 :

$$(22) \quad R_{\kappa\sigma\mu\nu}(X) = \frac{1}{2} \left(\partial_{\kappa\mu}|_X(g_{\sigma\nu}) + \partial_{\sigma\nu}|_X(g_{\kappa\mu}) - \partial_{\sigma\mu}|_X(g_{\kappa\nu}) - \partial_{\kappa\nu}|_X(g_{\sigma\mu}) \right)$$

En permutant les indices κ et σ dans (22) il vient :

$$(23) \quad R_{\sigma\kappa\mu\nu}(X) = \frac{1}{2} \left(\partial_{\sigma\mu}|_X(g_{\kappa\nu}) + \partial_{\kappa\nu}|_X(g_{\sigma\mu}) - \partial_{\kappa\mu}|_X(g_{\sigma\nu}) - \partial_{\sigma\nu}|_X(g_{\kappa\mu}) \right) = -R_{\kappa\sigma\mu\nu}(X)$$

De même, en effectuant les trois permutations circulaires sur les indices (σ, μ, ν) dans (22), et en additionnant membre à membre les identités obtenues, nous retrouvons la première identité de Bianchi, i.e. $R_{\sigma[\kappa\mu\nu]}(X) = 0$. Enfin, en permutant les blocs d'indices $\kappa\sigma$ et $\mu\nu$ dans (22), nous utilisons la symétrie des $g_{\mu\nu}$ et la commutation des dérivations partielles (holonomes : Lemme de Schwarz), ce qui permet de retrouver la bisymétrie $R_{\sigma\kappa\mu\nu}(X) = R_{\mu\nu\sigma\kappa}(X)$.

7. Le tenseur Ricci et l'équation d'Einstein

Le tenseur de Ricci est obtenu (au signe prêt) comme la seule contraction non nulle du tenseur de Riemann. Ainsi, dans des coordonnées normales et réduites en X ,

$$R^\tau{}_{\mu\nu\sigma} = \partial_\nu(\mathbf{C}^\tau{}_{\mu\sigma}) - \partial_\sigma(\mathbf{C}^\tau{}_{\mu\nu})$$

et pour $\tau = \mu$, les symboles de Christoffel réduits (c.f. [Oli15a, Lemme 15.2]) s'écrivent

$$R^\mu{}_{\mu\nu\sigma} = \partial_\nu \mathbf{C}^\mu{}_{\mu\sigma} - \partial_\sigma \mathbf{C}^\mu{}_{\mu\nu} = \partial_\nu \partial_\sigma \left(\log \sqrt{|\det g|} \right) - \partial_\sigma \partial_\nu \left(\log \sqrt{|\det g|} \right) = 0$$

(où on a utilisé le Lemme de Schwarz). Pour $\tau = \nu$ il vient $R_{\mu\sigma} = R^\nu{}_{\mu\nu\sigma} = \partial_\nu(\mathbf{C}^\nu{}_{\mu\sigma}) - \partial_\sigma(\mathbf{C}^\nu{}_{\mu\nu})$. Enfin, pour $\tau = \sigma$ nous avons $R^\sigma{}_{\mu\nu\sigma} = -R^\sigma{}_{\mu\sigma\nu} = -R_{\mu\nu}$. En résumé,

$$(R^\mu{}_{\mu j\sigma}, R^\nu{}_{\mu\nu\sigma}, R^\sigma{}_{\mu j\sigma}) = (0, R_{\mu\sigma}, -R_{\mu\nu})$$

Le **tenseur de Ricci** – noté **Ric** – est le tenseur $(0, 2)$ dont les composantes sont

$$R_{\mu\nu} = R^\sigma{}_{\mu\sigma\nu}$$

Le **scalaire de Ricci** R est alors défini comme la trace du tenseur de Ricci, en ce sens que

$$R := g^{\mu\nu} R_{\nu\mu} = R^\mu{}_{\mu}$$

Enfin, le **tenseur d'Einstein** – noté **Ein** – est le tenseur $(0, 2)$ dont les composantes sont :

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

Proposition 7.1. (i) : Le tenseur de Ricci est symétrique, i.e.

$$R_{\mu\nu} = R_{\nu\mu}$$

(ii) : La divergence du tenseur de Ricci est égale à la moitié du gradient du scalaire de Ricci, i.e.

$$\operatorname{div}(\operatorname{Ric})_\nu = \nabla_\mu (g^{\mu\eta} R_{\eta\nu}) = \nabla_\mu (R^\mu{}_\nu) = \frac{1}{2} \nabla_\nu (R) = \frac{1}{2} \partial_\nu (R)$$

(iii) : La divergence du tenseur d'Einstein est nulle, i.e.

$$\operatorname{div}(\operatorname{Ein})_\nu = \nabla_\mu (g^{\mu\eta} E_{\eta\nu}) = \nabla_\mu (E^\mu{}_\nu) = 0$$

(iv) : Le tenseur d'Einstein et le tenseur de Ricci s'annule simultanément, i.e. ;

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

Preuve. (i) : La symétrie du tenseur de Ricci découle de la **bisymétrie** du tenseur de Riemann covariant (c.f. Proposition 5.1-(iii)) : ainsi nous avons successivement

$$R_{\mu\nu} = R^\alpha{}_{\mu\alpha\nu} = g^{\alpha\beta}R_{\beta\mu\alpha\nu} = g^{\beta\alpha}R_{\alpha\nu\beta\mu} = R^\beta{}_{\nu\beta\mu} = R_{\nu\mu}$$

(ii) : D'une part (c.f. Proposition 5.2), la connexion de Levi-Civita ∇ étant sans torsion, nous pouvons utiliser la **seconde identité de Bianchi** soit, sous forme condensée $R^\tau{}_{\sigma[\mu\nu;\epsilon]} = 0$ (c.f. Remarque 5.3). D'autre part, ∇ étant aussi métrique, la divergence de la métrique est identiquement nulle (c.f. [Oli15a, Théorème 14.3]) et le relèvement/abaissement des indices commute avec les dérivations partielles covariantes ∇_μ . Ainsi, nous avons successivement :

$$\begin{aligned} R^\tau{}_{\sigma[\mu\nu;\epsilon]} &= 0 \\ \nabla_\epsilon(R^\tau{}_{\sigma\mu\nu}) + \nabla_\nu(R^\tau{}_{\sigma\epsilon\mu}) + \nabla_\mu(R^\tau{}_{\sigma\nu\epsilon}) &= 0 \\ \nabla_\epsilon(R_{\sigma\nu}) - \nabla_\nu(R_{\sigma\epsilon}) + \nabla_\mu(R^\mu{}_{\sigma\nu\epsilon}) &= 0 \quad (\text{avec } \tau = \mu) \\ \nabla_\epsilon(g_{\sigma\eta}R^\eta{}_{\nu}) - \nabla_\nu(g_{\sigma\eta}R^\eta{}_{\epsilon}) + \nabla_\mu(R^\mu{}_{\sigma\nu\epsilon}) &= 0 \\ \nabla_\epsilon(R^\eta{}_{\nu}) - \nabla_\nu(R^\eta{}_{\epsilon}) + \nabla_\mu(g^{\eta\sigma}g^{\mu\rho}R_{\rho\sigma\nu\epsilon}) &= 0 \\ \nabla_\epsilon(R^\eta{}_{\nu}) - \nabla_\nu(R^\eta{}_{\epsilon}) - \nabla_\mu(g^{\mu\rho}g^{\eta\sigma}R_{\rho\sigma\nu\epsilon}) &= 0 \\ \nabla_\epsilon(R^\eta{}_{\nu}) - \nabla_\nu(R^\eta{}_{\epsilon}) - \nabla_\mu(g^{\mu\rho}R^\eta{}_{\rho\nu\epsilon}) &= 0 \\ \nabla_\epsilon(R^\nu{}_{\nu}) - \nabla_\nu(R^\nu{}_{\epsilon}) - \nabla_\mu(g^{\mu\rho}R_{\rho\epsilon}) &= 0 \quad (\text{avec } \eta = \nu) \end{aligned}$$

soit encore $\nabla_\epsilon(R) = 2\nabla_\nu(R^\nu{}_{\epsilon})$: on conclut du fait que $\nabla_\epsilon(R) = \partial_\epsilon(R)$.

(iii) : L'expression de la divergence du tenseur de Ricci entraîne que

$$\nabla_\mu(R^\mu{}_{\nu}) = \frac{1}{2}\delta^\mu{}_{\nu}\nabla_\mu(R)$$

Ainsi, nous pouvons calculer la ν -ème composante de la divergence de Ein en écrivant :

$$\begin{aligned} \text{div}(\text{Ein})_\nu &= \nabla_\mu(E^\mu{}_{\nu}) = \nabla_\mu(R^\mu{}_{\nu}) - \frac{1}{2}\nabla_\mu(\delta^\mu{}_{\nu}R) \\ &= \frac{1}{2}\delta^\mu{}_{\nu}\nabla_\mu(R) - \frac{1}{2}\nabla_\mu(g^{\mu\epsilon}g_{\epsilon\nu}R) = \frac{1}{2}\delta^\mu{}_{\nu}\nabla_\mu(R) - \frac{1}{2}(g^{\mu\epsilon}g_{\epsilon\nu})\nabla_\mu(R) = 0 \end{aligned}$$

(iv) : Par définition du tenseur d'Einstein $R_{\mu\nu} = 0$ entraîne $E_{\mu\nu} = 0$. Pour la réciproque nous partons de nouveau de la définition de Ein d'où $R^\mu{}_{\nu} - 1/2\delta^\mu{}_{\nu}R = E^\mu{}_{\nu}$: pour $\mu = \nu$ il vient $R/2 = E^\nu{}_{\nu} = E$ de sorte que $R_{\mu\nu} = g_{\mu\nu}E + E_{\mu\nu}$. Finalement $E_{\mu\nu} = 0$ entraîne $R_{\mu\nu} = 0$. \square

Le travail d'Einstein sur la Relativité Générale, se déroule sur une dizaine d'année et prend sa finale dans son exposé magistral de 1916 [Ein16] (c.f. [Eis02] pour une histoire de la Relativité Générale). Chez Gallilé et Newton, les mouvements inertiels (vus d'un repère inertiel/galliléen) sont rectilignes et uniformes. Comme conséquence de la gravitation universelle de Newton, une distribution de masse ρ (définie en tout point de l'espace) engendre un champ de force conservatif, i.e. dérivant d'un potentiel Φ . La

relation entre la distribution de masse ρ et le potentiel Φ est donnée par l'équation de Poisson, soit

$$(24) \quad \Delta(\Phi) = 4\pi G\rho$$

Ici, Δ est l'opérateur laplacien, $\pi = 3.14\dots$ est la constante mathématique et G la constante de la gravitation universelle. L'équation de Poisson découle de la version gravitationnelle du Théorème de Gauss et du Théorème du flux-divergence. Dans la Relativité Générale d'Einstein, la gravitation n'est plus considérée comme une force : elle est identifiée à la courbure de l'espace-temps. Les mouvements inertiels suivent (par définition) des géodésiques inertielles de l'espace-temps (variété lorentzienne munie de la connexion de Levi-Civita) dont la courbure est entièrement déterminée par la métrique. Les deux dernières années de mise au point de sa théorie de la gravitation ont été très difficiles pour Einstein. Durant cette période, il correspond régulièrement avec David Hilbert sur ce sujet précis, risquant de se faire doubler au dernier moment, pour la formulation des équations de la gravitation, par le grand mathématicien du moment.

Le tenseur énergie-impulsion ¹¹ T de la Relativité Générale est un tenseur de valence $(0, 2)$ décrivant la distribution spatio-temporelle de masse-énergie et dont la définition découle du tenseur énergie-impulsion de l'électrodynamique relativiste (c.f. [Wei72] et [Car13] par exemple). Ainsi (et nous l'admettons) T est un tenseur symétrique (i.e. $T_{\mu\nu} = T_{\nu\mu}$), s'annulant uniformément dans les portions d'espace vide de matière-énergie et qui de plus, satisfait le principe de conservation de l'énergie, en ce sens que

$$\operatorname{div}(T)_\nu = \nabla_\mu(g^{\mu\eta}T_{\eta\nu}) = \nabla_\mu(T^\mu{}_\nu) \equiv 0$$

Les équations d'Einstein affirment

$$(25) \quad R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}T_{\mu\nu}$$

où on reconnaît (à gauche) le tenseur d'Einstein et où (à droite) nous trouvons $\pi = 3.14\dots$, la constante de la gravitation G et la vitesse de la lumière c . Il est important de remarquer que l'égalité entre le tenseur d'Einstein et le tenseur énergie-impulsion dans (20) est consistante relativement à la symétrie de T et au principe de conservation de l'énergie $\operatorname{div}(T) = 0$ (c.f. Proposition 7.1) : ceci n'est qu'une partie de l'ensemble des justifications du contenu physique de cette équation (voir de nouveau [Wei72] pour une analyse systématique de l'équation d'Einstein et [Car13] pour une heuristique très vivante).

Remarquons pour finir que le tenseur énergie-impulsion T s'annule dans les portions de l'espace qui sont vides de toute matière-énergie : ainsi, le fait (c.f. (iv)-Proposition 7.1) que le tenseur d'Einstein et le tenseur de Ricci s'annulent simultanément implique que les équations d'Einstein dans le vide se réduisent à

$$(26) \quad R_{\mu\nu} = 0$$

8. Interprétation géodésique du tenseur de Riemann

Soit X un point de M et U, V deux champs de vecteurs. Nous rappelons que $y \mapsto \exp_X(tU|_X)$ dénote la géodésique de M passant initialement par X à la vitesse $U|_X$. Pour

11. À ne pas confondre avec le tenseur de torsion qui est nul pour la connexion de Levi-Civita.

ε fixé (petit), nous notons $H = H(\varepsilon) := \exp_X(\varepsilon U|_X)$, de sorte qu'en composante ¹²

$$(27) \quad H^\sigma = X^\sigma + \varepsilon U^\sigma(X) - \frac{\varepsilon^2}{2} U^\mu(X) U^\nu(X) \Gamma^\sigma_{\mu\nu}(X) + \mathcal{O}(\varepsilon^3)$$

D'autre part, si $\hat{V}(H)$ est le vecteur de $T_H M$ obtenu par le transport parallèle de $V|_X$ le long de $t \mapsto \exp_X(tU|_X)$ en H , alors :

$$(28) \quad \hat{V}^\sigma(H) = V^\sigma(X) - \varepsilon U^\mu(X) V^\nu(X) \Gamma^\sigma_{\mu\nu}(X) + \mathcal{O}(\varepsilon^2)$$

Par suite, pour $I = I(\varepsilon) := \exp_H(\varepsilon \hat{V}(H))$ il vient :

$$\begin{aligned} I^\sigma &= H^\sigma + \varepsilon \hat{V}^\sigma(H) - \frac{\varepsilon^2}{2} \hat{V}^\mu(H) \hat{V}^\nu(H) \Gamma^\sigma_{\mu\nu}(H) + \mathcal{O}(\varepsilon^3) \\ &= H^\sigma + \varepsilon \hat{V}^\sigma(H) - \frac{\varepsilon^2}{2} V^\mu(X) V^\nu(X) \Gamma^\sigma_{\mu\nu}(X) + \mathcal{O}(\varepsilon^3) \end{aligned}$$

soit encore, compte tenu (27) et (28) :

$$(29) \quad \begin{aligned} I^\sigma &= X^\sigma + \varepsilon (U^\sigma(X) + V^\sigma(X)) - \varepsilon^2 U^\mu(X) V^\nu(X) \Gamma^\sigma_{\mu\nu}(X) \\ &\quad - \frac{\varepsilon^2}{2} (U^\mu(X) U^\nu(X) + \hat{V}^\mu(X) \hat{V}^\nu(X)) \Gamma^\sigma_{\mu\nu}(X) + \mathcal{O}(\varepsilon^3) \end{aligned}$$

Maintenant, en renversant les rôles des champs de vecteurs U et V nous définissons $K = K(\varepsilon)$ (de manière analogue à H) et $J = J(\varepsilon)$ (de manière analogue à I), ce qui donne symétriquement à (29)

$$(30) \quad \begin{aligned} J^\sigma &= X^\sigma + \varepsilon (V^\sigma(X) + U^\sigma(X)) - \varepsilon^2 V^\mu(X) U^\nu(X) \Gamma^\sigma_{\mu\nu}(X) \\ &\quad - \frac{\varepsilon^2}{2} (V^\mu(X) V^\nu(X) + \hat{U}^\mu(X) \hat{U}^\nu(X)) \Gamma^\sigma_{\mu\nu}(X) + \mathcal{O}(\varepsilon^3) \end{aligned}$$

Finalement en combinant (29) et (30), nous obtenons

$$(31) \quad I^\sigma - J^\sigma = \varepsilon^2 U^\mu(X) V^\nu(X) (\Gamma^\sigma_{\nu\mu}(X) - \Gamma^\sigma_{\mu\nu}(X)) + \mathcal{O}(\varepsilon^3)$$

À partir de maintenant, nous supposons que la connexion D est sans torsion : cette hypothèse simplificatrice entraîne par (31) que $I^\sigma - J^\sigma = I^\sigma(\varepsilon) - J^\sigma(\varepsilon) = \mathcal{O}(\varepsilon^3)$. En d'autres termes, l'approximation $I(\varepsilon) \approx J(\varepsilon)$ est raisonnable lors d'un calcul au second ordre en ε . Considérons alors un troisième champ de vecteur, soit W . Par définition $\hat{W}(H)$ (resp. $\hat{W}(I)$) est le vecteur de $T_H M$ (resp. $T_I M$) obtenu par le transport parallèle de $W|_X$ (resp. $\hat{W}(H)$) le long de la géodésique $t \mapsto \exp_X(tU|_X)$ (resp. $t \mapsto \exp_H(t\hat{V}(\varepsilon))$). En échangeant les rôles de U et V nous définissons les vecteurs $\hat{W}(K) \in T_K M$ et $\hat{W}(J) \in T_J M$ de manière analogue aux vecteur $\hat{W}(H)$ et $\hat{W}(I)$ respectivement.

Afin de calculer le développement en $\mathcal{O}(\varepsilon^3)$ de $\hat{W}(I)$ (et symétriquement pour $\hat{W}(J)$) nous notons $\gamma(t) := \exp_X(tU|_X)$, ce qui signifie que

$$(32) \quad (\forall \tau) \quad \ddot{\gamma}^\tau(t) + \dot{\gamma}^\mu(t) \dot{\gamma}^\nu(t) \Gamma^\tau_{\mu\nu}(\gamma(t)) = 0 \quad \text{avec} \quad (\gamma(0), \dot{\gamma}(0)) = (X, U|_X)$$

12. Rappelons qu'ici nous supposons que la variété M est un ouvert de \mathbb{R}^n et que nous supposons être dans le domaine d'un système de coordonnées locales $\mathbf{x} = (x^1, \dots, x^n)$: ainsi, pour tout point X du domaine des coordonnées locales, nous notons (X^1, \dots, X^n) les coordonnées cartésiennes de $\mathbf{x}(X)$. Par suite si $v = v^\mu \partial_\mu|_X$ est un vecteur de $T_X M \cong \mathbb{R}^n$ (suffisamment petit), et si (Y^1, \dots, Y^n) sont les coordonnées cartésiennes de $\mathbf{x}(Y)$ pour $Y = X + v$, alors $Y^\mu = X^\mu + v^\mu$ ($\forall \mu$).

Considérons alors le champ de vecteurs $t \mapsto w(t)$ obtenu par le transport parallèle du vecteur $W|_X = w(0) \in T_X M$ le long de γ : en d'autres termes, nous avons

$$(33) \quad (\forall \tau) \quad \dot{w}^\tau(t) + \dot{\gamma}^\mu(t)w^\nu(t)\Gamma^\tau_{\mu\nu}(\gamma(t)) = 0 \quad \text{avec} \quad w^\tau(0) = W^\tau(X)$$

Nous appliquons la **méthode d'Euler au second ordre** en considérant le développement de Taylor de $w^\tau(t)$ en $t = 0$, soit :

$$(34) \quad w^\tau(t) = w^\tau(0) + t\dot{w}^\tau(0) + \frac{t^2}{2}\ddot{w}^\tau(0) + \mathcal{O}(t^3)$$

Comme $\dot{\gamma}^\tau(0) = U^\tau(X)$ et que $w^\tau(0) = W^\tau(X)$, l'équation différentielle du problème de Cauchy en (33) nous donne directement

$$(35) \quad \dot{w}^\tau(0) = -U^\mu(X)W^\sigma(X)\Gamma^\tau_{\mu\sigma}(X)$$

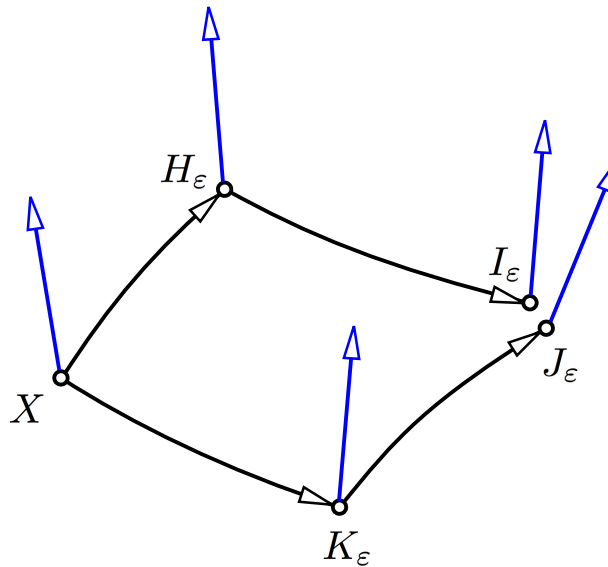


FIGURE 3. Soient U et V deux champs de vecteurs et X un point de M . Pour ε suffisamment petit fixé, nous définissons les points $H = \exp_X(\varepsilon U|_X)$, $I = \exp_H(\varepsilon \hat{V}(\varepsilon))$, $K = \exp_X(\varepsilon V|_X)$ et $J = \exp_K(\varepsilon \hat{U}(\varepsilon))$: ici $\hat{V}(\varepsilon)$ (resp. $\hat{U}(\varepsilon)$) est obtenu par le transport parallèle de $V|_X$ (resp. $U|_X$) le long de la géodésique issue de X et de vitesse initiale $U|_X$ (resp. $V|_X$). Alors,

$$I - J = \mathcal{O}(\varepsilon^3)$$

dès que **la connexion D est sans torsion**. Étant donné W un troisième champ de vecteurs, $\hat{W}(I)$ est obtenu par le transport parallèle de $W|_X$ le long de la géodésique issue de X et de vitesse initiale $U|_X$, suivi du transport parallèle le long de la géodésique issue de H et de vitesse initiale $\hat{V}(\varepsilon)$. Le vecteur $\hat{W}(J)$ étant obtenu symétriquement, en échangeant les rôles de U et V , on peut alors montrer que

$$\hat{W}(I) - \hat{W}(J) = \varepsilon^2 R(U, V)W|_X + \mathcal{O}(\varepsilon^3)$$

Nous pouvons utiliser cette dernière expression de $\dot{w}^\tau(0)$ avec une deuxième application de (33), de sorte qu'avec l'expression de $\dot{\gamma}(0)$ donnée par (32) il vient :

$$\begin{aligned}
\ddot{w}^\tau(0) &= -\left. \frac{d}{dt} \right|_{t=0} \left(\dot{\gamma}^\nu(t) w^\sigma(t) \Gamma^\tau_{\nu\sigma}(\gamma(t)) \right) \\
&= -\left(-U^\mu(X) U^\epsilon(X) \Gamma^\nu_{\mu\epsilon}(X) W^\sigma(X) \Gamma^\tau_{\nu\sigma}(X) \right. \\
&\quad \left. - U^\nu(X) U^\mu(X) \hat{W}^\epsilon(X) \Gamma^\sigma_{\mu\epsilon}(X) \Gamma^\tau_{\nu\sigma}(X) + U^\nu(X) W^\sigma(X) U^\mu(X) \partial_\mu|_X (\Gamma^\tau_{\nu\sigma}) \right) \\
&= -U^\mu(X) \left(-U^\nu(X) \Gamma^\epsilon_{\mu\nu}(X) W^\sigma(X) \Gamma^\tau_{\epsilon\sigma}(X) \right. \\
&\quad \left. - U^\nu(X) \hat{W}^\sigma(X) \Gamma^\epsilon_{\mu\sigma}(X) \Gamma^\tau_{\nu\epsilon}(X) + U^\nu(X) W^\sigma(X) \partial_\mu|_X (\Gamma^\tau_{\nu\sigma}) \right) \\
&= -U^\mu(X) U^\nu(X) W^\sigma(X) \left(-\Gamma^\epsilon_{\mu\nu}(X) \Gamma^\tau_{\epsilon\sigma}(X) - \Gamma^\epsilon_{\mu\sigma}(X) \Gamma^\tau_{\nu\epsilon}(X) + \partial_\mu|_X (\Gamma^\tau_{\nu\sigma}) \right)
\end{aligned}$$

soit encore :

$$(36) \quad \ddot{w}^\tau(0) = -U^\mu(X) U^\nu(X) W^\sigma(X) \Theta^\tau_{\sigma\mu\nu}(X)$$

où nous avons posé

$$(37) \quad \Theta^\tau_{\sigma\mu\nu} = \left(\partial_\mu (\Gamma^\tau_{\nu\sigma}) - \Gamma^\epsilon_{\mu\sigma} \Gamma^\tau_{\nu\epsilon} \right) - \Gamma^\epsilon_{\mu\nu} \Gamma^\tau_{\epsilon\sigma}$$

À partir du développement de Taylor de w^τ en (34) et compte tenu de la concition initiale sur $w^\tau(0)$ en (33) et des expressions de $\dot{w}^\tau(0)$ et $\ddot{w}^\tau(0)$ en (35) et (36), il vient :

$$(38) \quad \hat{W}^\tau(H) = W^\tau(X) - \varepsilon U^\mu(X) W^\sigma(X) \Gamma^\tau_{\mu\sigma}(X) - \frac{\varepsilon^2}{2} U^\mu(X) U^\nu(X) W^\sigma(X) \Theta^\tau_{\sigma\mu\nu}(X) + \mathcal{O}(\varepsilon^3)$$

De même (en prenant cette fois $\gamma(t) = \exp_H(t\hat{V}(H))$) nous obtenons

$$(39) \quad \hat{W}^\tau(I) = \hat{W}^\tau(H) - \varepsilon \hat{V}^\mu(H) \hat{W}^\sigma(H) \Gamma^\tau_{\mu\sigma}(H) - \frac{\varepsilon^2}{2} \hat{V}^\mu(H) \hat{V}^\nu(H) \hat{W}^\sigma(H) \Theta^\tau_{\sigma\mu\nu}(H) + \mathcal{O}(\varepsilon^3)$$

Pour obtenir le développement de $\hat{W}^\tau(I)$ à partir du point X , nous pouvons d'une part remplacer $\hat{W}^\tau(H)$ par son expression en (38) et d'autre part,

$$-\frac{\varepsilon^2}{2} \hat{V}^\mu(H) \hat{V}^\nu(H) \hat{W}^\sigma(H) \Theta^\tau_{\sigma\mu\nu}(H) = -\frac{\varepsilon^2}{2} V^\mu(X) V^\nu(X) \hat{W}^\sigma(X) \Theta^\tau_{\sigma\mu\nu}(X) + \mathcal{O}(\varepsilon^3)$$

Le terme restant s'obtient comme suit :

$$\begin{aligned}
-\varepsilon \hat{V}^\mu(H) \hat{W}^\sigma(H) \Gamma^\tau_{\mu\sigma}(H) &= -\varepsilon V^\mu(X) W^\sigma(X) \Gamma^\tau_{\mu\sigma}(X) \\
&\quad -\varepsilon^2 \left(-U^\nu(X) V^\epsilon(X) \Gamma^\mu_{\nu\epsilon}(X) W^\sigma(X) \Gamma^\tau_{\mu\sigma}(X) \right. \\
&\quad \quad -V^\mu(X) U^\nu(X) W^\epsilon(X) \Gamma^\sigma_{\nu\epsilon}(X) \Gamma^\tau_{\mu\sigma}(X) \\
&\quad \quad \left. +V^\mu(X) W^\sigma(X) U^\nu(X) \partial_\nu|_X (\Gamma^\tau_{\mu\sigma}(X)) \right) + \mathcal{O}(\varepsilon^3) \\
&= -\varepsilon V^\mu(X) W^\sigma(X) \Gamma^\tau_{\mu\sigma}(X) \\
&\quad -\varepsilon^2 U^\nu(X) \left(-V^\mu(X) \Gamma^\epsilon_{\nu\mu}(X) W^\sigma(X) \Gamma^\tau_{\epsilon\sigma}(X) \right. \\
&\quad \quad -V^\mu(X) W^\sigma(X) \Gamma^\epsilon_{\nu\sigma}(X) \Gamma^\tau_{\mu\epsilon}(X) \\
&\quad \quad \left. +V^\mu(X) W^\sigma(X) \partial_\nu|_X (\Gamma^\tau_{\mu\sigma}(X)) \right) + \mathcal{O}(\varepsilon^3) \\
&= -\varepsilon V^\mu(X) W^\sigma(X) \Gamma^\tau_{\mu\sigma}(X) \\
&\quad -\varepsilon^2 U^\nu(X) V^\mu(X) W^\sigma(X) \left(-\Gamma^\epsilon_{\nu\mu}(X) \Gamma^\tau_{\epsilon\sigma}(X) - \Gamma^\epsilon_{\nu\sigma}(X) \Gamma^\tau_{\mu\epsilon}(X) \right. \\
&\quad \quad \left. +\partial_\nu|_X (\Gamma^\tau_{\mu\sigma}(X)) \right) + \mathcal{O}(\varepsilon^3)
\end{aligned}$$

soit (modulo quelques indices muets et) en utilisant les symboles $\Theta^\tau_{\sigma\mu\nu}$ définis en (37)

$$\begin{aligned}
-\varepsilon \hat{V}^\nu(H) \hat{W}^\sigma(H) \Gamma^\tau_{\nu\sigma}(H) &= -\varepsilon V^\nu(X) W^\sigma(X) \Gamma^\tau_{\nu\sigma}(X) \\
&\quad -\varepsilon^2 U^\mu(X) V^\nu(X) W^\sigma(X) \Theta^\tau_{\sigma\mu\nu}(X) + \mathcal{O}(\varepsilon^2)
\end{aligned}$$

Finalement en utilisant le développement en (39) de $\hat{W}^\tau(I)$ devient :

$$\begin{aligned}
(40) \quad \hat{W}^\tau(I) &= W^\tau(X) - \varepsilon \left(U^\mu(X) + V^\mu(X) \right) W^\sigma(X) \Gamma^\tau_{\mu\sigma}(X) \\
&\quad -\frac{\varepsilon^2}{2} \left(U^\mu(X) U^\nu(X) - 2U^\mu(X) V^\nu(X) \right. \\
&\quad \quad \left. +V^\mu(X) V^\nu(X) \right) W^\sigma(X) \Theta^\tau_{\sigma\mu\nu}(X) + \mathcal{O}(\varepsilon^3)
\end{aligned}$$

Comme symétriquement

$$\begin{aligned}
(41) \quad \hat{W}^\tau(J) &= W^\tau(X) - \varepsilon \left(V^\mu(X) + U^\mu(X) \right) W^\sigma(X) \Gamma^\tau_{\mu\sigma}(X) \\
&\quad -\frac{\varepsilon^2}{2} \left(V^\mu(X) V^\nu(X) - 2V^\mu(X) U^\nu(X) \right. \\
&\quad \quad \left. +U^\mu(X) U^\nu(X) \right) W^\sigma(X) \Theta^\tau_{\sigma\mu\nu}(X) + \mathcal{O}(\varepsilon^3)
\end{aligned}$$

nous obtenons, par soustraction membre à membre de (40) et (41) :

$$\begin{aligned}
\hat{W}^\tau(I) - \hat{W}^\tau(J) &= -\varepsilon^2 \left(V^\mu(X) U^\nu(X) - U^\mu(X) V^\nu(X) \right) W^\sigma(X) \Theta^\tau_{\sigma\mu\nu}(X) + \mathcal{O}(\varepsilon^3) \\
&= \varepsilon^2 U^\mu(X) V^\nu(X) W^\sigma(X) \left(\Theta^\tau_{\sigma\mu\nu}(X) - \Theta^\tau_{\sigma\nu\mu}(X) \right) + \mathcal{O}(\varepsilon^3)
\end{aligned}$$

Or, en prenant en compte l'expression des valeurs prises par Θ données en (37), il vient

$$\begin{aligned}
\Theta^\tau_{\sigma\mu\nu} - \Theta^\tau_{\sigma\nu\mu} &= \left(\partial_\mu (\Gamma^\tau_{\nu\sigma}) - \Gamma^\epsilon_{\mu\sigma} \Gamma^\tau_{\nu\epsilon} \right) - \Gamma^\epsilon_{\nu\mu} \Gamma^\tau_{\epsilon\sigma} \\
&\quad - \left(\partial_\nu (\Gamma^\tau_{\mu\sigma}) - \Gamma^\epsilon_{\nu\sigma} \Gamma^\tau_{\mu\epsilon} \right) + \Gamma^\epsilon_{\mu\nu} \Gamma^\tau_{\epsilon\sigma} \\
&= \left(\partial_\mu (\Gamma^\tau_{\nu\sigma}) - \Gamma^\epsilon_{\mu\sigma} \Gamma^\tau_{\nu\epsilon} \right) - \left(\partial_\nu (\Gamma^\tau_{\mu\sigma}) - \Gamma^\epsilon_{\nu\sigma} \Gamma^\tau_{\mu\epsilon} \right) = R^\tau_{\sigma\mu\nu}(X)
\end{aligned}$$

(où nous utilisons l'absence de torsion de D i.e. $\Gamma_{\mu\nu}^{\epsilon} = \Gamma_{\nu\mu}^{\epsilon}$) et finalement

$$(42) \quad \hat{W}^{\tau}(I) - \hat{W}^{\tau}(J) = \varepsilon^2 U^{\mu}(X) V^{\nu}(X) W^{\sigma}(X) R^{\tau}_{\sigma\mu\nu}(X) + \mathcal{O}(\varepsilon^3)$$

9. Courbure de Gauss

9.1. **Rayon de courbure d'une courbe plane.** Soit $(C) := r([a; b])$ le support d'une chemin (lisse) $r : [a; b] \rightarrow \mathbb{R}^2$ (avec $a < 0 < b$) dont le paramétrage est l'abscisse curviligne que nous notons s : en d'autres termes $\langle \dot{r}(s) | \dot{r}(s) \rangle \equiv 1$. La dérivée de $\langle \dot{r}(s) | \dot{r}(s) \rangle$ par rapport à s est donc identiquement nulle, ce qui revient à dire que $2 \langle \ddot{r}(s) | \dot{r}(s) \rangle \equiv 0$: ainsi, l'accélération $\ddot{r}(s)$ est toujours normale à la (tangente de la) courbe dirigée par la vitesse $\dot{r}(s)$. Considérons maintenant que (e_1, e_2, e_3) désigne la base canonique de \mathbb{R}^3 et que le plan du mouvement de (C) soit celui engendré par (e_1, e_2) : alors, tout paramètre s est associé à la base de Serret-Frenet $(\tau(s), n(s))$ telle que

$$(43) \quad \tau(s) = \dot{r}(s) \quad \text{et} \quad n(s) = e_3 \times \tau(s)$$

Par définition, la **courbure (algébrique)** de (C) en $r(s)$ est le réel¹³ $\kappa(r(s)) = \kappa(s)$ t.q.

$$(44) \quad \ddot{r}(s) = \kappa(s)n(s)$$

son **rayon de courbure (algébrique)** valant $1/\kappa(s)$. Notons $\alpha(s)$ une mesure de l'angle orienté $(e_1, \tau(s))$ de sorte que $\tau(s) = \cos(\alpha(s))e_1 + \sin(\alpha(s))e_2$ et $n(s) = -\sin(\alpha(s))e_1 + \cos(\alpha(s))e_2$. Il vient alors $d\tau/ds = \dot{\alpha}(s)n_s$ et $dn/ds = -\dot{\alpha}(s)\tau(s)$. Par définition de la courbure, $\kappa(s)$ s'identifie donc à $\dot{\alpha}(s)$, ce qui nous permet de conclure que

$$(45) \quad \frac{d\tau}{ds} = \kappa(s)n(s) \quad \text{et} \quad \frac{dn}{ds} = -\kappa(s)\tau(s)$$

Enfin, la dérivée troisième de $r(s)$ s'obtient en dérivant (44) et compte de (45), il vient :

$$(46) \quad \ddot{\ddot{r}}(s) = -\kappa^2(s)\tau(s) + \dot{\kappa}(s)n(s)$$

Pour une interprétation géométrique intuitive de la courbure et du rayon de courbure, nous considérons sans (trop de) pertes de généralités que $r(s)$ est défini sur un voisinage de $t = 0$ avec $r(0) = (0, 0)$ et les deux conditions $\dot{r}(0) = (1, 0)$ et¹⁴ $\kappa(0) > 0$. Pour comprendre la signification de la courbure $\kappa(0)$ (et du rayon de courbure $1/\kappa(0)$), considérons les cercles tangents, soit (C_a) de centre $(0, a)$ et de rayon a et dont l'équation cartésienne est $X^2 + Y^2 - 2aY = 0$. Si le point $(0, 0)$ est un point d'inflexion de (C) , alors il n'existe pas de voisinage de $(0, 0)$ ni de rayon a pour lesquels la courbe (C) soit à l'intérieur du cercle (C_a) . En excluant ce cas, l'idée est justement d'obtenir le minimum des $a > 0$ pour lequel la courbe (C) soit – dans un voisinage de $(0, 0)$ – à l'intérieur du cercle (C_a) : si nous notons $(x(s), y(s))$ les coordonnées cartésiennes du point de (C) de paramètre s , alors cela est équivalent à minimiser a de sorte que

$$\Psi(s) := x(s)^2 + y(s)^2 - 2ay(s) \geq 0$$

13. Nous notons simplement $\kappa(s)$ au lieu de la notation complète $\kappa(r(s))$, car le paramétrage de $t \mapsto r(s)$ est curviligne : la notation complète (que nous utiliserons par la suite) est plus rigoureuse, car la courbure ponctuelle est un propriété de la courbe qui ne dépend pas du paramétrage choisi.

14. Le cas $\kappa(0) < 0$ est symétrique, le cas d'une courbure nulle en $r(0)$ nécessitant une analyse spéciale (cela arrive par exemple lorsque $r(0)$ est un point d'inflexion de (C) , ou encore lorsque r est *approximativement* plat en 0).

pour $|s|$ suffisamment petit. Ici $(\tau(0); \mathbf{n}(0)) = (\mathbf{e}_1, \mathbf{e}_2)$ et il suit de (43), (44) et (46) que :

$$\begin{aligned}(\mathbf{x}(0), \dot{\mathbf{x}}(0), \ddot{\mathbf{x}}(0), \ddot{\mathbf{x}}(0)) &= (0, 1, 0, -\kappa(0)^2) \\(\mathbf{y}(0), \dot{\mathbf{y}}(0), \ddot{\mathbf{y}}(0), \ddot{\mathbf{y}}(0)) &= (0, 0, \kappa(0), \dot{\kappa}(0))\end{aligned}$$

En écrivant les développements limités de $\mathbf{x}(s)$ et $\mathbf{y}(s)$, nous obtenons

$$\Psi(s) = s^2 - 2a \left(\frac{s^2}{2} \kappa(0) + \frac{s^3}{6} \dot{\kappa}(0) \right) + o(s^3) = s^2(1 - a\kappa(0)) - a \frac{s^3}{3} \dot{\kappa}(0) + o(s^3)$$

et par suite, $\Psi(s)$ est positif pour $|s|$ suffisamment petit dès que a vérifie $1 \geq a\kappa(0)$: le maximum d'un tel a correspond bien au rayon de courbure de (C) en $\mathbf{r}(0)$.

Nous terminons ce paragraphe par le cas particulier de la courbure en un point $(t, \varphi(t))$ du graphe d'une fonction $t \mapsto \varphi(t)$ définie (et lisse) dans un voisinage de 0 (cela nous permettra une interprétation simple de la courbure de Gauss d'une surface en un point). L'abscisse curviligne s étant définie à partir du point $(0, \varphi(0))$ en posant

$$s(t) = \int_0^t (1 + \dot{\varphi}^2(x))^{1/2} dx$$

nous pouvons vérifier que la base de Serret-Frenet, vérifie

$$\tau(s(t)) = \frac{1}{\sqrt{1 + \dot{\varphi}^2(t)}} \mathbf{e}_1 + \frac{\dot{\varphi}(t)}{\sqrt{1 + \dot{\varphi}^2(t)}} \mathbf{e}_2 \quad \text{et} \quad \mathbf{n}(s(t)) = -\frac{\dot{\varphi}(t)}{\sqrt{1 + \dot{\varphi}^2(t)}} \mathbf{e}_1 + \frac{1}{\sqrt{1 + \dot{\varphi}^2(t)}} \mathbf{e}_2$$

Alors, d'une part

$$\begin{aligned}\frac{ds}{dt} \frac{d\tau}{ds}(s(t)) &= \frac{d\tau \circ s}{dt} = -\frac{\dot{\varphi}(t)\ddot{\varphi}(t)}{(1 + \dot{\varphi}^2(t))^{3/2}} \mathbf{e}_1 + \left(\frac{\ddot{\varphi}(t)}{(1 + \dot{\varphi}^2(t))^{1/2}} - \frac{\dot{\varphi}(t)^2 \ddot{\varphi}(t)}{(1 + \dot{\varphi}^2(t))^{3/2}} \right) \mathbf{e}_2 \\ &= \frac{\ddot{\varphi}(t)}{(1 + \dot{\varphi}^2(t))} \mathbf{n}(s(t))\end{aligned}$$

et comme d'autre part $ds = (1 + \dot{\varphi}^2(t))^{1/2} dt$ nous pouvons conclure que

$$\frac{d\tau}{ds}(s(t)) = \frac{dt}{ds} \left(\frac{\ddot{\varphi}(t)}{(1 + \dot{\varphi}^2(t))} \right) \mathbf{n}(s(t)) = \frac{\ddot{\varphi}(t)}{(1 + \dot{\varphi}^2(t))^{3/2}} \mathbf{n}(s(t))$$

En d'autres termes, la courbure au point $(t, \varphi(t))$ vaut

$$\kappa(t, \varphi(t)) = \frac{\ddot{\varphi}(t)}{(1 + \dot{\varphi}^2(t))^{3/2}}.$$

Remarque 9.1. Si $\dot{\varphi}(0) = 0$ alors nous avons simplement $\kappa(0, \varphi(0)) = \ddot{\varphi}(0)$.

9.2. Courbure de Gauss. Soit $(x^1, x^2) \mapsto f(x^1, x^2)$ une application numérique définie sur un ouvert \mathcal{V} de \mathbb{R}^2 t.q. $f(0, 0) = 0$. Nous considérons que $X = (0, 0, 0)$ est un point de la surface $\Sigma := \{(x^1, x^2, f(x^1, x^2)) ; (x^1, x^2) \in \mathcal{V}\}$ et, pour simplifier, nous supposons que $\partial_1 f(0, 0) = \partial_2 f(0, 0) = 0$. L'approximation différentielle de f en $(0, 0)$ donne $f(x^1, x^2) \approx x^i x^j H_{ij}$ où H de composantes $H_{ij} := \partial_{ij} f(0, 0)$ est appelée la matrice hessienne de f en $(0, 0)$. Il n'est pas évident que **la courbure de Gauss de Σ en X** , définie en posant ¹⁵

$$(47) \quad K := \det(H)$$

15. La demi trace $\text{tr}(H)/2$ est appelée la courbure moyenne de Σ en X : cette notion a été introduite par Sophie Germain à propos des vibrations d'une membrane (cette notion intervient dans l'étude des surfaces minimales).

soit une quantité covariante (voir dans [Spi99] la présentation du Théorème Egregium de Gauss). Commençons avec le changement de variables $(x^1, x^2) = A_\theta(u^1, u^2)$, où A_θ est la rotation d'angle ¹⁶ θ (pour l'instant arbitraire) et soit $f_\theta : A_\theta^*(\mathcal{V}) \rightarrow \mathbb{R}$ t.q. $f_\theta(u^1, u^2) := f \circ A_\theta(u^1, u^2)$ (de sorte que $f_0 = f$). Alors le hessien H_θ de f_θ en $(0, 0)$ s'obtient en écrivant

$$\frac{\partial^2 f_\theta}{\partial u^j \partial u^i} = \frac{\partial x^\mu}{\partial u^i} \frac{\partial x^\nu}{\partial u^j} \left(\frac{\partial^2 f}{\partial x^\nu \partial x^\mu} \circ A_\theta \right) \quad \text{où} \quad \begin{pmatrix} \partial x^1 / \partial u^1 & \partial x^1 / \partial u^2 \\ \partial x^2 / \partial u^1 & \partial x^2 / \partial u^2 \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$$

Finalement en $(u^1, u^2) = (0, 0)$ nous obtenons la relation matricielle $H_\theta = A_\theta^* H_0 A_\theta$. La matrice A_θ étant orthogonale, le déterminant $\det(H_\theta)$ ne dépend pas du paramètre θ et coïncide donc avec la courbure de Gauss $K = \det(H_0) = \det(H)$ définie en (47). En particulier, la matrice $H = H_0$ étant symétrique, elle possède une base orthonormée directe de vecteurs propres : si κ_{11} et κ_{22} sont les valeurs propres en question, il existe $\theta \in \mathbb{R}$ (déterminée modulo 2π) t.q.

$$H_\theta = \begin{pmatrix} \kappa_{11} & 0 \\ 0 & \kappa_{22} \end{pmatrix} \quad \text{où} \quad \kappa_{ii} = \frac{\partial^2 f_\theta}{\partial u^i \partial u^i}(0, 0)$$

Les valeurs propres κ_{11} et κ_{22} sont appelées les courbures principales de Σ en X : d'après la Remarque 9.1, elle **coïncident respectivement avec les courbures en 0 des applications partielles $f_\theta(\cdot, 0)$ et $f_\theta(0, \cdot)$** . Ainsi, la courbure gaussienne de Σ en X définie en (47) correspond au produit $\kappa_{11}\kappa_{22}$ des courbures principales de f_θ en $(0, 0)$. Cette analyse préliminaire montre une covariance *restreinte* de la courbure de Gauss : cependant, elle nous permet d'introduire les systèmes de coordonnées locales dits réduits comme suit.

Nous allons généraliser notre présentation par une localisation : pour cela, considérons maintenant que Σ est une variété de dimension 2, plongée dans \mathbb{R}^3 : l'heuristique précédente justifie – et nous admettons – que pour tout $X \in \Sigma$, **il existe un système de coordonnées locales $(\mathcal{U}, \mathbf{x})$ réduit en X** , en ce sens que $X \in \mathcal{U}$ et pour tout $Y \in \mathcal{U}$

$$\mathbf{x}(Y) = (u^1, u^2) \iff Y = X + \Phi(u^1, u^2, f(u^1, u^2)) =: \mathbf{X}(u^1, u^2)$$

où Φ est une isométrie de \mathbb{R}^3 et où l'application numérique $f : \mathbf{x}(\mathcal{U}) \rightarrow \mathbb{R}$ est telle que $f(0, 0) = \partial_1 f(0, 0) = \partial_2 f(0, 0) = 0$ et possède un hessien diagonale en $(0, 0)$ avec

$$\frac{\partial^2 f}{\partial u^i \partial u^i}(0, 0) = \kappa_{ii}$$

Il s'agit maintenant de voir comment la courbure de Gauss en X définie en posant

$$K(X) := \kappa_{11}\kappa_{22}$$

(dont la valeur dépend a priori du système de coordonnées réduit) peut-être interprétée comme une grandeur covariante sur Σ . Sans perte de généralités, nous pouvons d'abord revenir (localement en X) au cas considéré dans l'heuristique préliminaire, en prenant $X = (0, 0, 0)$, l'isométrie Φ étant supposée réduite à l'identité de \mathbb{R}^3 . Ici, le point clef est de munir Σ de sa structure riemannienne naturelle – soit g – induite par la structure euclidienne de l'espace ambiant (i.e. \mathbb{R}^3). Cette structure riemannienne est caractérisée par le fait que $g(V, W) = \langle V|W \rangle$ pour tout couple (V, W) de champs de vecteurs sur Σ . Afin d'identifier la métrique, nous commençons par remarquer que les vecteurs ∂_1 et ∂_2

16. A_θ est identifiée à sa matrice dans la base canonique de \mathbb{R}^2 d'où les identification $A_\theta^{-1} = A_\theta^*$.

de la base holonôme de \mathbf{x} , s'expriment simplement dans la base canonique de \mathbb{R}^3 à l'aide des dérivées partielles usuelles de $f(u^1, u^2)$ en coordonnées cartésiennes. En effet,

$$\partial_1 = \mathbf{e}_1 + \left(\frac{\partial f}{\partial u^1} \right) \mathbf{e}_3 = \quad \text{et} \quad \partial_2 = \mathbf{e}_2 + \left(\frac{\partial f}{\partial u^2} \right) \mathbf{e}_3$$

de sorte que pour $V = V^1 \partial_1 + V^2 \partial_2$ et $W = W^1 \partial_1 + W^2 \partial_2$ il vient $\langle V|W \rangle = V^i W^j g_{ij}$ avec

$$(48) \quad g_{ii} = 1 + \left(\frac{\partial f}{\partial u^i} \right)^2 \quad \text{et} \quad g_{12} = g_{21} = \frac{\partial f}{\partial u^1} \frac{\partial f}{\partial u^2}$$

et (c.f. [Oli15a, Appendice 17]) les expressions des symboles de Christoffel sont :

$$\mathbf{C}^k_{ij} = \frac{\partial^2 f}{\partial u^i \partial u^j} \frac{\partial f}{\partial u^k} \left(1 + \left(\frac{\partial f}{\partial u^1} \right)^2 + \left(\frac{\partial f}{\partial u^2} \right)^2 \right)^{-1}$$

Lemme 9.2. Soit (Σ, g) une variété riemannienne de dimension 2 plongée isométriquement dans \mathbb{R}^3 ; alors dans un système de coordonnées réduit en X nous avons

$$(i) : g_{\mu\nu}(X) = \delta_{\mu\nu} \quad (ii) : \partial_k|_X(g_{\mu\nu}) = 0 \quad (iii) : \mathbf{C}^k_{\mu\nu}(X) = 0$$

Proposition 9.3. Soit (Σ, g) une variété riemannienne de dimension 2 plongée isométriquement dans \mathbb{R}^3 et munie de la connexion de Levi-Civita; alors, dans un système de coordonnées réduit en X , la courbure de Gauss est reliée au tenseur de Riemann, en ce sens que

$$K(X) = -R_{1212}(X)$$

Preuve. Soit $(\mathcal{U}, \mathbf{x})$ un système de coordonnées locales, supposés réduit en un point X de Σ : sans perte de généralités, nous supposons que $X = (0, 0, 0) \in \mathcal{U}$ avec

$$\mathbf{x}(Y) = (u^1, u^2) \iff Y = (u^1, u^2, f(u^1, u^2)) =: \mathbf{X}(u^1, u^2)$$

où l'application numérique $f : \mathbf{x}(\mathcal{U}) \rightarrow \mathbb{R}$ est telle que $f(0, 0) = \partial_1 f(0, 0) = \partial_2 f(0, 0) = 0$ et possède un hessien diagonal en $(0, 0)$ avec

$$\frac{\partial^2 f}{\partial u^i \partial u^i}(0, 0) = \kappa_{ii}(X)$$

de sorte que la courbure de Gauss en X vaut $K = \kappa_{11}\kappa_{22}$. D'autre part, en combinant les composantes du tenseur de Riemann covariant en (21) aux identités du Lemma 9.2 (spécifiques aux coordonnées locales \mathbf{x} supposées réduites en X), nous obtenons :

$$(49) \quad R_{1212}(X) = \frac{1}{2} \left(\partial_{11}|_X(g_{22}) + \partial_{22}|_X(g_{11}) \right) - \partial_{12}|_X(g_{12})$$

Pour simplifier, notons $G_{ij} := g_{ij} \circ \mathbf{X}$; alors, avec (48), il vient :

$$\frac{\partial^2 G_{22}}{\partial u^1 \partial u^1} = \frac{\partial^2}{\partial u^1 \partial u^1} \left(\frac{\partial f}{\partial u^2} \right)^2 = 2 \frac{\partial}{\partial u^1} \left(\frac{\partial f}{\partial u^2} \frac{\partial^2 f}{\partial u^1 \partial u^2} \right) = 2 \left(\left(\frac{\partial^2 f}{\partial u^1 \partial u^2} \right)^2 + \frac{\partial f}{\partial u^2} \frac{\partial^3 f}{\partial u^1 \partial u^1 \partial u^2} \right)$$

et de même

$$\frac{\partial^2 G_{11}}{\partial u^2 \partial u^2} = 2 \left(\left(\frac{\partial^2 f}{\partial u^2 \partial u^1} \right)^2 + \frac{\partial f}{\partial u^1} \frac{\partial^3 f}{\partial u^2 \partial u^2 \partial u^1} \right)$$

Enfin

$$\begin{aligned} \frac{\partial^2 G_{12}}{\partial u^1 \partial u^2} &= \frac{\partial^2}{\partial u^1 \partial u^2} \left(\frac{\partial f}{\partial u^1} \frac{\partial f}{\partial u^2} \right) \\ &= \frac{\partial}{\partial u^1} \left(\frac{\partial^2 f}{\partial u^2 \partial u^1} \frac{\partial f}{\partial u^2} + \frac{\partial f}{\partial u^1} \frac{\partial^2 f}{\partial u^2 \partial u^2} \right) \\ &= \frac{\partial^3 f}{\partial u^1 \partial u^2 \partial u^1} \frac{\partial f}{\partial u^2} + \frac{\partial^2 f}{\partial u^2 \partial u^1} \frac{\partial^2 f}{\partial u^1 \partial u^2} + \frac{\partial^2 f}{\partial u^1 \partial u^1} \frac{\partial^2 f}{\partial u^2 \partial u^2} + \frac{\partial f}{\partial u^1} \frac{\partial^3 f}{\partial u^1 \partial u^2 \partial u^2} \end{aligned}$$

Utilisant de nouveau les propriétés des coordonnées réduites en X , il vient avec (49)

$$R_{1212}(X) = \partial_{12}|_X(g_{12}) = -\frac{\partial^2 f}{\partial u^1 \partial u^1}(0,0) \frac{\partial^2 f}{\partial u^2 \partial u^2}(0,0) = -\kappa_{11}(X)\kappa_{22}(X) = -K(X)$$

□

Les propriétés de symétrie du tenseur de Riemann covariant \mathcal{R} définie sur Σ (de dimension 2) assure que pour tout $X \in \Sigma$, la forme quadrilinéaire $\mathcal{R}|_X$ appartient à l'espace vectoriel $\mathcal{S}_B(T_X \Sigma)$ (c.f. § 2). Or nous savons (c.f. Remarque 2.2) que $\mathcal{S}_B(T_X \Sigma)$ est de dimension 1 (car $T_X \Sigma$ est de dimension 2). Nous pouvons même dire un peu plus en définissant sur Σ le tenseur $(U, V) \mapsto \Lambda(U, V)$ tel que $\Lambda|_X(U, V)$ soit le déterminant du couple de vecteurs $(U|_X, V|_X)$, alors il existe un champ scalaire $X \mapsto \kappa(X)$ t.q.

$$\mathcal{R} = -\kappa \Lambda \otimes \Lambda$$

Maintenant, nous considérons que x est un système de coordonnées réduit en un point X donné : alors, par définition des coordonnées réduites (c.f. Lemme 9.2), nous savons que $(\partial_1|_X, \partial_2|_X)$ est une base orthonormée de $T_X \Sigma$ et par suite :

$$R_{1212}(X) = \mathcal{R}|_X(\partial_1|_X, \partial_2|_X, \partial_1|_X, \partial_2|_X) = -\kappa(X)\Lambda(\partial_1|_X, \partial_2|_X)\Lambda(\partial_1|_X, \partial_2|_X) = -\kappa(X)$$

Il suit de la Proposition 9.3 que $\kappa(X)$ coïncide avec la courbure gaussienne de Σ en X , i.e. $\kappa(X) = K(X)$. Nous venons de démontrer une forme du Théorème Egregium :

Théorème 9.4 (Gauss). *Soit (Σ, g) une variété riemannienne de dimension 2 plongée isométriquement dans \mathbb{R}^3 et munie de la connexion de Levi-Civita ; alors, la courbure de Gauss $X \mapsto K(X)$ définie en (47) est le champ scalaire $X \mapsto \kappa(X)$ tel que*

$$(50) \quad \mathcal{R} = -K \Lambda \otimes \Lambda$$

10. Notes

Riemann généralise la notion de courbure introduite par Gauss dans son mémoire d'habilitation de 1854 [Rie68] (et publié à titre posthume en 1868). On trouve aussi l'expression des *composantes du tenseur de courbure* dans un article de 1861 en latin¹⁷ [Rie92] et consacré à la conduction de la chaleur dans un corps homogène (une traduction commentée se trouve dans [Spi99] ; voir aussi [FK90]). Dans ces travaux, Riemann produit une réflexion d'avant-garde dans un style qui rendra son travail obscur pour beaucoup de géomètres de son temps. Riemann n'aura pas le temps de développer ses idées : il meurt en 1866 à l'âge de 39 ans. Il est assez éclairant d'écouter Bishop & Goldberg à ce propos : « *Le travail de Riemann en 1854, était exprimé dans une forme si obscure, que Beltrami en 1869 et Lipschitz en 1872 n'ont utilisé le langage géométrique qu'avec une grande précaution.*

17. Soumis – sans succès – pour un prix de l'Académie des Sciences de Paris.

En fait, les développements étaient si lents que la notion de parallélisme due à Levi-Civita n'est apparue qu'en 1917. » Une mention spéciale doit être faite pour le travail de Spivak [Spi99] où grâce à une analyse détaillée des articles originaux de Gauss et Riemann, l'auteur parvient à faire ressortir les idées originales des deux grands mathématiciens.

Dans [Oli15a] nous couvrons grosso-modo le premier cycle de développement des idées (géométriques) de Riemann. Cela commence en 1869 (tout juste un an après la publication posthume du mémoire de 1854) avec le papier de Christoffel [Chr69], où est introduit la notion de connexion. Notons aussi comme point de référence le papier de Ricci (1888) [Ric88] sur la covariance et la contravariance, ainsi que le papier de synthèse [RTL00] écrit avec Levi-Civita et qui est considérée comme l'acte fondateur du calcul tensoriel (c.f. [Bou92]). Comme évoqué par Bishop & Goldberg, l'article 1917 de Levi-Civita [LC17] marque la fin d'un premier cycle de développement des idées de Riemann. Ici nous avons choisi la présentation du tenseur de courbure par l'utilisation systématique de la notion de connexion et où la connexion de Levi-Civita est donnée par le Théorème d'unicité de Koszul (c.f. [Kos50]). Ces résultats/méthodes datent des années 50, soit à peu près un siècle après les travaux Riemann : ils présentent l'avantage d'une redoutable efficacité permettant d'accéder directement aux bases de la théorie. L'inconvénient de cette approche est de sacrifier l'intuition géométrique : pour compléter cette dernière lacune, nous nous référons à l'ouvrage assez extraordinaire de Roger Penrose [Pen04] (Roger Penrose est un des spécialistes incontestés de la Relativité Générale).

L'entrée en scène d'Einstein dans ce concert de mathématiciens tient presque du romanesque, comme en témoigne ce petit mot qu'il écrit à Sommerfeld (1912)¹⁸ : « *Je travaille maintenant exclusivement sur la gravitation et je pense que je peux surmonter toutes les difficultés avec l'aide d'un de mes amis mathématicien. Mais une chose est certaine : de toute ma vie je n'ai presque jamais travaillé aussi durement et je suis désormais pénétré d'un grand respect pour les mathématiques que j'avais jusqu'alors, dans ma simplicité d'esprit, considéré la partie la plus subtile comme un luxe. Comparé à ce problème, la théorie originale de la relativité est un jeu d'enfant.* ». L'ami dont parle Einstein est Marcel Grossmann et les mathématiques en question sont la géométrie de Gauss et Riemann et le calcul tensoriel de Ricci et Levi-Civita. Leur collaboration aboutira à un article en deux parties [EG13] (la partie physique due à Einstein et la partie mathématique due à Grossmann) qui marque une étape importante vers la forme définitive de la Relativité Générale exposée dans [Ein16].

Nous laissons le dernier mot à Jacques Hadamard sur l'importance du travail d'Einstein sur la gravitation dans le développement de la géométrie moderne¹⁹ : « *À mon avis il y avait une crise en géométrie infinitésimale et l'avènement de la relativité venait de la surmonter. Ceci était le début d'un éclairage inattendu sur les précédents travaux en géométrie montrant l'importance du calcul différentiel absolu de Ricci et de Levi-Civita, ce calcul étant resté presque sans éveiller d'intérêt depuis sa création.* »

RÉFÉRENCES

[Bou92] J.-P. Bourguignon. Transport parallèle et connexions en géométrie et en physique, in "1830-1930 : a century of geometry". Springer-Verlag, Lect. Notes in Physics, 402 :150–164, 1992.

18. Citation tirée de [Eis02].

19. Citation tirée de [VM05].

- [Car13] S. Carroll. *Spacetime and Geometry : An Introduction to General Relativity*. Pearson New International Edition (second edition of the 2003 Addison-Wesley edition), 2013.
- [Chr69] E.B. Christoffel. Über die Transformation der homogenen Differentialausdrücke zweiten Grades. *J. reine angew. Math.*, 70 :46–70, 1869.
- [EG13] A. Einstein and M. Grossmann. Entwurf einer verallgemeinerten Relativitätstheorie und eine Theorie der Gravitation. I. Physikalischer Teil von A. Einstein II. Mathematischer Teil von M. Grossmann (Outline of a Generalized Theory of Relativity and of a Theory of Gravitation. I. Physical Part by A. Einstein II. Mathematical Part by M. Grossmann. *Z. Math. & Phys.*, 62 :225–244, 245–261, 1913.
- [Ein16] A. Einstein. Grundlage der allgemeinen Relativitätstheorie (The Foundation of the General Theory of Relativity). *Annalen der Physik*, 49(IV) :769–822, 1916.
- [Eis02] J. Eisenstaedt. *Einstein et la relativité générale*. CNRS Édition - Paris, 2002.
- [FK90] R. Farwell and C. Knee. The Missing Link : Riemann's "Commentatio," Differential Geometry and Tensor Analysis. *Hist. Math.*, 17 :223–255, 1990.
- [Kos50] J. L. Koszul. Homologie et cohomologie des algèbres de Lie. *Bulletin de la Société Mathématique de France*, 78 :65–127, 1950.
- [LC17] T. Levi-Civita. Nozione di parallelismo in una varietà qualunque e conseguente specificazione geometrica della curvatura riemanniana. *Rend. Circ. Mat. Palermo (1884-1940)*, 42 :173–204, 1917.
- [Oli15a] E. Olivier. 100 ans de Relativité Générale I/III : covariance. *Bull. Inf. App. & App.*, 100 :23–63, 2015.
- [Oli15b] E. Olivier. 100 ans de Relativité Générale III/III : déflexion de la lumière. *Bull. Inf. App. & App.*, 102 :139–167, 2015.
- [Pen04] R. Penrose. *The Road to Reality : A Complete Guide to the Laws of the Universe*. Science : Astrophysics. A.A. Knopf, 2004.
- [Ric88] G. Ricci. Delle derivazioni covarianti e contravarianti e del loro uso nella analisi applicata. *Stud. ed. Univ. Padova, Padova*, 3, 12 :23 p., 1888.
- [Rie68] B. Riemann. *Über die Hypothesen, welche der Geometrie zu Grunde liegen. (Habilitationsschrift – 1854)*, volume XIII. Abhandl. Kgl. Gesell. Wiss. Göttingen, 1868.
- [Rie92] B. Riemann. Commentatio mathematica, qua respondere tentatur quaestioni ab III^{ma} Academia Parisiensi propositae (1861). *Riemann's Gesamm. Math. Werke*, XXII :391–423, 1892.
- [RTL00] G. Ricci and Tullio T. Levi-Civita. Méthodes de calcul différentiel absolu et leurs applications. *Mathematische Annalen (Springer)*, 54 :125–201, 1900.
- [Spi99] M. Spivak. *A comprehensive introduction to differential geometry*. Number vol. 2 in A Comprehensive Introduction to Differential Geometry. Publish or Perish, inc. (Sec. Ed.), 1999.
- [VM05] T. Shaposhnikova V. Maz'ya. *Jacques Hadamard : Un mathématicien universel (traduction de l'anglais par G. Tronel)*. EDP Science (Ed. originale de l'Am. Math. Soc. - 1998), 2005.
- [Wal10] R.M. Wald. *General Relativity*. University of Chicago Press, 2010.
- [Wei72] S. Weinberg. *Gravitation and Cosmology – Principles and Applications of the General Theory of Relativity*. John Wiley, 1972.



VOUZZAVEDIBISAR : Der Dirigent - Le directeur (par Annagret TAESCHNER)

BIAA

Bulletin d'Informatique Approfondie & Applications
Computation - Information
Volumes 2015

100 - 101 - 102



Couverture : dessin de Michel Avezard (Zevard) interprété par la MMIAGe (1985-1987)

ÉDITORIAL : Charles CORGE et INFORMATIQUE QUANTIQUE

Jean - Michel KNIPPEL

Résumé. – Charles Corge est maintenant un chercheur retraité du Commissariat à l'énergie atomique. En 2011 paraissait un ouvrage sur l'informatique quantique. Eric Olivier dans sa trilogie sur 100 ans de relativité générale en a fait un de ses ouvrages de référence. Instantanément me revenait à l'esprit l'ouvrage de Charles Corge "Eléments d'informatique" aux éditions Larousse... de 1975.

Charles Corge fut l'un des tous premiers à mettre un ordinateur en ligne sur des expériences de physique nucléaire dans les années soixante. C'est pour moi l'auteur de 3 ouvrages fondamentaux de l'informatique :

Eléments d'informatique. Informatique et démarche de l'esprit. Larousse. 1975.

Machines de Turing et automates cellulaires. Du trait gravé au trait animé. Ellipses. 2008.

Informatique quantique qu'est-ce et pour quoi faire ? Ellipses. 2011.

Ces livres sont tous aussi importants qu'un autre triptyque de l'informatique qui reste une mine de références et de principes de base. Il faut garder en tête que le terme "Informatique" fut créé en 1962. Charles Corge fait partie de cette série informatique des éditions Larousse *université* :

Jean-Pierre Meinadier. *Structure et fonctionnement des ordinateurs.* 1971.

Charles Corge. *Eléments d'informatique. Informatique et démarche de l'esprit.* 1975.

Pierre Morvan, Michel Lucas. *Images et ordinateur. Introduction à l'infographie interactive.* 1976.

Que peut apporter de plus l'informatique quantique ? Charles Corge répond ainsi dans un entretien dans le journal en ligne Atlantico.fr¹ : « *L'informatique quantique apporte donc une nouvelle manière d'appréhender le traitement de l'information et le calcul numérique. Elle offre de nouveaux outils qui permettent de mieux comprendre les implications de la physique quantique et de mieux les tester* ».

Bien sûr, l'application pratique de la physique quantique à des problèmes scientifiques concrets nécessite l'utilisation de mathématiques difficiles, et même très difficiles quand on essaie de marier la physique quantique à la relativité d'Einstein.²

N'oublions pas que dans les années 1930 à Princeton se rencontraient Einstein, Gödel et Von Neumann. Le jeune doctorant Alan Turing expliquait son concept de calcul universel à son professeur de thèse, Alonzo Church. Mais la guerre ne rassembla pas ces hommes d'exception et peut être là est un élément du non développement de l'informatique quantique à cette époque. Einstein avait une vue tolérante du monde, Gödel, à partir des années 1940, se consacre principalement à la philosophie, mais ne publie rien de ses réflexions, Von Neumann est cynique et pessimiste sur l'avenir du monde et participe aux calculs pour optimiser le lancer des bombes atomiques sur le Japon. Quant à

1. Informatique quantique : pourquoi ce pourrait être un gigantesque pas en avant pour l'ère du digital

2. Swen Ortoli, Jean-Pierre Pharabod. *Le cantique des quantiques. Le monde existe-t-il ?* . La Découverte. 1984

Turing, il a préféré retourner en Angleterre participer à la mise au point d'un procédé et d'une machine permettant le déchiffrement du code secret utilisé par les sous-marins allemands pour correspondre avec leur base, l'*Enigma*.

Il se peut que des calculateurs quantiques, un jour, puissent supporter l'aléatoire, la généralité et les calculs sur les systèmes de numération et de répartition. Toutefois, il faudra être patient, car lire une information quantique, c'est la détruire et la copier n'est pas toujours possible.

The Imitation Game : et si vous préféreriez l'original ?

Jean LASSÈGUE ¹²

Résumé. – 2012 a été l'année du centenaire de la naissance d'Alan Turing. Cette icône de l'informatique a été célébrée dignement. A la fin du mois de janvier 2015, sort un film biographique sur Turing, un blockbuster, « The Imitation Game ». Binaire se devait d'accompagner l'événement et a donc demandé leurs réactions à plusieurs amis. Nous démarrons avec Jean Lassègue, philosophe des sciences. Il remet en jeu quelques idées reçues. Serge Abiteboul.



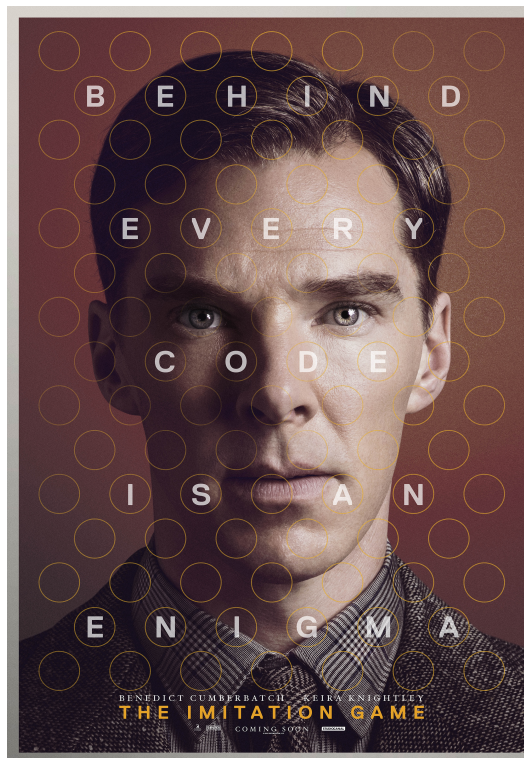
J'avais eu l'occasion de passer un an à Oxford il y a presque trente ans. J'étais alors tombé par hasard sur la biographie de Turing écrite par le mathématicien Andrew Hodges, *Alan Turing, The Enigma*. Je fus immédiatement subjugué par ses qualités littéraires : la beauté de cette prose, les exergues du poète Walt Whitman, mais aussi la culture britannique – des rites propres aux écoles secondaires à la campagne anglaise – me parlaient d'un monde que je connaissais bien mieux que la logique mathématique que j'avais commencé à étudier cette année-là. Je lus le livre d'une traite et le relus un certain nombre de

1. CNRS - Institut Marcel Mauss, École des Hautes Études en Sciences Sociales.

2. Paru dans un blog fantastique d'informaticien(ne)s de la Société Informatique de France « binaire.blog.lemonde.fr ».

fois dans les années qui suivirent, pendant lesquelles Turing fut pour moi tout à la fois une sorte d'interlocuteur, une grille d'interprétation et une source d'interrogation sur la façon dont la science s'inscrit dans le siècle.

Me voici de retour à Oxford. Entre temps, Turing est devenu l'une des figures les plus célèbres de l'histoire de l'informatique et, ce 18 décembre, on joue, au cinéma de Walton Street où j'allais de temps en temps à l'époque, le film de Morten Tyldum « The Imitation Game » qui porte sur la biographie de Turing et qui sortira en France le 28 janvier 2015. Quoi de plus normal, vingt-huit ans après avoir quitté Oxford, que d'aller voir ce film, et de renouer, sur le mode d'une fiction, avec ce qui allait devenir l'un des thèmes majeurs de mes préoccupations pendant les années qui suivirent ? Le film allait me rappeler encore une fois, s'il le fallait, combien il est difficile de parler d'un scientifique sans tomber dans l'hagiographie ou le grand spectacle, et ce, même quand le film est basé sur le livre si réussi d'Andrew Hodges.



Je ne compte pas faire une critique détaillée du film, décidément trop « théâtral » pour moi, même s'il m'a paru bien servi par les acteurs et qu'il peut sans doute permettre à un spectateur qui ignorerait tout de la vie de Turing de se faire une première idée du rôle capital qu'il joua pendant la guerre quand, en décryptant les messages codés envoyés aux sous-marins allemands faisant le blocus de l'Angleterre, il réussit à déjouer le blocus et raccourcit de ce fait la guerre de deux ans en sauvant des milliers de vies humaines. Cependant, réussir à faire tourner l'intrigue autour du couple Alan Turing / Joan Clarke, mathématicienne avec qui il fut fiancé à l'époque où ils faisaient tous les deux partie de

l'équipe de décryptage, tient pour moi d'un véritable prodige quand on sait, comme le montre d'ailleurs le film, que Turing était homosexuel et que son homosexualité lui valut condamnation pénale et contribua à son suicide. D'autres films récents, relevant plus du genre « documentaire » (voir en fin de blog) me paraissent plus proches de la vérité historique.

Je m'en tiendrai à un aspect particulier du film de Tyldum parce qu'il témoigne d'une attitude générale : sa fidélité stricte à l'interprétation de la biographie écrite par Andrew Hodges. Or il se trouve que, malgré les qualités éminentes que je reconnais volontiers au livre de Hodges dans lequel j'ai tant appris, je suis devenu, au fil du temps, fondamentalement en désaccord avec son interprétation. Le film « The Imitation Game », *par fidélité à Hodges*, me semble donc fondamentalement *infidèle à Turing* et c'est de cela dont je voudrais parler. Il ne s'agit pas seulement d'un débat entre spécialistes sur quelques vagues points de détail de l'histoire intellectuelle de Turing qui n'intéressent qu'eux et qui n'ont aucune portée. Il s'agit au contraire d'un point fondamental qui distingue deux façons radicalement différentes de concevoir la nature de l'informatique en général et le cadre philosophique et épistémologique de ses résultats.

L'interprétation canonique du parcours intellectuel de Turing

Voilà comment je vois la façon dont Hodges interprète la vie et l'œuvre de Turing, appuyée en cela par toute une tradition dérivée de l'empirisme logique dont l'audience est aujourd'hui mondiale, comme les cérémonies, hommages, colloques et événements pour le centenaire de Turing l'ont montré de par le monde en 2012. La notion fondamentale à étudier serait celle d'« intelligence » et la question pertinente à se poser serait celle de savoir si elle peut être conçue de façon mécanique ou pas. La réponse de Turing à cette question serait alors conçue comme le travail consistant à faire évoluer la mécanisation du renseignement ('intelligence' pris au sens anglais de collecte des données par les services secrets, ce qui renvoie au travail de Turing sur le décryptage au moyen de la machine Enigma des codes allemands pendant la guerre) à l'intelligence artificielle ('intelligence' pris au sens de l'esprit, ce qui renvoie à la construction du premier ordinateur et l'idée que le fonctionnement du cerveau est analogue à celui d'un ordinateur) : l'originalité – immense – de Turing serait ainsi d'avoir accompli ce passage du « renseignement » à l'« esprit » en s'en tenant strictement au paradigme mécanique tel qu'il a été pleinement réalisé par l'ordinateur. Dans cette optique, le « jeu de l'imitation », rebaptisé « Test de Turing » pour les besoins de la cause, deviendrait capital parce qu'il serait un « test » – entendez *algorithme mécanisable* - qui *prouverait* que la notion d'intelligence peut se concevoir comme détachée du support biologique de l'humanité et peut dès lors se transférer à un ordinateur, à peu près comme un logiciel peut tourner sur n'importe quel type de machine, pourvu qu'il s'agisse d'un ordinateur digital.

C'est à cette interprétation que je m'oppose parce qu'elle reconduit la différence logiciel et matériel (en reproduisant ce faisant un dualisme de l'âme et du corps) au lieu de tenter de penser cette différence et de la concevoir comme rendant possible une dynamique de leur rapport.

Un indice devrait tout d'abord nous mettre immédiatement la puce à l'oreille : qu'en est-il, dans cette interprétation désormais canonique, du travail de Turing en biologie théorique ? Il passe à la trappe, purement et simplement. Or je rappelle d'une part que les recherches de Turing en biologie théorique ont occupé toutes les dernières années de sa vie entre 1950 et 1954 une fois qu'il eut *définitivement* abandonné toute recherche fondamentale en informatique et d'autre part qu'il considérait ses résultats dans ce domaine comme *aussi fondamentaux* que ceux de son article de 1936 fondant la théorie de la calculabilité. S'agit-il donc seulement d'un passe-temps secondaire que l'on pourrait oublier, le temps d'un film grand public, ou minorer, comme dans l'interprétation canonique du parcours intellectuel de Turing ? C'est *impossible* si l'on veut rendre justice à ce que Turing disait de son propre travail en biologie théorique. Il faut donc reprendre complètement le cadre interprétatif proposé par Andrew Hodges et revenir à ce qui fait le nerf de la preuve de son argumentation, le jeu de l'imitation.

Le jeu de l'imitation

Contrairement à la façon dont il est présenté dans le film de Tyldum, suivant en cela la majorité des interprètes puisqu'une entrée « Turing test » se trouve depuis longtemps dans le dictionnaire anglais Collins³, le « jeu de l'imitation » ne consiste pas à montrer qu'il n'est pas possible de distinguer les réponses d'un homme des réponses d'un ordinateur convenablement programmé à qui on poserait des questions pendant une durée de jeu de cinq minutes et en cachant à l'interrogateur tout indice tenant à l'apparence physique des joueurs, c'est-à-dire en se limitant à des réponses imprimées. Le jeu est plus complexe car il est constitué de *deux étapes distinctes*, indispensables pour tenter d'obtenir le résultat escompté : l'indifférence entre les réponses humaines et les réponses de la machine et, partant, la « preuve » que l'intelligence mécanique est possible. Dans la formulation du jeu décrite par Turing dans son seul article de philosophie publié en 1950 dans la revue *Mind*, trois joueurs participent au jeu. L'un, appelé l'interrogateur, est séparé des deux autres, un homme et une femme, et doit tenter de deviner qui est l'homme et qui est la femme – bref doit tenter de déterminer quelle est la différence physique *maximale* entre deux êtres humains. Une fois en position d'échec pendant un certain temps, relativement court, la seconde étape du jeu consiste à remplacer le joueur masculin par un ordinateur convenablement programmé sans prévenir l'interrogateur et à se demander si celui-ci sera capable de déceler qu'il n'a plus affaire au même joueur mais que celui-ci a été remplacé par un ordinateur, bref que la différence entre humains et ordinateurs dans la deuxième étape du jeu ne sera pas plus décelable que la différence des sexes entre les humains dans la première.

Je soutiens que le jeu tel qu'il est décrit par Turing *ne peut pas parvenir au résultat escompté* pour une raison très simple : pour réussir à monter une partie du jeu, il faudrait à la fois prendre en compte la différence physique entre homme et femme dans la première étape et la différence physique entre un être humain et un ordinateur dans la seconde

3. Test de Turing : test visant la capacité de l'ordinateur à penser, requérant que la substitution cachée d'un participant par un ordinateur dans un dialogue par téletype soit indétectable pour le participant humain qui reste. On va voir combien cette description est trompeuse.

– car il faut avoir la capacité de choisir deux joueurs physiquement les plus opposés – et ne pas prendre en compte cette différence physique – puisqu'il s'agit de parvenir à la conclusion que la mise en échec de l'interrogateur a une portée universelle qui rend toute différence physique entre les joueurs indifférente. Bref, *la conclusion à laquelle le jeu doit parvenir détruit les conditions de possibilité de sa propre construction*. Autrement dit, pour que tout lecteur de l'article puisse parvenir à faire sienne la mise en échec de l'interrogateur et la conclusion que l'intelligence artificielle est bien réelle, il faudrait que tout lecteur puisse à la fois se dire que, s'il était à la place de l'interrogateur dans le jeu, il serait lui aussi mis en échec dans les deux étapes du jeu tout en se disant aussi qu'il doit cependant être *toujours* capable de faire physiquement la différence entre les deux joueurs (masculin et féminin dans la première étape, féminin et mécanique dans la seconde) pour monter une partie. Bref, la différence physique entre les joueurs doit en même temps être à *tout jamais* indécélable tout en étant pour *toujours* présumée pour que le jeu puisse fonctionner. La position exigée de la part du lecteur quant à la détermination du rapport d'identité ou de différence entre humain et ordinateur est donc un *indécidable* au sens technique que ce terme revêt depuis l'article de Turing de 1936 puisque les conditions d'accès à cette information rendent impossibles l'accès à l'information en question.

Déterminisme prédictif ou pas

Comment concevoir alors le jeu de l'imitation et plus globalement l'article de Turing de 1950 ? En remarquant qu'il y a une *toute autre superposition dans cet article que celle existant entre les deux sens de la notion d'intelligence* dont j'ai parlé plus haut. Turing remarque en effet que l'ordinateur en tant que machine physique est une machine « laplacienne » (relevant du déterminisme prédictif) mais que le monde physique ne l'est généralement pas pour des raisons ayant trait à la nature même de la matière, susceptible de comportements chaotiques (relevant du déterminisme non-prédictif). C'est donc la superposition de ces deux sens du déterminisme qui fait le fond de l'article de Turing et il faut le lire à ces deux niveaux : le niveau « grand public » (celui du film de Tyldum) dans lequel on débat (littéralement : *à n'en plus finir* pour des raisons ayant trait à l'indécidabilité dont j'ai parlé plus haut) pour savoir si et comment une machine peut être considérée comme « intelligente » et le niveau « de l'indécidable » dans lequel on sait pertinemment que la véritable question n'est pas là mais qu'elle se situe dans le rapport entre le calculable et l'incalculable et que *ce rapport est précisément celui que Turing a exploré tout au long de sa vie et ce, jusqu'en biologie* où les formes vivantes persistent – pour le temps de leur vie – à la frange du chaos qui mettra un terme à leur cohérence interne, comme l'a si profondément montré Giuseppe Longo (cf. F. Bailly & G. Longo, *Mathématiques et sciences de la nature ; La singularité physique du vivant*, Hermann, 2007 et G. Longo, M. Montévil, *Perspectives on Organisms : Biological Time, Symmetries and Singularities*, Springer, Berlin, 2014).

Il y a donc une continuité théorique totale dans le parcours de Turing, que l'on pourrait résumer dans cet aphorisme quasi-mallarméen dans son aspect paradoxal : *jamais un surcroît de programmation n'abolira le non-programmé, comme le prouve la théorie de la programmation*. Turing est en effet parvenu à montrer les limitations internes à la théorie de la programmation à partir de cette théorie même. Un problème capital se pose alors : comment

réussir à manifester et à explorer plus avant ce non-programmable si c'est seulement dans son rapport au programmable qu'il devient pensable ?

Où se situe le non-programmable ?

Turing passera le reste de sa vie, après son article de 1936, à essayer de répondre à cette question et la réponse qu'il a élaborée continue d'occuper le champ de la recherche aujourd'hui. Celle-ci me paraît être la suivante : la *production* de formes cohérentes que ce soit dans la *pensée* (l'invention du concept de machine de Turing, par exemple) ou dans la *nature* (l'apparition des formes vivantes, par exemple) est une manifestation de ce non-programmable.

Je soutiens que l'article de 1950 dans lequel Turing propose à la sagacité de son lecteur le « jeu de l'imitation » est une méditation sur les deux notions de la « pensée » et de la « nature » et plus encore sur leurs rapports. Or, pour avancer plus avant sur cette question difficile, on ne peut pas envisager la « pensée » ou la « nature » comme des notions fixées une fois pour toutes dont on pourrait étudier les produits complètement constitués (telle pensée, telle forme vivante) : il faut au contraire envisager la « pensée » et la « nature » comme des processus d'*individuation progressive de formes*. Or Turing a proposé, dans les dernières années de sa vie, un modèle de développement des formes vivantes à partir de brisures de symétrie dans la matière physique et c'est, à mon sens, en poursuivant une idée analogue *sur le mode du désir* qu'il construit le jeu de l'imitation quand il s'agit de rendre compte de l'invention du concept de « machine de Turing » : il s'agit de savoir si on peut remplacer la pensée d'un homme par un ordinateur à partir d'une brisure de symétrie dans la matière, c'est-à-dire d'une différence physique.

Vu sous cet angle, le jeu de l'imitation prend une tout autre tournure que celle de savoir si la « pensée » est un concept universel, indifféremment incarné dans l'être humain ou l'ordinateur : il consiste en la description d'un processus d'individuation d'une forme de la pensée (l'invention du concept de « machine de Turing » chez l'individu Turing) et du rapport ambivalent que cette forme entretient avec les deux modalités (programmable et non-programmable) de son incarnation possible le jeu devenant alors typique d'un « double entendre » (comme on dit en anglais !). De ce point de vue, la différence physique entre les joueurs et, en tout premier lieu, comme Turing n'a pas manqué de le voir, la différence des sexes, joue un rôle capital dans la dynamique du jeu puisque c'est son possible dépassement, pour les formes vivantes que nous sommes, qui en fait le moteur. L'aphorisme quasi-mallarméen dont je parlais plus haut prend alors la forme suivante : *jamais un surcroît de programmation n'abolira la différence des sexes, comme le suggère le jeu de l'imitation*.

Aussi peut-on dire que si le jeu parvenait à ses fins, la différence des sexes serait effectivement dépassée et la sexualité personnelle de Turing définitivement cachée. Turing apprit, à ses dépens, *que ce n'était pas le cas* et il y a presque de la prophétie dans le jeu de l'imitation, comme le film de Tyldum le laisse d'ailleurs entendre, lui qui finit par

avouer la nature de sa sexualité après un interrogatoire de police et qui fut condamné à une castration chimique.

La signification générale du parcours de Turing ne se situe donc pas, selon moi, dans le passage d'un sens de la notion d'intelligence à un autre (on en resterait à une perspective algorithmique visant à étendre indéfiniment le périmètre du calculable, ce que l'informatique n'a pas manqué de faire depuis qu'elle existe) mais dans le passage de la *forme* au sens du formalisme et de ses limitations internes à la *forme* au sens de la *production* des formes, idéales comme celles de la machine de Turing ou naturelles comme celles des formes biologiques, en se plaçant d'emblée du point de vue du rapport programmable/non-programmable. Et c'est évidemment dans cette production que se situe l'énigme de *l'invention des formes* que ce soit celle produite par Turing dans le concept de machine de Turing ou que ce soit celle produite par la nature. Il reste encore un film à faire sur le sujet, et il n'y a aucune raison de ne pas espérer le voir réalisé un jour.

Filmographie

- Le modèle Turing,
de Catherine Bernstein, produit par CNRS Images
- La drôle de guerre d'Alan Turing,
de Denis van Waerebeck, récemment passé sur Arte
- The strange Life and Death of Dr Turing,
de Christopher Sykes, produit par la BBC (en ligne sur YouTube).

Bibliographie

- F. Bailly & G. Longo,
Mathématiques et sciences de la nature ; La singularité physique du vivant, Hermann, 2007
- G. Longo & M. Montévil,
Perspectives on Organisms : Biological Time, Symmetries and Singularities, Springer, Berlin, 2014

100 ans de Relativité Générale (III/III) : déflexion de la lumière

Eric OLIVIER^{1,2}

Résumé. – « Il pourrait n'y avoir aucune déflexion, ce qui signifierait que la lumière n'est aucunement sujette à la gravitation. Ou une *demi-déflexion*, ce qui signifierait que la lumière est sujette à la gravitation, ainsi que Newton l'a suggéré et qu'elle suit la simple loi newtonienne. Ou, il pourrait y avoir une pleine déflexion, confirmant la loi d'Einstein au lieu de celle de Newton. » (Eddington, à propos des mesures de déflexion des rayons lumineux rasant la surface solaire – cité par Chandrasekhar.)

1. Introduction

L'espace est rapporté au repère cartésien $(S; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$; le *principe d'inertie de Newton* affirme que le mouvement d'une particule P , repérée (à l'instant t) par le *rayon vecteur* $\mathbf{r}(t)$ (issu de S), est *causé* par une force proportionnelle à l'accélération $\ddot{\mathbf{r}}(t) = d^2\mathbf{r}/dt^2$ de P : par définition, le coefficient de proportionnalité est appelé la *masse inerte* de P . Nous nous intéressons au cas où la particule P est prise dans le champ de gravitation d'une *masse grave* M concentrée en S . Le *principe d'équivalence* affirmant l'égalité de la *masse inerte* et de la *masse grave* nous permet d'écrire l'équation différentielle vérifiée par $t \mapsto \mathbf{r}(t)$, soit :

$$\ddot{\mathbf{r}} = \left(\frac{GM}{\rho^2} \right) \frac{\mathbf{r}}{\rho}$$

Ici, G désigne la constante de la gravitation et ρ est la norme euclidienne du rayon vecteur \mathbf{r} . Nous insistons sur un point essentiel, conséquence directe du principe d'équivalence : la masse (grave/inerte) de P n'affecte pas la forme de son mouvement dans le champ de gravitation de S . Newton a montré que les mouvements possibles de P dans le champ de gravitation de la masse M en S sont de deux natures : soit l'orbite de P est bornée et P décrit une ellipse – dont S est l'un des foyers – soit l'orbite de P est non bornée et P décrit une parabole ou une hyperbole – S étant toujours un foyer de la conique. Ainsi, la trajectoire parabolique réalise la transition entre le mouvement borné et le mouvement non borné. Pour voir cela, considérons une configuration un peu particulière. Supposons qu'à l'instant $t = 0$ la particule P est dans le plan $(S; \mathbf{e}_1, \mathbf{e}_2)$, avec $\mathbf{r}(0) = r_* \mathbf{e}_1$ et $\dot{\mathbf{r}}(0) = v \mathbf{e}_2$. Alors (nous verrons que) l'orbite de P reste dans le plan $(S; \mathbf{e}_1, \mathbf{e}_2)$, la vitesse de libération de P étant alors

$$(1) \quad c_* := \sqrt{\frac{2GM}{r_*}}$$

Nous nous intéressons uniquement au cas où le point S_* où se trouve P à l'instant $t = 0$ est le *périastre*, c'est-à-dire le point de l'orbite de P qui est le plus proche du foyer S de la trajectoire conique : on peut montrer que cette situation n'a lieu que pour $v \geq c_*/\sqrt{2}$: pour

1. GDAC-I2M UMR 7373 CNRS Université d'Aix-Marseille
2. eric.olivier@univ-amu.fr

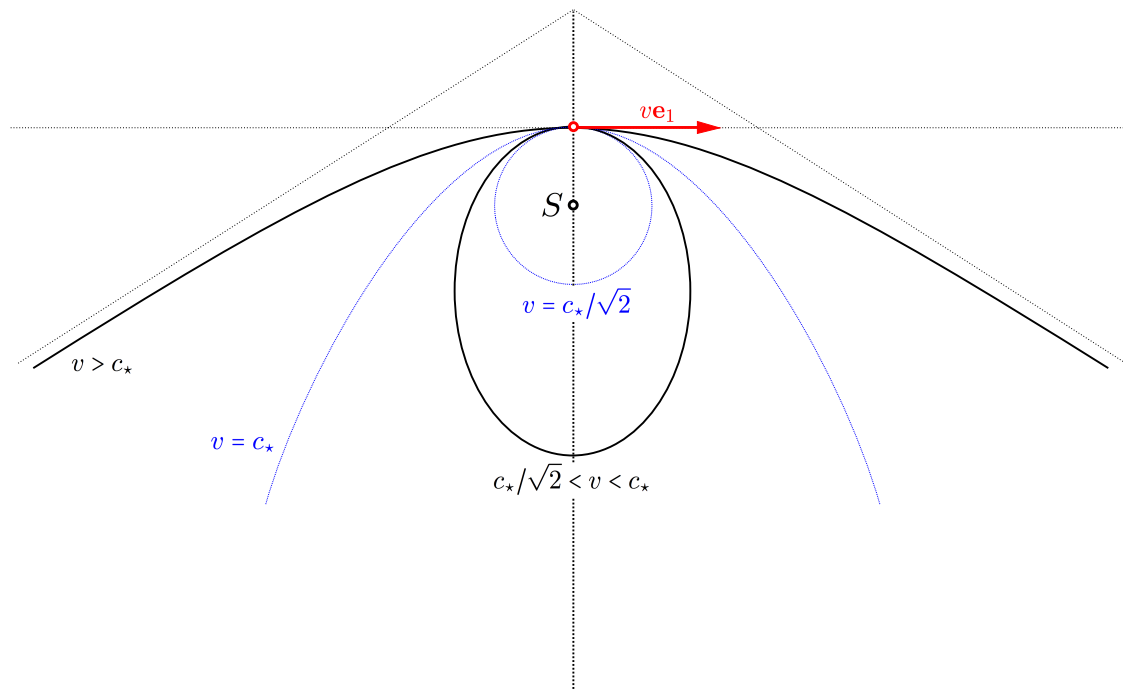


FIGURE 1. Famille des orbites coniques dans le champ de gravitation d'une masse M en fonction de la vitesse v au périastre de l'orbite et de la vitesse de libération $c_* := \sqrt{2GM/r_*}$ (où r_* est périastre); lorsque $v < c_*/\sqrt{2}$ le point initial n'est plus le périastre mais l'apoastre de l'orbite (nécessairement elliptique).

$v = c_*/\sqrt{2}$ l'orbite de P est un cercle et pour $0 \leq v < c_*/\sqrt{2}$, le point S_* devient l'apoastre³, i.e. le point de l'orbite de P qui est le plus éloigné du foyer S . Lorsque $c_*/\sqrt{2} \leq v < c_*$ (resp. $v > c_*$), P décrit une ellipse (resp. une hyperbole) : pour $v = c_*$ l'orbite est exactement une parabole (c.f. Fig 1). Supposons maintenant que la particule P vienne de l'infini avec la vitesse de la lumière – soit $c = 2.99792458 \times 10^8$ m/s – et que son orbite soit hyperbolique avec $\dot{r}(-\infty) = cu$ où u est un vecteur unitaire donné : alors la mesure β (en radians) de l'angle orienté $(u, \dot{r}(+\infty))$ – l'angle de déflexion – dépend du paramètre d'impact b (i.e. la distance de S à l'asymptote de la trajectoire de P en $t = -\infty$), mais ne dépend pas de la masse de la particule : nous verrons que

$$\tan\left(\frac{\beta}{2}\right) = \frac{GM}{bc^2}$$

En 1907 Einstein remarque⁴ que cette formule implique formellement qu'un rayon de lumière – trajectoire d'un corpuscule lumineux sans masse et dont la vitesse c est celle de la lumière – doit être courbée par l'attraction gravitationnelle de la masse M en S : il serait alors possible de mesurer la déflexion d'un rayon lumineux rasant la surface

3. Lorsque le centre du champ gravitationnel S est le Soleil on parle de périhélie et d'aphélie.

4. À l'instar de John Michell (1724-1793) et Pierre-Simon de Laplace (1749-1827). L'histoire de la Relativité générale est exposée dans [Eis02]; on y trouve aussi en particulier beaucoup d'informations sur la déflexion des rayons lumineux (voir aussi [Bro11, Chap. 8 - § 8.9] pour la position de Newton et l'opinion à ce sujet de H. Weil).

solaire, afin de vérifier la valeur newtonienne théorique, soit

$$\beta \simeq \frac{2GM_{\odot}}{c^2 r_{\odot}} \simeq 0.87''$$

Ici M_{\odot} et r_{\odot} désignent respectivement la masse et le rayon solaire (ces quantités étant exprimées en unités du système international des mesures physiques). Cette remarque d'Einstein sur l'action de la gravitation sur la lumière, se situe au moment où il commence à réfléchir sur la Relativité Générale. Il est aussi étonnant de remarquer que plusieurs années plus tard, la Relativité Générale [Ein16] donnera un contexte rigoureux justifiant que le rayon lumineux (rasant le soleil) est bien courbé par la masse solaire M_{\odot} , mais que cette déflexion vaut *exactement* le double de la déflexion newtonienne, soit

$$\beta \simeq \frac{4GM_{\odot}}{c^2 r_{\odot}} \simeq 1.75''$$

Cette formule a été obtenue par Schwarzschild en 1916 [Sch16] comme conséquence d'une première résolution explicite des équations d'Einstein de la Relativité Générale. La première vérification expérimentale de cette déflexion a été réalisée par Eddington, lors de l'éclipse de soleil de 1919 observée en Amérique du sud. Mentionnons qu'un point clef du calcul de Schwarzschild est la détermination (dans la relativité générale) d'une constante appelé le **rayon de Schwarzschild** pour une masse M , soit

$$r_{\bullet} = \frac{2GM}{c^2}$$

Étrangement, la formule de la vitesse c_{\star} de libération d'une masse M en (1) entraîne que le rayon r_{\star} pour le quel c_{\star} coïncide avec la lumière, correspond au rayon de Schwarzschild de la masse M : en effet r_{\bullet} coïncide avec la solution en ρ de l'équation

$$\sqrt{\frac{2GM}{\rho}} = c$$

L'interprétation (heuristique) est frappante : si la masse M est concentrées à l'intérieur de la sphère de rayon r_{\bullet} , alors un *photon émis* par cet objet ne peut se *libérer* de l'attraction gravitationnelle : John Wheeler a baptisé une telle masse sphérique **un trou noir** (voir les repères historiques au § 10). Le but de cette troisième note sur les 100 ans de la Relativité Générale, est de donner un contexte rigoureux (et raisonnablement abordable) permettant de voir comment la Relativité Générale *explique* le doublement de la valeur de l'angle *newtonien* de déflexion du rayon lumineux rasant la surface solaire.

2. Première constante du mouvement newtonien : l'énergie

Soit (x, y, z) les coordonnées cartésiennes d'un point d'un ouvert \mathcal{U} de \mathbb{R}^3 et soit $\Phi(x, y, z)$ la valeur en ce point d'une fonction numérique Φ . On dit que l'accélération $t \mapsto \ddot{\mathbf{r}}(t)$ de P **dérive du potentiel** Φ défini⁵ sur \mathcal{U} lorsque $\ddot{\mathbf{r}}(t)$ coïncide avec l'opposé du gradient $\text{grad}(\Phi)$ évalué en $\mathbf{r}(t)$. Le signe "-" affectant **le gradient du potentiel** Φ est une convention cohérente avec la définition de l'*énergie lagrangienne* (c.f. Proposition 2.1). Soit $\langle \cdot | \cdot \rangle$ le produit scalaire associé à la structure euclidienne canonique de \mathbb{R}^3 ; dans la suite,

5. Dans le cas du potentiel newtonien de la force de gravitation nous aurons $\mathcal{U} = \mathbb{R}^3 \setminus \{S\}$.

nous noterons $v = \langle \dot{\mathbf{r}} | \dot{\mathbf{r}} \rangle^{1/2}$ la vitesse de P , où $\dot{\mathbf{r}} := d\mathbf{r}/dt$ est le vecteur vitesse. La proposition suivante constitue en fait la [description hamiltonienne](#) d'un système mécanique d'un corps soumis à une force dérivant d'un potentiel.

Proposition 2.1. *Si l'accélération de P dérive d'un potentiel Φ alors*

$$\frac{d}{dt} \left(\frac{\mathbf{v}^2}{2} + \Phi(\mathbf{r}) \right) = 0$$

en d'autres termes *l'énergie (massique) $E = \mathbf{v}^2/2 + \Phi(\mathbf{r})$ est une constante du mouvement.*

Preuve. Pour dériver le produit scalaire sur l'orbite de P , nous écrivons

$$\begin{aligned} \mathbf{v}^2(t+h) &\simeq \left\langle \dot{\mathbf{r}}(t) + h\ddot{\mathbf{r}}(t) \middle| \dot{\mathbf{r}}(t) + h\ddot{\mathbf{r}}(t) \right\rangle \\ &\simeq \left\langle \dot{\mathbf{r}}(t) \middle| \dot{\mathbf{r}}(t) \right\rangle + 2h \left\langle \dot{\mathbf{r}}(t) \middle| \ddot{\mathbf{r}}(t) \right\rangle \simeq \mathbf{v}(t)^2 - 2h \left\langle \dot{\mathbf{r}}(t) \middle| \text{grad}_{|\mathbf{r}(t)} \Phi(\mathbf{r}(t)) \right\rangle \end{aligned}$$

par suite

$$\frac{1}{2} \frac{d\mathbf{v}^2}{dt} = - \left\langle \dot{\mathbf{r}}(t) \middle| \text{grad}_{|\mathbf{r}(t)} \Phi(\mathbf{r}(t)) \right\rangle$$

Or, en notant $(r_1(t), r_2(t), r_3(t))$ les composantes cartésiennes de $\mathbf{r}(t)$, la formule de dérivée des fonctions composées donne :

$$\frac{d}{dt} \Phi(\mathbf{r}) = \frac{\partial \Phi}{\partial x}(\mathbf{r}(t)) \dot{r}_1(t) + \frac{\partial \Phi}{\partial y}(\mathbf{r}(t)) \dot{r}_2(t) + \frac{\partial \Phi}{\partial z}(\mathbf{r}(t)) \dot{r}_3(t) = \left\langle \dot{\mathbf{r}}(t) \middle| \text{grad}_{|\mathbf{r}(t)} \Phi(\mathbf{r}(t)) \right\rangle$$

□

3. Deuxième constante du mouvement newtonien : le moment cinétique

Un point de $\mathcal{U} = \mathbb{R}^3 \setminus \{S\}$ est aussi bien repéré par ses coordonnées cartésiennes (x, y, z) que par ses coordonnées sphériques (ρ, θ, ϕ) , la correspondance étant donné par un difféomorphisme local ψ , avec

$$(x, y, z) = (\rho \sin(\phi) \cos(\theta), \rho \sin(\phi) \sin(\theta), \rho \sin(\phi)) =: \psi(\rho, \theta, \phi)$$

En considérant maintenant que $(x, y, z) \mapsto \Phi(x, y, z)$ est un potentiel sur \mathcal{U} , nous définissons le potentiel sphérique⁶ Ψ associé à Φ en posant :

$$(2) \quad \Psi(\rho, \theta, \phi) := \Phi \circ \psi(\rho, \theta, \phi) = \Phi(\rho \sin(\phi) \cos(\theta), \rho \sin(\phi) \sin(\theta), \rho \sin(\phi))$$

De plus, si $(\mathbf{e}_\rho, \mathbf{e}_\theta, \mathbf{e}_\phi)$ désigne le repère sphérique mobile en $X = \psi(\rho, \theta, \phi)$, alors

$$(3) \quad \text{grad}_X(\Phi) = \frac{\partial \Psi}{\partial \rho} \mathbf{e}_\rho + \frac{1}{\rho} \frac{\partial \Psi}{\partial \theta} \mathbf{e}_\theta + \frac{1}{\rho \sin(\theta)} \frac{\partial \Psi}{\partial \phi} \mathbf{e}_\phi$$

Une idée très générale (et formalisée dans le [Théorème de Noether](#)) est que toute symétrie d'un système mécanique induit une constante du mouvement. Sans entrer dans le détail de ce résultat (fondamental), nous allons en vérifier directement une conséquence. En

6. On peut tout aussi bien considérer qu'il n'y a qu'un seul potentiel – disons Φ – et que c'est le système des coordonnées qui change : ainsi on pourrait tout aussi bien écrire $\Phi(x, y, z)$ dans les coordonnées cartésiennes ou encore $\Phi(\rho, \theta, \phi)$ dans les coordonnées sphériques : cette dernière convention est beaucoup utilisée dans les textes de sciences physiques et donne lieu à pas mal d'ambiguïté d'écriture. Cette pratique trouve cependant une forme très rigoureuse en géométrie différentielle par la définition même de variété différentielle grâce à l'atlas des cartes locales.

plus de notre hypothèse – i.e. le champ d'accélération de P dérive d'un potentiel Φ – **supposons que Φ soit à symétrie sphérique** : cela signifie que *le potentiel sphérique* défini en (2) ne dépend ni de θ ni de ϕ en ce sens que $\partial\Psi/\partial\theta = \partial\Psi/\partial\phi = 0$. Nous déduisons de l'expression du gradient de Φ en coordonnées sphériques en (3) que

$$\ddot{\mathbf{r}} = -\frac{\partial\Psi}{\partial\rho}\mathbf{e}_\rho$$

Nous retrouvons le fait (intuitivement évident) qu'à un potentiel à symétrie sphérique correspond un **champ d'accélération centrale**.

Proposition 3.1. *Si la particule P est soumise à un champ d'accélération dérivant d'un potentiel à symétrie sphérique, alors le moment cinétique (massique) $\mathbf{r} \times \dot{\mathbf{r}}$ est constant sur l'orbite.*

Preuve. Le champ d'accélération dérivant d'un potentiel à symétrie centrale, nous savons que $\ddot{\mathbf{r}}$ est colinéaire à \mathbf{r} et par suite⁷

$$\frac{d}{dt}(\mathbf{r} \times \dot{\mathbf{r}}) = \dot{\mathbf{r}} \times \dot{\mathbf{r}} + \mathbf{r} \times \ddot{\mathbf{r}} = 0$$

□

Dans la suite nous notons $\boldsymbol{\sigma}_0 = \sigma_0\mathbf{e}_3$ (avec $\sigma_0 = |\boldsymbol{\sigma}_0| > 0$) le moment cinétique de P : cela revient à convenir que l'orbite de P se déroule dans le plan $(S; \mathbf{e}_1, \mathbf{e}_2)$. En particulier, il est possible de décrire le mouvement de P par ses coordonnées polaires $(\rho(t), \theta(t))$ où $\rho(t)$ est la norme du vecteur position $\mathbf{r}(t)$ et où $\theta(t)$ désigne la mesure (en radians) de l'angle orienté $(\mathbf{e}_1, \mathbf{r}(t))$. Les composantes polaires du mouvement nous permettent d'exprimer la vitesse de P dans le repère mobile polaire, soit $(S; \mathbf{e}_r, \mathbf{e}_\theta)$: en effet

$$\dot{\mathbf{r}} = \frac{d}{dt}\rho(\cos(\theta)\mathbf{e}_1 + \sin(\theta)\mathbf{e}_2) = \dot{\rho}(\cos(\theta)\mathbf{e}_1 + \sin(\theta)\mathbf{e}_2) + \rho\dot{\theta}(-\sin(\theta)\mathbf{e}_1 + \cos(\theta)\mathbf{e}_2)$$

ce qui donne finalement

$$(4) \quad \dot{\mathbf{r}} = \dot{\rho}\mathbf{e}_\rho + \rho\dot{\theta}\mathbf{e}_\theta$$

Par suite nous obtenons

$$\sigma_0\mathbf{e}_3 = \mathbf{r} \times \dot{\mathbf{r}} = \mathbf{r} \times (\dot{\rho}\mathbf{e}_\rho + \rho\dot{\theta}\mathbf{e}_\theta) = \rho\dot{\theta}(\mathbf{r} \times \mathbf{e}_\theta) = \rho^2\dot{\theta}(\mathbf{e}_\rho \times \mathbf{e}_\theta) = \rho^2\dot{\theta}\mathbf{e}_3$$

Théorème 3.2 (Loi des aires de Kepler). *Si la particule P est soumise à un champ d'accélération dérivant d'un potentiel à symétrie sphérique, alors la vitesse aréolaire est une constante du mouvement égale à la norme σ_0 du moment cinétique : en d'autres termes*

$$(5) \quad \sigma_0 = \rho^2\dot{\theta}$$

7. Nous utilisons les propriétés algébriques du produit vectoriel (cross product) avec les approximations différentielles pour écrire

$$\begin{aligned} \mathbf{r}(t+h) \times \dot{\mathbf{r}}(t+h) &= \left(\mathbf{r}(t) + h\dot{\mathbf{r}}(t) + \frac{h^2}{2}\ddot{\mathbf{r}}(t) + \dots \right) \times \left(\dot{\mathbf{r}}(t) + h\ddot{\mathbf{r}}(t) + \frac{h^2}{2}\dddot{\mathbf{r}}(t) + \dots \right) \\ &= \mathbf{r}(t) \times \dot{\mathbf{r}}(t) + h(\dot{\mathbf{r}}(t) \times \dot{\mathbf{r}}(t) + \mathbf{r}(t) \times \ddot{\mathbf{r}}(t)) \\ &\quad + \frac{h^2}{2}(2\dot{\mathbf{r}}(t) \times \ddot{\mathbf{r}}(t) + \mathbf{r}(t) \times \dddot{\mathbf{r}}(t) + \ddot{\mathbf{r}}(t) \times \dot{\mathbf{r}}(t)) + \dots \\ &= \mathbf{r}(t) \times \dot{\mathbf{r}}(t) + h\mathbf{r}(t) \times \ddot{\mathbf{r}}(t) + \frac{h^2}{2}(\dot{\mathbf{r}}(t) \times \ddot{\mathbf{r}}(t) + \mathbf{r}(t) \times \dddot{\mathbf{r}}(t)) + \dots \end{aligned}$$

4. Troisième constante du mouvement : le vecteur d'excentricité

Nous considérons maintenant que notre particule P est plongée dans le [champ de gravitation newtonien](#) de la masse M concentrée en S : la particule est donc soumise à un champ de force dérivant du potentiel $(\rho, \theta, \phi) \mapsto \Psi(\rho) = -GM/\rho$. Dans ce cas particulier, il existe un invariant supplémentaire appelé [vecteur d'excentricité](#) (ou encore [vecteur de Runge-Lenz](#)), qui spécifie la forme conique des trajectoires de P (ceci peut aussi être vu comme conséquence du Théorème de Noether).

Proposition 4.1. *Si la particule P d'orbite $t \mapsto \mathbf{r}(t)$ de vitesse $t \mapsto \dot{\mathbf{r}}(t)$ et de moment cinétique $\boldsymbol{\sigma}_0 = \sigma_0 \mathbf{e}_3$ est plongée dans le [champ de gravitation newtonien](#) d'une masse M centrée en S , alors*

$$\frac{d}{dt} \left(\frac{\dot{\mathbf{r}} \times \boldsymbol{\sigma}_0}{GM} \right) = \dot{\theta} \mathbf{e}_\theta = \frac{d}{dt} \mathbf{e}_\rho$$

et le vecteur d'excentricité⁸ est par définition :

$$\boldsymbol{\varepsilon}_0 = \frac{\dot{\mathbf{r}} \times \boldsymbol{\sigma}_0}{GM} - \mathbf{e}_\rho$$

Preuve. Rappelons que le moment cinétique (massique) $\boldsymbol{\sigma}_0 = \mathbf{r} \times \dot{\mathbf{r}} = \sigma_0 \mathbf{e}_3$ est un vecteur constant de direction orthogonale au plan de l'orbite de P . Par dérivation directe

$$\frac{d}{dt} (\dot{\mathbf{r}} \times \boldsymbol{\sigma}_0) = - \left(\frac{GM}{\rho^2} \right) \mathbf{e}_\rho \times (\rho^2 \dot{\theta}) \mathbf{e}_3 = GM \dot{\theta} (\mathbf{e}_\rho \times \mathbf{e}_3) = GM \dot{\theta} \mathbf{e}_\theta$$

□

Comme $\langle \boldsymbol{\sigma}_0 | \boldsymbol{\varepsilon}_0 \rangle = 0$, le vecteur $\boldsymbol{\varepsilon}_0$ appartient au plan vectoriel engendré par \mathbf{e}_ρ et \mathbf{e}_θ . On définit alors ω comme la mesure de l'angle orienté $(\mathbf{e}_r, \boldsymbol{\varepsilon}_0)$ (noter que $\dot{\theta} = \dot{\omega}$).

Proposition 4.2. *(i) : L'orbite de P dans le champ de gravitation de S de masse M et de moment cinétique $\boldsymbol{\sigma}_0 = \sigma_0 \mathbf{e}_3$ (avec $\sigma_0 = |\boldsymbol{\sigma}_0| > 0$) est une conique dont un foyer est S et d'équation polaire*

$$(6) \quad \rho = \frac{p}{1 + \varepsilon_0 \cos(\omega)}$$

l'excentricité $\varepsilon_0 \geq 0$ est la norme du vecteur de Runge-Lenz $\boldsymbol{\varepsilon}_0$ et la valeur du [paramètre](#) est

$$(7) \quad p = \frac{\sigma_0^2}{GM}$$

(ii) : la valeur du [périastre](#) (i.e. le minimum de $\rho(t)$) vaut

$$(8) \quad r_* := \frac{p}{1 + \varepsilon_0}$$

Preuve. (i) : Par définition de la norme ε_0 du vecteur de Runge-Lenz $\boldsymbol{\varepsilon}_0$ et par définition de l'angle polaire ω , il vient $\langle \mathbf{r} | \boldsymbol{\varepsilon}_0 \rangle = \rho \varepsilon_0 \cos(\omega)$. Mais d'autre part⁹

$$\langle \mathbf{r} | \boldsymbol{\varepsilon}_0 \rangle = \frac{1}{GM} \langle \mathbf{r} | \dot{\mathbf{r}} \times \boldsymbol{\sigma}_0 \rangle - \rho = \frac{1}{GM} \langle \boldsymbol{\sigma}_0 | \mathbf{r} \times \dot{\mathbf{r}} \rangle - \rho = \frac{1}{GM} \langle \boldsymbol{\sigma}_0 | \boldsymbol{\sigma}_0 \rangle - \rho = \frac{\sigma_0^2}{GM} - \rho$$

L'identification des deux expressions du produit scalaire $\langle \mathbf{r} | \boldsymbol{\varepsilon}_0 \rangle$ donne l'équation polaire (6) de l'orbite de P autour de S . C'est l'équation d'une conique d'excentricité ε_0 et de paramètre $p = \sigma_0^2/(GM)$, dont le centre de force S occupe l'un des foyers.

(ii) : Par (6) ρ atteint son minimum r_* lorsque $\omega = 0$ et donc $r_* = p/(1 + \varepsilon_0)$.

8. Le vecteur de Runge-Lenz proprement dit est $\dot{\mathbf{r}} \times \boldsymbol{\sigma}_0 - GM \mathbf{e}_r$.

9. En utilisant une propriété classique du produit mixte $\det(u, v, w) = [u, v, w] = \langle u | v \times w \rangle$.

□

La valeur de l'excentricité ε_0 détermine la nature conique de l'orbite de P . Si $\varepsilon_0 > 1$ (resp. $0 \leq \varepsilon_0 < 1$), la trajectoire est une hyperbole (resp. une ellipse – l'orbite de P étant circulaire lorsque $\varepsilon_0 = 0$) ; si $\varepsilon_0 = 1$, la trajectoire est une parabole.

Proposition 4.3. *Soit r_* le périastre de l'orbite (conique) de P de paramètre p et d'excentricité $\varepsilon_0 = \langle \varepsilon_0 | \varepsilon_0 \rangle^{1/2}$ et $(\mathbf{e}_{r_*}, \mathbf{e}_{\theta_*} = \mathbf{e}_3 \times \mathbf{e}_{r_*})$ le repère polaire au(x) moment(s) du périastre : alors*

$$\begin{aligned} (i) & : E = \frac{1}{2} \left(\frac{\sigma_0}{r_*} \right)^2 - \frac{GM}{r_*} \quad (\text{énergie de la particule}) \\ (ii) & : \varepsilon_0 = \left(\frac{p}{r_*} - 1 \right) = \left(\frac{\sigma_0^2}{GM r_*} - 1 \right) = \left(1 + \frac{2r_* E}{GM} \right) = \sqrt{1 + 2E \left(\frac{\sigma_0}{GM} \right)^2} \\ (iii) & : \varepsilon_0 = \left(\frac{p}{r_*} - 1 \right) \mathbf{e}_{r_*} = \left(\frac{\sigma_0^2}{GM r_*} - 1 \right) \mathbf{e}_{r_*} = \left(1 + \frac{2r_* E}{GM} \right) \mathbf{e}_{r_*} \end{aligned}$$

Preuve. (i) : Rappelons d'abord que l'énergie $E = \mathbf{v}^2/2 - MG/\rho$ est une constante du mouvement ; d'autre part, l'expression polaire de la vitesse en (4) et la loi des aires en (5) assurant que $\sigma_0 = \rho^2 \dot{\theta}$ entraînent que $\mathbf{v}^2 = \dot{\rho}^2 + \rho^2 \dot{\theta}^2 = \dot{\rho}^2 + \sigma_0^2/\rho^2$ et par suite,

$$E = \frac{1}{2} \left(\dot{\rho}^2 + \left(\frac{\sigma_0}{\rho} \right)^2 \right) - \frac{GM}{\rho}$$

Si t_* est un temps de passage au périhélie alors $\dot{\rho}(t_*) = 0$, d'où l'expression de E en (i).

(ii) : L'identité $r_* = p/(1 + \varepsilon_0)$ donnée en (8) entraîne que $\varepsilon_0 = p/r_* - 1$, la deuxième égalité étant obtenue avec la valeur de p donnée par (7). Pour la troisième expression nous l'expression de l'énergie de P en fonction de r_* et donnée en (i). Nous obtenons $\sigma_0^2 = 2r_*^2 E + 2GM r_*$; mais nous avons déjà démontré que $\varepsilon_0 = p/r_* - 1$ et en utilisant l'expression de p en (7), il vient :

$$\varepsilon_0 = \frac{\sigma_0^2}{GM r_*} - 1 = \frac{2r_* E}{GM} + 1$$

établissant ainsi l'avant dernière expression. Il est alors possible d'écrire

$$(\varepsilon_0 + 1)(\varepsilon_0 - 1) = \left(\frac{\sigma_0^2}{GM r_*} \right) \left(\frac{2r_* E}{GM} \right) = \frac{2\sigma_0^2 E}{G^2 M^2}$$

soit finalement

$$\varepsilon_0 = \sqrt{1 + \frac{2\sigma_0^2 E}{G^2 M^2}}$$

(iii) : Le moment cinétique $\boldsymbol{\sigma}_0$ vaut $\sigma_0 \mathbf{e}_3$, où $\sigma_0 = \rho^2 \dot{\theta}$ est la vitesse aréolaire (loi des aires de Kepler). Rappelons que t_* est un instant t pour lequel $r_* = \rho(t)$, ce qui donne $\dot{\rho}(t_*) = 0$. Par suite $\dot{\mathbf{r}}(t_*) = r_* \dot{\theta}(t_*) \mathbf{e}_{\theta_*} = \sigma_0/r_* \mathbf{e}_{\theta_*}$. Le vecteur d'excentricité étant aussi une constante du mouvement

$$\varepsilon_0 = \frac{\dot{\mathbf{r}}(t_*) \times \boldsymbol{\sigma}_0}{GM} - \mathbf{e}_{r_*} = \frac{(\sigma_0/r_* \mathbf{e}_{\theta_*}) \times (\sigma_0 \mathbf{e}_3)}{GM} - \mathbf{e}_{r_*} = \left(\frac{\sigma_0^2}{GM r_*} - 1 \right) \mathbf{e}_{r_*}$$

(où nous avons utilisé le fait que $\mathbf{e}_{\theta_*} \times \mathbf{e}_3 = \mathbf{e}_{r_*}$). Ceci démontre les deux premières expressions de ε_0 en fonction de la vitesse aréolaire σ_0 , la dernière expressions découlant de (i).

□

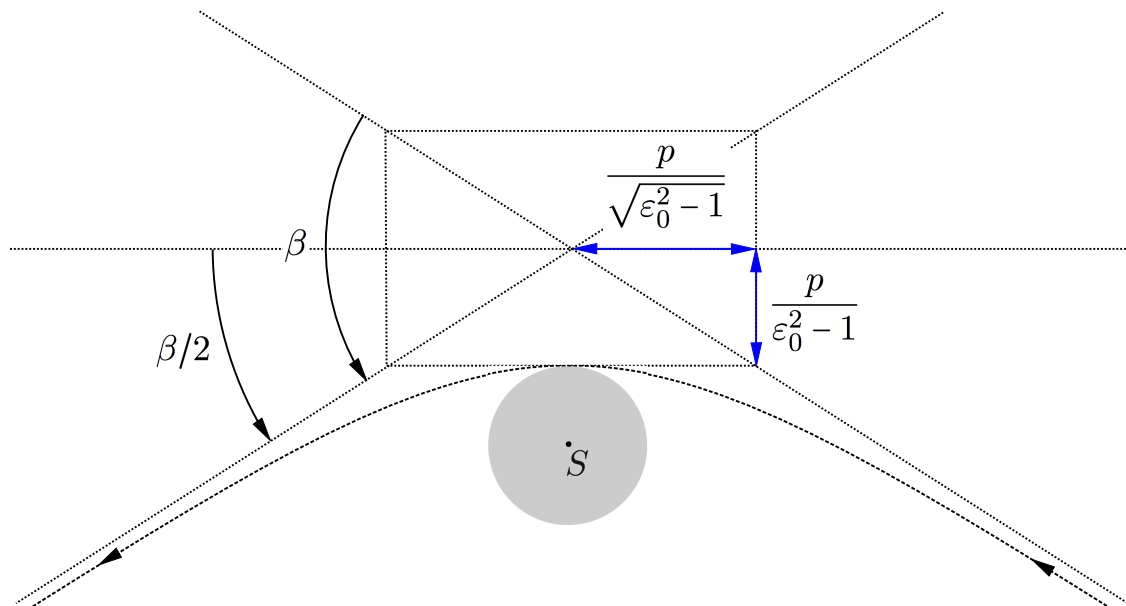


FIGURE 2. Angle de déflexion d'un rayon lumineux hyperbolique : si nous notons β cet angle de déflexion et si ε_0 (reps. p) est l'excentricité (reps. le paramètre) de l'orbite hyperbolique, alors

$$\tan\left(\frac{\beta}{2}\right) = \frac{p}{\varepsilon_0^2 - 1} \frac{\sqrt{\varepsilon_0^2 - 1}}{p} = \frac{1}{\sqrt{\varepsilon_0^2 - 1}}$$

5. Déflexion newtonienne I

À partir de maintenant c'est la masse solaire M_\odot qui est supposée centrée en S . La masse de P n'intervenant pas dans l'analyse de son orbite conique $t \mapsto \mathbf{r}(t)$, nous pouvons *formellement* considérer le cas où P est un *photon* rasant la surface du soleil. Plus précisément, la gravitation newtonienne nous autorise à nous placer dans la condition où l'orbite de P est hyperbolique avec un périastre r_* égal rayon solaire r_\odot , la *vitesse de P pout $t = \pm\infty$ étant égale à la vitesse de la lumière*¹⁰, i.e. $v(\pm\infty) = c$. Nous nous intéressons à la déflexion angulaire β de la trajectoire de P , i.e. la moitié de la mesure de l'angle orienté $(\dot{\mathbf{r}}(-\infty); \dot{\mathbf{r}}(+\infty))$: la trajectoire hyperbolique de P étant de paramètre p et d'excentricité ε_0 , nous savons (voir Fig. 2) que

$$(9) \quad \tan\left(\frac{\beta}{2}\right) = \frac{p}{\varepsilon_0^2 - 1} \cdot \frac{\sqrt{\varepsilon_0^2 - 1}}{p} = \frac{1}{\sqrt{\varepsilon_0^2 - 1}}.$$

Proposition 5.1. Soit P une particule soumise au champ gravitationnel du soleil S et décrivant une trajectoire hyperbolique dont la vitesse à l'infini est la vitesse de la lumière c ; b désigne le paramètre d'impact, i.e. la distance de S à l'une quelconque des deux asymptotes à la trajectoire de P . Si $r_\bullet = 2GM_\odot/c^2$ désigne le rayon de Schwarzschild du soleil et r_\odot le rayon solaire, alors la

10. Dans le cadre de la relativité galiléenne, cela entraîne que la vitesse de P à son passage au périastre est supérieure à la vitesse de la lumière.

déflexion angulaire ¹¹ β de la trajectoire de P rasant la surface solaire vérifie

$$(10) \quad 2 \tan\left(\frac{\beta}{2}\right) = \frac{2GM_{\odot}}{c^2 b} = \frac{r_{\bullet}}{b} = \frac{r_{\bullet}}{r_{\odot}} \frac{1}{\sqrt{1 + r_{\bullet}/r_{\odot}}}$$

Preuve. L'hypothèse $(v(+\infty), r(+\infty)) = (c, +\infty)$ donne les valeurs de l'énergie ainsi que de la vitesse aréolaire de P avec respectivement $2E = c^2$ et $\sigma_0 = cb$, où b est le paramètre d'impact. Grâce aux expressions de l'excentricité ε_0 (partie (ii) de la Proposition 4.3) et à l'expression du rayon de Schwarzschild r_{\bullet} de la masse solaire M_{\odot} , il vient

$$(11) \quad \varepsilon_0^2 = 1 + \frac{\sigma_0^2(2E)}{(GM_{\odot})^2} = 1 + \left(\frac{bc^2}{GM_{\odot}}\right)^2 = 1 + \left(\frac{2b}{r_{\bullet}}\right)^2$$

soit encore $\varepsilon_0^2 - 1 = (2b/r_{\bullet})^2$ et d'après (9) :

$$\tan\left(\frac{\beta}{2}\right) = \frac{1}{\sqrt{\varepsilon_0^2 - 1}} = \frac{1}{2} \frac{r_{\bullet}}{b}$$

La déflexion du rayon lumineux rasant la surface solaire étant très faible, il est raisonnable que le paramètre d'impact soit approximativement égal à la périhélie, soit $b \simeq r_{\star} = r_{\odot}$. Pour voir cela, nous partons de l'expression de l'énergie de P (c.f. (i)-Proposition 4.3) : comme $E = c^2/2$ et que $\sigma_0 = cb$, nous obtenons avec $r_{\star} = r_{\odot}$ et $r_{\bullet} = 2GM_{\odot}/c^2$

$$\frac{c^2}{2} = \frac{1}{2} \frac{c^2 b^2}{r_{\odot}^2} - \frac{c^2 r_{\bullet}}{2 r_{\odot}} = \frac{c^2}{2} \left(\frac{b^2}{r_{\odot}^2} - \frac{r_{\bullet}}{r_{\odot}} \right)$$

Après simplification, il vient $b^2 = r_{\odot}^2(1 + r_{\bullet}/r_{\odot})$ et finalement

$$2 \tan\left(\frac{\beta}{2}\right) = \frac{r_{\bullet}}{r_{\odot}} \frac{1}{\sqrt{1 + r_{\bullet}/r_{\odot}}}$$

□

5.1. Approximation de la déflexion. Le fait que le rayon de Schwarzschild du soleil soit très petit devant le rayon solaire (avec $r_{\bullet} \simeq 2.5$ km), justifie l'approximation $\beta \simeq r_{\bullet}/r_{\odot}$. Avec les constantes physiques (exprimées dans les unités internationales), soient

- la masse du soleil : $M_{\odot} = 1.9884 \times 10^{30}$ kg
- le rayon du soleil : $r_{\odot} = 6.95800 \times 10^8$ m
- la vitesse de la lumière : $c = 2.99792458 \times 10^8$ m/s
- la constant de gravitation de Newton : $G = 6.67384 \times 10^{-11}$ m³ · kg⁻¹ · s⁻²

nous obtenons

$$(12) \quad \frac{r_{\bullet}}{r_{\odot}} = \frac{2GM_{\odot}}{c^2 r_{\odot}} \simeq \frac{2 \times 6.67384 \times 10^{-11} \times 1.9884 \times 10^{30}}{(2.99792458 \times 10^8)^2 \times 6.95800 \times 10^8} \simeq 4,24 \times 10^{-6}$$

ce qui signifie que β vaut approximativement 4.24×10^{-6} rad. Pour la valeur en seconde d'arc, il faut se rappeler que $180 \times 3600''$ (secondes d'arc) valent π radians : ainsi $1''$ (1 seconde) d'arc vaut $4,85 \times 10^{-6}$ rad, et finalement

$$\beta \simeq (4,24/4,85)'' \simeq 0,87''$$

6. Déflexion newtonienne II

La particule P est soumise au champ gravitationnel de la masse solaire M_{\odot} centrée en S ; son orbite – supposée hyperbolique de centre S – se déroule dans le plan $(S; \mathbf{e}_1, \mathbf{e}_2)$. Par la symétrie du mouvement hyperbolique de P nous savons que $\mathbf{v}(-\infty) = \mathbf{v}(+\infty)$ et nous supposons que cette vitesse commune est la vitesse de la lumière notée c . Nous notons $\delta(\dot{\mathbf{r}}) = \dot{\mathbf{r}}(-\infty) - \dot{\mathbf{r}}(+\infty)$ la variation du vecteur vitesse entre les instants $t = -\infty$ et $t = +\infty$ (par passage à la limite) : sans perte de généralité nous pouvons supposer que $\delta(\dot{\mathbf{r}})$ est positivement proportionnel à \mathbf{e}_1 . Les mesures (en radians) des angles orientés $(-\dot{\mathbf{r}}(+\infty), \delta(\dot{\mathbf{r}}))$ et $(\delta(\dot{\mathbf{r}}), \dot{\mathbf{r}}(-\infty))$ sont égales (modulo 2π) à une valeur commune α : par suite, la position angulaire de P étant spécifiée par la mesure $\theta(t)$ de l'angle orienté $(\delta(\dot{\mathbf{r}}), \mathbf{r}(t)) = (\mathbf{e}_1, \mathbf{r}(t))$, nous obtenons

$$\theta(+\infty) = -\theta(-\infty) = \pi - \alpha$$

Nous nous intéressons à la déflexion angulaire de la trajectoire de P , c'est-à-dire à la mesure β de l'angle orienté $(\dot{\mathbf{r}}(-\infty), \dot{\mathbf{r}}(+\infty))$ (voir Fig. 3). Nous allons donner une nouvelle démonstration de la Proposition 5.1 affirmant que la déflexion angulaire du rayon lumineux rasant la surface du soleil vérifie $2 \tan(\beta/2) = r_{\bullet}/b$ (d'où $\beta \simeq r_{\bullet}/b \simeq r_{\bullet}/r_{\odot}$).

Preuve de la Proposition 5.1 (bis). Nous allons déterminer $\tan(\beta/2)$ en calculant $(\delta(\dot{\mathbf{r}})|_{\mathbf{e}_1})$ de deux manières différentes. D'abord, par définition de α nous avons d'une part $\delta(\dot{\mathbf{r}}) =$

11. Sans perte de généralité nous pouvons supposer que $\beta > 0$.

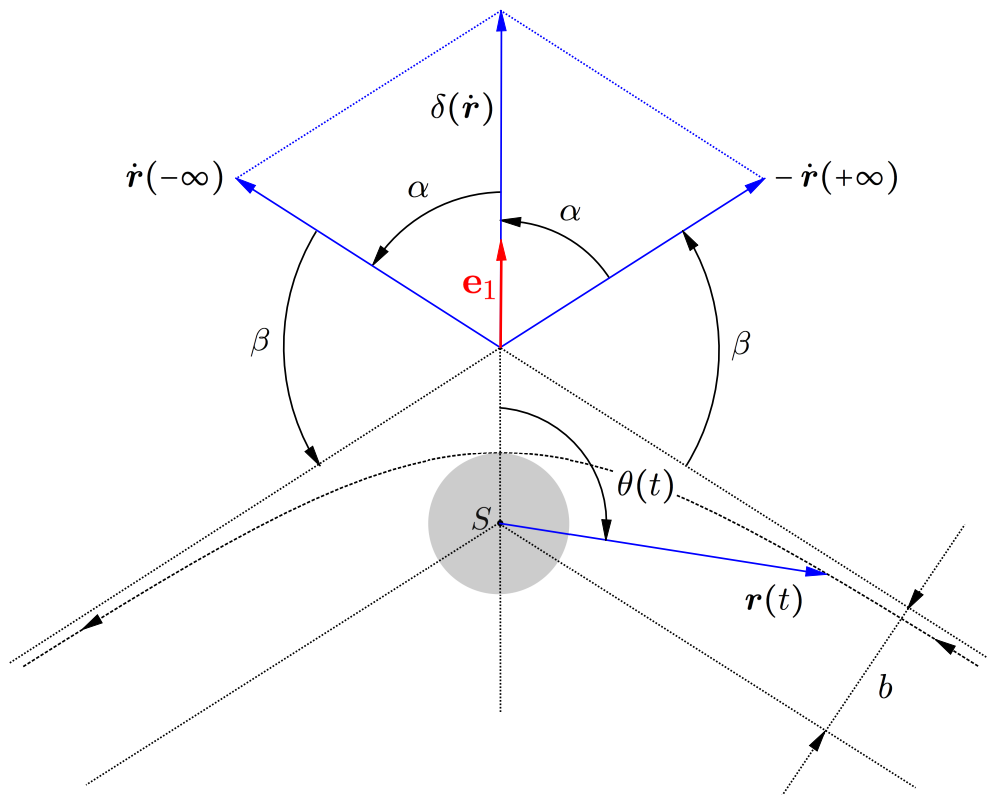


FIGURE 3. Déflexion newtonienne hyperbolique.

$2c \cos(\alpha) \mathbf{e}_1$; comme $\pi = 2\alpha + \beta$ il vient $\alpha = \pi/2 - \beta/2$ (voir Fig. 3). et donc

$$(13) \quad \langle \delta(\dot{\mathbf{r}}) | \mathbf{e}_1 \rangle = 2c \sin(\beta/2)$$

Mais d'autre part, en utilisant le fait $bc = \sigma_0 = \rho^2 \dot{\theta}$, il vient

$$\begin{aligned} \langle \delta(\dot{\mathbf{r}}) | \mathbf{e}_1 \rangle &= - \left\langle \int_{-\infty}^{+\infty} \frac{GM_\odot}{\rho^2(t)} \cdot \frac{\mathbf{r}(t)}{\rho(t)} dt \middle| \mathbf{e}_1 \right\rangle \\ &= - \int_{-\infty}^{+\infty} \frac{GM_\odot}{\rho^2(t)} \cdot \cos(\theta(t)) dt \\ &= - \int_{-\infty}^{+\infty} \frac{GM_\odot}{\rho^2(t) \dot{\theta}(t)} \dot{\theta}(t) \cos(\theta(t)) dt = - \frac{GM_\odot}{bc} \int_{\pi-\alpha}^{\alpha-\pi} \cos(\theta) d\theta = \frac{2GM_\odot}{cb} \sin(\alpha) \end{aligned}$$

soit encore en utilisant de nouveau le fait que $\alpha = \pi/2 - \beta/2$

$$(14) \quad \langle \delta(\dot{\mathbf{r}}) | \mathbf{e}_1 \rangle = \frac{2GM_\odot}{cb} \cos(\beta/2)$$

Par les identités en (13) et (14), nous pouvons alors conclure que :

$$\tan\left(\frac{\beta}{2}\right) = \frac{GM_\odot}{c^2 b} = \frac{1}{2} \frac{r_\bullet}{b}$$

□

7. Déflexion newtonienne III

7.1. Heuristique de l'argument par la méthode Binet. Considérons de nouveau la déflexion de la trajectoire hyperbolique de notre particule P plongée dans le champ de gravitation de la masse solaire M_\odot concentrée en S . Le calcul que nous allons maintenant développer suit les lignes de l'argument analogue dans le cadre de la Relativité Générale et aboutissant à la déflexion double de Schwarzschild – c.f. § 9 infra. Nous changeons légèrement les conditions de repérage du paragraphe précédent : l'orbite hyperbolique de P s'effectue dans le plan $(S; \mathbf{e}_1, \mathbf{e}_2)$, le vecteur vitesse de P en $t = -\infty$ étant $\dot{\mathbf{r}}(-\infty) = -c\mathbf{e}_1$. Ainsi $(\rho(t), \theta(t))$ sont les coordonnées polaires de P , où $\rho(t)$ est la norme de $\mathbf{r}(t)$ et où $\theta(t)$ est la mesure de l'angle orienté $(\mathbf{e}_1, \mathbf{r}(t))$. Dans le cas considéré ici (i.e. rayon lumineux rasant la surface solaire), la périhélie r_\star coïncide avec le rayon r_\odot du soleil. Sans perte de généralité, nous pouvons supposer que $r_\odot = \rho(0)$ (de sorte que $\dot{\rho}(t) = 0$ ssi $t = 0$) et en posant $\theta_\star := \theta(0)$, nous obtenons $\theta(-\infty) = 0$ et $\theta(+\infty) = 2\theta_\star$ (c.f. Fig. 4). Avec ces conditions de repérage, l'angle de déflexion de P est $\beta := 2\theta_\star - \pi$, ce qui donne :

$$(15) \quad \frac{\beta}{2} + \frac{\pi}{2} = \theta_\star = \int_0^{\theta_\star} d\theta$$

7.2. Méthode de Binet. La particule P est ici soumise à un champ d'accélération dérivant du potentiel newtonien a symétrie sphérique, soit $(\rho, \theta, \phi) \mapsto \Psi(\rho) = -GM_\odot/\rho$. L'orbite hyperbolique $t \mapsto \mathbf{r}(t)$ de P , qui s'effectue dans le plan $(S; \mathbf{e}_1, \mathbf{e}_2)$, est caractérisée par l'énergie (le hamiltonien) et la vitesse aréolaire, soient :

$$(16) \quad E = \frac{\mathbf{v}^2}{2} - \frac{GM_\odot}{\rho} \quad \text{et} \quad \sigma_0 = \rho^2 \frac{d\theta}{dt}$$

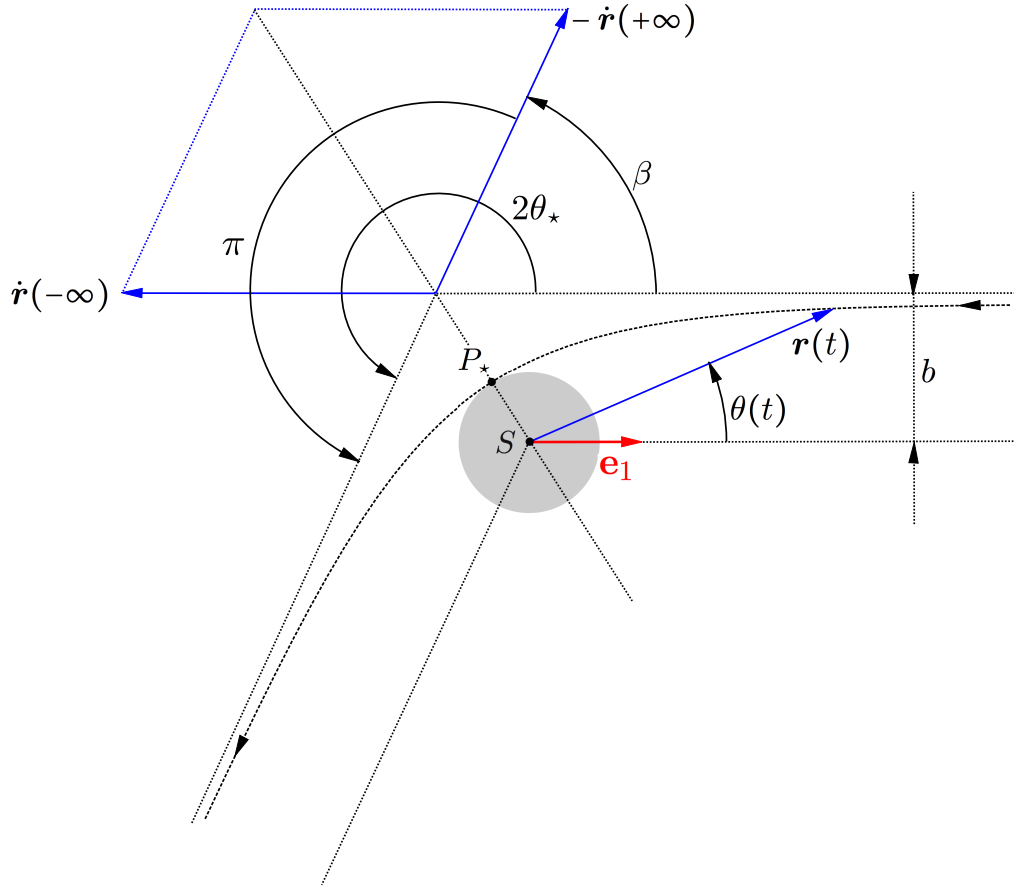


FIGURE 4. Déflexion newtonienne de l'orbite hyperbolique d'une particule P : le point P_* représente la périhélie de l'orbite de P , i.e. le point de l'orbite qui est le plus proche du Soleil.

Lemme 7.1. Soit $t \mapsto (\rho(t), \theta(t))$ le paramétrage polaire de l'orbite hyperbolique de la particule P prise dans le champ de gravitation du soleil, de vitesse $-c\mathbf{e}_1$ en $t = -\infty$ et atteignant sa périhélie au temps $t = 0$ avec $\rho(0) = r_\odot$ (et donc $d\theta/dt = 0$ en $t = 0$) ; alors (i) : l'énergie de P est $E = c^2/2$ et

$$(17) \quad \left(\frac{1}{\rho^2} \frac{d\rho}{d\theta} \right)^2 = \left(\frac{c}{\sigma_0} \right)^2 - \left(\frac{1}{\rho} \right)^2 + \left(\frac{c}{\sigma_0} \right)^2 \frac{r_\bullet}{\rho}$$

(ii) : avec le changement de variable de Binet $x := r_\odot/\rho$ nous avons pour $0 \leq \theta \leq \theta_*$

$$(18) \quad d\theta = \left(1 - x^2 + \frac{r_\bullet}{r_\odot} x \right)^{-1/2} dx$$

Preuve. (i) : La base polaire mobile $(\mathbf{e}_\rho, \mathbf{e}_\theta)$ étant orthonormée directe, l'expression polaire de la vitesse en (4) entraîne que $\mathbf{v}^2 = (d\rho/dt)^2 + \rho^2(d\theta/dt)^2$. L'énergie de P dans le champ de gravitation de S donnée en (16) s'écrit alors

$$(19) \quad E = \frac{1}{2} \left(\frac{d\rho}{dt} \right)^2 + \frac{\rho^2}{2} \left(\frac{d\theta}{dt} \right)^2 - \frac{GM_\odot}{\rho} = \frac{1}{2} \left(\frac{d\rho}{dt} \right)^2 + \frac{1}{2} \left(\frac{\sigma_0}{\rho} \right)^2 - \frac{c^2}{2} \frac{r_\bullet}{\rho}$$

L'hypothèse du mouvement hyperbolique de P assure que $\rho(\pm\infty) = +\infty$: avec une vitesse $v(+\infty)$ égale à la vitesse de la lumière c , il suit de (16) que $E = c^2/2$ et (19) donne

$$c^2 - \left(\frac{\sigma_0}{\rho}\right)^2 + c^2 \frac{r_\bullet}{\rho} = \left(\frac{d\rho}{dt}\right)^2 = \left(\frac{d\theta}{dt} \frac{d\rho}{d\theta}\right)^2 = \left(\frac{\sigma_0}{\rho^2} \frac{d\rho}{d\theta}\right)^2$$

ce qui donne (17)

(ii) : Dans notre configuration (rayon lumineux hyperbolique rasant la surface solaire), la périhélie a lieu à $t = 0$ avec $\rho(0) = r_\odot$ (rayon solaire). Nous appliquerons la méthode de Binet pour $-\infty \leq t \leq 0$ en posant $x = r_\odot/\rho$, de sorte que sur l'intervalle de temps considéré, x croît de 0 à 1. Dans ce cas, nous pouvons écrire à l'aide de (17) :

$$\left(\frac{dx}{d\theta}\right)^2 = \left(\frac{dx}{d\rho} \frac{d\rho}{d\theta}\right)^2 \left(\frac{r_\odot}{\rho^2} \frac{d\rho}{d\theta}\right)^2 = \left(\frac{cr_\odot}{\sigma_0}\right)^2 - \left(\frac{r_\odot}{\rho}\right)^2 + \left(\frac{cr_\odot}{\sigma_0}\right)^2 \frac{r_\bullet}{\rho} = \left(\frac{cr_\odot}{\sigma_0}\right)^2 - x^2 + \left(\frac{cr_\odot}{\sigma_0}\right)^2 \frac{r_\bullet}{r_\odot} x$$

Or $\sigma_0 = cb$ et par suite avec l'approximation $b \simeq r_\odot$, nous obtenons finalement :

$$\left(\frac{dx}{d\theta}\right)^2 \simeq 1 - x^2 + \frac{r_\bullet}{r_\odot} x$$

Nous obtenons (18) du fait que θ est une fonction croissante de x lorsque $t \leq 0$. □

Proposition 7.2. *L'angle de la déflexion newtonienne d'un rayon de lumière rasant la surface solaire est (approximativement) $\beta \simeq r_\bullet/r_\odot$.*

Preuve. Rappelons que le rayon de Schwarzschild de la masse solaire M_\odot est $r_\bullet = GM_\odot/c^2$: ce rayon est très petit relativement au rayon solaire r_\odot : grâce aux valeurs mesurées des grandeurs en question, nous savons – voir (12) – que $r_\bullet/r_\odot \simeq 4,24 \times 10^{-6}$. Il est ainsi possible d'approximer (18) en écrivant :

$$d\theta = \left(1 - \left(x^2 - \frac{r_\bullet}{r_\odot} + \left(\frac{1}{2} \frac{r_\bullet}{r_\odot}\right)^2\right) + \left(\frac{1}{2} \frac{r_\bullet}{r_\odot}\right)^2\right)^{-1/2} dx \simeq \left(1 - \left(x - \frac{1}{2} \frac{r_\bullet}{r_\odot}\right)^2\right)^{-1/2} dx$$

Comme x varie de 0 à 1 lorsque θ varie de $\theta(-\infty) = 0$ à $\theta(0) = \theta_*$, il vient avec (15) :

$$\begin{aligned} \frac{\beta}{2} + \frac{\pi}{2} = \theta_* &= \int_0^{\theta_*} d\theta \simeq \int_0^1 \left(1 - \left(x - \frac{1}{2} \frac{r_\bullet}{r_\odot}\right)^2\right)^{-1/2} dx \\ &\simeq \text{Arcsin}\left(1 - \frac{1}{2} \frac{r_\bullet}{r_\odot}\right) - \text{Arcsin}\left(-\frac{1}{2} \frac{r_\bullet}{r_\odot}\right) \simeq \frac{\pi}{2} + \frac{1}{2} \frac{r_\bullet}{r_\odot} \end{aligned}$$

□

8. La métrique de Schwarzschild

Rappelons (c.f. [Oli15b, § 7]) que le $(0, 2)$ -tenseur énergie-impulsion T de la Relativité Générale décrit la distribution spatio-temporelle de matière-énergie. C'est un tenseur symétrique (i.e. en composantes $T_{\mu\nu} = T_{\nu\mu}$), s'annulant uniformément dans les portions d'espace vide de matière-énergie et dont la divergence est nulle. Les équations d'Einstein s'écrivent en composantes sous la forme suivante :

$$(20) \quad R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \frac{8\pi G}{c^4} T_{\mu\nu}$$

À droite, nous reconnaissons les composantes du tenseur d'Einstein, où $R_{\mu\nu}$ sont les composantes du tenseur de Ricci et où R est le scalaire de Ricci. Le fait (c.f. [Oli15b, (iv)-Proposition 7.1]) que le tenseur d'Einstein et le tenseur de Ricci s'annulent simultanément implique que les équations d'Einstein dans le vide se réduisent à

$$(21) \quad R_{\mu\nu} = 0$$

La métrique de Schwarzschild est la première solution exacte du système d'équations (21), pour laquelle la courbure n'est pas identiquement nulle. Elle décrit la géométrie de l'espace-temps courbé par l'effet gravitationnel d'une masse à symétrie sphérique et stationnaire (il n'y a pas d'ondes gravitationnelles) dans le voisinage vide autour de la masse centrale.

Théorème 8.1 (Schwarzschild). Notons $(t, \rho, \theta, \varphi)$ les coordonnées sphériques de l'espace-temps (d'origine S) et pour $r_0 \geq 0$ donné, soit $\mathcal{U}(r_0)$ l'ouvert défini par la condition $\rho > r_0$: alors

(i) : l'équation d'Einstein dans le vide (annulation uniforme du tenseur de Ricci) admet sur $\mathcal{U}(r_0)$ une solution (métrique lorentzienne) à symétrie sphérique stationnaire, et coïncidant asymptotiquement avec la métrique de Minkowski pour $r \rightarrow +\infty$, soit :

$$(22) \quad ds^2 = -B(\rho)d\mathbf{t}^2 + A(\rho)d\rho^2 + \rho^2 d\theta^2 + \rho^2 \sin^2 \theta d\varphi^2 \quad \text{où} \quad B(\rho) = \frac{1}{A(\rho)} = 1 - \frac{r_0}{\rho}$$

Ici, $\mathbf{t} = ct$ est le temps renormalisé de l'espace-temps (avec c la célérité de la lumière) et ds est l'intervalle d'espace-temps d'un déplacement infinitésimal $(d\mathbf{t}, d\rho, d\theta, d\varphi)$ (le cas $r_0 = 0$ correspond à l'espace-temps sans gravitation, i.e. l'espace de Minkowski) ;

(ii) : les géodésiques non relativistes de cette métrique s'approximent, dans la limite du champ faible, par les orbites newtoniennes du champ de gravitation créé par une masse ponctuelle M (en S), lorsque r_0 est égal au rayon de Schwarzschild¹², soit :

$$(23) \quad r_\bullet := \frac{2GM}{c^2}$$

dans ces conditions, la métrique de Schwarzschild sur $\mathcal{U}(r_\bullet)$ est :

$$(24) \quad ds^2 = -\left(1 - \frac{r_\bullet}{\rho}\right) d\mathbf{t}^2 + \left(1 - \frac{r_\bullet}{\rho}\right)^{-1} d\rho^2 + \rho^2 d\theta^2 + \rho^2 \sin^2 \theta d\varphi^2$$

La première solution des équations d'Einstein est l'espace-temps de la Relativité Restreinte, i.e. \mathbb{R}^4 muni de la métrique de Minkowski, soit (en coordonnées sphérique)

$$ds^2 = -d\mathbf{t}^2 + d\rho^2 + \rho^2 d\varphi + \rho^2 \sin^2 \theta d\theta^2$$

Il s'agit là d'une solution sans gravitation ! Afin de trouver une solution proche de l'espace Minkowski, mais intégrant de la gravitation, Schwarzschild recherche l'analogie – pour la Relativité Générale – du champ newtonien de gravitation engendré par une masse centrale M (c'est la plus simple des situations imaginables). L'espace temps étant rapporté aux coordonnées sphériques $(\mathbf{t}, \rho, \theta, \varphi)$, nous considérons l'ouvert $\mathcal{U}(r_0) := \{\rho > r_0\}$ où $r_0 > 0$ est (pour l'instant) arbitraire. Plus précisément Schwarzschild recherche une solution de la forme

$$(25) \quad ds^2 = -B(\rho) d\mathbf{t}^2 + A(\rho) d\rho^2 + \rho^2 d\theta^2 + \rho^2 \sin^2 \theta d\varphi^2$$

12. Encore appelé l'horizon des événements.

où $\rho \mapsto A(\rho)$ et $\rho \mapsto B(\rho)$ sont des fonctions positives qu'il reste à déterminer (au signe près, nous utilisons les notations de Weinberg dans [Wei72, eq-(8.2.1) p. 179]). Nous reviendront sur l'introduction du rayon r_o , mais disons que le cas limite $r_o = 0$ doit correspondre à l'espace de Minkowski et que pour $r_o > 0$, la distribution de matière-énergie (à l'origine de l'effet gravitationnel) est supposée confinée dans le domaine $\{\rho \leq r_o\}$.

8.1. Preuve du Théorème 8.1 – partie (i). Rappelons que (dans le système de coordonnées locales d'une carte) les symboles de Christoffel s'expriment en fonction des dérivées partielles des composantes de la métrique, avec

$$\mathbf{C}^{\sigma}_{\mu\nu} = \frac{1}{2}g^{\sigma\epsilon}(\partial_{\mu}(g_{\nu\epsilon}) + \partial_{\nu}(g_{\mu\epsilon}) - \partial_{\epsilon}(g_{\mu\nu}))$$

Le tenseur de Ricci s'écrit alors :

$$(26) \quad R_{\mu\nu} = \partial_{\sigma}(\mathbf{C}^{\sigma}_{\mu\nu}) - \partial_{\nu}(\mathbf{C}^{\sigma}_{\sigma\mu}) + \mathbf{C}^{\sigma}_{\sigma\eta}\mathbf{C}^{\eta}_{\mu\nu} - \mathbf{C}^{\sigma}_{\nu\eta}\mathbf{C}^{\eta}_{\sigma\mu} = R_{\nu\mu}$$

Grâce à la symétrie des symboles de Christoffel (i.e. $\mathbf{C}^{\sigma}_{\mu\nu} = \mathbf{C}^{\sigma}_{\nu\mu}$) l'expression des composantes du tenseur de Ricci en (26) nous donne $R_{\mu\nu} = R_{\nu\mu}$: la symétrie du tenseur de Ricci est essentielle quant à la cohérence de l'équation d'Einstein (le tenseur énergie-impulsion étant symétrique par définition). La forme spéciale de la métrique recherchée nous permet de calculer explicitement les composantes du tenseur de Ricci. Le fait que $R_{\mu\nu} = R_{\nu\mu}$ réduit le nombre de composantes cherchées à $10 = 4 + 3 + 2 + 1$: ces composantes sont figurées dans la matrice suivante :

$$\begin{pmatrix} R_{tt} & R_{t\rho} & R_{t\theta} & R_{t\varphi} \\ * & R_{\rho\rho} & R_{\rho\theta} & R_{\rho\varphi} \\ * & * & R_{\theta\theta} & R_{\theta\varphi} \\ * & * & * & R_{\varphi\varphi} \end{pmatrix}$$

Les calculs sont assez fastidieux mais pas très difficiles : nous allons par exemple expliciter la composante R_{tt} . Pour cela nous développons l'expression en (26) pour $(\mu, \nu) = (t, t)$ à partir de la métrique sous la forme incomplète donnée en (25). La métrique étant diagonale dans le système de coordonnées sphérique (i.e. $g_{\mu\nu} = 0$ dès que $\mu \neq \nu$) et stationnaire (i.e. $\partial_t(g_{\mu\nu}) = 0$ pour tout μ, ν), nous pouvons utiliser l'expression de la contraction des symboles de Christoffel (i.e. $\mathbf{C}^{\sigma}_{\sigma\nu} = 1/2\partial_{\nu}(g)/g$, où $g = \det(g)$ – c.f. [Oli15a, Lemme 15.2]) pour écrire :

$$\begin{aligned} R_{tt} &= \partial_{\sigma}(\mathbf{C}^{\sigma}_{tt}) - \partial_t(\mathbf{C}^{\sigma}_{\sigma t}) + \mathbf{C}^{\sigma}_{\sigma\eta}\mathbf{C}^{\eta}_{tt} - \mathbf{C}^{\sigma}_{t\eta}\mathbf{C}^{\eta}_{\sigma t} \\ &= \partial_{\sigma}(\mathbf{C}^{\sigma}_{tt}) - \partial_t\left(\frac{1}{2g}\partial_t(g)\right) + \left(\frac{1}{2g}\partial_{\alpha}(g)\right)\mathbf{C}^{\alpha}_{tt} - \mathbf{C}^{\sigma}_{t\alpha}\mathbf{C}^{\alpha}_{t\sigma} \\ &= \partial_{\sigma}(\mathbf{C}^{\sigma}_{tt}) + \left(\frac{1}{2g}\partial_{\rho}(g)\right)\mathbf{C}^{\rho}_{tt} - \mathbf{C}^{\sigma}_{t\alpha}\mathbf{C}^{\alpha}_{t\sigma} \end{aligned}$$

Du fait que $g = -\rho^4 AB \sin^2(\theta)$, il vient :

$$\frac{1}{2g}\partial_{\rho}(g) = \frac{1}{2\rho^4 AB} \left(4\rho^3 AB + \rho^4 \left(\frac{dA}{d\rho} \right) B + \rho^4 A \left(\frac{dB}{d\rho} \right) \right) = \frac{2}{\rho} + \frac{1}{2A} \left(\frac{dA}{d\rho} \right) + \frac{1}{2B} \left(\frac{dB}{d\rho} \right)$$

ce qui donne

$$(27) \quad R_{tt} = \partial_{\alpha}(\mathbf{C}^{\alpha}_{tt}) + \left(\frac{2}{\rho} + \frac{1}{2A} \left(\frac{dA}{d\rho} \right) + \frac{1}{2B} \left(\frac{dB}{d\rho} \right) \right) \mathbf{C}^{\alpha}_{tt} - \mathbf{C}^{\sigma}_{t\alpha}\mathbf{C}^{\alpha}_{t\sigma}$$

Pour les symboles de Christoffel restants, nous avons d'une part,

$$\mathbf{C}^{\sigma}_{t\alpha} = \frac{1}{2}g^{\sigma\epsilon}(\partial_t(g_{\alpha\epsilon}) + \partial_\alpha(g_{t\epsilon}) - \partial_\epsilon(g_{t\alpha})) = \frac{1}{2}g^{\sigma\sigma}(\partial_\alpha(g_{t\sigma}) - \partial_\sigma(g_{t\alpha}))$$

ce qui donne

$$\mathbf{C}^t_{t\rho} = \frac{1}{2}g^{tt}\partial_\rho(g_{tt}) = \frac{1}{2B}\left(\frac{dB}{d\rho}\right) \quad \text{et} \quad \mathbf{C}^\rho_{tt} = -\frac{1}{2}g^{\rho\rho}\partial_\rho(g_{tt}) = \frac{1}{2A}\left(\frac{dB}{d\rho}\right)$$

avec $\mathbf{C}^\sigma_{t\alpha} = 0$ dès que $(\sigma, \alpha) \notin \{(\rho, t), (t, \rho)\}$. D'autre part,

$$\mathbf{C}^\alpha_{tt} = \frac{1}{2}g^{\alpha\epsilon}(\partial_t(g_{t\epsilon}) + \partial_t(g_{t\epsilon}) - \partial_\epsilon(g_{tt})) = -\frac{1}{2}g^{\alpha\epsilon}\partial_\epsilon(g_{tt}) = \frac{1}{2}g^{\alpha\rho}\left(\frac{dB}{d\rho}\right)$$

soit encore

$$\mathbf{C}^\rho_{tt} = \frac{1}{2A(\rho)}\left(\frac{dB}{d\rho}\right)$$

et $\mathbf{C}^\alpha_{tt} = 0$ dès que $\alpha \neq \rho$. Finalement, l'équation (27) devient

$$\begin{aligned} R_{tt} &= \frac{1}{2}\frac{d}{d\rho}\left(\frac{1}{A}\left(\frac{dB}{d\rho}\right)\right) + \left(\frac{2}{\rho} + \frac{1}{2A}\left(\frac{dA}{d\rho}\right) + \frac{1}{2B}\left(\frac{dB}{d\rho}\right)\right)\left(\frac{1}{2A}\left(\frac{dB}{d\rho}\right)\right) - \frac{1}{2AB}\left(\frac{dB}{d\rho}\right)^2 \\ &= -\frac{1}{2A^2}\left(\frac{dA}{d\rho}\right)\left(\frac{dB}{d\rho}\right) + \frac{1}{2A}\left(\frac{d^2B}{d\rho^2}\right) \\ &\quad + \frac{1}{\rho A}\left(\frac{dB}{d\rho}\right) + \frac{1}{4A^2}\left(\frac{dA}{d\rho}\right)\left(\frac{dB}{d\rho}\right) + \frac{1}{4AB}\left(\frac{dB}{d\rho}\right)^2 - \frac{1}{2AB}\left(\frac{dB}{d\rho}\right)^2 \\ &= \frac{1}{2A}\left(\frac{d^2B}{d\rho^2}\right) + \frac{1}{\rho A}\left(\frac{dB}{d\rho}\right) - \frac{1}{4A^2}\left(\frac{dA}{d\rho}\right)\left(\frac{dB}{d\rho}\right) - \frac{1}{4AB}\left(\frac{dB}{d\rho}\right)^2 \end{aligned}$$

soit encore

$$R_{tt} = \frac{1}{2A}\left(\frac{d^2B}{d\rho^2}\right) + \frac{1}{\rho A}\left(\frac{dB}{d\rho}\right) - \frac{1}{4A}\left(\frac{dB}{d\rho}\right)\left(\frac{1}{A}\left(\frac{dA}{d\rho}\right) + \frac{1}{B}\left(\frac{dB}{d\rho}\right)\right)$$

Les autres composantes du tenseur de Ricci s'obtiennent par un calcul analogue nécessitant la liste complète des expressions des symboles de Christoffel. Nous pouvons énoncer la proposition suivantes (c.f. Weinberg in [Wei72, eq. (8.1.13) p. 178]).

Proposition 8.2. *Soit g une métrique – définie sur un ouvert de l'espace-temps à symétrie sphérique – dont l'élément de volume associé s'écrit sous la forme*

$$ds^2 = -B(\rho) dt^2 + A(\rho) d\rho^2 + \rho^2 d\theta^2 + \rho^2 \sin^2 \theta d\varphi^2$$

Alors (dans la carte des coordonnées sphériques) (i) : les symboles de Christoffel non nuls sont :

$$(28) \quad \begin{aligned} \mathbf{C}^\rho_{\rho\rho} &= \frac{1}{2A(\rho)}\frac{dA}{d\rho} & \mathbf{C}^\rho_{\theta\theta} &= -\frac{\rho}{A(\rho)} \\ \mathbf{C}^\rho_{\varphi\varphi} &= -\frac{\rho \sin^2(\theta)}{A(\rho)} & \mathbf{C}^\rho_{tt} &= \frac{1}{2A(\rho)}\frac{dB}{d\rho} \\ \mathbf{C}^\theta_{\rho\theta} = \mathbf{C}^\theta_{\theta\rho} &= \frac{1}{\rho} & \mathbf{C}^\theta_{\varphi\varphi} &= -\sin(\theta)\cos(\theta) \\ \mathbf{C}^\varphi_{\rho\varphi} = \mathbf{C}^\varphi_{\varphi\rho} &= \frac{1}{\rho} & \mathbf{C}^\varphi_{\theta\varphi} = \mathbf{C}^\varphi_{\varphi\theta} &= -\cos(\theta) \\ \mathbf{C}^t_{\rho t} = \mathbf{C}^t_{t\rho} &= \frac{1}{2B(\rho)}\frac{dB}{d\rho} \end{aligned}$$

et (ii) : $R_{\mu\nu} = 0$ dès que $\mu \neq \nu$, les valeurs des composantes diagonales étant

$$(29) \quad R_{tt} = \frac{1}{2A} \left(\frac{d^2 B}{d\rho^2} \right) + \frac{1}{\rho A} \left(\frac{dB}{d\rho} \right) - \frac{1}{4A} \left(\frac{dB}{d\rho} \right) \left(\frac{1}{A} \left(\frac{dA}{d\rho} \right) + \frac{1}{B} \left(\frac{dB}{d\rho} \right) \right)$$

$$(30) \quad R_{\rho\rho} = -\frac{1}{2B} \left(\frac{d^2 B}{d\rho^2} \right) + \frac{1}{\rho A} \left(\frac{dA}{d\rho} \right) + \frac{1}{4B} \left(\frac{dB}{d\rho} \right) \left(\frac{1}{A} \left(\frac{dA}{d\rho} \right) + \frac{1}{B} \left(\frac{dB}{d\rho} \right) \right)$$

$$(31) \quad R_{\theta\theta} = 1 + \frac{\rho}{2A} \left(\frac{1}{A} \left(\frac{dA}{d\rho} \right) - \frac{1}{B} \left(\frac{dB}{d\rho} \right) \right) - \frac{1}{A}$$

$$(32) \quad R_{\varphi\varphi} = \sin^2(\theta) \left(1 + \frac{\rho}{2A} \left(\frac{1}{A} \left(\frac{dA}{d\rho} \right) - \frac{1}{B} \left(\frac{dB}{d\rho} \right) \right) - \frac{1}{A} \right) = \sin^2(\theta) R_{\theta\theta}$$

Les équations d'Einstein dans l'ouvert $\mathcal{U}(r_0)$ (considéré par hypothèse comme un espace vide de matière-énergie), se réduisent à $R_{\mu\nu} = 0$. Si de plus, nous nous plaçons dans les coordonnées sphériques, alors d'après la Proposition 8.2, nous savons que la matrice $(R_{\mu\nu})$ est diagonale, ce qui nous laisse avec quatre équations dont une est redondante. En résumé, il nous reste le système des trois équations suivantes (les équations (33) et (34) sont respectivement obtenues en multipliant (29) et (30) par ρA^2 et ρAB) :

$$(33) \quad 0 = \frac{\rho A}{2} \left(\frac{d^2 B}{d\rho^2} \right) + A \left(\frac{dB}{d\rho} \right) - \frac{\rho A}{4} \left(\frac{dB}{d\rho} \right) \left(\frac{1}{A} \left(\frac{dA}{d\rho} \right) + \frac{1}{B} \left(\frac{dB}{d\rho} \right) \right)$$

$$(34) \quad 0 = -\frac{\rho A}{2} \left(\frac{d^2 B}{d\rho^2} \right) + B \left(\frac{dA}{d\rho} \right) + \frac{\rho A}{4} \left(\frac{dB}{d\rho} \right) \left(\frac{1}{A} \left(\frac{dA}{d\rho} \right) + \frac{1}{B} \left(\frac{dB}{d\rho} \right) \right)$$

$$(35) \quad 0 = 1 + \frac{\rho}{2A} \left(\frac{1}{A} \left(\frac{dA}{d\rho} \right) - \frac{1}{B} \left(\frac{dB}{d\rho} \right) \right) - \frac{1}{A}$$

La somme de (33) et (34) donne $A(dB/d\rho) + (dA/d\rho)B = 0$, d'où $A(\rho)B(\rho) = K$ avec K une constante non-nulle ; or la coïncidence à l'infini de la métrique de Schwarzschild avec celle de Minkowski entraîne que $A(+\infty) \sim 1 \sim B(+\infty)$, ce qui assure que $K = 1$ et par suite

$$(36) \quad A(\rho)B(\rho) = 1$$

Il découle de (36) que $AdB/d\rho + BdA/d\rho = 0$ et (35) se transforme en une équation différentielle en B , soit $1 - \rho dB/d\rho - B = 0$, i.e.

$$\frac{d}{d\rho}(\rho B(\rho)) = 1$$

Nous en déduisons qu'il existe une constante réelle r_0 t.q.

$$B(\rho) = \left(1 - \frac{r_0}{\rho} \right)$$

et (25) devient :

$$(37) \quad ds^2 = - \left(1 - \frac{r_0}{\rho} \right) dt^2 + \left(1 - \frac{r_0}{\rho} \right)^{-1} d\rho^2 + \rho^2 d\theta^2 + \rho^2 \sin^2 \theta d\varphi^2$$

8.2. Preuve du Théorème 8.1 – partie (ii). Il reste à déterminer la constante r_0 . C'est un point délicat car il s'agit d'interpoler la théorie de la gravitation de Newton, la Relativité Restreinte et la Relativité Générale. Disons grossièrement que si $\rho \mapsto \Phi(\rho)$ est la fonction telle que $-g_{tt} = \exp(2\Phi(\rho)/c^2)$, alors l'approximation du champ faible affirme (c.f. Remarque 9.4 infra) que dans la limite $\rho \rightarrow +\infty$ la valeur de $\Phi(\rho)$ s'approche du potentiel

newtonien engendré par la masse M concentrée au centre de symétrie : en d'autres termes $\Phi(\rho) \simeq -GM/\rho$ quand $\rho \rightarrow +\infty$ et l'approximation linéaire de l'exponentielle donne

$$1 - \frac{r_0}{\rho} = \exp\left(\frac{2\Phi(\rho)}{c^2}\right) \simeq \exp\left(-\frac{2GM}{c^2\rho}\right) \simeq 1 - \frac{2GM}{c^2\rho}$$

Ceci n'est possible que si r_0 coïncide avec le rayon de Schwarzschild $r_\bullet = 2GM/c^2$.

Le point (ii) du Théorème 8.1 représente – ici – le trait d'union entre la gravitation selon Newton et Einstein ; de plus il fait apparaître le rayon de Schwarzschild r_\bullet comme une singularité (apparente : c.f. [Wei72, Chap 8.8]) de la métrique. Cette question doit être approfondie : pour cela nous nous référons à [Ton64, eq VI-42][Wei72, Chap 3.4 & 8.2][Car13, eq. 5.22].

9. Déflexion de la lumière pour la Relativité Générale

9.1. **Équation des géodésiques inertielles.** Nous nous intéressons aux géodésiques inertielle dans la projection spatiale se trouve dans le plan $\varphi = \pi/2$ de $\mathcal{U}(r_\bullet)$. Nous sommes alors ramenés à un espace-temps à 3 dimensions avec une dimension temporelle t et deux dimensions spatiales (ρ, θ) en paramétrisation polaires. La métrique de Schwarzschild se réduit alors à

$$(38) \quad ds^2 = -B(\rho)dt^2 + A(\rho)d\rho^2 + \rho^2d\theta^2$$

avec $B(\rho) = 1 - r_\bullet/\rho$ et $A(\rho) = 1/B(\rho)$. Une géodésique $u \mapsto \gamma(u) = (\gamma^t(u), \gamma^\rho(u), \gamma^\theta(u))$ pour cette métrique doit satisfaire le système des trois équations suivantes :

$$(39) \quad \frac{d^2\gamma^t}{du^2} + \mathbf{C}^t_{\mu\nu}(\gamma(u)) \frac{d\gamma^\mu}{du} \frac{d\gamma^\nu}{du} = 0$$

$$(40) \quad \frac{d^2\gamma^\rho}{du^2} + \mathbf{C}^\rho_{\mu\nu}(\gamma(u)) \frac{d\gamma^\mu}{du} \frac{d\gamma^\nu}{du} = 0$$

$$(41) \quad \frac{d^2\gamma^\theta}{du^2} + \mathbf{C}^\theta_{\mu\nu}(\gamma(u)) \frac{d\gamma^\mu}{du} \frac{d\gamma^\nu}{du} = 0$$

Par exemple, pour l'équation (40) – et compte tenu des expressions des symboles de Christoffel données en (28) – nous obtenons (avec les termes non nuls) :

$$\begin{aligned} (40) \quad &\iff \frac{d^2\gamma^\rho}{du^2} + \mathbf{C}^\rho_{\mu\nu}(\gamma(u)) \frac{d\gamma^\mu}{du} \frac{d\gamma^\nu}{du} = 0 \\ &\iff \frac{d^2\gamma^\rho}{du^2} + \mathbf{C}^\rho_{\rho\rho}(\gamma(u)) \frac{d\gamma^\rho}{du} \frac{d\gamma^\rho}{du} + \mathbf{C}^\rho_{\theta\theta}(\gamma(u)) \frac{d\gamma^\theta}{du} \frac{d\gamma^\theta}{du} + \mathbf{C}^\rho_{tt}(\gamma(u)) \frac{d\gamma^t}{du} \frac{d\gamma^t}{du} = 0 \\ &\iff \frac{d^2\gamma^\rho}{du^2} + \frac{1}{2A(\gamma^\rho(u))} \frac{dA}{d\rho}(\gamma^\rho(u)) \frac{d\gamma^\rho}{du} \frac{d\gamma^\rho}{du} \\ &\quad + \frac{\gamma^\rho(u)}{A(\gamma^\rho(u))} \frac{d\gamma^\theta}{du} \frac{d\gamma^\theta}{du} \\ &\quad + \frac{1}{2A(\gamma^\rho(u))} \frac{dB}{d\rho}(\gamma^\rho(u)) \frac{d\gamma^t}{du} \frac{d\gamma^t}{du} = 0 \end{aligned}$$

Les équations (41) et (39) s'établissent d'une manière analogue ; dans la proposition suivante, nous notons simplement $\rho(u) := \gamma^\rho(u)$, $\theta(u) := \gamma^\theta(u)$ et $t(u) := \gamma^t(u)$.

Proposition 9.1. *Le système des équations d'une géodésique $u \mapsto (\mathbf{t}(u), r(u), \theta(u))$ de l'espace-temps de Schwarzschild est formé des trois équations suivantes*

$$(42) \quad \frac{d^2 \mathbf{t}}{du^2} + \frac{1}{B} \frac{dB}{d\rho} \frac{d\rho}{du} \frac{d\mathbf{t}}{du} = 0$$

$$(43) \quad \frac{d^2 \rho}{du^2} + \frac{1}{2A} \frac{dA}{d\rho} \left(\frac{d\rho}{du} \right)^2 - \frac{\rho}{A} \left(\frac{d\theta}{du} \right)^2 + \frac{c^2}{2A} \frac{dB}{d\rho} \left(\frac{d\mathbf{t}}{du} \right)^2 = 0$$

$$(44) \quad \frac{d^2 \theta}{du^2} + \frac{2}{\rho} \frac{d\rho}{du} \frac{d\theta}{du} = 0$$

9.2. **Trois constantes de l'orbite géodésique.** En multipliant (44) par ρ^2 il vient

$$\rho^2 \frac{d^2 \theta}{du^2} + 2\rho \frac{d\rho}{du} \frac{d\theta}{du} = 0 \iff \frac{d}{du} \left(\rho^2 \frac{d\theta}{du} \right) = 0$$

et nous obtenons une première constante du mouvement (**le moment cinétique**), soit

$$(45) \quad \omega_0 = \rho^2 \frac{d\theta}{du}$$

Pour la deuxième constante, nous multiplions (42) par B , ce qui donne

$$B \frac{d^2 \mathbf{t}}{du^2} + \left(\frac{dB(\rho)}{d\rho} \frac{d\rho}{du} \right) \frac{d\mathbf{t}}{du} = 0 \iff \frac{d}{du} \left(B \frac{d\mathbf{t}}{du} \right) = 0$$

d'où une première constante du mouvement, soit

$$(46) \quad \kappa_0 = B \frac{d\mathbf{t}}{du}$$

Nous utilisons le fait qu'il est licite d'effectuer un changement linéaire du paramètre de l'équation d'une géodésique inertielle : ainsi, nous pouvons fixer $\kappa_0 = 1$ dans (46) de sorte que cette dernière équation devient simplement :

$$(47) \quad \frac{d\mathbf{t}}{du} = \frac{1}{B(r(u))} \quad \text{i.e.} \quad \frac{d\mathbf{t}}{du} = \frac{c}{B(r(u))}$$

Ainsi, pour les paramètres u t.q. $r(u) \simeq +\infty$, l'approximation $B(r(u)) \simeq 1$ combinée l'équation (47) entraîne que $\mathbf{t} \simeq u$, i.e. $t \simeq cu$.

Finalement, en partant de l'équation (43) multipliée par $2A dr/du$ il vient avec (45) :

$$\begin{aligned} 2A \frac{d\rho}{du} \left(\frac{d^2 \rho}{du^2} + \frac{1}{2A} \frac{dA}{d\rho} \left(\frac{d\rho}{du} \right)^2 - \frac{2}{2A} \frac{\omega_0^2}{\rho^3} + \frac{1}{2AB^2} \frac{dB}{d\rho} \right) &= 0 \\ A \left(2 \frac{d\rho}{du} \left(\frac{d^2 \rho}{du^2} \right) + \left(\frac{dA}{d\rho} \frac{d\rho}{du} \right) \left(\frac{d\rho}{du} \right)^2 - \frac{2\omega_0^2}{\rho^3} \frac{d\rho}{du} + \frac{1}{B^2} \frac{dB}{d\rho} \frac{d\rho}{du} \right) &= 0 \\ \frac{d}{du} \left(A \left(\frac{d\rho}{du} \right)^2 + \frac{\omega_0^2}{\rho^2} - \frac{1}{B} \right) &= 0 \end{aligned}$$

Nous obtenons ainsi une troisième constante du mouvement (**l'énergie**), soit

$$(48) \quad \mathcal{E} = \frac{c^2}{2} \left(\left(\frac{d\rho}{du} \right)^2 + \left(\frac{\omega_0}{\rho} \right)^2 B - 1 \right) \frac{1}{B}$$

Remarque 9.2. *Le facteur $c^2/2$ est (artificiellement) introduit afin de décomposer l'énergie \mathcal{E} suivant ses contributions radiale (**énergie potentielle effective**) et co-radiale, en écrivant*

$$\mathcal{E} = \frac{c^2}{2} \left(\frac{d\rho}{du} \right)^2 \frac{1}{B} + \mathcal{E}_r(\rho) \quad \text{où} \quad \mathcal{E}_r(\rho) := \frac{c^2}{2} \left(\left(\frac{\omega_0}{\rho} \right)^2 B - 1 \right) \frac{1}{B}$$

Nous verrons (c.f. Remarque 9.4) comment utiliser cette remarque pour justifier l'identité entre la constante r_\circ de la métrique de Schwarzschild donnée en (37) et le rayon de Schwarzschild r_\bullet .

9.3. Paramétrage spatial de l'orbite géodésique. Soit $(S; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ le repère cartésien de la partie spatiale de l'espace-temps, où S est sensé représenter le soleil. Rappelons que la projection spatiale $u \mapsto \mathbf{r}(u)$ de la géodésique $u \mapsto \gamma(u) = (\mathbf{t}(u), \mathbf{r}(u))$ est supposée se dérouler le plan $(S; \mathbf{e}_1, \mathbf{e}_2)$. Nous admettons que cette projection est une quasi-hyperbole, i.e. une perturbation du mouvement hyperbolique newtonien déjà étudié : en particulier, nous admettons l'existence et l'unicité de la périhélie spatiale et nous notons P_\star le point de $(S; \mathbf{e}_1, \mathbf{e}_2)$ correspondant. Le système de repérage spatial est identique à celui utilisé au § 7 (méthode de Binet pour la déflexion newtonienne des rayons lumineux) où $(\rho(u), \theta(u))$ désigne les coordonnées polaires de la position $\mathbf{r}(u)$ (i.e. $\rho(u)$ est la norme de $\mathbf{r}(u)$ et $\theta(u)$ est la mesure de l'angle orienté $(\mathbf{e}_1, \mathbf{r}(u))$) : voir Fig. 4). Dans le cas considéré ici (i.e. modélisation d'un rayon lumineux rasant la surface solaire) nous avons $(\rho(u), \theta(u)) = (+\infty, 0)$ lorsque $u \rightarrow -\infty$. Par analogie avec l'orbite hyperbolique du cas newtonien, il est aussi raisonnable¹³ d'admettre la symétrie (géométrique et dynamique) de l'orbite spatiale $u \mapsto \mathbf{r}(u)$ par rapport à l'axe de (SP_\star) de la périhélie P_\star : nous supposons $\mathbf{r}(0) = P_\star$, de sorte que $\rho(u)$ décroît de $+\infty$ jusqu'à la valeur minimale $\rho(0) = r_\circ$ (rayon solaire) puis croît de r_\circ à $+\infty$. De même pour $\theta_\star := \theta(0)$, l'angle $\theta(u)$ varie de $\theta(-\infty) = 0$ à $\theta(+\infty) = 2\theta_\star$ (c.f. Fig. 4). Alors, l'angle de déflexion β de la géodésique par rapport à sa direction initiale vérifie

$$(49) \quad \frac{\beta}{2} + \frac{\pi}{2} = \theta_\star = \int_0^{\theta_\star} d\theta$$

L'estimation de θ_\star – et donc de β – sera obtenu par une adaptation de la méthode d'intégration de Binet utilisée dans le cas newtonien (c.f. § 7).

9.4. Estimation de l'angle de déflexion. L'analyse de l'orbite spatiale $u \mapsto \mathbf{r}(u)$ est basée sur l'existence (supposée) du paramètre d'impact (fini, non nul), c'est-à-dire (compte tenu de notre système de repérage) la valeur de la limite b de $\rho \sin \theta = \langle \mathbf{r} | \mathbf{e}_2 \rangle$ lorsque $u \rightarrow -\infty$. Nous pouvons supposer (sans perte de généralité) que $b > 0$ (voir Fig. 4).

Proposition 9.3. Soit b le paramètre d'impact de l'orbite spatiale $u \mapsto \mathbf{r}(u)$ et v_0 la vitesse initiale, i.e. $d\mathbf{r}/du = -v_0\mathbf{e}_1$ lorsque $u \rightarrow -\infty$; alors (i) : $2/c^2\mathcal{E} = v_0^2 - 1$ et (ii) : $\omega_0 = bv_0$.

Preuve. (i) : Soit $(\mathbf{e}_\rho(u), \mathbf{e}_\theta(u))$ la base mobile polaire associée à $\mathbf{r}(u)$, i.e. $\mathbf{e}_\rho := \cos(\theta)\mathbf{e}_1 + \sin(\theta)\mathbf{e}_2$ et $\mathbf{e}_\theta := -\sin(\theta)\mathbf{e}_1 + \cos(\theta)\mathbf{e}_2$. Alors, d'une part, d'après (45) :

$$\frac{d\mathbf{e}_\rho}{du} = \frac{d\theta}{du} \frac{d\mathbf{e}_\rho}{d\theta} = \frac{\omega_0}{\rho^2} \mathbf{e}_\theta \quad \text{et} \quad \frac{d\mathbf{e}_\theta}{d\theta} = \frac{d\theta}{du} \frac{d\mathbf{e}_\theta}{d\theta} = -\frac{\omega_0}{\rho^2} \mathbf{e}_\rho$$

13. En fait cela n'a rien d'évident dans l'absolue : en particulier, si le paramètre d'impact est inférieur ou égal au rayon de Schwarzschild, l'orbite spatiale semble posséder une évolution irréversible : on pourrait imaginer que le rayon lumineux est absorbé par le trou noir (l'inversion du temps ne permettrait pas alors au rayon lumineux de ressortir du trou noir). Le paradoxe – bien que lié à la gravitation – est analogue à celui des jumeaux de Langevin de la relativité restreinte : il disparaît en réalisant que le rayon lumineux met – dans le référentiel du laboratoire – un temps infini pour rentrer dans le trou noir : cela s'explique si on garde à l'esprit que l'orbite spatiale correspond à la projection d'une géodésique inertielle de l'espace-temps.

Mais d'autre part, comme $\mathbf{r} = r\mathbf{e}_\rho$ il vient

$$(50) \quad \frac{d\mathbf{r}}{du} = \frac{d\rho}{du}\mathbf{e}_\rho + \frac{\omega_0}{\rho}\mathbf{e}_\theta$$

Comme $\mathbf{r}(0)$ est le point de la périhélie, l'application $u \mapsto \rho(u)$ est une fonction décroissante pour $u < 0$. Une conséquence de (50) est que $-d\rho/du$ vaut v_0 pour $u = -\infty$, de sorte qu'en partant de (48) et du fait que $B(\rho) \rightarrow 1$ lorsque $\rho \rightarrow +\infty$:

$$2/c^2\mathcal{E} = \lim_{u \rightarrow -\infty} \left(\left(\frac{d\rho}{du} \right)^2 + B \left(\frac{\omega_0}{\rho} \right)^2 - 1 \right) \frac{1}{B} = v_0^2 - 1$$

(ii) : Rappelons que le paramètre d'impact de l'orbite spatiale $u \mapsto \mathbf{r}(u)$ correspond à la limite b de $\langle \mathbf{r} | \mathbf{e}_2 \rangle$, lorsque $u \rightarrow -\infty$; par une deuxième application (50)

$$\mathbf{r} \times \frac{d\mathbf{r}}{du} = \rho\mathbf{e}_\rho \times \left(\frac{d\rho}{du}\mathbf{e}_\rho + \frac{\omega_0}{\rho}\mathbf{e}_\theta \right) = \omega_0\mathbf{e}_3$$

Du fait que $d\mathbf{r}/du = -v_0\mathbf{e}_1$ (resp. $\langle \mathbf{r} | \mathbf{e}_2 \rangle = b$) lorsque $u \rightarrow -\infty$ nous tirons alors

$$\omega_0\mathbf{e}_3 = \lim_{u \rightarrow -\infty} \mathbf{r}(u) \times \frac{d\mathbf{r}}{du} = \lim_{u \rightarrow -\infty} \left(\langle \mathbf{r} | \mathbf{e}_1 \rangle \mathbf{e}_1 + \langle \mathbf{r} | \mathbf{e}_2 \rangle \mathbf{e}_2 \right) \times \frac{d\mathbf{r}}{du} = -bv_0\mathbf{e}_2 \times \mathbf{e}_1 = bv_0\mathbf{e}_3$$

□

Remarque 9.4. L'heuristique de l'identification entre, d'une part la constante r_\circ de la métrique de Schwarzschild donnée en (37) et d'autre part, le rayon de Schwarzschild r_\bullet est classiquement basée sur l'approximation (newtonienne) du champ faible engendré par la matière-énergie concentrée au centre de la symétrie sphérique (c.f. § 8.2). Une autre manière de voir cela est de repartir de la forme initiale de la métrique en (37) et de comparer l'énergie potentielle effective d'orbites (spatiales) analogues – i.e. rayon lumineux (quasi)-hyperboliques ayant le même paramètre d'impact b – dans le cas newtonien et pour la Relativité Générale, lorsqu'on se trouve loin du centre de symétrie. Dans le cas newtonien, $\sigma_0 = cb$ de sorte que l'énergie totale du système donnée en (19) s'écrit

$$E = \frac{1}{2} \left(\frac{d\rho}{dt} \right)^2 + E_r(\rho) \quad \text{où} \quad E_r(\rho) = \frac{c^2}{2} \left(-\frac{r_\bullet}{\rho} + \frac{b^2}{\rho^2} \right)$$

Mais d'autre part, pour que la Relativité Générale (c.f. Remarque 9.2), l'énergie du système est

$$\mathcal{E} = \frac{c^2}{2} \left(\frac{d\rho}{du} \right)^2 \frac{1}{B} + \mathcal{E}_r(\rho)$$

où en utilisant le fait que $\omega_0 = b$ (c.f. (ii)-Proposition 9.3)

$$\begin{aligned} \mathcal{E}_r(\rho) &= \frac{c^2}{2} \left(-1 + \frac{b^2}{\rho^2} \left(1 - \frac{r_\circ}{\rho} \right) \right) \left(1 - \frac{r_\circ}{\rho} \right)^{-1} \\ &= \frac{c^2}{2} \left(-1 + \frac{b^2}{\rho^2} - \frac{b^2 r_\circ}{\rho^3} \right) \left(1 + \frac{r_\circ}{\rho} + \frac{r_\circ^2}{\rho^2} + \dots \right) = \frac{c^2}{2} \left(-1 - \frac{r_\circ}{\rho} + \frac{b^2}{\rho^2} + \dots \right) \end{aligned}$$

Par suite l'identité $r_\circ = r_\bullet$ est la condition nécessaire et suffisante pour que :

$$\mathcal{E}_r(\rho) = -\frac{c^2}{2} + E_r(\rho) + \mathcal{O} \left(\frac{1}{\rho^3} \right)$$

Lemme 9.5. Si $v_0 = 1$ (cas du rayon lumineux) alors (i) : $\mathcal{E} = 0$, de sorte que

$$(51) \quad \left(\frac{1}{\rho^2} \frac{d\rho}{d\theta} \right)^2 = \frac{B(r_\circ)}{r_\circ^2} - \frac{B(\rho)}{\rho^2}$$

(ii) : avec le changement de variable de Binet $x := r_{\odot}/\rho$ nous avons pour $0 \leq \theta \leq \theta_*$

$$(52) \quad d\theta = \left(1 - \frac{r_{\bullet}}{r_{\odot}} \left(x + \frac{1}{1+x}\right)\right)^{-1/2} \frac{dx}{(1-x^2)^{1/2}}$$

(iii) : pour $0 \leq \theta \leq \theta_*$ nous avons $0 \leq x := r_{\odot}/\rho \leq 1$ et en posant $x := \cos(\xi)$ il vient :

$$(53) \quad d\theta = - \left(1 - \frac{r_{\bullet}}{r_{\odot}} \left(\cos(\xi) + \frac{1}{1+\cos(\xi)}\right)\right)^{-1/2} d\xi$$

Preuve. (i) : Partant de (48) avec $\mathcal{E} = 0$ et (45) affirmant que $d\theta/du = \omega_0/\rho^2$ il vient :

$$(48) \iff \left(\frac{d\rho}{d\theta} \frac{d\theta}{du}\right)^2 = 1 - \frac{B\omega_0^2}{\rho^2} \iff \left(\frac{1}{\rho^2} \frac{d\rho}{d\theta}\right)^2 = \frac{1}{\omega_0^2} - \frac{B}{\rho^2}$$

Nous obtenons (51) du fait que $d\rho/d\theta = 0$ pour $\theta = \theta_*$ et $r_{\odot} = \rho(\theta_*)$ donne $1/\omega_0^2 = B(r_{\odot})/r_{\odot}^2$.

(ii) : Nous effectuons un changement de variable du type Binet en posant $x := r_{\odot}/\rho$; alors, avec $B(\rho) = 1 - r/r_{\bullet}$ il vient, grâce à (51)

$$\begin{aligned} \left(\frac{dx}{d\theta}\right)^2 &= \left(\frac{r_{\odot}}{\rho^2} \frac{d\rho}{d\theta}\right)^2 = \left(1 - \frac{r_{\bullet}}{r_{\odot}}\right) - \left(1 - \frac{r_{\bullet}}{r_{\odot}}x\right)x^2 \\ &= (1-x^2) - \frac{r_{\bullet}}{r_{\odot}}(1-x^3) = \left(1 - \frac{r_{\bullet}}{r_{\odot}} \left(\frac{1-x^3}{1-x^2}\right)\right)(1-x^2) \end{aligned}$$

ce qui donne (52) du fait que $x \mapsto \theta(x)$ est une application décroissante pour $\theta_* \leq \theta \leq 2\theta_*$.

(iii) : Le fait que $x = r_{\odot}/\rho$ varie de 0 à 1 lorsque ρ varie de $-\infty$ à r_{\odot} nous autorise à poser $x = \cos(\xi)$ pour ξ variant de $\pi/2$ à 0, ce qui donne (53). □

Proposition 9.6. Pour la Relativité Générale, l'angle β de déflexion d'un rayon de lumière rasant la surface solaire est le double de l'angle de déflexion newtonien i.e. $\beta \simeq 2r_{\bullet}/r_{\odot}$.

Preuve. Rappelons que le rayon de Schwarzschild de la masse solaire M_{\odot} est $r_{\bullet} = GM_{\odot}/c^2$: ce rayon est très petit relativement au rayon solaire r_{\odot} : grâce aux valeurs mesurées des grandeurs en question, nous savons – voir (12) – que $r_{\bullet}/r_{\odot} \simeq 4,24 \times 10^{-6}$. Il est ainsi possible d'approximer (53) en écrivant (pour $0 \leq \xi \leq \pi/2$) :

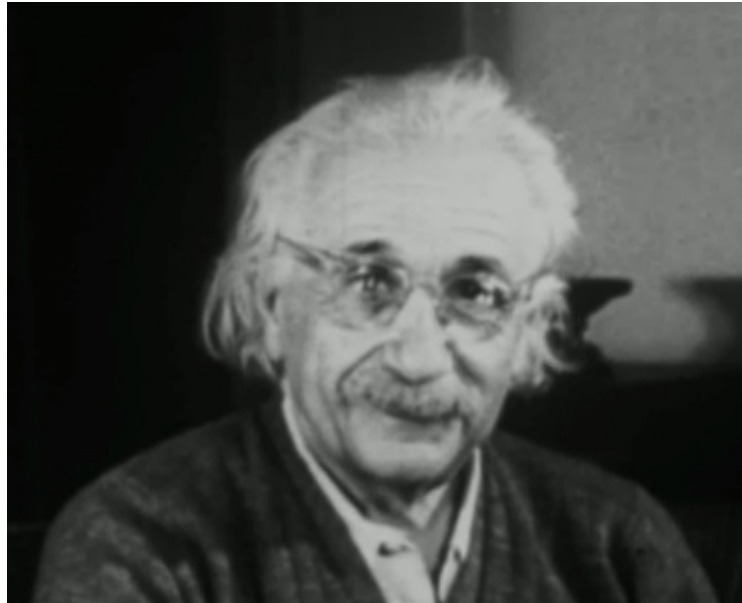
$$d\theta \simeq - \left(1 + \frac{1}{2} \frac{r_{\bullet}}{r_{\odot}} \left(\cos(\xi) + \frac{1}{1+\cos(\xi)}\right)\right) d\xi$$

Par suite, en partant de (49) il vient :

$$\frac{\beta}{2} + \frac{\pi}{2} = \theta_* = \int_0^{\theta_*} d\theta = \int_{\pi/2}^0 \frac{d\theta}{d\xi} d\xi \simeq \left[\xi + \frac{1}{2} \frac{r_{\bullet}}{r_{\odot}} \left(\sin(\xi) + \tan\left(\frac{\xi}{2}\right)\right) \right]_0^{\pi/2} = \frac{\pi}{2} + \frac{r_{\bullet}}{r_{\odot}}$$

□

L'application numérique donne $\beta \simeq 1.75''$.



10. Notes

Au début des années 30, l'étude de la structure des étoiles s'est développée parallèlement aux recherches sur la mécanique interne des noyaux atomiques et des réactions de fusion/fission nucléaires. Il faut noter que l'ensemble des travaux portant sur les corps célestes massifs ont une teneur très spéculative du fait qu'ils se trouvent à la frontière entre la mécanique quantique et la relativité générale – deux théories dont l'incompatibilité est un des défis de la physique théorique actuelle. En 1931 (c.f. [Cha31a, Cha31c, Cha31b] et [Cha35]), Subrahmanyan Chandrasekhar, jeune étudiant indien de 20 ans, en route vers Cambridge, aboutit à la conclusion troublante qu'une naine blanche¹⁴ dont la masse excède une certaine masse limite (la masse de Chandrasekhar) doit – selon ses calculs – s'effondrer sous l'effet de la seule gravitation, jusqu'à atteindre une densité infinie. Au même moment (février-mars 1931) à Copenhague, le physicien soviétique Lev Landau évoque, avec Niels Bohr et Léon Rosenfeld, la possible existence d'étoiles dont la matière serait si comprimée que les noyaux atomiques fusionneraient pour ne former plus qu'un seul noyau (énorme!) de densité comparable à celle des noyaux d'atomes. Landau publie ses résultats en 1932 [Lan32], au moment même de la découverte du neutron par James Chadwick : le savant soviétique est en fait sur la piste d'un nouveau type de corps céleste qu'on appellera bientôt les étoiles à neutrons. Oppenheimer n'est pas convaincu par le modèle de Landau et propose à son étudiant George Volkoff d'étudier un modèle d'un astre à symétrie sphérique. L'analyse d'Oppenheimer et Volkoff publiée en 1939 [OV39] donne une estimation de la masse maximale d'un cœur de neutrons stable, celle-ci s'avérant être de l'ordre de la moitié de la masse de Chandrasekhar. De même que dans le scénario de Chandrasekhar, au-delà de la masse critique d'Oppenheimer-Volkoff, une étoile en fin de vie s'effondrerait inéluctablement sur elle-même pour donner un corps

14. Objet céleste très dense issu de l'effondrement d'une étoile d'une masse inférieure à 10 masses solaires.

céleste très dense dont le champ de gravitation s’approcherait de celui décrit dans le modèle de Schwarzschild. La terminologie de *trou noir* pour décrire ces objets singuliers, à été introduite dans les années 60 par John Wheeler. L’histoire de l’astronomie moderne (post Einsteinienne) est magnifiquement racontée par K. Thorne [Tho01] (élève de Wheeler) ; on pourra aussi trouver de belles histoires dans les ouvrages de vulgarisation de haut niveau de J. P. Luminet [Lum87, Lum14], tout comme dans l’analyse décapante que fait J.P. Petit de la cosmologie moderne dans son petit livre [Pet01]. Pour garder la tête froide, mentionnons la critique des dérives de la sciences (particulièrement en cosmologie) faite par C. Magnan dans son livre [Mag05] ; notons au passage que le site-web ¹⁵ de C. Magnan regorge de belles *leçons* de physique, parmi lesquelles la déviation des rayons lumineux et la précession de la périhélie de Mercure en Relativité Générale.

Les *trous noirs* restent des objets théoriques d’ordre spéculatif (Einstein lui même n’était pas convaincu de leur existence). En 1975 [Haw75] Stefen Hawking utilise des effets quantiques pour montrer comment un trous noir pourrait émettre un rayonnement entraînant son *évaporation* : un tel rayonnement – appelé rayonnement de Hawking – serait tellement faible qu’il serait noyé dans le rayonnement fossile du Big-Bang à 3 Kelvin et donc – pour l’heure – indétectable. Beaucoup de progrès théoriques et observationnels ont été réalisés quant à la recherche de preuves d’existence de trous noirs et il existe aujourd’hui des centaines de candidats potentiel. Un des résultats les plus marquants dans cette direction est celui de Reinhard Genzel, Andrea Ghez et collaborateurs [EG96, GMB99, GDM⁺03] : en étudiant les trajectoires d’une série d’étoiles dans le voisinage (proche) du centre de notre galaxie (sur une quinzaine d’années) ils ont prouvé l’existence d’une masse centrale de l’ordre de 3 millions de masses solaire et n’émettant aucun rayonnement, c’est-à-dire d’un trou noir *super-massif*. Pour leur *découverte* du trou noir central de la Voix Lactée Genzel et Ghez ont reçu le prix Crafoord 2012.

RÉFÉRENCES

- [BG68] R.L. Bishop and S.I. Goldberg. *Tensor Analysis on Manifolds*. Dover Books on Mathematics. Dover Publications, 1968.
- [Bro11] K. Brown. *Reflections on Relativity*. Lulu Enterprises Incorporated, 2011.
- [Car13] S. Carroll. *Spacetime and Geometry : An Introduction to General Relativity*. Pearson New International Edition (second edition of the 2003 Addison-Wesley edition), 2013.
- [Cha31a] S. Chandrasekhar. The density of white dwarf stars. *Philosophical Magazine (7th series)*, 11 :592–596, 1931.
- [Cha31b] S. Chandrasekhar. The highly collapsed configurations of a stellar mass. *Monthly Notices of the Royal Astronomical Society*, 91 :456–466, 1931.
- [Cha31c] S. Chandrasekhar. The maximum mass of ideal white dwarfs. *Astrophysical Journal*, 74 :81–82, 1931.
- [Cha35] S. Chandrasekhar. The highly collapsed configurations of a stellar mass (second paper). *Monthly Notices of the Royal Astronomical Society*, 95 :207–225, 1935.
- [EG96] A. Eckart and R. Genzel. Observations of stellar proper motions near the galactic centre. *Nature*, 383 :415 – 417, 1996.
- [Ein16] A. Einstein. Grundlage der allgemeinen Relativitätstheorie (The Foundation of the General Theory of Relativity). *Annalen der Physik*, 49(IV) :769–822, 1916.
- [Eis02] J. Eisensteadt. *Einstein et la relativité générale*. CNRS Édition - Paris, 2002.

15. <http://www.lacosmo.com/>

- [GDM⁺03] A. M. Ghez, G. Duchêne, K. Matthews, S. D. Hornstein, A. Tanner, J. Larkin, M. Morris, E. E. Becklin, S. Salim, T. Kremenek, D. Thompson, B. T. Soifer, G. Neugebauer, and I. McLean. The first measurement of spectral lines in a short-period star bound to the galaxy's central black hole : A paradox of youth. *The Astrophysical Journal Letters*, 586(2) :L127–L131, 2003.
- [GMB99] A. M. Ghez, M. Morris, and E. E. Becklin. High proper motion stars in the vicinity of sgr a* : Evidence for a supermassive black hole at the center of our galaxy. *The Central Parsecs of the Galaxy, ASP Conference Series – Edited by Heino Falcke, Angela Coter, Wolfgang J. Duschl, Fulvio Melia, and Marcia J. Rieke*, 186 :18–22, 1999.
- [Haw75] S. W. Hawking. Particle creation by black holes (erratum *ibid.*, 46, 206 (1976)). *Communications in Mathematical Physics*, 43 :199–220, 1975.
- [Haw03] S. W. Hawking. *Sur les épaules des géants*. Dunod, 2003.
- [Lan32] L. Landau. On the theory of stars. *Physikalische Zeitschrift der Sowjetunion*, 2 :285–288, 1932.
- [Lum87] J.-P. Luminet. *Les trous noirs*. Pierre Belfond, 1987.
- [Lum14] J.P. Luminet. *L'Invention Du Big Bang*. Collection Points. Série Sciences. Éd. Points, 2014.
- [Mac83] E. Mach. The science of mechanics, Trans. by T. J. McCormack (2nd ed.), Open Course Publishing Co., 1883.
- [Mag05] C. Magnan. *La science pervertie. Épistémologie et philosophie des sciences*. Editions L'Harmattan, 2005.
- [Nak03] M. Nakahara. *Geometry, Topology and Physics, Second Edition*. Graduate student series in physics. Taylor & Francis, 2003.
- [Oli15a] E. Olivier. 100 ans de Relativité Générale I/III : covariance. *Bull. Inf. App. & App.*, 100 :23–63, 2015.
- [Oli15b] E. Olivier. 100 ans de Relativité Générale II/III : courbure. *Bull. Inf. App. & App.*, 101 :97–123, 2015.
- [OV39] J. R. Oppenheimer and G. M. Volkoff. On massive neutron cores. *Physical Review*, 55 no 4 :374–381, 1939.
- [Pet01] J.P. Petit. *On a perdu la moitié de l'univers*. Collection Pluriel. Hachette Littératures, 2001.
- [Sch16] K. Schwarzschild. Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie. *Sitzber. Deut. Akad. Wiss.*, pages 189–196, 1916.
- [Tho01] K.S. Thorne. *Trous noirs et distorsions du temps : l'héritage sulfureux d'Einstein (traduit de l'anglais par A. Bouquet)*. Nouvelle bibliothèque scientifique. Flammarion, 2001.
- [Ton64] M.A. Tonnelat. *Les vérifications expérimentales de la relativité générale*. Masson et Cie, 1964.
- [VM05] T. Shaposhnikova V. Maz'ya. *Jacques Hadamard : Un mathématicien universel (traduction de l'anglais par G. Tronel)*. EDP Science (Ed. originale de l'Am. Math. Soc. - 1998), 2005.
- [Wei72] S. Weinberg. *Gravitation and Cosmology – Principles and Applications of the General Theory of Relativity*. John Wiley, 1972.

11. Postscriptum : Textes choisis

Galileo Galilei – 1632

Dialogue sur les deux grands systèmes du Monde

« Le mouvement est mouvement et agit comme mouvement pour autant qu'il est en rapport avec des choses qui en sont dépourvues ; mais pour toutes les choses qui y participent également, il n'agit pas, il est comme s'il n'était pas. »

Galileo Galilei – 1636

*Discours et démonstrations
concernant deux sciences nouvelles*

« Une vitesse quelconque imprimée à un corps se conserve rigoureusement aussi longtemps que les causes extérieures d'accélération ou de ralentissement sont écartées, condition qui se réalise seulement dans le plan horizontal ; car dans les plans déclinés il existe déjà une cause d'accélération, tandis que dans les plans qui montent il existe une cause de ralentissement. D'où il suit que le mouvement sur le plan horizontal est perpétuel. »

René Descartes

Principes philosophiques – 1644

« Chaque chose, en tant qu'elle est simple et indivisible, demeure, pour autant qu'il est en elle, toujours dans le même état, à moins que des causes extérieures ne la changent. Que tout mouvement est par lui-même droit ; et ainsi, ce qui se meut en cercle, tendra toujours à s'éloigner du centre du cercle qu'il décrit. »

René Descartes

Principes philosophiques – 1644

« Dieu de sa toute puissance a créé la matière avec le mouvement et le repos et conserve maintenant en l'Univers, par son concours ordinaire, autant de mouvement et de repos qu'il y en a mis en le créant. »

Blaise Pascal

Pensées

« Je ne puis pardonner à Descartes ; il aurait bien voulu, dans toute sa philosophie, se pouvoir passer de Dieu ; mais il n'a pas pu s'empêcher de lui faire donner une chiquenaude, pour mettre le monde en mouvement ; après cela, il n'a plus que faire de Dieu. »

Johan Kepler

Première loi (1609) – « Les planètes décrivent autour du soleil des orbites elliptiques dont le soleil est un foyer. »

Deuxième loi (1609) – « Quand une planète parcourt son orbite, le rayon Soleil-planète balaie des aires égales en des intervalles de temps égaux. »

Troisième loi (1618) – « Le carré de la période de révolution est proportionnel au cube de la distance au Soleil. »

Halley et Newton – 1684

Anecdote retranscrite par S. Hawking dans [Haw03]

« Hooke déclara qu'il avait déduit de la loi de Kepler sur les ellipses la preuve que la gravité était une force, mais qu'il ne la dévoilerait à Halley et Wren que lorsqu'il serait prêt à rendre cette information publique. Furieux, Halley se rendit à Cambridge, raconta à Newton les prétentions de Hooke, et proposa la problème suivant : « *Quelle serait la forme de l'orbite d'une planète autour du Soleil, si elle était causée par une force dirigée vers le Soleil, et variant inversement au carré de la distance ?* » La réponse de Newton fût renversante. « *Ce serait une ellipse* » répondit-il immédiatement, et il raconta à Halley qu'il avait résolu le problème quatre ans auparavant, mais qu'il avait égaré la preuve dans son bureau. »

Isaac Newton

Principia Mathematica (1687)

« 1. Tout corps persévère dans son état de repos ou de mouvement rectiligne uniforme, sauf si des forces imprimées le contraignent d'en changer. »

« 2. Le changement du mouvement est proportionnel à la force exercée sur lui, et orienté dans la direction de la ligne droite sur laquelle cette force s'exerce. »

« 3. À chaque action est toujours opposée une réaction égale ; autrement dit, les actions mutuelles de deux corps l'un sur l'autre sont toujours égales, et orientées en sens contraire. »

Isaac Newton

Principia Mathematica (1687) – Trad. EO

« À cause de l'analogie entre la propagation des rayons de lumière et le mouvement des corps, j'ai pensé qu'il ne serait pas inutile d'ajouter la proposition suivante pour le traitement de l'optique ; non pas dans le but de considérer la nature de la lumière, ou encore d'essayer de déterminer sa possible composition corpusculaire, mais simplement pour noter que les trajectoires des corps sont très similaires aux trajectoires des rayons lumineux. »

Gottfried Leibnitz

« Un corps n'est jamais mû naturellement que par un autre corps qui le pousse en le touchant. Toute autre opération sur les corps est ou miraculeuse ou imaginaire. »

Pierre Louis de Maupertuis -1744*(Principe de la moindre quantité d'action pour la mécanique)*

« L'Action est proportionnelle au produit de la masse par la vitesse et par l'espace. Maintenant, voici ce principe, si sage, si digne de l'Être suprême : lorsqu'il arrive quelque changement dans la Nature, la quantité d'Action employée pour ce changement est toujours la plus petite qu'il soit possible. »

Pierre Simon de Laplace - 1796*(Exposition du système du monde)*

« Il existe donc dans les espaces célestes des corps obscurs aussi considérables, et peut-être en aussi grand nombre que les étoiles. Un astre lumineux de même densité que la Terre, dont le diamètre serait deux cent cinquante fois plus grand que celui du Soleil, ne laisserait en vertu de son attraction parvenir aucun de ses rayons jusqu'à nous ; il est donc possible que les plus grands corps lumineux de l'univers soient par cela même invisibles. »

Ernst Mach - 1883*(extrait de [Mac83] : tiré de [Wei72, p. 16])*

« Newton's experiment with the rotating vessel of water simply informs us, that the relative rotation of water with respect to the sides of the vessel produces no noticeable centrifugal forces, but that such forces are produced by its relative motion with respect to the mass of the Earth and the other celestial bodies. No one is competent to say how the experiment would turn out if the sides of the vessel increased in thickness and mass, until they were several leagues thick. »

Albert Einstein - 1912*(Lettre à Sommerfeld - extrait : tiré de [Eis02])*

« Je travaille maintenant exclusivement sur la gravitation et je pense que je peux surmonter toutes les difficultés avec l'aide d'un de mes amis mathématicien^(*). Mais une chose est certaine : de toute ma vie je n'ai presque jamais travaillé aussi durement et je suis désormais pénétré d'un grand respect pour les mathématiques que j'avais jusqu'alors, dans ma simplicité d'esprit, considéré la partie la plus subtile comme un luxe. Comparé à ce problème, la théorie originale de la relativité est un jeu d'enfant. » ^(*) Marcel Grossmann.

Mikio Nakahara*(À propos de l'équation d'Einstein [Nak03] – Trad. EO)*

« Historiquement, quand Einstein proposa la première formulation de la relativité générale, il écrivit d'abord la proportionnalité du tenseur de Ricci et du tenseur énergie-impulsion. Plus tard, il réalisa que la divergence du tenseur énergie impulsion était nulle, ce qui n'était pas le cas du tenseur de Ricci. »

David Hilbert*(À propos de la contribution d'Albert Einstein en relativité générale – tiré de [Tho01, p. 121])*

« N'importe quel enfant dans les rues de Göttingen en connaît plus qu'Einstein sur la géométrie à quatre dimensions. Malgré tout, c'est Einstein qui fit le travail [Formulation des lois de la gravité en Relativité Générale] et non les mathématiciens. »

Bishop & Goldberg – 1968*(Tiré de [BG68] – Trad. EO)*

« Les prémices de l'analyse tensorielle remontent à Lagrange avec sa présentation générale des systèmes dynamiques, puis à Riemann qui est le premier à considérer la géométrie en dimension quelconque. Le travail de Riemann en 1854, était exprimé dans une forme si obscure, que Beltrami en 1869 et Lipschitz en 1872 n'ont utilisé le langage géométrique qu'avec une grande précaution. En fait, les développements étaient si lents que la notion de parallélisme due à Levi-Civita n'est apparue qu'en 1917. »

Jacques Hadamard - 1922*(Tiré de [VM05])*

« À mon avis il y avait une crise en géométrie infinitésimale et l'avènement de la relativité venait de la surmonter. Ceci était le début d'un éclairage inattendu sur les précédents travaux en géométrie montrant l'importance du calcul différentiel absolue de Ricci et de Levi-Civita, ce calcul étant resté presque sans éveiller d'intérêt depuis sa création. Quelle que soit l'opinion qu'on puisse avoir sur le bien fondé des nouvelles hypothèses, elles donnent à la géométrie un rôle nouveau et lui ouvre une période nouvelle, non une brève nouveauté qui trop souvent influence le mathématicien rivé à ses propres préoccupations, mais une nouveauté infiniment féconde qui jaillit sur la nature des choses. »

Albert Einstein - 1919*(Conceptions scientifiques : trad. M. Solovine)*

« Le théorie évite tous les défauts internes que nous avons reprochés au fondements de la mécanique classique. Elle est suffisante, autant que nous sachions, pour la représentation des faits observés de la mécanique céleste. Mais elle ressemble à un édifice dont une aile est bâtie de marbre fin (premier membre de l'équation) et l'autre de bois de qualité inférieure (second membre de l'équation). »

Arthur Eddington - 1919*(Cité par Chandrasekhar : tiré de [Eis02])*

« Il pourrait n'y avoir aucune déflexion, ce qui signifierait que la lumière n'est aucunement sujette à la gravitation. Ou une *demi-déflexion*, ce qui signifierait que la lumière est sujette à la gravitation, ainsi que

Newton l'a suggéré et qu'elle suit la simple loi newtonienne. Ou, il pourrait y avoir une pleine déflexion, confirmant la loi d'Einstein au lieu de celle de Newton. »

Albert Einstein - 1921

(Article dans *Nature* - extrait : tiré de [Eis02])

« Les champs électriques et gravitationnels ont-ils des caractéristiques réellement si différentes qu'il n'y ait aucune unité formelle à laquelle il puissent être réduits ? Les champs gravitationnels jouent-ils un rôle dans la constitution de la matière et, à l'intérieur du noyau atomique, le continuum doit-il être de façon appréciable considéré comme non euclidien ? »

Albert Einstein - 1939

(À propos de la solution de Schwarzschild tiré de [Tho01, p. 123])

« Le principal résultat de cette étude [écrivait Einstein dans un article technique de 1939] est une compréhension claire de la raison pour laquelle les "singularités de Schwarzschild" n'existent pas dans la réalité physique. »

Albert Einstein

(*Conceptions scientifiques* : trad. M. Solovine)

« La conception esquissée plus haut [modification de la solution de Schwarzschild] correspond à priori, à la structure atomique de la matière, parce que le « pont^(*) » est par sa nature un élément discret. Nous voyons de plus que la constante m de la masse de la particule neutre doit nécessairement être positive, puisque aucune solution sans singularité ne peut correspondre à la solution de Schwarzschild pour une valeur négative de m . Seul l'examen du problème à plusieurs ponts peut montrer si, oui ou non, cette méthode théorique fournit une explication de l'égalité, empiriquement démontrée, des masses de particules qu'on trouve dans la nature, et si elle rend compte des faits que la mécanique quantique a si merveilleusement saisis.

[...] Elle [la théorie des ponts] constitue néanmoins – c'est un fait certain – la première tentative d'élaborer d'une manière cohérente une théorie du champ qui offre une possibilité pour expliquer les propriétés de la matière. En faveur de cette tentative, il faut aussi ajouter qu'elle est basée sur les équations du champ relativistes les plus simples connues aujourd'hui. »

(*) dit aussi *Pont d'Einstein-Rosen* ou encore appelé *trou de ver* (J. Wheeler).

Subrahmanyan Chandrasekhar

(Extrait de la conférence Nobel - 1983)

« A black hole partitions the three-dimensional space into two regions : an inner region which is bounded by a smooth two-dimensional surface called the event horizon; and an outer region, external to the

event horizon, which is asymptotically flat; and it is required (as a part of the definition) that no point in the inner region can communicate with any point of the outer region. This incommunicability is guaranteed by the impossibility of any light signal, originating in the inner region, crossing the event horizon. The requirement of asymptotic flatness of the outer region is equivalent to the requirement that the black hole is isolated in space and that far from the event horizon the space-time approaches the customary space-time of terrestrial physics.(...)

Black holes are macroscopic objects with masses varying from a few solar masses to millions of solar masses. To the extent they may be considered as stationary and isolated, to that extent, they are all, every single one of them, described exactly by the Kerr solution. This is the only instance we have of an exact description of a macroscopic object. Macroscopic objects, as we see them all around us, are governed by a variety of forces, derived from a variety of approximations to a variety of physical theories. In contrast, the only elements in the construction of black holes are our basic concepts of space and time. They are, thus, almost by definition, the most perfect macroscopic objects there are in the universe. And since the general theory of relativity provides a single unique two-parameter family of solutions for their descriptions, they are the simplest objects as well. »

Steven Weinberg - 1971

(Extrait de la préface de [Wei72])

« In learning general relativity, and then teaching it to classes at Berkeley and M.I.T., I became dissatisfied with what seemed to be the usual approach to the subject. I found that in most textbooks geometric ideas were given a starring role, so that a student who asked why the gravitational field is represented a metric tensor or why freely falling particles move on geodesics, or why the field equations are generally covariant, would come away with in impression that this had some thing to do with the fact that space-time is Riemannian manifold.

Of course, this *was* Einstein's point of view, and his preeminent genius necessarily shapes our understanding of the theory he created. However, I believe that the geometrical approach has driven a wedge between general relativity and the theory of elementary particles. As long long as it could be hopped, as Einstein did hope, that matter would eventually be understood in geometrical terms, it made sense to give a Riemannian geometry a primary role in describing the theory of gravitation. But the passage of time has taught us not to expect that the strong, weak, and electromagnetic interactions can be understood in geometrical terms, and to great emphasis on geometry can only obscure the deep connection between gravitation and the rest of physics. »

Carlo Rovelli

(Le Monde : Hors série juillet 2010)

« Aujourd’hui nous appelons photons les quanta de lumière introduits par Einstein en 1905. Ils forment la base conceptuelle de la mécanique quantique, la

théorie qui décrit le mieux, actuellement, la dynamique de la matière et de la lumière. La relativité générale trouve ses limites à chaque fois que les phénomènes quantiques deviennent assez importants pour interférer avec la courbure de l’espace : le centre d’un trou noir et le Big Bang sont deux exemples de ces situations. »



VOUZZAVEDIBISAR : Videac. Télétruc. Janvier 1982 (par Jean-Pierre PETIT)



B.I.A.A.

BULLETIN D'INFORMATIQUE APPROFONDIE ET APPLICATIONS

Revue fondée par Edmond Bianco

Publication trimestrielle de l'Université d'Aix-Marseille

ISSN 0291-5413

Le bulletin d'informatique approfondie et applications est une revue pluridisciplinaire destinée à éclairer les connaissances fondamentales informatiques. Les fondements sont un domaine vaste allant de la structure intérieure de l'ordinateur, où se matérialise la machine universelle, à l'algorithme qui devient programme, pour aboutir à la notion de système. Nous contribuons ainsi à ce que les autres disciplines plus anciennes (sciences humaines et de la société, sciences de la matière et de l'énergie, sciences mathématiques, sciences de la nature, sciences de la terre, sciences de l'univers, sciences de la vie, etc.) n'aient pas tendance à considérer l'informatique comme un simple outil définitivement figé. Il importe de continuer à maîtriser les développements fondamentaux de l'informatique pour que nos disciplines puissent en tirer un meilleur parti.

Notre publication est ouverte à l'ensemble de la communauté scientifique. Le périodique est diffusé vers les bibliothèques universitaires de France et vers quelques bibliothèques des cinq continents.

COMITÉ SCIENTIFIQUE

Pr. Patrick Abellard (Université du Sud, Toulon)
Françoise Adreit (Université de Toulouse I)
France Chappaz (Université de Provence)
Georges Chappaz (Université d'Aix-Marseille)
M'hamed Charifi (Consultant autonome)
Jean - Paul Coste (Université de Provence)
Pr. Roger Cusin (Université de la Méditerranée)
Jean - Claude Fumanal (Université Paul Cézanne)
Alain de Gantès (Université d'Aix-Marseille)
Jean Gonella (Université d'Aix-Marseille)
Pr. Bernard Goossens (Université de Perpignan)
Sami Hilala (Université d'Aix-Marseille)
Patrick Isoardi (Université d'Avignon et Pays de Vaucluse)
Robert Jacquier (Université Paul Cézanne)
Jean - Michel Knippel (Université d'Aix-Marseille)
Jean - Philippe Lehmann (Université d'Avignon et Pays de Vaucluse)
Pr. Agathe Merceron (Technische Fachhochschule, Berlin)
Nadia Mesli (Université d'Aix-Marseille)
Eric Olivier (Université d'Aix-Marseille)
Patrick Sanchez (C.N.R.S., Marseille)
Rolland Stutzmann (I.U.T. de Strasbourg Sud)
Pr. André Tricot (E.S.P.E., Toulouse)

CORRESPONDANT(E)S

Pr. Mohamed Tayeb Laskri (Université Badji Mokhtar, Afrique)
Sylvie Monjal (Cégep de Sainte Foy, anciennement Académie de Québec, Amériques)
Moussa HadjAli (Université Virtuelle de Syrie, Asie)
José Rouillard (Université des Sciences et Technologies de Lille, Europe)
Kalina Yacef (Université de Sydney, Océanie)