

Mais où est donc passée la relativité générale ?

Version 3.0

Pascal Picard \*

29 mars 2015

---

\* Je suis amateur de Mathématiques et de Physique Théorique, convaincu que ces sciences sont accessibles à tous, à condition de les expliquer progressivement et de les introduire par les prérequis nécessaires. Depuis quelques années, je m'attelle à écrire des textes théoriques sous forme de pièces de théâtre. Trois personnages y bavardent : Béatrix est la Candide, c'est elle qui pose les questions ; Euristide est un peu philosophe, un peu physicien, il est l'intuitif du groupe ; Mathine est notre mathématicienne, c'est elle qui présente les calculs et les démonstrations. Ces textes sont mis à disposition gratuitement sur Internet. Le prérequis pour la lecture des documents est le programme de Terminale S en France.

à Pascale, Marine et Morgane

Tous mes chaleureux remerciements :

- au site <http://www.sciences.ch> qui a permis la publication de ce document
- à mes précieux relecteurs qui ont courageusement lu et annoté de nombreuses pages : Olivier, Christian et Mina
- à quelques lecteurs anonymes sur Internet qui m'ont signalé des erreurs

---

## Table des matières

<b>I</b>	<b>Acte 0 - Introduction</b>	<b>14</b>
<b>II</b>	<b>Acte I - Les notions fondamentales</b>	<b>15</b>
1	Scène I-1 : La géométrie	15
2	Scène I-2 : La notion de fonction	32
3	Scène I-3 : La notion de limite d'une fonction	34
4	Scène I-4 : Le concept de voisinage	36
5	Scène I-5 : La dérivée	39
6	Scène I-6 : La dérivée partielle	44
7	Scène I-7 : L'intégration	46
8	Scène I-8 : Les équations différentielles	48
9	Scène I-9 : Les vecteurs	53
10	Scène I-10 : Le calcul matriciel	60
<b>III</b>	<b>Acte II - La cinématique</b>	<b>63</b>
11	Scène II-1 : Les mouvements des corps	63
12	Scène II-2 : Le mouvement rectiligne du point	63
13	Scène II-3 : La définition de la vitesse	64
14	Scène II-4 : La relation entre vitesse et distance	65
15	Scène II-5 : La définition de l'accélération	66
16	Scène II-6 : Le mouvement d'un point	68
17	Scène II-7 : La vitesse d'un point	70
18	Scène II-8 : L'accélération d'un point	72
<b>IV</b>	<b>Acte III - La mécanique newtonienne</b>	<b>75</b>
19	Scène III-1 : Les interactions des corps	75
20	Scène III-2 : La première loi de Newton	77
21	Scène III-3 : La seconde loi de Newton	78

---

22 Scène III-4 : La troisième loi de Newton	79
23 Scène III-5 : Le mouvement résultant des forces	80
24 Scène III-6 : Les mouvements liés d'un corps	83
25 Scène III-7 : La quantité de mouvement	84
26 Scène III-8 : La transmission de l'élan	86
27 Scène III-9 : L'impulsion	87
28 Scène III-10 : Les corps de masse variable	87
V Le travail et l'énergie	90
29 Scène IV-1 : La notion d'énergie	90
30 Scène IV-2 : Le travail	90
31 Scène IV-3 : Le travail d'une force	90
32 Scène IV-4 : L'énergie potentielle de déformation	92
33 Scène IV-5 : L'énergie cinétique	93
34 Scène IV-6 : L'énergie potentielle	95
35 Scène IV-7 : La conservation de l'énergie	96
VI Acte V - La notion de relativité du mouvement	97
36 Scène V-1 : Les systèmes de référence galiléens	97
37 Scène V-2 : Le référentiel non galiléen	97
38 Scène V-3 : Le référentiel en rotation	99
VII Acte VI - Le mouvement du corps solide	101
39 Scène VI-1 : La translation et la rotation	101
40 Scène VI-2 : L'équilibre sur un axe fixe	105
41 Scène VI-3 : La dynamique autour un axe fixe	106
42 Scène VI-4 : Le moment de la quantité de mouvement	109
43 Scène VI-5 : L'énergie cinétique en rotation	109
44 Scène VI-6 : Le mouvement du centre d'inertie	110

---

45 Scène VI-7 : Le mouvement des planètes	110
<b>VIII Acte VII - Le principe variationnel</b>	<b>113</b>
46 Scène VII-1 : Les variations	113
47 Scène VII-2 : Les extrémales	115
48 Scène VII-3 : L'équation d'Euler-Lagrange	118
49 Scène VII-4 : L'équation de Lagrange	120
50 Scène VII-5 : Le principe de moindre action	120
<b>IX Acte VIII - Le calcul tensoriel</b>	<b>122</b>
51 Scène VIII-1 : L'écriture indicielle	122
52 Scène VIII-2 : Le changement de coordonnées	125
53 Scène VIII-3 : La notion de variance	132
54 Scène VIII-4 : Introduction du tenseur	133
55 Scène VIII-5 : La symétrie	137
56 Scène VIII-6 : La métrique de l'espace euclidien	138
57 Scène VIII-7 : Les composantes covariantes	139
<b>X Acte IX - La relativité restreinte</b>	<b>141</b>
58 Scène IX-1 : La vitesse des interactions	141
59 Scène IX-2 : La notion d'intervalle	143
60 Scène IX-3 : Le temps propre	149
61 Scène IX-4 : La transformation de Lorentz	152
62 Scène IX-5 : La transformation de la vitesse	157
63 Scène IX-6 : La notation tensorielle	160
64 Scène IX-7 : La mécanique relativiste	163
<b>XI Acte X - Les variétés</b>	<b>168</b>
65 Scène X-1 : La variété différentiable	168
66 Scène X-2 : Notion d'espace tangent	176

---

67 Scène X-3 : Notion d'espace cotangent	179
68 Scène X-4 : Le tenseur métrique	180
69 Scène X-5 : Le pullback	188
70 Scène X-6 : Le pushforward	189
71 Scène X-7 : Les courbes intégrales	195
72 Scène X-8 : Le champ vectoriel de Killing	198
<b>XII Acte XI - La courbure</b>	<b>200</b>
73 Scène XI-1 : La notion de courbure	200
74 Scène XI-2 : La dérivée covariante	202
75 Scène XI-3 : Le transport parallèle	211
76 Scène XI-4 : Les géodésiques	214
77 Scène XI-5 : Le tenseur de Riemann	218
78 Scène XI-6 : Courbure intrinsèque et extrinsèque	224
<b>XIII Acte XII - La gravitation</b>	<b>226</b>
79 Scène XII-1 : Le principe d'équivalence	226
80 Scène XII-2 : Les équations du champ d'Einstein	232
<b>XIV Acte XIII - Les applications de la relativité générale</b>	<b>243</b>
81 Scène XIII-1 : Le champ de gravitation faible	243
82 Scène XIII-2 : Les ondes gravitationnelles	248
83 Scène XIII-3 : Le champ à symétrie centrale	259
84 Scène XIII-4 : Les trous noirs	280
<b>XV Acte XIV - La cosmologie relativiste</b>	<b>282</b>
85 Scène XIV-1 : Les principes cosmologiques	282
86 Scène XIV-2 : Modèles Friedmann-Robertson-Walker	282
87 Scène XIV-3 : Cosmologies homogènes et isotropes	285
88 Scène XIV-4 : La loi de Hubble	290

## Table des figures

Fig. 1 - La méthode	14
Fig. 2 - Décomposition d'un déplacement suivant les axes N et E	16
Fig. 3 - Le principe du système de coordonnées	17
Fig. 4 - Système d'axes de coordonnées	18
Fig. 5 - Système d'axes de coordonnées oblique	18
Fig. 6 - Système de coordonnées en 3 dimensions	19
Fig. 7 - Distance entre 2 points	20
Fig. 8 - Triangle rectangle	21
Fig. 9 - Théorème de Pythagore	22
Fig. 10 - Système de coordonnées sur une surface plate	23
Fig. 11 - Système de coordonnées d'une surface curviligne	24
Fig. 12 - Trajectoire à angles droits carrée dans un plan	25
Fig. 13 - Trajectoire à angles droits sur un triangle sphérique	25
Fig. 14 - Direction d'une voiture dans un virage	26
Fig. 15 - Les directions tangentes au virage	27
Fig. 16 - Le plan tangent d'une surface en un point	28
Fig. 17 - Les coordonnées du point $P$	29
Fig. 18 - Le changement de système de coordonnées	29
Fig. 19 - Les coordonnées polaires en 2 dimensions	31
Fig. 20 - Les coordonnées polaires en 3 dimensions	32
Fig. 21 - Exemple de fonction à deux variables	34
Fig. 22 - $f(x)$ tend vers $+\infty$ quand $x$ tend vers 0	35
Fig. 23 - $f(x)$ tend vers 0 quand $x$ tend vers $+\infty$	36
Fig. 24 - Analyse d'une fonction au voisinage de 0	37
Fig. 25 - Deux points au voisinage de $c$	41
Fig. 26 - Vitesse lorsque $x$ tend vers $c$	42

---

Fig. 27 - Vitesse lorsque $x_1$ tend vers $c$	42
Fig. 28 - Représentation de la colline	45
Fig. 29 - Variation d'une fonction lorsqu'on fixe une coordonnée	46
Fig. 30 - Interprétation de $v(x)dx$	47
Fig. 31 - Analyse de la vitesse à l'instant $t$	50
Fig. 32 - Analyse de la vitesse à l'instant $t$	51
Fig. 33 - Découpage de la surface en petits rectangles	52
Fig. 34 - Le vecteur AB	54
Fig. 35 - Un même vecteur en différents points du plan	55
Fig. 36 - Les vecteurs de coordonnées en 2 dimensions	56
Fig. 37 - Les vecteurs de coordonnées en 3 dimensions	56
Fig. 38 - Les vecteurs $u$ et $v$ dans le plan	57
Fig. 39 - La somme de deux vecteurs	58
Fig. 40 - Illustration de deux vecteurs opposés	59
Fig. 41 - La multiplication d'un vecteur par un nombre	59
Fig. 42 - Schéma du produit de deux matrices	62
Fig. 43 - Représentation de la trajectoire d'un point	64
Fig. 44 - La vitesse comme la pente d'une droite	65
Fig. 45 - La distance est l'intégrale de la vitesse	66
Fig. 46 - Représentation de la vitesse	67
Fig. 47 - Représentation de l'accélération	68
Fig. 48 - Trajectoire d'un point dans l'espace	69
Fig. 49 - Le vecteur position	69
Fig. 50 - Interprétation de la différentielle du vecteur position	70
Fig. 51 - La vitesse est un vecteur tangent à la trajectoire	71
Fig. 52 - La vitesse d'une voiture est tangente à la route	71
Fig. 53 - Représentation de la différentielle de la vitesse	73
Fig. 54 - Évolution du vecteur accélération	74

---



---

Fig. 55 - Illustration des forces	76
Fig. 56 - L'équilibre d'un corps posé	77
Fig. 57 - L'action et la réaction	79
Fig. 58 - La trajectoire du ballon	82
Fig. 59 - Mouvement d'une bille sur un plan incliné	83
Fig. 60 - Equilibre d'un système de forces	85
Fig. 61 - Le mouvement d'une fusée	88
Fig. 62 - Le produit scalaire de deux vecteurs	91
Fig. 63 - Le travail d'une force le long d'un trajet	92
Fig. 64 - L'énergie potentielle de déformation	93
Fig. 65 - La force d'inertie	97
Fig. 66 - L'équilibre des forces dans un référentiel en mouvement	98
Fig. 67 - Le référentiel en rotation	99
Fig. 68 - La translation	101
Fig. 69 - La rotation	102
Fig. 70 - Vitesse angulaire, vitesse linéaire	102
Fig. 71 - Le produit vectoriel	103
Fig. 72 - Rotation d'un rayon vecteur	104
Fig. 73 - Le bras de levier	105
Fig. 74 - Les moments de forces par rapport à l'axe de rotation	106
Fig. 75 - La deuxième loi de Kepler	111
Fig. 76 - Analyse de la longueur d'une courbe	113
Fig. 77 - La droite comme chemin le plus court	116
Fig. 78 - Coordonnées cartésiennes orthogonales	125
Fig. 79 - Coordonnées obliques	126
Fig. 80 - Coordonnées polaires	126
Fig. 81 - Coordonnées par réseau de courbes	127
Fig. 82 - Coordonnées de Gauss	128

---

Fig. 83 - Méridiens et parallèles sur la Terre	129
Fig. 84 - Champs de vecteurs de coordonnées	130
Fig. 85 - Les composantes covariantes d'un vecteur	140
Fig. 86 - Mouvement rectiligne uniforme, droite d'univers	143
Fig. 87 - Mouvement rectiligne non uniforme	144
Fig. 88 - Référentiels $G$ et $G'$	145
Fig. 89 - Le cône de lumière	147
Fig. 90 - L'horizon des événements	148
Fig. 91 - L'aberration de la lumière	158
Fig. 92 - Ligne d'univers de l'horloge au repos	164
Fig. 93 - Intérieur d'un disque	170
Fig. 94 - Intérieur d'une courbe close	171
Fig. 95 - Carte non injective	171
Fig. 96 - Carte d'un sous-ensemble $U$	172
Fig. 97 - Atlas d'un ensemble $M$	173
Fig. 98 - La projection de Mercator	174
Fig. 99 - Dérivées partielles sur une surface curviligne	175
Fig. 100 - Les vecteurs de coordonnées varient en fonction du point	176
Fig. 101 - Plan tangent d'une sphère en un point	177
Fig. 102 - Analyse des variations de $f$ en un point	178
Fig. 103 - Coordonnées normales riemaniennes	185
Fig. 104 - Le pullback	188
Fig. 105 - Transport par pullback d'une fonction sur l'ellipsoïde	189
Fig. 106 - Le pushforward	189
Fig. 107 - Transport d'un champ de vecteurs par pushforward	190
Fig. 108 - Transport d'un tenseur covariant par pullback	191
Fig. 109 - Transport d'un tenseur contravariant par pushforward	192
Fig. 110 - Comparaison de vecteurs sur une sphère	195

---

Fig. 111 - Difféomorphisme sur une sphère	196
Fig. 112 - Champ vectoriel sur une sphère	197
Fig. 113 - La dérivée de Lie	198
Fig. 114 - La sphère est un espace curviligne	201
Fig. 115 - Différentielle d'un vecteur dans un espace plat	203
Fig. 116 - Différentielle d'un vecteur dans un espace curviligne	204
Fig. 117 - Différentielle d'un vecteur a	207
Fig. 118 - Espace curviligne avec torsion	210
Fig. 119 - Transport parallèle le long d'une courbe	211
Fig. 120 - Transport parallèle suivant 2 courbes	212
Fig. 121 - Transport parallèle sur une sphère	212
Fig. 122 - Courbe paramétrée	213
Fig. 123 - Transport parallèle le long de l'équateur	214
Fig. 124 - Courbe géodésique, courbe non géodésique	215
Fig. 125 - Géodésique d'une sphère : arc de grand cercle	218
Fig. 126 - Transport parallèle le long d'une courbe plate fermée	219
Fig. 127 - Transport parallèle sur une courbe curviligne fermée	219
Fig. 128 - Modèle de la courbe fermée au moyen de 2 vecteurs	220
Fig. 129 - Transport du vecteur u	221
Fig. 130 - La courbure négative	225
Fig. 131 - Propagation du son d'une voiture	229
Fig. 132 - L'effet Doppler quand la voiture se rapproche	229
Fig. 133 - L'effet Doppler quand la voiture s'éloigne	230
Fig. 134 - Modèle de 2 lampes émettrices	231
Fig. 135 - Représentation du gradient d'un champ	233
Fig. 136 - La déviation géodésique	250
Fig. 137 - Mode de polarisation en + de l'onde de gravitation	252
Fig. 138 - Mode de polarisation en × de l'onde de gravitation	253

---

Fig. 139 - Modèle de l'étoile binaire	256
Fig. 140 - Les 3 champs de vecteurs de Killing d'un champ à symétrie sphérique	261
Fig. 141 - Relation d'orthogonalité des vecteurs de Killing	262
Fig. 142 - Foliation de l'espace	262
Fig. 143 - Coordonnées polaires sur la sphère	263
Fig. 144 - Une étoile selon Schwarzschild	271
Fig. 145 - Déflexion de la lumière au voisinage du Soleil	277
Fig. 146 - Le périhélie	278
Fig. 147 - La foliation de l'espace-temps	283
Fig. 148 - Le Big Bang	288
Fig. 149 - Emission de deux signaux lumineux	291

---

**Résumé**

**BEATRIX** : La théorie de la relativité... Quelle grandeur, quelle beauté, quelle pureté!

**EURISTIDE** : C'est en effet une jolie théorie; partant de principes simples et intuitifs, elle permet de construire un solide modèle du monde macroscopique, de l'infiniment grand de notre univers, des phénomènes étonnants des corps lancés à grande vitesse et des phénomènes à peine imaginables des corps extrêmement massifs. Cette théorie n'est pas la théorie du quotidien de l'homme, mais elle est toutefois utilisée couramment, très couramment, sans qu'on le sache toujours : les GPS ont besoin des calculs de la relativité restreinte et de la relativité générale pour parvenir à la précision de l'ordre du mètre qui est requise pour diriger une automobile sur un parcours routier.

**MATHINE** : C'est en effet une théorie actuelle, et le GPS est la preuve tangible de son actualité. Elle ne parvient toutefois pas à expliquer le monde infiniment petit des atomes et des particules de la physique quantique. C'est un des grands mystères qui occupe encore aujourd'hui nos physiciens.

## Première partie

# Acte 0 - Introduction

La scène se joue dans une petite ville de province. Trois amis sont réunis chez l'un d'entre eux. C'est le soir. La soirée leur appartient. Ils ont l'intention de parler tard dans la nuit. De quoi parleront-ils, ce soir ? De physique. et plus précisément, de la relativité générale.

Béatrix est la plus jeune d'entre eux. Elle est curieuse de tout. Elle veut comprendre la théorie de la relativité générale. Elle sera notre candide. C'est une bonne occasion pour elle ce soir, parce qu'elle est entourée de deux experts. Euristide est physicien, et s'attachera à interpréter les lois physiques pour qu'elles frappent notre intuition. Mathine, elle, est mathématicienne, et connaît sur le bout des doigts le support mathématique de la théorie de la relativité.

Tout au long de cette conversation, nos amis oscilleront entre théorie et réalité. Ils expliqueront la réalité par la théorie, et donneront ainsi un sens à la réalité; ils confirmeront la théorie par la réalité, et donneront ainsi un sens à la théorie. Cette approche nous sera de grande utilité pour comprendre au fond la raison d'être des théories qu'ils vont nous exposer ici.

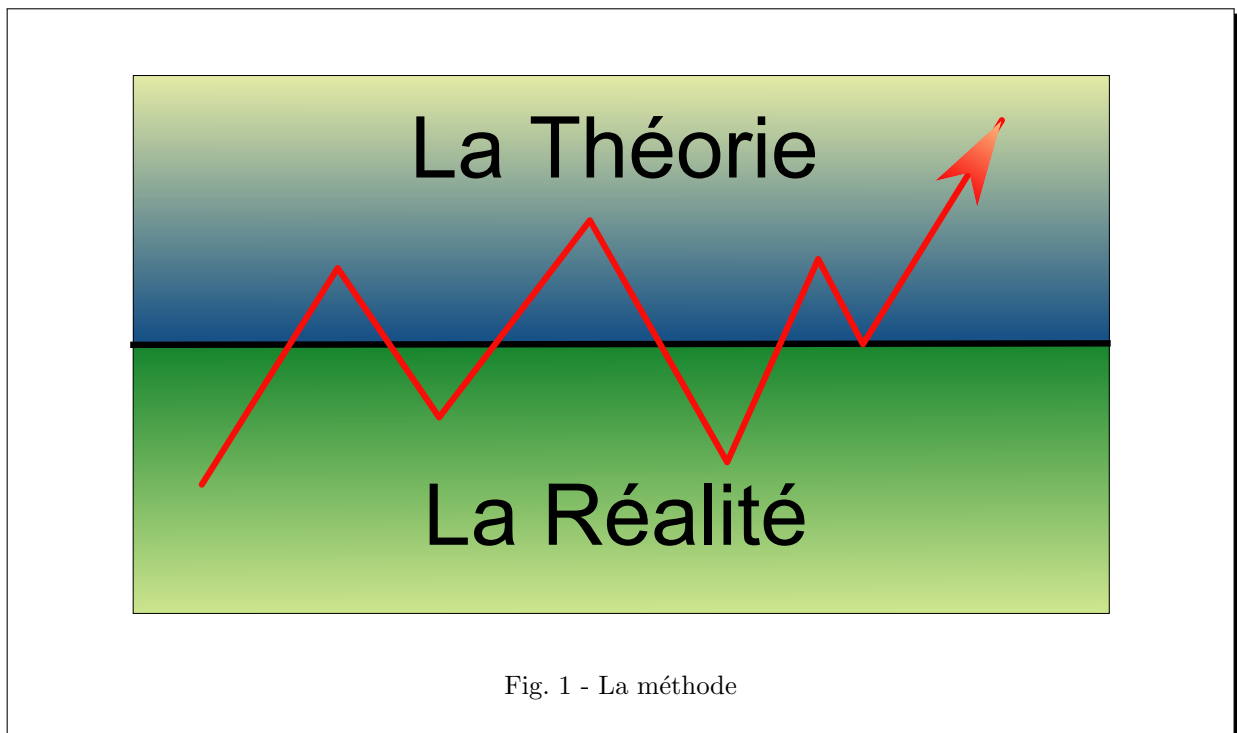


Fig. 1 - La méthode

---

## Deuxième partie

# Acte I - Les notions fondamentales

## 1 Scène I-1 : La géométrie

BEATRIX : Aujourd'hui, je voudrais tout savoir sur la relativité générale! Croyez-vous que cela soit possible?

MATHINE : Hum... Pourquoi pas. Mais il va falloir introduire au préalable de nombreuses notions mathématiques : le calcul des variations, les tenseurs, les variétés, la courbure, les théories de jauge, et j'en passe... Sans compter qu'il faudra aussi vérifier que tu as bien assimilé les autres bases mathématiques qui sont indispensables. Le plus simple est que je t'explique à nouveau ces bases moi-même pour m'assurer que tu les as bien comprises.

EURISTIDE : Oui, je suis d'accord. Et ce sera l'occasion de t'imprégner des concepts dont nous aurons besoin. Mais pour réussir, nous aurons besoin de toute ton attention et de ta curiosité. Nous devons, ensemble, commenter les équations que nous présenterons, pour leur donner vie et écouter leur chant. C'est important : toutes les formules physiques et mathématiques ont un chant qu'il nous faut écouter; ce n'est qu'à ce prix que nous pouvons comprendre leur vraie signification.

BEATRIX : D'accord. Je suis prête. Par quoi commençons-nous?

EURISTIDE : Commençons par faire un peu de géométrie. La géométrie, c'est le chapitre des mathématiques qui décrit et construit les modèles de l'espace qui nous entoure.

Une des notions fondamentales de la géométrie, c'est la notion de **dimension**. L'espace qui nous entoure est un espace à trois dimensions, tu le sais certainement. Mais cela vaut la peine de bien comprendre cette idée de dimension. Nous en aurons besoin par la suite.

Supposons que nous soyons une minuscule fourmi, qui se déplace exclusivement sur un fil tendu, que nous pourrions supposer étendu infiniment des deux côtés. Nous n'avons la possibilité de nous déplacer que dans la direction du fil tendu, vers l'avant ou vers l'arrière. Toute autre trajectoire est impossible. Nous disons que nous n'avons qu'un seul **degré de liberté**. Nous disons alors que nous sommes plongés dans un **espace de dimension 1**.

Supposons toujours que nous soyons cette fourmi infiniment petite.

Nous considérons maintenant que nous pouvons nous déplacer sur une feuille de carton plane infiniment étendue en toutes directions. Nous avons maintenant la possibilité de nous déplacer devant nous, vers l'arrière, à droite ou à gauche. Tous nos déplacements, quels qu'ils soient, peuvent être considérés comme combinaison d'un déplacement vers le Nord (ou le Sud) d'une part, et d'un déplacement vers l'Est (ou l'Ouest) d'autre part. En effet, je représente cela sur la figure ci-dessous :

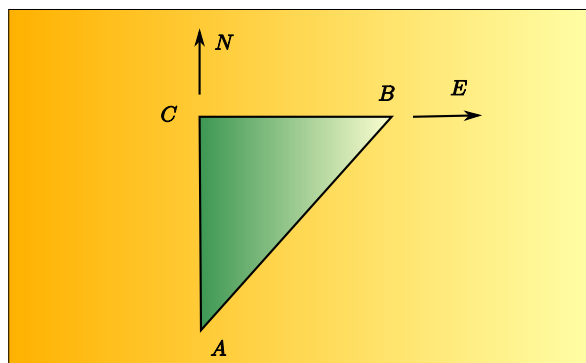


Fig. 2 - Décomposition d'un déplacement suivant les axes N et E

Supposons que nous soyons initialement situés au point  $A$ . Le fait de nous déplacer du point  $A$  vers le point  $B$  peut être virtuellement décomposé comme la combinaison du déplacement  $AC$  vers le Nord, suivi du déplacement  $CB$  vers l'Est. Il en va de même pour tous les déplacements. Cette interprétation toujours possible montre que nous avons deux degrés de liberté pour nos déplacements dans cet espace constitué d'une feuille de carton. Nous disons que nous sommes baignés dans un **espace de dimension 2**.

Supposons maintenant que nous soyons un papillon. Nous pouvons alors nous déplacer comme la fourmi sur des feuilles de carton, mais nous pouvons aussi nous élever au-dessus de ces feuilles, ou voler en-dessous de ces dernières. Nous avons en résumé, maintenant la possibilité de nous déplacer en avant, en arrière, à droite, à gauche, vers le haut, vers le bas. Nous avons à présent trois degrés de liberté pour nos déplacements. Nous disons que nous sommes baignés dans un **espace de dimension 3**. C'est l'espace qui nous entoure et que nous connaissons tous bien.

BEATRIX : Nous avons vu l'espace de dimension 1, puis 2, puis 3. Peut-on continuer, et inventer un espace de dimension 4, puis 5, puis...

MATHINE : Bien sûr, nous pouvons généraliser et imaginer des espaces de dimension  $n$ , où  $n$  peut prendre une valeur entière (c'est-à-dire un entier comme 1, 2, 3, etc.) quelconque. Pour formaliser cela, nous allons devoir introduire la notion de coordonnées. Je laisse Euristide expliquer cela.

EURISTIDE : Pour définir des coordonnées, nous allons commencer par regarder ce que cela signifie dans un espace de dimension 2, donc sur notre fameuse feuille de carton. Nous commençons par choisir un point sur la feuille; nous appellerons ce point l'**origine**, et nous allons le noter  $O$ .

A partir de ce point, nous allons tracer deux droites distinctes : par exemple, l'une vers le Nord que nous appellerons  $ON$  et l'une vers l'Est que nous appellerons  $OE$  :



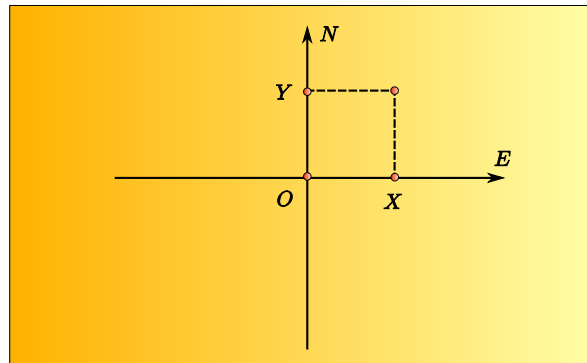


Fig. 3 - Le principe du système de coordonnées

Si nous considérons un point  $P$  sur notre espace de dimension 2, nous pouvons tracer une droite parallèle à l'axe Nord passant par le point  $P$ , et une droite parallèle à l'axe Est, passant également par le point  $P$ . Ce faisant, nos deux dernières droites rencontrent les axes passant par  $O$  en deux points  $X$  et  $Y$  respectivement, comme cela est représenté sur la figure ci-dessus.

Avec ce système, nous pouvons définir le point  $P$  au moyen des deux longueurs  $OX$  et  $OY$ . Nous disons que nous avons défini un **système de coordonnées** dans notre espace à 2 dimensions, et l'origine  $O$  et les axes  $OE$  et  $ON$  constituent le référentiel de ce système de coordonnées. En général, les axes  $OE$  et  $ON$  sont notés  $OX$ ,  $OY$ . Nous adopterons cette notation dorénavant.

**MATHINE** : On formalise cela en mathématiques en disant que nous définissons sur l'espace de dimension 2 un système de coordonnées par le point d'origine  $O$  et deux droites distinctes  $OX$  et  $OY$ , appelés **axes** du système de coordonnées.

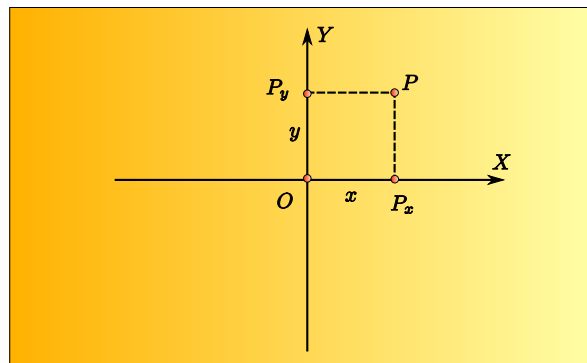


Fig. 4 - Système d'axes de coordonnées

Il faut bien noter que l'exemple ici représente un jeu d'axes de coordonnées perpendiculaires. Ce n'est pas toujours le cas : il est possible de construire un système de coordonnées avec des axes obliques.

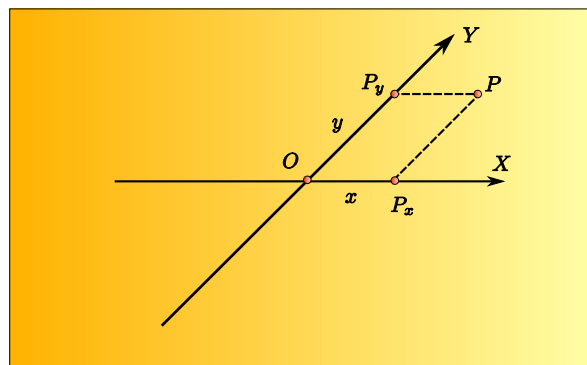


Fig. 5 - Système d'axes de coordonnées oblique

Un point  $P$  est caractérisé par sa projection  $P_y$  sur l'axe  $OY$  parallèlement à  $OX$  et sa projection  $P_x$  sur l'axe  $OX$  parallèlement à  $OY$ , comme indiqué sur la figure ci-dessus. Si la distance de  $O$  à  $P_y$  est notée  $y$  et la distance de  $O$  à  $P_x$  est notée  $x$ , le point  $P$  est alors entièrement déterminé par les deux nombres  $x$  et  $y$ . Le couple  $(x, y)$  est alors appelé **couple des coordonnées de  $P$  dans le système de coordonnées  $(OX, OY)$** .

Cette notation est tellement pratique qu'on fait bien souvent un pas supplémentaire en confondant  $P$  avec le couple  $(x, y)$ . On parle alors du point  $(x, y)$ .

BEATRIX : Ah oui, j'ai compris. Nous sommes dans un espace de dimension 2, dans lequel les points sont des couples  $(x, y)$ .

Si nous étions dans un espace de dimension 3, les points seraient des triplets  $(x, y, z)$ . Sur un dessin, cela ressemblerait à :

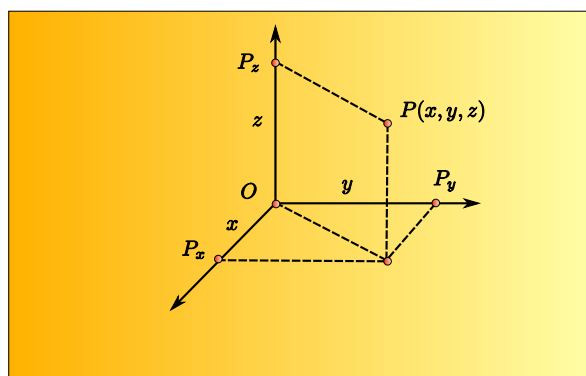


Fig. 6 - Système de coordonnées en 3 dimensions

MATHINE : Exactement ! Tu as parfaitement compris. Dans un espace de dimension 2, les points sont représentés par des couples. Dans un espace de dimension 3, les points sont représentés par des triplets. Et dans un espace de dimension 4, nous représenterons les points par des quadruplets  $(x, y, z, t)$ , etc.

Nous pouvons maintenant généraliser facilement : dans un espace de dimension  $n$ , nous représenterons les points par une série de  $n$  coordonnées, que nous appelons des  $n$ -uplets  $(x_1, x_2, \dots, x_n)$ .

EURISTIDE : Si maintenant nous voulons regarder les choses d'un point de vue plus physique, le plus simple est de partir de ce qui nous entoure. Si nous avons besoin de 3 données indépendantes pour décrire un point dans le monde (la hauteur, la largeur, et la longueur), alors nous pouvons construire un espace de dimension 3 pour modéliser ce monde. C'est ce que fait la géométrie euclidienne telle que nous l'avons apprise à l'école.

Notre espace ambiant est donc un espace de dimension 3. Les physiciens ont imaginé d'introduire dans cette description du monde matériel une dimension supplémentaire, qui est le temps. En introduisant le temps dans l'espace de dimension 3, nous obtenons un espace de dimension 4. C'est ce que nous appellerons **l'espace-temps**.

BEATRIX : Ça y est ! Voilà déjà l'espace-temps qui pointe le bout de son nez ! Mais c'est vrai que pour connaître parfaitement un événement, il faut le situer dans l'espace (donc avec trois coordonnées) et

le situer dans l'histoire (au moyen d'une coordonnée appelée temps.) Par exemple, je dis : "J'ai fêté mon 20ème anniversaire à Pontivy au 30 rue Nationale."

MATHINE : Revenons pour l'instant à l'espace usuel de dimension 3. Le formalisme des triplets permet de calculer directement la distance entre deux points à partir des composantes de ces triplets.

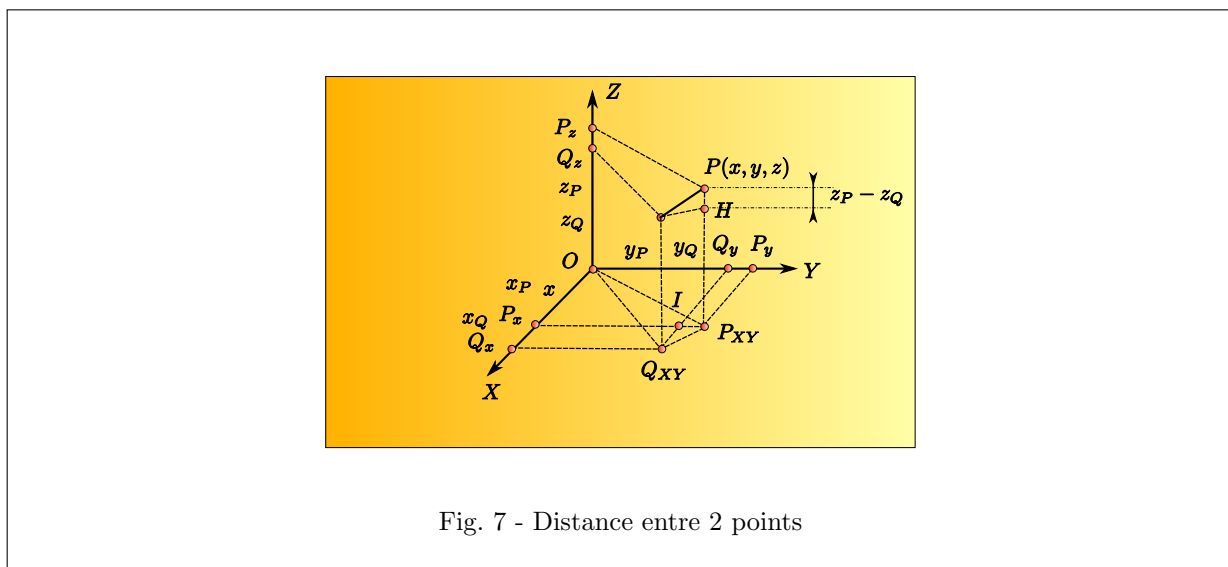


Fig. 7 - Distance entre 2 points

On suppose ici que les axes sont perpendiculaires entre eux deux à deux. On trouve donc dans la figure que j'ai dessinée à l'instant un grand nombre de triangles rectangles pour lesquels nous pouvons appliquer le fameux théorème de pythagore qui dit que le carré de la longueur de l'hypothénuse d'un triangle rectangle est égal à la somme des carrés des longueurs des deux autres côtés.

BEATRIX : Je crois que j'ai un peu oublié l'inoubliable théorème de Pythagore...

MATHINE : Le théorème de Pythagore donne une propriété simple mettant en relation les côtés d'un triangle rectangle.

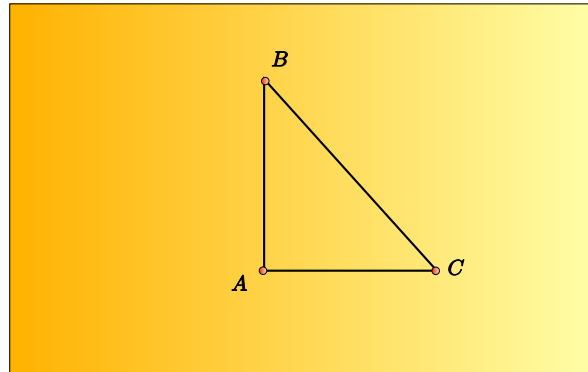
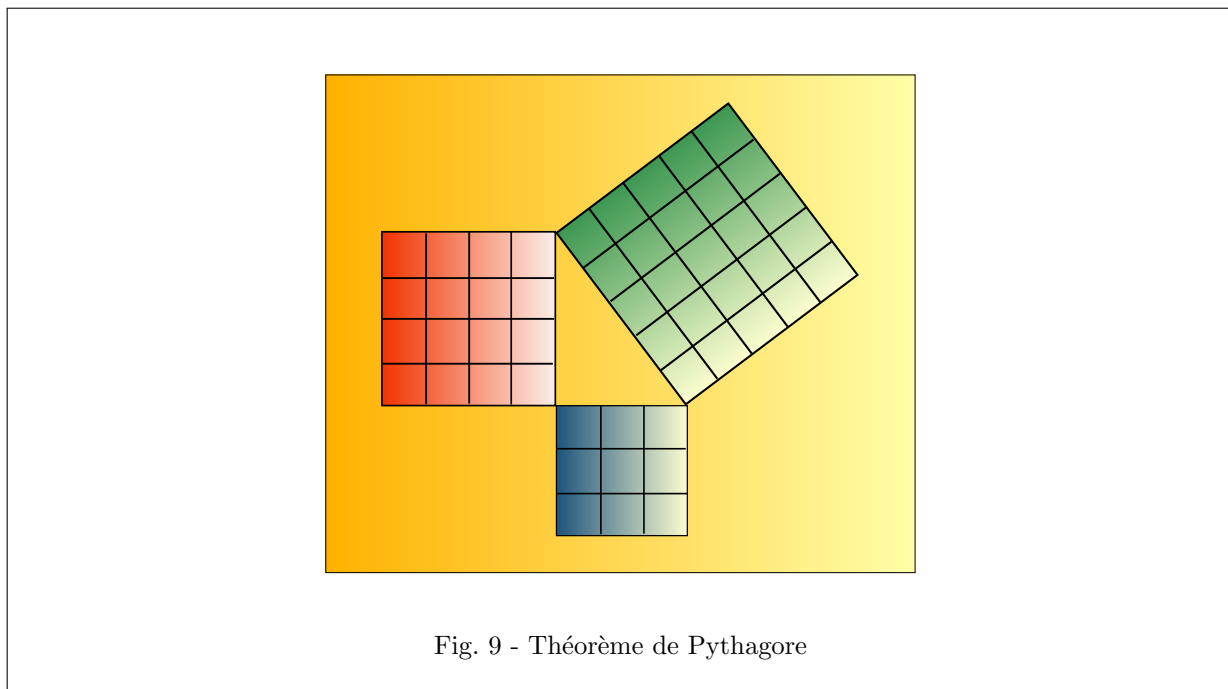


Fig. 8 - Triangle rectangle

Pour un triangle rectangle  $ABC$  rectangle en  $A$ , on a :

$$BC^2 = AB^2 + AC^2. \quad (1)$$

BEATRIX : Ah! Ça y est, je me rappelle ce théorème. Je me souviens que mon prof. de maths. m'avait expliqué ce théorème avec le dessin suivant :



Les deux côtés perpendiculaires du triangle rectangle ont pour longueur respectivement 3 cm et 4 cm. On construit sur chacun de ces côtés un carré. Pour le côté de 3 cm, le carré a pour surface  $3 \times 3 = 9$  carrés de  $1 \text{ cm}^2$ . Pour le côté de 4 cm, le carré a pour surface  $4 \times 4 = 16$  carrés de  $1 \text{ cm}^2$ . Le nombre total de carrés de  $1 \text{ cm}^2$  de ces deux grands carrés construits est de  $9 + 16 = 25$ . Et on constate qu'en construisant un grand carré sur le côté le plus grand du triangle (l'hypoténuse), on obtient un carré possédant  $5 \times 5 = 25$  petits carrés de  $1 \text{ cm}^2$ . On a ainsi vérifié qu'il y a autant de petits carrés dans le grand carré construit sur l'hypoténuse que de petits carrés réunis dans les deux grands carrés construits sur les deux autres côtés du triangle. C'est magique !

**MATHINE** : Voilà, c'est bien cela le théorème de Pythagore.

En étudiant la figure précédente, on trouve plusieurs triangles rectangles sur lesquels nous allons appliquer le théorème de Pythagore

On a, dans le triangle rectangle  $PQH$ , si  $H$  est la projection de  $P$  sur le plan horizontal où se trouve  $Q$  :

$$PQ^2 = (Q_{XY}P_{XY})^2 + (z_P - z_Q)^2. \quad (2)$$

Puis, dans le triangle rectangle  $Q_{XY}P_{XY}I$  :

$$(Q_{XY}P_{XY})^2 = (x_Q - x_P)^2 + (y_P - y_Q)^2. \quad (3)$$

D'où :

$$PQ^2 = (x_Q - x_P)^2 + (y_P - y_Q)^2 + (z_P - z_Q)^2. \quad (4)$$

En notant que :

$$(y_P - y_Q)^2 = (y_Q - y_P)^2 \quad (5)$$

$$(z_P - z_Q)^2 = (z_Q - z_P)^2, \quad (6)$$

on finit par obtenir la relation suivante :

$$PQ^2 = (x_Q - x_P)^2 + (y_Q - y_P)^2 + (z_Q - z_P)^2, \quad (7)$$

ce qui donne finalement

$$PQ = \sqrt{(x_Q - x_P)^2 + (y_Q - y_P)^2 + (z_Q - z_P)^2}. \quad (8)$$

EURISTIDE : Autrement dit, pour calculer la distance entre deux points dans un système de coordonnées à axes perpendiculaires, il suffit de faire la somme des carrés des différences de leurs coordonnées, et de prendre la racine carrée de tout cela.

BEATRIX : Oui, c'est finalement plutôt simple.

EURISTIDE : Mais attention, c'est beaucoup plus compliqué si les axes ne sont pas perpendiculaires. Mais restons pour l'instant à un niveau de complexité raisonnable.

Nous allons maintenant aborder une question délicate : la **courbure d'un espace**. Nous verrons plus tard avec Mathine comment formaliser cette notion au moyen d'outils mathématiques plus puissants que ce que nous connaissons maintenant. Pour l'instant, nous allons regarder cette notion de courbure sous un angle intuitif.

Considérons un espace de dimension 2 (notre feuille de carton de tout à l'heure) référencé par un système de coordonnées dessiné sur la feuille de carton sous forme d'une grille représentant des axes de coordonnées et les différentes parallèles à ces axes, permettant de repérer un point dans l'espace.

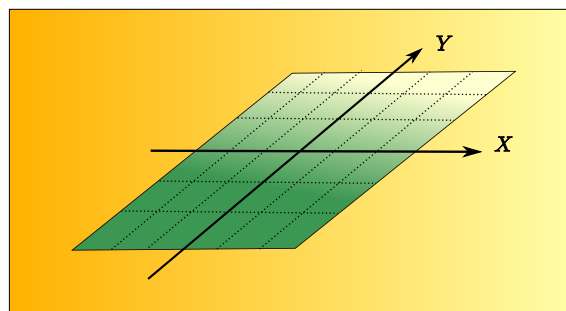


Fig. 10 - Système de coordonnées sur une surface plate

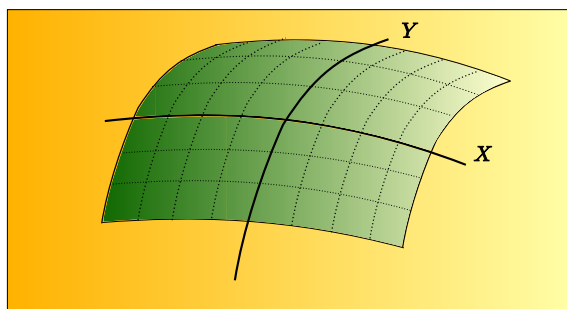


Fig. 11 - Système de coordonnées d'une surface curviligne

Les choses ont quelque peu changé dans ce nouvel espace de dimension 2. D'abord, lorsque je parcours l'axe des  $Y$ , ma trajectoire n'est plus droite, mais courbe. Ensuite, la direction  $OY$  n'est plus constante en tous les points de l'espace. Elle varie en fonction de l'endroit où nous nous trouvons, par exemple sur la courbe  $OY$  qui était à l'origine la droite  $OY$ , et que nous avons incurvée par la déformation de la pâte à modeler. Nous sommes en présence d'un espace curviligne, et la géométrie n'y est plus aussi simple que dans un espace plat.

Un exemple très connu d'espace curviligne est celui de la surface de la Terre. Nous, humains, sommes infiniment petits à la surface de la Terre, et par conséquent, nous pouvons considérer que nous sommes à la surface d'un espace curviligne de dimension 2.

**BEATRIX** : Et quelles sont les conséquences pour nous, humains infiniment petits à la surface de cette Terre ?

**EURISTIDE** : Nous verrons plus en détail les conséquences précisément lorsque nous travaillerons sur la courbure au sens mathématique du terme. Mais je vais prendre un exemple qui conduit à un paradoxe en apparence, qui nous permettra de mieux illustrer et de comprendre en quoi la géométrie à la surface d'une sphère est très différente de celle d'un espace plat.

Nous allons d'abord considérer que nous sommes sur une surface plate : je suis en un point  $A$  et je me déplace sur 1 km dans une direction quelconque jusqu'à un point  $B$ . Puis je tourne à droite d'un angle de  $90^\circ$  à partir de ce point  $B$ , et je me déplace sur 1 km jusqu'à un point  $C$ . Puis je tourne encore à droite d'un angle de  $90^\circ$  à partir de ce point  $C$ , et je me déplace sur 1 km jusqu'à un point  $D$ . Si je tourne encore à droite d'un angle de  $90^\circ$  à partir de ce point  $D$  et si je me déplace sur 1 km, je rejoindrai le point  $A$ , mon point de départ.



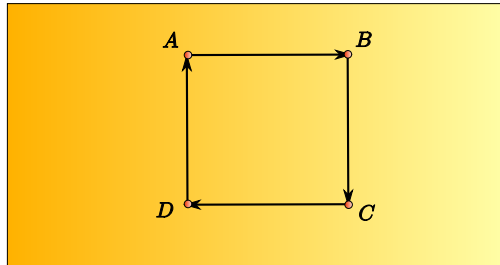


Fig. 12 - Trajectoire à angles droits carrée dans un plan

Nous savons donc que sur un plan, lorsque nous nous déplaçons en effectuant un angle de  $90^\circ$  tous les kilomètres, nous décrivons un carré avant de revenir à notre point de départ. Considérons maintenant que nous sommes sur la Terre au pôle Nord que nous appellerons pour l'occasion point *A*. Effectuons un premier trajet vers le Sud, jusqu'à un point *B* situé sur l'équateur. Tournons à droite à partir de ce point *B* d'un angle de 90 degrés et déplaçons-nous jusqu'à un point *C* situé également sur l'équateur. Tournons encore une fois à droite à partir de ce point *C* d'un angle de 90 degrés et avançons. Oh stupeur ! nous avons rejoint le pôle Nord, donc le point *A* ! Et ceci en 3 étapes au lieu de 4 comme précédemment.

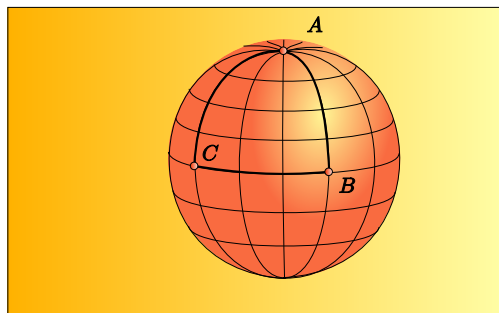


Fig. 13 - Trajectoire à angles droits sur un triangle sphérique

Nous nous sommes déplacés sur un triangle, dont les 3 angles sont des angles droits ! Surprenant, non ?

BEATRIX : C'est troublant... Ainsi, la somme des angles d'un triangle sur la surface de la Terre n'est pas de  $180^\circ$ , comme d'habitude ?

MATHINE : Non, et c'est l'une des caractéristiques de ces surfaces curvilignes : la somme des angles d'un triangle n'est pas égale à 180 degrés. Si la courbure est convexe, comme dans le cas de la Terre, la somme des angles est supérieure à 180 degrés. Si la courbure est concave, comme dans le cas d'une fosse, la somme des angles serait inférieure à 180 degrés.

Nous avons pu illustrer le concept d'espace plat et d'espace curviligne pour des espaces de dimension 2. Il est possible bien sûr de généraliser cette notion à un espace de dimension 3 ou plus. Mais il est hélas impossible de représenter visuellement ce qu'est un espace de dimension 3 (ou de dimension supérieure) curviligne. C'est pourquoi nous aurons besoin plus tard d'un formalisme mathématique plus puissant pour caractériser un espace curviligne.

EURISTIDE : Lorsque la notion d'espace curviligne intervient, mais aussi dans d'autres cas, nous avons besoin d'un outil pour caractériser la direction que prend une courbe ou une surface en un point donné.

Illustrons cela par un exemple : je suis en voiture sur une petite route de campagne, en Picardie par exemple. La route est droite, et je vois bien que la direction que pointe l'avant de ma voiture est la même que celle de la route. Quelques kilomètres plus loin, la route amorce une longue courbe.

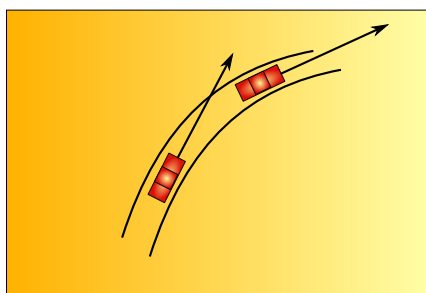


Fig. 14 - Direction d'une voiture dans un virage

En un point quelconque de cette courbe, je m'aperçois que l'avant de ma voiture ne pointe plus dans la direction de la route, mais suivant une direction qui semble quitter la route. Si je suivais cette direction, je quitterais la route, mais en fait, à chaque instant, ma voiture, dont les roues sont tournées, pivote et son avant pointe dans une nouvelle direction. Tant que la route aura une forme courbée, ma voiture pointera ainsi suivant une direction qui semblera frôler et quitter la route. Si nous représentons ce phénomène de façon plus stylisée, nous avons le schéma suivant :

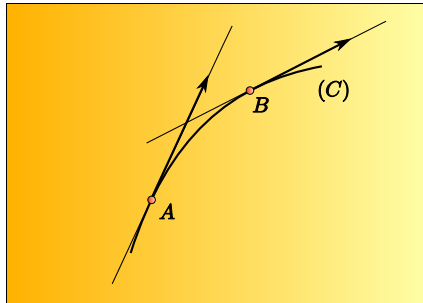


Fig. 15 - Les directions tangentes au virage

Sur les points  $A$  et  $B$ , nous avons représenté deux droites qui sont appelées **tangentes de la courbe** aux points  $A$  et  $B$ .

**BEATRIX:** Si je comprends bien, la tangente d'une courbe en un point, c'est la direction que semble prendre cette courbe à l'endroit où se trouve le point. Les trajectoires droites ont la caractéristique d'avoir la même tangente en tout point, tandis que les trajectoires courbes ont justement une tangente qui varie d'un point à un autre.

**MATHINE :** Exactement, Béatrix. Nous étudierons un peu plus tard l'outil mathématique permettant de décrire ces droites tangentes.

**EURISTIDE :** En attendant, passons à l'étape suivante. Nous avons vu à l'instant la tangente d'une courbe. Nous pouvons considérer que nous avons étudié la tangente d'un espace de dimension 1. La tangente d'un espace de dimension 1 est donc une droite, c'est-à-dire elle-même un espace de dimension 1. L'étape suivante, c'est de regarder ce qu'il en est de la tangente à une surface ( $S$ ), qui est un espace de dimension 2.

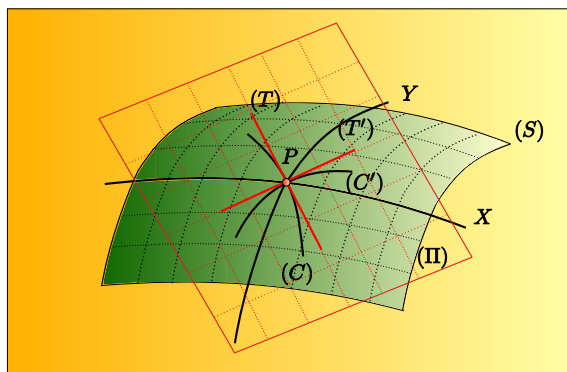


Fig. 16 - Le plan tangent d'une surface en un point

En un point  $P$  de cette surface  $(S)$ , nous pouvons considérer deux courbes quelconques distinctes  $(C)$  et  $(C')$ . Si nous considérons les tangentes respectives de ces deux courbes  $(T)$  et  $(T')$ , et si nous avons choisi les courbes  $(C)$  et  $(C')$  de telle façon que ces tangentes soient distinctes, nous avons deux droites  $(T)$  et  $(T')$  sécantes au point  $P$ , et qui déterminent par conséquent un plan  $(\Pi)$  qui est appelé le **plan tangent** en  $P$  à la surface  $(S)$ . Le plan possède la propriété de contenir toutes les tangentes des courbes appartenant à la surface  $(S)$  et passant par le point  $P$ .

La tangente à une surface, qui est un espace de dimension 2, est un plan, qui est aussi un espace de dimension 2.

**MATHINE** : Nous pouvons généraliser : l'espace tangent en un point  $P$  à un espace  $(E)$  de dimension  $n$  est un espace de dimension  $n$ . Il contient tous les espaces tangents de dimension  $n - 1$ , tangents aux espaces de dimension  $n - 1$  passant par le point  $P$  et appartenant à l'espace  $(E)$ . C'est un peu compliqué à dire, mais en fait c'est une mécanique toute simple : un espace tangent est de même dimension que l'espace auquel il est tangent.

Cette notion d'espace tangent nous sera extrêmement utile pour la relativité générale, parce qu'elle nous permettra de construire des coordonnées cartésiennes dans un espace curviligne.

**EURISTIDE** : Puisque nous parlons de coordonnées, nous devons revoir ensemble ces notions, car la relativité générale est émaillée de nombreux systèmes de coordonnées et de nombreux changements de coordonnées qui seront parfois indispensables pour simplifier des expressions, changer le point de vue sous lequel nous représentons mathématiquement les phénomènes physiques, et voire éliminer des paradoxes, des divisions par zéro ou autres impasses mathématiques.

**MATHINE** : Oui, tu as raison, Euristide, un des fondements de la géométrie, c'est le système de coordonnées. On appelle **système de coordonnées** dans un espace de dimension  $n$ , comme nous l'avons vu précédemment, la donnée d'une origine, de  $n$  axes distincts et d'une unité de mesure sur chacun de ces axes. Par exemple, en dimension 2, nous définissons sur la figure :

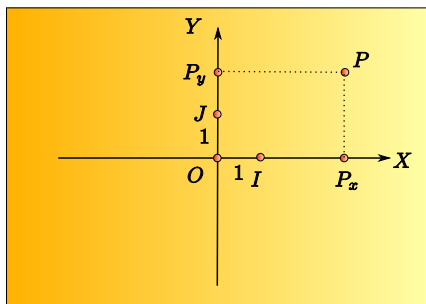


Fig. 17 - Les coordonnées du point  $P$

l'origine  $O$ , les axes  $OX$  et  $OY$  et les points  $I$  sur  $OX$  et  $J$  sur  $OY$  tels que  $OI = 1$  et  $OJ = 1$ . Dans ce système de coordonnées, le point  $P$  sur la figure possède les coordonnées  $(3, 2)$ . Effectuer un changement de coordonnées, c'est définir un nouveau système de coordonnées et exprimer le point  $P$ , par exemple, dans le nouveau système de coordonnées en fonction de ses coordonnées dans l'ancien système de coordonnées. Prenons un exemple : partant du système de coordonnées précédent, je définis un nouveau système de coordonnées ayant pour origine  $O'$  de coordonnées dans l'ancien système  $(1, 1)$ , pour axe  $O'X'$  parallèle à l'axe  $OX$ , pour mesure  $O'I'$  telle que  $O'I' = 0.5$  dans l'ancien système, et pour axe  $O'Y'$  un axe faisant  $45^\circ$  avec l'axe  $O'X'$  et pour mesure  $O'J'$  telle que  $O'J' = \sqrt{2}$  dans le système précédent.

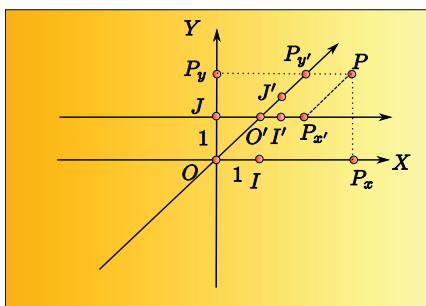


Fig. 18 - Le changement de système de coordonnées

On peut calculer les coordonnées de  $P$  dans le nouveau système de coordonnées au moyen de la formule de transformation de l'ancien système de coordonnées dans le nouveau système de coordonnées. En exploitant la figure ci-dessus, on déduit que  $P_{y'}$  est la longueur du segment  $O'P_{y'}$ .

Considérons le triangle rectangle  $P_{y'}O'P_{x'}$ . Il est rectangle en  $P_{x'}$ , donc on peut utiliser le théorème de Pythagore :

$$O'P_{x'}^2 + P_{x'}P_{y'}^2 = O'P_{y'}^2. \quad (9)$$

Or :

$$O'P_{x'} = P_{x'}P_{y'}, \quad (10)$$

puisque l'inclinaison de l'axe  $O'Y'$  est de 45 degrés, donc :

$$O'P_{y'}^2 = 2P_{x'}P_{y'}^2, \quad (11)$$

soit :

$$O'P_{y'} = \sqrt{2}(JP_y), \quad (12)$$

ou encore :

$$O'P_{y'} = \sqrt{2}(P_y - OJ). \quad (13)$$

Pour calculer  $P_{x'} = O'P_{x'}$ , il faut retrancher à  $P_x$  les longueurs  $JO' = OI$  et  $P_{y'}P = P_{x'}P_{y'} = JP_y = P_y - OJ$ .  
Donc :

$$O'P_{x'} = P_x - OI - P_y + OJ = P_x - P_y. \quad (14)$$

Ces longueurs, exprimées en unités du nouveau système de coordonnées, doivent être doublées, puisque  $O'P_{y'} = 2O'J'$  et  $O'P_{x'} = 2O'I'$ . Donc :

$$P_{y'} = 2\sqrt{2}(P_y - OJ) = 2\sqrt{2}(P_y - 1) \quad (15)$$

$$\begin{aligned} P_{x'} &= 2(P_x - OI - P_y + OJ) \\ &= 2(P_x - P_y). \end{aligned} \quad (16)$$

Donc les coordonnées dans le nouveau système de coordonnées se déduisent de celles de l'ancien système de coordonnées par :

$$P_{x'} = 2(P_x - P_y) \quad (17)$$

$$P_{y'} = 2\sqrt{2}(P_y - 1). \quad (18)$$

En appliquant ce résultat au point  $P(3, 2)$  dans l'ancien système de coordonnées, on obtient, dans le nouveau système de coordonnées :

$$P_{x'} = 2 \times (3 - 2) = 2 \quad (19)$$

$$P_{y'} = 2\sqrt{2}(2 - 1) = 2\sqrt{2}. \quad (20)$$

Dans la suite, pour simplifier les notations, nous exprimerons les coordonnées par  $(x, y)$  pour l'ancien système de coordonnées et  $(x', y')$  pour le nouveau système de coordonnées, ce qui permet de rendre les formules plus lisibles :

$$x' = 2(x - y) \quad (21)$$

$$y' = 2\sqrt{2}(y - 1) \quad (22)$$

**EURISTIDE** : Nous avons vu plusieurs choses importantes ici. D'abord, nous savons maintenant que pour changer de coordonnées, il suffit de connaître la loi de transformation des anciennes coordonnées dans

les nouvelles. Cette loi peut être déduite lorsqu'on connaît, pour le nouveau système la position de l'origine, l'orientation des axes et la longueur unité de chacun des axes. Nous avons vu aussi que les systèmes de coordonnées peuvent être variés, avec ou sans axes perpendiculaires entre eux, avec ou sans longueurs unité identiques sur les deux axes.

MATHINE : D'ailleurs, les systèmes de coordonnées ayant cette particularité d'avoir des axes perpendiculaires sont très usités, puisqu'ils sont le choix le plus naturel qu'on adopte lorsqu'on souhaite repérer un espace. Ces systèmes de coordonnées sont dits **cartésiens**.

Si un système cartésien possède en outre des unités de longueur égales, alors il est dit **orthonormé**.

BEATRIX : En fait, si je comprends bien, un bon exemple de système de coordonnées orthonormé, c'est une feuille de papier quadrillé, ou un papier millimétré sur lesquels on aurait sélectionné pour origine un point à l'intersection du quadrillage.

Et un bon exemple de système de coordonnées non orthonormé, c'est un papier à échelle logarithmique où les axes  $x$  et  $y$  ne possèdent pas la même mesure.

Enfin, un exemple de système de coordonnées non cartésien est constitué par les parallèles et les méridiens qui permettent de se repérer à la surface de la Terre.

MATHINE : C'est exact. Mais on peut imaginer des systèmes de coordonnées encore plus insolites. Par exemple, toujours dans le plan, nous pouvons repérer le plan autour de l'origine, par un angle et une longueur, en considérant qu'un point quelconque est placé à une distance  $r$  de l'origine, et que la droite qui le raccorde à l'origine fait un angle  $\theta$  avec une droite de référence (horizontale par exemple.)

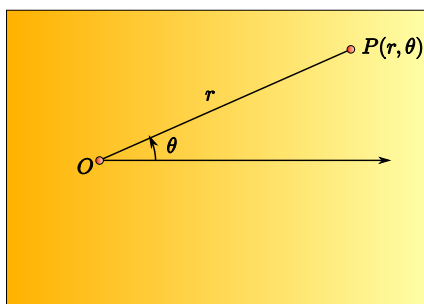


Fig. 19 - Les coordonnées polaires en 2 dimensions

De telles coordonnées sont appelées coordonnées polaires (parce qu'elles se rapportent au point  $O$  appelé le pôle.) Dans un espace de dimension 3, pour un point  $P$ , et un pôle  $O$ , nous pouvons définir également des coordonnées polaires, en considérant l'angle  $\theta$  que fait l'intersection du plan vertical issu de la droite  $OP$  avec le plan horizontal passant par  $O$ , avec une droite de référence horizontale issue de  $O$ , et en considérant l'angle  $\phi$  que fait la même intersection avec la droite  $OP$ .

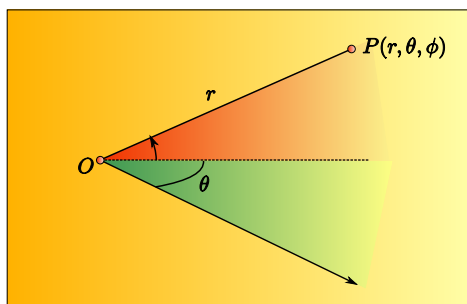


Fig. 20 - Les coordonnées polaires en 3 dimensions

## 2 Scène I-2 : La notion de fonction

EURISTIDE : Tout au long de cette discussion sur la relativité générale, nous utiliserons le concept omniprésent de la fonction au sens mathématique.

BEATRIX : Je connais les fonctions : c'est  $y = f(x)$ .

EURISTIDE : Oui, ce que tu cites est un exemple de fonction, où  $x$  est la variable, et  $y$  l'image. Les fonctions sont au coeur même de la physique. Chaque fois que nous aurons besoin de décrire le comportement d'un système physique dépendant de paramètres variables, nous utiliserons le concept de fonction mathématique.

Une fonction mathématique est une opération transformant un ensemble de données appelé source en un ensemble de données appelé image de la fonction. En d'autres termes, la fonction définit une loi de transformation des données qui lui sont fournies en entrée en un ensemble de données qu'elle restitue en sortie.

BEATRIX : En quelque sorte, une fonction c'est une machine à transformer. On lui donne quelque chose à manger, et elle nous rend cette chose transformée.

MATHINE : C'est bien ça.

Les fonctions sont désignées généralement par des lettres telles que  $f$  ou  $g$ . Les données en entrée sont symbolisées par des lettres appelées variables qui sont généralement notées par la lettre  $x$ . L'image de la



fonction est le plus souvent notée par la lettre  $y$ . Une fonction est donc généralement exprimée sous une forme telle que :

$$y = f(x). \quad (23)$$

Cette notation signifie que la fonction  $f$  transforme les différentes valeurs que peut prendre la variable  $x$  en un ensemble de valeurs image représenté par les valeurs de  $y$ .

Prenons un exemple. Nous pouvons nous intéresser à la fonction qui transforme tout nombre réel en son double. Nous pouvons noter une telle fonction sous la forme suivante, où  $x$  peut parcourir l'ensemble de tous les nombres réels :

$$y = \text{double}(x) = 2x. \quad (24)$$

Si nous donnons le nombre 2 à manger à la fonction, elle nous rend le nombre 4. Si nous lui donnons 2,5, elle nous rend 5.

Si nous considérons maintenant la fonction qui transforme un angle en son cosinus, nous écrirons une fonction un peu plus compliquée, sous la forme suivante, où  $\alpha$  peut parcourir les nombres réels compris entre 0 et  $2\pi$  :

$$y = \cos \alpha. \quad (25)$$

**BEATRIX** : Le nombre de fonctions que l'on peut imaginer est incroyablement grand, non? En fait, j'ai l'impression qu'il est infini...

**MATHINE** : Oui, il est infini. Il est même infiniment infini. Définir une fonction, c'est définir un ensemble source : il y a une infinité de façons de choisir cet ensemble source. Ensuite, pour définir une fonction, il faut aussi choisir un ensemble image, dans lequel nous allons affecter à chaque élément de l'ensemble source, un élément de l'ensemble image. Il y a aussi une infinité de choix possibles pour cette affectation si les ensembles source et image sont infinis. C'est donc une infinité de choix dans un ensemble que l'on peut choisir d'une infinité de façons possibles. C'est ce que j'appelle être infiniment infini.

**BEATRIX** : C'est vertigineux!

**MATHINE** : Nous pouvons bien entendu imaginer également des fonctions à plusieurs variables, qui sont également tout aussi fréquemment utilisées en physique.

Par exemple, la transformation qui, à tout couple de coordonnées  $(x, y)$  d'un point  $P$  dans un espace de dimension 2 fait correspondre sa distance à l'origine  $O$  est une fonction à deux variables dont l'image est l'ensemble des nombres réels. On peut écrire, si  $d$  est la distance dont nous parlons :

$$d = f(x, y) = \sqrt{x^2 + y^2}. \quad (26)$$

On peut imaginer également des fonctions dont la source est l'image comportent plusieurs variables. Par exemple, la fonction qui fait correspondre à un point donné le symétrique par rapport à l'origine  $O$  dans un espace de dimension 2, se traduit par :

$$(x', y') = f(x, y) = (-x, -y). \quad (27)$$

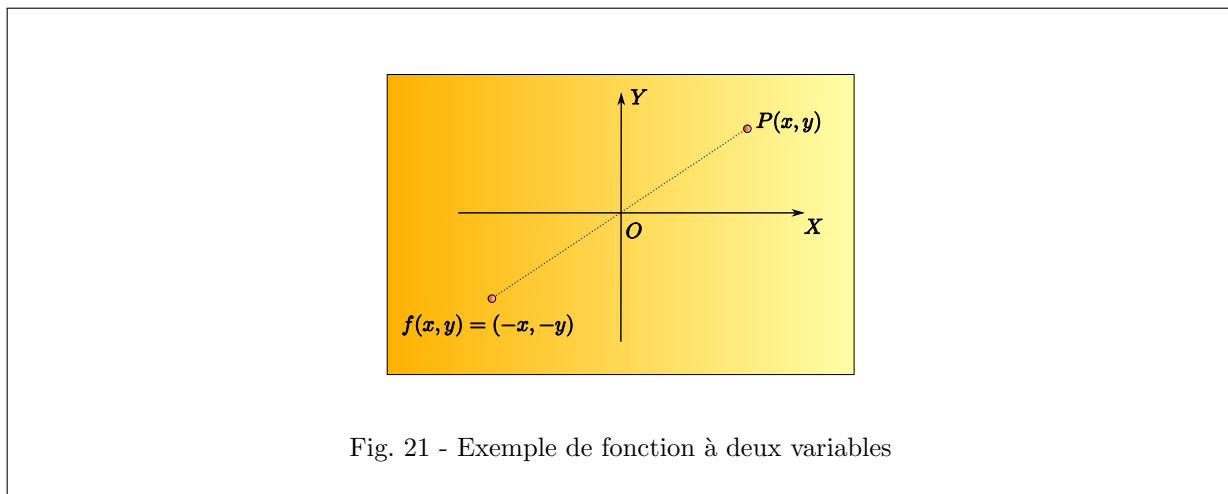


Fig. 21 - Exemple de fonction à deux variables

BEATRIX : En résumé, les fonctions sont omnivores. Elles peuvent manger de tout, apparemment.

EURISTIDE : Elles sont même cannibales, puisqu'elles peuvent également transformer des fonctions, c'est-à-dire leurs propres congénères.

### 3 Scène I-3 : La notion de limite d'une fonction

MATHINE : Certaines fonctions possèdent une ou plusieurs valeurs de variable pour lesquelles elles ne sont pas définies. Par exemple, cela peut se produire parce que cette valeur de variable provoque une division par zéro, ou entraînerait de prendre la racine carrée d'une valeur négative, etc. Dans ce cas, il est intéressant d'étudier le comportement de la fonction lorsque la variable concernée s'approche de cette valeur. On dit alors qu'on étudie le comportement de la fonction à sa limite.

EURISTIDE : Cette notion de limite d'une fonction est fondamentale pour la physique. Et en particulier, nous en aurons beaucoup l'usage pour la relativité générale, parce qu'un certain nombre de phénomènes expliqués par cette théorie découlent du comportement à la limite de certaines fonctions plus ou moins compliquées.

BEATRIX : Alors, je suis toute ouïe...

MATHINE : Prenons l'exemple de la fonction :

$$y = f(x) = \frac{1}{x}. \quad (28)$$

Cette fonction n'est évidemment pas définie pour  $x = 0$ , à cause de la division par zéro.

Considérons une valeur de  $x$  arbitrairement petite que nous noterons  $x_0$ . Ce peut être, pour fixer les idées, 0,0001. La valeur de  $y$  pour cette valeur de la variable  $x_0$  est :

$$y_0 = \frac{1}{x_0}. \quad (29)$$

En l'occurrence, cela donnerait, pour  $x_0 = 0,0001$ ,  $y = 10000$ . Considérons maintenant une valeur de variable  $x_1$  plus grande que  $x_0$ . Pour fixer les idées, prenons  $x_1 = 0,001$ . Nous avons :

$$x_1 > x_0, \quad (30)$$

donc :

$$\frac{1}{x_1} < \frac{1}{x_0}, \quad (31)$$

puisque, bien évidemment,  $1000 < 10000$ .

Nous constatons donc que lorsque la variable  $x$  décroît en s'approchant de 0 (on dit **en tendant vers 0**), l'image de la fonction croît en tendant vers l'infiniment grand.

Nous pouvons formaliser cette propriété en démontrant que si nous choisissons un nombre arbitrairement grand  $A$ , correspondant à l'image par  $f$  d'une valeur de la variable  $x$ , que nous notons  $x_1$  :

$$x_1 = \frac{1}{A}, \quad (32)$$

alors, nous pouvons toujours trouver un nombre plus grand que  $A$  correspondant à l'image par  $f$  d'une valeur de la variable  $x$ , notée  $x_0$ , et qui soit plus proche de 0 que  $x_1$ . Il suffit en effet de choisir  $A + 1$ , et nous pouvons bien sûr vérifier immédiatement que  $x_0 = \frac{1}{A+1}$  est compris entre 0 et  $x_1$  :

$$0 < \frac{1}{A+1} < \frac{1}{A}. \quad (33)$$

Ceci s'exprime de la façon suivante : aussi grande que je choisisse une image par  $f$  d'un nombre proche de 0, je peux toujours trouver un nombre plus grand qui soit l'image d'un nombre encore plus proche de 0.

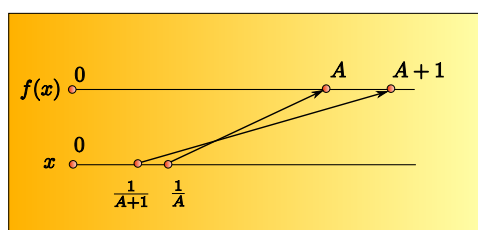


Fig. 22 -  $f(x)$  tend vers  $+\infty$  quand  $x$  tend vers 0

On dit que la fonction  $f$  tend vers  $+\infty$  (prononcer "plus l'infini") lorsque  $x$  tend vers 0 en restant positif (on dit "par valeur positive"), et ceci se note :

$$\lim_{x \rightarrow 0^+} \frac{1}{x} = +\infty. \quad (34)$$

De façon similaire, nous pouvons considérer la limite d'une fonction lorsque  $x$  tend vers  $+\infty$ . Un raisonnement analogue au précédent nous permet de déterminer que la limite de notre fonction  $f(x) = \frac{1}{x}$  lorsque  $x$  tend vers  $+\infty$  est 0.

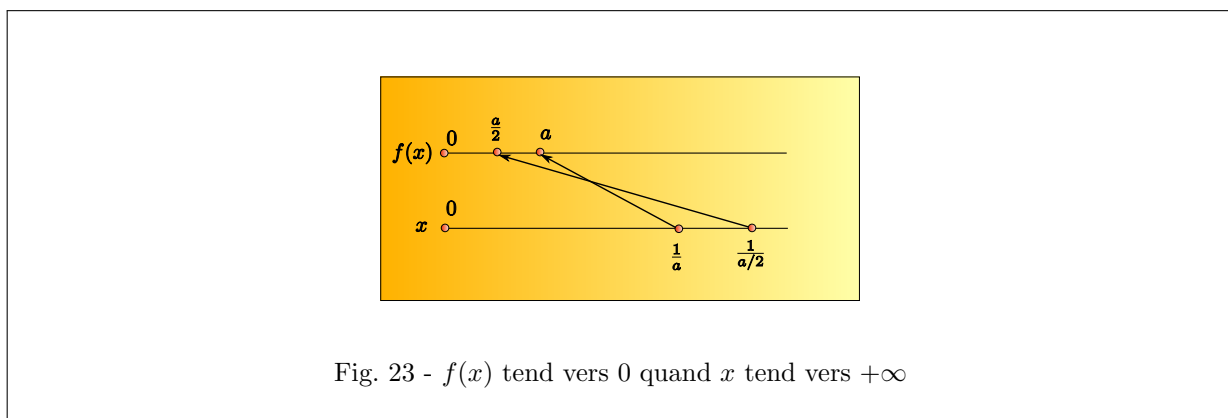
En effet, choisissons un nombre arbitrairement petit  $a$ , correspondant à l'image par  $f$  d'une valeur de la variable  $x$  que nous notons  $x_0$  :

$$x_0 = \frac{1}{a}. \quad (35)$$

Nous pouvons toujours trouver un nombre plus petit que  $a$ , correspondant à l'image par  $f$  d'une valeur de la variable  $x$  notée  $x_1$ , et qui soit plus grande que  $x_0$ . Il suffit en effet de choisir  $a/2$ , et nous pouvons vérifier immédiatement que :

$$\frac{1}{a/2} > \frac{1}{a}. \quad (36)$$

On exprime ceci en disant qu'aussi petite que je choisisse une image par  $f$  d'un nombre grand, je peux toujours trouver un nombre plus petit qui soit l'image d'un nombre encore plus grand.



On dit que la fonction  $f$  tend vers 0 lorsque  $x$  tend vers  $+\infty$  :

$$\lim_{x \rightarrow +\infty} \frac{1}{x} = 0. \quad (37)$$

## 4 Scène I-4 : Le concept de voisinage

EURISTIDE : Nous devons aborder maintenant une notion dont nous aurons souvent besoin par la suite : le voisinage. Cette notion permet d'observer une fonction au voisinage d'un point et de simplifier

son expression en considérant que tout près du point considéré (on dit dans son **voisinage**), la fonction "ressemble" à une autre fonction plus simple.

BEATRIX : C'est en fait une notion d'approximation d'une fonction, non ?

EURISTIDE : Non, en réalité, c'est plus subtil que cela, et surtout plus rigoureux qu'une banale approximation. Comment l'expliquer... Ce procédé consiste à dire que lorsque l'intervalle dans lequel on observe le comportement d'une fonction tend à diminuer pour se rapprocher infiniment du point considéré tout seul, sans jamais l'atteindre, alors, à cette limite, la fonction est équivalente à une fonction plus simple.

BEATRIX : C'est compliqué !

EURISTIDE : Nous allons prendre un exemple pour bien comprendre. Supposons que nous nous intéressions à la fonction représentée par le graphe suivant :

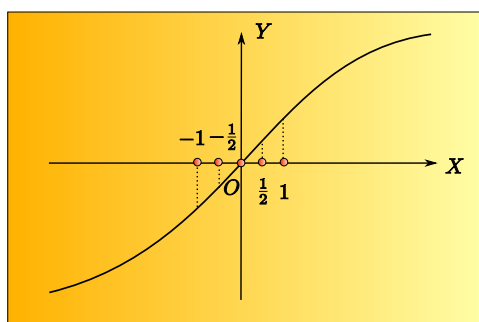


Fig. 24 - Analyse d'une fonction au voisinage de 0

Nous allons rester dans le domaine intuitif pour l'instant : la formulation mathématique de cette fonction importe peu. Ce qui compte, c'est sa forme. Elle est incurvée lorsque  $x$  est relativement grand, et lorsque  $x$  possède une valeur négative relativement importante. Schématiquement, pour  $x > 1$ , la fonction est franchement incurvée, et pour  $x < -1$  également. En revanche, elle est presque droite pour  $x > -1$  et  $x < 1$ . Donc, dans l'intervalle  $[-1, 1]$ , la fonction ressemble à une droite.

C'est encore plus vrai quand l'intervalle est restreint à  $[-\frac{1}{2}, \frac{1}{2}]$ . Et plus l'intervalle diminue autour de 0, plus la fonction ressemble à une droite d'équation  $y = x$ . En d'autres termes, lorsque  $\epsilon$  tend vers 0, la fonction  $y = f(x)$  sur l'intervalle  $[-\epsilon, \epsilon]$  tend vers  $y = x$ . Nous disons que la fonction  $y = f(x)$  est équivalente à  $y = x$  au voisinage de 0.

BEATRIX : Je comprends mieux : il s'agit de la forme de la fonction à la limite d'un intervalle qui se rétrécit indéfiniment. C'est pour cela qu'on parle de voisinage. Et comme il s'agit d'une limite, la notion est bien rigoureuse, ce n'est pas une approximation.

EURISTIDE : Voilà.

BEATRIX : Mais là, nous avons regardé intuitivement comment cela fonctionnait. C'est bien, mais comment allons-nous formaliser cela dans un langage mathématique et comment trouvons-nous la fonction plus simple à laquelle est équivalente notre fonction de départ ?

MATHINE : C'est là que les mathématiques interviennent. Et plus particulièrement un mathématicien appelé Taylor. Taylor a découvert une formule permettant de représenter n'importe quelle fonction au voisinage du point 0 sous la forme d'une somme infinie d'expressions du type  $ax^n$ , où  $a$  est un nombre et  $x^n$  est la fonction qui à  $x$  fait correspondre  $x^n$ .

Par exemple, Taylor a démontré que la fonction  $y = \cos x$  pouvait, au voisinage de 0, s'écrire :

$$y = 1 - \frac{x^2}{2} + \frac{x^4}{24} - \dots + (-1)^n \frac{x^{2n}}{(2n)!} + \dots \quad (38)$$

Cette expression s'appelle le développement de Taylor de la fonction  $y = \cos x$  au voisinage de  $x = 0$ . L'avantage de cette expression, et tout son génie, c'est que les termes de la somme à droite de l'égalité sont de plus en plus petits ; lorsque  $x$  est plus petit que 1 (ce qui est évidemment le cas lorsque  $x$  est dans le voisinage de 0), nous avons :

$$1 > \frac{x^2}{2} > \frac{x^4}{24} > \dots > \frac{x^{2n}}{(2n)!} > \dots \quad (39)$$

Il y a mieux : plus  $x$  est proche de 0, plus l'écart se creuse entre ces nombres qui deviennent vertigineusement de plus en plus petits quand les puissances de  $x$  augmentent.

Prenons un exemple : si  $x = 0,1$ , alors  $\frac{x^2}{2} = 0,005$  et  $\frac{x^4}{24} \approx 0,00004$ . Si  $x = 0,01$ , alors  $\frac{x^2}{2} = 0,00005$  et  $\frac{x^4}{24} \approx 0,00000004$  ! On dit que  $\frac{x^4}{24}$  est négligeable devant  $\frac{x^2}{2}$  lorsque  $x$  est au voisinage de 0.

Plus formellement, on dit qu'une fonction  $f(x)$  est négligeable par rapport à une fonction  $g(x)$  au voisinage de 0 si en choisissant un nombre  $\epsilon$  aussi petit que l'on veut, on peut toujours trouver une valeur de  $x$  telle que :

$$|f(x)| \leq \epsilon |g(x)|. \quad (40)$$

Autrement dit, on peut bien choisir de multiplier  $g(x)$  par un nombre aussi petit que l'on veut, ou pourra toujours trouver, en prenant  $x$  de plus en plus petit, une valeur de  $f(x)$  qui est encore plus petite. C'est la façon formelle de dire que  $f(x)$  est négligeable devant  $g(x)$ .

Ceci s'écrit :

$$\boxed{f(x) = o(g(x))}. \quad (41)$$

Il faut prononcer : "f de x est égal à petit 'o' de g de x".

Donc, dans le développement de Taylor de la fonction  $y = \cos x$ , au voisinage de 0, on voit que les termes de la somme sont négligeables devant tous les termes précédents de la somme.

Et c'est bien utile !

Cette expression permet de déduire que la fonction  $y = \cos x$  est équivalente à la fonction 1, mais aussi à la fonction :

$$y = 1 - \frac{x^2}{2}, \quad (42)$$

au voisinage de 0.

Cette notion d'équivalence est formalisée de la façon suivante. On considère  $f$  et  $g$  deux fonctions dont on veut montrer qu'elles sont équivalentes au voisinage de 0. Nous choisissons un nombre  $\epsilon$  aussi petit que l'on veut ; alors, si on peut toujours trouver une valeur de  $x$  aussi petite que l'on veut telle que :

$$|f(x) - g(x)| < \epsilon, \quad (43)$$

autrement dit, telle que l'écart entre les deux fonctions équivalentes soit plus petit encore que ce nombre arbitrairement petit  $\epsilon$ , alors les deux fonctions sont équivalentes.

En l'occurrence, pour la fonction  $f(x) = \cos x$ , nous disons que  $f(x)$  est équivalente à la fonction  $g(x) = 1$ . C'est dire que  $f(x)$ , au voisinage de 0, peut s'écrire :

$$f(x) = 1 + o(1). \quad (44)$$

Cette expression signifie que  $f(x)$  est égale à 1 plus des termes qui sont négligeables devant 1 au voisinage de 0. C'est bien le cas, puisque les termes suivants dans le développement de Taylor en  $x^2$ ,  $x^4$  et  $x^6$  sont infiniment petits lorsque  $x$  est au voisinage de 0.

Mais nous pouvons écrire aussi :

$$f(x) = 1 - \frac{x^2}{2} + o(x^2). \quad (45)$$

Cela signifie que  $f(x)$  est égale à  $1 - \frac{x^2}{2}$  plus des termes qui sont négligeables devant  $x^2$  au voisinage de 0.

**BEATRIX** : Je comprends... C'est fantastique ! J'imagine que nous pouvons utiliser ces propriétés miraculeuses pour simplifier des fonctions horriblement compliquées lorsqu'on veut les étudier au voisinage d'un point.

**EURISTIDE** : Oui. Et nous aurons beaucoup d'occasions de le faire. L'étude d'un champ de gravitation au voisinage du centre d'un trou noir, au voisinage de son horizon des événements, etc., toutes sortes de choses que nous verrons tout à l'heure. Mais je vois que tu as parfaitement saisi toute la magie et cependant toute la rigueur de cette démarche. Nous pouvons donc continuer.

## 5 Scène I-5 : La dérivée

**BEATRIX** : Alors, quelle est l'étape suivante ?

**EURISTIDE** : Nous allons continuer à étudier le comportement des fonctions. C'est normal que nous insistions un peu sur les fonctions : comme je te l'ai dit tout à l'heure, les fonctions sont omniprésentes en physique. Nous avons commencé par observer le comportement d'une fonction à la limite, en analysant la valeur vers laquelle elle tendait lorsque sa variable tendait vers une valeur donnée. Ceci nous permettait notamment de formaliser la valeur d'une fonction à l'une de ses "extrémités" qu'elle ne peut jamais atteindre. Puis nous avons voulu analyser la forme d'une fonction en un point ; c'est ce que nous avons fait en introduisant la notion de voisinage et celle de fonctions équivalentes au voisinage d'un point. Ceci nous a permis de formaliser le fait qu'une fonction au voisinage d'un point ressemblait à une fonction plus simple. L'étape suivante, c'est maintenant de nous intéresser à la vitesse à laquelle une fonction évolue. Par exemple, est-elle croissante, à quelle vitesse est-elle croissante ? Croît-elle de façon constante ? Décroît-elle ?

Est-elle au contraire constante ?

L'outil que nous allons utiliser pour cela est la **dérivée** d'une fonction.

MATHINE : Lorsque nous étudions une fonction  $f(x)$ , nous sommes intéressés par la manière dont cette fonction évolue lorsque la variable  $x$  varie. Prenons par exemple la fonction  $f(x)$  représentant la position d'un coureur le long d'une piste de course rectiligne. Supposons plus précisément que  $f(x)$  représente la distance parcourue par le coureur à l'instant  $x$ . Pour  $x = 0$ , le coureur se trouve sur la ligne de départ, et il a parcouru une distance nulle. Donc :

$$f(0) = 0. \quad (46)$$

Pour  $x = a$ , le coureur a atteint la ligne d'arrivée, et il a parcouru la distance :

$$f(a) = d. \quad (47)$$

Si nous nous intéressons aux performances de notre coureur, nous avons envie de connaître sa vitesse. Pour cela, nous savons que nous pouvons diviser la distance parcourue  $d$  par le temps écoulé pour parcourir cette distance, et nous obtiendrons une vitesse que nous pouvons exprimer par :

$$v = \frac{d}{a}. \quad (48)$$

Si le coureur n'a pas bougé entre l'instant  $x = 0$  et l'instant  $x = a$ , alors  $d = 0$ , et par conséquent sa vitesse est bien nulle.

BEATRIX : Si l'athlète court très vite, il parcourt plus de distance en un temps donné  $a$ , donc  $v$  est plus grande. Inversement, s'il court moins vite, il parcourt moins de distance dans le même temps, et par conséquent  $v$  est plus petite. Ça marche, la formule  $v = \frac{d}{a}$  est bien l'expression de la vitesse du coureur.

MATHINE : Oui, mais nous allons voir que ce n'est pas tout à fait suffisant. D'abord, généralisons un tout petit peu l'étude précédente, pour traiter le cas où notre coureur est sur la ligne de départ à l'instant  $x = a$  et où il atteint la ligne d'arrivée à l'instant  $x = b$ . Dans l'intervalle de temps entre  $a$  et  $b$ , il parcourt la distance  $f(b) - f(a)$ . Par conséquent, sa vitesse sera exprimée comme suit :

$$v = \frac{f(b) - f(a)}{b - a}. \quad (49)$$

BEATRIX : D'accord. Mais je trouve qu'il y a un hic.

MATHINE : C'est intéressant. Lequel ?

BEATRIX : Nous avons exprimé la vitesse du coureur en regardant la distance qu'il a parcourue et en divisant cette distance par le temps écoulé. C'est bien. Mais nous obtenons une vitesse moyenne, dans ce cas. Notre coureur a certainement fait varier sa vitesse : par exemple, il a commencé par accélérer, donc sa vitesse a augmenté, puis sa vitesse a été constante pendant quelque temps, et enfin sa vitesse a commencé à décroître, parce qu'il était fatigué. Moi, ce qui m'intéresse, c'est sa vitesse à un instant donné. En fait, ce qu'il nous faudrait, c'est la fonction qui représente la vitesse du coureur.



MATHINE : D'accord. Allons-y. Supposons que nous nous intéressions à la vitesse de notre coureur à l'instant  $x = c$  situé quelque part entre l'instant  $x = a$  et l'instant  $x = b$ .

Nous pouvons affiner notre évaluation de la vitesse du coureur en considérant deux points autour de  $x = c$ , soit  $x = x_0$  et  $x = x_1$ , tels que :

$$a < x_0 < c < x_1 < b. \quad (50)$$

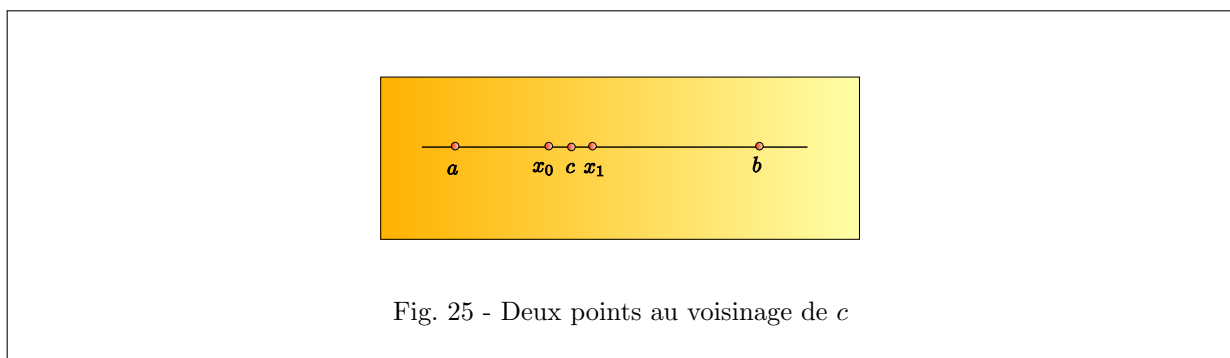


Fig. 25 - Deux points au voisinage de  $c$

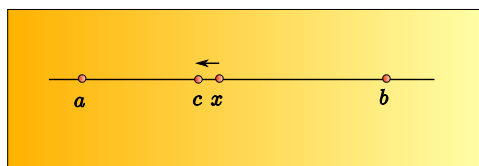
Nous pouvons alors calculer la vitesse moyenne autour de l'instant  $c$  comme nous l'avons fait précédemment :

$$v = \frac{f(x_1) - f(x_0)}{x_1 - x_0} \quad (51)$$

Cette vitesse est la vitesse moyenne du coureur dans un intervalle de temps qui peut être choisi tout petit autour de l'instant  $c$ . C'est donc une très bonne approximation de la vitesse du coureur. Mais ce n'est qu'une approximation, et ce n'est donc pas encore la vitesse exacte de notre champion à cet instant.

BEATRIX : Là, je ne sais pas pourquoi, mais mon petit doigt me dit que nous allons diminuer à l'infiniment petit l'intervalle de temps autour de l'instant  $c$ , pour obtenir par le calcul précédent une approximation de plus en plus exacte de la vitesse du coureur à l'instant  $c$ .

MATHINE : Oui, c'est une très bonne idée. Nous allons utiliser la limite. Pour cela, nous allons changer un peu l'expression précédente pour des raisons pratiques. Nous allons considérer que  $x_0 = c$ , et que  $x_1 = x$ , où  $x$  est la variable temps que nous allons faire tendre vers  $c$ .

Fig. 26 - Vitesse lorsque  $x$  tend vers  $c$ 

On a une fonction vitesse dépendant de  $x$  :

$$v(x) = \frac{f(x) - f(c)}{x - c}. \quad (52)$$

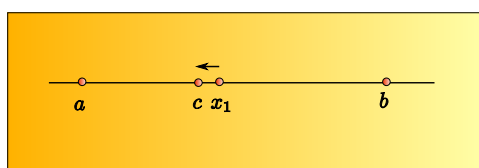
En choisissant  $x$  de plus en plus proche de  $c$ , c'est-à-dire en faisant tendre  $x$  vers  $c$ , nous obtenons une limite :

$$v(c) = \lim_{x \rightarrow c} \frac{f(x) - f(c)}{x - c}. \quad (53)$$

Cette limite est la vitesse exacte du coureur à l'instant  $c$ .

Nous pouvons définir la fonction vitesse du coureur en prenant  $c$  pour variable. Pour les besoins des conventions d'écriture des mathématiciens (qui sont des gens maniaques!), si  $c$  est variable, nous allons plutôt l'appeler  $x$ , et nous écrivons la fonction vitesse sous la forme :

$$v(x) = \lim_{x_1 \rightarrow x} \frac{f(x_1) - f(x)}{x_1 - x}. \quad (54)$$

Fig. 27 - Vitesse lorsque  $x_1$  tend vers  $c$ 

Si nous abandonnons l'exemple du coureur et que nous considérons plus généralement une fonction  $f$ , cette limite est la définition de ce qu'on appelle la dérivée de la fonction  $f$ .

L'expression de la vitesse est une des applications possibles de la notion de dérivée, mais il y en a d'autres. Par exemple, si la fonction représente un volume d'eau déversé par un robinet dans une baignoire (tu sais

bien, les fameux problèmes de robinets), la dérivée représentera le débit du robinet. La dérivée de la fonction  $f(x)$  peut être notée de différentes manières :

$$f'(x) \quad (55)$$

$$\dot{f}(x) \quad (56)$$

$$\frac{df}{dx}. \quad (57)$$

On peut démontrer facilement, en utilisant la définition de la dérivée sous forme d'une limite d'une fraction, un certain nombre de résultats d'opérations sur ces dérivées. En particulier, la dérivée de la somme de deux fonctions est la somme des dérivées des deux fonctions, ce qui s'écrit, si  $f$  et  $g$  sont deux fonctions :

$$\boxed{(f + g)' = f' + g'}. \quad (58)$$

La dérivée du produit d'une fonction par une constante est égale au produit de la dérivée de la fonction par cette même constante, ce qui s'écrit, si  $f$  est une fonction et  $\alpha$  une constante :

$$\boxed{(\alpha f)' = \alpha f'}. \quad (59)$$

Enfin, la dérivée du produit de deux fonctions est un peu plus complexe à calculer. On doit faire l'addition de deux termes, l'un dans lequel on a pris la dérivée d'une des fonctions et l'autre dans lequel on a pris la dérivée de la deuxième fonction. Si  $f$  et  $g$  sont deux fonctions :

$$\boxed{(fg)' = f'g + fg'}. \quad (60)$$

**EURISTIDE** : Certaines de ces règles de calcul sont assez intuitives : quand j'additionne deux fonctions, leurs vitesses de croissance vont logiquement s'additionner également. Par conséquent, leurs dérivées s'ajoutent. De même, quand je multiplie une fonction par un nombre constant, sa croissance est également multipliée par ce nombre constant. Par conséquent, la dérivée sera multipliée par la constante. La dernière loi, pour la multiplication, est moins intuitive ; mais on peut la retrouver en considérant la limite de la différence de la fonction suivante, où  $\epsilon$  est un nombre qui tend vers 0 :

$$\begin{aligned} & \frac{f(x + \epsilon)g(x + \epsilon) - f(x)g(x)}{\epsilon} \\ &= \frac{f(x + \epsilon)g(x + \epsilon) - f(x)g(x + \epsilon) + f(x)g(x + \epsilon) - f(x)g(x)}{\epsilon} \\ &= \frac{(f(x + \epsilon) - f(x))g(x + \epsilon)}{\epsilon} + \frac{(g(x + \epsilon) - g(x))f(x)}{\epsilon} \end{aligned} \quad (61)$$

En considérant que  $g(x + \epsilon) = g(x) + (g(x + \epsilon) - g(x))$ , on trouve :

$$\begin{aligned} \frac{f(x + \epsilon)g(x + \epsilon) - f(x)g(x)}{\epsilon} &= \frac{f(x + \epsilon) - f(x)}{\epsilon} g(x) + \frac{g(x + \epsilon) - g(x)}{\epsilon} f(x) \\ &+ \frac{(f(x + \epsilon) - f(x))(g(x + \epsilon) - g(x))}{\epsilon}. \end{aligned} \quad (62)$$

On sait que  $\epsilon$  tend vers 0. Le dernier terme est la multiplication de deux termes tendant vers 0, divisé par un terme tendant vers 0. Il est donc, lorsque  $\epsilon$  tend vers 0, infiniment plus petit que les deux autres termes de la même somme (qui sont le rapport de deux termes qui tendent vers 0.) Il est donc possible de négliger

le dernier terme par rapport aux deux premiers dans cette expression, et nous pouvons donc écrire, lorsque  $\epsilon$  est petit :

$$\frac{f(x + \epsilon)g(x + \epsilon) - f(x)g(x)}{\epsilon} \approx \frac{f(x + \epsilon) - f(x)}{\epsilon}g(x) + \frac{g(x + \epsilon) - g(x)}{\epsilon}f(x) \quad (63)$$

Et par conséquent, en passant à la limite lorsque  $\epsilon = 0$ , on obtient la formule recherchée pour la dérivée d'un produit de fonctions.

BEATRIX : La dernière expression de la dérivée  $\frac{df}{dx}$  rappelle, sous une forme condensée, l'expression que nous avons vue tout à l'heure :

$$\frac{f(x_1) - f(x)}{x_1 - x}, \quad (64)$$

si on décide que dans "df", le "d" veut dire différence.

MATHINE : C'est un peu ça. "df" signifie en fait **différentielle** de  $f$ , c'est-à-dire une variation infinitésimale (infinitésimal signifie infiniment petit) de  $f$ . Donc l'expression :

$$\frac{df}{dx} \quad (65)$$

représente le rapport d'une variation infiniment petite de  $f$  par une variation infiniment petite de  $x$ . C'est une façon symbolique de représenter la limite lorsque  $x_1$  tend vers  $x$  de :

$$\frac{f(x_1) - f(x)}{x_1 - x}. \quad (66)$$

En synthèse,  $df$  est la différentielle de  $f$  et  $dx$  est la différentielle de  $x$ .

EURISTIDE : Justement, la différentielle est beaucoup utilisée en physique. Il est toujours utile d'analyser le comportement d'un système pour des variations infiniment petites de ses paramètres. Dans ce cas, le comportement est souvent plus simple, et plus facile à mettre en équations. On verra cela plus tard à propos des équations différentielles.

## 6 Scène I-6 : La dérivée partielle

BEATRIX : On a vu jusque là des dérivées de fonctions plutôt simples comme la vitesse d'un coureur sur une ligne droite. Mais si le coureur évolue sur une courbe quelconque (une route sinueuse par exemple), il nous faudra, je crois, une fonction un peu plus compliquée, qui définira sa position selon deux variables  $x$  et  $y$  le long de l'axe Nord et Est par exemple. Son comportement va donc être décrit par une fonction ou plusieurs fonctions à plusieurs variables. Comment calcule-t-on alors la dérivée d'une fonction à plusieurs variables ?

EURISTIDE : C'est une bonne question. La réponse est simple : nous allons considérer la dérivée de la fonction par rapport à une de ses variables, en figeant les autres variables de la fonction.

MATHINE : En langage mathématique, cela donne quelque chose de plutôt intuitif. Considérons une fonction  $f(x, y)$  dépendant des variables  $x$  et  $y$ . Nous voulons étudier la dérivée de cette fonction par rapport à la variable  $x$ , en supposant que la variable  $y$  reste constante. Cette dérivée, dite **dérivée partielle de  $f$  par rapport à  $x$** , s'écrit :

$$\frac{\partial f}{\partial x} = \left. \frac{df}{dx} \right|_{y=Cte} . \quad (67)$$

De façon analogue, la dérivée partielle de  $f$  par rapport à  $y$  s'écrit :

$$\frac{\partial f}{\partial y} = \left. \frac{df}{dy} \right|_{x=Cte} . \quad (68)$$

EURISTIDE : Pour mieux visualiser ce qu'est une dérivée partielle, on va prendre l'exemple d'une fonction :

$$z = f(x, y) \quad (69)$$

représentant l'altitude  $z$  d'un point sur la carte topographique d'une colline.

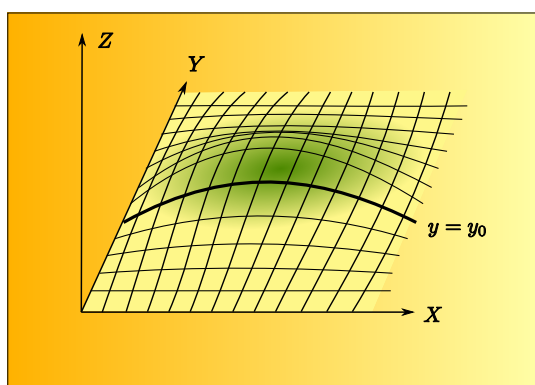


Fig. 28 - Représentation de la colline

Si nous fixons  $y$  à une valeur  $y_0$ , la fonction  $\frac{\partial f}{\partial x}$  représente la vitesse de variation de la courbe indiquée ci-dessous, qui est la coupe de notre colline suivant le plan où  $y = y_0$ .

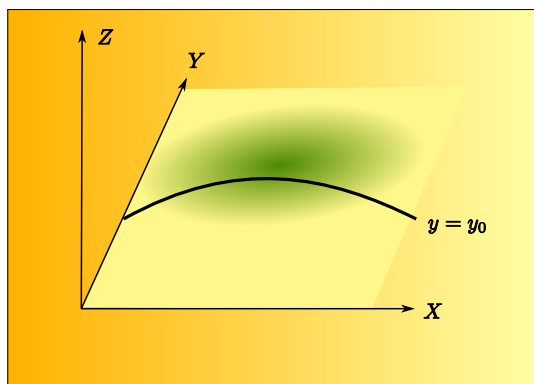


Fig. 29 - Variation d'une fonction lorsqu'on fixe une coordonnée

BEATRIX : En d'autres termes, si j'ai bien compris, c'est la pente de la colline, quand on se déplace sur elle suivant l'axe des  $X$  en une position  $y = y_0$  constante.

MATHINE : C'est exactement cela.

## 7 Scène I-7 : L'intégration

EURISTIDE : Après avoir étudié les dérivées et les différentielles, nous allons nous intéresser à l'opération inverse, qui permet de déduire une fonction à partir de sa dérivée. C'est une question qui revient régulièrement en physique. Revenons une fois encore à l'exemple de notre fameux coureur à pied.

Supposons que nous connaissions son point de départ  $a$ , et la fonction  $v(x)$  de sa vitesse en tout point de sa trajectoire à partir de  $a$ , jusqu'à son point d'arrivée en  $b$ . Nous voulons tracer le parcours du coureur au fil du temps  $x$ . C'est pourquoi nous sommes intéressés par la détermination de la fonction  $f(x)$  qui décrit la position du coureur à tout instant entre  $x = a$  et  $x = b$ . C'est intuitivement possible, puisque dès lors que nous connaissons le point de départ du coureur, en considérant sa vitesse en chaque point, on saura déduire de sa position précédente sa nouvelle position à l'instant d'après en multipliant sa vitesse par un intervalle de temps pendant lequel cette vitesse est bien celle du coureur, et ainsi de suite, de proche en proche.

On sent bien, quand on dit qu'il faut procéder de proche en proche, qu'il faut le faire pour des positions du coureur infiniment voisines les unes des autres.

BEATRIX : Ah ! Je sens que la différentielle de tout à l'heure va pointer le bout de son nez !

EURISTIDE : Oui, nous allons donc naturellement utiliser la différentielle  $dx$  de la variable  $x$ . Supposons que notre coureur se déplace pendant un intervalle de temps  $dx$ . La quantité qu'il faudra ajouter à la position précédente  $x$  du coureur sera le produit  $v(x)dx$ , qui correspond bien à la distance infinitésimale parcourue par le coureur à la vitesse  $v(x)$  pendant le temps infiniment petit  $dx$ .

Ainsi, pour déterminer la position du coureur en un instant quelconque  $x = c$ , il faudra effectuer la somme de tous les termes infiniment petits de la forme  $v(x)dx$  depuis la valeur  $x = a$  jusqu'à la valeur  $x = c$ .

Nous venons de définir ainsi intuitivement ce qu'est une intégrale. Il s'agit d'une somme d'un nombre infiniment grand de termes qui sont eux-mêmes infiniment petits de la forme  $v(x)dx$ , où  $v$  est la fonction de la variable  $x$ .

Cette somme un peu particulière se note par le signe :

$$\int_a^c \cdot \quad (70)$$

Cette notation se prononce "intégrale de  $a$  à  $c$ " ou "somme de  $a$  à  $c$ ". Ainsi, la position de notre coureur au point  $c$  s'écrira :

$$\int_a^c v(x)dx. \quad (71)$$

Et voilà ! Une intégrale, c'est aussi simple que cela !

MATHINE : En utilisant la représentation graphique de  $v(x)$ , nous allons bien comprendre la raison pour laquelle il est nécessaire de faire une somme infinie de termes infiniment petits :

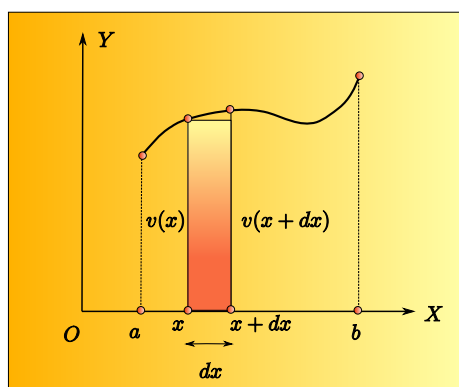


Fig. 30 - Interprétation de  $v(x)dx$

L'expression  $v(x)dx$  représente la surface du rectangle très mince de largeur  $dx$  et de hauteur  $v(x)$ . C'est le rectangle gris dans la figure ci-dessus. Ce petit rectangle ne représente pas tout à fait la surface qu'il nous faut considérer sous la courbe entre  $v(x)$  et  $v(x+dx)$ . Il manque un petit morceau presque triangulaire qui se trouve sous la courbe au dessus du rectangle gris.

Mais, il s'avère que ce petit morceau triangulaire tend à disparaître lorsque la largeur du rectangle  $dx$  tend vers 0. Et par hypothèse,  $dx$  représente justement une donnée qui tend vers 0. Donc, si nous prenons la surface  $v(x)dx$  avec  $dx$  infiniment petit, nous pouvons l'assimiler à la surface se trouvant sous la courbe entre  $v(x)$  et  $v(x + dx)$ .

En effectuant la somme des expressions  $v(x)dx$  entre  $x = a$  et  $x = b$ , nous obtenons la totalité de la surface qui se trouve sous la courbe  $v(x)$  entre les points d'abscisse  $a$  et  $b$ .

BEATRIX : C'est super ! L'intégrale, ce n'est finalement pas plus compliqué qu'un moyen de calculer la surface qui se trouve sous une courbe.

MATHINE : Nous avons considéré jusqu'ici des fonctions à une variable. Nous pouvons généraliser les définitions précédentes, et considérer une fonction  $g(x, y, z)$  à trois variables  $x$ ,  $y$  et  $z$ . Nous pouvons intégrer cette fonction par rapport à ses trois variables  $x$ ,  $y$  et  $z$ , en écrivant une intégrale dite **intégrale triple**, sous la forme, par exemple :

$$\int_{x_0}^{x_1} \int_{y_0}^{y_1} \int_{z_0}^{z_1} g(x, y, z) dx dy dz. \quad (72)$$

BEATRIX : Les fameuses intégrales triples qui font peur à tout le monde.

MATHINE : Oui, les intégrales triples. Mais elles ne sont pas méchantes, en fait. Quand on a compris ce qu'est une intégrale, qu'elle soit simple, double, triple ou multiple, on y retrouve le même mécanisme. Il suffit de le répéter 1 fois, 2 fois ou plusieurs fois la recette de l'intégrale.

Avant de quitter le monde des intégrales, nous allons considérer une dernière notion. Dans certains cas, il peut être utile de considérer des intégrales pour lesquelles les bornes de variation des variables ne sont pas définies, c'est-à-dire qu'elles sont elles-mêmes des variables. Nous construisons alors à partir d'une fonction  $g(x)$  une fonction  $f(x)$  qui est appelée primitive de la fonction  $g$ . Cette fonction primitive dépend alors de la variable  $x$ , et comme nous n'avons pas choisi de borne de départ non plus, cette primitive dépendra aussi du choix de cette borne de départ. Nous allons considérer que la borne de départ est une constante qui est indéfinie. On dit que la primitive est une fonction définie à une constante près. La primitive  $f$  de la fonction  $g$  se note :

$$f(x) = \int g(x) dx + C, \quad (73)$$

où  $C$  est une constante.

Pour une fonction  $h(x, y, z)$  à trois variables, nous aurons une primitive sous forme d'une intégrale triple ressemblant à ceci :

$$f(x) = \int \int \int h(x, y, z) dx dy dz + C. \quad (74)$$

## 8 Scène I-8 : Les équations différentielles

EURISTIDE : Nous avons vu ce qu'est une fonction, sa dérivée, son intégrale. Nous avons aussi vu ce qu'est la différentielle d'une variable ou d'une fonction. Maintenant, nous avons aussi très souvent besoin de déterminer, comme je te l'ai expliqué tout à l'heure, le comportement d'un système physique en analysant son fonctionnement ou son comportement pour des variations infiniment petites de ses variables.



BEATRIX : Pour que faire ?

EURISTIDE : Parce que lorsque nous examinons un système dans l'infiniment petit, il est plus simple. Pourquoi ? Comme précédemment, pour l'intégrale, dans l'infiniment petit, on peut négliger les variations des fonctions qui sont plus petites encore que la variation déjà infiniment petite de la variable. C'est ce qui permet de négliger le petit triangle blanc de la figure tout à l'heure dans l'intégrale. C'est ce qui permet de résoudre des problèmes complexes en passant par l'infiniment petit. On obtient alors une équation décrivant le comportement du système physique dans l'infiniment petit, comportant des variations infiniment petites de ses variables. Une telle équation est appelée **équation différentielle**.

MATHINE : On va prendre un exemple simple pour illustrer cela et bien comprendre ce dont il s'agit. Notre objectif va être de déduire la fonction décrivant la position d'un coureur, à partir d'une fonction décrivant sa vitesse.

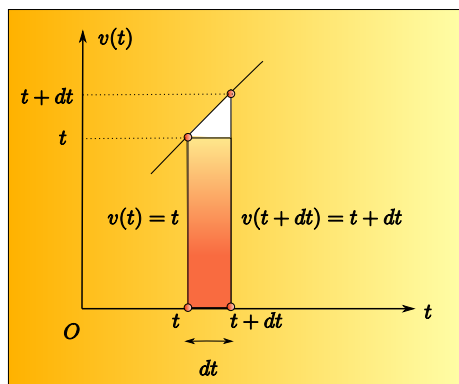
Revenons donc à notre cher coureur. Supposons que notre champion favori ait décidé d'accélérer sa course régulièrement en prenant une vitesse égale numériquement à l'instant présent lorsqu'il est exprimé par exemple en secondes. En d'autres termes, au départ, sa vitesse est de 0 mètres par seconde, au bout de 1 seconde, sa vitesse est de 1 mètre par seconde, au bout de 2 secondes, sa vitesse est de 2 mètres par seconde, etc..

Il est a priori difficile d'exprimer directement l'équation du mouvement du coureur, c'est-à-dire de savoir à quel endroit se trouve le coureur à un instant donné. Ce n'est pas un problème simple.

Mais il est possible de le résoudre, précisément en utilisant la magie des équations différentielles.

Observons ce qu'il se passe lorsque nous faisons varier le temps infiniment peu. Analysons le mouvement du coureur à l'instant  $t$ , et étudions l'évolution de la position  $y$  du coureur après un laps de temps infiniment petit  $dt$ . Nous savons qu'il se trouvera dans une position que nous pouvons noter  $y + dy$  à l'instant  $t + dt$ .

Sa vitesse à l'instant présent est  $t$  par définition (rappelle-toi : le coureur prend une vitesse égale numériquement à l'instant présent.) Sa vitesse à l'instant  $t + dt$  est  $t + dt$ . Nous pouvons représenter graphiquement l'évolution de la vitesse du coureur :

Fig. 31 - Analyse de la vitesse à l'instant  $t$ 

La distance parcourue pendant l'intervalle de temps  $dt$  peut être assimilée à  $tdt$ , puisque nous négligeons de nouveau le petit triangle blanc se trouvant sous la droite et au-dessus du rectangle gris de côtés  $t$  et  $dt$ , puisque  $dt$  tend vers 0 et dans ce cas le petit triangle blanc tend à disparaître.

Par conséquent, la position du coureur après l'instant  $dt$  est :

$$y + dy = y + tdt. \quad (75)$$

En simplifiant les deux membres de l'égalité par  $y$ , on obtient une véritable équation différentielle :

$$dy = tdt, \quad (76)$$

qui peut s'écrire aussi :

$$\frac{dy}{dt} = t. \quad (77)$$

Pour trouver la fonction décrivant le mouvement du coureur, il faut maintenant intégrer cette équation différentielle. Il se trouve que la dérivée de la fonction :

$$f(t) = \frac{1}{2}t^2, \quad (78)$$

est précisément :

$$\frac{df}{dt} = t. \quad (79)$$

Donc la primitive de la fonction  $g(t) = t$  est :

$$\int tdt = \frac{1}{2}t^2 + C, \quad (80)$$

où  $C$  est une constante, puisque nous savons que les primitives sont définies à une constante près. Donc la solution à l'équation différentielle ci-dessus est :

$$f(t) = \frac{1}{2}t^2 + C. \quad (81)$$

Cette constante  $C$  peut être déterminée si nous connaissons la position du coureur à l'instant  $t = 0$  par exemple. En l'occurrence, si sa position est 0 à cet instant  $t = 0$ , alors  $C = 0$ .

BEATRIX : C'est bien, mais il y a une chose qui me chiffonne. Cette histoire de négliger le petit triangle, ça fait un peu bricolage et tour de passe-passe. En fait, on fait des approximations, donc je ne vois pas très bien la rigueur de tout cela...

MATHINE : Il faut reconnaître que c'est un peu troublant. Mais en fait, tout cela est parfaitement rigoureux, et nous avons parfaitement raison de négliger le petit triangle blanc. Il y a d'excellentes raisons pour que nous procédions ainsi.

Reprenons l'exemple précédent et calculons la surface  $ds$  du petit triangle blanc :

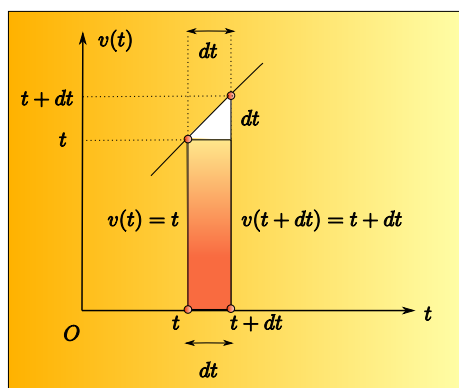


Fig. 32 - Analyse de la vitesse à l'instant  $t$

Le petit triangle blanc est un triangle rectangle qui a pour base  $dt$  et pour hauteur  $dt$ . Il possède donc une surface  $ds$  exprimée par :

$$ds = \frac{1}{2}(dt)^2. \quad (82)$$

Donc, si on prend cette expression en compte, la distance parcourue par le coureur pendant l'intervalle de temps  $dt$  sera :

$$t dt + \frac{1}{2}(dt)^2. \quad (83)$$

Il faut maintenant que nous trouvions un moyen calculatoire pour faire tendre  $dt$  vers 0. Le meilleur moyen est de considérer que nous découpons l'intervalle de temps total  $T$  sur lequel nous calculons le déplacement du coureur en un nombre  $n$  de petits intervalles de temps, et c'est ce nombre  $n$  que nous ferons tendre vers l'infini, pour que la durée des tronçons eux-mêmes tendent vers 0. Et dans ce cas,  $dt = \frac{T}{n}$ . Je représente cette opération sur le schéma ci-dessous :

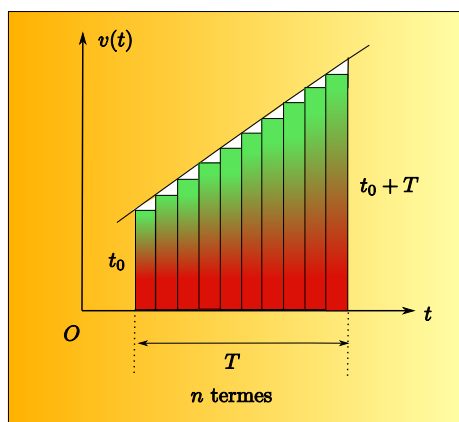


Fig. 33 - Découpage de la surface en petits rectangles

On considère que le départ du coureur se fait à l'instant  $t_0$ . Le premier rectangle a pour surface :

$$t_0 \frac{T}{n}. \quad (84)$$

Le deuxième rectangle a pour surface :

$$\left(t_0 + \frac{T}{n}\right) \frac{T}{n}. \quad (85)$$

Et ainsi de suite, le  $n$ -ième rectangle a pour surface :

$$\left(t_0 + n \frac{T}{n}\right) \frac{T}{n}. \quad (86)$$

A ces  $n$  rectangles, pour obtenir la distance parcourue par le coureur entre l'instant  $t_0$  et l'instant  $t_0 + T$ , il faut ajouter les  $n$  petits triangles se trouvant au-dessus de chacun des rectangles énumérés précédemment. Chacun de ces petits triangles a pour surface :

$$\frac{1}{2} \left(\frac{T}{n}\right)^2. \quad (87)$$

Donc, la distance parcourue par le coureur est la somme des rectangles et des  $n$  petits triangles :

$$\begin{aligned} d &= t_0 \frac{T}{n} + \left(t_0 + \frac{T}{n}\right) \frac{T}{n} + \dots + \left(t_0 + \frac{nT}{n}\right) \frac{T}{n} + \frac{n}{2} \left(\frac{T}{n}\right)^2 \\ &= \frac{(n+1)t_0 T}{n} + (1+2+\dots+n) \frac{T^2}{n^2} + \frac{n}{2} \left(\frac{T}{n}\right)^2 \\ &= \frac{nt_0 T}{n} + \frac{t_0 T}{n} + \frac{n(n+1) T^2}{2 n^2} + \frac{T^2}{2n} \\ &= t_0 T + \frac{t_0 T}{n} + \frac{T^2}{2} + \frac{T^2}{2n} + \frac{T^2}{2n} \\ &= t_0 T + \frac{t_0 T}{n} + \frac{T^2}{2} + \frac{T^2}{n} \end{aligned} \quad (88)$$

Maintenant, c'est l'étape la plus importante. Nous allons faire tendre  $n$  vers l'infini. Lorsque  $n$  tend vers l'infini, les termes ayant un  $n$  au dénominateur, et pas de  $n$  au numérateur tendent vers 0. Donc lorsque  $n$  tend vers l'infini, on a :

$$\lim_{n \rightarrow \infty} d = t_0 T + \frac{T^2}{2}. \quad (89)$$

Regardons maintenant quelle est l'expression de  $d$  si nous négligeons les petits triangles. Dans ce cas  $d$  devient :

$$\begin{aligned} d &= t_0 \frac{T}{n} + \left(t_0 + \frac{T}{n}\right) \frac{T}{n} + \dots + \left(t_0 + \frac{nT}{n}\right) \\ &= \frac{(n+1)t_0 T}{n} + (1+2+\dots+n) \frac{T^2}{n^2} \\ &= \frac{nt_0 T}{n} + \frac{t_0 T}{n} + \frac{n(n+1) T^2}{2 n^2} \\ &= t_0 T + \frac{t_0 T}{n} + \frac{T^2}{2} + \frac{T^2}{2n} \end{aligned} \quad (90)$$

Lorsque  $n$  tend vers l'infini, les termes possédant un  $n$  au dénominateur et pas de  $n$  au numérateur tendent vers 0, et la limite lorsque  $n$  tend vers l'infini est exactement la même expression que lorsque nous avons introduit les surfaces des petits triangles. C'est bien l'illustration que nous pouvons négliger ces petits triangles dans notre calcul. Et tout cela est parfaitement rigoureux, parce qu'il faut bien comprendre que lorsque nous manipulons des termes différentiels comme  $dt$ , nous travaillons virtuellement à la limite lorsque  $dt$  tend vers 0.

Pourquoi les petits triangles n'ont-ils aucun impact sur la somme à la limite lorsque  $dt$  tend vers 0? Pour le comprendre, revenons à l'expression de la distance  $ds$  parcourue par le coureur durant l'intervalle de temps  $dt$ , lorsque nous ne négligeons pas le petit triangle blanc :

$$ds = t dt + \frac{1}{2}(dt)^2. \quad (91)$$

Cette expression comporte deux termes : l'un est de l'ordre de grandeur de  $dt$ , l'autre est de l'ordre de grandeur de  $(dt)^2$ . Je parle d'ordre de grandeur, parce que nous savons que lorsque  $dt$  est très petit, l'expression  $(dt)^2$  est largement beaucoup plus petite. Et d'ailleurs, nous avons vu, lorsque nous étudions les voisinages, que lorsque  $dt$  tend vers 0,  $(dt)^2$  tend vers 0 beaucoup plus vite que  $dt$ . C'est la raison pour laquelle ce terme  $\frac{1}{2}(dt)^2$  (qui représente la surface du petit triangle blanc) est négligé dans les calculs ; c'est parce qu'il est infiniment plus petit que  $\frac{1}{2}(dt)^2$  est par conséquent négligeable devant  $t dt$  lorsque  $dt$  tend vers 0.

BEATRIX : C'est limpide! C'est magique!

## 9 Scène I-9 : Les vecteurs

EURISTIDE : Le domaine des mathématiques que nous avons abordé avec les fonctions, leurs dérivées, les intégrales et les primitives, les différentielles et le calcul différentiel, s'appelle **l'analyse**. C'est un nom plutôt bien adapté, puisque ces outils permettent précisément aux physiciens d'analyser le comportement d'un système physique.

Nous allons aborder maintenant un domaine tout différent parmi les outils mathématiques. Nous aurons besoin très souvent tout au long de cette conversation, d'un concept mathématique permettant de modéliser la notion de direction, de sens et d'amplitude. Il y a de nombreuses grandeurs qui possèdent ces trois caractéristiques.

BEATRIX : Oui, je pense tout de suite aux forces : lorsque je pousse de la main le verre posé sur la table, je lui applique une force. Cette force possède une direction (mettons parallèlement au bord droit de la table), un sens (mettons vers la fenêtre qui se trouve en face de moi), et une amplitude (ou une intensité) qui représente la puissance avec laquelle je pousse le verre.

EURISTIDE : La vitesse est également une grandeur possédant ces trois caractéristiques. Ces concepts de force ou de vitesse solidement ancrés dans notre intuition, sont modélisés par un objet que nous appelons **un vecteur**. Un vecteur est donc une entité possédant les trois caractéristiques : une direction, un sens et une intensité (que l'on appelle **module** dans le cas d'un vecteur.) La direction est représentée par une droite supportant le vecteur. Le module est représenté par la longueur d'un segment sur cette droite porteuse. Le sens est représenté par une flèche placée à l'extrémité du segment.

Ainsi, si nous considérons deux points  $A$  et  $B$  dans l'espace à deux dimensions d'un plan, nous pouvons tracer un vecteur  $\vec{v} = \overrightarrow{AB}$ , dont la direction sera la droite  $(AB)$ , le sens sera dirigé de  $A$  vers  $B$ , et le module sera la longueur du segment  $[AB]$ . Les vecteurs sont représentés par une expression ou une lettre surmontée d'une flèche. Nous pouvons représenter cela dans le plan :

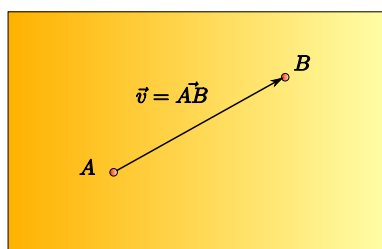
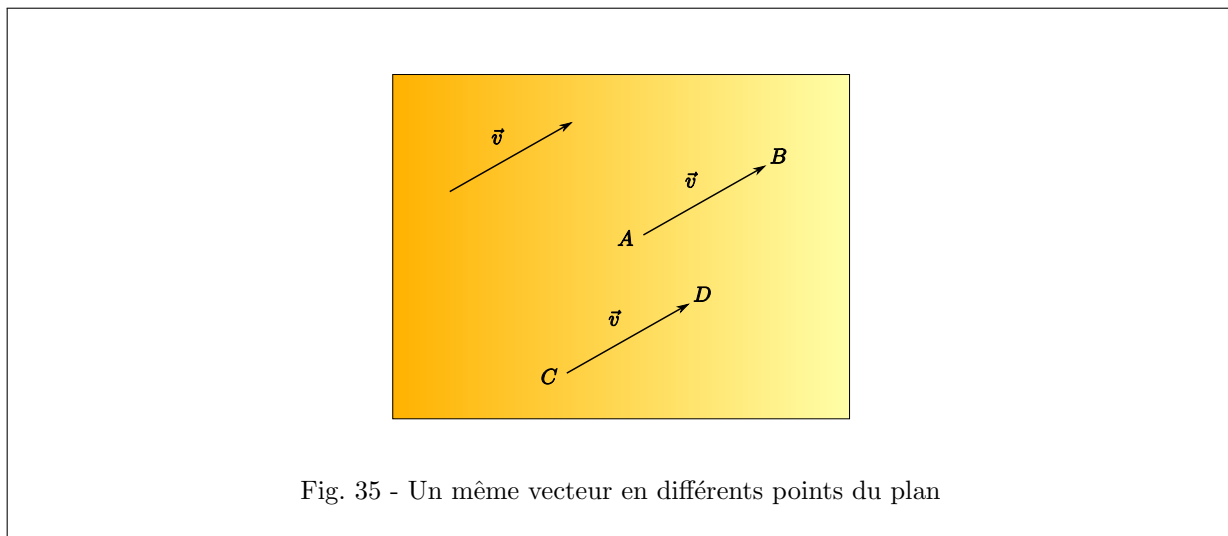


Fig. 34 - Le vecteur AB

Il faut considérer les vecteurs comme des entités qui peuvent "s'accrocher" en quelque sorte à des points :



Ainsi, je peux accrocher le vecteur  $\vec{v}$  de la figure ci-dessus au point  $A$  et au point  $C$ . Les vecteurs  $\vec{v}$ ,  $\overrightarrow{AB}$  et  $\overrightarrow{CD}$  sont égaux. Il s'agit de la représentation d'un seul et même vecteur accroché à plusieurs points. C'est exactement ce que nous faisons quand nous appliquons une force à un objet matériel.

**MATHINE** : En mathématiques et en physique, on utilise beaucoup la notion de **vecteur unitaire**. Un vecteur unitaire est un vecteur dont le module est la longueur 1 du système de coordonnées. D'où son nom de vecteur unitaire. Un tel vecteur est utilisé souvent pour représenter une direction et un sens, lorsque le module du vecteur n'a pas de signification physique particulière. Dans ce cas, nous n'avons pas besoin du module, et donc nous le positionnons à 1 arbitrairement. Les vecteurs unitaires sont aussi très employés pour représenter les axes de coordonnées d'un système de coordonnées. Dans ce cas, en dimension 3 par exemple, on les note  $\vec{i}$ ,  $\vec{j}$ ,  $\vec{k}$ , ou bien  $\vec{e}_1$ ,  $\vec{e}_2$ ,  $\vec{e}_3$ .

Par exemple, dans un plan (espace de dimension 2), on aura :

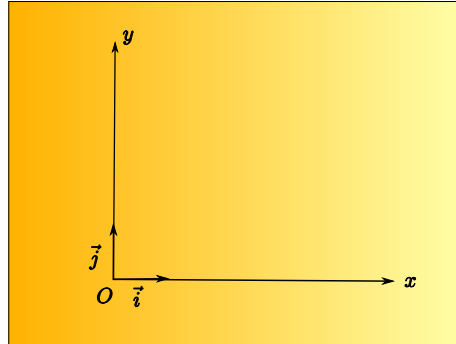


Fig. 36 - Les vecteurs de coordonnées en 2 dimensions

Et pour un espace de dimension 3, nous aurons :

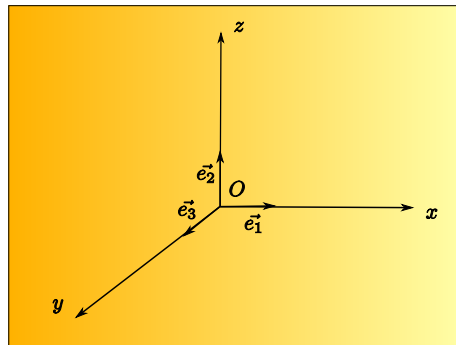


Fig. 37 - Les vecteurs de coordonnées en 3 dimensions

BEATRIX : Je crois qu'on peut faire des additions ou des multiplications avec les vecteurs, non ?

MATHINE : Oui, les vecteurs peuvent s'additionner. Le principe est simple : je prends deux vecteurs  $\vec{u}$  et  $\vec{v}$ . Dans un système de coordonnées, je peux représenter un vecteur au moyen de ses coordonnées.



Regardons ce qu'il en est dans le plan :

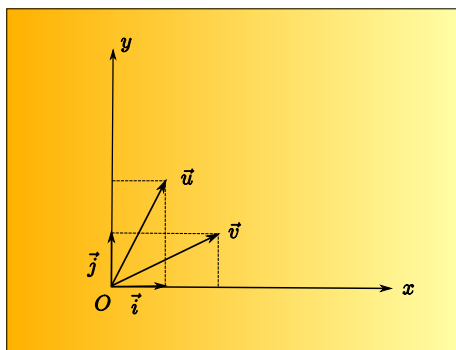


Fig. 38 - Les vecteurs  $u$  et  $v$  dans le plan

Comme nous avons vu que les vecteurs peuvent être accrochés sur n'importe quel point du plan, on choisit de les accrocher à l'origine du système de coordonnées pour obtenir leurs coordonnées.

Alors, les coordonnées du vecteur seront les coordonnées du point se trouvant à l'extrémité du côté de la flèche du vecteur. Ainsi, sur la figure ci-dessus, le vecteur  $\vec{u}$  a pour coordonnées  $(1, 2)$  et le vecteur  $\vec{v}$  a pour coordonnées  $(2, 1)$ .

Comme nous l'avons fait pour les points d'un espace de dimension  $n$ , on peut définir des vecteurs dans l'espace de dimension  $n$  qui sont assimilés à des  $n$ -uplets  $(x_1, x_2, \dots, x_n)$ .

**BEATRIX** : Je comprends. Maintenant que nous avons des coordonnées pour les vecteurs, nous allons pouvoir calculer des additions, des multiplications, etc. de vecteurs.

**MATHINE** : Oui. On va regarder de près ce qu'est l'addition de deux vecteurs. La définition mathématique, c'est que si  $\vec{u}$  a pour coordonnées  $(x_1, y_1)$  et si  $\vec{v}$  a pour coordonnées  $(x_2, y_2)$ , alors le vecteur  $\vec{u} + \vec{v}$  a pour coordonnées :

$$(x_1 + x_2, y_1 + y_2). \quad (92)$$

**EURISTIDE** : C'est la définition mathématique. Maintenant, regardons ce que cela signifie avec nos deux vecteurs  $\vec{u}(1, 2)$  et  $\vec{v}(2, 1)$ .

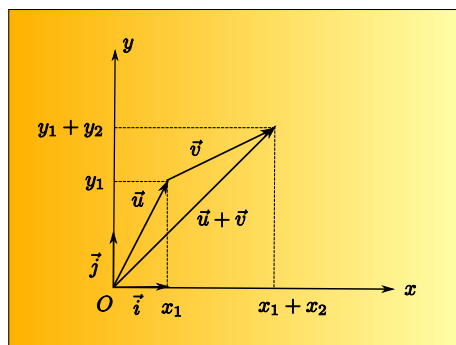


Fig. 39 - La somme de deux vecteurs

Pour additionner le vecteur  $\vec{v}$  au vecteur  $\vec{u}$ , nous additionnons les coordonnées  $x$  d'une part, et les coordonnées  $y$  d'autre part. Graphiquement parlant, cela revient à accrocher le vecteur  $\vec{v}$  à l'extrémité du vecteur  $\vec{u}$  comme cela est indiqué sur la figure ci-dessus. Ainsi, le vecteur  $\vec{u} + \vec{v}$  a pour coordonnées  $(3, 3)$ . On voit ainsi quelque chose de très important pour l'addition des vecteurs : pour additionner deux vecteurs, il ne faut pas croire qu'il suffit d'additionner leurs modules. On voit sur la figure précédente que le mécanisme d'addition tient compte des directions des deux vecteurs. Et si les vecteurs ne sont pas de même direction, le module de leur somme est inférieur à la somme de leurs modules. C'est ce qu'on voit très clairement sur la figure précédente.

On retrouve ce résultat intuitivement en pensant aux forces. Imaginons deux déménageurs qui tirent une grosse caisse de bois posée au sol. Si les deux déménageurs tirent dans la même direction, et dans le même sens, leurs forces vont s'additionner en intensité. S'ils tirent dans des directions différentes, nous savons bien que la résultante de leurs efforts ne sera pas la somme de leurs forces, mais sera plus petite, voire nulle s'ils tirent exactement de la même façon dans une même direction, mais dans des sens opposés.

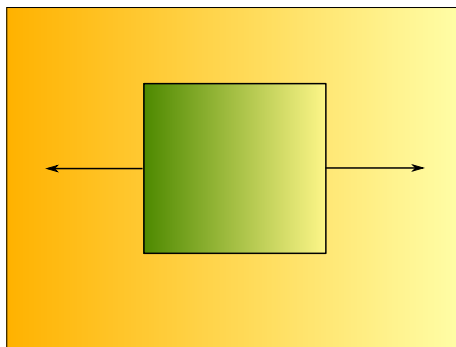


Fig. 40 - Illustration de deux vecteurs opposés

MATHINE : On peut donc additionner des vecteurs, comme nous l'avons vu. On peut aussi les multiplier par un nombre. Par exemple, si le vecteur  $\vec{u}$  a pour coordonnées  $(x_1, x_2)$ , le vecteur  $2\vec{u}$  aura pour coordonnées  $(2x_1, 2x_2)$ .

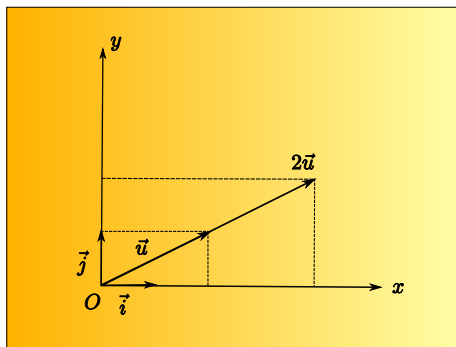


Fig. 41 - La multiplication d'un vecteur par un nombre

En associant aux vecteurs l'addition et la multiplication par un nombre (on dit par un scalaire), on obtient une structure mathématique qu'on appelle **espace vectoriel**. Un espace vectoriel possède un certain nombre de dimensions comme nous l'avons vu. On parle donc d'espace vectoriel de dimension 1, de dimension 2, de dimension  $n$ .

## 10 Scène I-10 : Le calcul matriciel

EURISTIDE : Je crois qu'il est temps de parler de notre dernière notion de base : les matrices. Nous avons beaucoup parlé de coordonnées et nous avons vu les différents calculs réalisés à partir de ces coordonnées, notamment quand nous avons voulu changer de système de coordonnées.

Nous allons introduire un outil appelé **calcul matriciel**, qui permet de représenter facilement les coordonnées, leurs transformations et les différents calculs que nous pouvons effectuer sur ces coordonnées.

MATHINE : Oui, les mathématiciens ont pour habitude de représenter un couple de coordonnées  $(x, y)$  dans un plan par exemple, sous forme d'un tableau :

$$\begin{pmatrix} x \\ y \end{pmatrix}. \quad (93)$$

Un tel tableau est appelé **matrice** à deux lignes et une colonne. Nous avons parlé, il y a quelque temps déjà, des transformations de coordonnées.

Par exemple, si nous passons des coordonnées :

$$\begin{pmatrix} x \\ y \end{pmatrix}, \quad (94)$$

aux coordonnées :

$$\begin{pmatrix} x' \\ y' \end{pmatrix}, \quad (95)$$

au moyen de la transformation suivante :

$$x' = 2(x - y) \quad (96)$$

$$y' = 2\sqrt{2}(y - 1), \quad (97)$$

les mathématiciens utilisent une représentation matricielle pour noter une telle transformation :

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 2 & -2 \\ 0 & 2\sqrt{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ -2\sqrt{2} \end{pmatrix} \quad (98)$$

Cette notation se traduit de la façon suivante :

$$x' = 2 \times x + (-2) \times y + 0 \quad (99)$$

$$y' = 0 \times x + (2\sqrt{2}) \times y + (-2\sqrt{2}), \quad (100)$$

ce qui correspond bien à l'expression initiale de la transformation.

Imaginons maintenant plus généralement une transformation sous la forme :

$$x' = a_1x + b_1y \quad (101)$$

$$y' = a_2x + b_2y. \quad (102)$$

Cette transformation va s'écrire sous forme matricielle de la façon suivante :

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \quad (103)$$

La matrice :

$$\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}, \quad (104)$$

est appelée **matrice de transformation** des coordonnées. Il s'agit d'une matrice carrée  $2 \times 2$ .

Si on décrit une transformation des coordonnées dans un espace à  $n$  dimensions, on aura une matrice carrée  $n \times n$  :

$$\begin{pmatrix} x'_1 \\ x'_2 \\ \dots \\ x'_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}. \quad (105)$$

On peut effectuer divers calculs avec les matrices.

L'addition de deux matrices peut se faire à condition que les deux matrices aient les mêmes caractéristiques (même nombre de lignes et même nombre de colonnes). Il suffit alors d'ajouter les termes des deux matrices :

$$\boxed{\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} + \begin{pmatrix} c_1 & d_1 \\ c_2 & d_2 \end{pmatrix} = \begin{pmatrix} a_1 + c_1 & b_1 + d_1 \\ a_2 + c_2 & b_2 + d_2 \end{pmatrix}.} \quad (106)$$

La multiplication par un scalaire  $\alpha$  s'écrit :

$$\boxed{\alpha \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = \begin{pmatrix} \alpha a_1 & \alpha b_1 \\ \alpha a_2 & \alpha b_2 \end{pmatrix}.} \quad (107)$$

Et nous avons vu précédemment une multiplication de matrices dans la présentation des transformations de coordonnées :

$$\boxed{\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_1x + b_1y \\ a_2x + b_2y \end{pmatrix}.} \quad (108)$$

La technique de multiplication des matrices consiste à "combiner" les lignes de la première matrice avec les colonnes de la seconde matrice. Il est donc indispensable, pour qu'une multiplication soit faisable, que la seconde matrice ait autant de lignes que la première matrice possède de colonnes. La matrice résultant du calcul de la multiplication aura autant de colonnes que la seconde matrice, et autant de lignes que la première matrice.

Par exemple, pour la multiplication d'une matrice à deux lignes et deux colonnes par une matrice de deux lignes et deux colonnes, nous obtenons une matrice de deux lignes et deux colonnes :

$$\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \begin{pmatrix} c_1 & d_1 \\ c_2 & d_2 \end{pmatrix} = \begin{pmatrix} a_1c_1 + b_1c_2 & a_1d_1 + b_1d_2 \\ a_2c_1 + b_2c_2 & a_2d_1 + b_2d_2 \end{pmatrix}. \quad (109)$$

**BEATRIX** : D'un côté, c'est compliqué. D'un autre côté, il faut reconnaître que cette notation simplifie la présentation des changements de coordonnées. C'est pour cela qu'elle existe, je suppose.

**MATHINE** : C'est vrai. Mais on reviendra à de multiples reprises sur ces notations matricielles, en particulier lorsque nous parlerons des tenseurs et du calcul tensoriel. En attendant, un petit schéma simple peut montrer comment s'opère la multiplication de matrices.

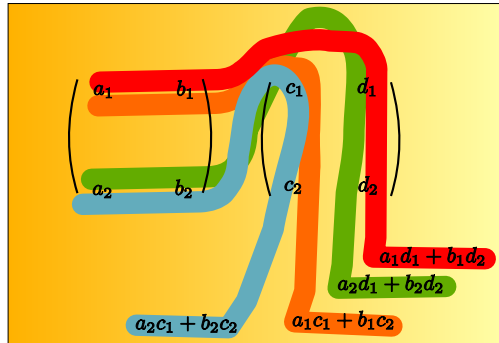


Fig. 42 - Schéma du produit de deux matrices

Nous introduisons également le **déterminant d'une matrice** ; pour une matrice de rang  $2 \times 2$ , le déterminant s'écrit :

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1. \quad (110)$$

Pour une matrice de rang  $3 \times 3$  :

$$\begin{aligned} & \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \\ &= a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & c_1 \\ b_3 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} \\ &= a_1 b_2 c_3 - a_1 b_3 c_2 - a_2 b_1 c_3 + a_2 b_3 c_1 \\ &+ a_3 b_1 c_2 - a_3 b_2 c_1. \end{aligned} \quad (111)$$

---

## Troisième partie

# Acte II - La cinématique

## 11 Scène II-1 : Les mouvements des corps

EURISTIDE : Nous avons maintenant en mains les outils indispensables pour la suite du programme. Bien sûr, nous verrons plus tard d'autres outils mathématiques, au fur et à mesure que nous progresserons dans nos discussions. Mais les bases sont déjà acquises.

BEATRIX : Qu'allons-nous voir maintenant ?

EURISTIDE : Nous allons poursuivre notre parcours d'initiation vers la relativité générale. Et nous allons passer aux bases de la mécanique, en commençant par la science du mouvement : la cinématique.

Pour étudier le mouvement des corps, on ramène en pensée les corps à un ensemble de points. On peut considérer un point comme une petite partie du corps, dont les dimensions sont infiniment petites par rapport au corps entier et par rapport au parcours de ce point.

Ainsi, nous pourrions représenter tout corps comme un ensemble de points liés entre eux. Nous pourrions par conséquent déduire l'étude du mouvement des corps de celle du mouvement des points.

BEATRIX : J'imagine que le mouvement d'un point est plus facile à modéliser que celui d'un corps. On va donc commencer par l'étude la plus simple.

## 12 Scène II-2 : Le mouvement rectiligne du point

EURISTIDE : En effet. Considérons un point  $M$  se déplaçant sur une trajectoire rectiligne (une droite). Pour cela, on peut choisir un point particulier de la droite, que nous appellerons  $O$ , et on peut observer la façon dont la distance du point  $M$  évolue par rapport au point  $O$  qui va nous servir de repère.

MATHINE : Pour formaliser cela, on note la distance  $OM$  par le symbole  $y$ . Par conséquent, nous nous intéressons aux variations de  $y$  en fonction du temps. Nous notons la variable temps par le symbole  $t$ . Nous allons introduire une fonction  $f$  qui représentera les variations de  $x$  en fonction de  $t$ , ce qui s'écrit :

$$y = f(t). \tag{112}$$

EURISTIDE : Pour avoir une représentation claire du comportement du point  $M$ , il est utile de dessiner une représentation graphique de la fonction  $f$  comme suit :

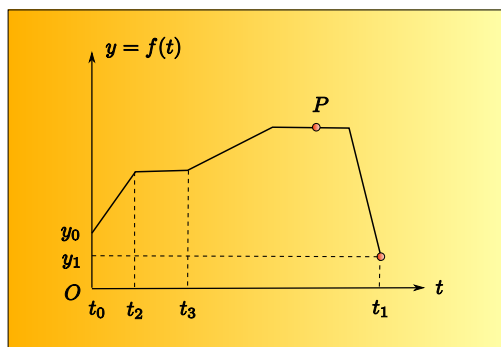


Fig. 43 - Représentation de la trajectoire d'un point

Cette représentation graphique nous apprend plusieurs choses concernant le mouvement du point  $M$  :

- Lorsque la courbe tracée est horizontale, le point est immobile, comme ici entre les instants  $t_2$  et  $t_3$ .
- Lorsque la courbe tracée croît, le point  $M$  s'éloigne de  $O$ .
- Lorsque la courbe tracée décroît, le point  $M$  se rapproche de  $O$ .
- Plus la courbe tracée croît rapidement, plus le point s'éloigne rapidement de  $O$ .
- La plus grande valeur  $P$  de  $x$  sur la courbe est atteinte pour l'endroit le plus éloigné du point  $M$  par rapport au point  $O$ .

BEATRIX : Ce que je retiens ici, c'est qu'on apprend beaucoup de l'observation de la représentation graphique d'une fonction. Il suffit d'interpréter d'un point de vue physique ce que signifient les variations et valeurs de la courbe sur le graphique.

### 13 Scène II-3 : La définition de la vitesse

EURISTIDE : Nous avons vu sur la courbe précédente que plus la courbe croissait rapidement, plus le point  $M$  s'éloignait rapidement de  $O$ . Nous retrouvons ici ce que nous avons vu à propos de la dérivée d'une fonction : la vitesse du point  $M$  est la dérivée de la fonction  $f$  en fonction du temps, ou aussi, de façon équivalente, la pente de la courbe de la fonction  $f$ .

MATHINE : Nous allons donc pouvoir décrire cette vitesse au moyen des outils que nous avons étudiés précédemment. Si nous appelons  $v(t)$  la vitesse du point  $M$ , nous pouvons écrire :

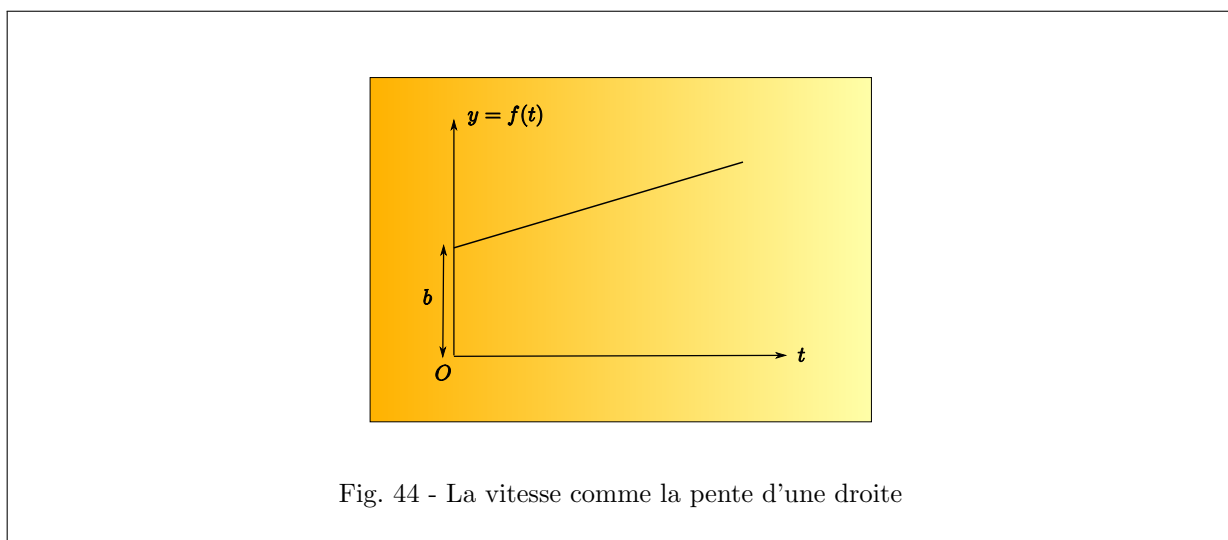
$$v(t) = \frac{df}{dt}. \quad (113)$$



Prenons l'exemple d'un point  $M$  qui se déplace sur la droite que nous avons définie suivant la fonction :

$$f(t) = at + b, \quad (114)$$

où les quantités  $a$  et  $b$  sont des valeurs constantes.



EURISTIDE : Nous pouvons interpréter aisément les constantes  $a$  et  $b$ .  $b$  est évidemment la position du point  $M$  par rapport à  $O$  lorsque  $t = 0$ .

BEATRIX : Ah oui, en effet :  $f(t) = b$  pour  $t = 0$ . Par ailleurs, je suppose que  $a$  représente la pente de la droite du graphique ci-dessus.

MATHINE : Oui, c'est exact. Et on peut retrouver ce résultat en calculant la dérivée de la fonction  $f(t)$  :

$$\frac{df(t)}{dt} = \frac{d(at + b)}{dt} = a. \quad (115)$$

C'est donc que  $a$  représente bien la vitesse du point  $M$  ou la pente de la droite sur le graphique ci-dessus.

## 14 Scène II-4 : La relation entre vitesse et distance

BEATRIX : Alors, en utilisant ce que nous avons appris tout à l'heure, on doit pouvoir retrouver la distance  $OM$  à tout instant à partir de la fonction vitesse  $v(t)$  et à partir de la position initiale du point.

EURISTIDE : Oui, la fonction  $f(t)$  est l'intégrale de la fonction  $v(t)$ . Je vais laisser à Mathine le soin de nous montrer cela.

MATHINE : Avec plaisir. Mais tu as presque tout dit. La distance  $d$  parcourue par le point  $M$  entre les instants  $t_0$  et  $t_1$  s'écrit :

$$d = \int_{t_0}^{t_1} v(t) dt. \quad (116)$$

EURISTIDE : Et comme nous l'avons vu, cette distance est représentée graphiquement comme la surface grise sur le graphique ci-dessous, représentant la fonction  $v(t)$  :

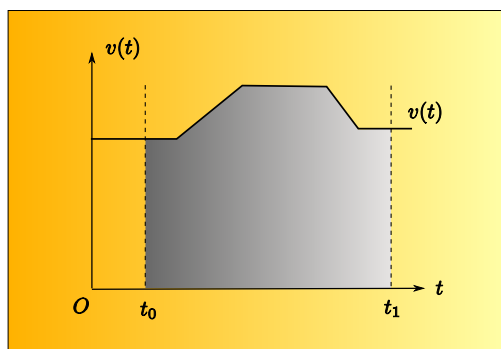


Fig. 45 - La distance est l'intégrale de la vitesse

La surface grise représente la distance parcourue par le point  $M$  entre les instants  $t_0$  et  $t_1$ .

## 15 Scène II-5 : La définition de l'accélération

BEATRIX : Nous avons représenté la fonction vitesse à plusieurs reprises. On voit que cette fonction vitesse peut croître et décroître plus ou moins rapidement. La vitesse de croissance de la vitesse, c'est l'accélération. La vitesse de décroissance de la vitesse, c'est la décélération, je crois.

EURISTIDE : Oui, nous devons d'ailleurs définir une fonction qui s'appelle **l'accélération**. Plus un point en mouvement augmente sa vitesse rapidement, plus son accélération est grande et positive. Si un point possède une vitesse constante, son accélération est nulle. Enfin, si un point possède une vitesse en diminution, il ralentit et son accélération est négative.

BEATRIX : Autrement dit, la décélération, c'est une accélération négative.

MATHINE : Exactement. L'accélération se calcule comme la dérivée par rapport au temps de la vitesse  $v(t)$ , donc si nous notons  $w(t)$  la fonction accélération, nous avons :

$$w(t) = \frac{dv(t)}{dt}. \quad (117)$$

Si on représente graphiquement une fonction vitesse :

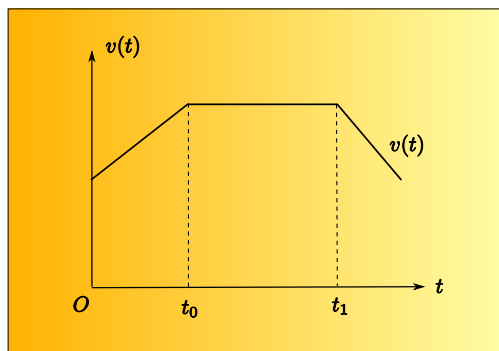


Fig. 46 - Représentation de la vitesse

Dans l'intervalle de temps  $[0, t_0]$ , la vitesse est croissante, donc l'accélération du point est positive. Dans l'intervalle de temps  $[t_0, t_1]$ , la vitesse est constante, donc l'accélération est nulle. Dans l'intervalle de temps  $t > t_1$ , la vitesse est décroissante, donc l'accélération est négative.

Nous pouvons représenter le graphe de la fonction accélération de ce point de la façon suivante :

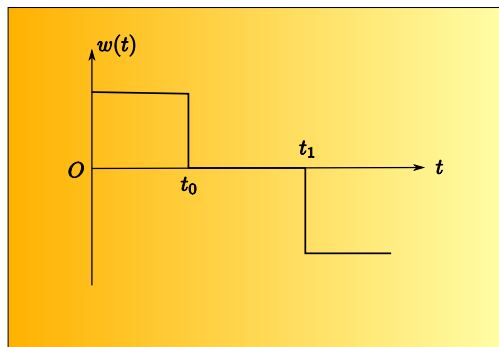


Fig. 47 - Représentation de l'accélération

## 16 Scène II-6 : Le mouvement d'un point

EURISTIDE : Jusqu'ici, nous avons étudié le mouvement d'un point sur une trajectoire très simple rectiligne. Il nous faut donc passer maintenant à un cas plus général, celui du mouvement d'un point dans l'espace.

MATHINE : La position d'un point dans l'espace est représentée par trois coordonnées qui dépendent d'un paramètre : le temps  $t$ . Les coordonnées du point en mouvement dans l'espace sont exprimées sous la forme de fonctions de la variable temps :

$$\begin{aligned}x &= x(t) \\y &= y(t) \\z &= z(t).\end{aligned}\tag{118}$$

EURISTIDE : Par conséquent, le mouvement du point dans l'espace est l'évolution de la position de ce point entre deux instants, mettons  $t_0$  et  $t_1$ . On peut représenter graphiquement ce mouvement par une courbe dans l'espace, comme suit. La courbe représentée ici est appelée **trajectoire** du point.

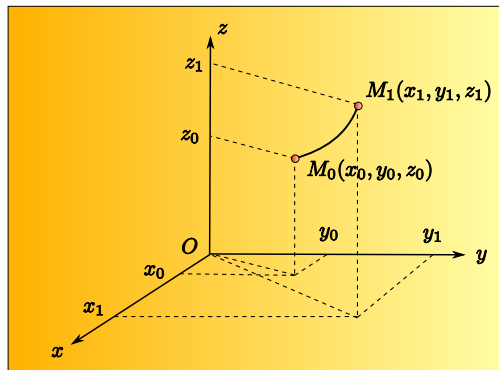


Fig. 48 - Trajectoire d'un point dans l'espace

**MATHINE** : Pour représenter la position du point à chaque instant, nous allons utiliser un vecteur, noté  $\vec{s}(t)$ , qui est le vecteur joignant l'origine  $O$  du système de coordonnées à la position du point sur sa trajectoire. Ce vecteur est appelé **vecteur position** du point.

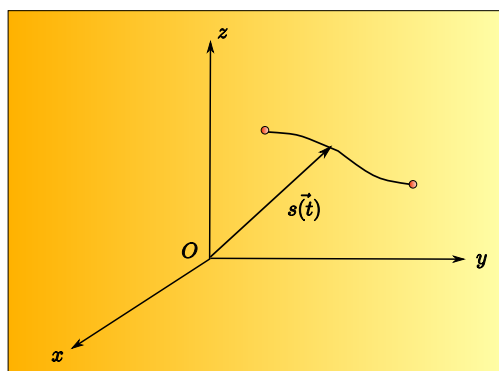


Fig. 49 - Le vecteur position

Avec cette convention,  $\overrightarrow{s(t)}$  possède les coordonnées suivantes :

$$s(t) \quad \begin{cases} x(t) \\ y(t) \\ z(t). \end{cases} \quad (119)$$

## 17 Scène II-7 : La vitesse d'un point

BEATRIX : Nous savons maintenant représenter et modéliser le mouvement quelconque d'un point dans l'espace. Il nous faut maintenant un moyen de représenter la vitesse de ce point. Comment fait-on ?

EURISTIDE : Tu te rappelles que tout à l'heure, nous avons défini la vitesse comme la dérivée par rapport au temps de la fonction définissant le mouvement du point. Ici, nous allons procéder de la même façon. Nous allons partir de la définition de la dérivée qui est le rapport d'une variation infinitésimale de la fonction mouvement du point par rapport à une variation infinitésimale du paramètre  $t$ .

Il se trouve que la fonction qui représente la position du point est maintenant un vecteur  $\overrightarrow{s(t)}$ .

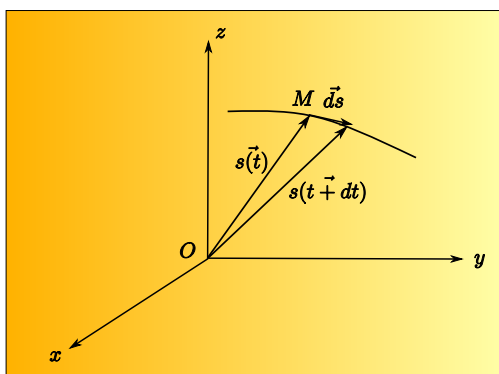


Fig. 50 - Interprétation de la différentielle du vecteur position

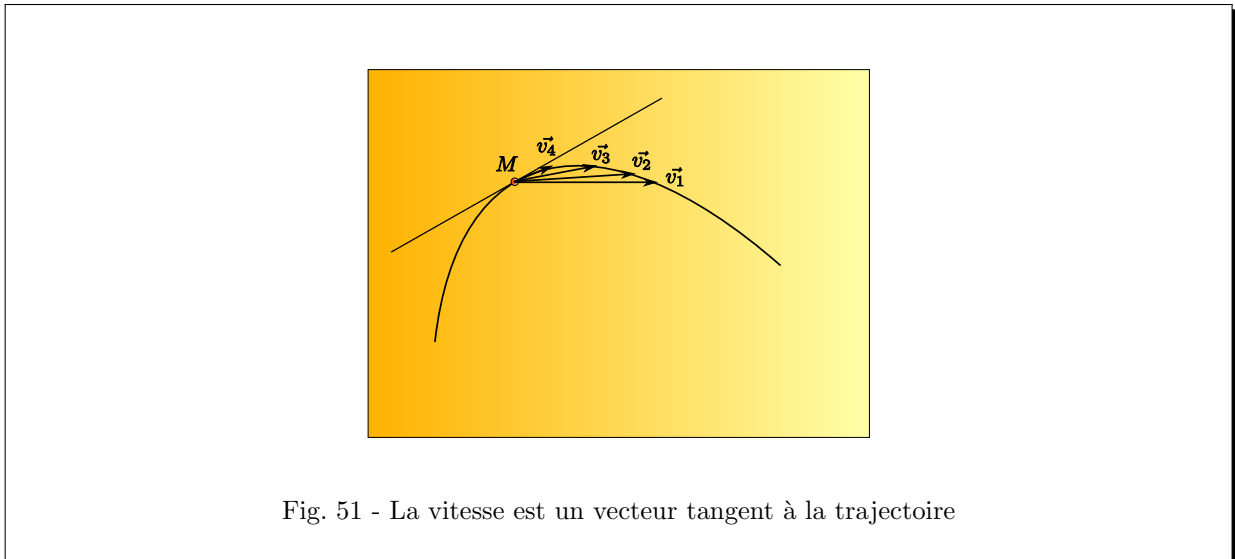
Faisons varier  $dt$  de façon infiniment petite. Le point sur trajectoire se déplace d'une position  $\overrightarrow{s(t)}$  au point  $M$  vers une position infiniment proche, à la position  $\overrightarrow{s(t+dt)}$ . L'écart entre les deux positions infiniment proches est :

$$\overrightarrow{s(t+dt)} - \overrightarrow{s(t)}. \quad (120)$$

Mais nous connaissons la règle d'addition des vecteurs. Si on appelle  $\overrightarrow{ds}$  le vecteur infiniment petit qui relie les deux points voisins de la trajectoire, nous avons :

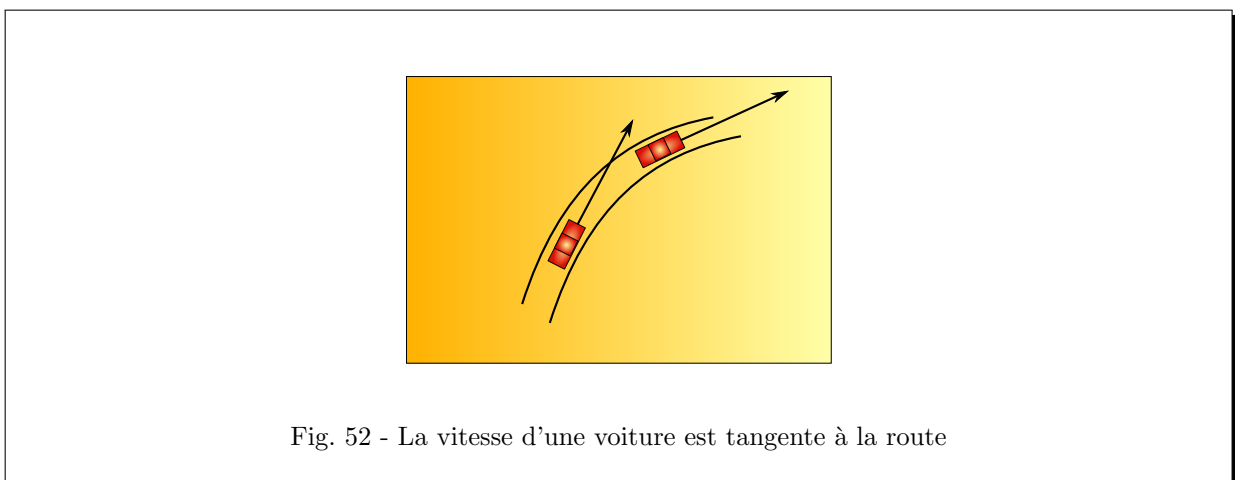
$$\overrightarrow{ds} = \overrightarrow{s(t+dt)} - \overrightarrow{s(t)}. \quad (121)$$

Comme  $dt$  tend vers 0, le vecteur  $\vec{ds}$  tend vers un vecteur de module nul, mais sa direction, quant à elle, tend à se rapprocher d'une droite qui est tangente à la trajectoire au point  $M$ , comme nous le voyons sur la figure ci-dessous pour les vecteurs  $\vec{v}_1$ ,  $\vec{v}_2$ ,  $\vec{v}_3$  et  $\vec{v}_4$  :



La vitesse du point sera exprimée comme le rapport de  $\vec{ds}$  à la variation infinitésimale du temps  $dt$ . Par conséquent, la vitesse pourra être représentée par un vecteur qui sera tangent à la trajectoire au point  $M$  et dont le module sera la grandeur de cette vitesse.

**BEATRIX** : Je comprends maintenant. Nous avons parlé de la voiture dans un virage au début de notre conversation. A chaque instant, nous disions qu'elle semblait se diriger vers l'extérieur de la route suivant une tangente à la courbe de la route.



C'est bien l'illustration que ce vecteur vitesse est tangent à la trajectoire.

MATHINE : C'est exactement cela. La vitesse du point s'exprime au moyen du vecteur vitesse :

$$\boxed{\vec{v}(t) = \frac{d\vec{s}(t)}{dt}}. \quad (122)$$

BEATRIX : Mais comment allons-nous calculer cette dérivée d'un vecteur. C'est un peu nouveau comme notion, non ?

MATHINE : Oui, c'est nouveau. Mais en réalité, nous avons tous les outils pour la calculer.

Pour cela, on va exprimer le vecteur  $\vec{s}(t)$  en fonction de ses coordonnées dans un système de coordonnées  $(O, \vec{i}, \vec{j}, \vec{k})$  :

$$\vec{s}(t) = x(t)\vec{i} + y(t)\vec{j} + z(t)\vec{k}. \quad (123)$$

Nous pouvons maintenant prendre la dérivée de cette expression :

$$\frac{d\vec{s}(t)}{dt} = \frac{d}{dt}(x(t)\vec{i} + y(t)\vec{j} + z(t)\vec{k}). \quad (124)$$

Il se trouve que la dérivée d'une somme de fonctions est la somme des dérivées des fonctions, donc :

$$\frac{d\vec{s}(t)}{dt} = \frac{d(x(t)\vec{i})}{dt} + \frac{d(y(t)\vec{j})}{dt} + \frac{d(z(t)\vec{k})}{dt}. \quad (125)$$

Il se trouve également que la dérivée d'une fonction multipliée par une constante est égale à cette constante multipliée par la dérivée de la fonction, donc :

$$\frac{d\vec{s}(t)}{dt} = \frac{dx(t)}{dt}\vec{i} + \frac{dy(t)}{dt}\vec{j} + \frac{dz(t)}{dt}\vec{k}. \quad (126)$$

EURISTIDE : Nous avons donc trouvé l'expression de la vitesse, qui est en fait un vecteur dont les coordonnées sont les dérivées des fonctions coordonnées du vecteur position. Rien de surprenant, en fait.

BEATRIX : Oui, c'est ce qu'on pouvait attendre. Quid de l'accélération ?

## 18 Scène II-8 : L'accélération d'un point

MATHINE : Il n'y a pas de surprise non plus. L'accélération va s'exprimer comme la dérivée du vecteur vitesse.

$$\vec{w}(t) = \frac{d\vec{v}(t)}{dt}. \quad (127)$$



Pour pouvoir calculer cela, nous allons introduire une nouvelle notation ; considérons une fonction  $f(t)$  donc nous connaissons la dérivée :

$$v(t) = \frac{df(t)}{dt}. \quad (128)$$

Nous pouvons considérer la dérivée de cette fonction  $v(t)$  à son tour :

$$\frac{dv(t)}{dt} = \frac{d\left(\frac{df(t)}{dt}\right)}{dt}. \quad (129)$$

Pour simplifier la notation, nous introduisons une notion de dérivée seconde, qui s'écrit :

$$\frac{d\left(\frac{df(t)}{dt}\right)}{dt} = \frac{d^2 f}{dt^2}. \quad (130)$$

Cette expression signifie qu'on dérive deux fois successivement la fonction  $f$  par rapport au temps  $t$ . En utilisant cette convention, on obtient :

$$\vec{w}(t) = \frac{d\vec{v}(t)}{dt} = \frac{d^2 \vec{s}(t)}{dt^2} = \frac{d^2 x}{dt^2} \vec{i} + \frac{d^2 y}{dt^2} \vec{j} + \frac{d^2 z}{dt^2} \vec{k}. \quad (131)$$

EURISTIDE : Ceci étant établi, on peut se demander maintenant quelle est la direction de ce vecteur accélération. Pour cela, on représente sur la figure ci-dessous la variation infinitésimale d'un vecteur vitesse :

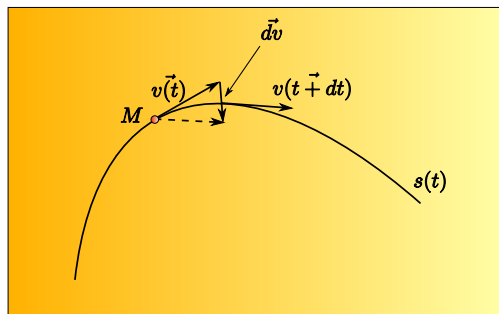


Fig. 53 - Représentation de la différentielle de la vitesse

Comme nous l'avons fait précédemment, lorsqu'on fait varier  $t$  de façon infinitésimale de  $t$  en  $t + dt$ , le vecteur  $\vec{v}(t)$  varie infiniment peu pour devenir  $\vec{v}(t + dt)$ .

La différence entre  $\vec{v}(t + dt)$  et  $\vec{v}(t)$  est un petit vecteur, comme indiqué sur la figure,  $\vec{dv}$ , qui tend vers 0 lorsque  $dt$  tend vers 0, mais dont la direction tend à se rapprocher d'une direction déterminée par rapport à la tangente à la courbe en  $M$ .

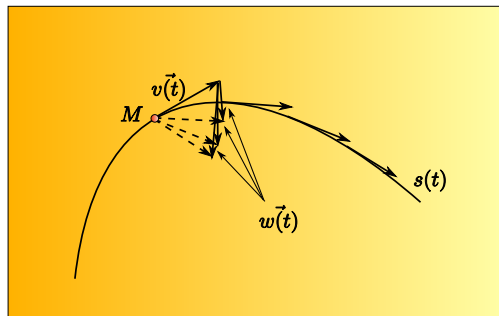


Fig. 54 - Évolution du vecteur accélération

Par exemple, pour un mouvement accéléré le long d'une ligne droite, le vecteur accélération sera dirigé le long de cette même droite. Pour un mouvement circulaire dont la vitesse a un module constant, mouvement un peu similaire localement au point représenté sur la figure précédente, le vecteur accélération est orthogonal à la trajectoire.

---

## Quatrième partie

# Acte III - La mécanique newtonienne

## 19 Scène III-1 : Les interactions des corps

BEATRIX : Maintenant que j'ai compris comment représenter et modéliser le mouvement d'un point, je serais curieuse de comprendre par quels mécanismes les points matériels se mettent en mouvement.

EURISTIDE : C'est l'objet de la science que l'on appelle la dynamique. Les lois de la dynamique ont été étudiées par Newton et ont pour but d'établir les lois liant les mouvements des corps aux causes ayant produit ces mouvements.

BEATRIX : On a parlé des forces tout à l'heure. Ne serait-ce pas elles qui constituent ces causes ?

EURISTIDE : Oui, pour Newton les corps sont soumis à des forces telles que l'attraction, la gravitation, etc. Lorsque l'ensemble des forces soumise à un corps sont en équilibre (c'est-à-dire lorsqu'elles s'annulent), le corps ne modifie pas sa vitesse initiale. Sa vitesse est donc constante. Lorsque les forces soumise à un corps ne sont pas en équilibre, le corps voit sa vitesse modifiée en fonction de la direction, du sens et de l'intensité de la somme résultante de ces forces.

Le poids est un exemple de force. Le poids appliqué à un corps entraîne différents effets possibles : le basculement du plateau d'une balance, la chute d'un corps, la déformation d'un ressort auquel est suspendu ce corps.

BEATRIX : On avait dit qu'une force était représentée par un vecteur en mathématiques. On pourra donc utiliser ce formalisme dans nos calculs de théorie de la dynamique.

EURISTIDE : Oui, une force est représentée tout naturellement par un vecteur, possédant une intensité qui est le module du vecteur, possédant une direction qui est la direction du vecteur, possédant un sens qui est le sens du vecteur.

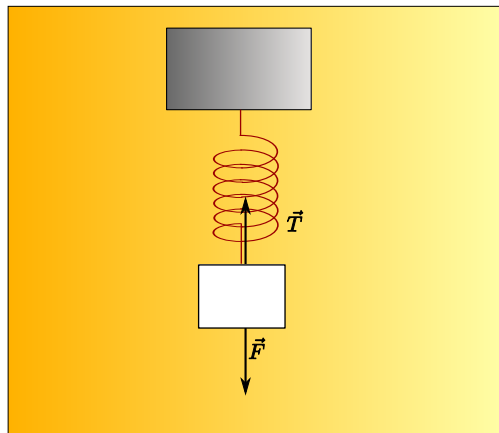


Fig. 55 - Illustration des forces

Pour qu'un corps soit en équilibre (ce qui veut dire que sa vitesse est constante, nulle ou non nulle), il faut et il suffit que la somme des forces qui lui sont appliquées soit nulle.

BEATRIX : Mais, si je pose un livre sur une table, je ne vois qu'une seule force appliquée à ce livre, c'est son poids. Ce poids n'est pas nul. Et pourtant, le livre est bien en équilibre, puisqu'il possède une vitesse nulle.

EURISTIDE : Lorsque tu dis que le livre est posé sur la table, cela signifie que la table exerce une réaction sur le livre qui s'appuie par son poids sur la table.

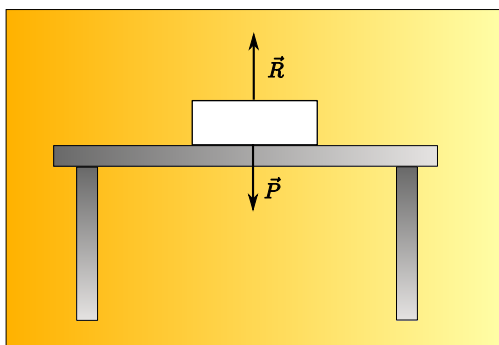


Fig. 56 - L'équilibre d'un corps posé

Cette force de réaction, par nature, est exactement d'intensité égale au poids du livre et de sens opposé. Par conséquent, la somme des forces appliquées au livre (qui sont  $\vec{P}$  et  $\vec{R}$ ) est :

$$\vec{P} + \vec{R} = \vec{0}. \quad (132)$$

Donc, elle est bien nulle, ce qui est cohérent avec l'observation que nous faisons, où nous constatons que le livre est en équilibre sur la table.

MATHINE : D'une manière plus générale, un corps soumis aux forces  $\vec{F}_1, \vec{F}_2, \dots, \vec{F}_n$  est en équilibre lorsque :

$$\boxed{\vec{F}_1 + \vec{F}_2 + \dots + \vec{F}_n = \vec{0}.} \quad (133)$$

## 20 Scène III-2 : La première loi de Newton

BEATRIX : Autrement dit, c'est la somme des forces appliquées à un corps qui détermine le mouvement de ce corps. Si elles sont nulles, le corps reste au repos. Si elles ne sont pas nulles, le corps n'est plus au repos.

MATHINE : Ce que tu viens d'énoncer est la première loi de Newton, qui est exprimée habituellement de la façon suivante :

**Tout corps se trouve à l'état de repos ou de mouvement rectiligne uniforme tant que des forces appliquées ne provoquent pas des variations de cet état.**

EURISTIDE : Autrement dit, un corps qui n'est pas soumis à des forces sera au repos ou en mouvement rectiligne à une vitesse constante. Dès lors qu'on lui appliquera un ensemble de forces dont la somme n'est pas nulle, il modifiera sa vitesse et sa trajectoire.

## 21 Scène III-3 : La seconde loi de Newton

BEATRIX : D'accord, nous savons que les forces modifient le mouvement d'un corps. Mais la question, c'est : comment ? Autrement dit, quel est l'impact d'une force sur la vitesse d'un corps.

MATHINE : La réponse à cette question a été donnée par la seconde loi de Newton, énoncée comme suit :

**La variation de la quantité de mouvement est proportionnelle à la force agissante et s'effectue suivant la direction de la droite de l'action de cette force.**

BEATRIX : Que cela veut-il dire ?

EURISTIDE : Ce que Newton appelle la quantité de mouvement, c'est le produit de la masse par la vitesse :

$$m \vec{v}, \quad (134)$$

si  $m$  est la masse du corps, et  $\vec{v}$  est son vecteur vitesse.

Intuitivement, la quantité de mouvement représente une quantité proportionnelle à la vitesse d'autant plus grande que le corps est lourd. C'est donc une mesure de l'élan d'un corps de masse  $m$  lancé à la vitesse  $v$ .

MATHINE : Donc, la quantité de mouvement est un vecteur, qu'on note souvent par la lettre  $p$  :

$$\vec{p} = m \vec{v}. \quad (135)$$

EURISTIDE : Ce que dit la seconde loi de Newton, c'est que la variation dans le temps de ce vecteur  $\vec{p}$  est proportionnelle à la force agissant sur le corps. Et la force agissant sur le corps est égale à la somme des forces  $\vec{F}$  s'appliquant au corps.

MATHINE : Ce qui s'écrit :

$$\boxed{\vec{F} = \frac{d}{dt}(m \vec{v})}. \quad (136)$$

BEATRIX : Intuitivement, c'est compréhensible. Cela signifie que la force  $\vec{F}$  provoque une variation de la vitesse, et c'est bien ce à quoi on s'attendait.

D'ailleurs, si la force est nulle, cela signifie que :

$$\frac{d}{dt}(m \vec{v}) = \vec{0}, \quad (137)$$

donc que  $m \vec{v}$  est constante, et par conséquent, si la masse  $m$  est constante, alors  $\vec{v}$  est constante, autrement dit, le corps est au repos. Tout cela se tient !

MATHINE : Pour un corps dont la masse est constante, on peut écrire :

$$\vec{F} = m \frac{d\vec{v}}{dt} = m\vec{w}, \quad (138)$$

si on convient de noter par  $\vec{w}$  l'accélération du corps, qui est la dérivée de la vitesse du corps par rapport au temps.

Si le corps ne possède pas une masse constante (par exemple, une fusée dont la masse de carburant diminue lors de son voyage), alors :

$$\vec{F} = \frac{dm}{dt} \vec{v} + m\vec{w}. \quad (139)$$

## 22 Scène III-4 : La troisième loi de Newton

EURISTIDE : La troisième loi de Newton parle de l'action et de la réaction. Si je tiens dans ma main un poids de balance, le poids applique sur ma main une force proportionnelle à sa masse. La troisième loi de Newton exprime le fait que notre main applique au poids, en réaction, une force opposée à cette force appliquée par le poids de la balance à notre main. Cette force est appelée **réaction**.

BEATRIX : Autrement dit, si j'applique une force sur un corps, ce corps produit une réaction opposée à cette force.

EURISTIDE : C'est ça. La troisième loi de Newton s'exprime comme suit :

**A l'action est toujours opposée une réaction égale ; autrement dit, les interactions mutuelles de deux corps sont égales entre elles et dirigées dans des sens opposés.**

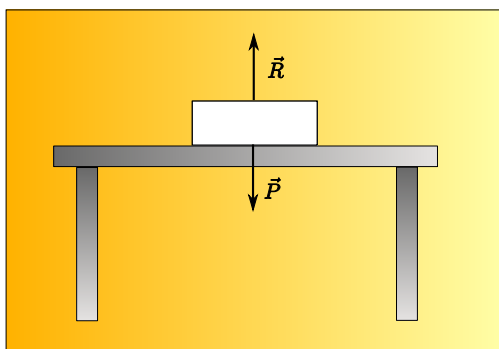


Fig. 57 - L'action et la réaction

BEATRIX : Et on retrouve l'histoire du livre sur la table, pour lequel la réaction de la table est égale et opposée au poids du livre.

## 23 Scène III-5 : Le mouvement résultant des forces

EURISTIDE : Nous allons maintenant appliquer ces lois de la dynamique de Newton au cas simple d'un corps se trouvant soumis uniquement à son poids sous l'effet de la gravitation terrestre.

BEATRIX : Pour un corps de masse constante, on a vu avec la seconde loi de la dynamique que la force était égale à l'accélération du corps multipliée par la masse. Donc ici, l'accélération du corps sera constante, puisque la force est constante.

EURISTIDE : Oui, c'est bien ça. Et un corps dont l'accélération est constante a un mouvement en forme de parabole. C'est-à-dire que dans le cas le plus général, sa trajectoire est une parabole. Mathine va nous décrire cela plus en détail, mais on voit déjà qu'on retrouve le fait qu'un ballon shooté par un joueur décrit une parabole avant de retomber au sol.

MATHINE : Oui, alors regardons ce que cela donne en équations. Supposons qu'un joueur de football ait shooté dans le ballon. Nous allons maintenant regarder le comportement du ballon durant sa chute. Le ballon est soumis uniquement à son poids :

$$\vec{F} = m\vec{g}, \quad (140)$$

où  $\vec{g}$  est la gravitation terrestre, dirigée vers le bas. D'après la seconde loi de la dynamique, on a :

$$\vec{F} = \frac{d}{dt}(m\vec{v}), \quad (141)$$

où  $\vec{v}$  est la vitesse du corps.

La masse du corps étant constante, on obtient :

$$\vec{F} = m\vec{g} = m\frac{d\vec{v}}{dt} = m\vec{w}, \quad (142)$$

où  $\vec{w}$  est l'accélération du corps.

On en déduit que :

$$\vec{g} = \vec{w}. \quad (143)$$

BEATRIX : En d'autres termes, un corps soumis à son poids seulement subit une accélération égale à la gravitation. La gravitation est appelée parfois accélération de la gravitation. On comprend pourquoi, maintenant.



MATHINE : Nous allons maintenant étudier de plus près la trajectoire de notre ballon, en considérant que sa coordonnée suivant l'axe  $y$  est nulle, dans un système de coordonnées  $(O, \vec{i}, \vec{j}, \vec{k})$  tel que  $x$  est l'un des axes horizontaux, et  $z$  l'axe vertical.

Alors, en écrivant l'expression des vecteurs  $\vec{g}$  et  $\vec{w}$  dans ce système de coordonnées, on obtient l'expression ci-après, en considérant que  $\vec{g}$  est un vecteur vertical dirigé vers le bas, donc égal à  $-g\vec{k}$ , et  $\vec{w}$  est la dérivée du vecteur vitesse, donc possède pour coordonnées les dérivées des coordonnées du vecteur vitesse  $(v_x, v_y, v_z)$ . Or, la coordonnée  $v_y$  suivant  $\vec{j}$  est nulle, donc :

$$-g\vec{k} = \frac{dv_x}{dt}\vec{i} + \frac{dv_z}{dt}\vec{k}. \quad (144)$$

Le membre de gauche ne possède qu'une composante en  $\vec{k}$ , pas de composante en  $\vec{i}$ . Ce doit être donc également le cas pour le membre de droite de l'égalité. Donc, les deux termes en  $\vec{k}$  sont égaux et le terme en  $\vec{i}$  est nul :

$$\begin{aligned} \frac{dv_x}{dt} &= 0 \\ \frac{dv_z}{dt} &= -g \end{aligned} \quad (145)$$

Par intégration de la première équation, on déduit que :

$$v_x = v_{x_0}, \quad (146)$$

donc que la vitesse du ballon suivant  $x$  est constante, d'où :

$$x = v_{x_0}t + x_0, \quad (147)$$

où  $v_{x_0}$  est la vitesse initiale du ballon suivant  $x$  (donc l'élan initial fourni horizontalement par le joueur), et  $x_0$  est la position initiale du ballon suivant l'axe  $x$ .

De la seconde équation, nous déduisons, par intégration :

$$v_z = -gt + z_0, \quad (148)$$

où  $v_z$  est la vitesse initiale du ballon suivant  $z$  (donc l'élan initial fourni verticalement par le joueur.)

Or :

$$v_z = \frac{dz}{dt}, \quad (149)$$

donc :

$$\frac{dz}{dt} = -gt + v_{z_0}. \quad (150)$$

En intégrant l'équation une nouvelle fois, on obtient :

$$z = -\frac{1}{2}gt^2 + v_{z_0}t + z_0, \quad (151)$$

où  $z_0$  est la position initiale du ballon suivant l'axe  $z$ .

En résumé, nous avons trouvé deux relations qui décrivent le mouvement du ballon :

$$\begin{aligned} x &= v_{x_0}t + x_0 \\ z &= -\frac{1}{2}gt^2 + v_{z_0}t + z_0. \end{aligned} \quad (152)$$

BEATRIX : Les équations définissent le mouvement du ballon. Mais elles ne sont pas très parlantes...

EURISTIDE : On peut quand même en déduire quelques informations intéressantes. Tout d'abord, on voit que si le ballon possède une vitesse initiale nulle suivant l'axe  $x$ , il ne se déplacera que suivant l'axe vertical. C'est bien ce que l'on observe lorsqu'on jette un ballon à la verticale de soi. Ensuite, on voit que le ballon décrit suivant l'axe des  $z$  un mouvement parabolique en fonction du temps. C'est dire que ce mouvement, s'il est ascendant, atteindra après un instant, un point culminant qui correspondra au point maximal de la parabole. C'est également ce qu'on observe : le ballon que je lance finit par retomber au sol après avoir atteint un point maximal.

MATHINE : On va maintenant déduire l'équation de la courbe décrite par un ballon lancé avec une vitesse initiale non nulle suivant l'axe  $x$ .

Pour cela, on suppose que  $v_{x_0} \neq 0$  et on élimine le paramètre  $t$  entre les deux équations précédentes. Ce qui donne :

$$t = \frac{x - x_0}{v_{x_0}}. \quad (153)$$

Puis,

$$\begin{aligned} z &= -\frac{1}{2}g \frac{(x - x_0)^2}{v_{x_0}^2} + v_{z_0} \frac{x - x_0}{v_{x_0}} + z_0 \\ &= -\frac{g}{2v_{x_0}^2}x^2 + \frac{g}{v_{x_0}^2}x_0x + \frac{v_{z_0}}{v_{x_0}}x - \frac{v_{z_0}}{v_{x_0}}x_0 + z_0 - \frac{x_0^2 g}{2v_{x_0}^2}. \end{aligned} \quad (154)$$

Soit :

$$z = -\frac{g}{2v_{x_0}^2}x^2 + \left( \frac{gx_0}{v_{x_0}^2} + \frac{v_{z_0}}{v_{x_0}} \right)x + \left( z_0 - \frac{g}{2v_{x_0}^2}x_0^2 - \frac{v_{z_0}}{v_{x_0}}x_0 \right). \quad (155)$$

C'est l'équation d'une parabole située dans le plan  $(x, z)$ , parabole dont la convexité est orientée vers le haut, puisque le coefficient de  $x^2$  est négatif :

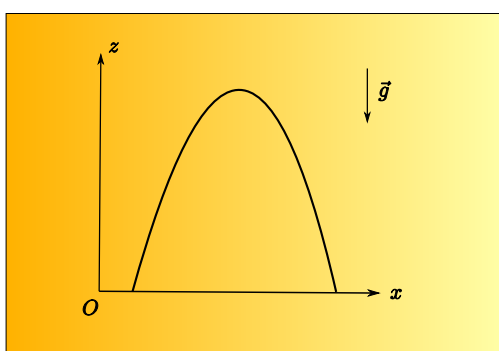


Fig. 58 - La trajectoire du ballon

EURISTIDE : Il s'agit bien là de la trajectoire suivie par un ballon lors d'un shoot effectué avec une vitesse horizontale et une vitesse initiale verticale données. Le ballon s'élève, atteint un point culminant, puis retombe plus loin sur la pelouse du stade.

MATHINE : Dans le cas où  $v_{x_0}$  est nulle, nous revenons au jeu d'équations initial, que nous pouvons écrire :

$$\begin{aligned} x &= x_0 \\ z &= -\frac{1}{2}gt^2 + v_{z_0}t + z_0. \end{aligned} \quad (156)$$

Le mouvement s'effectue sur une verticale au dessus du point d'abscisse  $x = x_0$ .

BEATRIX : C'est bien la trajectoire d'une balle de tennis que nous projetons verticalement pour effectuer un service au tennis. La balle s'élève au dessus de nous, ralentit et atteint un point culminant, puis retombe.

## 24 Scène III-6 : Les mouvements liés d'un corps

EURISTIDE : Voilà. Nous connaissons maintenant le mouvement d'un corps libre soumis à son poids.

Dans certains systèmes physiques, deux ou plusieurs corps peuvent être assujettis à respecter certaines contraintes. Par exemple, un corps roulant sur un plan incliné restera "attaché" à ce plan incliné lors de son mouvement, et donc introduira dans les équations du mouvement une contrainte qui modifiera le mouvement par rapport au mouvement libre.

MATHINE : Etudions plus spécifiquement le cas de figure présenté par le schéma ci-dessous :

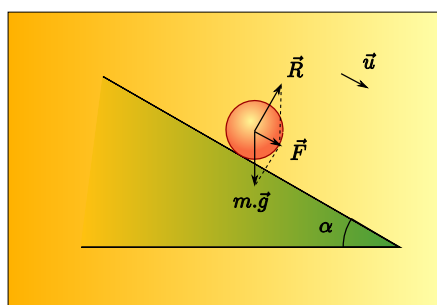


Fig. 59 - Mouvement d'une bille sur un plan incliné

Soit donc une bille de masse  $m$  placée au repos sur un plan incliné faisant un angle  $\alpha$  avec le plan horizontal.

Pour résoudre ce genre de problème, il faut penser à faire le bilan des forces appliquées au système physique. En l'occurrence, si nous faisons le bilan des forces appliquées à la bille, nous avons son poids  $m\vec{g}$  et la réaction  $\vec{R}$  du plan sur la bille. Cette réaction  $\vec{R}$  est perpendiculaire au plan incliné. Le poids  $m\vec{g}$  est vertical.

Donc la résultante  $\vec{F}$  de ces deux forces n'est pas nulle (contrairement à ce qui se passait sur un plan horizontal), et c'est ce qui va provoquer le déplacement de la bille vers la partie basse du plan incliné, conformément à la première loi de la dynamique de Newton :

$$\vec{F} = mg \sin \alpha \vec{u}, \quad (157)$$

où  $\vec{u}$  est un vecteur unitaire parallèle au plan incliné et dirigé dans le sens de la descente.

Alors, d'après la seconde loi de la dynamique, on a :

$$mg \sin \alpha \vec{u} = m \frac{d\vec{v}}{dt}. \quad (158)$$

Par conséquent, le corps prend une accélération constante dirigée vers le bas du plan incliné et d'intensité :

$$w = mg \sin \alpha. \quad (159)$$

BEATRIX : La présence du sinus (toujours inférieur à 1) dans l'expression de l'intensité de l'accélération montre que le corps accélère moins vite que s'il est libre (sans plan incliné). D'ailleurs, plus le plan est fortement incliné, plus l'angle  $\alpha$  est proche de  $\pi/2$ , et plus le sinus est proche de la valeur 1, et par conséquent, plus l'accélération se rapproche de l'accélération de la gravité.

Pour  $\alpha$  proche de 0,  $w$  tend vers 0. C'est ce que l'on constate : si le plan n'est presque pas incliné, la bille possède une accélération très faible.

## 25 Scène III-7 : La quantité de mouvement

EURISTIDE : Les physiciens ont introduit un certain nombre de notions très utiles caractérisant les systèmes physiques et leur comportement. En particulier, lorsque certaines grandeurs interviennent souvent dans les équations de la dynamique, il est utile de leur donner un nom et de se pencher sur leur signification physique.

Par exemple, on appelle **quantité de mouvement** d'un corps, le produit de sa masse par sa vitesse. Cette grandeur, et sa dérivée sont présentes dans les équations de la dynamique :

$$\boxed{\vec{p} = m \vec{v}}. \quad (160)$$

On a parlé tout à l'heure de la quantité de mouvement à propos de la deuxième loi de la dynamique : c'est une sorte de vitesse pesante qui représente l'impact de la masse sur la vitesse d'un corps. Il s'agit donc de son élan.

BEATRIX : D'accord. Mais quelle est son utilité ?

MATHINE : Nous verrons dans un instant que dans le phénomène de prise d'élan, la quantité de mouvement prend tout son sens. Mais pour l'instant, on peut déduire des lois de la dynamique une loi pleine de sens sur la quantité de mouvement.

Etudions, en utilisant les seconde et troisième loi de la dynamique, le comportement d'un système physique constitué de corps soumis à des forces en interaction, mais qui ne sont pas soumis à des forces externes à ce système :

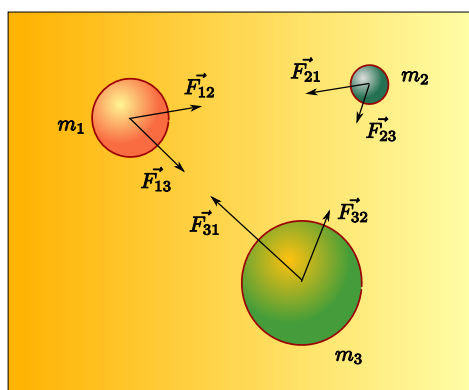


Fig. 60 - Equilibre d'un système de forces

Considérons trois corps de masses respectives  $m_1$ ,  $m_2$  et  $m_3$ .

Supposons que le corps 1 soit soumis par le corps 2 à une force  $F_{12}$ , et par le corps 3 à une force  $F_{13}$ . De même, nous définissons les forces  $F_{31}$  et  $F_{32}$  appliquées au corps 3 par les corps 1 et 2, et  $F_{21}$  et  $F_{23}$  les forces appliquées au corps 2.

La seconde loi de la dynamique s'écrit ici :

$$\begin{aligned} m_1 \vec{w}_1 &= \vec{F}_{12} + \vec{F}_{13} \\ m_2 \vec{w}_2 &= \vec{F}_{21} + \vec{F}_{23} \\ m_3 \vec{w}_3 &= \vec{F}_{31} + \vec{F}_{32}. \end{aligned} \quad (161)$$

En additionnant les trois équations membre à membre, on obtient :

$$m_1 \vec{w}_1 + m_2 \vec{w}_2 + m_3 \vec{w}_3 = \vec{F}_{12} + \vec{F}_{13} + \vec{F}_{21} + \vec{F}_{23} + \vec{F}_{31} + \vec{F}_{32}. \quad (162)$$

Or, par hypothèse, nous avons dit que le système n'était soumis à aucune force externe. C'est donc que la somme des forces internes est nulle. Donc :

$$m_1 \vec{w}_1 + m_2 \vec{w}_2 + m_3 \vec{w}_3 = \vec{0}. \quad (163)$$

C'est-à-dire :

$$\frac{d}{dt}(m_1 \vec{v}_1 + m_2 \vec{v}_2 + m_3 \vec{v}_3) = \vec{0}. \quad (164)$$

Cette équation signifie que la dérivée de la quantité de mouvement totale du système est nulle, donc que la quantité de mouvement du système est constante.

EURISTIDE : C'est une loi fondamentale : la quantité de mouvement d'un système de corps ne peut se modifier sous l'effet de forces intérieures.

$$\boxed{\sum_i \frac{d}{dt} m_i \vec{v}_i = \vec{0}.} \quad (165)$$

## 26 Scène III-8 : La transmission de l'élan

BEATRIX : J'ai bien compris le principe. Mais comment l'utilise-t-on en pratique ?

EURISTIDE : On trouve des applications immédiates de ce principe dans le mécanisme de la prise d'élan. Si tu as déjà pratiqué le skate-board, tu as certainement déjà utilisé le principe de conservation de la quantité de mouvement. Si tu cours sur le sol avant de grimper sur le skate-board qui est à l'arrêt initialement, ton élan provoque le démarrage du skate-board. Que s'est-il passé ?

BEATRIX : Ma vitesse lors de ma course à pied a été transmise au skate-board...

EURISTIDE : Ce n'est pas ta vitesse qui est transmise au skate-board, mais ta quantité de mouvement. En effet, si l'on considère le système constitué de toi-même et le skate-board, on voit qu'il n'est soumis à aucune force extérieure. Dans la situation initiale, la quantité de mouvement du système est ta masse multipliée par ta vitesse. Dans la situation finale, toi et le skate-board avez une même vitesse et comme la quantité de mouvement doit être conservée, on peut en déduire la vitesse que doit prendre le skate-board.

MATHINE : Mettons cela en équations. Supposons qu'un homme de masse  $m_0$  et de vitesse constante  $\vec{v}_0$  monte sur un skate-board de masse  $m_1$  et de vitesse initiale nulle.

Appliquons le principe de conservation de la quantité de mouvement au système constitué de l'homme et du skate-board réunis. Avant que l'homme ne monte sur le skate-board, la quantité de mouvement de l'homme est :

$$m_0 \vec{v}_0 = \vec{0}. \quad (166)$$

Celle du skate-board est nulle.

Donc la quantité de mouvement du système est :

$$\vec{p} = m_0 \vec{v}_0. \quad (167)$$

Après que l'homme est monté sur le skate-board, supposons que ce dernier a acquis la vitesse  $v_1$ . La quantité de mouvement du système constitué de l'homme et du skate-board est alors :

$$\vec{p} = (m_0 + m_1) \vec{v}_1. \quad (168)$$

Donc, d'après la loi de conservation de la quantité de mouvement, on doit avoir :

$$m_0 \vec{v}_0 = (m_0 + m_1) \vec{v}_1. \quad (169)$$

On en déduit l'expression de la vitesse du skate-board, après la prise d'élan :

$$\vec{v}_1 = \frac{m_0}{m_0 + m_1} \vec{v}_0. \quad (170)$$

EURISTIDE : On reconnaît un certain nombre de propriétés bien connues : plus le skate-board est lourd, plus la vitesse résultant de l'élan est faible, puisque le poids du skate-board apparaît au dénominateur de la fraction ci-dessus. En d'autres termes, l'élan est d'autant moins efficace que la masse du skate-board est importante. On voit aussi que sur un skate-board dont la masse est négligeable par rapport à celle du coureur, la vitesse est sensiblement égale à celle du coureur initialement.

## 27 Scène III-9 : L'impulsion

BEATRIX : C'est tout à fait intuitif. Y-a-t'il d'autres grandeurs à commenter ?

EURISTIDE : Il y a **l'impulsion d'une force**. C'est le produit de la force appliquée à un corps par l'intervalle de temps pendant lequel celle-ci est appliquée. Le nom est bien le reflet du rôle de cette impulsion en physique. J'applique une force à un corps pendant un certain temps. C'est l'impulsion que je communique au corps, et celle-ci va déterminer la vitesse du corps.

MATHINE : La seconde loi de la dynamique de Newton s'écrit d'ailleurs :

$$\boxed{d(m\vec{v}) = \vec{F} dt.} \quad (171)$$

La grandeur  $\vec{F} dt$  est appelée impulsion de la force  $\vec{F}$  pendant le laps de temps  $dt$ .

On voit que l'accroissement de la quantité de mouvement durant un intervalle de temps est égal à l'impulsion de la force appliquée au cours du même intervalle de temps.

## 28 Scène III-10 : Les corps de masse variable

EURISTIDE : Le concept de l'impulsion est utile pour traiter de l'étude des systèmes dont la masse est variable dans le temps.

MATHINE : Considérons une fusée à propergol liquide, éjectant par sa tuyère des particules de masse  $\mu$  chaque seconde à la vitesse relative  $\vec{c}$  par rapport à la fusée, et dirigée vers l'arrière de cette dernière :

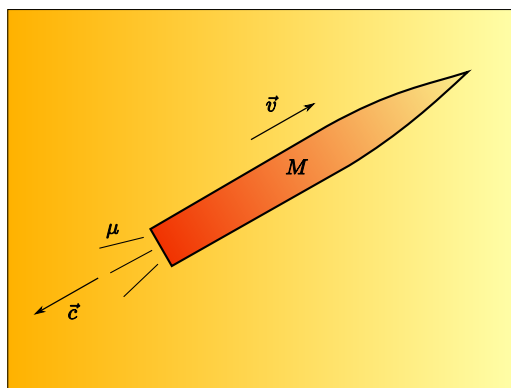


Fig. 61 - Le mouvement d'une fusée

Supposons que cette fusée possède une vitesse  $\vec{v}$  et une masse  $M$ .  
 A l'instant  $t$ , la fusée possède la quantité de mouvement suivante :

$$\vec{p} = M \vec{v}. \quad (172)$$

A l'instant  $t + dt$ , la fusée a perdu la quantité de carburant suivante :

$$dM = \mu dt, \quad (173)$$

et sa vitesse est devenue :

$$\vec{v} + d\vec{v}. \quad (174)$$

Donc sa quantité de mouvement est :

$$(M - \mu dt)(\vec{v} + d\vec{v}). \quad (175)$$

Notons  $\vec{v}_1$  la vitesse du carburant éjecté. La quantité de mouvement du carburant éjecté est :

$$\mu dt \vec{v}_1. \quad (176)$$

Donc la quantité de mouvement à l'instant  $t$  est  $M \vec{v}$ , et à l'instant  $t + dt$  est :

$$(M - \mu dt)(\vec{v} + d\vec{v}) + \mu dt \vec{v}_1. \quad (177)$$

Ecrivons maintenant la loi de l'accroissement de la quantité de mouvement de l'ensemble du système constitué de la fusée et du carburant :

$$(M - \mu dt)(\vec{v} + d\vec{v}) + \mu dt \vec{v}_1 - M \vec{v} = \vec{F} dt, \quad (178)$$

où  $\vec{F}$  est la force extérieure appliquée à la fusée (poids et frottement de l'air par exemple).  
 Alors, en simplifiant l'expression, on obtient :

$$M d\vec{v} + \mu dt (\vec{v}_1 - \vec{v}) = \vec{F} dt. \quad (179)$$



Or  $\vec{v}_1 - \vec{v}$  exprime la vitesse relative du carburant éjecté, c'est-à-dire  $\vec{c}$ .

D'où :

$$M d\vec{v} + \mu dt \vec{c} = \vec{F} dt, \quad (180)$$

ou encore :

$$M \frac{d\vec{v}}{dt} = \vec{F} - \mu \vec{c}. \quad (181)$$

Si la fusée évolue dans l'espace, alors elle n'est soumise à aucun poids et à aucun frottement puisque l'espace est vide et sans pesanteur. Donc  $\vec{F} = \vec{0}$ .

Donc, en prenant les modules des vecteurs (puisque les vecteurs sont tous colinéaires, c'est-à-dire sur la même direction), nous avons :

$$M \frac{dv}{dt} = \mu c. \quad (182)$$

A l'instant initial, la masse de la fusée est  $M_0$ . A l'instant  $t$  elle est :

$$M = M_0 - \mu t. \quad (183)$$

D'où :

$$\frac{dv}{c} = \frac{\mu dt}{M_0 - \mu t}. \quad (184)$$

D'où, en intégrant :

$$\begin{aligned} \frac{1}{c}(v - v_0) &= \int_0^t \frac{\mu dt}{M_0 - \mu t} = \ln |M_0| - \ln |M_0 - \mu t| \\ &= \ln \left\| \frac{M_0}{M_0 - \mu t} \right\| \\ &= \ln \frac{M_0}{M}. \end{aligned} \quad (185)$$

**EURISTIDE** : Cette expression nous montre bien ce qui se passe. La vitesse augmente avec le temps au fur et à mesure que la masse de la fusée diminue en perdant une masse de son carburant par la tuyère.

---

## Cinquième partie

# Le travail et l'énergie

### 29 Scène IV-1 : La notion d'énergie

BEATRIX : Il y a une notion sous-jacente à toutes les notions que nous avons vues : la vitesse, la transmission de la quantité de mouvement, l'impulsion. N'est-il pas question d'énergie dans tous les cas ?

EURISTIDE : L'énergie est un concept important que nous devons bien comprendre. Nous en aurons besoin lorsque nous aborderons la relativité générale.

Lorsqu'un corps se déplace, nous considérons intuitivement que ce corps possède une certaine énergie.

Lorsque ce même corps mobile heurte un obstacle, son mouvement s'arrête, et il se peut que le corps qu'il a heurté ait maintenant un mouvement également, donc une certaine énergie.

Il s'avère que le premier corps a transmis son énergie au second corps lors du choc.

Cette notion d'énergie est essentielle pour la compréhension des phénomènes physiques. L'énergie est en effet le lien fondamental entre toutes les formes de mouvement. Nous savons qu'il est possible de transformer l'énergie électrique en énergie mécanique dans un moteur électrique, nous savons aussi qu'il est possible de transformer l'énergie mécanique en énergie électrique, comme dans un alternateur. De nombreux autres exemples de transformation de l'énergie existent dans la nature.

### 30 Scène IV-2 : Le travail

BEATRIX : Mais comment cette énergie est-elle transmise d'un corps à un autre ?

EURISTIDE : Pour répondre à cette question, nous introduisons la notion de travail. Le **travail** est le produit de la force par le déplacement. Lorsqu'une force est appliquée par un corps (ma main, par exemple) sur un autre corps, et que le deuxième corps se déplace sous la poussée du premier, ce deuxième corps possède une énergie résultante communiquée par le travail correspondant de la force appliquée.

BEATRIX : Un peu compliqué... Mais si je résume, le travail d'une force, c'est ce qui représente le transfert d'énergie réalisé par l'application de la force sur une trajectoire donnée.

### 31 Scène IV-3 : Le travail d'une force

EURISTIDE : Il est évident intuitivement que si une force est exercée dans la même direction et le même sens que le déplacement d'un corps, l'énergie de ce corps va être accrue et le travail correspondant sera par conséquent positif.

Inversement, si la force est exercée dans un sens opposé au déplacement, le travail sera négatif.

MATHINE : Nous allons avoir besoin d'une notion mathématique dont nous n'avons pas encore parlé. Il s'agit du produit scalaire. Si nous considérons deux vecteurs  $\vec{u}$  et  $\vec{v}$ , de modules respectifs  $u$  et  $v$ , on appelle **produit scalaire** des vecteurs  $\vec{u}$  et  $\vec{v}$  la grandeur :

$$\vec{u} \cdot \vec{v} = u \times v \times \cos(\vec{u}, \vec{v}). \quad (186)$$

où l'on prend le cosinus de l'angle entre  $\vec{u}$  et  $\vec{v}$ . On généralise en  $n$  dimensions immédiatement avec la même formule. Si l'un des vecteurs est nul, le produit scalaire est évidemment nul. Il se trouve que si deux vecteurs sont orthogonaux, alors leur produit scalaire est nul puisque si l'angle  $(\vec{u}, \vec{v}) = \frac{\pi}{2}$ , alors  $\cos \frac{\pi}{2} = 0$ . Et la propriété réciproque est vraie également : si le produit de deux vecteurs non nuls est nul, alors ils sont orthogonaux.

Il se trouve que le produit scalaire des vecteurs  $\vec{u}$  et  $\vec{v}$  correspond au produit des longueurs du vecteur  $\vec{u}$  et de la projection orthogonale du vecteur  $\vec{v}$  sur le vecteur  $\vec{u}$  :

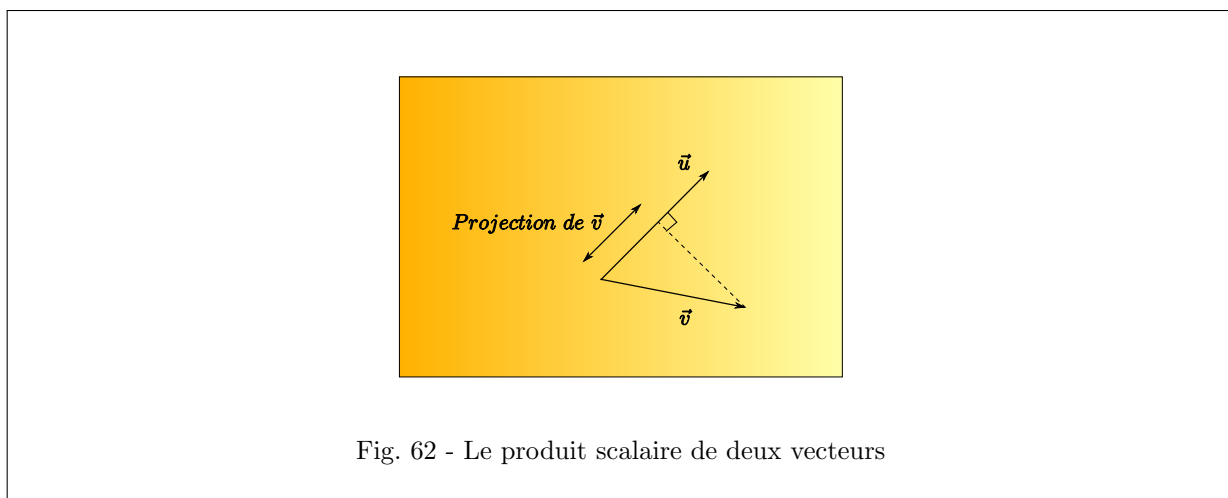


Fig. 62 - Le produit scalaire de deux vecteurs

A noter que si  $\vec{u}$  et  $\vec{v}$  forment un angle obtus, alors la valeur de  $\vec{u} \cdot \vec{v}$  est négative. Revenons maintenant au travail d'une force. Le travail d'une force est mathématiquement égal au produit scalaire du vecteur force par le vecteur déplacement, sur un trajet infiniment petit :

$$dA = \vec{F} \cdot d\vec{s}. \quad (187)$$

Ainsi, le travail d'une force le long d'un chemin  $(x_1, x_2)$  est égal à l'intégrale de ce produit scalaire le long du chemin considéré :

$$A = \int_{x_1}^{x_2} dA, \quad (188)$$

soit :

$$A = \int_{x_1}^{x_2} \vec{F} \cdot d\vec{s}. \quad (189)$$

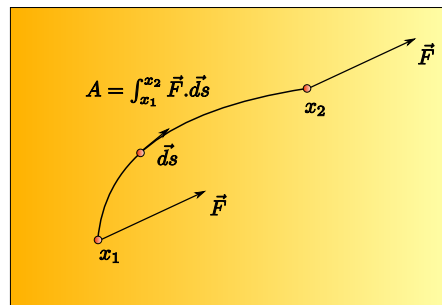


Fig. 63 - Le travail d'une force le long d'un trajet

BEATRIX : On voit ici une intégrale d'un vecteur ! Nous n'avions jamais vu cela jusqu'alors, non ?

MATHINE : C'est vrai. L'intégrale d'un vecteur est une notion un peu nouvelle pour nous. Mais, il suffit d'explicitier les coordonnées des vecteurs dans un repère orthonormé pour retrouver des intégrales de nombres réels.

BEATRIX : Et on constate au passage que le travail d'une force agissant perpendiculairement au déplacement d'un corps est nul, puisque le produit scalaire de la force et du vecteur déplacement dans ce cas est nul. On le savait : on ne peut pas déplacer un corps en le poussant perpendiculairement à la direction dans laquelle il peut se déplacer. Par exemple, pour pousser un wagon sur des rails, il vaut mieux pousser parallèlement aux rails, plutôt que perpendiculairement !

## 32 Scène IV-4 : L'énergie potentielle de déformation

EURISTIDE : Voyons maintenant les différentes manières sous lesquelles l'énergie d'un système mécanique se fait connaître.

Considérons d'abord un ressort déformé sous l'effet d'une force  $\vec{F}$ . Supposons que le ressort s'allonge d'une longueur  $x_0$  sous l'effet de cette force :

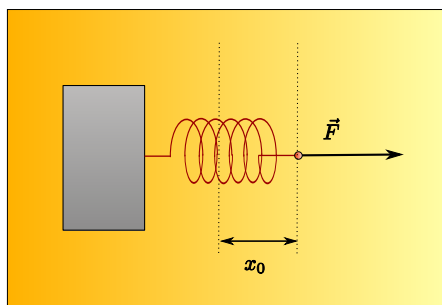


Fig. 64 - L'énergie potentielle de déformation

Le travail dépensé par cette force sur cette distance porte le nom **d'énergie potentielle de déformation**.

MATHINE : Pour un ressort, on a déterminé expérimentalement que la force  $F$  est proportionnelle à la déformation, donc :

$$F = kx, \quad (190)$$

où  $k$  est une constante, correspondant à l'élasticité du ressort.

Par conséquent, l'énergie potentielle de déformation du ressort allongé d'une longueur  $x_0$  est :

$$A = \int_0^{x_0} F dx = \int_0^{x_0} kx dx = \frac{1}{2} kx_0^2. \quad (191)$$

### 33 Scène IV-5 : L'énergie cinétique

BEATRIX : Et comment s'appelle l'énergie d'un corps en mouvement ?

EURISTIDE : Lorsqu'une force est appliquée sur un corps initialement immobile, et que ce corps est en mouvement de ce fait, alors le corps en mouvement a acquis une réserve d'énergie supplémentaire. Cette réserve d'énergie est appelée énergie cinétique.

MATHINE : Considérons un corps de masse  $m$  et de vitesse initiale nulle, atteignant sous l'effet de la force  $\vec{F}$ , la vitesse  $v_0$ .

D'après la seconde loi de la dynamique, on a :

$$m \frac{d\vec{v}}{dt} = \vec{F}. \quad (192)$$

En multipliant les deux membres de cette égalité par le déplacement infinitésimal  $\vec{ds}$ , on obtient :

$$m \frac{d\vec{v}}{dt} \cdot \vec{ds} = \vec{F} \cdot \vec{ds} = dA, \quad (193)$$

où  $dA$  est l'accroissement infinitésimal du travail.

On peut écrire l'équation précédente sous la forme :

$$m \vec{v} \cdot d\vec{v} = dA, \quad (194)$$

puisque :

$$\frac{d\vec{s}}{dt} = d\vec{v}. \quad (195)$$

On passe au module, en considérant que le corps se déplace sur une ligne droite :

$$mvdv = dA. \quad (196)$$

En intégrant cette équation entre  $v = 0$  et  $v = v_0$ , nous obtenons :

$$\boxed{A = m \int_0^{v_0} v dv = \frac{1}{2} m v_0^2.} \quad (197)$$

Ceci est l'expression de l'énergie cinétique d'un corps de masse  $m$  se déplaçant à la vitesse  $v_0$ .

**BEATRIX** : Dans ce calcul, on voit bien qu'il s'agit du calcul d'un transfert d'énergie, parce qu'on a calculé le travail total qui a été fourni pour conduire le corps de l'immobilité à la vitesse  $v_0$ . C'est donc ce travail qui s'est accumulé dans le corps sous forme d'énergie cinétique. L'accumulation du travail fourni est représentée mathématiquement par le calcul de l'intégrale qui a été utilisée pour parvenir au résultat.

## 34 Scène IV-6 : L'énergie potentielle

EURISTIDE : Bravo! Belle interprétation! Tu es mûre pour passer au chapitre suivant, celui de l'énergie potentielle.

BEATRIX : On a parlé tout à l'heure de l'énergie potentielle de déformation d'un ressort. S'agit-il de quelque chose d'autre?

EURISTIDE : Il s'agit de l'énergie potentielle acquise par un corps lorsqu'il prend de l'altitude. Lorsqu'une force est appliquée à un corps et que ce corps est soulevé (c'est-à-dire élevé à une certaine altitude au dessus du sol), alors ce corps acquiert une réserve d'énergie supplémentaire. Cette réserve d'énergie est appelée énergie potentielle, car elle donne au corps la potentialité de tomber en chute libre si la force qui soulève le corps est annulée.

BEATRIX : Il faut qu'on applique la même méthode que précédemment, en calculant le travail nécessaire à l'élévation du corps à une certaine altitude.

MATHINE : Oui, c'est ce que nous allons faire. Considérons un corps de masse  $m$  monté à la hauteur  $h$  sous l'effet de la force  $\vec{F}$ , et soumis à la force de gravitation  $\vec{F}_g$ .

D'après la seconde loi de la dynamique :

$$\vec{F} + \vec{F}_g = m \frac{d\vec{v}}{dt}. \quad (198)$$

D'où :

$$\vec{F} \cdot d\vec{s} + \vec{F}_g \cdot d\vec{s} = m \frac{d\vec{v}}{dt} \cdot d\vec{s}. \quad (199)$$

Ce qui peut s'écrire :

$$\vec{F} \cdot d\vec{s} = -\vec{F}_g \cdot d\vec{s} + d\left(\frac{mv^2}{2}\right). \quad (200)$$

$d\vec{s}$  représente le déplacement vertical du corps vers le haut.  $\vec{F}_g$  représente la force de gravitation dirigée vers le bas. Donc en passant au module les vecteurs qui sont colinéaires, on obtient :

$$-\vec{F}_g \cdot d\vec{s} = F_g ds. \quad (201)$$

Or, la gravitation  $F_g$  s'écrit aussi :

$$F_g = mg. \quad (202)$$

Donc, en intégrant sur la trajectoire entre  $A$  et  $B$ , verticale, de longueur  $h$ , on obtient :

$$\int_A^B \vec{F} \cdot d\vec{s} = A = mgh + \frac{mv_B^2}{2} - \frac{mv_A^2}{2}. \quad (203)$$

Donc, si le corps est immobile au départ en  $A$  et à l'arrivée en  $B$ , le travail de la force est :

$$\boxed{\int_A^B \vec{F} \cdot d\vec{s} = mgh.} \quad (204)$$

Cette expression est l'énergie potentielle du corps qui s'élève de la hauteur  $h$ , dans le champ de gravitation terrestre.

## 35 Scène IV-7 : La conservation de l'énergie

BEATRIX : Je comprends. Il y a un échange d'énergie lorsqu'un corps possède une énergie potentielle et qu'il se met à se mouvoir.

Lorsqu'un corps est en mouvement libre dans le champ de gravitation terrestre, son énergie mécanique, somme de son énergie potentielle et de son énergie cinétique, reste constante. Si le corps est en chute libre sous l'effet de la gravitation, son énergie potentielle est convertie en énergie cinétique, au fur et à mesure que le corps accélère dans sa chute. Inversement, si le corps est projeté vers le haut, son énergie cinétique est convertie en énergie potentielle, au fur et à mesure que le corps ralentit jusqu'à son point culminant où il possède une énergie cinétique nulle et une énergie potentielle maximale. Il commence alors à redescendre, pour convertir son énergie potentielle en énergie cinétique.

EURISTIDE : Tu as parfaitement compris les mécanismes de transfert d'énergie. On va pouvoir passer à la suite.



## Sixième partie

# Acte V - La notion de relativité du mouvement

## 36 Scène V-1 : Les systèmes de référence galiléens

MATHINE : Revenons aux systèmes de coordonnées. Jusqu'ici, nous avons implicitement considéré que le système de référence était au repos. C'est une condition pour que les lois de la dynamique de Newton s'appliquent.

Nous avons vu qu'un référentiel au repos était un référentiel immobile, ou un référentiel dont la vitesse est constante et la trajectoire rectiligne par rapport à un référentiel immobile. Ainsi, les lois de la dynamique restent valables dans tout référentiel immobile ou en mouvement rectiligne à vitesse constante par rapport à celui-ci. De tels référentiels sont dits **galiléens**.

## 37 Scène V-2 : Le référentiel non galiléen

BEATRIX : Mais alors, que se passe-t-il si le référentiel n'est pas galiléen ? Comment les lois de la dynamique sont-elles modifiées ?

EURISTIDE : On va regarder ce que deviennent les lois de la dynamique dans un référentiel non galiléen. Un exemple de référentiel non galiléen, c'est celui attaché à un wagon doué d'un mouvement rectiligne accéléré, d'accélération  $\vec{a}$  constante. Nous allons donc observer les lois de la physique à l'intérieur de ce wagon.

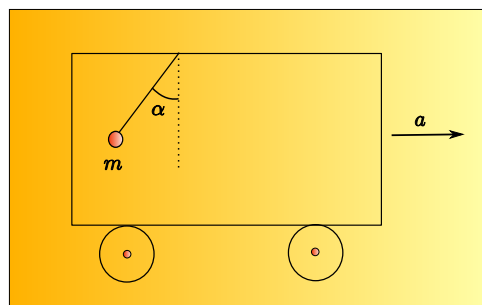


Fig. 65 - La force d'inertie

L'expérience montre que le pendule de masse  $m$  suspendu au plafond du wagon fait un angle avec la verticale. Nous connaissons cet effet qui tend à nous presser contre notre siège lorsqu'un train ou une automobile accélère. Soit  $\alpha$  l'angle que fait le pendule avec la verticale. Sous l'effet de l'accélération du wagon (c'est-à-dire sous l'effet de l'accélération du référentiel attaché au wagon), on considère que le pendule est soumis à une force imaginaire, dite **force d'inertie**  $\vec{F}$ , dont la valeur est :

$$\vec{F} = -m\vec{a}. \quad (205)$$

On voit que cette force présente une certaine analogie avec le poids  $\vec{P}$  :

$$\vec{P} = m\vec{g}. \quad (206)$$

Ici,  $\vec{F}$  remplace  $\vec{P}$  et  $-\vec{a}$  remplace  $\vec{g}$ . Autrement dit, on considère que le corps est soumis à une force opposée à l'accélération du référentiel, qui constitue une tendance du corps à rester immobile, c'est-à-dire à résister à l'accélération. D'où son nom de force d'inertie.

MATHINE : Ainsi, la position d'équilibre du pendule de masse  $m$  est assurée sous l'effet des forces suivantes :  $\vec{F}$  la force d'inertie,  $\vec{N}$  la force de tension du fil du pendule,  $\vec{P}$  le poids du pendule.

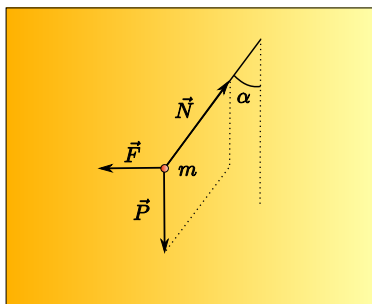


Fig. 66 - L'équilibre des forces dans un référentiel en mouvement

Donc, on a :

$$\vec{F} + \vec{N} + \vec{P} = \vec{0}. \quad (207)$$

Soit :

$$-m\vec{a} + \vec{N} + m\vec{g} = \vec{0}. \quad (208)$$

Ce qui, après projection des vecteurs sur l'horizontale, donne :

$$ma = mg \tan \alpha \quad (209)$$

ou encore :

$$\tan \alpha = \frac{a}{g}. \quad (210)$$

Cette expression donne l'angle  $\alpha$  du pendule par rapport à la verticale.

**BEATRIX** : Donc, en résumé, le fait que le référentiel soit accéléré fait comme si le corps était soumis à une force de sens opposé. Ensuite, on applique les lois de la dynamique, comme d'habitude, avec cette force fictive dans le bilan des forces du système.

### 38 Scène V-3 : Le référentiel en rotation

**EURISTIDE** : Et voilà, c'est aussi simple que cela. Et nous verrons, lorsque nous aborderons la relativité générale, que cette propriété est tout à fait en ligne avec celle-ci. La dualité accélération-force est une pierre angulaire de la relativité générale.

**MATHINE** : Mais pour l'instant, traitons le cas d'un pendule placé sur un disque en rotation. Soit  $\omega$  la vitesse angulaire du disque (et donc du référentiel). La vitesse angulaire correspond au nombre de tours effectués par seconde du disque. le pendule se trouve à une distance  $R$  du centre du disque.

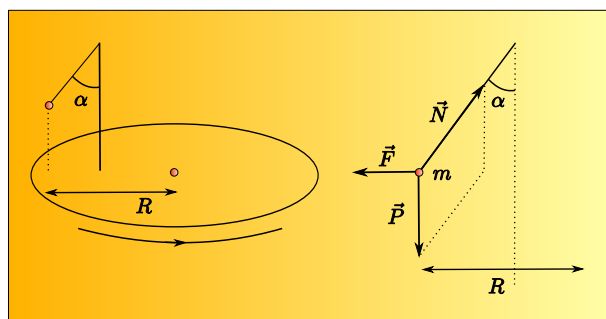


Fig. 67 - Le référentiel en rotation

La position à l'instant  $t$  du pendule sur le plan horizontal avec un système de coordonnées centré sur l'axe de rotation est donnée par :

$$\begin{aligned} x &= R \cos \omega t \\ y &= R \sin \omega t. \end{aligned} \quad (211)$$

Alors, la vitesse a pour coordonnées la dérivée par rapport au temps des coordonnées précédentes :

$$\begin{aligned}v_x &= -R\omega \sin \omega t \\v_y &= R\omega \cos \omega t.\end{aligned}\tag{212}$$

Et l'accélération est la dérivée par rapport au temps des coordonnées de la vitesse, soit :

$$\begin{aligned}a_x &= -R\omega^2 \cos \omega t \\a_y &= -R\omega^2 \sin \omega t.\end{aligned}\tag{213}$$

Par conséquent, le module de l'accélération s'écrit :

$$\sqrt{a_x^2 + a_y^2} = \omega^2 R.\tag{214}$$

Par conséquent, le pendule est soumis à une accélération  $\omega^2 R$ , et donc nous pouvons appliquer la formule vue à la scène précédente et écrire :

$$\tan \alpha = \frac{\omega^2 R}{g}.\tag{215}$$

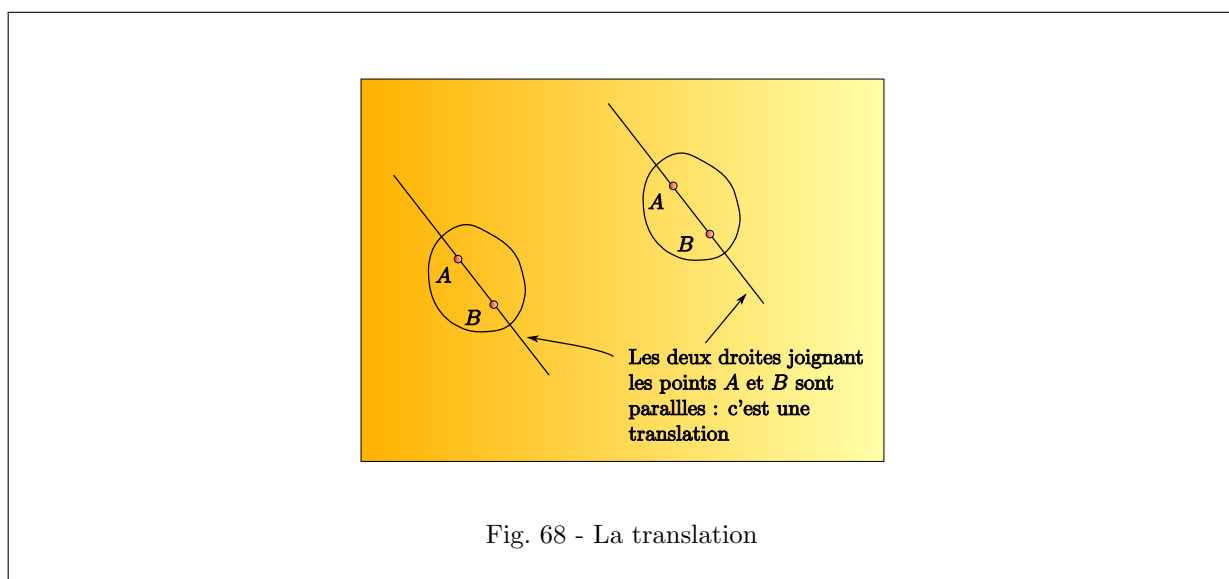
## Septième partie

# Acte VI - Le mouvement du corps solide

## 39 Scène VI-1 : La translation et la rotation

BEATRIX : On a vu beaucoup de choses maintenant sur le mouvement d'un point. Mais un corps solide ne se résume pas à un point. Le corps solide peut tourner sur lui-même, par exemple. Les lois que nous venons de voir ne sont pas suffisantes pour décrire le mouvement d'un corps.

EURISTIDE : Oui, tu as raison. Il est temps de s'intéresser au corps solide. On appelle corps solide un ensemble de points solidaires les uns des autres. Un corps solide peut effectuer des translations et des rotations. On appelle translation d'un corps solide, le mouvement dans lequel chaque ligne joignant deux points quelconques du corps conserve une direction invariable dans l'espace.



On appelle rotation d'un corps solide le mouvement dans lequel les trajectoires de tous les points d'un corps solide sont des circonférences cocentriques dont les centres sont disposés sur une même droite, dite axe de rotation.

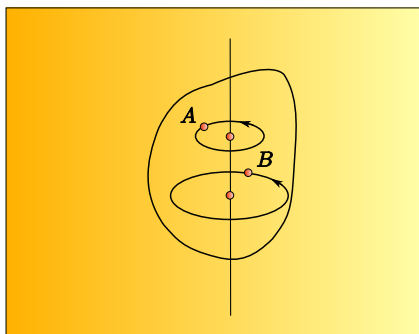


Fig. 69 - La rotation

Tout mouvement d'un corps solide peut se décomposer en translation et rotation. La vitesse de chaque point est alors définie comme la somme de deux vitesses : la vitesse de translation de l'axe de rotation, et la vitesse linéaire du point résultant du mouvement rotatif du solide autour de son axe.

**MATHINE** : Pour calculer cette vitesse linéaire due à la rotation du corps solide, on doit effectuer ce qu'on appelle le produit vectoriel de la vitesse de rotation  $\vec{\omega}$  et du vecteur allant orthogonalement de l'axe de rotation au point considéré. Nous définirons tout à l'heure le produit vectoriel. En attendant, nous voyons ici représentés les différents vecteurs du système en rotation :  $\vec{\omega}$  la vitesse angulaire (ou vitesse de rotation),  $\vec{R}$  le vecteur rayon allant du centre de rotation jusqu'au point considéré  $M$ , et  $\vec{v}$  le vecteur vitesse linéaire du point  $M$ .

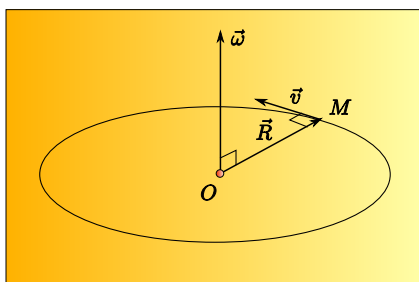


Fig. 70 - Vitesse angulaire, vitesse linéaire

Ce qu'on appelle **vitesse angulaire** est un vecteur dont la direction est portée par l'axe de rotation et dont l'intensité est la vitesse de rotation du corps exprimée en mètres<sup>-1</sup>. La vitesse angulaire est un vecteur parallèle à l'axe de rotation, dirigé de telle façon que si celui-ci est dirigé vers nos yeux, nous voyons que le corps solide tourne dans le sens contraire des aiguilles d'une montre.

BEATRIX : Le vecteur vitesse angulaire donne en fait la direction et le sens de progression d'une vis qu'on ferait tourner pour la faire pénétrer dans un matériau.

MATHINE : Le produit vectoriel de deux vecteurs  $\vec{u}$  et  $\vec{v}$  est un produit un peu spécial permettant de transformer deux vecteurs en un vecteur dont la direction est perpendiculaire au plan des deux vecteurs d'origine et dont le sens est tel que si on positionne le vecteur résultant vers le haut, on passe en tournant sur l'axe de ce troisième vecteur du premier vecteur vers le second vecteur, en tournant dans le sens contraire des aiguilles d'une montre. On note  $\vec{u} \wedge \vec{v}$  le produit vectoriel de  $\vec{u}$  et  $\vec{v}$ .

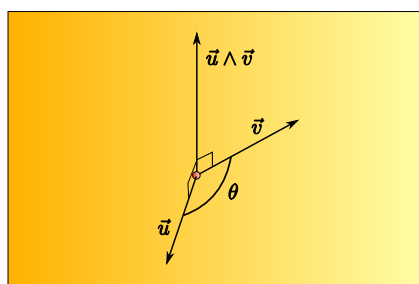


Fig. 71 - Le produit vectoriel

BEATRIX : Et quel est le module de ce produit vectoriel ?

MATHINE : Le module dépend de l'angle entre  $\vec{u}$  et  $\vec{v}$ . Plus cet angle est proche de l'angle droit  $\pi/2$ , et plus le module est important. Plus précisément :

$$|\vec{u} \wedge \vec{v}| = uv \sin \theta \quad (216)$$

où  $\theta$  est l'angle que font entre eux les vecteurs  $\vec{u}$  et  $\vec{v}$ .

BEATRIX : Donc, le produit vectoriel de deux vecteurs parallèles est nul, et le module du produit vectoriel de deux vecteurs orthogonaux est égal au produit de leurs modules.

MATHINE : En l'occurrence, dans notre cas de figure, les vecteurs  $\vec{\omega}$  et  $\vec{R}$  sont orthogonaux, donc le module de  $\vec{\omega} \wedge \vec{R}$  est le produit des modules de  $\vec{\omega}$  et de  $\vec{R}$ . Par conséquent, le module de la vitesse  $\vec{v} = \vec{\omega} \wedge \vec{R}$  est bien :

$$v = \omega R. \quad (217)$$

BEATRIX : Donc, plus la vitesse angulaire est grande, plus la vitesse linéaire est grande. On le savait. Mais aussi, plus le rayon est grand, plus la vitesse linéaire est grande. On s'en aperçoit bien lorsqu'on se trouve sur un tourniquet pour les enfants dans un jardin public. Lorsque je suis assise près du centre du tourniquet, mon corps tourne sur lui-même, mais sa vitesse linéaire est faible. Je ne sens pas le vent dans mes cheveux. Il en est autrement lorsque je m'assieds à la périphérie du tourniquet. Ma vitesse linéaire est beaucoup plus grande, et mes cheveux volent au vent. Et pourtant, la vitesse angulaire du tourniquet est toujours la même.

EURISTIDE : Fondamentalement, la vitesse linéaire d'un point en rotation autour d'un axe est bien le produit vectoriel du vecteur vitesse angulaire  $\vec{\omega}$  avec le vecteur rayon  $\vec{R}$  : parce que c'est bien physiquement une quantité proportionnelle à  $\omega R$ , sa direction est bien orthogonale à l'axe de rotation d'une part et au rayon d'autre part, et enfin son sens est bien compatible avec la règle du produit vectoriel.

Enfin, si d'aventure le vecteur rayon faisait un angle  $\alpha$  différent de l'angle droit avec l'axe de rotation :

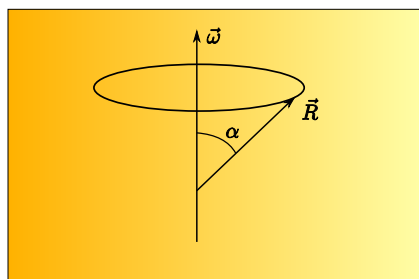


Fig. 72 - Rotation d'un rayon vecteur

alors, c'est la distance à l'axe, donc la projection de  $\vec{R}$  sur le plan perpendiculaire à l'axe de rotation qui déterminerait le module du vecteur vitesse linéaire. Donc on aurait :

$$v = R\omega \sin \alpha, \quad (218)$$

ce qui est tout à fait en ligne avec la définition d'un produit vectoriel.

BEATRIX : C'est très clair.



## 40 Scène VI-2 : L'équilibre sur un axe fixe

EURISTIDE : Un point qu'il est intéressant de regarder maintenant, c'est la condition pour qu'un corps pouvant se mouvoir en rotation, soit en équilibre.

Pour cela, on va considérer une roue pouvant tourner autour de son axe. Intuitivement, nous savons bien que nous avons d'autant plus de facilité à imprimer un mouvement de rotation à la roue, que nous saisissons celle-ci loin de son axe. C'est la notion de bras de levier que nous voyons illustrée ici.

Pour formaliser cela, nous allons introduire deux nouvelles notions.

On appelle **bras de levier d'une force** relativement à l'axe considéré d'un corps, la distance la plus courte entre l'axe de rotation et la ligne d'action de la force. On appelle **moment d'une force relativement à un axe** la grandeur physique numériquement égale au produit de la force par son bras de levier.

Cette distance est matérialisée par un segment qui est perpendiculaire à l'axe de rotation d'une part et à la force d'autre part.

Le moment représente, en quelque sorte, l'impact potentiel de la force sur la rotation du corps. Plus le bras de levier de la force est grand, plus son impact est grand pour entraîner la rotation du corps; et par conséquent, plus le moment de cette force relativement à l'axe de rotation est grand. Inversement, si le bras de levier de la force est très faible, la force n'a presque pas d'impact sur la rotation; le moment de cette force est très faible.

BEATRIX : Pourquoi appelle-t-on cela un "moment"? Cela n'a rien à voir avec un instant, n'est-ce pas?

EURISTIDE : Ce mot "moment" vient du latin "momentum", qui signifie "influence". En l'occurrence, le moment de la force représente l'influence de cette force sur la rotation du corps.

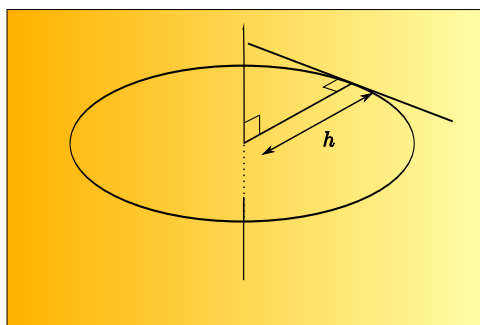


Fig. 73 - Le bras de levier

Ainsi, un corps solide ayant un axe fixe de rotation est en équilibre lorsque la somme des moments des forces par rapport à l'axe est nulle. En effet, si on applique au corps une force tendant à faire tourner le

corps dans un sens, et si on applique à ce même corps une autre force tendant à le faire tourner dans le sens inverse, alors ces deux forces vont lutter l'une contre l'autre. Si leurs impacts sur la rotation (donc leurs moments) se contrecarrent exactement, le corps sera en équilibre et n'entrera pas en rotation. Cela revient à dire que leurs moments sont égaux et opposés. Par conséquent, leur somme sera nulle.

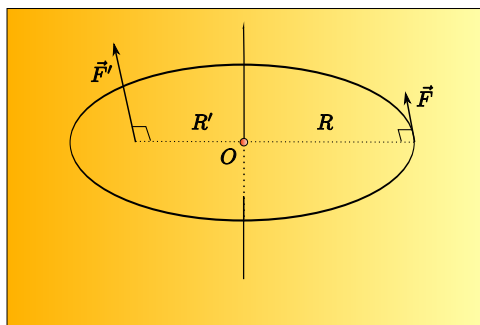


Fig. 74 - Les moments de forces par rapport à l'axe de rotation

MATHINE : Pour le schéma ci-dessus, il faut la condition suivante, pour assurer l'équilibre du corps :

$$R' \vec{F}' + R \vec{F} = \vec{0}. \quad (219)$$

Soit :

$$\vec{F}' = -\frac{R}{R'} \vec{F}. \quad (220)$$

BEATRIX : Cette formule correspond à ce qu'on attendait : si une force  $\vec{F}$  appliquée sur la circonférence de la roue tend à faire tourner cette roue, il faut appliquer à une distance moindre de l'axe, une force plus importante pour contrecarrer cette rotation.

EURISTIDE : C'est l'illustration de l'effet du bras de levier.

## 41 Scène VI-3 : La dynamique autour un axe fixe

BEATRIX : Maintenant que nous avons vu les conditions d'équilibre d'un corps autour d'un axe, je suis intéressée par la dynamique d'un corps en rotation autour d'un axe.

EURISTIDE : Nous allons maintenant établir, par similitude avec ce qui se passe pour le point, la relation qui existe entre l'accélération de la rotation d'un corps autour d'un axe fixe et les forces qui lui sont appliquées. Lorsque nous avons étudié la dynamique du point, nous avons utilisé la vitesse et l'accélération du point. Pour le corps en rotation autour d'un axe, nous avons déjà introduit précédemment la vitesse angulaire; nous pouvons, par analogie, introduire la notion d'accélération angulaire, comme la dérivée par rapport au temps de la vitesse angulaire. Nous utiliserons également le moment des forces par rapport à l'axe de rotation, puisque nous avons vu que c'est essentiellement le moment des forces qui intervient pour mesurer l'impact des forces à faire pivoter le solide autour d'un axe fixe. Nous avons vu que la dynamique du point fait intervenir la vitesse, l'accélération et les forces. Nous verrons que la dynamique du corps solide en rotation autour d'un axe fixe fait intervenir la vitesse angulaire à la place de la vitesse, l'accélération angulaire à la place de l'accélération, et le moment des forces par rapport à l'axe de rotation à la place des forces, car c'est en effet l'impact des forces sur la rotation du solide qui nous importe dans cette étude.

MATHINE : Commençons par formaliser l'accélération d'un tel corps en rotation. Nous avons défini la vitesse angulaire  $\vec{\omega}$ . On appelle l'accélération angulaire la dérivée par rapport au temps de la vitesse angulaire; on la note  $\vec{\beta}$  :

$$\vec{\beta} = \frac{d\vec{\omega}}{dt}. \quad (221)$$

Considérons maintenant une particule quelconque du corps de masse  $\Delta m_i$  située à une distance  $r_i$  de l'axe de rotation. Ecrivons la seconde équation de la dynamique pour cette particule :

$$\Delta m_i \frac{d\vec{v}_i}{dt} = \Delta m_i r_i \frac{d\vec{\omega}}{dt} = (\vec{f}_i)_{in} + (\vec{f}_i)_{ex}, \quad (222)$$

où  $(\vec{f}_i)_{in}$  est la force interne au corps appliquée à la particule, et  $(\vec{f}_i)_{ex}$  est la force externe au corps appliquée à cette même particule.

On peut écrire, en multipliant les deux membres de l'égalité par  $r_i$ , la distance de la masse à l'axe, pour faire apparaître le moment des forces (puisque, rappelons-le, le moment d'une force est le produit de la force par sa distance à l'axe) :

$$\Delta m_i r_i^2 \frac{d\vec{\omega}}{dt} = r_i (\vec{f}_i)_{in} + r_i (\vec{f}_i)_{ex}. \quad (223)$$

Nous avons là l'équation de la dynamique pour une seule particule. Ecrivons la somme de ces équations pour l'ensemble des particules du solide :

$$\frac{d\vec{\omega}}{dt} \sum_i \Delta m_i r_i^2 = \sum_i r_i (\vec{f}_i)_{in} + \sum_i r_i (\vec{f}_i)_{ex}. \quad (224)$$

Le corps est solide, donc par définition, la somme des moments des forces intérieures qui sont appliquées à ses particules s'annule (sinon le corps se disloquerait). Donc :

$$\sum_i r_i (\vec{f}_i)_{in} = \vec{0}. \quad (225)$$

Par ailleurs, l'expression :

$$\boxed{\sum_i r_i (\vec{f}_i)_{ex} = \vec{M}} \quad (226)$$

est appelée **moment rotatif** des forces extérieures appliquées au corps et notée  $\vec{M}$ . Son module est  $M$ .

EURISTIDE : On retrouve la notion de moment (d'influence, étymologiquement) qui indique que l'expression mesure l'impact des forces pour la rotation du corps solide.

MATHINE : L'expression :

$$\boxed{\sum_i \Delta m_i r_i^2 = I} \quad (227)$$

est appelée **moment d'inertie par rapport à l'axe de rotation** considéré, et noté  $I$ .

EURISTIDE : On retrouve encore une fois la notion de moment. Cette expression intuitivement l'impact de la masse du corps sur l'inertie de la rotation du corps solide. En d'autres termes, le moment d'inertie joue le rôle que jouait la masse dans la dynamique du point.

MATHINE : Ainsi, nous pouvons écrire l'équation fondamentale de la dynamique pour un corps solide en rotation autour d'un axe fixe :

$$\boxed{\vec{M} = I \frac{d\vec{\omega}}{dt}} \quad (228)$$

EURISTIDE : Ce qui s'exprime : le moment des forces extérieures entraînant le corps dans la rotation autour d'un axe donné fixe, est égal au moment d'inertie du corps par rapport à cet axe, multiplié par l'accélération angulaire du corps.

BEATRIX : On peut faire le parallèle avec la loi de la dynamique pour un point :

$$\vec{F} = m \frac{d\vec{v}}{dt} \quad (229)$$

Dans le cas d'un corps en rotation, nous voyons que le moment d'inertie joue le rôle de la masse  $m$  (ce qui explique son nom d'ailleurs), et le moment des forces extérieures joue le rôle de la force  $F$ .

EURISTIDE : Cette relation entre moment d'inertie et moment des forces permet de comprendre quelques phénomènes que nous connaissons bien. Ce moment d'inertie  $I$  est d'autant plus grand que la masse du corps en rotation se trouve concentrée à la périphérie du corps. Comme nous avons :

$$\frac{d\omega}{dt} = \frac{M}{I}, \quad (230)$$

il faut une force dont le moment est d'autant plus grand pour parvenir à accélérer un corps dont la masse est importante à la périphérie. De la même façon, il est aussi plus difficile de ralentir un tel corps lancé en rotation (ralentir en cinématique signifie rendre l'accélération négative.) Cela se comprend avec cette même relation, et est parfaitement illustré dans le principe des volants d'inertie, qui sont des corps ayant un moment d'inertie très important. On a vu que le moment d'inertie s'écrit :

$$I = \sum_i \Delta m_i r_i^2. \quad (231)$$

Il est donc d'autant plus grand que la masse est importante bien sûr, mais aussi et surtout lorsque cette masse se trouve éloignée de l'axe de rotation. C'est la raison pour laquelle les volants d'inertie sont en général des sortes de roues dont la circonférence est constituée d'un métal très lourd.

## 42 Scène VI-4 : Le moment de la quantité de mouvement

MATHINE : Toujours par analogie avec la dynamique du point, on peut définir le moment de la quantité de mouvement par l'expression :

$$N = \sum_i \Delta m_i r_i^2 \omega = I\omega, \quad (232)$$

où la somme s'effectue pour l'ensemble des particules du solide de masse  $m_i$  à la distance  $r_i$  de l'axe de rotation du solide.

EURISTIDE : On se souvient que la quantité de mouvement représentait en dynamique du point l'élan du point. Ce moment de la quantité de mouvement représente donc l'élan du corps dans sa rotation autour de son axe.

MATHINE : Ainsi, la loi fondamentale de la dynamique pour un corps solide (donc de masse fixe) en rotation autour d'un axe immobile s'écrit également :

$$M = I \frac{d\omega}{dt} = \frac{dN}{dt}. \quad (233)$$

EURISTIDE : Ceci s'exprime comme suit : la dérivée par rapport au temps du moment de la quantité de mouvement du corps par rapport à l'axe de rotation est égale au moment des forces par rapport au même axe.

BEATRIX : On retrouve encore une analogie avec la loi de la dynamique pour un point :

$$\vec{F} = \frac{d\vec{p}}{dt} \quad (234)$$

où  $\vec{p}$  est la quantité de mouvement du point. Pour le cas d'un corps en rotation, nous voyons que le moment de la quantité de mouvement joue le rôle de la quantité de mouvement, et le moment des forces extérieures joue le rôle de la force  $\vec{F}$ .

## 43 Scène VI-5 : L'énergie cinétique en rotation

MATHINE : Continuons avec l'énergie cinétique. L'énergie cinétique d'un corps en rotation est la somme des énergies cinétiques des particules du corps. Pour chaque particule, on a :

$$T_i = \frac{\Delta m_i v_i^2}{2} = \omega^2 \frac{\Delta m_i r_i^2}{2}. \quad (235)$$

Donc, en faisant la somme de ces  $T_i$  pour l'ensemble des points du corps, on trouve que l'énergie cinétique du corps en rotation est :

$$T = \frac{I\omega^2}{2}. \quad (236)$$

BEATRIX : Ici encore, on reconnaît la formule de l'énergie cinétique d'un point :

$$T = \frac{mv^2}{2}, \quad (237)$$

où l'on a remplacé comme d'habitude maintenant, la masse par le moment d'inertie et la vitesse par la vitesse angulaire.

C'est finalement assez simple. Les lois de la dynamique du point s'appliquent pour les solides en rotation à condition de remplacer la masse du point par le moment d'inertie du solide, la vitesse linéaire du point par la vitesse angulaire du solide, l'accélération linéaire du point par l'accélération angulaire du solide, la quantité de mouvement par le moment de la quantité de mouvement, les forces extérieures par le moment des forces extérieures.

EURISTIDE : Souvenons-nous, le mot "moment" provient du latin "momentum", qui signifiait "influence". Les différents moments que nous avons vu ici représentent donc l'influence des masses, des forces, des quantités de mouvement sur la rotation du corps autour de son axe fixe.

BEATRIX : Oui, c'est clair. En dynamique, moment égale influence.

## 44 Scène VI-6 : Le mouvement du centre d'inertie

EURISTIDE : La rotation du corps autour de lui-même s'effectue autour de son centre d'inertie. Mais qu'en est-il du déplacement de ce centre d'inertie ? Le centre d'inertie d'un corps solide se meut comme si toutes les forces appliquées au corps étaient appliquées en ce point, donc la masse du corps y est entièrement concentrée.

Ainsi, la quantité de mouvement  $\vec{K}$  du corps est égale à sa masse  $m$  multipliée par la vitesse de son centre d'inertie.

## 45 Scène VI-7 : Le mouvement des planètes

EURISTIDE : Un exemple d'application de la mécanique de Newton est la détermination du mouvement des planètes du système solaire. En 1609, Johannes Kepler est parti des observations de Tycho-Brahé pour déduire des lois de Newton, les trois lois du mouvement des planètes.

MATHINE : La première loi de Kepler dit que :

**La trajectoire de chaque planète est une ellipse dont le Soleil occupe l'un des foyers.**

Cette loi résulte directement de la loi de gravitation de Newton.

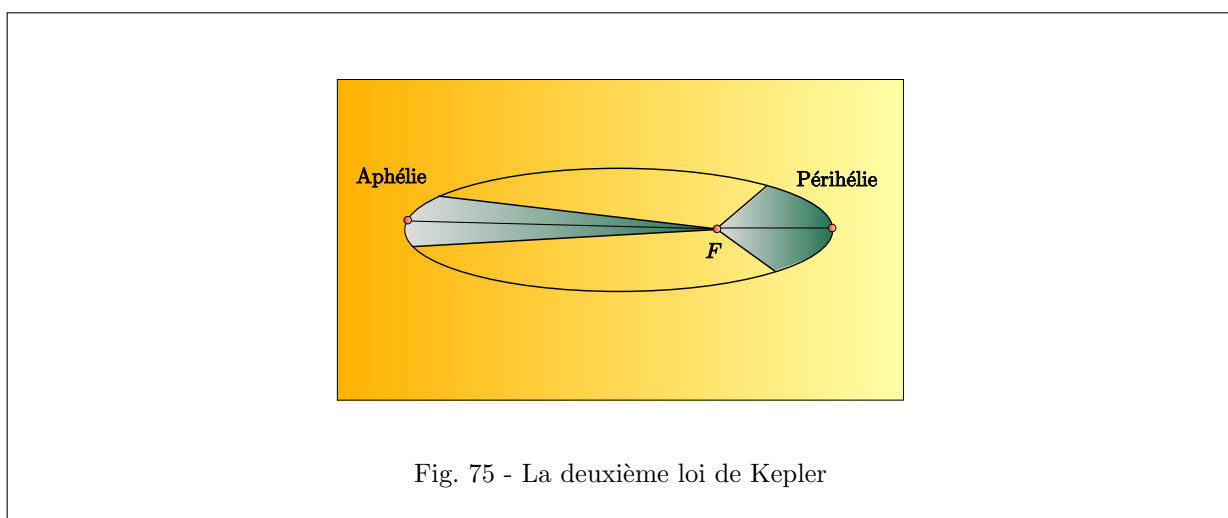
EURISTIDE : Il se trouve que dans notre système solaire, les planètes se déplacent sur des ellipses qui sont très proches de cercles.

Voici quelques définitions. On appelle **écliptique** le plan formé par l'ellipse de la trajectoire de la Terre autour du Soleil. On appelle **périhélie** le point de l'orbite d'une planète le plus rapproché du Soleil. On appelle **aphélie** le point le plus éloigné de cette orbite.

BEATRIX : Nous connaissons donc maintenant la trajectoire des planètes. Qu'en est-il de leur vitesse sur cette trajectoire ?

MATHINE : La deuxième loi de Kepler nous dit que :

**Les aires balayées par les rayons vecteurs de chaque planète en des temps égaux sont identiques.**



Ce qu'on appelle rayon-vecteur est un vecteur issu d'un des foyers de l'ellipse parcourue par la planète et aboutissant au point courant de la trajectoire.

EURISTIDE : Autrement dit, les deux aires hachurées sur la figure ci-dessus sont égales.

Cette loi détermine la vitesse des planètes sur leurs trajectoires. Cela signifie que les planètes ralentissent à l'aphélie et accélèrent au périhélie. Cette loi résulte de la conservation du moment cinétique.

MATHINE : La troisième loi de Kepler dit enfin que :

Les carrés des durées de révolution  $T_1$  et  $T_2$  de deux planètes sont proportionnels aux cubes des

grands axes  $a_1$  et  $a_2$  de leurs orbites :

$$\boxed{\frac{T_1^2}{T_2^2} = \frac{a_1^3}{a_2^3}}, \quad (238)$$

et :

$$\boxed{\frac{T^2}{a^3} = Constante} \quad (239)$$

**BEATRIX** : Cette loi signifie en particulier que plus la planète est lointaine, plus sa vitesse de rotation autour du Soleil est grande. Mais la vitesse de rotation augmente plus vite que la distance de la planète au Soleil.



## Huitième partie

# Acte VII - Le principe variationnel

## 46 Scène VII-1 : Les variations

MATHINE : Nous avons fini de regarder et de réviser les principes de base de la mécanique. Nous allons maintenant regarder une notion qui nous sera utile pour déterminer les équations du champ gravitationnel d'Einstein. Il s'agit du principe de moindre action. Nous verrons que les systèmes physiques, quels qu'ils soient, se comportent afin de minimiser une fonction décrivant leur niveau d'énergie : cette fonction est appelée action.

Pour cela, nous allons commencer par étudier ce que sont les variations de fonctions particulières qui transforment des courbes en nombres. On dit que ces fonctions particulières sont définies sur l'espace des courbes dans un espace de dimension  $n$ . Ces fonctions un peu spéciales sont appelées **fonctionnelles**.

BEATRIX : C'est un peu abstrait pour l'instant. Peut-on prendre un exemple pour mieux comprendre ce que c'est ?

MATHINE : La longueur d'une courbe dans le plan euclidien est un exemple de fonctionnelle. Elle transforme une courbe en un nombre qui est sa longueur. Les fonctionnelles sont donc des fonctions particulières qui transforment des courbes en nombres réels.

Soit par exemple une courbe  $\gamma$  définie entre les points  $t = t_0$  et  $t = t_1$  :

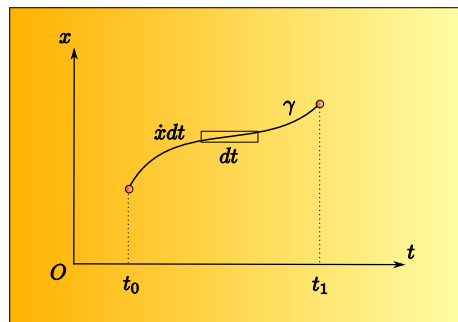


Fig. 76 - Analyse de la longueur d'une courbe

On va maintenant calculer la longueur de cette courbe, ce qui permettra d'explicitier cette fonctionnelle. Considérons un déplacement infinitésimal le long de la courbe  $\gamma$ . Le déplacement a une durée  $dt$ . Si on

considère la dérivée de  $x$  par rapport à  $t$ , notée  $\dot{x} = \frac{dx}{dt}$ , la longueur du déplacement infinitésimal par rapport à la coordonnée  $x$  est  $\dot{x}dt$ .

La longueur  $dl$  sur la courbe  $\gamma$ , parcourue pendant ce déplacement infinitésimal est, au premier ordre, la diagonale du petit rectangle de hauteur  $\dot{x}$  et de largeur  $dt$ . En appliquant le théorème de pythagore à ce petit rectangle, on trouve :

$$dl = \sqrt{(\dot{x}dt)^2 + (dt)^2} = \sqrt{\dot{x}^2 + 1}dt. \quad (240)$$

Donc, finalement, la longueur de la courbe est exprimée par :

$$\Phi(\gamma) = \int_{t_0}^{t_1} \sqrt{1 + \dot{x}^2}dt. \quad (241)$$

$\Phi(\gamma)$  est un exemple de fonctionnelle.

Nous allons nous intéresser, à partir de ceci, au comportement de la fonctionnelle lorsque nous faisons varier la courbe  $\gamma$  de façon infinitésimale. Ceci ressemble au calcul des différentielles que nous avons fait au début de notre discussion lors des révisions des notions de base. En effet, nous avons une fonction  $f(t)$ , et nous calculons sa différentielle en faisant varier  $t$  de façon infinitésimale, ce qui se notait  $dt$ . Le calcul correspondant est appelé **calcul variationnel**. Il se trouve que ce calcul possède des propriétés tout à fait remarquables et utiles, qui nous permettront de déduire rapidement et efficacement les équations décrivant le comportement d'un système physique.

Nous allons donc analyser la variation d'une fonctionnelle  $\Phi$  lorsque son paramètre (qui est une courbe, ne l'oublions pas) varie de façon infinitésimale. Pour cela, considérons une courbe  $\gamma$  et une courbe  $\gamma + h$  infiniment voisine de celle-ci.

Nous nous intéressons alors à la variation de cette fonctionnelle lorsque la courbe varie de  $\gamma$  à  $\gamma + h$ .

Par analogie avec ce que nous connaissons des différentielles des fonctions réelles, nous pouvons considérer que nous avons une différentielle d'une fonctionnelle si nous parvenons à écrire la différence :

$$\Phi(\gamma + h) - \Phi(\gamma) \quad (242)$$

sous la forme d'une somme de deux termes :

$$\Phi(\gamma + h) - \Phi(\gamma) = F(h) + o(h), \quad (243)$$

où le premier terme est une fonction du premier degré en  $h$ , et est appelé différentielle de la fonctionnelle, et le deuxième terme est une expression négligeable devant ce premier terme. Rappelons que la notation  $o(h)$  signifie que  $o(h)$  est une expression négligeable par rapport à  $h$  lorsque  $h$  est proche de 0. Par conséquent, en faisant tendre  $h$  vers 0, l'expression  $\Phi(\gamma + h) - \Phi(\gamma)$  tendra vers l'expression  $F(h)$  qui ne contient que des termes du premier degré en  $h$ . Ceci revient à obtenir une fonction s'approximant à  $\Phi(\gamma + h) - \Phi(\gamma)$  lorsque  $h$  tend vers 0. On démontre (par une démonstration un peu longue que je ne ferai pas ici) que si on considère une fonctionnelle, écrite sous la forme suivante :

$$\Phi(\gamma) = \int_{t_0}^{t_1} L(x, \dot{x}, t)dt, \quad (244)$$

où  $L$  est une fonction dépendant de  $x$ , de la dérivée de  $x$  par rapport à  $t$  et de  $t$ , alors sa différentielle est :

$$F(h) = \int_{t_0}^{t_1} \left( \frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} \right) h dt + \left[ \frac{\partial L}{\partial \dot{x}} h \right]_{t_0}^{t_1}. \quad (245)$$

L'expression :

$$\left[ \frac{\partial L}{\partial \dot{x}} h \right]_{t_0}^{t_1} \quad (246)$$

représente l'expression :

$$\left(\frac{\partial L}{\partial \dot{x}}\right)\Big|_{t=t_1} - \left(\frac{\partial L}{\partial \dot{x}}\right)\Big|_{t=t_0}. \quad (247)$$

Si on reprend le cas particulier de la fonctionnelle longueur d'une courbe :

$$L = \sqrt{1 + \dot{x}^2}, \quad (248)$$

alors la fonctionnelle :

$$\Phi(\gamma) = \int_{t_0}^{t_1} \sqrt{1 + \dot{x}^2} dt, \quad (249)$$

est bien l'expression de la longueur d'un courbe. En appliquant la formule précédente, nous trouvons que :

$$\begin{aligned} \frac{\partial L}{\partial x} &= 0 \\ \frac{\partial L}{\partial \dot{x}} &= \frac{\dot{x}}{\sqrt{1 + \dot{x}^2}}. \end{aligned} \quad (250)$$

Donc la différentielle de la fonctionnelle est :

$$F(h) = \int_{t_0}^{t_1} -\frac{d}{dt} \left( \frac{\dot{x}}{\sqrt{1 + \dot{x}^2}} \right) h dt + \left[ \frac{\dot{x}h}{\sqrt{1 + \dot{x}^2}} \right]_{t_0}^{t_1}. \quad (251)$$

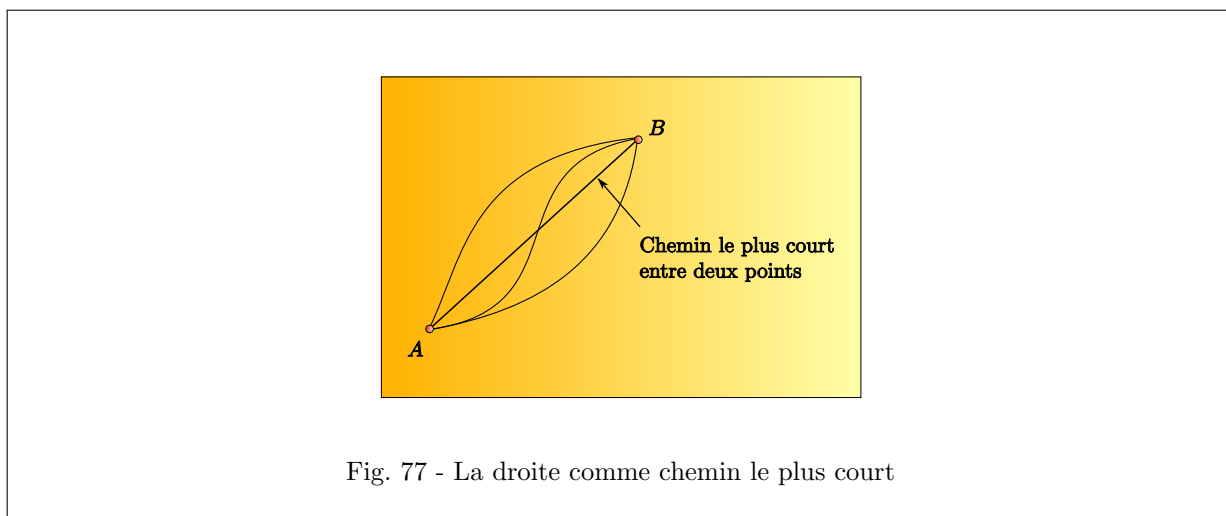
**EURISTIDE** : Cette fonctionnelle  $F(h)$  représente la façon dont varie la longueur d'une courbe lorsqu'on fait varier de façon infinitésimale la courbe en lui ajoutant un accroissement infiniment petit  $h$ . Il faut noter que  $h$  est ici une courbe, donc une fonction de  $t$ , tout comme  $\gamma$ .

## 47 Scène VII-2 : Les extrémales

**BEATRIX** : C'est bien compliqué! Mais à quoi cela va-t-il nous servir?

**EURISTIDE** : Nous avons maintenant une classe de fonctionnelles dont nous connaissons la différentielle. Nous allons naturellement nous intéresser aux courbes qui annulent cette différentielle. Cela correspondra aux courbes qui rendent ces fonctionnelles extrémales (c'est-à-dire minimales ou maximales.) Ceci peut se comprendre par analogie avec ce que nous avons appris en cinématique : lorsque nous lançons une balle verticalement au dessus de nous, nous savons que la vitesse de la balle finit par s'annuler. Le point où la vitesse de la balle s'annule correspond à sa position maximale en hauteur.

Nous cherchons à exprimer la même notion pour les courbes. Là où les courbes qui annulent la différentielle d'une fonctionnelle correspondent à une ou des courbes pour lesquelles le critère exprimé par la fonctionnelle est maximisé ou minimisé. Si la fonctionnelle représente la longueur des courbes passant par deux points fixes, alors nous pourrions déduire avec ce que nous savons maintenant, l'équation des courbes qui sont le chemin le plus court entre les deux points considérés. Nous devrions trouver des équations de droites.



On voit aussi immédiatement l'intérêt de cette démarche en mécanique, puisque nous pourrons ainsi identifier une fonctionnelle exprimant l'énergie d'un système en fonction de sa trajectoire, et chercher la trajectoire de ce système qui minimise la valeur de l'énergie. La trajectoire recherchée sera l'extrémale de cette fonctionnelle (en l'occurrence un minimum de la fonctionnelle.)

**MATHINE** : Plus formellement, on appelle extrémale d'une fonctionnelle  $\Phi(\gamma)$ , une courbe  $\gamma$  telle que  $F(h, \gamma) = 0$  quel que soit  $h$ .

**EURISTIDE** : Cette notion d'extrémale va permettre de définir la notion d'optimum, c'est-à-dire de déterminer la courbe optimisant le chemin d'un point à un autre sous certaines conditions définies par la fonctionnelle.

**MATHINE** : Je vais encore vous asséner un résultat un peu compliqué. Une condition nécessaire et suffisante pour qu'une courbe  $\gamma$  définie entre les points  $t = t_0$  et  $t = t_1$  soit extrémale de la fonctionnelle :

$$\Phi(\gamma) = \int_{t_0}^{t_1} L(x, \dot{x}, t) dt \quad (252)$$

pour toutes les courbes passant par les points  $t = t_0$  et  $t = t_1$  est que sur la courbe  $x(t)$ , on ait :

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}} \right) - \frac{\partial L}{\partial x} = 0. \quad (253)$$

En effet, dans ce cas :

$$F(h) = \int_{t_0}^{t_1} \left( \frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} \right) h dt - \left[ \frac{\partial L}{\partial \dot{x}} h \right]_{t_0}^{t_1}. \quad (254)$$

Or les courbes passent toutes par les points  $t = t_0$  et  $t = t_1$ , donc  $h(t_0) = h(t_1) = 0$ .

Donc :

$$\left[ \frac{\partial L}{\partial \dot{x}} h \right]_{t_1}^{t_1} = 0. \quad (255)$$

Puis, dire que  $F(h)$  est nul pour tout  $h$ , c'est dire que :

$$\int_{t_0}^{t_1} \left( \frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} \right) h dt \quad (256)$$

est nul pour tout  $h$ .

Il faut donc que l'expression :

$$\frac{\partial L}{\partial x} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} \quad (257)$$

dans l'intégrale soit nulle. D'où le résultat.

**BEATRIX** : Ah, je comprend mieux ! Nous avons maintenant une méthode directe pour exprimer les courbes extrémales de la fonctionnelle au moyen de cette dernière équation. Celle-ci crée une équation différentielle, qu'il nous suffira ensuite de résoudre pour trouver l'équation des courbes qui sont extrémales. Si on prenait l'exemple de la fonctionnelle exprimant la longueur des courbes entre  $t = t_0$  et  $t = t_1$ , on devrait certainement trouver l'équation d'une droite, puisque, c'est bien connu, la droite est le chemin le plus court entre deux points.

**MATHINE** : Et bien, allons-y ! Prenons cet exemple.

Nous avons vu tout à l'heure que la fonctionnelle longueur d'une courbe s'écrivait :

$$\Phi(\gamma) = \int_{t_0}^{t_1} L(x, \dot{x}, t) dt, \quad (258)$$

avec :

$$L(x, \dot{x}, t) = \sqrt{1 + \dot{x}^2}. \quad (259)$$

Alors :

$$\begin{aligned} \frac{\partial L}{\partial x} &= 0 \\ \frac{\partial L}{\partial \dot{x}} &= \frac{\dot{x}}{\sqrt{1 + \dot{x}^2}}. \end{aligned} \quad (260)$$

Donc, la condition :

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}} \right) - \frac{\partial L}{\partial x} = 0 \quad (261)$$

s'écrit :

$$\frac{\partial L}{\partial \dot{x}} = \frac{\dot{x}}{\sqrt{1 + \dot{x}^2}} = \text{Constante}, \quad (262)$$

donc :

$$\dot{x} = \text{Constante} = c. \quad (263)$$

D'où enfin :

$$x = ct + a, \quad (264)$$

où  $a$  est une constante. Cette dernière équation constitue bien l'équation d'une droite. Les constantes  $a$  et  $c$  peuvent être déterminées en exprimant que la droite passe par des points prédéterminés pour  $t = t_0$  et  $t = t_1$ .

## 48 Scène VII-3 : L'équation d'Euler-Lagrange

EURISTIDE : L'équation dont nous sommes partis :

$$\boxed{\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}} \right) - \frac{\partial L}{\partial x} = 0} \quad (265)$$

est très utile en physique pour construire l'équation différentielle d'un système physique. Elle porte le nom d'équation d'Euler-Lagrange pour la fonctionnelle :

$$\Phi = \int_{t_0}^{t_1} L(x, \dot{x}, t) dt. \quad (266)$$

Elle caractérise donc les courbes extrémales de la fonctionnelle  $\Phi$ . Il suffit de bien choisir la fonctionnelle  $\Phi$  pour obtenir l'équation d'un système physique qui minimise ou maximise cette fonctionnelle.

BEATRIX : C'est presque magique. L'inconvénient de cette magie, c'est qu'il est difficile de donner une signification intuitive à l'équation d'Euler-Lagrange. C'est particulièrement abstrait.

EURISTIDE : C'est vrai. Mais on peut toutefois dire deux trois choses sur cette équation.

BEATRIX : Je n'en attendais pas moins de vous.

EURISTIDE : Cette équation nous dit en fait que :

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}} \right) = \frac{\partial L}{\partial x}. \quad (267)$$

En français, cela donne : le taux de variation de  $L$  quand on fait varier  $x$ , est égal à la vitesse de variation du taux de variation de  $L$  quand on fait varier la dérivée de  $x$ . Déjà, on voit que si la fonction de Lagrange  $L$  est constante le long de la courbe, alors :

$$\frac{\partial L}{\partial x} = 0, \quad (268)$$

donc, d'après l'équation d'Euler-Lagrange :

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}} \right) = 0. \quad (269)$$

Donc, dans ce cas, le taux de variation de la fonction de Lagrange en fonction de la dérivée de  $x$  est constant. Reprenons l'exemple de la fonctionnelle longueur. La fonction de Lagrange  $L$  représente l'élément de longueur infinitésimal. L'équation d'Euler-Lagrange nous dit que l'élément de longueur infinitésimal varie le long de la courbe, comme la vitesse de variation de l'élément de longueur en fonction de la pente de la courbe. On sait déjà que l'élément de longueur ne dépend que de la variable  $\dot{x}$ . Donc  $\frac{\partial L}{\partial x} = 0$ . Donc, d'après l'équation d'Euler-Lagrange,  $\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}} \right)$ . Donc la vitesse de la variation de l'élément de longueur en fonction de la pente de la courbe est nulle, donc l'élément de longueur varie de façon constante avec la pente de la courbe. Donc, si la pente de la courbe augmentait, alors l'élément de longueur augmenterait également de façon proportionnelle.

Mais on a vu que cet élément de longueur ne dépendait que de  $\dot{x}$ . Donc, l'équation d'Euler-Lagrange applique une contrainte qui fait que la pente de la courbe  $\dot{x}$  est forcément constante le long de la courbe.

BEATRIX : Et une courbe de pente constante, c'est une droite. Bravo !

EURISTIDE : Maintenant le cas général est plus difficile à interpréter. Supposons que  $L$  soit l'expression de l'énergie d'un système mécanique. L'équation d'Euler-Lagrange dit que l'énergie du système mécanique le long de sa trajectoire ( $\frac{\partial L}{\partial \dot{x}}$ ) varie exactement comme la vitesse de variation dans le temps de la variation de l'énergie en fonction de la vitesse sur la trajectoire.

BEATRIX : C'est un peu obscur, non ?

EURISTIDE : Ca s'éclaircit, si on prend l'intégrale de l'équation correspondante :

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}} \right) = \frac{\partial L}{\partial x}, \quad (270)$$

c'est-à-dire :

$$\frac{\partial L}{\partial \dot{x}} = \int \frac{\partial L}{\partial x} dt. \quad (271)$$

L'équation d'Euler-Lagrange appliquée à l'énergie nous dit maintenant que la variation de l'énergie en fonction de la vitesse le long de la trajectoire du système permet une accumulation (et accumulation égale intégrale!) d'énergie le long de la trajectoire. En d'autres termes, l'équation d'Euler-Lagrange va appliquer au système une contrainte telle que si la vitesse du système le long de sa trajectoire varie, cela induira une variation de l'énergie, et le système suivra une trajectoire telle que l'énergie accumulée le long de cette trajectoire sera exactement égale à la variation d'énergie induite par la variation de la vitesse. En termes encore plus simples, l'équation d'Euler-Lagrange nous dit que si l'énergie cinétique du système varie, le système est contraint à suivre une trajectoire le long de laquelle il accumule une énergie potentielle exactement égale à cette variation de l'énergie cinétique. Donc, finalement, tout système mécanique parcourt une trajectoire qui lui permet de transférer ses variations d'énergie cinétique en énergie potentielle.

BEATRIX : Tiens, nous avons déjà vu cela en mécanique tout à l'heure. Je comprends maintenant mieux le rôle de cette fameuse équation d'Euler-Lagrange.

EURISTIDE : On voit tout l'intérêt de cette équation. On ne peut pas s'en passer, et nous verrons toute sa puissance quand il s'agira de rechercher les équations d'Einstein pour la relativité générale.

MATHINE : Nous devons aussi généraliser cette équation à un espace de dimension  $n$ . Le côté abstrait de la chose ne va pas s'arranger. . . .

Pour cela, on considère un espace à  $n$  dimensions, et une courbe de l'espace à  $n + 1$  dimensions, que nous définissons pour le paramètre  $t$  et le vecteur  $\vec{x}$  de l'espace à  $n$  dimensions, entre les points  $(t_0, \vec{x}_0)$  et  $(t_1, \vec{x}_1)$ .

On considère également une fonction  $L$  à  $2n + 1$  variables réelles.

On démontre qu'une condition nécessaire et suffisante pour qu'une courbe  $\gamma$  soit extrémale de la

fonctionnelle :

$$\Phi(\gamma) = \int_{t_0}^{t_1} L(\vec{x}, \dot{\vec{x}}, t) dt \quad (272)$$

sur l'ensemble des courbes passant par deux points donnés  $(t_0, \vec{x}_0)$  et  $(t_1, \vec{x}_1)$  est que l'équation d'Euler-Lagrange :

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{\vec{x}}} \right) - \frac{\partial L}{\partial \vec{x}} = \vec{0} \quad (273)$$

soit vérifiée sur cette courbe.

Il faut noter que si  $\vec{x} = (x_1, x_2, \dots, x_n)$  :

$$\frac{\partial L}{\partial \vec{x}} \quad (274)$$

est une notation synthétique du vecteur :

$$\left( \frac{\partial L}{\partial x_1}, \frac{\partial L}{\partial x_2}, \dots, \frac{\partial L}{\partial x_n} \right). \quad (275)$$

## 49 Scène VII-4 : L'équation de Lagrange

**EURISTIDE** : Nous avons constitué dans la discussion précédente l'outillage suffisant pour pouvoir maintenant aborder l'application du principe variationnel aux équations de la dynamique de Newton. L'objectif est de trouver une fonctionnelle dont les extrémales seront données par les équations de la dynamique de Newton.

Les équations de la dynamique de Newton peuvent s'écrire sous la forme :

$$\frac{d}{dt} (m_i \vec{v}_i) + \frac{\partial U}{\partial \vec{r}_i} = \vec{0}, \quad (276)$$

où  $U$  est l'énergie potentielle du système, et où les vecteurs  $\vec{r}_i$  sont les rayons-vecteurs des différentes particules du système mécanique. Les rayons-vecteurs sont les vecteurs ayant pour origine l'origine du repère des coordonnées, et pour aboutissement chaque particule.

Sous cette forme, ces équations présentent une analogie évidente avec l'équation :

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{\vec{x}}} \right) - \frac{\partial L}{\partial \vec{x}} = \vec{0}. \quad (277)$$

**BEATRIX** : On va pouvoir constituer la fonctionnelle dont l'extrémale est déterminée par cette équation. Il faut faire intervenir l'énergie potentielle  $U$  et sans doute l'énergie cinétique.

## 50 Scène VII-5 : Le principe de moindre action

**EURISTIDE** : Considérons la fonction  $L$  suivante :

$$L = T - U, \quad (278)$$



où  $T$  est l'énergie cinétique et  $U$  est l'énergie potentielle.

Alors :

$$L = \frac{1}{2} m_i \vec{r}_i^2 - U, \quad (279)$$

où  $\vec{r}_i$  est la vitesse des particules du corps considéré.

On a bien :

$$\begin{aligned} \frac{\partial L}{\partial \vec{x}} &= m_i \vec{r}_i \\ \frac{\partial L}{\partial \vec{x}} &= -\frac{\partial U}{\partial \vec{r}_i} \end{aligned} \quad (280)$$

D'où, il s'ensuit bien que :

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \vec{x}} \right) - \frac{\partial L}{\partial \vec{x}} = \frac{d}{dt} (m_i \vec{r}_i) + \frac{\partial U}{\partial \vec{r}_i} = \vec{0} \quad (281)$$

correspond aux lois de la dynamique de Newton.

**BEATRIX** : Par conséquent, on peut interpréter cela en disant que les mouvements d'un système mécanique coïncident avec les extrémales de la fonctionnelle :

$$\Phi(\gamma) = \int_{t_0}^{t_1} (T - U) dt. \quad (282)$$

**EURISTIDE** : L'équation d'Euler-Lagrange appliquée à cette fonctionnelle  $\Phi(\gamma)$  est appelée **principe de moindre action de Hamilton**. Tout système physique tend à minimiser son action  $\Phi(\gamma)$  :

$$\Phi(\gamma) = \int_{t_0}^{t_1} (T - U) dt. \quad (283)$$

Son mouvement sera donc une extrémale de cette fonctionnelle. Et nous savons comment la déterminer, grâce à l'équation d'Euler-Lagrange.

**BEATRIX** : En conclusion, je dirais que l'Univers est un fainéant !

**EURISTIDE** : En effet. Et c'est une caractéristique de l'action d'un système physique, que d'être minimale lorsqu'elle est calculée sur une trajectoire de longueur minimale entre deux points fixes. Et c'est une propriété de notre Univers que de vouloir minimiser cette action, et par conséquent d'être fainéant comme tu le dis.

## Neuvième partie

# Acte VIII - Le calcul tensoriel

## 51 Scène VIII-1 : L'écriture indicielle

MATHINE : Nous allons maintenant introduire un chapitre des mathématiques dont nous aurons besoin intensivement en relativité générale. Il s'agit du calcul tensoriel.

Nous allons commencer par introduire une convention pour l'écriture des sommes, qui a été mise au point par Einstein.

Les notations de sommes de dérivées partielles ou de diverses coordonnées constituent généralement des notations rapidement lourdes. Une solution consiste à utiliser le signe de sommation  $\Sigma$ , comme nous l'avons fait à de multiples reprises, et qui permet de diminuer la taille des expressions algébriques. Par exemple, on écrit :

$$a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 = \sum_{i=1}^4 a_ix_i. \quad (284)$$

Afin de s'affranchir du signe  $\Sigma$ , Einstein a mis en place une convention consistant à considérer que la présence dans une expression mathématique d'un même indice en haut et en bas, indique qu'il y a sommation sur cet indice. Ainsi, en utilisant cette convention, on peut écrire :

$$a_1x^1 + a_2x^2 + a_3x^3 + a_4x^4 = a_ix^i, \quad (285)$$

ou encore :

$$\sum_{i=1}^n a_iy^i = a_iy^i. \quad (286)$$

Dans cette notation, il ne faut pas confondre les indices placés en haut avec des exposants de la fonction mathématique "puissance". C'est le contexte du calcul qui permet d'éviter la confusion.

Par exemple, on peut aussi écrire :

$$a_i^1b_2^i = \sum_{i=1}^n a_i^1b_2^i, \quad (287)$$

ou encore, dans un autre exemple, la différentielle totale de  $u(x, y, z)$  s'écrit :

$$du = \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy + \frac{\partial u}{\partial z}dz. \quad (288)$$

En notant :

$$\begin{aligned} x &= x^0 \\ y &= x^1 \\ z &= x^2, \end{aligned} \quad (289)$$

on peut écrire avec la convention d'Einstein :

$$du = \frac{\partial u}{\partial x^i}dx^i. \quad (290)$$

On voit que dans cette notation, un indice haut au dénominateur d'une fraction correspond à un indice bas au numérateur d'une fraction, ou en dehors d'une fraction.

Nous aurons également besoin de deux symboles importants, nous permettant d'exprimer aisément certaines sommes.

Le premier d'entre eux est défini par :

$$\delta_{ij} = \begin{cases} 1 & \text{pour } i = j \\ 0 & \text{pour } i \neq j \end{cases} . \quad (291)$$

Par exemple, pour exprimer la distance infinitésimale parcourue le long d'une courbe dans un espace de dimension  $n$ , il faut écrire :

$$ds^2 = \sum_{i=1}^n (dx^i)^2 = \delta_{ij} dx^i dx^j . \quad (292)$$

Ce symbole est appelé symbole de Kronecker.

Le deuxième symbole dont nous aurons besoin est défini par :

$$\epsilon^{ij} = \begin{cases} 0 & \text{pour } i = j \\ 1 & \text{pour } i = 1, j = 2 \\ -1 & \text{pour } i = 2, j = 1 \end{cases} . \quad (293)$$

Ce qui donne les valeurs suivantes :

$$\begin{aligned} \epsilon^{11} &= 0 \\ \epsilon^{22} &= 0 \\ \epsilon^{12} &= 1 \\ \epsilon^{21} &= -1. \end{aligned} \quad (294)$$

Ce symbole s'appelle symbole d'antisymétrie.

On voit immédiatement que la relation suivante est vérifiée :

$$\epsilon^{ij} = -\epsilon^{ji} . \quad (295)$$

Prenons un exemple illustrant l'utilisation de symbole d'antisymétrie :

$$\begin{aligned} \epsilon^{ij} a_i^1 a_j^2 &= \epsilon^{11} a_1^1 a_1^2 + \epsilon^{12} a_1^1 a_2^2 + \epsilon^{21} a_2^1 a_1^2 + \epsilon^{22} a_2^1 a_2^2 \\ &= a_1^1 a_2^2 - a_2^1 a_1^2 \\ &= \begin{vmatrix} a_1^1 & a_2^1 \\ a_1^2 & a_2^2 \end{vmatrix} . \end{aligned} \quad (296)$$

**BEATRIX** : C'est assez pratique. C'est l'expression simplifiée du déterminant d'une matrice  $2 \times 2$ .

**MATHINE** : On peut étendre cette notion de symbole d'antisymétrie à plus de 2 indices. Pour illustrer cette généralisation, regardons ce qu'est le symbole d'antisymétrie pour 3 indices :

$$\epsilon^{ijk} = \begin{cases} 0 & \text{si deux quelconques des indices ont même valeur} \\ 1 & \text{si les indices résultent d'un nombre pair de permutations de 123} \\ -1 & \text{si les indices résultent d'un nombre impair de permutations de 123} \end{cases} . \quad (297)$$

Ainsi par exemple :

$$\begin{aligned}
 \epsilon^{221} &= 0 \\
 \epsilon^{123} &= 1 \quad (0 \text{ permutations}) \\
 \epsilon^{132} &= -1 \quad (1 \text{ permutation}) \\
 \epsilon^{231} &= 1 \quad (2 \text{ permutations}).
 \end{aligned} \tag{298}$$

On peut écrire alors :

$$\begin{aligned}
 \begin{vmatrix} a_1^1 & a_1^2 & a_1^3 \\ a_2^1 & a_2^2 & a_2^3 \\ a_3^1 & a_3^2 & a_3^3 \end{vmatrix} &= a_1^1(a_2^3a_3^3 - a_2^2a_3^2) \\
 &- a_2^1(a_1^2a_3^3 - a_2^2a_3^1) \\
 &+ a_3^1(a_1^2a_2^2 - a_2^2a_1^1) \\
 &= a_1^1a_2^2a_3^3 - a_1^1a_3^2a_2^3 \\
 &- a_2^1a_1^2a_3^3 + a_2^1a_3^2a_1^1 \\
 &+ a_3^1a_1^2a_2^3 - a_3^1a_2^2a_1^1 \\
 &= \epsilon^{123}a_1^1a_2^2a_3^3 + \epsilon^{132}a_1^1a_3^2a_2^3 \\
 &+ \epsilon^{213}a_2^1a_1^2a_3^3 + \epsilon^{231}a_2^1a_3^2a_1^1 \\
 &+ \epsilon^{312}a_3^1a_1^2a_2^3 + \epsilon^{321}a_3^1a_2^2a_1^1 \\
 &= \epsilon^{ijk} a_i^1 a_j^2 a_k^3.
 \end{aligned} \tag{299}$$

Voilà. Cette notation avec des indices en haut et en bas peut paraître un peu compliquée. Mais on verra qu'elle nous sera extrêmement utile lorsque nous aborderons les tenseurs.

## 52 Scène VIII-2 : Le changement de coordonnées

EURISTIDE : En utilisant cette notation bien utile, nous allons maintenant regarder de plus près ce qu'est un changement de coordonnées dans un espace de dimension  $n$ .

BEATRIX : Mais avant, il peut être utile de rappeler ce que c'est qu'un système de coordonnées. Je me souviens qu'on peut avoir différents systèmes de coordonnées : cartésiens, non cartésiens, curvilignes ou non.

MATHINE : Oui, on va passer en revue différents systèmes de coordonnées pour fixer les idées sur leurs différents types.

Les coordonnées cartésiennes orthogonales constituent le système de référence naturel que nous utilisons tous spontanément, sur un plan de ville ou sur une carte géographique, l'un des axes représentant la direction Ouest-Est et l'autre la direction Sud-Nord. Nous notons  $x^1$  le premier axe et  $x^2$  le second axe, pour nous habituer dès à présent à employer la notation des indices haut qui nous sera nécessaire dans les développements sur les tenseurs :

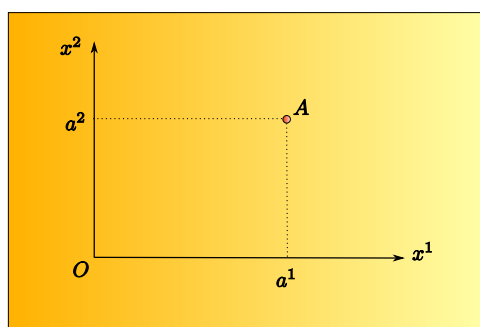


Fig. 78 - Coordonnées cartésiennes orthogonales

On peut aussi utiliser un système de coordonnées obliques, par exemple en perspective. Les axes de coordonnées sont rectilignes, mais sont placés suivant un angle  $\alpha$  qui n'est pas droit. De même que précédemment, le premier axe est noté  $x^1$  et le second  $x^2$ .

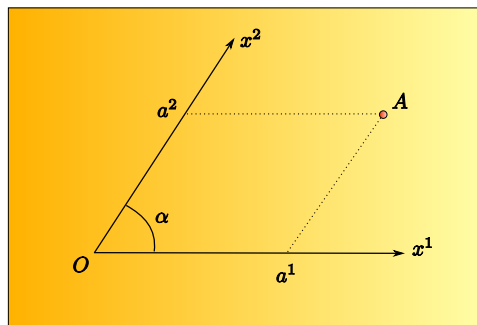


Fig. 79 - Coordonnées obliques

Les coordonnées polaires sont très utilisées pour décrire les phénomènes présentant une symétrie par rapport à un point (on dit une symétrie centrale), par exemple les mouvements d'un point tournant autour d'un point fixe, ou les mouvements d'un mobile oscillant autour d'un axe. La première coordonnée d'un point  $A$  notée  $x^1$  est la distance du point à l'origine  $O$ ; la seconde coordonnée, notée  $x^2$  est l'angle que fait la droite  $OA$  avec un axe fixe de référence :

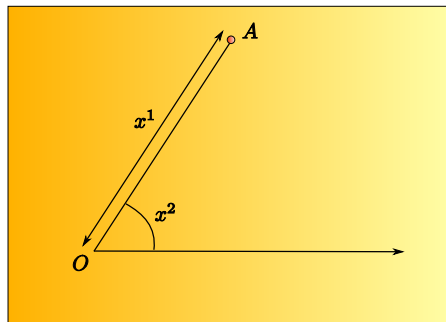


Fig. 80 - Coordonnées polaires

Il est aussi possible d'utiliser deux réseaux de courbes dans le plan pour définir un système de coordonnées. Par exemple, on peut définir un système de coordonnées par les deux réseaux de courbes :

$$\begin{aligned} x^1 &= a \cosh u^1 \cos u^2 \\ x^2 &= a \sinh u^1 \sin u^2 \end{aligned} \quad (300)$$

Nous voyons ci-après la représentation graphique de ces réseaux de courbes :

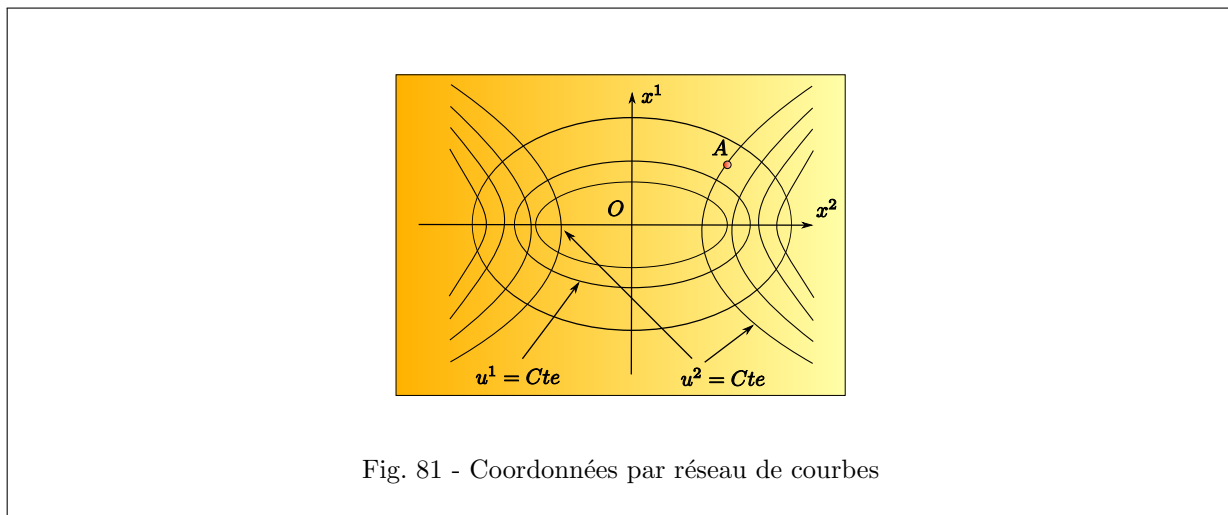


Fig. 81 - Coordonnées par réseau de courbes

On voit ici que lorsque  $u^1$  est constante, et seul  $u^2$  varie, les points de coordonnées  $(x^1, x^2)$  décrivent une ellipse d'équation :

$$\begin{aligned} x^1 &= a \cosh u_{\text{Cte}}^1 \cos u^2 \\ x^2 &= a \sinh u_{\text{Cte}}^1 \sin u^2. \end{aligned} \quad (301)$$

Et lorsque  $u^2$  est constante et seul  $u^1$  varie, les points décrivent une hyperbole d'équation :

$$\begin{aligned} x^1 &= a \cos u_{\text{Cte}}^2 \cosh u^1 \\ x^2 &= a \sin u_{\text{Cte}}^2 \sinh u^1. \end{aligned} \quad (302)$$

BEATRIX : Autrement dit, au lieu de repérer un point dans le plan par sa distance par rapport à des axes perpendiculaires, on repère ce point comme étant l'intersection de deux courbes, ici en l'occurrence une ellipse lorsqu'on fixe la coordonnée  $u^2$  et une hyperbole lorsqu'on fixe la coordonnée  $u^1$ .

MATHINE : On dit que  $u^1 = \text{Cte}$  et  $u^2 = \text{Cte}$  constituent deux réseaux de courbes.

On peut maintenant généraliser cette notion de réseau de courbes, et introduire l'idée de système de coordonnées curvilignes.

Pour cela, considérons une surface quelconque, définie dans l'espace muni d'un système de trois coordonnées cartésiennes orthogonales par l'équation :

$$F(x^1, x^2, x^3) = 0. \quad (303)$$

On considère alors dans ce même système de coordonnées deux fonctions  $u^1$  et  $u^2$  définies par :

$$\begin{aligned} u^1 &= u^1(x^1, x^2, x^3) \\ u^2 &= u^2(x^1, x^2, x^3) \end{aligned} \quad (304)$$

telles que les lignes de réseau  $u^1 = Cte$  et  $u^2 = Cte$  appartiennent à la surface  $F(x^1, x^2, x^3) = 0$  et ne se coupent qu'en un seul point.

On obtient ce qu'on appelle alors des coordonnées de Gauss sur la surface considérée.

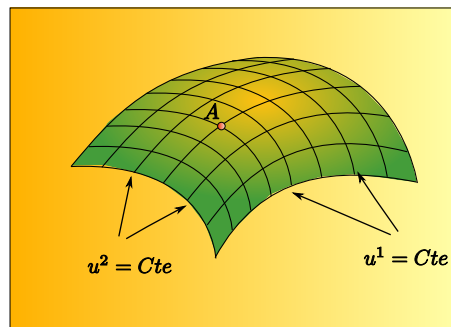


Fig. 82 - Coordonnées de Gauss

Un point  $A$  est déterminé de façon unique comme l'intersection de deux courbes des réseaux  $u^1 = Cte$  et  $u^2 = Cte$ . A noter que nous utilisons toujours la notation des indices hauts.

**BEATRIX** : Pour mieux comprendre, peut-on prendre l'exemple du système de coordonnées à la surface de la Terre : les méridiens et les parallèles ?

**EURISTIDE** : C'est un excellent exemple. On va considérer que la Terre est une sphère de centre  $O$  et de rayon  $R$  :



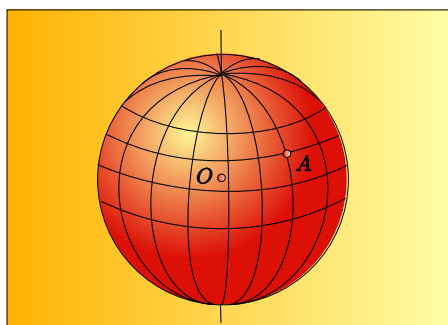


Fig. 83 - Méridiens et parallèles sur la Terre

Les coordonnées curvilignes utilisées habituellement à la surface de la Terre sont constituées du réseau des grands cercles que sont les méridiens et des cercles horizontaux que sont les parallèles. (On appelle grand cercle d'une sphère un cercle de la surface dont le centre est le centre de la sphère. Ils ont donc tous pour rayon le rayon de la sphère.) Les méridiens sont les grands cercles dont le plan est vertical et passe par les pôles. Les parallèles sont des cercles parallèles à l'équateur. L'équateur est le seul parallèle à être un grand cercle.

Un point  $A$  à la surface de la Terre est donc défini par l'intersection d'un méridien et d'une parallèle. Comme il y a deux intersections entre un méridien et une parallèle, on distingue ces intersections en précisant la position par rapport au méridien dit de Greenwich (passant par la ville du même nom) : par exemple,  $35^\circ$  Est ou Ouest.

**MATHINE** : Nous allons maintenant regarder comment décrire un changement de coordonnées curvilignes.

Mais d'abord, regardons ce que sont les vecteurs unitaires de coordonnées dans un tel système. On peut attacher à un système de coordonnées un jeu de vecteurs de coordonnées  $(a_i)$ . Chacun de ces vecteurs est tangent en tout point au réseau des courbes de coordonnées.

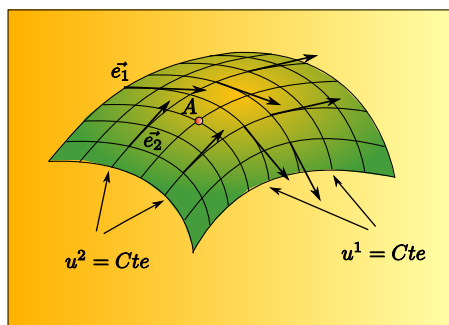


Fig. 84 - Champs de vecteurs de coordonnées

Il est évident que dans ce cas, les vecteurs de coordonnées  $\vec{e}_i$  sont en fait des fonctions-vecteurs, c'est-à-dire des fonctions qui retournent un vecteur en fonction de l'endroit où l'on se trouve dans l'espace. Ces fonctions-vecteurs ne sont donc pas des constantes lorsqu'on se déplace le long des courbes du système de coordonnées, mais sont des fonctions de leur position sur ces courbes. Ils restent toutefois tangents en tout point aux courbes de coordonnées comme le montre la figure précédente.

Considérons un changement de coordonnées curvilignes qui va décrire comment les coordonnées  $(x^1, x^2, x^3)$  sont construites à partir des coordonnées  $(y^1, y^2, y^3)$ . Ce changement de coordonnées est défini par :

$$\begin{aligned}x^1 &= x^1(y^1, y^2, y^3) \\x^2 &= x^2(y^1, y^2, y^3) \\x^3 &= x^3(y^1, y^2, y^3),\end{aligned}\tag{305}$$

ou encore, inversement, les fonctions suivantes définissent comment les coordonnées  $(y^1, y^2, y^3)$  sont construites à partir des coordonnées  $(x^1, x^2, x^3)$  :

$$\begin{aligned}y^1 &= y^1(x^1, x^2, x^3) \\y^2 &= y^2(x^1, x^2, x^3) \\y^3 &= y^3(x^1, x^2, x^3).\end{aligned}$$

Supposons que les  $\vec{e}_i$  soient les fonctions-vecteurs de coordonnées du système  $x^i$  et  $\vec{e}_j$  les fonctions-vecteurs des coordonnées du système  $y^j$ . Il faut noter que nous avons utilisé pour les fonctions-vecteurs de coordonnées  $\vec{e}_i$  et  $\vec{e}_j$  des indices bas et pour les coordonnées elles-mêmes des indices haut. Patience. Nous verrons la raison fondamentale de cela plus tard. En attendant considérons que nous aurons besoin de faire la somme des produits des  $x^i$  par les  $\vec{e}_i$ , et pour cela nous aurons besoin d'utiliser la convention d'Einstein, donc d'indices haut et bas.

Supposons enfin que, pour ce qui concerne les fonctions-vecteurs, le changement de coordonnées s'exprime par :

$$\vec{e}_i = \beta_i^j \vec{e}_j,\tag{306}$$

et inversement :

$$\vec{e}_j = \alpha_j^i \vec{e}_i.\tag{307}$$

Notons que nous avons utilisé ici la notation d'Einstein pour la sommation des indices. Nous savons que, par définition, les fonctions-vecteurs  $\vec{e}_j$  sont tangents aux courbes  $y^j$ . De même, les fonctions-vecteurs  $\vec{e}_i$  sont tangents aux courbes  $x^i$ . Comme la relation qui définit le changement de coordonnées doit dépendre du point dans l'espace, nous allons nous intéresser à une zone infiniment petite de l'espace où cette relation pourra être figée. Donc une variation infinitésimale de la coordonnée  $x^i$  sera assimilable à une variation infinitésimale selon la direction de la fonction-vecteur de coordonnées  $\vec{e}_i$ .

La relation :

$$\vec{e}_i = \beta_i^j \vec{e}_j, \quad (308)$$

nous permet de déduire que la variation infinitésimale se fera par conséquent suivant la direction  $\beta_i^j \vec{e}_j$ .

Par conséquent, si nous considérons une variation infinitésimale d'une fonction-vecteur  $\vec{a}$  :

$$\vec{da} = dx^i \vec{e}_i, \quad (309)$$

nous avons :

$$\vec{da} = dx^i \beta_i^j \vec{e}_j. \quad (310)$$

Et comme les fonctions-vecteurs de coordonnées  $\vec{e}_j$  sont tangentes aux courbes  $y^j$ , nous savons aussi que :

$$\vec{da} = dy^j \vec{e}_j, \quad (311)$$

et par conséquent, par identification des coefficients des fonctions-vecteurs de coordonnées  $\vec{e}_j$  pour la fonction-vecteur  $\vec{da}$ , nous obtenons que la variation de la coordonnée  $y^j$  se fera suivant la relation :

$$dy^j = \beta_i^j dx^i. \quad (312)$$

Inversement, nous pouvons aussi définir des coefficients  $\alpha_j^i$  tels que :

$$\vec{e}_j = \alpha_j^i \vec{e}_i, \quad (313)$$

et par conséquent en appliquant la même méthode que précédemment, nous trouvons :

$$dx^i = \alpha_j^i dy^j. \quad (314)$$

Par ailleurs, nous connaissons l'expression générale des différentielles totales :

$$\begin{aligned} dx^i &= \frac{\partial x^i}{\partial y^j} dy^j \\ dy^j &= \frac{\partial y^j}{\partial x^i} dx^i. \end{aligned} \quad (315)$$

Par conséquent, en identifiant les termes dans les deux expressions de  $dy^j$  et de  $dx^i$ , nous obtenons :

$$\begin{aligned} \alpha_j^i &= \frac{\partial x^i}{\partial y^j} \\ \beta_i^j &= \frac{\partial y^j}{\partial x^i}. \end{aligned} \quad (316)$$

Et en conséquence, nous pouvons écrire maintenant les lois de transformation des fonctions-vecteurs de coordonnées dans un changement de coordonnées :

$$\begin{aligned} \vec{e}_j &= \frac{\partial x^i}{\partial y^j} \vec{e}_i \\ \vec{e}_i &= \frac{\partial y^j}{\partial x^i} \vec{e}_j. \end{aligned} \quad (317)$$

EURISTIDE : C'est un résultat très important qu'il faut un peu commenter. Nous venons de voir que lorsque nous effectuons un changement de coordonnées curvilignes, les coefficients permettant de calculer ce changement de coordonnées, qui transforment les fonctions-vecteurs de coordonnées du premier système dans le second système, sont les dérivées partielles des composantes du premier système par rapport aux composantes du second système de coordonnées. C'était assez prévisible, en fait, puisque pour déduire la nouvelle coordonnée, nous allons faire varier  $x^i$  de façon infinitésimale et observer la variation infinitésimale conséquente de chacun des  $y^j$ . Cette variation est définie, d'après l'expression de la différentielle totale, par l'expression :

$$\frac{\partial x^i}{\partial y^j} dy^j. \quad (318)$$

Nous voyons immédiatement que les coefficients de transformation ne sont pas des constantes. C'est ce qui caractérise un système de coordonnées curvilignes, par rapport à un système de coordonnées rectiligne classique.

BEATRIX : En fait, le changement de coordonnées résulte de la somme des variations des coordonnées sur chacun des "axes" curvilignes. Pour obtenir ces variations, on doit forcément utiliser la dérivée partielle qui nous indique la variation partielle par rapport à chacune des coordonnées.

### 53 Scène VIII-3 : La notion de variance

MATHINE : Nous avons vu dans l'analyse du changement de coordonnées que :

$$\vec{e}_j = \frac{\partial x^i}{\partial y^j} \vec{e}_i. \quad (319)$$

On dit que ces fonctions-vecteurs de coordonnées se transforment suivant une loi covariante, où les variations de coordonnées  $y^j$  (au dénominateur) appartenant au nouveau système de coordonnées, induisent la transformation des vecteurs  $\vec{e}_j$ , dans le nouveau système de coordonnées.

EURISTIDE : La notion de loi covariante signifie ici intuitivement que les fonctions-vecteurs de coordonnées du nouveau système de coordonnées varient en même temps que les coordonnées appartenant au nouveau système.

On dit que la fonction-vecteur de coordonnées  $\vec{e}_j$  est une grandeur **covariante**.

Considérons maintenant une fonction-vecteur  $\vec{a}$  dont les coordonnées dans le système de fonctions-vecteurs de coordonnées  $\vec{e}_i$  sont notées  $a^i$  et celles dans le système de fonctions-vecteurs de coordonnées  $\vec{e}'_j$  dont les coordonnées sont notées  $a'^j$ .

Alors, nous avons, par définition :

$$\vec{a} = a^i \cdot \vec{e}_i, \quad (320)$$

et :

$$\vec{a} = a'^j \cdot \vec{e}'_j. \quad (321)$$

Par conséquent, en identifiant les deux expressions de la fonction-vecteur  $\vec{a}$  :

$$a^i \vec{e}_i = a'^j \vec{e}'_j, \quad (322)$$

d'où :

$$\vec{e}_j = \frac{a^i}{a'^j} \vec{e}_i, \quad (323)$$

et par conséquent, par identification avec l'expression :

$$\vec{e}_j = \frac{\partial x^i}{\partial y^j} \vec{e}_i, \quad (324)$$

nous obtenons :

$$\frac{a^i}{a'^j} = \frac{\partial x^i}{\partial y^j}, \quad (325)$$

puis :

$$a'^j = \frac{\partial y^j}{\partial x^i} a^i. \quad (326)$$

Les coordonnées d'une fonction-vecteur, contrairement aux fonctions-vecteurs elles-mêmes, se transforment suivant une loi contravariante, où ce sont les variations de coordonnées du système d'origine  $x^i$  (au dénominateur) qui induisent la transformation des coordonnées du vecteur dans le système résultant.

On dit que les coordonnées  $a^i$  d'un vecteur sont des grandeurs **contravariantes**.

Par convention, les grandeurs covariantes ont des indices bas, et les grandeurs contravariantes ont des indices hauts.

BEATRIX : Ah, je comprends mieux pourquoi nous manipulons depuis tout à l'heure des indices hauts et bas. Tout cela est très cohérent. Dans les termes d'une somme avec un terme à indice bas, et un terme à indice haut, suivant la convention d'Einstein, il y a donc forcément un terme covariant et un terme contravariant. Et on a vu que les fonctions-vecteurs en général sont covariantes, et leurs coordonnées en général sont contravariantes.

## 54 Scène VIII-4 : Introduction du tenseur

EURISTIDE : Bien, ceci étant dit, nous sommes donc mûrs maintenant pour parler des tenseurs.

BEATRIX : Qu'est-ce qu'un tenseur ?

EURISTIDE : Il faut voir un **tenseur** comme une fonction capable de transformer des fonctions-vecteurs en un ensemble de fonctions-vecteurs ou un nombre.

Un tenseur qui fournit en sortie une fonction-vecteur est dit tenseur une fois contravariant.

Un tenseur qui fournit en sortie deux fonctions-vecteurs est dit tenseur deux fois contravariant, ou contravariant de valence 2.

Un tenseur qui transforme une fonction-vecteur en un nombre, est dit tenseur covariant.

Un tenseur qui transforme deux fonctions-vecteurs en un nombre, est dit tenseur deux fois covariant, ou covariant de valence 2.

On peut imaginer des tenseurs mixtes. Par exemple, un tenseur qui transforme une fonction-vecteur en une fonction-vecteur est un tenseur mixte de valence 2 une fois covariant et une fois contravariant. Un tenseur qui transforme deux fonctions-vecteurs en trois fonctions-vecteurs est un tenseur de valence 5 deux fois covariant et trois fois contravariant.

Un tenseur de valence 0 est une fonction réelle. On peut donc considérer qu'un tenseur est une machine à transformer un nombre ou des fonctions-vecteurs en nombre ou en un ensemble de fonctions-vecteurs. Un tenseur de valence 1 une fois contravariant transforme un nombre en fonction-vecteur. Un tenseur de valence 1 une fois covariant transforme une fonction-vecteur en nombre.

BEATRIX : Pouvez-vous donner un exemple simple de tenseur ?

EURISTIDE : Un tenseur peut être utilisé par exemple pour effectuer la dilatation par un coefficient  $k$  d'une fonction-vecteur  $\vec{a}$ . En notant  $T$  le tenseur et  $\vec{b}$  la fonction-vecteur résultante, on a :

$$\vec{b} = T\vec{a}. \quad (327)$$

Supposons que  $\vec{b}$  ait pour coordonnées  $b^\beta$ .

Supposons que  $\vec{a}$  ait pour coordonnées  $a^\alpha$ .

Le tenseur  $T$  possède un jeu de composantes permettant d'effectuer la transformation recherchée. On peut représenter un tenseur par une matrice, dont les composantes sont les composantes du tenseur. Supposons que le tenseur  $T$  ait pour composantes  $T_\alpha^\beta$ , alors on peut écrire :

$$b^\beta = T_\alpha^\beta a^\alpha. \quad (328)$$

Il faut noter que nous avons attribué au tenseur  $T$  des composantes cohérentes avec la convention d'Einstein. Pour réaliser avec ce tenseur la fonction de dilatation par le coefficient  $k$ , il faut que :

$$\begin{aligned} T_\alpha^\beta &= 0 \quad \text{pour } \alpha \neq \beta \\ T_\alpha^\beta &= k \quad \text{pour } \alpha = \beta. \end{aligned} \quad (329)$$

Le tenseur ainsi défini permet de transformer toute fonction-vecteur  $\vec{a}$  en une fonction-vecteur dilatée par le coefficient  $k$ .

Cet exemple très simple ne présente pas réellement l'intérêt de la notion de tenseur, puisque la dilatation peut être facilement réalisée en multipliant les fonctions-vecteurs par le nombre  $k$  tout simplement. Mais le tenseur permet de réaliser une fonction de dilatation plus complexe, par exemple en dilatant les vecteurs par le coefficient  $k$  le long de l'axe  $x^1$ , par le coefficient  $l$  le long de l'axe  $x^2$  et par le coefficient  $m$  le long de l'axe  $x^3$ . Dans ce cas, la multiplication simple par un nombre des fonctions-vecteurs ne suffit plus, et le tenseur permet de réaliser cette fonction plus complexe. En l'occurrence, un tenseur dans un espace à trois dimensions implémentant cette fonction de dilatation complexe s'écrirait :

$$T = \begin{pmatrix} k & 0 & 0 \\ 0 & l & 0 \\ 0 & 0 & m \end{pmatrix}. \quad (330)$$

Alors, pour une fonction-vecteur  $\vec{a}$  de coordonnées  $(a^1, a^2, a^3)$ , on obtient la dilatation suivante :

$$T \begin{pmatrix} a^1 \\ a^2 \\ a^3 \end{pmatrix} = \begin{pmatrix} k & 0 & 0 \\ 0 & l & 0 \\ 0 & 0 & m \end{pmatrix} \begin{pmatrix} a^1 \\ a^2 \\ a^3 \end{pmatrix} = \begin{pmatrix} ka^1 \\ la^2 \\ ma^3 \end{pmatrix}. \quad (331)$$

BEATRIX : D'accord. Résumons. Un tenseur, c'est une machine qui avale des fonctions-vecteurs et recrache des fonctions-vecteurs. La covariance, c'est le nombre de fonctions-vecteurs avalées, et la contravariance, c'est le nombre de fonctions-vecteurs recrachées. La valence, c'est la somme de la covariance et de la

contravariance.

Mais je serais curieuse de comprendre pourquoi on appelle ces deux caractéristiques précisément covariance et contravariance. Il doit y avoir une raison, dans le cadre des changements de coordonnées, non ?

MATHINE : Oui, il y a de très bonnes raisons pour cela. On va considérer un tenseur une fois covariant et une fois contravariant. Nous avons vu qu'il transformait une fonction-vecteur en une fonction-vecteur. Donc, si on note  $T$  le tenseur, on peut écrire :

$$\vec{b} = T\vec{a}. \quad (332)$$

Et en utilisant les coordonnées des vecteurs et du tenseur :

$$b^\beta = T_\alpha^\beta a^\alpha, \quad (333)$$

où nous avons attribué de nouveau au tenseur  $T$  un ensemble de composantes cohérentes avec la règle d'écriture des sommes d'Einstein. En effet, pour passer des composantes  $a^\alpha$  aux composantes  $b^\beta$ , nous avons besoin d'un ensemble de coefficients, et l'application de la règle de sommation d'Einstein impose l'écriture des indices  $\alpha$  en bas et  $\beta$  en haut.

Ecrivons maintenant une transformation de coordonnées. Nous utiliserons dorénavant les lettres grecques pour l'ancien système de coordonnées, et les lettres romaines pour le nouveau système de coordonnées. Les quantités  $b^\beta$  et  $a^\alpha$  sont contravariantes, donc :

$$\begin{aligned} b'^j &= \frac{\partial x^j}{\partial y^\beta} b^\beta \\ a'^i &= \frac{\partial x^i}{\partial y^\alpha} a^\alpha. \end{aligned} \quad (334)$$

En injectant ces données dans l'expression de la transformation réalisée par le tenseur  $T_\alpha^\beta$ , on obtient :

$$b'^j \frac{\partial y^\beta}{\partial x^j} = T_\alpha^\beta a'^i \frac{\partial y^\alpha}{\partial x^i}. \quad (335)$$

Autrement dit :

$$b'^j = T_\alpha^\beta \frac{\partial y^\alpha}{\partial x^i} \frac{\partial x^j}{\partial y^\beta} a'^i. \quad (336)$$

Tout se passe donc comme si le tenseur  $T_\alpha^\beta$  qui est une fois covariant et une fois contravariant avait été transformé en un tenseur également une fois covariant et une fois contravariant  $T'^j_i$  dont l'expression est :

$$T'^j_i = T_\alpha^\beta \frac{\partial y^\alpha}{\partial x^i} \frac{\partial x^j}{\partial y^\beta}. \quad (337)$$

EURISTIDE : En d'autres termes, notre tenseur se transforme de façon covariante pour son indice  $\alpha$  et de façon contravariante pour son indice  $\beta$ . C'est d'ailleurs ce qui caractérise un tenseur. Si une grandeur est transformée comme indiqué dans l'équation précédente lors d'un changement de coordonnées, alors c'est un tenseur. Si cette grandeur ne se transforme pas ainsi, alors ce n'est pas un tenseur.

BEATRIX : On a donc bien fait de l'appeler une fois covariant et une fois contravariant. En plus, on a bien fait de lui mettre un indice haut et un indice bas. En bref, tout cela est très cohérent : la notation d'Einstein, la covariance et la contravariance, et l'appellation des tenseurs. Bravo, les matheux !

EURISTIDE : On confirme donc qu'on peut assimiler un tenseur une fois covariant à une fonction-vecteur, puisqu'il possède les mêmes propriétés de transformation qu'une fonction-vecteur.

MATHINE : On peut effectuer des opérations sur les tenseurs. On peut ajouter deux tenseurs de mêmes caractéristiques, en additionnant simplement leurs composantes, et on obtient un tenseur de mêmes caractéristiques que les deux premiers.

BEATRIX : Qu'est-ce au juste que les caractéristiques d'un tenseur ?

MATHINE : Un tenseur possède deux caractéristiques : sa valence (somme de la contravariance et de la covariance) et le nombre de dimensions de l'espace de référence dans lequel nous travaillons. Ainsi, la valence détermine le nombre d'indices des composantes du tenseur (et la covariance donne le nombre d'indices bas et la contravariance le nombre d'indices hauts), et le nombre de dimensions de l'espace de référence détermine le nombre de valeurs entre 1 et  $n$  que peuvent prendre les indices.

Si deux tenseurs possèdent les mêmes caractéristiques, on peut donc les additionner. Regardons comment cela se passe ; nous allons caractériser les tenseurs que nous utiliserons en exprimant leur loi de changement de coordonnées, exprimant explicitement et concrètement leur valence ; soit deux tenseurs  $A_{\alpha\beta}$  et  $B_{\alpha\beta}$ , dont nous observons les transformations de coordonnées suivantes :

$$\begin{aligned} A'_{ij} &= \frac{\partial x^\alpha}{\partial y^i} \frac{\partial x^\beta}{\partial y^j} A_{\alpha\beta} \\ B'_{ij} &= \frac{\partial x^\alpha}{\partial y^i} \frac{\partial x^\beta}{\partial y^j} B_{\alpha\beta}. \end{aligned} \quad (338)$$

Alors :

$$A'_{ij} + B'_{ij} = \frac{\partial x^\alpha}{\partial y^i} \frac{\partial x^\beta}{\partial y^j} (A_{\alpha\beta} + B_{\alpha\beta}). \quad (339)$$

BEATRIX : Donc pour additionner deux tenseurs, on additionne tout simplement terme à terme leurs composantes, à condition qu'ils aient les mêmes caractéristiques.

MATHINE : La multiplication de tenseurs quelconques se fait comme suit :

$$\begin{aligned} A'_{ij} &= \frac{\partial x^\alpha}{\partial y^i} \frac{\partial x^\beta}{\partial y^j} A_{\alpha\beta} \\ B'^k &= \frac{\partial y^k}{\partial x^\gamma} B^\gamma \\ A'_{ij} B'^k &= \frac{\partial x^\alpha}{\partial y^i} \frac{\partial x^\beta}{\partial y^j} \frac{\partial y^k}{\partial x^\gamma} A_{\alpha\beta} B^\gamma. \end{aligned} \quad (340)$$

BEATRIX : Même chose, je suppose : pour multiplier deux tenseurs, on multiplie terme à terme leurs composantes, en prenant garde que les tenseurs n'aient pas d'indice de même nom.



MATHINE : Justement, il y a une opération très importante sur les tenseurs qui s'appelle la **contraction**. Cette opération permet d'abaisser la valence d'un tenseur en donnant le même nom à deux indices différents. Alors, la règle d'Einstein est appliquée et les composantes correspondantes du tenseur sont sommées entre elles.

Soit un tenseur  $A_{\alpha\beta}^{\gamma}$ , deux fois covariant et une fois contravariant, caractérisé par la transformation de coordonnées suivante :

$$A'^k_{ij} = \frac{\partial x^\alpha}{\partial y^i} \frac{\partial x^\beta}{\partial y^j} \frac{\partial y^k}{\partial x^\gamma} A_{\alpha\beta}^{\gamma}. \quad (341)$$

En posant  $\beta = \gamma$ , on obtient un nouveau tenseur une fois covariant :

$$A'^k_{ij} = \frac{\partial x^\alpha}{\partial y^i} \frac{\partial x^\beta}{\partial y^j} \frac{\partial y^k}{\partial x^\beta} A_{\alpha\beta}^{\beta}. \quad (342)$$

Or, nous savons que :

$$\frac{\partial x^\beta}{\partial y^j} \frac{\partial y^k}{\partial x^\beta} = \delta_j^k. \quad (343)$$

Donc :

$$A'^k_{ij} = \delta_j^k \frac{\partial x^\alpha}{\partial y^i} A_{\alpha\beta}^{\beta}. \quad (344)$$

Donc la présence du terme  $\delta_j^k$  indique que tous les termes où  $j \neq k$  sont nuls, donc il reste :

$$A'^j_{ij} = \frac{\partial x^\alpha}{\partial y^i} A_{\alpha\beta}^{\beta}. \quad (345)$$

La multiplication contractée est une opération comparable permettant d'effectuer une contraction d'un tenseur en le multipliant par un autre tenseur avec un ou plusieurs indices en commun. Si :

$$C_{\alpha\beta}^{\gamma} = A_{\alpha\beta} B^{\gamma}, \quad (346)$$

alors, par exemple :

$$C_{\alpha\beta}^{\beta} = A_{\alpha\beta} B^{\beta}. \quad (347)$$

## 55 Scène VIII-5 : La symétrie

BEATRIX : C'est assez simple finalement. Il suffit de jouer avec les indices.

MATHINE : Pour finir l'énumération des propriétés des tenseurs, regardons les propriétés de symétrie des tenseurs. Soit  $A^{\alpha\beta}$  un tenseur.

On dit que ce tenseur est symétrique si :

$$A^{\alpha\beta} = A^{\beta\alpha}. \quad (348)$$

On dit que ce tenseur est antisymétrique si :

$$A^{\alpha\beta} = -A^{\beta\alpha}. \quad (349)$$

On définit les mêmes notions pour d'autres tenseurs de caractéristiques différentes, suivant les mêmes règles.

## 56 Scène VIII-6 : La métrique de l'espace euclidien

EURISTIDE : Nous allons maintenant étudier un tenseur particulier attaché à la notion de métrique dans un espace. Nous verrons plus tard qu'une métrique dans un espace est un outil permettant de spécifier les caractéristiques topologiques de cet espace (longueurs, courbure, etc.)

MATHINE : Oui, mais commençons par rappeler ce qu'est un produit scalaire dans un système de coordonnées cartésiennes. Si on considère deux fonctions-vecteurs  $\vec{a}$  et  $\vec{b}$  de coordonnées respectives  $a^\alpha$  et  $b^\beta$ , leur produit scalaire dans un système de coordonnées cartésien, est :

$$\vec{a} \cdot \vec{b} = \delta_{\alpha\beta} a^\alpha b^\beta. \quad (350)$$

En effet, le produit scalaire de  $\vec{a}$  par  $\vec{b}$  est la somme des produits deux à deux des coordonnées des deux fonctions-vecteurs.

Donc :

$$\vec{a} \cdot \vec{b} = a^1 b^1 + a^2 b^2 + \dots \quad (351)$$

Ce qui peut s'écrire au moyen de  $\delta_{\alpha\beta}$  (les termes  $\delta_{\alpha\beta}$  où  $\alpha \neq \beta$  étant nuls, et les termes  $\delta_{\alpha\beta}$  où  $\alpha = \beta$  étant égaux à 1) :

$$\vec{a} \cdot \vec{b} = \delta_{11} a^1 b^1 + \delta_{22} a^2 b^2 + \dots + \delta_{12} a^1 b^2 + \delta_{21} a^2 b^1 + \dots \quad (352)$$

La norme d'un vecteur  $\text{nor}(\vec{a})$  est :

$$\text{nor}(\vec{a}) = \vec{a} \cdot \vec{a} = \delta_{\alpha\beta} a^\alpha a^\beta, \quad (353)$$

et le module du vecteur (c'est-à-dire sa longueur) est :

$$|\vec{a}| = \sqrt{\delta_{\alpha\beta} a^\alpha a^\beta}. \quad (354)$$

Enfin, suivant le même principe, l'accroissement infiniment petit de distance est tel que :

$$\boxed{ds^2 = \delta_{\alpha\beta} dx^\alpha dx^\beta}. \quad (355)$$

EURISTIDE : Nous allons maintenant quitter le système de coordonnées cartésien et nous placer, par transformation de coordonnées, dans un système de coordonnées curvilignes. Nous allons voir que les choses changent pour ce qui concerne le produit scalaire, la norme et l'accroissement infinitésimal. Mais cependant, nous allons trouver, grâce aux tenseurs, une belle similitude dans l'expression de ces quantités, entre un système de coordonnées cartésien et un système de coordonnées curviligne.

MATHINE : Alors, allons-y. Considérons un changement de coordonnées défini par :

$$\begin{aligned} y^i &= y^i(x^1, x^2, \dots, x^n) \\ x^\alpha &= x^\alpha(y^1, y^2, \dots, y^n). \end{aligned} \quad (356)$$

On a, pour nos deux vecteurs-fonctions  $\vec{a}$  et  $\vec{b}$  :

$$\begin{aligned} a^\alpha &= \frac{\partial x^\alpha}{\partial y^i} a^i \\ b^\beta &= \frac{\partial x^\beta}{\partial y^j} b^j. \end{aligned} \quad (357)$$

Par conséquent :

$$\vec{a} \cdot \vec{b} = \delta_{\alpha\beta} a^\alpha b^\beta = \delta_{\alpha\beta} \frac{\partial x^\alpha}{\partial y^i} \frac{\partial x^\beta}{\partial y^j} a^i b^j. \quad (358)$$

En notant :

$$g_{ij} = \delta_{\alpha\beta} \frac{\partial x^\alpha}{\partial y^i} \frac{\partial x^\beta}{\partial y^j}, \quad (359)$$

nous obtenons :

$$\vec{a} \cdot \vec{b} = g_{ij} a^i b^j. \quad (360)$$

L'expression  $g_{ij}$  est appelée **tenseur euclidien fondamental**.

**EURISTIDE** : Ce tenseur, comme son nom l'indique, est fondamental. Il caractérise le système de coordonnées et permet de calculer les produits scalaires, les distances, etc.

Il permet également de calculer l'accroissement infinitésimal de la distance :

$$ds^2 = \delta_{\alpha\beta} dx^\alpha dx^\beta = \delta_{\alpha\beta} \frac{\partial x^\alpha}{\partial y^i} \frac{\partial x^\beta}{\partial y^j} dy^i dy^j = g_{ij} dy^i dy^j. \quad (361)$$

L'expression :

$$ds^2 = g_{ij} dy^i dy^j \quad (362)$$

est appelée forme quadratique fondamentale.

**BEATRIX** : Finalement, à quoi servent les tenseurs ?

**EURISTIDE** : Les tenseurs sont utilisés chaque fois que nous avons besoin de transformer des fonctions-vecteurs en d'autres fonctions-vecteurs ou en un nombre. Le tenseur euclidien fondamental permet de transformer les fonctions-vecteurs infinitésimales de coordonnées en longueur. Nous verrons que les tenseurs sont très utiles lorsque nous voulons modéliser la courbure d'un espace : en effet, nous verrons que la courbure a pour effet de déformer l'angle que font entre eux des couples de fonctions-vecteurs lorsqu'on les déplace parallèlement à une courbe. Le tenseur est donc tout indiqué.

## 57 Scène VIII-7 : Les composantes covariantes

**MATHINE** : Avant de terminer sur cette introduction aux tenseurs, je vais expliquer à quoi correspondent les composantes covariantes et contravariantes d'un vecteur.

Considérons une fonction-vecteur  $\vec{a}$  tel que :

$$\vec{a} = a^j \vec{e}_j. \quad (363)$$

Considérons maintenant l'expression :

$$a_i = \vec{a} \cdot \vec{e}_i. \quad (364)$$

En injectant l'expression de  $\vec{a}$  dans cette dernière expression, nous obtenons :

$$a_i = \vec{e}_j \cdot a^j \vec{e}_i. \quad (365)$$

Les grandeurs  $a_i$  sont appelées **composantes covariantes** de la fonction-vecteur  $\vec{a}$ . Nous avons vu, lors de sa définition, que le produit scalaire de deux vecteurs était la longueur de la projection orthogonale d'un des vecteurs sur l'autre. Donc l'expression précédente montre que les composantes covariantes sont les projections orthogonales de  $\vec{a}$  sur les  $n$  fonctions-vecteurs de base  $\vec{e}_i$ .

Dans un repère orthonormé, les composantes covariantes sont confondues avec les composantes contravariantes d'un vecteur. Mais il n'en va pas de même pour un repère quelconque non orthonormé.

EURISTIDE : Prenons un exemple pour y voir plus clair, et considérons les coordonnées obliques dans le plan  $(u^1, u^2)$  exprimées par :

$$\begin{aligned} x^1 &= u^1 + u^2 \cos \alpha \\ x^2 &= u^2 \sin \alpha. \end{aligned} \quad (366)$$

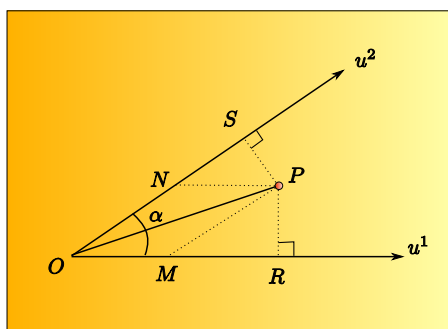


Fig. 85 - Les composantes covariantes d'un vecteur

Les composantes contravariantes classiques du vecteur  $\vec{OP}$  sont :

$$\begin{aligned} OM &= u^1 \\ ON &= u^2. \end{aligned}$$

Les composantes covariantes du vecteur  $\vec{OP}$  sont les projections orthogonales de  $OP$  sur  $u^1$  et  $u^2$ , soit :

$$\begin{aligned} OR &= u_1 \\ OS &= u_2. \end{aligned}$$

---

## Dixième partie

# Acte IX - La relativité restreinte

## 58 Scène IX-1 : La vitesse des interactions

EURISTIDE : Nous possédons maintenant les outils de base. Nous allons pouvoir aborder le domaine des théories de la relativité d'Einstein. Mais avant de nous lancer dans cette aventure passionnante, je propose que nous fassions un petit bilan de ce que nous avons appris dans ces longues discussions d'introduction.

BEATRIX : Nous avons commencé par faire des révisions des bases mathématiques : la notion d'espace géométrique, de systèmes de coordonnées permettant de s'y repérer, et la différence fondamentale entre un espace plat où les vecteurs de coordonnées sont constants et un espace curviligne où les fonctions-vecteurs de coordonnées ne sont pas constantes.

EURISTIDE : Puis nous avons introduit les notions de dérivée, permettant de calculer une vitesse ou une accélération, la notion d'intégration qui est l'opération inverse de la dérivée. Nous avons vu les équations différentielles qui traduisent le comportement d'un système dans l'infiniment petit et permettant de déduire son comportement macroscopique. Nous avons abordé la notion de vecteur également, qui est le support de quantités telles que les forces, les vitesses, les accélérations, etc.

Puis nous avons discuté de la cinématique qui décrit les mouvements des points dans l'espace. Puis la dynamique de Newton nous a permis de comprendre pourquoi les points se meuvent, en décrivant le comportement des points matériels ayant une masse et étant soumis à des forces extérieures. Nous avons étendu cette analyse en éclaircissant la notion d'énergie, et celle de travail d'une force, ainsi qu'en distinguant l'énergie cinétique et l'énergie potentielle, et en dégagant le principe fondamental de la conservation de l'énergie.

BEATRIX : Nous avons pu voir ensuite le comportement d'un système mécanique dans un référentiel accéléré ou en rotation. Nous avons pu ainsi comprendre la dualité qui existe entre accélération et force, dualité que nous retrouverons, je crois, en relativité générale.

EURISTIDE : Nous avons ensuite mis en évidence que le mouvement des solides est la composition du mouvement de leur centre d'inertie, avec leur mouvement propre de rotation autour d'un axe ou de plusieurs axes. Nous avons pu nous rendre compte que le mouvement d'un solide autour d'un axe fixe répond à des lois très similaires à celles de la dynamique du point, en remplaçant simplement les forces par leur moment par rapport à l'axe, la masse par le moment d'inertie et la quantité de mouvement par son moment.

BEATRIX : On a vu ensuite comment le principe de moindre action permet de trouver la trajectoire du mouvement d'un système sous forme d'une équation différentielle issue de l'équation d'Euler-Lagrange.

EURISTIDE : Et puis, nous avons terminé par le calcul tensoriel. Nous avons vu qu'un tenseur est une machine qui mange des fonctions-vecteurs et recrache des fonctions-vecteurs. Le nombre de fonctions-vecteurs mangés est la covariance. Le nombre de fonctions-vecteurs recrachés est la contravariance. Et nous avons abordé le tenseur euclidien fondamental qui caractérise le système de coordonnées d'un espace.

BEATRIX : Et voilà. Nous voilà bien outillés... On attaque donc la relativité maintenant ?

EURISTIDE : Oui, et nous allons commencer par la relativité restreinte, qui est la théorie développée par Einstein pour traduire le fait que la vitesse de la lumière est finie.

Nous avons étudié lors des discussions précédentes la notion de référentiels galiléens : ce sont des référentiels où les corps non soumis à une force extérieure sont doués d'un mouvement à une vitesse constante rectiligne. Si deux référentiels sont animés l'un par rapport à l'autre d'un mouvement rectiligne uniforme et si l'un d'entre eux est galiléen, alors l'autre l'est aussi.

Le **principe de relativité** dit que toutes les lois de la nature sont identiques dans tous les référentiels galiléens. Autrement dit, les équations décrivant les lois de la nature sont identiques et invariantes lorsqu'on passe d'un référentiel galiléen à un autre.

L'expérience montre que la notion d'interaction instantanée n'existe pas dans la nature. Toutes les interactions, quelles qu'elles soient, ont besoin d'un certain temps pour se propager d'un corps à un autre, d'un endroit à un autre. Il existe donc une vitesse limite de déplacement de tous les corps et de toute interaction, car sinon, un corps qui dépasserait cette vitesse limite constituerait lui-même une interaction dont la vitesse serait supérieure à la vitesse limite.

Le fait que les lois de la nature soient identiques dans tous les référentiels entraîne que cette limite de propagation des interactions est la même dans tous les référentiels galiléens. C'est donc une constante universelle.

Cette vitesse constante est aussi la vitesse de propagation de la lumière dans le vide. On la note  $c$ , et sa valeur est :

$$c = 2,99793.10^8 \text{ m.s}^{-1}. \quad (367)$$

Cette vitesse est très grande, et c'est pourquoi, en première approximation, les lois de la physique classique que nous connaissons, et qui considèrent que la vitesse de la lumière est infinie, ou qui considèrent que les interactions peuvent se propager instantanément, restent valables, et n'altèrent pas le niveau de précision recherché pour la plupart des résultats.

Le principe de relativité de Galilée associé à l'existence d'une vitesse limite de propagation des interactions est appelé principe de relativité d'Einstein. La théorie correspondante est appelée Relativité Restreinte d'Einstein.

BEATRIX : Mais en quoi cette vitesse limite change-t-elle quelque chose à la physique ?

EURISTIDE : Le fait que la vitesse de propagation des interactions soit finie et constante a des impacts importants sur la notion de temps. En effet, dans un monde où les interactions se propagent à vitesse infinie, on peut additionner les vitesses, comme nous l'avons fait à plusieurs reprises dans nos discussions précédentes. Par exemple, si un coureur se déplace à 5 km/h sur un tapis roulant se déplaçant lui-même à 2 km/h, sa vitesse est finalement de 7 km/h par rapport au sol immobile.

Considérons maintenant un rayon lumineux émis à partir d'une lampe posée sur un tapis roulant se déplaçant lui-même à 2 km/h. Si la loi d'addition des vitesses que nous connaissons était valable, notre rayon lumineux atteindrait la vitesse de  $c + 2$  km/h.

BEATRIX : Sa vitesse de propagation serait supérieure à la constante universelle  $c$ . C'est impossible !

EURISTIDE : Il y a donc une contradiction dans le raisonnement additif pour les vitesses. Les notions d'additions simples des vitesses ne s'appliquent plus. C'est-à-dire que les notions de vitesse absolue ou de temps absolu n'existent pas. Le temps semble s'écouler différemment dans différents référentiels et induit des lois plus complexes pour l'addition des vitesses.

## 59 Scène IX-2 : La notion d'intervalle

BEATRIX : Bon. Il va falloir déterminer ces lois maintenant.

EURISTIDE : On va devoir introduire un nouvel outil pour cela. Il s'agit de l'espace-temps. Un **événement** est défini par son lieu et son instant. Ainsi, dans un espace physique à trois dimensions, un événement possède 4 coordonnées :  $x$ ,  $y$ ,  $z$  et  $t$ , constituées des 3 coordonnées spatiales habituelles auxquelles on ajoute la coordonnée temporelle. Les points de coordonnées  $(x, y, z, t)$  dans l'espace-temps sont appelés **points d'univers** et représentent à la fois la position et l'instant d'un événement dans l'espace-temps à 4 dimensions.

Le déplacement d'une particule détermine dans l'espace-temps de dimension 4 ce qu'on appelle une **ligne d'univers**. C'est une courbe dans un espace à 4 dimensions. Les points de cette ligne déterminent la position de la particule à chaque instant. Une particule douée d'un mouvement rectiligne uniforme aura pour ligne d'univers une droite.

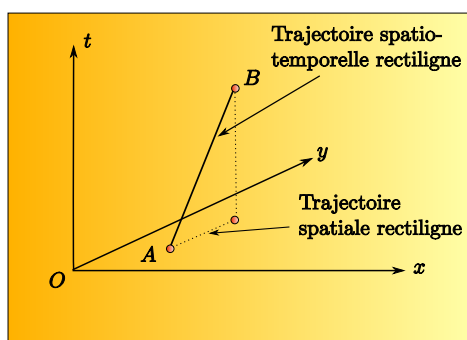


Fig. 86 - Mouvement rectiligne uniforme, droite d'univers

En revanche, dans cet espace-temps, une particule qui suit une trajectoire rectiligne, mais dont la vitesse n'est pas constante, n'aura pas pour ligne d'univers une droite.

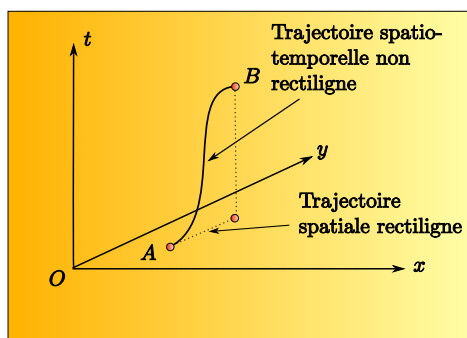
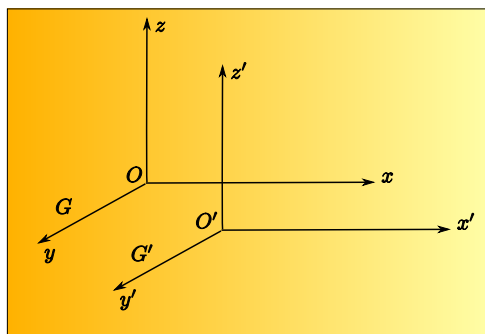


Fig. 87 - Mouvement rectiligne non uniforme

BEATRIX : Ah oui, je comprends maintenant tout l'intérêt de cet espace-temps. Les lignes d'univers droites correspondent au mouvement rectiligne à vitesse constante (dit mouvement rectiligne uniforme), et c'est précisément le mouvement des référentiels galiléens les uns par rapport aux autres. Dans l'espace-temps, la ligne d'univers droite (donc le plus court chemin entre deux événements) est la ligne d'univers des référentiels galiléens.

MATHINE : Nous allons maintenant traduire sous forme mathématique le fait que la vitesse de la lumière est finie et constante dans tous les référentiels. Considérons deux référentiels galiléens  $G$  et  $G'$ , se déplaçant l'un par rapport à l'autre à une vitesse constante  $\vec{v}$ .



Fig. 88 - Référentiels  $G$  et  $G'$ 

On va supposer, pour simplifier les premiers calculs, que les repères  $G$  et  $G'$  ont en commun les axes  $Ox$  et  $O'x'$ , et que leurs axes  $Oy$ ,  $O'y'$ , et  $Oz$ ,  $O'z'$  sont parallèles entre eux. On désigne par  $t$  le temps dans le référentiel  $G$  et par  $t'$  le temps dans le référentiel  $G'$ . La vitesse  $\vec{v}$  sera dirigée dans la direction  $Ox$ . Considérons un événement à l'instant  $t_1$  dans le référentiel  $G$  consistant en l'envoi d'un signal lumineux à partir du point  $(x_1, y_1, z_1)$  à la vitesse de la lumière  $c$ .

Un second événement sera l'arrivée de ce signal dans le référentiel  $G$  au point  $(x_2, y_2, z_2)$  à l'instant  $t_2$ .

**EURISTIDE** : Il y a deux façons de calculer la distance entre les deux points  $(x_1, y_1, z_1)$  et  $(x_2, y_2, z_2)$ . Soit en calculant directement la distance au sens géométrique dans l'espace à trois dimensions. Mathine nous en donnera la formule. Soit en disant que la distance parcourue par le signal lumineux est égale à la vitesse de la lumière  $c$  multipliée par le temps écoulé  $(t_2 - t_1)$ .

**BEATRIX** : Oui, d'accord. Ces deux méthodes de calcul devront donner la même résultat, ce qui nous donnera une équation.

**MATHINE** : En effet ; la distance géométrique des 2 événements dans l'espace à 3 dimensions est :

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}, \quad (368)$$

donc on doit avoir :

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} = c(t_2 - t_1). \quad (369)$$

Ce qui peut s'écrire, en élevant le tout au carré :

$$(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 = c^2(t_2 - t_1)^2. \quad (370)$$

Soit :

$$c^2(t_2 - t_1)^2 - (x_2 - x_1)^2 - (y_2 - y_1)^2 - (z_2 - z_1)^2 = 0. \quad (371)$$

Ou encore :

$$\sqrt{c^2(t_2 - t_1)^2 - (x_2 - x_1)^2 - (y_2 - y_1)^2 - (z_2 - z_1)^2} = 0. \quad (372)$$

EURISTIDE : La quantité :

$$\sqrt{c^2(t_1 - t_1)^2 - (x_2 - x_1)^2 - (y_2 - y_1)^2 - (z_2 - z_1)^2}, \quad (373)$$

est appelée **intervalle** entre deux événements. Elle représente une sorte de distance en 4 dimensions entre deux événements lorsqu'on considère le temps comme une de leurs coordonnées.

MATHINE : Dans le référentiel  $G'$ , on va trouver le même résultat, puisque la vitesse de la lumière est la même :

$$\sqrt{c^2(t'_2 - t'_1)^2 - (x'_2 - x'_1)^2 - (y'_2 - y'_1)^2 - (z'_2 - z'_1)^2} = 0. \quad (374)$$

EURISTIDE : Du fait que la vitesse de la lumière est finie et constante dans tous les référentiels galiléens, l'intervalle entre deux événements est toujours identique, quel que soit le référentiel galiléen dans lequel il est mesuré. Ce résultat est essentiel et traduit l'existence de cette constante universelle qu'est la vitesse de la lumière.

MATHINE : Notons  $l_{12}^2$  l'expression de la distance physique (dans l'espace à 3 dimensions) entre les deux points :

$$l_{12}^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2. \quad (375)$$

Attention, ici le "2" posé en indice haut sur  $l_{12}$  est bien un carré. Ce n'est pas un indice haut dans la notation d'Einstein. En revanche "12" en indice bas est bien une paire d'indices indiquant qu'on considère la distance physique entre les points d'indice 1 et 2.

Notons par  $t_{12}$  l'expression :

$$t_{12} = t_2 - t_1. \quad (376)$$

Le carré de l'intervalle entre les événements  $(t_1, x_1, y_1, z_1)$  et  $(t_2, x_2, y_2, z_2)$  dans le référentiel galiléen  $G$  s'exprime alors par :

$$s_{12}^2 = c^2 t_{12}^2 - l_{12}^2. \quad (377)$$

De la même façon, dans le référentiel galiléen  $G'$ , on obtient :

$$s'_{12}^2 = c^2 t'_{12}{}^2 - l'_{12}{}^2. \quad (378)$$

L'identité des intervalles dans tous les référentiels galiléens s'écrit alors :

$$c^2 t_{12}^2 - l_{12}^2 = c^2 t'_{12}{}^2 - l'_{12}{}^2. \quad (379)$$

EURISTIDE : Partant de cette notion d'événement et d'intervalle, nous allons maintenant parler des événements de genre temps, de genre espace et de genre lumière, ainsi que de l'idée importante de cône de lumière.

Considérons pour cela un événement survenant au centre 0 du système de coordonnées  $(0, x, y, z)$  d'un référentiel galiléen  $G$ . Supposons qu'en ce point 0 un éclair lumineux soit émis, en toutes directions.

Pour simplifier, considérons le comportement de cet éclair lumineux dans le plan  $(x, t)$ . Nous pourrions ainsi facilement représenter sur une feuille de papier les trajectoires. Comme les rayons lumineux se propagent

suivant la relation  $x = ct$ , dans ce plan, ils vont suivre des droites de pente  $1/c$  dans toutes les directions de l'axe  $x$ .

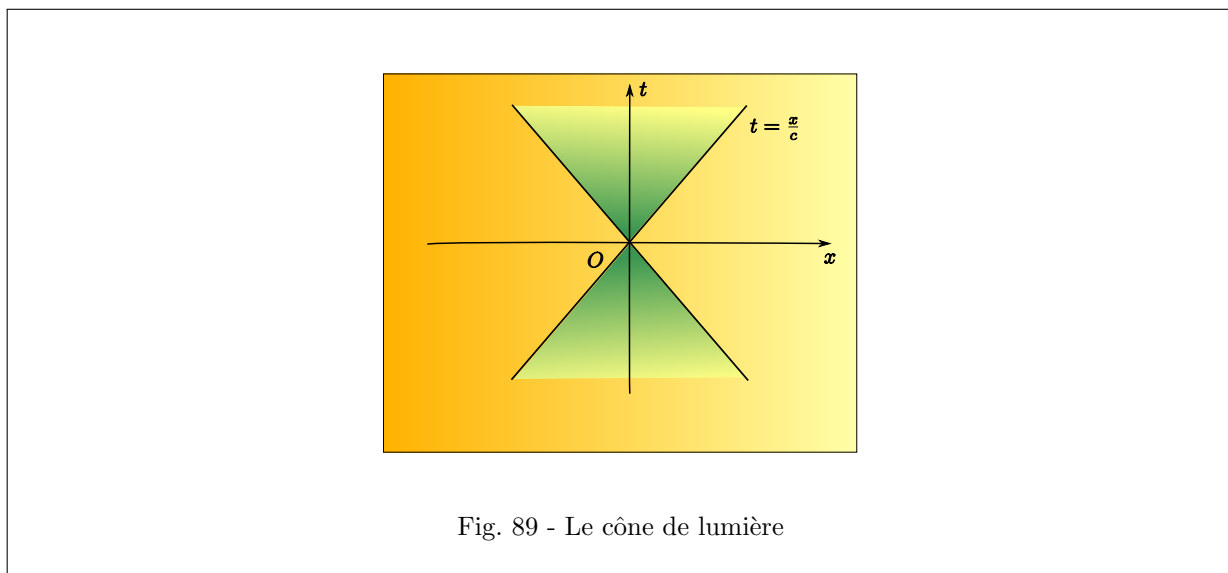


Fig. 89 - Le cône de lumière

**BEATRIX** : Dans l'espace à 2 dimensions, cela correspond à deux triangles tête-bêche se rejoignant en 0. Dans l'espace à 3 dimensions  $(x, y, z)$  cela correspondrait à un cône. Dans l'espace-temps à 4 dimensions, cela correspond à la généralisation du cône, objet qu'on pourrait appeler un hyper-cône.

**EURISTIDE** : Voilà. Maintenant, regardons d'un peu plus près ce cône. On a vu que la vitesse de la lumière est la vitesse limite de propagation de tout événement. Donc aucune ligne d'univers issue de l'origine 0 ne peut comporter de point se trouvant à l'extérieur de ce cône. En effet, si c'était le cas, un tel point correspondrait à un événement parvenu au point  $x$  en un temps  $t > \frac{x}{c}$ . Ce qui signifierait qu'il se serait propagé à une vitesse supérieure à celle de la lumière.

**BEATRIX** : Ce n'est pas possible, on l'a vu. Donc, toutes les lignes d'univers issues de 0 se trouvent cantonnées dans la zone grisée du schéma. Autrement dit, aucune trajectoire de point issue de  $O$  ne peut se trouver en dehors du cône de lumière.

**EURISTIDE** : Si on considère que l'événement à l'origine  $O$  est le présent, en outre, la partie supérieure du cône correspond aux événements du futur par rapport à l'événement origine  $O$ . La partie inférieure du cône correspond aux événements du passé par rapport au présent  $O$ . Tous les événements à l'intérieur du cône sont "accessibles" à un instant donné depuis l'événement  $O$ , puisqu'on peut les atteindre au moyen d'un déplacement à une vitesse inférieure à la vitesse de la lumière. Ces événements sont donc accessibles à un instant donné par la trajectoire d'un événement, ou de façon équivalente, on peut dire qu'ils sont visibles depuis l'événement présent  $O$ , car la lumière qui en est issue a pu parvenir au point  $O$  à l'instant présent.

Pour cette raison, on dit que ces événements sont du **genre temps**.

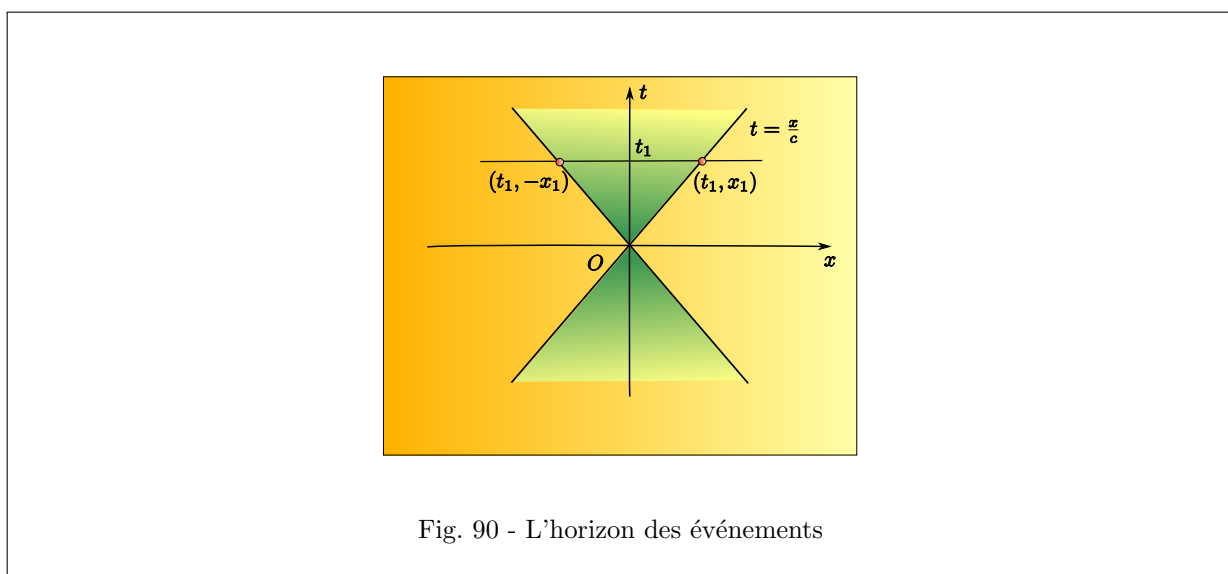
Tous les événements qui se trouvent à la surface du cône ne sont accessibles à partir de  $O$  qu'à la vitesse de la lumière. C'est pourquoi ces événements sont dits de **genre lumière**. Et le cône est par conséquent appelé **cône de lumière**.

Enfin, les événements qui se trouvent à l'extérieur du cône ne sont pas accessibles à partir de l'événement  $O$  puisqu'il faudrait se déplacer à une vitesse supérieure à celle de la lumière pour les rejoindre. Ils sont donc en un autre lieu de l'espace qui n'est pas visible du point  $O$ . On dit qu'ils sont du **genre espace**.

BEATRIX : Si je résume donc, le cône de lumière découpe l'espace-temps en trois zones. A l'intérieur du cône, les événements de genre temps qui sont accessibles ou visibles à partir de l'événement  $O$ . A la surface du cône, les événements de genre lumière, qui ne peuvent être accessibles depuis  $O$  que par un déplacement à la vitesse de la lumière. A l'extérieur du cône, les événements de genre espace, qui ne sont pas visibles à partir de l'événement  $O$ .

EURISTIDE : Et nous avons vu que l'intervalle entre deux événements ne dépend pas du référentiel galiléen dans lequel il est mesuré. Donc la notion de cône de lumière et le concept d'événement de genre temps, lumière et espace sont des propriétés absolues de l'espace-temps.

Cette notion est de grande importance en cosmologie. Supposons que le big-bang (création de l'univers) soit l'événement  $O$ . Tous les événements de l'univers sont évidemment issus de ce big bang. Par conséquent, tous les événements de l'univers se trouvent à l'intérieur du cône de lumière.



Considérons l'instant présent  $t_1$ , situé quelque part dans le futur de l'événement  $O$  qui est le big bang. Les événements se produisant à cet instant  $t_1$  sont dans le cône de lumière, et sur la droite parallèle à l'axe des  $x$  entre les points d'univers  $(t_1, -x_1)$  et  $(t_1, x_1)$ . Il n'existe pas d'événement au delà de cet horizon, car sinon ils auraient dû se propager à partir du big bang à une vitesse supérieure à celle de la lumière.

BEATRIX : Ce qui est impossible, nous l'avons vu !

EURISTIDE : Un observateur à un instant donné et en un lieu donné ne peut pas observer un événement plus lointain que ce que la lumière a pu parcourir depuis le big bang. Cette distance limite visible par un observateur porte le nom **d'horizon cosmologique**.

Plus généralement, deux événements ne peuvent être liés de cause à effet que si leur intervalle est du genre temps. Ceci est la conséquence immédiate de l'impossibilité que des interactions de toutes sortes (et pas seulement la lumière) se propagent plus vite que la vitesse de la lumière.

BEATRIX : Oui, c'est amusant d'y songer. Rien ne va plus vite que la lumière. Ni la pensée (la transmission de la pensée, si elle existait, n'irait pas plus vite que les signaux lumineux ou des signaux radio), ni la causalité (la transmission d'une cause à un effet.) Un extraterrestre situé à des millions d'années-lumière voit actuellement notre planète Terre comme elle était il y a plusieurs millions d'années, c'est-à-dire le temps que la lumière de notre planète ait parcouru cette énorme distance de millions d'années-lumière. Donc, notre extraterrestre voit actuellement une planète peuplée d'animaux sauvages ! La loi intergalactique bien connue qui interdit toute colonisation d'une planète peuplée d'êtres intelligents ne semble pas devoir s'appliquer...

## 60 Scène IX-3 : Le temps propre

EURISTIDE : Mais s'ils décidaient maintenant de coloniser la Terre, ils n'arriveraient en vue de notre planète que dans quelques millions d'années. Nous serions alors certainement très différents de ce que nous sommes aujourd'hui. Et eux aussi... Nous pouvons espérer que nous serions plus sages et considérerions de part et d'autre un programme de coopération, plutôt que des stratégies de colonisation ou de défense.

Mais revenons à cette propriété d'invariance des intervalles entre deux événements dans les référentiels galiléens. Cette propriété, nous allons le voir, conduit à des propriétés tout à fait surprenantes concernant l'écoulement du temps lui-même.

MATHINE : Considérons, en effet, un référentiel galiléen  $G$ , dans lequel une horloge se déplace suivant un mouvement rectiligne uniforme, à la vitesse  $v$ .

Considérons maintenant un référentiel  $G'$  attaché à l'horloge. L'horloge est donc immobile dans le référentiel  $G'$ , et  $G'$  est un référentiel galiléen animé d'une vitesse constante  $v$  par rapport à  $G$ .

Considérons deux instants  $t_1$  et  $t_2$  dans  $G$ . Dans le référentiel  $G$ , l'horloge parcourt la distance  $l_{12}$  telle que :

$$l_{12}^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2. \quad (380)$$

Dans le référentiel  $G'$ , l'horloge est immobile, donc :

$$l'_{12} = 0. \quad (381)$$

On sait que les intervalles des deux événements aux instants  $t_1$  et  $t_2$  dans  $G$  et dans  $G'$  sont égaux, donc :

$$c^2(t_2 - t_1)^2 - l_{12}^2 = c^2(t'_2 - t'_1)^2 - l'^2_{12}. \quad (382)$$

Ce qui s'écrit, puisque  $l'_{12} = 0$  :

$$c^2(t_2 - t_1)^2 - l_{12}^2 = c^2(t'_2 - t'_1)^2. \quad (383)$$

Or, l'horloge se déplace à la vitesse  $v$  constante, donc :

$$l_{12} = v(t_2 - t_1). \quad (384)$$

Ce qui peut encore s'écrire :

$$v^2(t_2 - t_1)^2 = l_{12}^2. \quad (385)$$

En insérant cette expression dans l'expression de l'égalité des intervalles ci-dessus, on obtient :

$$c^2(t'_2 - t'_1)^2 = c^2(t_2 - t_1)^2 - v^2(t_2 - t_1)^2, \quad (386)$$

ou encore :

$$\boxed{t'_2 - t'_1 = (t_2 - t_1) \sqrt{1 - \frac{v^2}{c^2}}}. \quad (387)$$

Le temps écoulé dans le référentiel  $G'$ , (c'est-à-dire  $t'_2 - t'_1$ ) est appelé **temps propre** de l'horloge.

**EURISTIDE** : Ce dernier résultat est tout à la fois fondamental et surprenant : puisque  $\sqrt{1 - \frac{v^2}{c^2}}$  est inférieur à 1, le temps s'écoule plus lentement pour un observateur lancé à une vitesse proche de celle de la lumière. Il s'écoule d'ailleurs d'autant plus lentement que sa vitesse s'approche de la vitesse de la lumière.

**BEATRIX** : Comment imaginer cela ? Le temps ne s'écoule-t-il pas toujours de la même façon ?

**EURISTIDE** : C'est une conséquence du fait que la vitesse de la lumière est constante et la même dans tous les référentiels galiléens. On peut le comprendre intuitivement, parce que même si le référentiel où je me trouve possède une vitesse propre proche de la vitesse de la lumière, nous savons que les rayons lumineux se propagent quand même à la vitesse de la lumière par rapport à moi, immobile dans mon référentiel lancé à grande vitesse. Comment expliquer qu'un rayon lumineux s'éloigne de moi à la vitesse de la lumière, alors que je suis moi-même lancé pratiquement à cette vitesse ? C'est donc que quelque chose a changé dans l'expression de cette vitesse, qui est un rapport de longueur à temps écoulé. Si le temps s'écoule plus lentement, alors la valeur de l'intervalle de temps au dénominateur est plus faible, donc la valeur numérique de la vitesse augmente. C'est donc une conséquence de cette universalité de la vitesse de la lumière qui veut que le temps s'écoule plus lentement lorsque je me déplace moi-même à une vitesse proche de cette vitesse de la lumière.

**BEATRIX** : Oui, je comprends mieux. Finalement, pour éviter le paradoxe de l'impossible addition des vitesses lorsqu'elles sont proches celle de la lumière, la Nature a trouvé l'astuce de la relativité du temps. La notion de vitesse est elle-même relative au temps et donc impactée par le ralentissement de l'écoulement du temps.

**MATHINE** : Traduisons cela en termes plus concrets. Supposons que la capitaine d'un vaisseau interstellaire quitte la Terre pour un voyage à grande vitesse, puis revienne sur la Terre. Pour fixer les idées, supposons qu'il se déplace à une vitesse  $v$ , moitié de celle de la lumière :

$$v = \frac{c}{2}. \quad (388)$$

Supposons qu'il effectue, du point de vue d'un observateur resté sur Terre, un voyage de 10 ans à l'aller, qu'il séjourne sur une planète distante pendant 10 ans, et qu'il revienne sur Terre après un voyage de retour de

10 ans. Nous supposons que sa vitesse durant le voyage aller est constante et son mouvement rectiligne, et que le référentiel  $G_1$  attaché à sa fusée est par conséquent galiléen. Nous supposons que le référentiel  $G_3$  attaché à sa fusée au retour est également galiléen. Nous supposons enfin que le référentiel  $G_2$  attaché à la planète sur laquelle il séjourne durant 10 ans se déplace à la même vitesse que le référentiel galiléen  $G$  de la Terre.

Supposons que le capitaine du vaisseau ait 30 ans à son départ de la Terre et qu'il ait un frère jumeau qui, lui, reste sur sa Terre natale pendant le voyage de son frère. Quel sera l'âge du capitaine après son retour ?

BEATRIX : Tiens, je n'imaginai pas qu'un jour, j'aurai à résoudre un véritable exercice me demandant l'âge du capitaine !

MATHINE : Et pourtant, la question vaut la peine d'être posée. Ecoute un peu. En utilisant la formule que nous avons déduite tout à l'heure :

$$t'_2 - t'_1 = (t_2 - t_1) \sqrt{1 - \frac{v^2}{c^2}}, \quad (389)$$

durant le voyage aller, le temps propre qui s'est écoulé pour le capitaine est :

$$\Delta t'_{Aller} = 10 \text{ ans} \times \sqrt{1 - \frac{v^2}{c^2}} = 10 \times \sqrt{\frac{3}{4}} = 5\sqrt{3} \text{ ans}. \quad (390)$$

Durant le séjour sur la planète lointaine, le temps propre écoulé pour le capitaine est :

$$\Delta t'_{Sejour} = 10 \text{ ans}. \quad (391)$$

Durant le voyage de retour, le temps propre écoulé pour le capitaine est, à l'instar de l'aller :

$$\Delta t'_{Retour} = 5\sqrt{3} \text{ ans}. \quad (392)$$

Par conséquent, lorsque le capitaine revient de son voyage, il a vieilli de :

$$\Delta t'_{Aller} + \Delta t'_{Sejour} + \Delta t'_{Retour} = 10(1 + \sqrt{3}) = 27,3 \text{ ans}. \quad (393)$$

Il a donc 57 ans et presque 4 mois à son retour sur Terre.

En revanche, son frère jumeau a déjà 60 ans, puisque durant ce voyage, il s'est écoulé 30 ans sur la Terre.

BEATRIX : Les deux frères ne sont donc plus vraiment jumeaux après cette aventure !

MATHINE : Nous avons jusque là raisonné avec une horloge entraînée dans un mouvement rectiligne uniforme. Considérons une horloge douée d'un mouvement quelconque, sa vitesse à chaque instant étant  $v(t)$ .

Dans une situation plus complexe comme celle-ci, où la vitesse varie dans le temps, nous avons toujours le même réflexe : analyser ce qui se passe lors de variations infiniment petites de la variable temps. Considérons un intervalle de temps infinitésimal  $dt$ . Durant ce temps, l'horloge mobile parcourt la distance  $ds$  :

$$ds = \sqrt{dx^2 + dy^2 + dz^2}. \quad (394)$$

A tout instant, le mouvement peut être considéré comme uniforme à l'échelle infinitésimale. Par conséquent, nous pouvons appliquer la notation que nous avons vue tout à l'heure, et le temps propre dans le référentiel attaché à l'horloge, considéré comme galiléen à tout instant, est tel que :

$$dt' = \sqrt{1 - \frac{v^2}{c^2}} dt. \quad (395)$$

En effectuant l'intégration de cette relation entre les instants  $t_1$  et  $t_2$ , on obtient l'expression de l'intervalle de temps propre :

$$t'(t_2) - t'(t_1) = t'_2 - t'_1 = \int_{t_1}^{t_2} \sqrt{1 - \frac{v^2}{c^2}} dt. \quad (396)$$

Nous avons noté ici  $t'_1$  et  $t'_2$  les valeurs du temps propre correspondant respectivement aux instants  $t_1$  et  $t_2$  dans le référentiel  $G$ .

**EURISTIDE** : On notera que pour  $v$  très petite par rapport à  $c$  (ce qui est notre lot quotidien), le temps propre est sensiblement voisin au temps écoulé dans le référentiel au repos, car dans ce cas :

$$1 - \frac{v^2}{c^2} \approx 1 \quad (397)$$

est sensiblement égal à 1, donc :

$$\sqrt{1 - \frac{v^2}{c^2}} \approx 1 \quad (398)$$

également.

**BEATRIX** : Dans la plus grande partie des expériences quotidiennes que nous vivons les vitesses de déplacement d'objets ou de nous-mêmes sont négligeables par rapport à  $c$ . C'est la raison pour laquelle les effets relativistes ne sont pas apparents dans notre quotidien.

**EURISTIDE** : C'est exact. Par exemple, si j'observe un objet lancé à la vitesse  $v$  supersonique de 3600 km/h (qui est déjà une vitesse très impressionnante pour l'homme), cela donne, après conversion des vitesses en km/s :

$$\begin{aligned} \sqrt{1 - \frac{v^2}{c^2}} &= \sqrt{1 - \frac{1}{(300000)^2}} \\ &\approx \sqrt{1 - \frac{1}{100000000000}}. \end{aligned} \quad (399)$$

Ce qui constitue une valeur extrêmement proche de 1, à un cent milliardième près.

L'effet relativiste est donc quasiment imperceptible dans ce cas de figure, alors que la vitesse à l'échelle humaine est déjà très grande.

## 61 Scène IX-4 : La transformation de Lorentz

**MATHINE** : Nous allons poursuivre l'analyse des conséquences du fait que la vitesse de la lumière est finie et constante dans tous les référentiels galiléens.

**EURISTIDE** : Oui, et nous allons voir que non seulement le temps propre se contracte à grande vitesse, mais aussi les longueurs.



BEATRIX : C'est normal (enfin, si on peut le dire!). On a vu tout à l'heure que lorsque les vitesses étaient proches de celle de la lumière, elles ne pouvaient pas s'ajouter directement ; elles doivent donc subir également une contraction. Et si la vitesse et le temps se contractent à grande vitesse, alors les distances doivent forcément se contracter également.

MATHINE : Oui, on peut considérer les choses comme cela. Regardons plus précisément ce qui se passe. Attention, nous allons enchaîner une série de calculs un peu compliqués, pour aboutir au résultat. Accroche-toi!

Considérons un référentiel  $G$  au repos et un référentiel  $G'$  doué d'une vitesse  $V$  uniforme le long de l'axe  $x$  de  $G$ . On supposera que les axes des deux référentiels  $G$  et  $G'$  sont parallèles entre eux : soit  $x$  parallèle à  $x'$ ,  $y$  parallèle à  $y'$  et  $z$  parallèle à  $z'$ .

Dans ces conditions, la conservation des intervalles dit que :

$$c^2t^2 - x^2 = c^2t'^2 - x'^2. \quad (400)$$

Par conséquent,  $x$  et  $ct$  peuvent être considérés comme des fonctions de  $x'$  et de  $ct'$ .

Ecrivons  $x$  sous la forme :

$$x = ax' + bct'. \quad (401)$$

où  $a$  et  $b$  sont des expressions que nous devons déterminer. Ecrivons de même  $ct$  sous la forme :

$$ct = dx' + ect', \quad (402)$$

où  $c$  et  $d$  sont des expressions à déterminer.

En utilisant l'expression de l'égalité des intervalles, nous pouvons écrire :

$$d^2x'^2 + 2dex'ct' + e^2c^2t'^2 - a^2x'^2 - b^2c^2t'^2 - 2abcx't' = c^2t'^2 - x'^2. \quad (403)$$

Soit :

$$(d^2 - a^2)x'^2 + (2de - 2ab)x'ct' + (e^2 - b^2)c^2t'^2 = c^2t'^2 - x'^2. \quad (404)$$

Nous pouvons maintenant identifier les termes en  $t'$  et  $x'$  des deux membres de l'égalité. Cela donne :

$$\begin{aligned} d^2 - a^2 &= -1 \\ 2de - 2ab &= 0 \\ e^2 - b^2 &= 1. \end{aligned} \quad (405)$$

D'où les trois équations régissant les valeurs de  $a$ ,  $b$ ,  $d$  et  $e$  :

$$\begin{aligned} ab &= de \\ a^2 - d^2 &= 1 \\ e^2 - b^2 &= 1. \end{aligned} \quad (406)$$

En mathématicienne avertie, quand je vois une expression telle que :

$$x^2 - y^2 = 1, \quad (407)$$

je pense tout de suite à ce qu'on appelle la trigonométrie hyperbolique. C'est un nom bien mystérieux pour désigner des fonctions cosinus hyperbolique (cosh) et sinus hyperbolique (sinh) qui ont pour propriété, justement, de vérifier :

$$\cosh^2 x - \sinh^2 x = 1. \quad (408)$$

Donc, nous allons effectuer un changement de variable et considérer une quantité  $\psi$  telle que :

$$a = \cosh \psi. \quad (409)$$

Alors, nous pouvons déduire des équations ci-dessus que :

$$d = \sinh \psi. \quad (410)$$

De la même façon, nous allons considérer une quantité  $\phi$  telle que :

$$e = \cosh \phi. \quad (411)$$

Alors, comme précédemment, les équations ci-dessus nous permettent de déduire que :

$$b = \sinh \phi. \quad (412)$$

Or, nous savons aussi que :

$$ab = de. \quad (413)$$

Donc, nous pouvons en déduire, en remplaçant  $a$ ,  $b$ ,  $d$  et  $e$  par leurs nouvelles expressions, que :

$$\cosh \psi \sinh \phi = \sinh \psi \cosh \phi. \quad (414)$$

D'où :

$$\tanh \psi = \frac{\sinh \psi}{\cosh \psi} = \tanh \phi = \frac{\sinh \phi}{\cosh \phi}, \quad (415)$$

où  $\tanh \psi$  est appelée tangente hyperbolique de  $\psi$ .

Donc,  $\psi$  et  $\phi$  ont même tangente hyperbolique. On démontre en effet (mais nous n'allons pas le faire ici) que si deux nombres ont même tangente hyperbolique, alors ils sont égaux. Nous admettrons donc que :

$$\psi = \phi. \quad (416)$$

Donc, les solutions des équations écrites plus haut sont :

$$\begin{aligned} e &= a \\ d &= b \\ a &= \cosh \psi \\ b &= \sinh \psi \end{aligned} \quad (417)$$

Par conséquent, nous pouvons donc écrire :

$$\begin{aligned} x &= x' \cosh \psi + ct' \sinh \psi \\ ct &= x' \sinh \psi + ct' \cosh \psi. \end{aligned} \quad (418)$$

Considérons maintenant dans le référentiel  $G$  le mouvement de l'origine des coordonnées de  $G'$ . Alors :

$$x' = 0, \quad (419)$$

donc :

$$\begin{aligned} x &= ct' \sinh \psi \\ ct &= ct' \cosh \psi. \end{aligned} \quad (420)$$

Par conséquent :

$$\frac{x}{ct} = \tanh \psi. \quad (421)$$

Or, nous savons que le référentiel  $G'$  a pour vitesse  $V$  par rapport à  $G$  :

$$\frac{x}{t} = V. \quad (422)$$

Donc :

$$\tanh \psi = \frac{V}{c}. \quad (423)$$

Or, on sait que :

$$\cosh^2 \psi - \sinh^2 \psi = 1, \quad (424)$$

donc, en divisant par  $\cosh^2 \psi$ , on obtient :

$$1 - \frac{\sinh^2 \psi}{\cosh^2 \psi} = \frac{1}{\cosh^2 \psi}, \quad (425)$$

d'où :

$$\cosh^2 \psi = \frac{1}{1 - \tanh^2 \psi}, \quad (426)$$

et :

$$\sinh^2 \psi = \cosh^2 \psi - 1 = \frac{\tanh^2 \psi}{1 - \tanh^2 \psi}. \quad (427)$$

Donc, finalement, en ne considérant que les valeurs positives qui ont un sens physique ici :

$$\begin{aligned} \sinh \psi &= \frac{\tanh \psi}{\sqrt{1 - \tanh^2 \psi}} = -\frac{\frac{V}{c}}{\sqrt{1 - \frac{V^2}{c^2}}} \\ \cosh \psi &= \frac{1}{\sqrt{1 - \tanh^2 \psi}} = \frac{1}{\sqrt{1 - \frac{V^2}{c^2}}}. \end{aligned} \quad (428)$$

Par conséquent, on obtient :

$$\begin{aligned} x &= x' \cosh \psi + ct' \sinh \psi \\ &= \frac{x'}{\sqrt{1 - \frac{V^2}{c^2}}} + \frac{ct' \frac{V}{c}}{\sqrt{1 - \frac{V^2}{c^2}}} \\ &= \frac{x'}{\sqrt{1 - \frac{V^2}{c^2}}} + \frac{Vt'}{\sqrt{1 - \frac{V^2}{c^2}}}, \end{aligned} \quad (429)$$

et :

$$\begin{aligned} t &= \frac{x'}{c} \sinh \psi + t' \cosh \psi \\ &= \frac{\frac{x'}{c} \frac{V}{c}}{\sqrt{1 - \frac{V^2}{c^2}}} + \frac{t'}{\sqrt{1 - \frac{V^2}{c^2}}} \\ &= \frac{\frac{V}{c^2} x'}{\sqrt{1 - \frac{V^2}{c^2}}} + \frac{t'}{\sqrt{1 - \frac{V^2}{c^2}}}, \end{aligned} \quad (430)$$

d'où finalement :

$$\begin{aligned}x &= \frac{x' + Vt'}{\sqrt{1 - \frac{V^2}{c^2}}} \\t &= \frac{t' + \frac{V}{c^2}x'}{\sqrt{1 - \frac{V^2}{c^2}}}.\end{aligned}\quad (431)$$

BEATRIX : Ca fait un peu tour de passe-passe, avec ces sinus et cosinus hyperboliques.

MATHINE : En réalité, on peut déduire la même chose sans utiliser ce que tu appelles un tour de passe-passe. Mais cela aurait été plus long et fastidieux. Cette technique de changement de variable (en l'occurrence remplacer  $a, b, d, e$  par  $\psi$  et  $\phi$ ) est une technique très courante et souvent très efficace en mathématiques. C'est pour cela que je l'ai choisie. J'admets que la trigonométrie hyperbolique, c'est un peu compliqué, mais c'est très efficace ici.

EURISTIDE : Les expressions que nous avons obtenues sont appelées **formules de transformation de Lorentz**, et expriment des lois de transformation des coordonnées dans le changement de référentiel galiléen.

Ces formules sont essentielles pour la suite. En les appliquant à différentes situations, on met en évidence un certain nombre de phénomènes de contraction relativiste.

BEATRIX : Allons-y, alors !

MATHINE : Voyons quelques exemples. Considérons une règle (pour mesurer) au repos dans le référentiel  $G$ , de longueur  $l = x_2 - x_1$ .

Considérons cette règle observée depuis le référentiel  $G'$  en mouvement à la vitesse  $V$  par rapport à  $G$ .

D'après les transformations de Lorentz, on a dans  $G'$  :

$$l' = x'_2 - x'_1, \quad (432)$$

avec :

$$\begin{aligned}x_1 &= \frac{x'_1 + Vt'}{\sqrt{1 - \frac{V^2}{c^2}}} \\x_2 &= \frac{x'_2 + Vt'}{\sqrt{1 - \frac{V^2}{c^2}}}.\end{aligned}\quad (433)$$

On en déduit immédiatement que :

$$x_2 - x_1 = \frac{x'_2 - x'_1}{\sqrt{1 - \frac{V^2}{c^2}}}. \quad (434)$$

Soit :

$$x'_2 - x'_1 = (x_2 - x_1)\sqrt{1 - \frac{V^2}{c^2}}. \quad (435)$$

Autrement dit, la règle est plus courte dans le référentiel en mouvement que dans le référentiel au repos.  
La valeur :

$$\sqrt{1 - \frac{V^2}{c^2}} \quad (436)$$

est appelée **contraction de Lorentz**, puisque c'est elle qui détermine la contraction des corps en mouvement.

BEATRIX : Donc, je résume. A grande vitesse, le temps et les longueurs sont contractés par la contraction de Lorentz. C'est assez logique : à partir du moment où, à grande vitesse, le temps s'écoule plus lentement, il faut bien que les longueurs soient contractées pour que la vitesse de la lumière demeure constante, puisqu'intuitivement, la distance  $L$  parcourue par un rayon lumineux en un intervalle de temps  $\Delta t$  est  $L = c\Delta t$ .  $\Delta t$  étant plus petit à grande vitesse, il faut bien que  $L$  soit contractée.

## 62 Scène IX-5 : La transformation de la vitesse

EURISTIDE : Nous avons étudié la transformation des coordonnées d'un référentiel galiléen à un autre, lorsque ce dernier est animé d'un mouvement rectiligne uniforme de vitesse  $V$  le long des axes  $x$  et  $x'$ . Nous allons maintenant regarder ce qu'il advient de la vitesse  $v$  d'un point dans le référentiel  $G$  si sa vitesse est  $v'$  dans le référentiel  $G'$ .

BEATRIX : Nous avons vu tout à l'heure que  $v$  n'est pas égal à la somme des vitesses  $v'$  et  $V$ , puisque la limite de la vitesse de la lumière l'interdit.

MATHINE : Supposons que  $v$  ait pour composantes  $(v_x, v_y, v_z)$  et  $v'$  pour composantes  $(v'_x, v'_y, v'_z)$ . Nous avons vu précédemment que :

$$\begin{aligned} dx &= \frac{dx' + V dt'}{\sqrt{1 - \frac{V^2}{c^2}}} \\ dy &= dy' \\ dz &= dz' \\ dt &= \frac{dt' + \frac{V}{c^2} dx'}{\sqrt{1 - \frac{V^2}{c^2}}}. \end{aligned} \quad (437)$$

Nous obtenons alors les composantes de la vitesse en divisant  $dx$ ,  $dy$  et  $dz$  par  $dt$  :

$$\begin{aligned} v_x &= \frac{dx}{dt} = \frac{dx' + V dt'}{dt' + \frac{V}{c^2} dx'} = \frac{v'_x + V}{1 + \frac{V}{c^2} v'_x} \\ v_y &= \frac{dy}{dt} = \frac{dy' \sqrt{1 - \frac{V^2}{c^2}}}{dt' + \frac{V}{c^2} dx'} = \frac{v'_y \sqrt{1 - \frac{V^2}{c^2}}}{1 + \frac{V}{c^2} v'_x} \\ v_z &= \frac{dz}{dt} = \frac{dz' \sqrt{1 - \frac{V^2}{c^2}}}{dt' + \frac{V}{c^2} dx'} = \frac{v'_z \sqrt{1 - \frac{V^2}{c^2}}}{1 + \frac{V}{c^2} v'_x}. \end{aligned} \quad (438)$$

BEATRIX : On a donc confirmation qu'en relativité restreinte, les vitesses ne s'ajoutent pas, contrairement à ce que nous avons appris en cinématique classique.

EURISTIDE : En revanche, nous voyons bien que lorsque  $V$ ,  $v'_x$ ,  $v'_y$ , et  $v'_z$  sont négligeables devant  $c$ , ce qui le cas usuel, les formules se ramènent aux formules classiques d'addition des vitesses, à savoir :

$$\begin{aligned}v_x &= v'_x + V \\v_y &= v'_y \\v_z &= v'_z.\end{aligned}\tag{439}$$

Cette transformation de la vitesse a pour conséquence un phénomène assez surprenant et observable de changement de direction de cette vitesse. Nous allons étudier le phénomène dit d'aberration de la lumière. En effet, si la lumière n'est pas émise dans une direction parallèle à celle du déplacement du référentiel  $G'$ , alors l'angle sous lequel on voit les rayons lumineux est modifié pour un observateur situé dans le référentiel  $G'$ . C'est cette déviation qu'on appelle **aberration de la lumière**.

MATHINE : Considérons donc toujours nos deux référentiels  $G$  et  $G'$ ,  $G'$  étant lancé à une vitesse  $V$  par rapport à  $G$  le long de leurs axes communs  $x$  et  $x'$ .

Considérons le rayon lumineux émis dans le plan  $x'y'$  dans le référentiel  $G'$  (donc  $v' = c$ ), sous l'angle  $\theta'$  par rapport à l'axe des  $x'$ .

Nous allons calculer l'angle  $\theta$  par rapport à l'axe des  $x$  sous lequel le rayon lumineux est vu dans le référentiel  $G$ .

Nous avons :

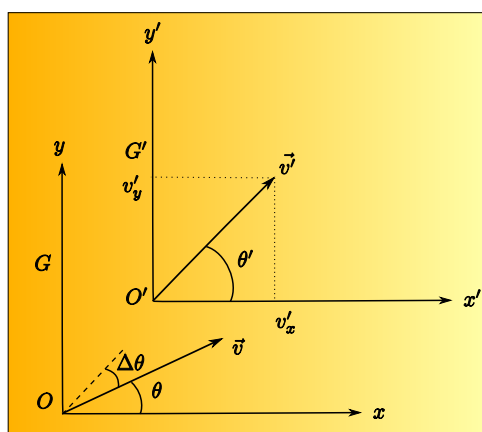


Fig. 91 - L'aberration de la lumière

$$\begin{aligned}v'_x &= c \cos \theta' \\v'_y &= c \sin \theta'.\end{aligned}\tag{440}$$

Et de même :

$$\begin{aligned}v_x &= c \cos \theta \\v_y &= c \sin \theta.\end{aligned}\tag{441}$$

Alors, les formules de transformation de la vitesse deviennent :

$$\begin{aligned}c \cos \theta &= \frac{c \cos \theta' + V}{1 + \frac{c \cos \theta' V}{c^2}} = c \frac{\cos \theta' + \frac{V}{c}}{1 + \frac{V}{c} \cos \theta'} \\c \sin \theta &= \frac{c \sin \theta' \sqrt{1 - \frac{V^2}{c^2}}}{1 + \frac{c \cos \theta' V}{c^2}} = c \frac{\sqrt{1 - \frac{V^2}{c^2}}}{1 + \frac{V}{c} \cos \theta'} \sin \theta'.\end{aligned}\tag{442}$$

Supposons maintenant que  $V$  soit petit par rapport à  $c$ . Nous allons pouvoir, dans des expressions bien choisies, négliger  $\frac{V}{c}$  par rapport à 1. Alors, dans un premier temps, en négligeant  $\frac{V^2}{c^2}$  par rapport à 1 :

$$\sin \theta \approx \frac{\sin \theta'}{1 + \frac{V}{c} \cos \theta'}.\tag{443}$$

D'où :

$$\begin{aligned}\sin \theta' - \sin \theta &= \sin \theta' \left( 1 - \frac{1}{1 + \frac{V}{c} \cos \theta'} \right) \\&= \sin \theta' \frac{\frac{V}{c} \cos \theta'}{1 + \frac{V}{c} \cos \theta'}.\end{aligned}\tag{444}$$

Dans cette expression, nous allons maintenant négliger l'expression  $\frac{V}{c} \cos \theta'$  qui est inférieure à  $\frac{V}{c}$  et donc négligeable devant 1. On obtient alors :

$$\sin \theta' - \sin \theta \approx \frac{V}{c} \sin \theta' \cos \theta'.\tag{445}$$

Or, nous savons, par une formule de trigonométrie bien connue des lycéens, que :

$$\sin \theta' - \sin \theta = 2 \cos \frac{\theta' + \theta}{2} \sin \frac{\theta' - \theta}{2}.\tag{446}$$

Puis, nous pouvons faire les approximations suivantes :

$$\cos \frac{\theta' + \theta}{2} \approx \cos \theta',\tag{447}$$

puisque l'aberration de la lumière est un phénomène faible, et donc  $\theta$  et  $\theta'$  sont très proches l'un de l'autre, donc  $\theta - \theta'$  est négligeable devant  $\theta'$  et donc  $\theta' + \theta = 2\theta' + (\theta - \theta') \approx 2\theta'$ . De plus :

$$\sin \frac{\theta' - \theta}{2} \approx \frac{1}{2}(\theta' - \theta),\tag{448}$$

puisque  $\theta' - \theta$  est voisin de 0 pour la même raison. D'où :

$$2 \cos \theta' \frac{1}{2}(\theta' - \theta) \approx \frac{V}{c} \sin \theta' \cos \theta',\tag{449}$$

soit :

$$\theta' - \theta \approx \frac{V}{c} \sin \theta'. \quad (450)$$

De plus, nous notons :

$$\theta' - \theta = \Delta\theta \quad (451)$$

que nous appellerons **angle d'aberration**. Alors :

$$\Delta\theta \approx \frac{V}{c} \sin \theta' \quad (452)$$

qui est la formule habituelle de l'aberration de la lumière.

**EURISTIDE** : Cette formule de l'aberration de la lumière vaut la peine d'être commentée. Tout d'abord, elle dépend de  $\frac{V}{c}$ , donc elle s'annule pratiquement, lorsque  $V$  est petite devant  $c$ . Il n'y a donc pas de phénomène d'aberration pour les vitesses usuelles qui sont très petites par rapport à  $c$ . Ensuite, elle dépend de  $\sin \theta'$ . Donc, elle s'annule lorsque  $\theta'$  est nul, c'est-à-dire qu'il n'y a pas d'effet d'aberration de la lumière lorsque le rayon est émis dans la direction du déplacement de la source lumineuse.

**BEATRIX** : De plus, on doit pouvoir détecter cet effet plus facilement que ceux des contractions de Lorentz, puisque l'angle d'aberration dépend ici du rapport  $\frac{V}{c}$ , et non pas  $\frac{V^2}{c^2}$ .

## 63 Scène IX-6 : La notation tensorielle

**MATHINE** : Nous avons vu dans nos discussions d'introduction la notation indicielle et les tenseurs. Il est temps maintenant d'utiliser ces outils.

Nous sommes en possession d'un espace de dimension 4, constitué du temps et des 3 dimensions habituelles de l'espace  $x$ ,  $y$ , et  $z$ .

Dans cet espace, nous considérons les coordonnées suivantes, exprimées en notation indicielle :

$$\begin{aligned} x^0 &= ct \\ x^1 &= x \\ x^2 &= y \\ x^3 &= z. \end{aligned} \quad (453)$$

**EURISTIDE** : Partant de ces coordonnées, il est naturel de considérer l'intervalle entre deux événements comme la norme d'un vecteur de l'espace-temps, muni de ce système de coordonnées, puisque l'intervalle est un invariant de l'espace-temps.

**MATHINE** : Soit  $A^i$  les composantes contravariantes d'un vecteur de l'espace-temps (donc  $i = 0, 1, 2, 3$ ). Alors l'intervalle entre l'origine de l'espace-temps et l'extrémité du vecteur  $A^i$  est :

$$c^2 t^2 - x^2 - y^2 - z^2 = (A^0)^2 - (A^1)^2 - (A^2)^2 - (A^3)^2. \quad (454)$$

Pour pouvoir utiliser la convention d'écriture des indices d'Einstein, nous avons besoin d'introduire les composantes covariantes du vecteur  $A^i$ . Nous avons vu que la norme d'un vecteur est exprimée par la



somme des produits des composantes covariantes par les composantes contravariantes. Nous voulons que les composantes covariantes  $A_i$  du vecteur de l'espace-temps soient telles que l'expression :

$$A^i A_i = A^0 A_0 + A^1 A_1 + A^2 A_2 + A^3 A_3 \quad (455)$$

représente la norme au carré, donc l'intervalle au carré entre l'origine de l'espace-temps et l'extrémité du vecteur  $A^i$ . Donc :

$$A^i A_i = A^0 A_0 + A^1 A_1 + A^2 A_2 + A^3 A_3 = (A^0)^2 - (A^1)^2 - (A^2)^2 - (A^3)^2. \quad (456)$$

Donc :

$$\begin{aligned} A_0 &= A^0 \\ A_1 &= -A^1 \\ A_2 &= -A^2 \\ A_3 &= -A^3. \end{aligned} \quad (457)$$

Et avec ces valeurs,  $A^i A_i$  est donc le carré de l'intervalle au sens relativiste.

**EURISTIDE** : Le signe de  $A^i A_i$  permet d'interpréter le genre de l'événement :

- Si  $A^i A_i > 0$ , l'événement est du genre temps ;
- Si  $A^i A_i = 0$ , l'événement est du genre lumière ;
- Si  $A^i A_i < 0$ , l'événement est du genre espace.

En effet : si  $A^i A_i > 0$ , cela signifie que  $c^2 \Delta t^2 > \Delta l^2$ , donc la distance physique entre les deux événements est accessible sans dépasser la vitesse de la lumière. Le point d'une extrémité est donc à l'intérieur du cône de lumière issu de l'autre extrémité.

Si  $A^i A_i = 0$ , cela signifie que  $c^2 \Delta t^2 = \Delta l^2$ , donc le parcours d'une extrémité à l'autre de la distance physique se fait à la vitesse de la lumière. Le point d'une extrémité est donc sur le cône de lumière issu de l'autre extrémité.

Si  $A^i A_i < 0$ , cela signifie que  $c^2 \Delta t^2 < \Delta l^2$ , donc les deux événements sont inaccessibles sans dépasser la vitesse de la lumière. Donc le point d'une extrémité est à l'extérieur du cône de lumière issu de l'autre extrémité.

**BEATRIX** : En fait, nous avons créé une norme un peu particulière dans l'espace-temps, puisqu'elle peut être négative, contrairement à la norme habituelle que nous connaissons dans un espace euclidien.

**EURISTIDE** : Oui, nous pouvons même dire que nous avons défini une métrique particulière dans l'espace-temps. Toute la particularité de cet espace-temps réside dans le signe "moins" appliqué aux composantes covariantes d'indices 1, 2 et 3. C'est ce qui donne à la métrique son comportement particulier, et qui aboutit aux transformations de Lorentz. Avec cette métrique, le produit scalaire de deux vecteurs  $A^i$  et  $B^i$  s'écrit :

$$A^i B_i = A^0 B^0 - A^1 B^1 - A^2 B^2 - A^3 B^3. \quad (458)$$

Or, nous avons vu que le produit scalaire d'un vecteur s'écrivait :

$$g_{ik} A^i B^k, \quad (459)$$

où  $g_{ik}$  est le tenseur euclidien fondamental, ou tenseur métrique de l'espace-temps. En prenant pour tenseur

métrique  $g_{ik}$  :

$$g_{ik} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad (460)$$

on obtient :

$$\begin{aligned} g_{ik}A^iB^k &= g_{00}A^0B^0 + g_{01}A^0B^1 + g_{02}A^0B^2 + g_{03}A^0B^3 \\ &+ g_{10}A^1B^0 + g_{11}A^1B^1 + g_{12}A^1B^2 + g_{13}A^1B^3 \\ &+ g_{20}A^2B^0 + g_{21}A^2B^1 + g_{22}A^2B^2 + g_{23}A^2B^3 \\ &+ g_{30}A^3B^0 + g_{31}A^3B^1 + g_{32}A^3B^2 + g_{33}A^3B^3. \end{aligned} \quad (461)$$

Or, tous les termes  $g_{ik}$  où  $i \neq k$  sont nuls, et  $g_{00} = 1$ ,  $g_{11} = -1$ ,  $g_{22} = -1$  et  $g_{33} = -1$ .

Donc :

$$g_{ik}A^iB^k = A^0B^0 - A^1B^1 - A^2B^2 - A^3B^3. \quad (462)$$

Et on retrouve donc bien :

$$A^iB_i = g_{ik}A^iB^k. \quad (463)$$

Et la norme d'un vecteur est :

$$A^iA_i = g_{ik}A^iA^k. \quad (464)$$

L'espace-temps de dimension 4 muni de cette métrique  $g_{ik}$  s'appelle **espace de Minkowski**, caractérisé par son tenseur euclidien fondamental  $g_{ik}$ .

Si nous considérons un vecteur de composantes contravariantes  $A^i$ , alors ses composantes covariantes sont :

$$A_i = g_{ik}A^k. \quad (465)$$

Considérons un point  $A^i$  représenté par son vecteur  $A^i$  issu de l'origine.

Le vecteur :

$$u^i = \frac{dA^i}{d\tau}, \quad (466)$$

où  $\tau$  est le temps propre du point  $A^i$ , est appelé **quadrivitesse** du point  $A^i$ .

Supposons que le point  $A^i$  ait pour vitesse  $\vec{v}(v_x, v_y, v_z)$  dans l'espace ordinaire de dimension 3, alors on a, comme nous l'avons vu :

$$d\tau = \sqrt{1 - \frac{v^2}{c^2}} dt. \quad (467)$$

Alors, la quadrivitesse a pour composantes :

$$\begin{aligned} u^0 &= \frac{cdt}{d\tau} = \frac{cdt}{\sqrt{c^2 - v^2} dt} = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \\ u^1 &= \frac{dx}{d\tau} = \frac{dx}{\sqrt{1 - \frac{v^2}{c^2}} dt} = \frac{v_x}{\sqrt{1 - \frac{v^2}{c^2}}} \\ u^2 &= \frac{dy}{d\tau} = \frac{dy}{\sqrt{1 - \frac{v^2}{c^2}} dt} = \frac{v_y}{\sqrt{1 - \frac{v^2}{c^2}}} \\ u^3 &= \frac{dz}{d\tau} = \frac{dz}{\sqrt{1 - \frac{v^2}{c^2}} dt} = \frac{v_z}{\sqrt{1 - \frac{v^2}{c^2}}}. \end{aligned} \quad (468)$$

BEATRIX : Et que représente donc la quadrivitesse ?

EURISTIDE : Il s'agit du vecteur vitesse relativiste exprimé dans l'espace de Minkowski. Cette vitesse relativiste comporte une première composante qui est la dérivée du temps par rapport au temps propre ; c'est donc la vitesse de variation du temps dans le référentiel au repos du point de vue d'un observateur attaché à la particule en mouvement. Et les trois autres composantes sont les composantes spatiales de la vitesse perçue par la particule dans son temps propre. Les  $(u^1, u^2, u^3)$  sont les composantes d'un vecteur dans l'espace de dimension 3 représentant la vitesse par rapport au temps propre, donc la vitesse perçue par un observateur attaché à la particule en mouvement. C'est donc une sorte de "vitesse propre".

## 64 Scène IX-7 : La mécanique relativiste

BEATRIX : Comme tout à l'heure, étant donné que nous avons traité de la cinématique relativiste, nous pouvons maintenant regarder la mécanique relativiste. Et je serais curieuse de connaître les lois de la mécanique dans cet espace de Minkowski. Elles doivent être différentes de celles d'un espace euclidien, non ?

EURISTIDE : Nous allons y venir, et nous allons utiliser pour cela le principe de moindre action que nous avons appris tout à l'heure. Rappelons-nous que tout système physique tend à se déplacer sur une trajectoire qui minimise ou maximise son action. C'est-à-dire que tout système tend à se déplacer sur une extrémale de l'action.

Considérons une horloge mobile dans un référentiel  $G$ . Considérons un référentiel  $G'$  attaché à l'horloge mobile.

Par invariance de l'intervalle, on a :

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 = c^2 dt'^2, \quad (469)$$

où  $dt$  est l'accroissement infinitésimal du temps dans  $G$  et  $dt'$  l'accroissement dans  $G'$ . Par conséquent :

$$dt' = \frac{ds}{c}. \quad (470)$$

Et donc, le temps  $S$  indiqué par l'horloge entre des événements  $a$  et  $b$ , quel que le soit le référentiel, est l'intégrale du temps :

$$S = \frac{1}{c} \int_a^b ds, \quad (471)$$

où l'intégration est effectuée sur la ligne d'univers entre les événements  $a$  et  $b$ .

Si l'horloge est au repos, sa ligne d'univers est une droite parallèle à l'axe des temps entre les événements  $a$  et  $b$ .

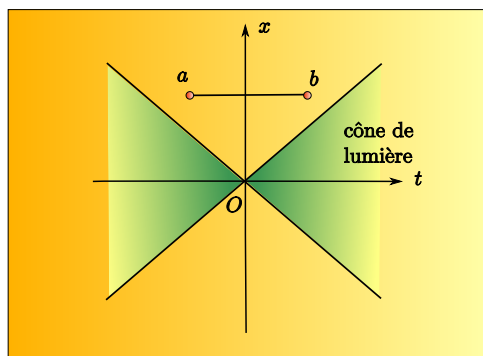


Fig. 92 - Ligne d'univers de l'horloge au repos

Si l'horloge décrit un mouvement non uniforme, alors sa ligne d'univers est une courbe passant par les événements  $a$  et  $b$ . Nous avons vu dans l'étude de la transformation de Lorentz du temps, que dans ce cas, le temps indiqué par l'horloge sera plus court que celui d'une horloge au repos.

Nous en déduisons que l'intégrale :

$$\int_a^b ds \quad (472)$$

atteint sa valeur maximale lorsqu'elle est calculée le long de la droite d'univers entre les deux événements  $a$  et  $b$ .

**BEATRIX** : Cette intégrale est donc proportionnelle à l'action du système, puisqu'elle se comporte exactement de la même façon caractéristique. En effet, nous avons vu en étudiant le principe variationnel tout à l'heure que la caractéristique de l'action, c'était précisément d'être minimale pour les trajectoires minimales, puisque l'univers est un fainéant.

**EURISTIDE** : C'est vrai, et l'action peut être considérée comme la consommation d'énergie du système physique demandée pour le déplacement le long de la trajectoire choisie.

**MATHINE** : Nous avons donc trouvé une fonction qui peut représenter l'action du système physique. Donc l'action  $S$  doit s'écrire au moyen d'une constante  $\alpha$  :

$$S = -\alpha \int_a^b ds. \quad (473)$$

Nous avons placé un signe moins, puisque l'action atteint sa valeur minimale lorsque l'intégrale atteint sa valeur maximale.

Or nous savons que :

$$ds^2 = c^2 dt^2 - (dx^2 + dy^2 + dz^2). \quad (474)$$

D'où :

$$ds^2 = c^2 dt^2 \left( 1 - \frac{v^2}{c^2} \right). \quad (475)$$

Soit, encore :

$$ds = c\sqrt{1 - \frac{v^2}{c^2}} dt. \quad (476)$$

Par conséquent, l'action  $S$  peut s'écrire maintenant :

$$S = - \int_a^b \alpha c \sqrt{1 - \frac{v^2}{c^2}} dt = \int_a^b L dt, \quad (477)$$

où  $L$  est la fonction de Lagrange du système mécanique, comme nous l'avons vu lors de la discussion sur le principe variationnel.

Or nous avons vu qu'en mécanique classique, la fonction de Lagrange s'écrit :

$$L = \frac{1}{2}mv^2. \quad (478)$$

Donc, en observant le comportement de l'expression relativiste de  $L$  lorsque  $c$  tend vers l'infini, nous pouvons, par identification avec l'expression non relativiste, en déduire la valeur de la constante  $\alpha$  :

$$L = -\alpha c \sqrt{1 - \frac{v^2}{c^2}} \approx -\alpha c + \frac{\alpha v^2}{2c} = \frac{1}{2}mv^2. \quad (479)$$

Or, les fonctions de Lagrange peuvent être définies à une constante près, puisque l'action elle-même peut être définie à une constante près. Donc par identification des termes non constants, nous obtenons :

$$\alpha = mc. \quad (480)$$

Donc, l'action d'une particule libre, représentant donc par définition, l'énergie dépensée dans son déplacement, s'écrit finalement :

$$S = \int_a^b L dt, \quad (481)$$

où :

$$L = -mc^2 \sqrt{1 - \frac{v^2}{c^2}}. \quad (482)$$

**EURISTIDE** : Cette expression de l'action nous montre qu'elle est nulle pour  $v = c$ . Comme  $L$  est négative, c'est donc qu'elle atteint sa valeur maximale pour la trajectoire d'un rayon lumineux.

**MATHINE** : On a vu que l'impulsion d'une particule s'exprime comme la dérivée par rapport à la vitesse de la fonction de Lagrange. Donc :

$$\vec{p} = \frac{\partial L}{\partial \vec{v}}. \quad (483)$$

Si nous développons l'expression en utilisant l'expression de  $L$  que nous avons déduite, nous avons, lorsque la masse est constante :

$$\vec{p} = \frac{m \vec{v}}{\sqrt{1 - \frac{v^2}{c^2}}}. \quad (484)$$

**BEATRIX** : Et nous retrouvons comme d'habitude deux résultats : pour  $v$  négligeable devant  $c$ , l'impulsion tend vers son expression classique :

$$\vec{p} = m \vec{v}. \quad (485)$$

Et pour  $v$  tendant vers  $c$ , l'impulsion tend vers l'infini, donc elle n'a pas de sens physique pour un rayon lumineux. Ce qui veut dire qu'un rayon lumineux ne se comporte pas comme une particule libre, au sens de la relativité restreinte.

MATHINE : Continuons notre périple dans la mécanique relativiste. La force exercée sur la particule est la dérivée par rapport au temps de l'impulsion. Donc :

$$\vec{F} = \frac{d\vec{p}}{dt}. \quad (486)$$

Passons à l'énergie maintenant. On appelle énergie d'une particule la somme de son énergie cinétique et de son énergie potentielle. Par conséquent :

$$E = T + U. \quad (487)$$

Or, nous avons vu, lors de l'étude du principe variationnel, que  $L = T - U$ , et nous savons que, par définition,  $T = \frac{1}{2} \vec{p} \cdot \vec{v}$ . Donc, on peut écrire :

$$E + L = 2T, \quad (488)$$

d'où :

$$E = 2T - L, \quad (489)$$

soit :

$$E = \vec{p} \cdot \vec{v} - L. \quad (490)$$

Par conséquent, pour une particule de masse constante :

$$\begin{aligned} E &= \frac{mv^2}{\sqrt{1 - \frac{v^2}{c^2}}} + mc^2 \sqrt{1 - \frac{v^2}{c^2}} \\ &= \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \left( mv^2 + mc^2 \left( 1 - \frac{v^2}{c^2} \right) \right) \\ &= \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}}. \end{aligned} \quad (491)$$

EURISTIDE : Bien. Considérons maintenant le cas particulier d'une particule au repos, dont la vitesse  $v$  est nulle. Alors :

$$\boxed{E = mc^2}. \quad (492)$$

BEATRIX : Tiens ! On retrouve la célèbre expression d'Einstein sur l'énergie !

EURISTIDE : Oui,  $E = mc^2$  est l'expression de l'énergie intrinsèque d'une particule de masse  $m$ .

MATHINE : Exprimons maintenant ceci dans le formalisme indiciel.

Nous savons que la quadrivitesse est :

$$u^i = \left( \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}, \frac{\vec{v}}{\sqrt{1 - \frac{v^2}{c^2}}} \right). \quad (493)$$

On appelle quadri-impulsion l'expression :

$$p^i = mcu^i. \quad (494)$$

Par conséquent :

$$\begin{aligned} p^i &= \left( \frac{mc}{\sqrt{1 - \frac{v^2}{c^2}}}, \frac{mc \vec{v}}{\sqrt{1 - \frac{v^2}{c^2}}} \right) \\ &= \left( \frac{E}{c}, \frac{mc \vec{v}}{\sqrt{1 - \frac{v^2}{c^2}}} \right) \\ &= \left( \frac{E}{c}, \vec{p} \right). \end{aligned} \quad (495)$$

EURISTIDE : Une petite remarque sur la méthode que nous avons utilisée et que nous emploierons de nouveau à plusieurs reprises : nous avons déduit la loi que nous recherchions en effectuant une analyse infinitésimale, puis une intégration. Nous avons ainsi trouvé une loi qui comportait des constantes inconnues. L'analyse de cette équation dans le cas où  $v$  est très petit devant  $c$ , et la comparaison aux lois de la mécanique classique ont permis de déterminer les constantes manquantes. Nous aurons de nouveau l'occasion de voir cette méthode dans les développements qui suivent.

---

## Onzième partie

# Acte X - Les variétés

### 65 Scène X-1 : La variété différentiable

EURISTIDE : Maintenant que nous avons vu la relativité restreinte, nous allons entrer d'un premier pas dans le domaine passionnant de la relativité générale. Pour aborder cette théorie sereinement, nous procéderons en trois étapes. Nous allons commencer par introduire une nouvelle notion mathématique : celle de variété différentiable. Puis, en utilisant cet outil, nous pourrions décrire de façon précise ce qu'est la courbure d'un espace. Enfin, nous décrirons à l'aide de ces méthodes mathématiques, les équations du champ de gravitation d'Einstein.

BEATRIX : Alors, commençons par le début... Qu'est-ce qu'une variété différentiable ?

EURISTIDE : Après avoir développé la relativité restreinte, Einstein s'est mis à la recherche d'une théorie de la gravitation. Son extraordinaire intuition l'a conduit à remplacer l'espace-temps de Minkowski par un espace-temps curviligne, où le système de coordonnées n'est plus un ensemble de droites, mais un ensemble de courbes. Einstein a imaginé un univers où la courbure de l'espace-temps était créée en réaction à l'énergie et à l'impulsion ; autrement dit, les corps massifs eux-mêmes provoquaient la courbure de l'espace-temps ! C'est cette théorie que nous allons développer maintenant, et c'est pour cela que nous allons avoir besoin d'outils qui permettent de décrire des espaces curvilignes.

BEATRIX : Donc je suppose que les variétés différentiables dont nous avons parlé vont permettre cette description ?

EURISTIDE : Oui, les variétés différentiables seront notre outil de base pour décrire l'espace-temps curviligne. Nous verrons que la relativité générale est finalement un concept très simple : les corps massifs, et plus généralement l'énergie et l'impulsion de ces corps ou d'autres phénomènes physiques provoquent une déformation de l'espace-temps. Cette déformation de l'espace-temps conduit à la déformation des trajectoires des particules libres qui étaient jusque là rectilignes et uniformes. La relativité générale, ce n'est que ça : décrire la trajectoire déformée des particules suite à la courbure de l'espace-temps par l'énergie et l'impulsion.

Mais laissons d'abord Mathine nous expliquer ce qu'est une variété différentiable.

MATHINE : Une variété différentiable est un des concepts les plus importants de la géométrie pour les mathématiques et la physique.

Nous allons considérer que nous sommes dans un espace de dimension  $n$ . Un exemple bien connu d'espace euclidien de dimension  $n$  est l'espace  $\mathbb{R}^n$ , constitué des éléments de la forme  $(x_1, x_2, \dots, x_n)$  dont les  $n$  composantes sont des nombres réels.

Nous savons tous qu'il n'est pas possible de poser à plat sur une table une pelure d'orange sans la déchirer. Cela résulte du fait que la peau d'une orange est une surface curviligne. Et cette propriété fait qu'elle ne



peut pas être représentée, en général et dans son intégralité par un espace plat tel que  $\mathbb{R}^n$ .

Une variété différentiable permet de rendre compte de l'idée qu'un espace qui est curviligne, donc non représentable dans sa globalité par l'espace plat  $\mathbb{R}^n$ , est constitué d'un assemblage de régions locales en nombre infini et de tailles infiniment petites, qui sont, elles, représentables par  $\mathbb{R}^n$ .

Cela veut dire qu'une variété différentiable est un espace constitué d'un assemblage de régions assimilables à un espace plat tel que  $\mathbb{R}^n$  et ces différentes régions sont cousues ensemble pour constituer l'espace complet.

BEATRIX : Pouvez-vous donner des exemples mathématiques de variétés différentiables qui ne soient pas des espaces plats ?

MATHINE : La surface d'une sphère (notée  $S^2$  en géométrie) est une variété différentiable de dimension 2. La circonférence d'un cercle (qui est notée  $S^1$  en géométrie) est une variété différentiable de dimension 1. Plus généralement, une "sphère" de dimension  $n$ , qui est notée  $S^n$  est également une variété différentiable, de dimension  $n$ .

EURISTIDE : Maintenant que nous avons vu intuitivement ce qu'est une variété différentiable, nous allons formaliser cette notion pour mieux l'appréhender et pour pouvoir l'utiliser dans nos développements mathématiques et physiques. N'est-ce pas, Mathine ?

MATHINE : Oui, bien sûr. Mais pour commencer, nous allons introduire quelques notions utiles à ce qui va suivre.

On appelle **carte** entre deux ensembles  $M$  et  $N$  une relation  $\phi$ , qui à chaque élément de  $M$  associe exactement un élément de  $N$ .

BEATRIX : C'est donc une fonction de  $M$  dans  $N$  :

$$\phi : M \rightarrow N. \quad (496)$$

MATHINE : Oui, c'est bien une fonction de  $M$  dans  $N$ .

EURISTIDE : On appelle carte cette fonction dans un contexte géométrique, parce que c'est bien une fonction de ce type qu'on utilise lorsqu'on possède la carte géographique d'un territoire : l'ensemble  $M$  est la surface de la région concernée sur la Terre ; l'ensemble  $N$  est la carte géographique. La fonction  $\phi$  associe à un point de la surface de la Terre dans la région considérée, un point unique sur la carte géographique. C'est ainsi que nous pouvons nous repérer : en regardant un point de la carte géographique, nous trouvons le point correspondant sur la Terre.

MATHINE : Définissons maintenant un ensemble ouvert : on appelle **balle ouverte** un ensemble de points  $x$  de  $\mathbb{R}^n$ , tels que :

$$|x - y| < r, \quad (497)$$

pour un point  $y$  fixe de  $\mathbb{R}^n$  et pour une constante  $r$  réelle, et où l'on note la distance de  $x$  à  $y$  par :

$$|x - y| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}. \quad (498)$$

Une balle ouverte est donc l'ensemble des points de  $\mathbb{R}^n$  situés à une distance strictement inférieure à  $r$  d'un point  $y$  fixé.

**BEATRIX** : Je suppose que l'intérieur d'un disque de centre  $y$  et de rayon  $r$ , à condition de ne pas prendre la circonférence, est une balle ouverte dans  $\mathbb{R}^2$  ?

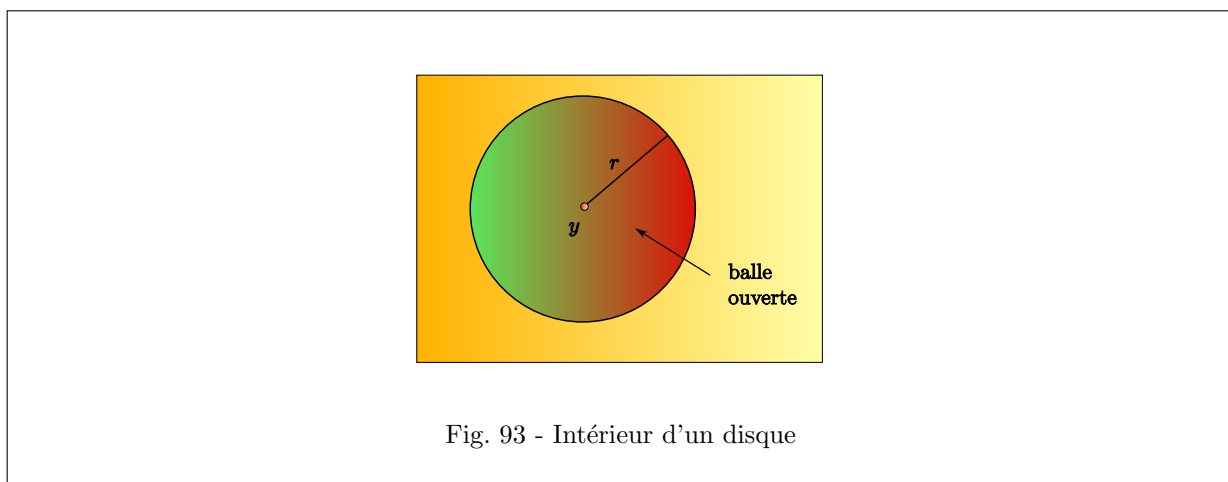


Fig. 93 - Intérieur d'un disque

**EURISTIDE** : Oui, et de même l'intérieur (duquel on exclut la surface) de la sphère  $S^2$  est une balle ouverte.

**MATHINE** : Maintenant que nous savons ce qu'est une balle ouverte, nous pouvons poursuivre. On appelle **ensemble ouvert** de  $\mathbb{R}^n$ , la réunion d'un nombre quelconque, éventuellement infini, de balles ouvertes.

**EURISTIDE** : Concrètement, un ensemble ouvert est en quelque sorte l'intérieur d'une surface close de dimension  $n - 1$ , ou la réunion de plusieurs tels intérieurs.

**BEATRIX** : Par exemple, en dimension 2, c'est l'intérieur d'une courbe close quelconque :

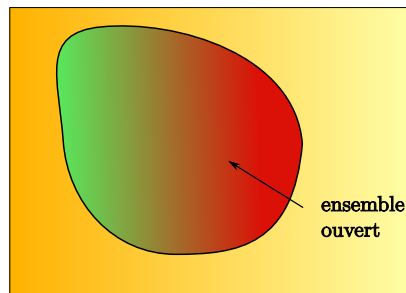


Fig. 94 - Intérieur d'une courbe close

En dimension 3, c'est l'intérieur d'une surface close quelconque, etc.

MATHINE : Comme je vois que c'est bien clair, nous passons à la suite. Attention, c'est un peu plus compliqué. On appelle **système de coordonnées**, un sous-ensemble  $U$  d'un ensemble  $M$  et une **carte injective** :

$$\phi : U \rightarrow \mathbb{R}^n, \quad (499)$$

tels que l'image  $\phi(U)$  soit un ensemble ouvert dans  $\mathbb{R}^n$ .

BEATRIX : Qu'est-ce qu'une carte injective ?

MATHINE : On dit qu'une carte est injective, s'il n'existe pas deux éléments de  $U$  qui aient la même image par  $\phi$ . Autrement dit, avec une carte injective, la situation suivante :

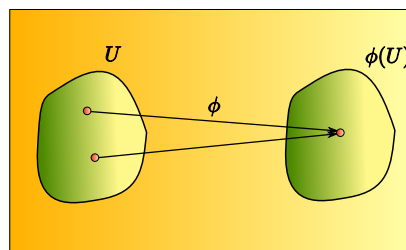


Fig. 95 - Carte non injective

n'est pas possible.

En d'autres termes, cela revient à dire qu'à tout point de  $U$ , on fait correspondre un point différent dans  $\phi(U)$ .  $\phi(U)$  est donc une sorte de représentation fidèle de  $U$  dans  $\mathbb{R}^n$ . D'où sa capacité à jouer le rôle d'un système de coordonnées : on représente le sous-ensemble  $U$  par une copie ou une carte injective se trouvant dans  $\mathbb{R}^n$ .

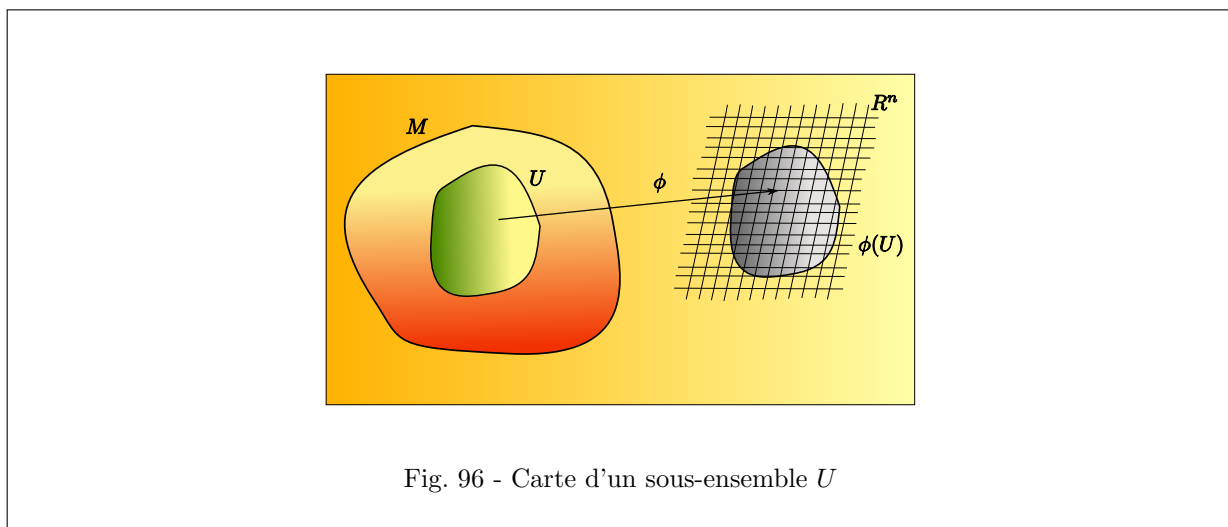


Fig. 96 - Carte d'un sous-ensemble  $U$

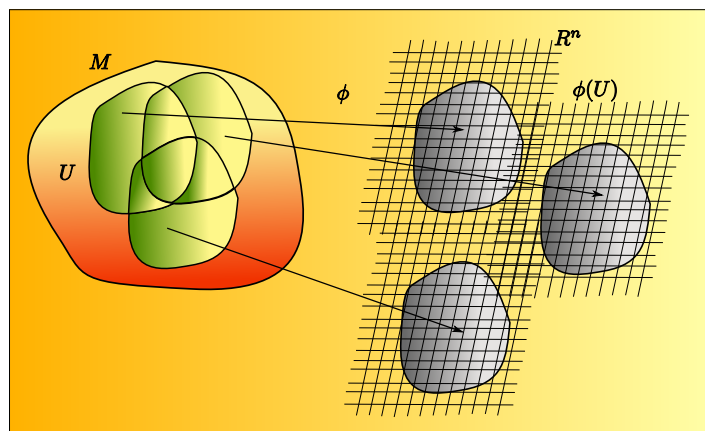
BEATRIX : Oui, en fait,  $\phi(U)$  est une sorte de carte géographique à  $n$  dimensions de la région  $U$ .

MATHINE : C'est exactement cela. Définissons maintenant ce qu'est l'assemblage de diverses cartes ou systèmes de coordonnées de cette nature.

Considérons un ensemble  $M$ , et une collection de sous-ensembles de  $M$  que nous noterons  $U_\alpha$ . On appelle **atlas**, une collection de systèmes de coordonnées  $\{(U_\alpha, \phi_\alpha)\}$  qui satisfait aux deux conditions suivantes :

- La réunion de tous les  $U_\alpha$  est égale à  $M$  ;
- Les systèmes de coordonnées sont cousus ensemble de façon continue.

Donc, l'atlas d'un ensemble  $M$  est un ensemble de systèmes de coordonnées qui recouvrent l'ensemble  $M$  et qui sont en relation continue sur leurs recouvrements.

Fig. 97 - Atlas d'un ensemble  $M$ 

**BEATRIX** : Le nom d'atlas est bien choisi, puisque c'est exactement ce qui se passe dans un atlas géographique. Chaque carte constitue une représentation fidèle de la région du monde qu'elle décrit. La réunion de toutes les cartes représente l'ensemble des régions du monde, avec d'éventuels recouvrements.

**MATHINE** : Et enfin, une **variété différentiable** est un ensemble  $M$  muni d'un atlas dit **maximal**, qui contient tout système de coordonnées compatible pour que deux ensembles équivalents munis d'atlas différents ne comptent pas comme deux variétés différentes.

**EURISTIDE** : Nous avons vu par conséquent, qu'une variété différentiable est un ensemble recouvert (ou représenté) par plusieurs systèmes de coordonnées.

**BEATRIX** : Pourquoi a-t-on plusieurs systèmes de coordonnées ?

**EURISTIDE** : Parce qu'en général, il n'est pas toujours possible de représenter ou de recouvrir un espace curviligne par un seul système de coordonnées. Prenons la sphère. Nous connaissons la projection de Mercator qui permet de représenter sur une feuille de papier la surface de la Terre (que nous considérerons approximativement sphérique ici.)

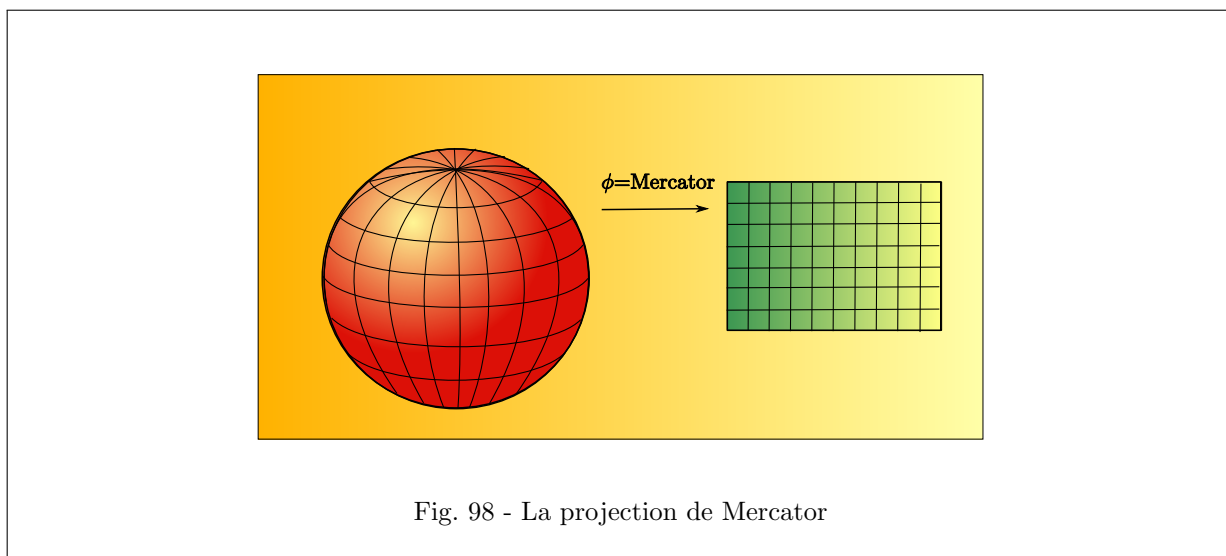


Fig. 98 - La projection de Mercator

Dans cette projection, trois éléments ne sont pas représentés : le pôle Nord, le pôle Sud et la ligne de changement de date (parce que le système de coordonnées doit être ouvert, ne l'oublions pas!)  
Donc, il faut au moins deux systèmes de coordonnées pour représenter une sphère intégralement.

**BEATRIX** : On voit bien qu'on a trouvé un outil pour représenter un espace curviligne, tel qu'une sphère, au moyen de plusieurs projections dans un espace euclidien comme  $\mathbb{R}^n$ . Mais comment allons-nous utiliser cela concrètement ?

**MATHINE** : Nous allons utiliser une notion mathématique extrêmement puissante, appelée le **plongement**. Comme nous avons vu qu'il y a une correspondante exacte (on dit une bijection) entre chaque région de la variété et une région de l'espace euclidien  $\mathbb{R}^n$ , nous allons pouvoir assimiler les points de la variété, localement dans chaque région, à un point dans l'espace euclidien  $\mathbb{R}^n$ .

Nous allons faire mieux encore. Nous allons pouvoir considérer que les propriétés que nous avons vues dans un espace euclidien (calcul différentiel, vecteurs, intégrales, cinématique, etc.) sont applicables à notre variété. Nous disons que nous **plongeons la variété localement dans un espace euclidien de dimension  $n$** .

De cette façon, nous allons pouvoir parler de la dérivée partielle d'une fonction de points dans la variété par rapport aux coordonnées dans  $\mathbb{R}^n$ , et écrire par exemple une expression telle que :

$$\frac{\partial f}{\partial x^\mu} \quad (500)$$

Cette expression représente la dérivée partielle par rapport à la composante  $x^\mu$  d'une fonction de la variété  $M$ . Si la fonction  $f$  représente une courbe sur une surface curviligne, alors  $\partial f / \partial x^\mu$  représente la variation de cette courbe dans la direction de cette coordonnée  $x^\mu$ .

**EURISTIDE** : Regardons concrètement ce que cela signifie pour une sphère. Considérons la projection de Mercator. Nous avons alors une représentation plane de régions de la sphère. Nous introduisons donc,

avec la notation des indices haut, deux coordonnées  $x^1$  pour les abscisses (correspondant aux parallèles de la sphère) et  $x^2$  pour les ordonnées (correspondant aux méridiens de la sphère) :

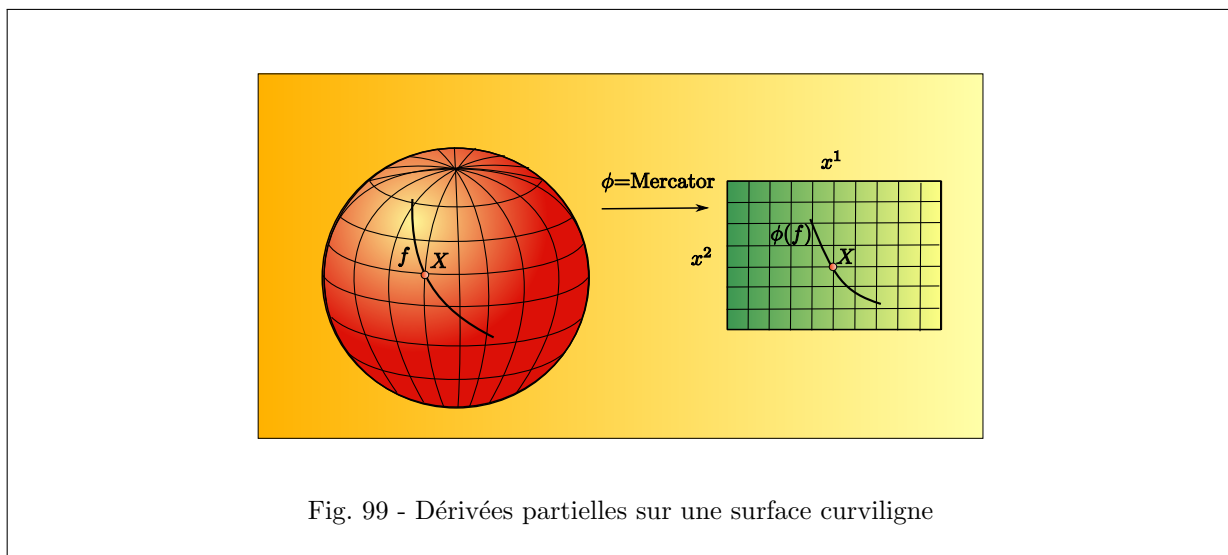


Fig. 99 - Dérivées partielles sur une surface curviligne

$\frac{\partial f}{\partial x^1}$  représente la pente de la courbe  $f$  dans la direction des parallèles à la surface de la sphère.  
 $\frac{\partial f}{\partial x^2}$  représente la pente de la courbe  $f$  dans la direction des méridiens à la surface de la sphère.

**BEATRIX** : On voit bien ici ce qui se passe. Le système de coordonnées, qui est ici représenté par une grille, peut être localement "décalqué" sur la surface curviligne et les axes  $x^1$  et  $x^2$  peuvent ainsi constituer les directions du système de coordonnées sur la surface curviligne. Evidemment, ce "décalage" est local et permet simplement de définir les deux directions  $x^1$  et  $x^2$  des coordonnées au voisinage du point  $X$  où le système de coordonnées a été ainsi "appliqué".

Il est clair que les directions  $x^1$  et  $x^2$  seront différentes en un point  $Y$  situé ailleurs sur la surface curviligne.

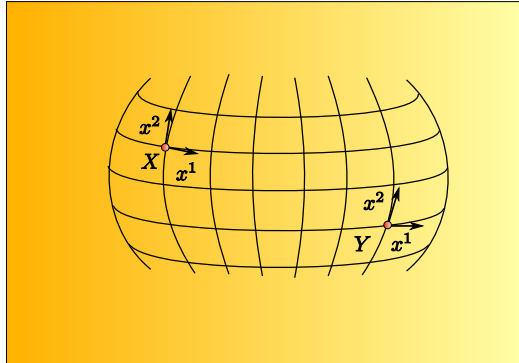


Fig. 100 - Les vecteurs de coordonnées varient en fonction du point

## 66 Scène X-2 : Notion d'espace tangent

MATHINE : Ce que tu viens de dire nous conduit tout naturellement à la notion d'espace tangent à une variété.

EURISTIDE : Intuitivement, il est assez facile d'imaginer ce qu'est l'espace tangent à une variété en un point en considérant l'exemple de la sphère.



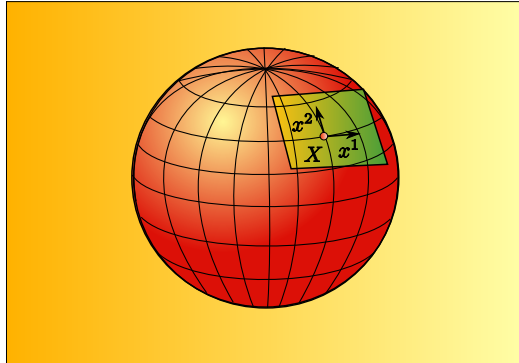


Fig. 101 - Plan tangent d'une sphère en un point

Comme nous l'avons vu, le système de coordonnées permet de définir en un point  $X$  de la sphère une direction  $x^1$  et une direction  $x^2$ . En considérant ces deux directions comme un système de coordonnées d'un plan passant par le point  $X$ , on obtient le plan tangent à la sphère. Nous venons de créer le **plan tangent à une variété**.

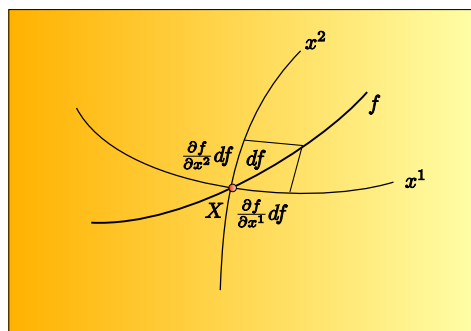
MATHINE : En restant dans le cas de la sphère, et en considérant une courbure  $f$  sur cette sphère, nous avons vu que l'expression :

$$\frac{\partial f}{\partial x^1} \quad (501)$$

représentait la pente de la courbe dans la direction  $x^1$  plaquée sur la surface de la sphère. Nous pouvons donc considérer que localement en un point  $X$  de la courbe,

$$\frac{\partial f}{\partial x^1} dx^1 \quad (502)$$

est la projection sur l'axe  $x^1$

Fig. 102 - Analyse des variations de  $f$  en un point

d'un accroissement infiniment petit de la courbe  $f$ , soit  $df$ .

Les mathématiciens, qui aiment bien effectuer des assimilations, vont donc pouvoir assimiler l'opérateur :

$$\frac{\partial}{\partial x^1} \quad (503)$$

à un vecteur représentant l'axe  $x^1$  dans le système de coordonnées local en  $X$  sur la variété, et de même :

$$\frac{\partial}{\partial x^2} \quad (504)$$

au vecteur représentant l'axe  $x^2$  dans le système de coordonnées local en  $X$ .

Par conséquent, nous venons de créer en tout point  $X$  d'une variété à 2 dimensions, un système de coordonnées :

$$\left( \begin{array}{c} \overrightarrow{\frac{\partial}{\partial x^1}}, \overrightarrow{\frac{\partial}{\partial x^2}} \end{array} \right) \quad (505)$$

qui définit naturellement l'espace tangent en  $X$  à la variété.

Et avec ce système de coordonnées, nous pouvons écrire :

$$\overrightarrow{df} = f dx^1 \overrightarrow{\frac{\partial}{\partial x^1}} + f dx^2 \overrightarrow{\frac{\partial}{\partial x^2}}. \quad (506)$$

De la même façon, pour une variété de dimension  $n$ , nous définissons l'espace tangent au moyen d'un système de coordonnées :

$$\left( \begin{array}{c} \overrightarrow{\frac{\partial}{\partial x^1}}, \overrightarrow{\frac{\partial}{\partial x^2}}, \dots, \overrightarrow{\frac{\partial}{\partial x^n}} \end{array} \right). \quad (507)$$

Pour simplifier les notations, nous notons :

$$\overrightarrow{\partial}_\mu = \overrightarrow{\frac{\partial}{\partial x^\mu}} \quad (508)$$

et le système de coordonnées de l'espace tangent par le symbole :

$$\left( \overrightarrow{\partial}_\mu \right). \quad (509)$$

EURISTIDE : Plus intuitivement, l'espace tangent en un point  $X$  est l'ensemble de tous les vecteurs tangents à toutes les courbes passant par le point  $X$ . Tous ces vecteurs sont une combinaison des vecteurs de base  $\vec{\partial}_\mu$ .

BEATRIX : En synthèse, au voisinage de chaque point  $X$  d'un espace curviligne, on peut définir, pour un système de coordonnées  $(x^1, x^2, \dots, x^n)$ , l'espace tangent par son système de coordonnées naturel :

$$\left( \frac{\vec{\partial}}{\partial x^1}, \frac{\vec{\partial}}{\partial x^2}, \dots, \frac{\vec{\partial}}{\partial x^n} \right) = \left( \vec{\partial}_\mu \right). \quad (510)$$

Un vecteur  $\vec{t}$  de cet espace tangent s'exprimera donc sous la forme :

$$\vec{t} = t^\mu \vec{\partial}_\mu. \quad (511)$$

Et bien entendu, contrairement à un système de coordonnées cartésien, les  $\vec{\partial}_\mu$  ne sont pas constants et sont des fonctions-vecteurs qui dépendent du point  $X$  de l'espace curviligne.

## 67 Scène X-3 : Notion d'espace cotangent

EURISTIDE : Nous savons définir maintenant en chaque point de la variété un espace plat tangent, avec un système de coordonnées cartésien. Il suffit de choisir l'atlas  $\{(x^1, x^2, \dots, x^n)\}$  pour que les systèmes de coordonnées qui lui sont associés soient cartésiens, et par conséquent, nous aurons un système de coordonnées local cartésien sur l'espace tangent en tout point de la variété. Nous aurons donc la possibilité d'appliquer localement au voisinage de chaque point de la variété, des propriétés d'un espace plat muni d'un système de coordonnées cartésien.

L'espace tangent à une variété en un point est une chose. Nous aurons besoin également de l'espace cotangent.

BEATRIX : L'espace cotangent ? Qu'est-ce ?

EURISTIDE : On a vu que l'espace tangent en un point  $X$  de la variété était l'ensemble de tous les vecteurs tangents à toutes les courbes de la variété passant par  $X$ . Il possède un système de coordonnées naturel, constitué des fonctions-vecteurs suivantes :

$$\left( \frac{\vec{\partial}}{\partial x^1}, \frac{\vec{\partial}}{\partial x^2}, \dots, \frac{\vec{\partial}}{\partial x^n} \right). \quad (512)$$

Ce qu'on appelle **espace cotangent** en un point  $X$  de la variété, est l'ensemble de toutes les différentielles  $df$  de toutes les courbes  $f$  passant par  $X$ .

MATHINE : Or on sait que, par définition d'une différentielle totale :

$$df = \frac{\partial f}{\partial x^i} dx^i. \quad (513)$$

Donc, un système de coordonnées naturel de cet espace cotangent est constitué des différentielles des fonctions coordonnées du système de coordonnées :

$$(dx^1, dx^2, \dots, dx^n). \quad (514)$$

EURISTIDE : On voit ici, dans cette somme :

$$df = \frac{\partial f}{\partial x^i} dx^i, \quad (515)$$

que dans l'expression de la différentielle d'une courbe  $f$ , pour l'espace tangent, les  $dx^i$  sont les composantes dans le système de coordonnées  $\left(\frac{\partial}{\partial x^i}\right)$ .

Inversement, pour l'espace cotangent, les  $\frac{\partial f}{\partial x^i}$  sont les composantes dans le système de coordonnées  $(dx^i)$ . Cette particularité fait qu'on considère l'espace cotangent comme le **dual** de l'espace tangent, et c'est cette symétrie qui fait qu'on l'a appelé précisément espace cotangent.

## 68 Scène X-4 : Le tenseur métrique

MATHINE : Nous avons vu dans l'espace euclidien le tenseur métrique, défini par :

$$ds^2 = g_{ij} dx^i dx^j. \quad (516)$$

Pour rappel, le rôle de ce tenseur métrique était de définir la façon dont on multiplie entre elles les coordonnées de deux fonctions-vecteurs pour former leur produit scalaire.

Dans une variété, nous pouvons également définir le tenseur métrique  $g_{ij}$  en tout point de la variété par la même expression ; celle-ci prend tout son sens, puisque  $dx^i$  et  $dx^j$  sont des fonctions-vecteurs de base de l'espace cotangent, comme nous venons de le voir. C'est donc que  $g_{ij}$  est bien un tenseur deux fois covariant, puisqu'il transforme deux fonctions-vecteurs en un nombre.

D'après sa définition, en manipulant les noms des indices :

$$ds^2 = g_{ij} dx^i dx^j = g_{ji} dx^j dx^i = g_{ji} dx^i dx^j, \quad (517)$$

ce tenseur est symétrique, c'est-à-dire que :

$$g_{ij} = g_{ji}. \quad (518)$$

On peut aussi définir la métrique inverse, qui est un tenseur deux fois contravariant, défini par :

$$g^{ij} g_{jk} = \delta_k^i. \quad (519)$$

EURISTIDE : Le tenseur métrique possède la propriété de coder à lui seul dans ses composantes toutes les caractéristiques de la variété : par exemple, sa courbure, la définition des longueurs, etc. En particulier, la variété est un espace plat lorsque ce tenseur possède des composantes constantes. Prenons l'exemple de la surface de la sphère. Considérons le système de coordonnées sphériques dans l'espace de dimension 3 :

$$\begin{aligned} x &= r \sin \theta \cos \phi \\ y &= r \sin \theta \sin \phi \\ z &= r \cos \theta. \end{aligned} \quad (520)$$

L'élément de ligne s'écrit :

$$ds^2 = dx^2 + dy^2 + dz^2. \quad (521)$$

Calculons la différentielle de cette expression, par rapport aux trois coordonnées  $r$ ,  $\theta$  et  $\phi$  :

$$\begin{aligned} ds^2 &= d(r \sin \theta \cos \phi)^2 + d(r \sin \theta \sin \phi)^2 + d(r \cos \theta)^2 \\ &= (dr \sin \theta \cos \phi + r \cos \theta \cos \phi d\theta - r \sin \theta \sin \phi d\phi)^2 \\ &+ (dr \sin \theta \sin \phi + r \cos \theta \sin \phi d\theta + r \sin \theta \cos \phi d\phi)^2 \\ &+ (dr \cos \theta - r \sin \theta d\theta)^2 \\ &= dr^2 \sin^2 \theta \cos^2 \phi + r^2 \cos^2 \theta \cos^2 \phi d\theta^2 + r^2 \sin^2 \theta \sin^2 \phi d\phi^2 \\ &+ 2rdr \cos \theta \sin \theta \cos^2 \phi d\theta - 2rdr \sin^2 \theta \cos \phi \sin \phi d\phi \\ &- 2r^2 \cos \theta \sin \theta \cos \phi \sin \phi d\theta d\phi \\ &+ dr^2 \sin^2 \theta \sin^2 \phi + r^2 \cos^2 \theta \sin^2 \phi d\theta^2 + r^2 \sin^2 \theta \cos^2 \phi d\phi^2 \\ &+ 2rdr \cos \theta \sin \theta \sin^2 \phi d\theta + 2rdr \sin^2 \theta \sin \phi \cos \phi d\phi \\ &+ 2r^2 \cos \theta \sin \theta \cos \phi \sin \phi d\theta d\phi \\ &+ dr^2 \cos^2 \theta + r^2 \sin^2 \theta d\theta^2 - 2rdr \cos \theta \sin \theta d\theta. \end{aligned} \quad (522)$$

Or :

$$\cos^2 \phi + \sin^2 \phi = 1, \quad (523)$$

et :

$$\cos^2 \theta + \sin^2 \theta = 1, \quad (524)$$

donc :

$$\begin{aligned} ds^2 &= dr^2 \sin^2 \theta \cos^2 \phi + dr^2 \sin^2 \theta \sin^2 \phi \\ &+ r^2 \cos^2 \theta \cos^2 \phi d\theta^2 + r^2 \cos^2 \theta \sin^2 \phi d\theta^2 \\ &+ r^2 \sin^2 \theta \sin^2 \phi d\phi^2 + r^2 \sin^2 \theta \cos^2 \phi d\phi^2 \\ &+ 2rdr \cos \theta \sin \theta \cos^2 \phi d\theta + 2rdr \cos \theta \sin \theta \sin^2 \phi d\theta \\ &+ 2rdr \sin^2 \theta \sin \phi \cos \phi d\phi - 2rdr \sin^2 \theta \cos \phi \sin \phi d\phi \\ &- 2rdr \cos \theta \sin \theta d\theta \\ &+ 2r^2 \cos \theta \sin \theta \cos \phi \sin \phi d\theta d\phi \\ &- 2r^2 \cos \theta \sin \theta \cos \phi \sin \phi d\theta d\phi \\ &+ dr^2 \cos^2 \theta + r^2 \sin^2 \theta d\theta^2 \\ &= dr^2 \sin^2 \theta + r^2 \cos^2 \theta d\theta^2 + r^2 \sin^2 \theta d\phi^2 \\ &+ dr^2 \cos^2 \theta \\ &+ r^2 \sin^2 \theta d\theta^2. \end{aligned} \quad (525)$$

D'où :

$$ds^2 = dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2. \quad (526)$$

Considérons le cas d'une sphère de rayon  $r = 1$  par exemple, alors  $dr = 0$ , puisque  $r$  est constant, et par conséquent :

$$ds^2 = d\theta^2 + \sin^2 \theta d\phi^2. \quad (527)$$

Donc, le tenseur métrique dans l'espace de dimension 2 de coordonnées  $(\theta, \phi)$ , pour la surface curviligne d'une sphère de rayon  $r = 1$  est :

$$g_{ij} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2 \theta \end{pmatrix}. \quad (528)$$

On vérifie bien que cet espace de dimension 2 n'est pas plat, puisque les composantes du tenseur métrique ne sont pas des constantes, mais dépendent de l'angle  $\theta$ .

BEATRIX : D'accord, je comprends. Dans un espace plat tel qu'un plan (à deux dimensions), on aurait :

$$ds^2 = dx^2 + dy^2, \quad (529)$$

donc :

$$g_{ij} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (530)$$

Dans ce cas, l'espace considéré est bien plat, puisque les composantes de son tenseur métrique sont des constantes.

EURISTIDE : Il est amusant de considérer la surface de la sphère de deux points de vue. Si on se considère comme observateur dans un espace de dimension 3, la surface de la sphère est un objet plongé dans un espace de dimension 3 décrit par le tenseur métrique :

$$g_{ik} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (531)$$

puisque :

$$ds^2 = dx^2 + dy^2 + dz^2. \quad (532)$$

Dans ce contexte, la surface de la sphère est un ensemble de points dans un espace plat à 3 dimensions. Si on considère maintenant la sphère intrinsèquement, c'est-à-dire du point de vue d'un être à deux dimensions infiniment petit plaqué à la surface de cette sphère, la métrique de cet espace à 2 dimensions, pour cet être infiniment plat qui y habite est :

$$g_{ik} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2 \theta \end{pmatrix}. \quad (533)$$

Il s'agit donc d'un espace curviligne.

BEATRIX : Je comprends. Suivant qu'on regarde la sphère de l'extérieur ou depuis sa surface, la surface de la sphère peut être vécue comme un objet plongé dans un espace plat, ou bien comme un espace curviligne.

MATHINE : Dans les deux cas, la sphère est une surface curviligne. Mais cette courbure est "vécue" de façon différente, suivant que nous sommes observateurs externes ou internes.

EURISTIDE : Il est difficile de se représenter ce qu'est un espace curviligne de dimension 3. Mais on peut extrapoler en 3 dimensions ce qui se passe en 2 dimensions. Si nous sommes un être imaginaire plongé dans un espace plat de dimension 4, nous pouvons observer de l'extérieur notre espace de dimension 3 et nous pouvons "voir" qu'il est "courbe". Si nous sommes un être de dimension 3, vivant **dans cet espace curviligne de dimension 3**, nous pouvons "sentir" qu'il est curviligne, parce que son tenseur métrique ne possède pas des composantes constantes.

BEATRIX : Et en quoi la variation des composantes du tenseur métrique va-t-elle nous affecter, nous les êtres vivants plongés dans cet espace curviligne de dimension 3 ?

EURISTIDE : Et bien, nous allons constater que l'expression des longueurs, par exemple, va dépendre de notre position dans l'espace. Nous aurons donc des lois physiques relatives aux longueurs, aux vitesses qui pourront dépendre de la position à laquelle nous nous trouvons dans l'espace. C'est le même phénomène qu'à la surface d'une sphère, où le calcul des longueurs va dépendre de la valeur de l'angle  $\theta$ , puisque le tenseur métrique dépend de la valeur de  $\sin^2 \theta$ .

MATHINE : Nous avons extrapolé ce qui se passe en 2 dimensions pour un espace curviligne à 3 dimensions. Nous pouvons continuer avec 4 dimensions. Dans le cadre de la relativité générale, Einstein a découvert que l'espace-temps (qui est un espace à 4 dimensions, comme nous l'avons vu) est un espace curviligne. C'est-à-dire que le calcul des longueurs dépend de la position (et du temps) dans cet espace-temps. La métrique de l'espace-temps en relativité générale comporte des composantes qui ne sont pas constantes. C'est ce qui donne la caractéristique curviligne à l'espace-temps de la relativité générale :

$$ds^2 = g_{ik} dx^i dx^k, \quad \text{où } i, k = 1, 2, 3, 4. \quad (534)$$

EURISTIDE : Nous avons vu en relativité restreinte que :

$$ds^2 = c dt^2 - dx^2 - dy^2 - dz^2. \quad (535)$$

Par conséquent, pour la relativité restreinte, le tenseur métrique est exprimé comme suit :

$$g_{ik} = \begin{pmatrix} c & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (536)$$

Donc la théorie de la relativité restreinte décrit un espace-temps plat, pour lequel le tenseur métrique possède des composantes constantes. Cet espace-temps n'est pas pour autant euclidien, puisque les composantes de  $g_{ik}$  ne sont pas toutes positives.

On distingue trois types de métriques, décrivant trois types d'espaces ou de variétés :

- La métrique riemannienne, possédant des composantes toutes positives ou nulles, qui décrit un espace dit riemanien ;
- La métrique lorentzienne, possédant des composantes toutes du même signe ou nulles, à l'exception d'une seule qui est du signe opposé. Cette métrique décrit un espace dit lorentzien ;
- Les autres métriques, qui décrivent un espace indéfini.

Dans le cadre qui nous intéresse, le monde physique a été successivement décrit dans l'histoire des sciences par trois types de métriques successives :

- Newton a décrit un espace-temps euclidien, c'est-à-dire un espace-temps riemanien, dont la métrique possède des composantes constantes :

$$ds^2 = dx^2 + dy^2 + dz^2. \quad (537)$$

- Einstein, dans la théorie de la relativité restreinte, a décrit un espace-temps lorentzien, dont la métrique lorentzienne possède des composantes constantes :

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2. \quad (538)$$

Cette métrique décrit un espace-temps plat, mais qui toutefois conduit aux phénomènes un peu particuliers de la contraction du temps et des longueurs que nous avons étudiés tout à l'heure. L'existence et la nature de ces phénomènes sont codés dans cette métrique.

- Enfin, dans la théorie de la relativité générale, Einstein décrit un espace-temps lorentzien plus général, dont les composantes de la métrique ne sont plus obligatoirement des constantes :

$$ds^2 = g_{ik} dx^i dx^k. \quad (539)$$

Cet espace-temps n'est donc pas plat, mais il possède intrinsèquement les propriétés d'un espace lorentzien, intégrant les propriétés de la relativité restreinte. Le fait qu'il ne soit pas plat donne à l'espace et au temps physiques des propriétés supplémentaires que nous étudierons tout à l'heure.

**BEATRIX** : C'est là qu'on voit que la métrique détermine tout. C'est elle qui construit notre monde. Chacune des composantes, variable ou constante, du tenseur métrique va déterminer la nature et les propriétés du monde dans lequel nous vivons. C'est merveilleux de penser que le monde physique, aussi complexe soit-il, peut être entièrement décrit au moyen d'un tenseur qui possède 16 composantes :

$$g_{ik} = \begin{pmatrix} g_{00} & g_{01} & g_{02} & g_{03} \\ g_{10} & g_{11} & g_{12} & g_{13} \\ g_{20} & g_{21} & g_{22} & g_{23} \\ g_{30} & g_{31} & g_{32} & g_{33} \end{pmatrix}. \quad (540)$$

**MATHINE** : C'est vrai, le tenseur métrique possède 16 composantes a priori. Mais elles ne sont pas toutes indépendantes parce que le tenseur métrique n'est pas totalement arbitraire : il doit répondre à un certain nombre d'exigences.

Notamment, nous savons que le tenseur métrique doit être symétrique, donc :

$$g_{ik} = g_{ki}. \quad (541)$$

Ensuite, comme nous voulons que l'espace soit lorentzien, cela signifie que nous voulons que le tenseur métrique puisse s'exprimer, localement en tout point de la variété, dans un système de coordonnées bien choisi, sous la forme d'une matrice diagonale, où tous les termes  $g_{ik}$  tels que  $i \neq k$  soient nuls et où, parmi tous les termes diagonaux, il y en a un seul qui n'est pas du même signe que les autres. Par exemple :

$$g_{ik} = \begin{pmatrix} h_{00} & 0 & 0 & 0 \\ 0 & -h_{11} & 0 & 0 \\ 0 & 0 & -h_{22} & 0 \\ 0 & 0 & 0 & -h_{33} \end{pmatrix}, \quad (542)$$

où les  $h_{jj}$  sont des fonctions positives.

**BEATRIX** : Donc, en synthèse, en choisissant correctement le système de coordonnées localement en chaque point de l'espace-temps, on peut obtenir un tenseur métrique qui possède une matrice de composantes diagonale.

**MATHINE** : Nous n'allons pas en faire la démonstration ici, mais on démontre que l'on peut faire encore mieux. On peut toujours trouver, localement en chaque point d'une variété curviligne, un système de coordonnées tel que chaque composante de la forme lorentzienne de la métrique ait des dérivées de premier



ordre par rapport aux coordonnées qui soit nulle. Il faut noter que dans ce cas là, les dérivées de second ordre ne s'annuleront pas toutes, a priori. Nous pourrions donc considérer que nous définissons en tout point de la variété un système de coordonnées dans lequel la variété ressemblera à un espace plat sur une région infiniment petite. Le système de coordonnées trouvé s'appelle **système de coordonnées normales riemanniennes**, et on dit que l'espace plat généré par ce système de coordonnées ressemble à la variété en ce point au premier ordre.

EURISTIDE : C'est un peu compliqué a priori, mais cela se comprend très bien en prenant l'exemple de notre surface sphérique favorite :

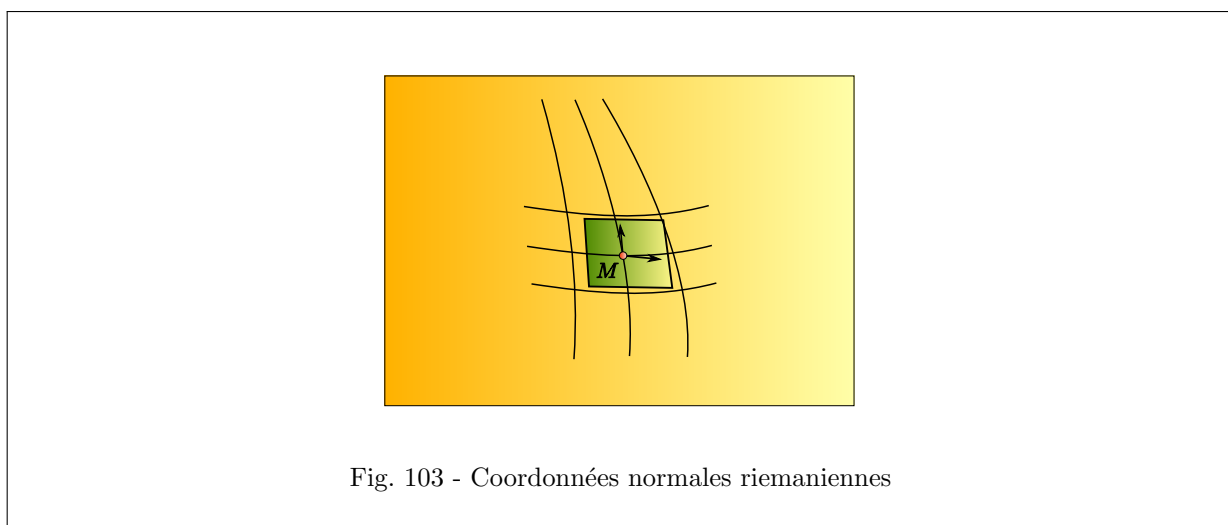


Fig. 103 - Coordonnées normales riemanniennes

En utilisant le système de coordonnées habituel associé à la projection de Mercator (méridiens, parallèles), nous pouvons considérer une région suffisamment petite autour d'un point  $M$  de la sphère. Sur cette région, nous pouvons considérer deux vecteurs qui sont respectivement tangents au méridien et au parallèle. Les vecteurs déterminent un espace localement plat, que l'on peut considérer comme ressemblant "au premier ordre" à la surface de la sphère au point  $M$ .

Nous connaissons l'expression générale du tenseur métrique de la surface de la sphère :

$$g_{ik} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2 \theta \end{pmatrix}. \quad (543)$$

Dire que l'espace est localement plat au point  $M$  de coordonnées  $(\theta_0, \phi_0)$ , revient à considérer que dans une petite région autour d'un point  $M$ , pour lequel  $\theta = \theta_0$ , nous considérons que  $\sin^2 \theta$  est constant égal à la valeur  $\sin^2 \theta_0$ . Si nous notons la constante :

$$\alpha_0 = \sin^2 \theta_0, \quad (544)$$

alors, nous obtenons une métrique d'un espace plat :

$$g_{ik} = \begin{pmatrix} 1 & 0 \\ 0 & \alpha_0 \end{pmatrix}, \quad (545)$$

où  $\alpha_0$  est une constante.

BEATRIX : En fait, intuitivement, cela revient à prendre un système de coordonnées sur le plan tangent à la surface de la sphère au point  $M$ , non ?

EURISTIDE : Oui. En résumé, nous avons vu que le monde physique est décrit par un tenseur métrique dont les composantes ne sont a priori pas constantes. Ce tenseur métrique peut être ramené localement en chaque point de l'espace-temps à un tenseur dont les composantes non diagonales sont toutes nulles, et dont les composantes diagonales sont des constantes au premier ordre, en choisissant de l'exprimer dans le système de coordonnées normales riemanniennes.

MATHINE : Nous allons maintenant regarder une grandeur dite **symbole de Levi-Civita**, qui nous a déjà été utile et qui le sera encore beaucoup par la suite.

Nous avons vu il y a quelque temps le symbole antisymétrique  $\epsilon^{ijk}$  défini, pour un espace de dimension 3, par :

$$\epsilon^{ijk} = \begin{cases} +1 & \text{si } ijk \text{ est une permutation paire de } 012 \\ -1 & \text{si } ijk \text{ est une permutation impaire de } 012 \\ 0 & \text{sinon.} \end{cases} \quad (546)$$

Nous pouvons généraliser ce symbole avec  $n$  indices, en définissant le symbole de Levi-Civita par :

$$\tilde{\epsilon}_{\mu_1\mu_2\dots\mu_n} = \begin{cases} +1 & \text{si } \mu_1\mu_2\dots\mu_n \text{ est une permutation paire de } 0, 1, \dots, n-1 \\ -1 & \text{si } \mu_1\mu_2\dots\mu_n \text{ est une permutation impaire de } 0, 1, \dots, n-1 \\ 0 & \text{sinon.} \end{cases} \quad (547)$$

Considérons deux systèmes de coordonnées dans une variété à  $n$  dimensions, et considérons une matrice  $n \times n$  notée :

$$M_{\mu'}^{\mu}. \quad (548)$$

Considérons le déterminant  $|M|$  de cette matrice. Comme nous l'avons vu lors de nos explications sur le calcul tensoriel, en utilisant le symbole antisymétrique, nous pouvons calculer le déterminant d'une matrice  $3 \times 3$ . En généralisant, avec le symbole de Levi-Civita, nous pouvons calculer le déterminant d'une matrice  $n \times n$  :

$$|M| = \tilde{\epsilon}_{\mu_1\mu_2\dots\mu_n} M_1^{\mu_1} M_2^{\mu_2} \dots M_n^{\mu_n}. \quad (549)$$

Par conséquent, en multipliant les deux membres par l'expression du symbole de Levi-Civita avec des indices primes :

$$\begin{aligned} \tilde{\epsilon}_{\mu'_1\mu'_2\dots\mu'_n} |M| &= \tilde{\epsilon}_{\mu'_1\mu'_2\dots\mu'_n} \tilde{\epsilon}_{\mu_1\mu_2\dots\mu_n} M_1^{\mu_1} M_2^{\mu_2} \dots M_n^{\mu_n} \\ &= \tilde{\epsilon}_{\mu_1\mu_2\dots\mu_n} \tilde{\epsilon}_{\mu'_1\mu'_2\dots\mu'_n} M_1^{\mu_1} M_2^{\mu_2} \dots M_n^{\mu_n} \\ &= \tilde{\epsilon}_{\mu_1\mu_2\dots\mu_n} M_{\mu'_1}^{\mu_1} M_{\mu'_2}^{\mu_2} \dots M_{\mu'_n}^{\mu_n}. \end{aligned} \quad (550)$$

Supposons maintenant que la matrice  $M_{\mu'}^{\mu}$  soit en fait définie par :

$$M_{\mu'}^{\mu} = \frac{\partial x^{\mu}}{\partial x^{\mu'}}. \quad (551)$$

Alors, la transformation ci-dessus s'écrit :

$$\tilde{\epsilon}_{\mu'_1\mu'_2\dots\mu'_n} = \left| \frac{\partial x^{\mu'}}{\partial x^{\mu}} \right| \tilde{\epsilon}_{\mu_1\mu_2\dots\mu_n} \frac{\partial x^{\mu_1}}{\partial x^{\mu'_1}} \frac{\partial x^{\mu_2}}{\partial x^{\mu'_2}} \dots \frac{\partial x^{\mu_n}}{\partial x^{\mu'_n}}. \quad (552)$$

Cette loi ressemble à une loi de transformation tensorielle dans un changement de coordonnées. Mais ce n'est pas tout à fait le cas, puisque le coefficient :

$$\left| \frac{\partial x^{\mu'}}{\partial x^{\mu}} \right| \quad (553)$$

apparaît dans la formule de changement de coordonnées.

Considérons alors la grandeur :

$$\epsilon_{\mu_1 \mu_2 \dots \mu_n} = \sqrt{|g|} \tilde{\epsilon}_{\mu_1 \mu_2 \dots \mu_n}. \quad (554)$$

Alors, nous avons, en substituant cette valeur dans l'expression précédente :

$$\frac{1}{\sqrt{|g(\mu')|}} \epsilon_{\mu'_1 \mu'_2 \dots \mu'_n} = \left| \frac{\partial x^{\mu'}}{\partial x^{\mu}} \right| \epsilon_{\mu_1 \mu_2 \dots \mu_n} \frac{\partial x^{\mu_1}}{\partial x^{\mu'_1}} \cdots \frac{\partial x^{\mu_n}}{\partial x^{\mu'_n}} \frac{1}{\sqrt{|g(\mu)|}}. \quad (555)$$

Or, nous savons que :

$$g_{\mu' \nu'} = \frac{\partial x^{\mu}}{\partial x^{\mu'}} \frac{\partial x^{\nu}}{\partial x^{\nu'}} g_{\mu \nu}. \quad (556)$$

Donc que :

$$|g(\mu')| = \left| \frac{\partial x^{\mu}}{\partial x^{\mu'}} \right|^2 |g(\mu)|. \quad (557)$$

Soit :

$$\frac{1}{\sqrt{|g(\mu')|}} = \left| \frac{\partial x^{\mu'}}{\partial x^{\mu}} \right| \frac{1}{\sqrt{|g(\mu)|}}. \quad (558)$$

En reportant ce résultat dans l'expression ci-dessus, on obtient :

$$\epsilon_{\mu'_1 \mu'_2 \dots \mu'_n} = \epsilon_{\mu_1 \mu_2 \dots \mu_n} \frac{\partial x^{\mu_1}}{\partial x^{\mu'_1}} \frac{\partial x^{\mu_2}}{\partial x^{\mu'_2}} \cdots \frac{\partial x^{\mu_n}}{\partial x^{\mu'_n}}. \quad (559)$$

Par conséquent,

$$\epsilon_{\mu_1 \mu_2 \dots \mu_n} = \sqrt{|g|} \tilde{\epsilon}_{\mu_1 \mu_2 \dots \mu_n} \quad (560)$$

se transforme comme un tenseur dans un changement de coordonnées, et est donc un tenseur  $n$  fois covariant, que l'on appelle **tenseur de Levi-Civita**. En revanche :

$$\tilde{\epsilon}_{\mu_1 \mu_2 \dots \mu_n} \quad (561)$$

n'est pas un tenseur. Mais le fait que sa loi de transformation dans un changement de coordonnées fasse intervenir le déterminant :

$$\left| \frac{\partial x^{\mu'}}{\partial x^{\mu}} \right|, \quad (562)$$

que l'on appelle **jacobien du changement de coordonnées**, fait qu'on appelle cette grandeur une **densité tensorielle** de poids 1. On appelle **poids** de la densité tensorielle, la puissance à laquelle est élevé le jacobien :

$$\left| \frac{\partial x^{\mu'}}{\partial x^{\mu}} \right|, \quad (563)$$

dans l'expression de changement de coordonnées.

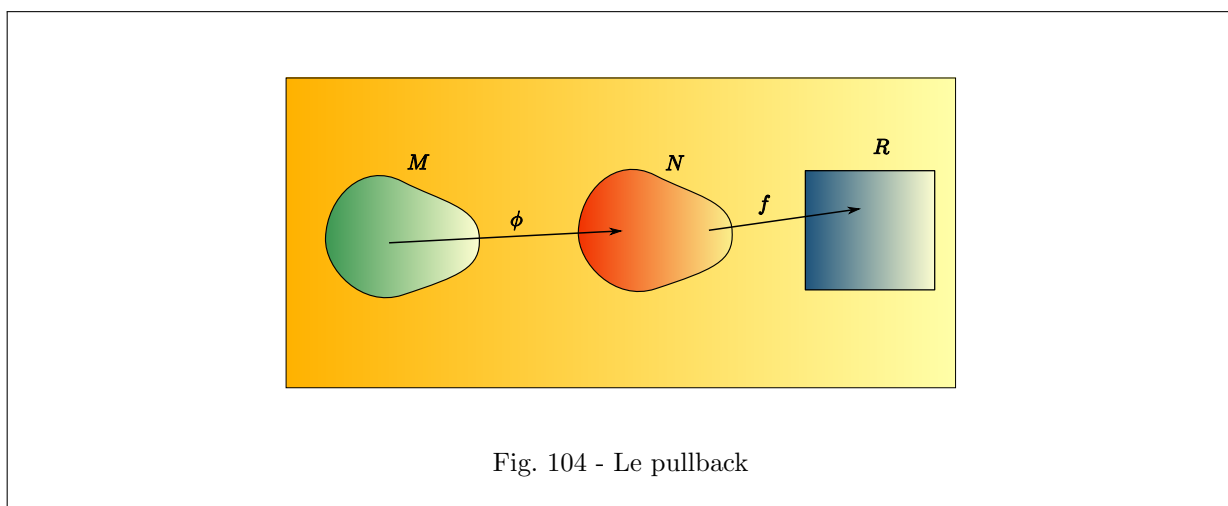
## 69 Scène X-5 : Le pullback

EURISTIDE : Nous verrons plus tard l'utilisation de ce tenseur de Levi-Civita. Nous avons maintenant l'avantage de disposer d'un tenseur qui est capable de façon générale d'appliquer à un tenseur par multiplication les propriétés du symbole antisymétrique.

MATHINE : Pour l'instant nous allons revenir à quelques notions de géométrie autour des variétés et de la notion de carte sur ces variétés.

Considérons deux variétés  $M$  et  $N$  ayant pour systèmes de coordonnées respectifs  $x^\mu$  et  $y^\alpha$ .

Considérons une carte de  $M$  dans  $N$ , que nous noterons  $\phi$  :



Et considérons, par ailleurs, une fonction  $f$  de  $N$  dans  $\mathbb{R}$ . Si nous considérons la fonction composée de  $\phi$  et de  $f$ , notée  $f \circ \phi$ , nous avons créé une fonction de  $M$  dans  $\mathbb{R}$ . Cette fonction permet, en quelque sorte, de tirer en arrière la fonction  $f$  de  $N$  dans  $\mathbb{R}$ , en la ramenant au moyen de  $\phi$ , de  $M$  dans  $\mathbb{R}$ .

C'est la raison pour laquelle on appelle cette fonction  $f \circ \phi$  le **pullback de  $f$  par  $\phi$**  (pullback signifiant "tirer en arrière" en anglais), et nous le notons :

$$\phi_* f = f \circ \phi. \quad (564)$$

BEATRIX : Je vois bien ce que c'est, mais à quoi cela sert-il ?

EURISTIDE : Supposons que nous disposions d'une sphère  $N$  et d'un ellipsoïde  $M$  qui sont deux variétés de dimension 2. Supposons que nous disposions d'une carte  $\phi$  de l'ellipsoïde sur la sphère, permettant de "cartographier" l'ellipsoïde au moyen de la sphère :

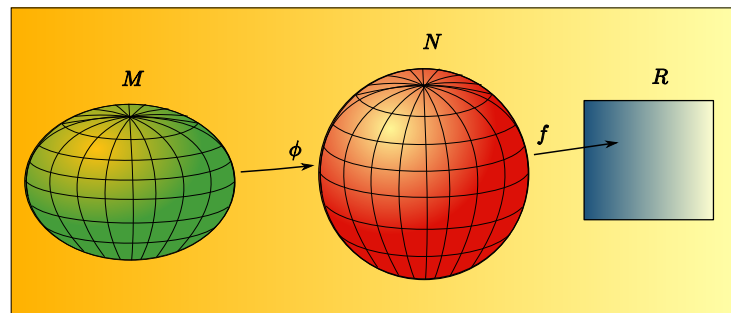


Fig. 105 - Transport par pullback d'une fonction sur l'ellipsoïde

Le pullback va permettre de calquer la fonction  $f$  définie sur la sphère sur la surface de l'ellipsoïde. Par exemple, si la fonction  $f$  représente un coloriage, une altitude ou une température sur la sphère  $N$ , nous allons, au moyen du pullback, reporter cette fonction point à point de la sphère dans le monde de la surface de l'ellipsoïde.

## 70 Scène X-6 : Le pushforward

MATHINE : Maintenant, nous pourrions avoir envie de reporter une fonction de la variété  $M$  sur la variété  $N$  :

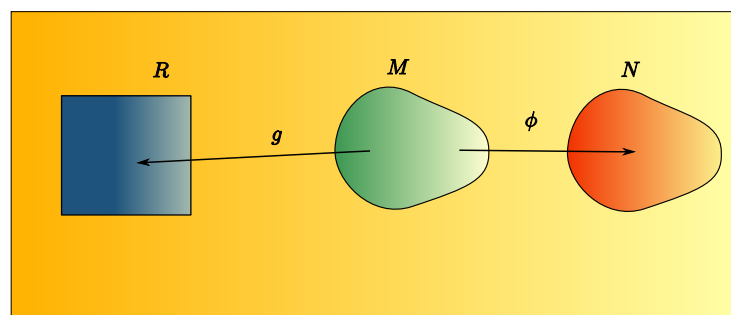


Fig. 106 - Le pushforward

Mais le schéma ci-dessus montre bien que les flèches des fonctions ne correspondent pas, et qu'il est impossible de composer la fonction  $\phi$  avec la fonction  $g$  pour créer une fonction de  $N$  dans  $\mathbb{R}$ .

Mais nous allons en fait tenter de transporter des vecteurs de la variété  $M$  dans la variété  $N$ . Et pour cela, nous allons devoir ruser.

Une fonction-vecteur dans la variété  $M$  peut être considéré comme un opérateur de dérivée qui transforme une fonction en nombre. En effet, nous avons vu que les fonctions-vecteurs de coordonnées sur la variété  $M$  pouvaient être définis comme l'opérateur :

$$\overrightarrow{\frac{\partial}{\partial x^\mu}} \quad (565)$$

Cet opérateur est applicable à n'importe quelle fonction définie sur  $M$ . Il en va de même pour tout vecteur  $\overrightarrow{V}$  défini sur  $M$ , qui pourra être considéré comme un opérateur de dérivée applicable aux fonctions définies sur  $M$  dans  $\mathbb{R}$ . Mais voici donc la ruse que je propose. Nous connaissons, grâce au pullback, un moyen de transporter une fonction définie sur  $N$  en une fonction définie sur  $M$  à valeurs dans  $\mathbb{R}$ . En effet, si  $f$  est une fonction définie sur  $N$ , alors :

$$\phi_* f \quad (566)$$

est une fonction définie sur  $M$  à valeurs dans  $\mathbb{R}$ . Nous pouvons donc considérer l'image de cette fonction par notre fonction-vecteur  $\overrightarrow{V}$  :

$$\overrightarrow{V}(\phi_* f). \quad (567)$$

Revenant à la fonction  $f$ , nous pouvons maintenant considérer que nous avons défini une nouvelle fonction-vecteur (que nous appellerons également champ de vecteurs)  $\overrightarrow{W}$  sur  $N$  au moyen de l'expression qui définit son application aux fonctions de  $N$  :

$$\overrightarrow{W}(f) = \overrightarrow{V}(\phi_* f). \quad (568)$$

La fonction-vecteur  $\overrightarrow{W}$  est appelé le **pushforward de  $\overrightarrow{V}$  par  $\phi$**  et est noté :

$$\phi_* \overrightarrow{V}(f) = \overrightarrow{W}(f) = \overrightarrow{V}(\phi_* f). \quad (569)$$

("pushforward" signifie "pousser en avant" en anglais.)

**BEATRIX** : Nous avons donc réussi à pousser un champ de vecteurs de  $M$  dans  $N$  en considérant que l'action du champ de vecteurs sur  $N$  est égale à l'action du champ de vecteurs d'origine dans  $M$  sur le pullback de la fonction de  $N$  dans  $M$  par  $\phi$ .

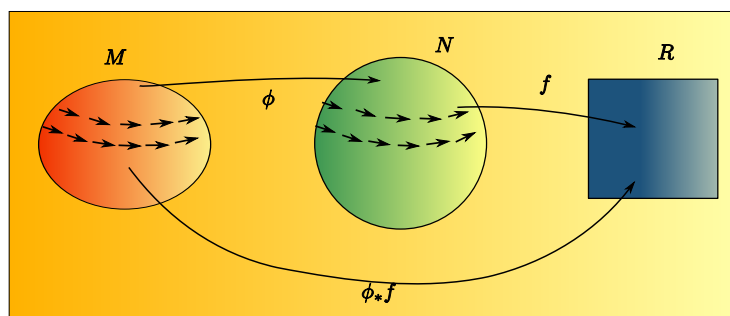


Fig. 107 - Transport d'un champ de vecteurs par pushforward

Cela revient, avec un procédé un peu compliqué, à déplacer un champ de vecteurs de la variété  $M$  sur la variété  $N$ , en veillant à ce que l'effet sur les fonctions définies dans  $N$  soit le même que sur leur pullback dans  $M$ .

EURISTIDE : Voilà. Autrement dit, si la fonction  $f$  est, par exemple, la fonction couleur à la surface de la sphère  $N$  comportant des points jaunes et des points verts, si un champ de vecteurs à la surface de l'ellipsoïde  $M$  permet de se déplacer des points de couleur jaune aux points de couleur verte pour la fonction couleur sur  $N$  appliquée par pullback à  $M$ , alors ce champ de vecteurs peut être appliqué sur  $N$ , en transformant les points jaunes de  $N$  en points verts.

MATHINE : Nous avons vu que nous pouvons tirer en arrière des fonctions. Or nous savons qu'un tenseur  $l$  fois covariant est assimilable à un ensemble de  $l$  fonctions à images dans  $\mathbb{R}$ . Par conséquent, nous pouvons également tirer en arrière un tenseur quelconque  $l$  fois covariant.

En effet, un tenseur  $T_{\mu_1 \dots \mu_l}$   $l$  fois covariant "mange"  $l$  fonctions-vecteurs et restitue 1 nombre réel.

Pour tirer en arrière le tenseur, il faut donc considérer l'action du tenseur tiré en arrière sur les fonctions-vecteurs de la variété  $M$ . Pour cela, nous allons pousser en avant les fonctions-vecteurs de  $M$  dans  $N$  par  $\phi$ , et considérer l'action du tenseur d'origine  $T_{\mu_1 \dots \mu_l}$  sur ces vecteurs poussés en avant.

Le résultat définira le pullback du tenseur  $T_{\mu_1 \dots \mu_l}$ .

En d'autres termes, si  $\overrightarrow{V^{(1)}}, \dots, \overrightarrow{V^{(l)}}$  sont des fonctions-vecteurs de la variété  $M$ , alors, on définit le pullback du tenseur  $T$ , soit  $\phi_* T$  par :

$$(\phi_* T)(\overrightarrow{V^{(1)}}, \overrightarrow{V^{(2)}}, \dots, \overrightarrow{V^{(l)}}) = T(\phi^* \overrightarrow{V^{(1)}}, \phi^* \overrightarrow{V^{(2)}}, \dots, \phi^* \overrightarrow{V^{(l)}}). \quad (570)$$

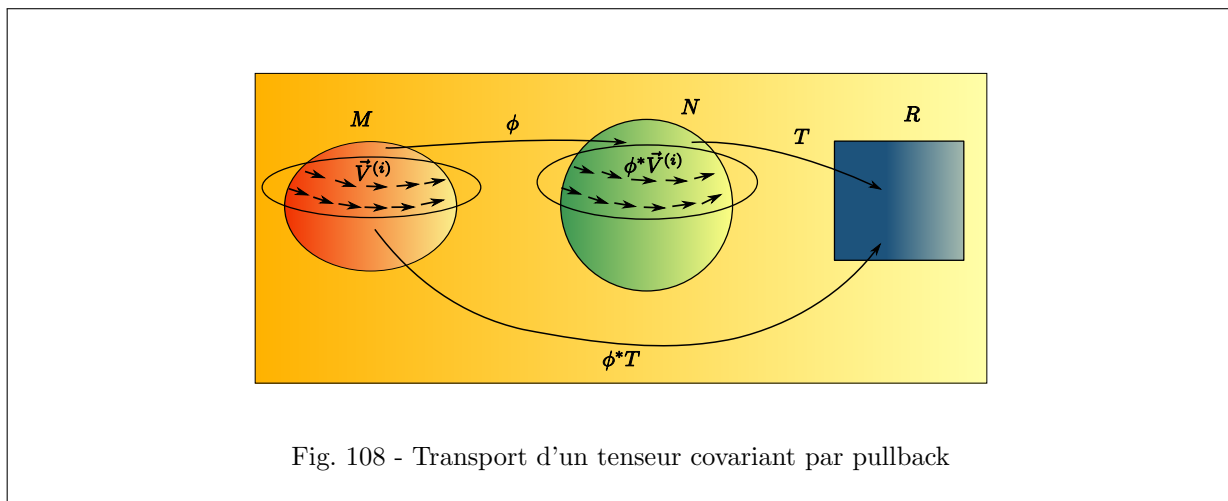


Fig. 108 - Transport d'un tenseur covariant par pullback

De façon similaire, nous savons qu'un tenseur  $k$  fois contravariant est assimilable à un ensemble de  $k$  champs vectoriels. Par conséquent, nous pouvons pousser en avant un tenseur quelconque  $k$  fois contravariant. En effet, un tenseur  $k$  fois contravariant  $T^{\mu_1 \dots \mu_k}$  "mange"  $k$  fonctions pour restituer  $k$  fonctions-vecteurs.

Pour pousser en avant le tenseur  $T$ , il faut considérer l'action de ce tenseur poussé en avant sur les fonctions de la variété  $N$ . Pour cela, nous allons tirer en arrière les fonctions de la variété  $N$  dans  $M$  par  $\phi$ , et considérer l'action du tenseur d'origine  $T^{\mu_1 \dots \mu_k}$  sur ces fonctions tirées en arrière.

Le résultat définira le pushforward du tenseur  $T^{\mu_1 \dots \mu_k}$ . En d'autres termes, si  $f^{(1)}, \dots, f^{(k)}$  sont des fonctions de la variété  $N$  dans  $\mathbb{R}$ , alors on définit le pushforward du tenseur  $T$ , soit  $\phi_* T$  par :

$$\phi_* T(f^{(1)}, f^{(2)}, \dots, f^{(k)}) = T(\phi_* f^{(1)}, \phi_* f^{(2)}, \dots, \phi_* f^{(k)}). \quad (571)$$

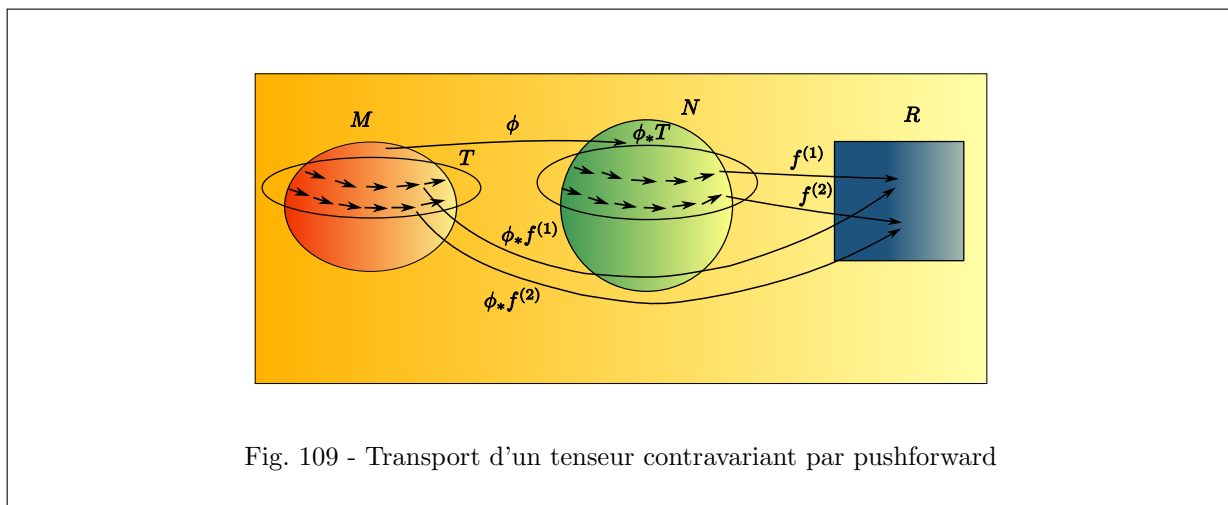


Fig. 109 - Transport d'un tenseur contravariant par pushforward

**EURISTIDE** : En plus concret, considérons la fonction couleur ayant deux valeurs possibles : jaune et vert. Considérons également la fonction lumière, pouvant avoir deux valeurs possibles : clair et sombre. Alors, le pullback de  $f^{(1)}$  (noté  $\phi_* f^{(1)}$ ) et  $f^{(2)}$  (noté  $\phi_* f^{(2)}$ ) permettent de dessiner sur l'ellipsoïde  $M$  les points avec une couleur jaune ou verte et une lumière claire ou sombre, en conformité avec ce qui est dessiné sur la sphère.

Considérons maintenant sur l'ellipsoïde, un tenseur qui fournit les champs vectoriels associés aux fonctions  $\phi_* f^{(1)}$  et  $\phi_* f^{(2)}$  permettant de transformer les points jaunes en points verts et les points sombres en points clairs. Alors, le pushforward du tenseur sur la variété  $N$  permet de définir sur la sphère  $N$  un tenseur transformant à son tour les points jaunes en points verts et les points sombres en points clairs.

**BEATRIX** : C'est maintenant lumineux! Et au passage, ça permet de mieux comprendre le rôle du tenseur qui permet un cumul de plusieurs transformations le long de champs de vecteurs.

**EURISTIDE** : On va prendre un exemple un peu plus complexe, et lié au tenseur métrique, pour exprimer plus directement ce qu'est un pullback. Supposons que la variété  $M$  soit la sphère  $S^2$  de dimension 2, et que la variété  $N$  soit l'espace  $\mathbb{R}^3$  de dimension 3.

Prenons pour coordonnées dans  $M$  :

$$x^\mu = (\theta, \phi). \quad (572)$$



Prenons pour coordonnées dans  $N$  :

$$y^\alpha = (x, y, z). \quad (573)$$

Alors, la carte  $\psi$  de  $M$  dans  $N$  est donnée par :

$$\psi(\theta, \phi) = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta). \quad (574)$$

Considérons une fonction  $f$  de  $N$  dans  $\mathbb{R}$  quelconque et une fonction-vecteur  $\vec{V}$  de  $M$  appliqué au pullback de cette fonction par  $f$ . Alors, nous considérons la fonction :

$$\vec{V}(\psi_* f). \quad (575)$$

Or, une fonction-vecteur peut s'exprimer comme un opérateur de dérivée partielle, appliqué à la fonction, donc :

$$\begin{aligned} \vec{V}(\psi_* f) &= V^\mu \frac{\overrightarrow{\partial}}{\partial x^\mu}(\psi_* f) \\ &= V^\mu \frac{\overrightarrow{\partial}}{\partial x^\mu}(f \circ \psi) \\ &= V^\mu \frac{\partial y^\alpha}{\partial x^\mu} \frac{\overrightarrow{\partial}}{\partial y^\alpha}(f \circ \psi). \end{aligned} \quad (576)$$

On peut donc considérer, d'après cette dernière expression, que le pullback d'une fonction est assimilable à un opérateur constitué par la matrice :

$$\left( \frac{\partial y^\alpha}{\partial x^\mu} \right), \quad (577)$$

puisque cette matrice transforme les coordonnées d'un vecteur de  $M$  en celles d'un vecteur de  $N$ , et donc permet de définir la fonction  $\psi_* f$ , pullback de  $f$ .

Par conséquent, pour trouver le tenseur métrique dans  $S^2$  (qui est un tenseur 2 fois covariant), nous pouvons effectuer un pullback de  $N$  dans  $M$  par  $\phi$ , en appliquant la matrice de transformation deux fois au tenseur 2 fois covariant  $g_{\alpha\beta}$  de  $\mathbb{R}^3$ .

Le tenseur métrique de  $\mathbb{R}^3$ , qui est un espace euclidien, est :

$$g_{\alpha\beta} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (578)$$

La matrice de transformation du pullback est :

$$\begin{aligned} \left( \frac{\partial y^\alpha}{\partial x^\mu} \right) &= \begin{pmatrix} \frac{\partial x}{\partial \theta} & \frac{\partial y}{\partial \theta} & \frac{\partial z}{\partial \theta} \\ \frac{\partial x}{\partial \phi} & \frac{\partial y}{\partial \phi} & \frac{\partial z}{\partial \phi} \end{pmatrix} \\ &= \begin{pmatrix} \cos \theta \cos \phi & \cos \theta \sin \phi & -\sin \theta \\ -\sin \theta \cos \phi & -\sin \theta \sin \phi & 0 \end{pmatrix}. \end{aligned} \quad (579)$$

D'où :

$$(\phi_* g)_{\mu\nu} = \left( \frac{\partial y^\alpha}{\partial x^\mu} \right) \left( \frac{\partial y^\beta}{\partial x^\nu} \right). \quad (580)$$

Or  $g \circ \psi = g$  puisque  $g$  est la matrice unité dans un espace plat euclidien. D'où :

$$(\phi_* g)_{\mu\nu} = \left( \frac{\partial y^\alpha}{\partial x^\mu} \right) \begin{pmatrix} \cos \theta \cos \phi & -\sin \theta \sin \phi \\ \cos \theta \sin \phi & \sin \theta \cos \phi \\ -\sin \theta & 0 \end{pmatrix}, \quad (581)$$

soit :

$$\begin{aligned}
 (\phi_*g)_{\mu\nu} &= \begin{pmatrix} \cos \theta \cos \phi & \cos \theta \sin \phi & -\sin \theta \\ -\sin \theta \sin \phi & \sin \theta \cos \phi & 0 \end{pmatrix} \begin{pmatrix} \cos \theta \cos \phi & -\sin \theta \sin \phi \\ \cos \theta \sin \phi & \sin \theta \cos \phi \\ -\sin \theta & 0 \end{pmatrix} \\
 &= \begin{pmatrix} \cos^2 \theta \cos^2 \phi + \cos^2 \theta \sin^2 \phi + \sin^2 \theta & 0 \\ 0 & \sin^2 \theta \sin^2 \phi + \sin^2 \theta \cos^2 \phi \end{pmatrix}.
 \end{aligned} \tag{582}$$

Soit :

$$(\phi_*g)_{\mu\nu} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2 \theta \end{pmatrix}. \tag{583}$$

BEATRIX : On a donc pu déduire directement le tenseur métrique de la sphère  $S^2$  à partir du pullback du tenseur métrique de l'espace plat  $\mathbb{R}^3$  de dimension 3. Et nous retrouvons ce que nous savons déjà :

$$ds^2 = d\theta^2 + \sin^2 \theta d\phi^2, \tag{584}$$

que nous avons calculé tout à l'heure beaucoup moins élégamment par substitution de l'expression de l'élément de ligne.

Je comprends maintenant tout l'intérêt de cette méthode des pullback et des pushforward. Nous pouvons transporter la structure d'une variété dans une autre à partir du moment où nous avons pu définir une carte d'une variété sur l'autre. Et concrètement, cela permet de plaquer les caractéristiques d'une variété que nous connaissons déjà par son tenseur métrique, sur une nouvelle variété que nous devons découvrir, mais dont nous connaissons au moins une carte.

MATHINE : Nous avons vu que nous pouvions effectuer un pullback de tenseurs  $l$  fois covariants, et un pushforward de tenseurs  $k$  fois contravariants.

Mais on peut se demander pourquoi on ne peut pas faire un pushforward ou un pullback de tenseurs quelconques mixtes (à la fois covariants et contravariants.) Nous avons vu que ce n'était pas possible à cause du sens des flèches dans les diagrammes indiquant les fonctions et la carte. Cela est dû notamment au sens de la flèche de la fonction  $\phi$  qui est la carte de la variété  $M$  sur la variété  $N$ .

Mais nous pouvons très bien imaginer que la fonction  $\phi$  est inversible, c'est-à-dire qu'il est possible de passer de la variété  $N$  à la variété  $M$  par une carte qui est la fonction inverse de  $\phi$ . Nous noterons cette carte inverse par  $\phi^{-1}$ . Dans ce cas, un tenseur covariant dans  $M$  pourra être poussé en avant dans  $N$  grâce à la carte  $\phi^{-1}$ , et de la même façon, un tenseur contravariant dans  $N$  pourra faire l'objet d'un pullback dans  $M$  grâce à la carte  $\phi^{-1}$ .

En résumé, si la carte est inversible (c'est-à-dire si elle est bijective, en faisant correspondre à tout élément de  $M$  un et un seul élément distinct de  $N$ ), alors on peut pousser en avant ou tirer en arrière n'importe quel tenseur, qu'il soit covariant, contravariant ou mixte.

Dans ce cas, le pushforward d'un tenseur  $T_{\nu_1 \dots \nu_l}^{\mu_1 \dots \mu_k}$   $l$  fois covariant et  $k$  fois contravariant pourra s'écrire :

$$(\phi^*T)_{\beta_1 \dots \beta_l}^{\alpha_1 \dots \alpha_k} = \frac{\partial y^{\alpha_1}}{\partial x^{\mu_1}} \cdots \frac{\partial y^{\alpha_k}}{\partial x^{\mu_k}} \frac{\partial x^{\nu_1}}{\partial y^{\beta_1}} \cdots \frac{\partial x^{\nu_l}}{\partial y^{\beta_l}} T_{\nu_1 \dots \nu_l}^{\mu_1 \dots \mu_k}. \tag{585}$$

Une fonction  $\phi$  bijective, donc inversible, porte le nom de **difféomorphisme**.

BEATRIX : Mais je suppose que le fait d'avoir une carte  $\phi$  qui soit un difféomorphisme impose certaines contraintes sur les variétés  $M$  et  $N$ . Par exemple, elles doivent être de même dimension, et avoir le même nombre d'éléments (même s'il est infini), etc.

EURISTIDE : C'est très vrai, cela restreint le champ d'application à des variétés de même dimension ; on dit que ces variétés sont équivalentes. Mais ce n'est pas grave, parce que nous allons utiliser ces difféomorphismes pour étudier la variation des tenseurs au sein d'une même variété, donc considérer que  $M$  et  $N$  sont une même et unique variété.

## 71 Scène X-7 : Les courbes intégrales

MATHINE : Grâce aux difféomorphismes, nous pouvons pousser en avant ou tirer en arrière n'importe quel tenseur.

Nous savons que nous ne pouvons pas facilement comparer les valeurs d'un tenseur en deux points distincts  $P$  et  $Q$ , car ces valeurs, ou composantes, sont exprimées dans deux systèmes de coordonnées locaux distincts.

EURISTIDE : Par exemple, à la surface de la sphère, nous ne pouvons pas comparer directement le vecteur tangent à un parallèle en un point  $P$ , et le vecteur tangent à un parallèle au point  $Q$ . Intuitivement, pour un être infiniment petit vivant à la surface de cette sphère (et n'ayant pas de notion d'une troisième dimension), ces deux vecteurs ont même direction, puisqu'ils sont tous deux parallèles à un parallèle, et pointent vers l'Est.

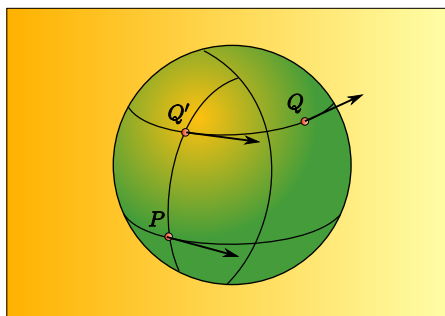


Fig. 110 - Comparaison de vecteurs sur une sphère

Pour faire cette comparaison rigoureusement, nous devons d'abord déplacer le vecteur qui se trouve au point  $Q$  le long de son parallèle, tout en le gardant parallèle à ce parallèle, pour arriver au point  $Q'$ , qui se trouve sur le même méridien que  $P$ . Puis, nous le déplaçons le long du méridien pour le conduire sur le point  $P$ , tout en gardant sa direction toujours perpendiculaire au méridien (c'est-à-dire parallèle aux parallèles.) Une fois que nous avons réalisé cette opération fastidieuse, nous pouvons enfin constater si le vecteur en  $Q$  coïncide bien avec le vecteur en  $P$ .

Mais attention ! Cette vision de l'égalité des vecteurs n'est valable que dans la variété de dimension 2 qui est la surface de la sphère, donc du point de vue d'un habitant infiniment petit vivant à la surface de cette sphère. Il est évident qu'un être en 3 dimensions regardant cette sphère de l'extérieur s'aperçoit immédiatement que les deux vecteurs plongés dans l'espace de dimension 3 ne sont pas identiques !

En déplaçant le vecteur qui se trouvait en  $Q$  vers le point  $P$ , nous avons fait l'équivalent d'un pullback, pour un difféomorphisme qui transformerait le point  $P$  en le point  $Q$ . On peut bien entendu procéder de même avec un tenseur quelconque.

**MATHINE** : Cette illustration va donc nous permettre d'introduire le fait que nous avons maintenant une technique pour déplacer, puis comparer deux tenseurs appliqués en deux points distincts de la variété. Considérons une variété  $M$ , un difféomorphisme  $\phi : M \rightarrow M$  sur cette variété.

Considérons un point  $P$  de la variété et considérons un champ tensoriel  $T_{\nu_1 \dots \nu_l}^{\mu_1 \dots \mu_k}(x)$ , c'est-à-dire une fonction qui définit un tenseur en chaque point de la variété. Soit :

$$T_{\nu_1 \dots \nu_l}^{\mu_1 \dots \mu_k}(P), \quad (586)$$

la valeur de ce tenseur au point  $P$ . Pour comparer la valeur de ce tenseur en  $P$  avec sa valeur en  $\phi(P)$ , nous pouvons exprimer la différence entre le tenseur au point  $P$  et la valeur du tenseur en  $\phi(P)$  tirée en arrière au moyen de son pullback par  $\phi$ . Ceci s'écrit :

$$T_{\nu_1 \dots \nu_l}^{\mu_1 \dots \mu_k}(P) - \phi_* (T_{\nu_1 \dots \nu_l}^{\mu_1 \dots \mu_k}(\phi(P))). \quad (587)$$

Nous allons maintenant considérer que nous n'avons pas un difféomorphisme  $\phi$ , mais une famille  $\phi_t$  continue à un paramètre de difféomorphismes. Lorsque  $t$  varie sur l'ensemble  $\mathbb{R}$ , le difféomorphisme va déplacer un point donné en différents autres points, qui constitueront une courbe.

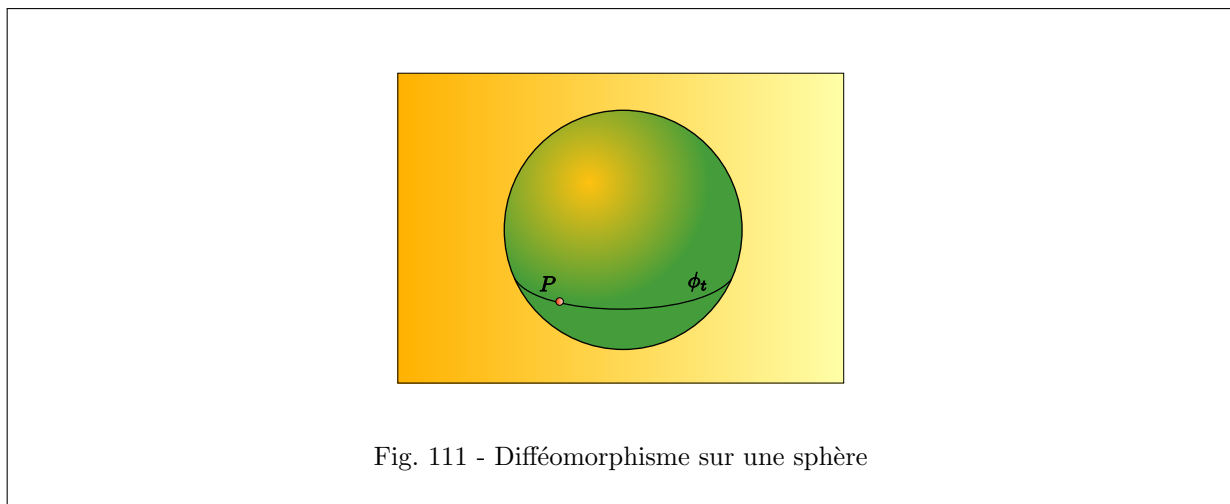


Fig. 111 - Difféomorphisme sur une sphère

Nous allons pouvoir associer à l'ensemble de ces courbes pour tous les points de la variété, une famille de vecteurs tangents aux courbes, comme suit :

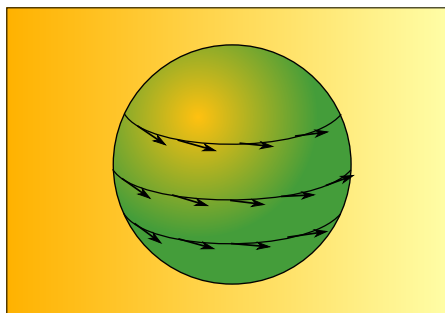


Fig. 112 - Champ vectoriel sur une sphère

Cette famille de vecteurs tangents  $V^\mu(x)$  constitue une fonction vectorielle qui à chaque point de la variété associe un vecteur. On appelle aussi cette famille de vecteurs un **champ vectoriel**.

Il y a donc une correspondance exacte entre la famille de difféomorphismes  $\phi_t$ , le champ vectoriel  $V^\mu(x)$  et les courbes générées par le difféomorphisme.

Inversement, on démontre que si on considère un champ vectoriel quelconque  $V^\mu(x)$ , alors on peut lui faire correspondre un ensemble de courbes auxquelles ces vecteurs sont tangents, et par conséquent une famille de difféomorphismes générant ces courbes.

Les courbes associées à un tel champ vectoriel sont appelées **courbes intégrales** du champ vectoriel, et le champ vectoriel est appelé **générateur** des difféomorphismes. Les difféomorphismes, quant à eux, représentent l'écoulement le long des courbes intégrales.

**BEATRIX** : Cela rappelle fortement des notions d'électromagnétisme, je trouve. Le champ magnétique est le champ vectoriel. Les courbes sont les lignes de champ, c'est-à-dire la direction dans laquelle les aimants dirigent leur axe. Le difféomorphisme correspond à l'écoulement du flux magnétique le long des lignes de champ.

**MATHINE** : Ce que nous avons décrit pour un champ vectoriel, nous pouvons évidemment le généraliser pour des tenseurs constituant un champ tensoriel. Maintenant que nous possédons un moyen de comparer les valeurs d'un champ tensoriel en deux points, nous allons pouvoir regarder à quelle vitesse varie un champ tensoriel entre deux points.

Nous allons donc, comme nous l'avons fait à propos de la vitesse en cinématique, concevoir la notion de dérivée d'un champ tensoriel, comme étant la limite de la variation par pullback de la valeur du tenseur entre deux points en relation par le difféomorphisme  $\phi_t$ , lorsque  $t$  tend vers 0, c'est-à-dire lorsque le difféomorphisme tend vers la fonction identité.

Autrement dit, nous considérons un tenseur  $T_{\nu_1 \dots \nu_l}^{\mu_1 \dots \mu_k}$ , en un point  $P$ , et une famille de difféomorphismes  $\phi_t$ . La variation de la valeur du tenseur est :

$$\Delta_t T_{\nu_1 \dots \nu_l}^{\mu_1 \dots \mu_k}(P) = \phi_{t*} (T_{\nu_1 \dots \nu_l}^{\mu_1 \dots \mu_k}(\phi_t(P))) - T_{\nu_1 \dots \nu_l}^{\mu_1 \dots \mu_k}(P). \quad (588)$$

Cette formule compliquée, c'est la différence entre la valeur du tenseur au point  $\phi_t(P)$ , tiré en arrière par le pullback par  $\phi_t$ , et la valeur du tenseur au point  $P$ .

On appelle **dérivée de Lie** du tenseur le long du champ vectoriel  $\vec{V}$  défini par le difféomorphisme  $\phi_t$ , la limite lorsque  $t$  tend vers 0 de cette différence, soit l'expression :

$$\mathcal{L}_{\vec{V}} T_{\nu_1 \dots \nu_l}^{\mu_1 \dots \mu_k} = \lim_{t \rightarrow 0} \left( \frac{\Delta_t T_{\nu_1 \dots \nu_l}^{\mu_1 \dots \mu_k}}{t} \right). \quad (589)$$

## 72 Scène X-8 : Le champ vectoriel de Killing

**BEATRIX** : Nous avons donc créé une nature toute particulière de dérivée, tenant compte du fait qu'il faut déplacer un tenseur ou un vecteur sur un point donné, avant de pouvoir le comparer à un autre tenseur ou vecteur qui se trouve en ce même point.

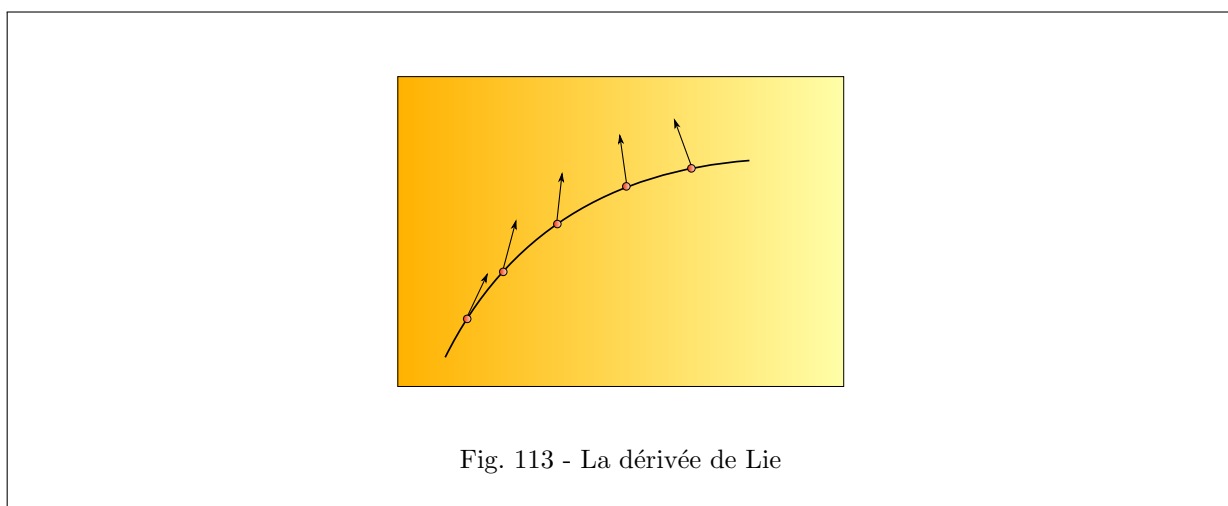


Fig. 113 - La dérivée de Lie

Si on considère un champ vectoriel qui évolue le long d'une courbe comme indiqué ci-dessus, alors, la dérivée de Lie de ce champ vectoriel le long de la courbe (ou le long du champ vectoriel qui génère cette courbe), va permettre de quantifier la vitesse à laquelle le vecteur s'écarte de la courbe, en compensant le fait que la courbe elle-même s'incurve.

**MATHINE** : C'est exactement l'illustration du concept de dérivée de Lie dans une variété de dimension 1, c'est-à-dire pour une courbe.

**EURISTIDE** : Nous allons terminer ces explications sur les variétés par une analyse de ce qu'est la symétrie dans une variété. Nous aurons beaucoup besoin de cette notion, car les symétries de l'univers permettront de simplifier grandement, dans certains cas de figure, les équations de la relativité générale. Il est donc important que nous ayons des outils pour modéliser la notion de symétrie dans une variété.

MATHINE : On dit qu'un difféomorphisme  $\phi$  est une **symétrie** d'un tenseur si le tenseur est invariant après avoir été tiré en arrière par  $\phi$ , autrement dit si :

$$\phi_* T = T. \quad (590)$$

Si nous avons une famille de symétries  $\phi_t$  générée par un champ vectoriel  $\vec{V}$ , alors nous aurons :

$$\phi_{t*} T = T, \quad (591)$$

ce qui, d'après la définition de la dérivée de Lie, nous conduira à :

$$\mathcal{L}_{\vec{V}} T = 0. \quad (592)$$

EURISTIDE : En application de cela, nous allons considérer les symétries du tenseur métrique  $g_{\mu\nu}$ . Un difféomorphisme  $\phi$  pour lequel :

$$\phi_* g_{\mu\nu} = g_{\mu\nu}, \quad (593)$$

est appelé **isométrie**.

Si nous avons une famille d'isométries  $\phi_t$  générée par un champ vectoriel  $V^\mu(x)$ , alors  $V^\mu$  est appelé **champ vectoriel de Killing**.

La condition pour que  $V^\mu$  soit un vecteur de Killing est que :

$$\mathcal{L}_{\vec{V}} g_{\mu\nu} = 0. \quad (594)$$

Par conséquent, la métrique est constante le long d'un champ de vecteurs de Killing.

---

## Douzième partie

# Acte XI - La courbure

### 73 Scène XI-1 : La notion de courbure

BEATRIX : Donc, si je comprends bien, les vecteurs de Killing caractérisent la symétrie de la métrique, donc celle de l'espace.

EURISTIDE : Oui, c'est cela. Et nous utiliserons cette notion puissante lorsque nous étudierons des exemples d'espaces à symétrie centrale, comme les étoiles massives, ou les trous noirs.

Nous allons maintenant attaquer le concept de courbure, qui, tu l'auras compris, est fondamental dans la construction de la théorie de la relativité générale.

BEATRIX : D'accord. Mais avant cela, je propose qu'on récapitule ce que nous avons appris sur les variétés. C'était un peu abstrait, et ça vaut la peine de faire un petit résumé de ce qu'on a vu.

EURISTIDE : Alors, je t'écoute ?

BEATRIX : Bien. Nous avons commencé par introduire le concept géométrique de variété, qui représente un espace curviligne dans sa généralité. En effet, un espace curviligne, du fait de sa courbure, ne peut pas toujours être déplié et représenté sur une surface plate ou plus généralement dans un espace plat. Le concept de variété permet donc d'associer à un espace quelconque, un jeu de cartes que l'on appelle un atlas, qui permet la représentation de cet espace curviligne par un ensemble de projections sur un espace plat de toutes les régions de l'espace. En chacune des régions de l'espace curviligne, on va pouvoir représenter localement cette région par une carte dessinée dans un espace plat. C'est typiquement ce qu'on fait avec des cartes géographiques de la Terre.

Ce faisant, en calquant le système de coordonnées de la carte sur la variété localement dans chaque région, nous avons pu créer deux choses : un système de coordonnées curvilignes sur la variété, et en chaque point, un espace plat tangent à la variété en ce point. L'espace tangent est l'ensemble de tous les vecteurs tangents aux courbes dessinées sur la variété et passant par ce point. Cet espace tangent permet de définir naturellement des directions d'axes de coordonnées sur la variété. On a défini dans la foulée, l'espace cotangent à la variété en un point. C'est l'espace des gradients de toutes les courbes passant par ce point.

Puis nous avons regardé ce que devenait le tenseur métrique dans une variété. On s'est aperçus que ce qui caractérisait le fait qu'un espace était curviligne était que le tenseur métrique possédait des composantes qui n'étaient pas constantes. Puis, en choisissant un système de coordonnées bien adéquat, nous avons réussi à obtenir un tenseur métrique sous forme de matrice diagonale. Le système de coordonnées en question s'appelle le système de coordonnées normales riemaniennes. C'est un système de coordonnées défini localement en chaque point de la variété, et permettant de faire l'approximation que l'espace en tout point est localement plat, à condition de considérer une portion de l'espace extrêmement petite.

Puis nous avons découvert les merveilles du pullback et du pushforward. C'est une technique permettant de rapprocher deux vecteurs distants sur une variété, afin qu'ils soient tous deux exactement au même point, et afin qu'on puisse les comparer. Parce qu'en effet, lorsque deux vecteurs sont distants dans un espace



curviligne, l'un a subi une rotation par rapport à l'autre du fait de la courbure de l'espace. Ce qui nous intéresse, c'est la différence relative entre les deux vecteurs dans l'espace curviligne, différence relative qui est calculée en se débarrassant de la courbure de l'espace du fait qu'on remet les deux vecteurs au même endroit dans l'espace.

Pour finir, nous avons vu que nous pouvions définir des courbes intégrales tangentes à tous les vecteurs d'un champ vectoriel, et de cette façon définir une dérivée le long d'une courbe intégrale appelée dérivée de Lie, représentant la variation relative d'un vecteur le long de cette courbe, en négligeant les effets de la courbure. Avec ce formalisme, nous avons pu définir un champ vectoriel de Killing, qui est l'ensemble des vecteurs définissant des courbes intégrales le long desquelles le tenseur métrique est constant au sens de la dérivée de Lie, c'est-à-dire le long desquelles le tenseur métrique est constant relativement à la courbure de l'espace. Ouf!

EURISTIDE : Bien! Merci pour ce résumé. Attaquons maintenant la courbure. La courbure est, nous l'avons vu, une propriété intrinsèque de l'espace, qui est "codée" dans son tenseur métrique. C'est une propriété qui se traduit par différents phénomènes observables, tels que la courbure des axes de coordonnées, l'impossibilité de représenter l'espace globalement par une carte plate dans un espace euclidien, les coordonnées des points de l'espace dans le système de coordonnées ne sont pas des constantes, mais des fonctions, les triangles sur les surfaces curvilignes ont une somme de leurs angles différente de 180 degrés. Nous allons nous attacher maintenant à définir une quantité mathématique qui caractérise précisément la courbure d'un espace.

BEATRIX : Oui, puisque la relativité générale traite de la courbure de l'espace-temps, il nous faudra un modèle mathématique pour la courbure, nous permettant d'écrire ainsi les équations de la relativité générale qui la feront intervenir.

EURISTIDE : C'est exactement ce que nous devons faire. Pour l'instant, rappelons d'un peu plus près la notion intuitive de courbure.

Un exemple d'espace curviligne est la surface d'une sphère appelée  $S^2$  en géométrie.

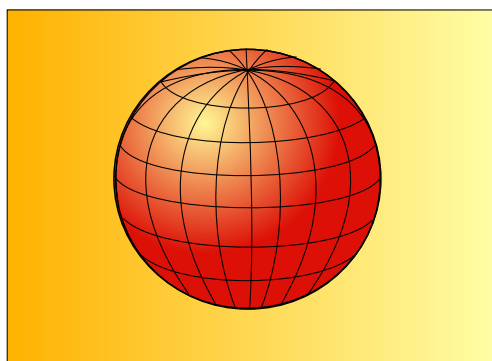


Fig. 114 - La sphère est un espace curviligne

BEATRIX : Comment sait-on que cet espace est curviligne ?

EURISTIDE : Il y a plusieurs façons de s'en apercevoir. D'abord, comme nous l'avons vu, la métrique s'écrit :

$$ds^2 = d\theta^2 + \sin^2 \theta d\phi^2. \quad (595)$$

Les coefficients de la métrique ne sont pas constants, et cette situation est caractéristique d'un espace curviligne.

BEATRIX : En d'autres termes, le calcul de la longueur infinitésimale dépend de la position sur la sphère. Au niveau de l'équateur ou le long d'un parallèle, nous avons  $\phi = \text{Constante}$ , donc  $d\phi = 0$  et par conséquent, l'élément de longueur infinitésimal est :

$$ds^2 = d\theta^2. \quad (596)$$

La mesure de cet élément de longueur ne dépend donc pas de sa position sur l'équateur.

En revanche, si  $\phi$  n'est pas constante, c'est-à-dire si on ne se déplace pas le long d'un parallèle, l'élément de longueur va dépendre de la position du point sur lequel on le mesure.

EURISTIDE : Intuitivement, cette propriété vient du fait que les méridiens se rejoignent au pôle. Par conséquent, les axes de coordonnées curviligne se "rapprochent" lorsque l'angle  $\phi$  augmente, un peu comme si l'espace se rétrécissait vers les pôles. C'est cette déformation qui provoque la variation du calcul de l'élément de ligne en fonction de l'angle  $\phi$ , et par suite, la caractéristique curviligne de l'espace qui se traduit dans le tenseur métrique.

MATHINE : Comme nous l'avons vu dans les discussions précédentes, cette caractéristique est "codée" dans le tenseur métrique. Elle peut être généralisée en dimension  $n$  quelconque.

## 74 Scène XI-2 : La dérivée covariante

EURISTIDE : Pour construire un modèle mathématique de la courbure, nous allons avoir besoin de la dérivée. mais, comme on l'a vu à plusieurs reprises, nous souhaitons avoir des expressions mathématiques tensorielles, ce qui permet d'écrire des expressions simples pour modéliser ce que nous voulons. Il se trouve que dans un espace plat, la dérivée partielle d'un vecteur ou d'un tenseur par rapport à une des des coordonnées :

$$\frac{\partial T^{\mu\nu}}{\partial x^\mu} = \partial_\mu T^{\mu\nu}, \quad (597)$$

est bien un tenseur.

MATHINE : En effet, d'après la règle de la chaîne pour les dérivées partielles et la loi de transformation tensorielle dans un changement de coordonnées, nous avons :

$$\begin{aligned} \frac{\partial T'^{ij}}{\partial x'^k} &= \frac{\partial T'^{ij}}{\partial x^\gamma} \frac{\partial x^\gamma}{\partial x'^k} \\ &= \frac{\partial T^{\alpha\beta}}{\partial x^\gamma} \frac{\partial x'^i}{\partial x^\alpha} \frac{\partial x'^j}{\partial x^\beta} \frac{\partial x^\gamma}{\partial x'^k}. \end{aligned} \quad (598)$$

Cette dernière expression signifie que :

$$\frac{\partial T^{\alpha\beta}}{\partial x^\gamma} \quad (599)$$

peut bien être considéré comme un tenseur deux fois contravariant et une fois covariant.

EURISTIDE : Mais il n'en va pas de même si l'espace est curviligne. En effet, considérons la différentielle d'un vecteur  $\vec{a}$  de coordonnées  $(a^1, a^2, \dots, a^n)$ .

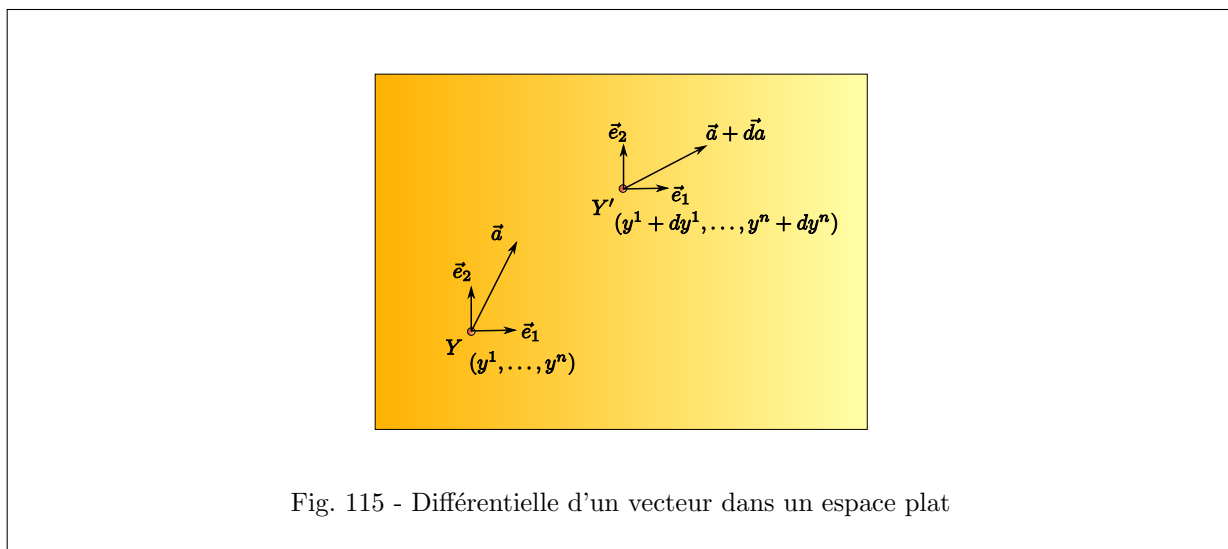


Fig. 115 - Différentielle d'un vecteur dans un espace plat

Nous pouvons considérer le vecteur  $\vec{a}$  exprimé au point  $Y$  de coordonnées  $(y^1, \dots, y^n)$ , et le vecteur  $\vec{a} + \vec{da}$  exprimé au point  $Y'$  de coordonnées  $(y^1 + dy^1, \dots, y^n + dy^n)$ .

Dans un espace plat, le système de coordonnées en  $Y'$  est le même qu'en  $Y$ . Donc, on peut exprimer le vecteur  $\vec{a} + \vec{da}$  au point  $Y'$  en fonction de ses coordonnées et des vecteurs de coordonnées  $(\vec{e}_1, \dots, \vec{e}_n)$ , par :

$$\vec{a} + \vec{da} = a^i \vec{e}_i + da^i \vec{e}_i. \quad (600)$$

D'où l'expression de  $\vec{da}$  :

$$\vec{da} = da^i \vec{e}_i. \quad (601)$$

Dans un espace curviligne, il se trouve que les fonctions-vecteurs de coordonnées ne sont pas les mêmes en  $Y$  et en  $Y'$ . C'est ce qui caractérise justement l'espace curviligne :

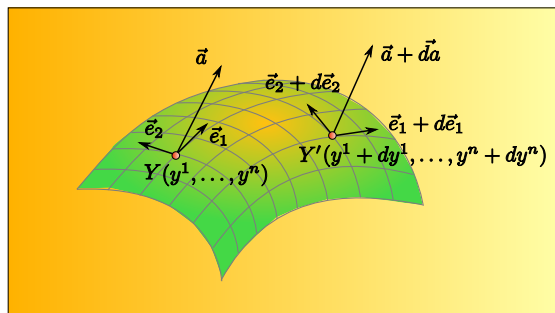


Fig. 116 - Différentielle d'un vecteur dans un espace curviligne

Nous sommes donc obligés de faire intervenir la dérivée des fonctions-vecteurs de coordonnées de base :

$$\vec{a} + \vec{da} = a^i \vec{e}_i + da^i \vec{e}_i + a^i \vec{de}_i. \quad (602)$$

D'où :

$$\vec{da} = da^i \vec{e}_i + a^i \vec{de}_i. \quad (603)$$

Pour pouvoir établir une expression des coordonnées du vecteur  $\vec{da}$  en fonction des fonctions-vecteurs de base, nous devons calculer les  $\vec{de}_i$  en fonction des  $\vec{e}_i$ .

**MATHINE** : Considérons donc l'expression des coordonnées  $\omega_i^j$  de  $\vec{de}_i$  exprimée dans le système de coordonnées  $\vec{e}_j$  :

$$\vec{de}_i = \omega_i^j \vec{e}_j. \quad (604)$$

Puisque nous travaillons avec des grandeurs infiniment petites, nous pouvons considérer que les grandeurs  $\omega_i^j$  sont des fonctions linéaires des coordonnées du point  $Y$ , donc exprimées elles-mêmes en fonction des  $dy^k$ , sous la forme :

$$\omega_i^j = \Gamma_{ki}^j dy^k. \quad (605)$$

Alors, on peut écrire :

$$\vec{da} = da^j \vec{e}_j + a^i \Gamma_{ki}^j dy^k \vec{e}_j. \quad (606)$$

Soit :

$$\vec{da} = (da^j + \Gamma_{ki}^j a^i dy^k) \vec{e}_j. \quad (607)$$

**EURISTIDE** : Ce sont les quantités supplémentaires  $\Gamma_{ki}^j a^k dy^k$  dans cette expression qui caractérisent la nature curviligne de l'espace. Par conséquent, la dérivée absolue de la composante  $a^j$  du vecteur  $\vec{a}$  dans un espace curviligne, notée  $\nabla a^j$  et telle que :

$$\vec{da} = \nabla a^j \vec{e}_j, \quad (608)$$

s'écrit :

$$\begin{aligned}\nabla a^j &= da^j + \Gamma_{ki}^j a^i dy^k \\ &= \frac{\partial a^j}{\partial y^k} dy^k + \Gamma_{ki}^j a^i dy^k.\end{aligned}\quad (609)$$

Soit :

$$\nabla a^j = \left( \frac{\partial a^j}{\partial y^k} + \Gamma_{ki}^j a^i \right) dy^k. \quad (610)$$

BEATRIX : Autrement dit, dans un espace plat, on a :

$$\vec{da} = da^j \vec{e}_j = \frac{\partial a^j}{\partial y^k} dy^k \vec{e}_j, \quad (611)$$

et dans un espace curviligne, on a :

$$\boxed{\vec{da} = \nabla a^j \vec{e}_j = \left( \frac{\partial a^j}{\partial y^k} + \Gamma_{ki}^j a^i \right) dy^k \vec{e}_j.} \quad (612)$$

Donc, dans un espace curviligne, nous avons remplacé la dérivée partielle par l'expression :

$$\frac{\partial a^j}{\partial y^k} + \Gamma_{ki}^j a^i. \quad (613)$$

La nature curviligne de l'espace se traduit par l'ajout d'une composante :

$$\Gamma_{ki}^j a^i, \quad (614)$$

dans l'expression de la dérivée partielle.

MATHINE : Regardons maintenant l'expression des dérivées partielles par rapport aux coordonnées. Dans un espace plat, nous avons l'expression suivante pour la dérivée d'un vecteur :

$$da^j = \frac{\partial a^j}{\partial y^k} dy^k. \quad (615)$$

Dans un espace curviligne, nous avons dû définir une nouvelle dérivée absolue :

$$\nabla a^j = \left( \frac{\partial a^j}{\partial y^k} + \Gamma_{ki}^j a^i \right) dy^k. \quad (616)$$

Par similitude avec l'expression précédente d'un espace plat, nous pouvons considérer que l'expression :

$$\frac{\partial a^j}{\partial y^k} + \Gamma_{ki}^j a^i, \quad (617)$$

est une sorte de dérivée partielle par rapport aux coordonnées, adaptée à un espace curviligne.

On appelle cette expression **la dérivée covariante** des composantes  $a^j$  du vecteur  $\vec{a}$  par rapport à la coordonnée  $y^k$ , et on la note :

$$\boxed{\nabla_k a^j = \frac{\partial a^j}{\partial y^k} + \Gamma_{ki}^j a^i.} \quad (618)$$

On obtient alors une nouvelle expression de la dérivée absolue d'un vecteur dans un espace curviligne :

$$\nabla a^j = \nabla_k a^j dy^k. \quad (619)$$

BEATRIX : Et cette expression permet d'obtenir une expression de la dérivée d'un vecteur dans un espace curviligne qui ressemble comme deux gouttes d'eau à celle de la dérivée d'un vecteur dans un espace plat.

En effet, dans un espace plat, on a :

$$\vec{da} = \partial_k a^j dy^k \vec{e}_j. \quad (620)$$

Et dans un espace curviligne :

$$\vec{da} = \nabla_k a^j dy^k \vec{e}_j. \quad (621)$$

MATHINE : Et voilà le travail! Et pour achever le tableau, nous voyons que l'expression  $\nabla_k a^j$  est un tenseur mixte une fois covariant et une fois contravariant, puisqu'il se transforme suivant la loi de transformation d'un tenseur dans un changement de coordonnées. En effet :

$$\begin{aligned} \vec{da} &= \nabla_k a^j dy^k \vec{e}_j \\ &= \nabla_\gamma a'^{\beta} dy'^{\gamma} \vec{e}'_{\beta} \\ &= \nabla_\gamma a'^{\beta} \frac{\partial y'^{\gamma}}{\partial y^k} dy^k \vec{e}'_{\beta}. \end{aligned} \quad (622)$$

Et par conséquent, puisque nous savons que :

$$\vec{e}_j = \frac{\partial y'^{\beta}}{\partial y^j} \vec{e}'_{\beta}, \quad (623)$$

en identifiant les coordonnées de  $\vec{da}$  dans les deux versions de son expression, on obtient :

$$\nabla_k a^j = \nabla_\gamma a'^{\beta} \frac{\partial y'^{\gamma}}{\partial y^k} \frac{\partial y^j}{\partial y'^{\beta}}. \quad (624)$$

BEATRIX : Nous sommes donc satisfaits : nous avons trouvé une expression tensorielle, similaire à celle qu'on trouve dans un espace plat, qui représente les coordonnées de la dérivée partielle d'un vecteur par rapport aux coordonnées dans un espace curviligne.

Cette notion constitue donc une généralisation de la notion de dérivée partielle pour des coordonnées curvilignes, je suppose.

EURISTIDE : Tout à fait. Les coefficients  $\Gamma_{ki}^j$  portent le nom de **symboles de Christoffel**. Ils constituent une caractérisation de la courbure de l'espace. Lorsque l'espace est plat, ils sont tous nuls.

BEATRIX : Nous avons effectué des calculs un peu compliqués pour obtenir ce résultat, mais je n'ai pas encore tout à fait compris ce que nous avons fait en réalité. Est-ce possible d'expliquer cela de façon plus illustrée et géométrique ?

EURISTIDE : Et bien, on va essayer... Tu vas voir, cela devrait devenir plus clair.

Considérons une surface curviligne de dimension 2, sur laquelle nous allons regarder la dérivée du vecteur  $\vec{a}$ .

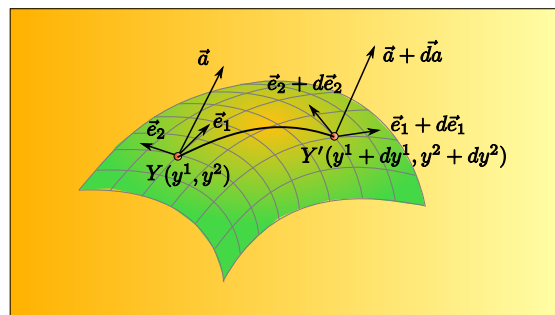


Fig. 117 - Différentielle d'un vecteur a

Lorsque nous considérons la variation  $\vec{a} + d\vec{a}$  du vecteur  $\vec{a}$ , nous faisons varier les coordonnées de son origine  $(y^1, y^2)$  en  $(y^1 + dy^1, y^2 + dy^2)$ .

Si nous étions dans un espace plat, nous aurions un moyen direct de comparer le vecteur  $\vec{a}$  et le vecteur  $\vec{a} + d\vec{a}$  (et par conséquent de calculer sa dérivée), parce que nous aurions en  $(y^1 + dy^1, y^2 + dy^2)$  un même système de vecteurs de coordonnées qu'en  $(y^1, y^2)$ .

Mais nous sommes dans un espace curviligne, et le système de coordonnées en  $(y^1 + dy^1, y^2 + dy^2)$  n'est pas le même qu'en  $(y^1, y^2)$  à cause de la courbure de la surface. Par conséquent, si nous comparons directement les vecteurs  $\vec{a} + d\vec{a}$  et  $\vec{a}$  dans leurs systèmes de coordonnées respectifs, nous oublions de prendre en compte la variation du système de coordonnées entre les points  $(y^1, y^2)$  et  $(y^1 + dy^1, y^2 + dy^2)$ .

BEATRIX : J'ai compris ! Il faut donc qu'on ramène le vecteur  $\vec{a} + d\vec{a}$  dans le système de coordonnées  $(y^1, y^2)$  avant d'effectuer la comparaison.

EURISTIDE : C'est exactement ce que nous avons fait en exprimant, au moyen des symboles de Christoffel  $\Gamma_{ki}^j$ , les coordonnées du vecteur  $\vec{a} + d\vec{a}$  dans le système de coordonnées qui se trouve en  $(y^1, y^2)$ . Ca ne te rappelle pas quelque chose ?

BEATRIX : Attendez... Si, bien sûr ! Nous avons fait un pullback du vecteur  $\vec{a} + d\vec{a}$  en  $(y^1, y^2)$  par la fonction  $\phi$ , qui à un point  $(y^1, y^2)$  fait correspondre le point infiniment voisin  $(y^1 + dy^1, y^2 + dy^2)$ .

**EURISTIDE** : C'est ce pullback qui nous permet de tirer en arrière le vecteur  $\vec{a} + \vec{da}$ , et de le mettre à côté du vecteur  $\vec{a}$  dans son système de coordonnées local, et par conséquent de pouvoir les comparer directement, puisqu'ils sont alors exprimés dans le même système de coordonnées. Ce sont les symboles de Christoffel qui permettent de réaliser ce pullback.

**MATHINE** : Nous avons vu la dérivée covariante des composantes d'un vecteur, qui sont des quantités contravariantes. Voyons ce qu'est la dérivée covariante d'une quantité covariante.

En utilisant le fait que la dérivée covariante d'une quantité non vectorielle est nulle, et que la somme des produits des composantes d'un vecteur est un nombre, on déduit l'expression de la dérivée covariante d'une quantité covariante. En effet :

$$a^i a_i = \text{Cte}, \quad (625)$$

donc :

$$\nabla_k (a^i a_i) = 0, \quad (626)$$

d'où :

$$(\nabla_k a^i) a_i + (\nabla_k a_i) a^i = 0, \quad (627)$$

donc :

$$(\nabla_k a_i) a^i = - (\nabla_k a^i) a_i = - \left( \frac{\partial a^i}{\partial y^k} + \Gamma_{kl}^i a^l \right) a_i. \quad (628)$$

Or :

$$\frac{\partial (a^i a_i)}{\partial y^k} = 0, \quad (629)$$

d'où :

$$\frac{\partial a^i}{\partial y^k} a_i + \frac{\partial a_i}{\partial y^k} a^i = 0, \quad (630)$$

soit :

$$\frac{\partial a^i}{\partial y^k} a_i = - \frac{\partial a_i}{\partial y^k} a^i. \quad (631)$$

En reportant dans l'expression de  $(\nabla_k a_i) a^i$ , on obtient :

$$(\nabla_k a_i) a^i = - \left( - \frac{\partial a_i}{\partial y^k} a^i + \Gamma_{kl}^i a^l a_i \right). \quad (632)$$

Et en changeant les noms des index du deuxième terme, on obtient :

$$(\nabla_k a_i) a^i = \frac{\partial a_i}{\partial y^k} a^i - \Gamma_{ki}^l a^l a_i. \quad (633)$$

D'où, finalement :

$$\nabla_k a_i = \frac{\partial a_i}{\partial y^k} - \Gamma_{ki}^l a_l. \quad (634)$$

On peut alors généraliser pour un tenseur quelconque, par exemple un tenseur une fois covariant et une fois contravariant :

$$\nabla_k T_i^j = \frac{\partial T_i^j}{\partial y^k} - \Gamma_{ki}^l T_l^j + \Gamma_{kl}^j T_i^l. \quad (635)$$

**BEATRIX** : Autrement dit, pour les indices covariants, on applique le symbole de Christoffel avec un signe moins, et pour les indices contravariants, on applique le signe plus.



MATHINE : Nous avons vu que les symboles de Christoffel représentaient la caractéristique de la courbure de l'espace. Ils sont donc des fonctions du tenseur métrique  $g_{ik}$ , dans la mesure où nous avons également vu que toutes les caractéristiques d'un espace curviligne étaient "codées" dans son tenseur métrique.

Nous allons maintenant faire le choix que notre espace curviligne possède la particularité que ses symboles de Christoffel sont **compatibles avec la métrique**. Cela signifie que la dérivée covariante de la métrique est nulle.

EURISTIDE : Ou plus intuitivement, cela revient à dire que le tenseur métrique est constant sur tout l'espace curviligne, à condition de convenir que cette propriété d'être constant est considérée en neutralisant les effets de la courbure de l'espace au moyen de la dérivée covariante. Nous considérerons dorénavant uniquement des espaces curvilignes (ou des variétés) dont les symboles de Christoffel sont compatibles avec la métrique.

MATHINE : Dans ce cas, nous pouvons écrire :

$$\nabla_k g_{ij} = 0. \quad (636)$$

Cette expression porte le nom de **théorème de Ricci**.

Pour retrouver l'expression complète des symboles de Christoffel (compatibles avec la métrique, bien sûr) en fonction de la métrique, nous allons exprimer cette égalité en utilisant le développement de l'expression de la dérivée covariante, et en faisant varier de 3 façons différentes les indices :

$$\begin{aligned} \nabla_k g_{ij} &= \partial_k g_{ij} - \Gamma_{ki}^l g_{lj} - \Gamma_{kj}^l g_{il} = 0 \\ \nabla_i g_{jk} &= \partial_i g_{jk} - \Gamma_{ij}^l g_{lk} - \Gamma_{ik}^l g_{jl} = 0 \\ \nabla_j g_{ki} &= \partial_j g_{ki} - \Gamma_{jk}^l g_{li} - \Gamma_{ji}^l g_{kl} = 0. \end{aligned} \quad (637)$$

Nous allons également faire l'hypothèse que nous sommes dans un espace curviligne tel que les symboles de Christoffel sont symétriques suivant les indices bas, c'est-à-dire tel que :

$$\Gamma_{ki}^j = \Gamma_{ik}^j. \quad (638)$$

EURISTIDE : Cela correspond à un espace exempt de **torsion**. Dans un espace ayant une torsion non nulle, il se trouve que lorsqu'on considère 2 vecteurs  $\vec{u}$  et  $\vec{v}$ , le vecteur  $\vec{v}$  étant positionné à l'extrémité du vecteur  $\vec{u}$ , si on tire en arrière le vecteur  $\vec{v}$  pour le ramener au point origine du vecteur  $\vec{u}$  et qu'on en fait la somme avec le vecteur  $\vec{u}$ , on n'obtient pas le même résultat que si on positionne le vecteur  $\vec{u}$  à l'extrémité du vecteur  $\vec{v}$ , et qu'on tire en arrière le vecteur  $\vec{u}$  pour le ramener au point origine du vecteur  $\vec{v}$  et qu'on en fait la somme avec le vecteur  $\vec{v}$ . En d'autres termes, le parallélogramme formé par les deux vecteurs  $\vec{u}$  et  $\vec{v}$  n'est pas conservé tel quel lorsque les vecteurs sont tirés en arrière dans un espace possédant une torsion. Nous considérerons dorénavant uniquement des espaces exempts de torsion.

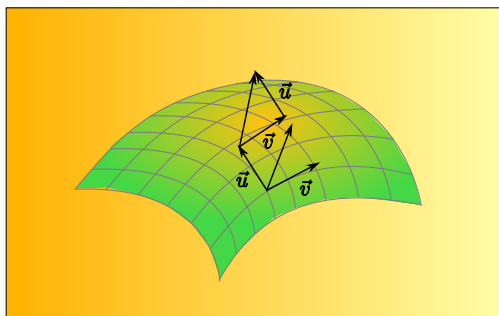


Fig. 118 - Espace curviligne avec torsion

MATHINE : Dans ces conditions, en soustrayant la seconde égalité et la troisième égalité de la première dans le jeu des trois équations ci-dessus, nous obtenons :

$$\nabla_k g_{ij} - \nabla_i g_{jk} - \nabla_j g_{ki} = 0. \quad (639)$$

D'où :

$$\begin{aligned} \partial_k g_{ij} - \Gamma_{ki}^l g_{lj} - \Gamma_{kj}^l g_{il} - \partial_i g_{jk} + \Gamma_{ij}^l g_{lk} + \Gamma_{ik}^l g_{jl} \\ - \partial_j g_{ki} + \Gamma_{jk}^l g_{li} + \Gamma_{ji}^l g_{kl} = 0. \end{aligned} \quad (640)$$

Soit :

$$\partial_k g_{ij} - \partial_i g_{jk} - \partial_j g_{ki} + 2\Gamma_{ij}^l g_{lk} = 0, \quad (641)$$

ce qui s'écrit :

$$\Gamma_{ij}^l g_{lk} g^{lk} = \frac{1}{2} g^{lk} (\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij}), \quad (642)$$

ou encore :

$$\Gamma_{ij}^l = \frac{1}{2} g^{lk} \left( \frac{\partial g_{jk}}{\partial y^i} + \frac{\partial g_{ki}}{\partial y^j} - \frac{\partial g_{ij}}{\partial y^k} \right). \quad (643)$$

En inversant les deux indices  $l$  et  $k$ , on obtient :

$$\Gamma_{ij}^k = \frac{1}{2} g^{kl} \left( \frac{\partial g_{jl}}{\partial y^i} + \frac{\partial g_{li}}{\partial y^j} - \frac{\partial g_{ij}}{\partial y^l} \right). \quad (644)$$

Et en faisant une rotation des trois indices ( $i$  vers  $k$ ,  $k$  vers  $j$  et  $j$  vers  $i$ ), on obtient finalement l'expression du symbole de Christoffel compatible avec la métrique, et pour un espace curviligne exempt de torsion :

$$\Gamma_{ki}^j = \frac{1}{2} g^{jl} \left( \frac{\partial g_{il}}{\partial y^k} + \frac{\partial g_{lk}}{\partial y^i} - \frac{\partial g_{ki}}{\partial y^l} \right). \quad (645)$$

## 75 Scène XI-3 : Le transport parallèle

EURISTIDE : Nous avons vu tout à l'heure que pour calculer la dérivée d'un vecteur  $\vec{a}$ , nous avons dû tirer le vecteur  $\vec{a} + d\vec{a}$  en arrière vers l'emplacement du vecteur  $\vec{a}$ . Cette notion de déplacement d'un vecteur va nous permettre d'introduire la notion de transport parallèle.

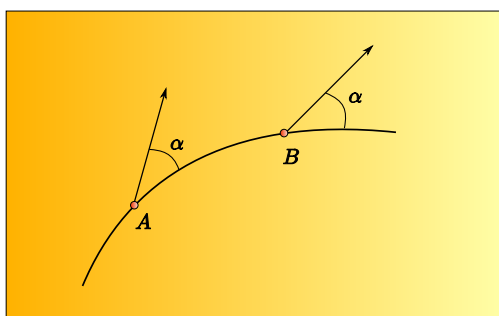


Fig. 119 - Transport parallèle le long d'une courbe

On dit qu'un vecteur est transporté parallèlement sur un arc de courbe, s'il passe d'un point  $A$  à un point  $B$  sans que l'angle qu'il fait avec la tangente à la courbe ne change durant ce trajet.

BEATRIX : Un tel transport parallèle entre deux points fixés effectué le long d'une courbe donne-t-il toujours le même résultat, quelle que soit la courbe utilisée ?

EURISTIDE : Non justement. Et c'est une des caractéristiques d'un espace curviligne, par rapport à un espace plat. Dans un espace curviligne, le résultat du transport parallèle d'un point à un autre dépend du chemin pris entre ces points.

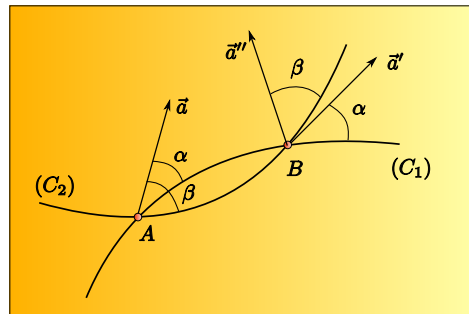


Fig. 120 - Transport parallèle suivant 2 courbes

Dans la figure ci-dessus, lorsqu'on déplace le vecteur  $\vec{a}$  parallèlement le long de la courbe  $(C_1)$ , on obtient le vecteur  $\vec{a}'$ . Lorsqu'on déplace le même vecteur  $\vec{a}$  le long de la courbe  $(C_2)$ , on obtient le vecteur  $\vec{a}''$  qui est différent de  $\vec{a}'$ .

Prenons un autre exemple : considérons une sphère et un vecteur  $\vec{a}$  situé à l'équateur et de direction parallèle au méridien au point d'origine où il se trouve.

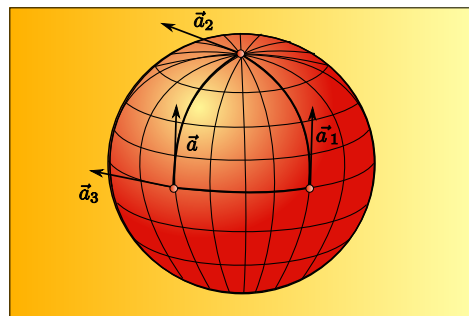


Fig. 121 - Transport parallèle sur une sphère

Faisons tourner ce vecteur d'un angle  $\theta$  autour de l'équateur en le transportant parallèlement à l'équateur. On obtient un vecteur  $\vec{a}_1$ . Transportons maintenant ce vecteur  $\vec{a}_1$  le long du méridien jusqu'au pôle. On obtient un vecteur  $\vec{a}_2$ . Transportons enfin ce vecteur  $\vec{a}_2$  parallèlement au méridien en direction du point d'origine du vecteur  $\vec{a}$  sur l'équateur. Nous obtenons un vecteur  $\vec{a}_3$  qui est différent du vecteur  $\vec{a}$ , et qui a en fait tourné d'un angle  $\theta$  par rapport au vecteur  $\vec{a}$ .

Cette propriété qui fait qu'un vecteur transporté parallèlement le long d'une courbe fermée ne retrouve pas sa direction d'origine est une caractéristique d'un espace curviligne.

BEATRIX : Je connais cette histoire sous forme d'une devinette. Un chasseur est situé en un point de la Terre. Il marche 1 km vers le Sud, puis 1 km vers l'Est, puis 1 km vers le Nord. A ce moment, il s'aperçoit qu'il est revenu à son point de départ, après avoir tourné deux fois vers la gauche à angle droit. Il aperçoit un ours, près de lui, et l'abat avant d'être dévoré. Quelle est la couleur de l'ours ?

Quand on analyse la situation décrite par le périple du chasseur, on s'aperçoit que le seul endroit sur Terre où il est possible de décrire un triangle dont les 3 angles sont droits et dont deux côtés sont dirigés vers le Sud, est le pôle Nord.

Par conséquent, l'ours est un ours polaire, donc il est de couleur blanche.

EURISTIDE : Une caractéristique de ces espaces curvilignes est aussi que la somme des angles d'un triangle ne fait pas toujours 180 degrés, contrairement à ce que nous avons appris en géométrie plane. En l'occurrence, ici, le triangle décrit par notre chasseur polaire comporte trois angles de 90 degrés, donc ayant pour somme 270 degrés.

MATHINE : Nous avons vu comment transporter parallèlement un vecteur. On peut, de la même façon, transporter parallèlement un tenseur de n'importe quel espace de dimension  $n$ .

EURISTIDE : On peut considérer que transporter parallèlement un vecteur ou un tenseur est la généralisation dans un espace curviligne de la notion de garder constant un vecteur ou un tenseur en le déplaçant le long d'une courbe. Essayons de formaliser cela maintenant. Je laisse la parole à Mathine pour cela.

MATHINE : Considérons une courbe  $x^\mu(\lambda)$  paramétrée par un paramètre  $\lambda$ . Cela signifie que chaque valeur de  $\lambda$  détermine un point de la courbe :

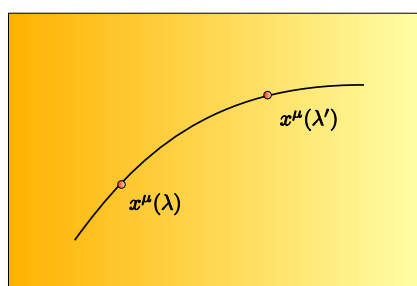


Fig. 122 - Courbe paramétrée

La spécification de la constance d'un tenseur  $T$  dans un espace plat s'écrit :

$$\frac{dT}{d\lambda} = \frac{dx^\mu}{d\lambda} \frac{\partial T}{\partial x^\mu} = 0. \quad (646)$$

Nous avons vu que l'équivalent, dans un espace curviligne, de la dérivée partielle d'un espace plat, est la dérivée covariante. Par conséquent, dire qu'un tenseur  $T$  est constant le long d'une courbe, dans un espace curviligne, c'est dire que :

$$\frac{dx^\mu}{d\lambda} \nabla_\mu T = 0. \quad (647)$$

Cette équation est une équation tensorielle appelée **équation du transport parallèle**. Elle définit les courbes de l'espace curviligne sur lesquelles le tenseur  $T$  est invariant.

EURISTIDE : Revenons sur l'exemple de la sphère  $S^2$  :

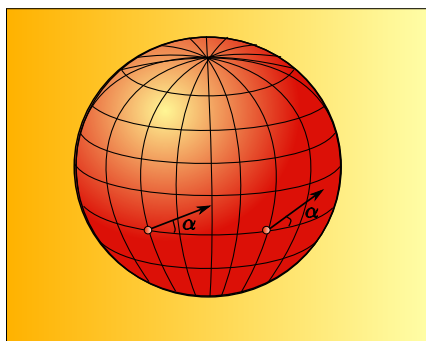


Fig. 123 - Transport parallèle le long de l'équateur

Nous voyons qu'un vecteur transporté parallèlement le long de l'équateur reste constant au sens de l'espace curviligne. C'est-à-dire qu'il garde un angle constant avec la courbe de l'équateur. L'équateur d'une sphère est donc une des courbes qui vérifient l'équation du transport parallèle.

## 76 Scène XI-4 : Les géodésiques

BEATRIX : En revanche, nous avons vu que le chemin constitué des déplacements successifs sur l'équateur, un méridien, puis un autre méridien, pour revenir au même point, comme dans l'histoire de l'ours blanc, ne permettait pas un transport parallèle. Cette courbe ne vérifie donc pas l'équation du transport parallèle.

MATHINE : Nous allons maintenant nous attarder sur une catégorie particulière de courbes qui respectent le transport parallèle. Il s'agit des géodésiques.

EURISTIDE : Les géodésiques sont à considérer comme une généralisation de la ligne droite d'un espace euclidien plat. Une ligne droite est une ligne qui représente le chemin le plus court entre deux points, mais c'est aussi une ligne qui transporte parallèlement son propre vecteur tangent. C'est un peu compliqué comme définition, mais c'est bien celle dont nous allons avoir besoin.

En transposant cette définition dans un espace curviligne, nous allons appeler géodésique une courbe qui transporte parallèlement son propre vecteur tangent.

BEATRIX : Donc un arc de l'équateur ( $C$ ) est une géodésique :

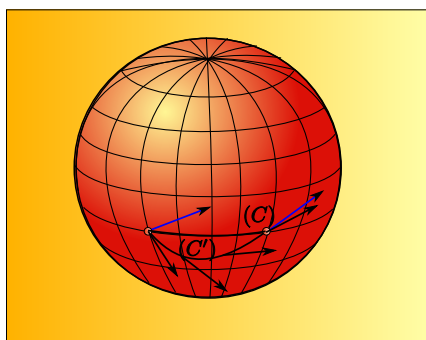


Fig. 124 - Courbe géodésique, courbe non géodésique

EURISTIDE : Mais tous les chemins d'une sphère ne sont pas des géodésiques. Si nous considérons la courbe ( $C'$ ) du schéma ci-dessus, nous constatons que le vecteur tangent n'est pas transporté parallèlement par cette courbe.

MATHINE : Si la courbe est une courbe paramétrée  $x^\mu(\lambda)$ , alors le vecteur tangent à cette courbe a pour coordonnées :

$$\frac{dx^\mu}{d\lambda}. \quad (648)$$

Par conséquent, le transport parallèle de ce vecteur le long de la courbe s'écrit au moyen de l'équation du transport parallèle, sous la forme :

$$\nabla_\rho \frac{dx^\mu}{d\lambda} = 0. \quad (649)$$

Soit :

$$\frac{\partial}{\partial x^\rho} \frac{dx^\mu}{d\lambda} + \Gamma_{\rho\sigma}^\mu \frac{dx^\sigma}{d\lambda} = 0. \quad (650)$$

Donc, en multipliant par  $\frac{dx^\rho}{d\lambda}$ , on obtient :

$$\boxed{\frac{d^2 x^\mu}{d\lambda^2} + \Gamma_{\rho\sigma}^\mu \frac{dx^\sigma}{d\lambda} \frac{dx^\rho}{d\lambda} = 0.} \quad (651)$$

Cette équation est connue sous le nom **d'équation géodésique**. C'est l'équation de l'équivalent des droites dans un espace curviligne.

BEATRIX : D'ailleurs, si les  $\Gamma_{\rho\sigma}^\mu$  sont nuls, on retombe sur l'équation d'une droite :

$$\frac{d^2 x^\mu}{d\lambda^2} = 0. \quad (652)$$

MATHINE : On va vérifier que ces géodésiques correspondent au plus court chemin entre deux points dans un espace curviligne, à l'instar de ce qui se passe dans un espace plat, où les lignes droites correspondent au plus court chemin entre deux points. Ceci achèvera de nous convaincre que les géodésiques jouent le rôle des droites dans un espace curviligne.

Nous avons vu en relativité restreinte que le plus court chemin entre deux points correspondait (paradoxalement !) à la trajectoire pour laquelle le temps propre était maximal.

BEATRIX : On voit donc les techniques variationnelles pointer le bout de leur nez. Comme il faut qu'on détermine la trajectoire pour laquelle le temps propre est maximal, nous allons étudier les variations du temps propre, je suppose.

MATHINE : C'est vrai, Béatrix. L'expression de la variation infinitésimale du temps propre, le long de la trajectoire, est, par invariance de l'intervalle, au sens relativiste :

$$\begin{aligned} ds^2 &= c^2 d\tau^2 \\ ds^2 &= g_{\mu\nu} dx^\mu dx^\nu. \end{aligned} \quad (653)$$

Nous obtenons l'équation suivante, en considérant un changement d'unités permettant d'exprimer que  $c = 1$  :

$$\frac{d\tau^2}{d\lambda^2} - g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} = 0. \quad (654)$$

Par conséquent, en multipliant par  $d\lambda^2$ , en prenant la racine carrée de l'expression, et en calculant l'intégrale du temps propre sur le chemin parcouru, on obtient :

$$\tau = \int \sqrt{g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} d\lambda. \quad (655)$$

On considère une variation infinitésimale du chemin :

$$\begin{aligned} x^\mu &\rightarrow x^\mu + \delta x^\mu \\ g_{\mu\nu} &\rightarrow g_{\mu\nu} + \delta x^\sigma \partial_\sigma g_{\mu\nu}. \end{aligned}$$



Alors :

$$\begin{aligned}
\tau + \delta\tau &= \int \sqrt{g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} + \partial_\sigma g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} \delta x^\sigma + 2g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{d(\delta x^\nu)}{d\lambda}} d\lambda \\
&= \int \sqrt{g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} \\
&\quad \times \sqrt{1 + \frac{\partial_\sigma g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} \delta x^\sigma + 2g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{d(\delta x^\nu)}{d\lambda}}{g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}}} d\lambda.
\end{aligned} \tag{656}$$

Or,  $\delta x^\sigma$  est infiniment petit, donc le terme ajouté à 1 sous la racine carrée est négligeable devant 1, et par conséquent, on peut écrire son développement de Taylor :

$$\begin{aligned}
\tau + \delta\tau &\approx \int \sqrt{g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} \left( 1 - \frac{1}{2} \frac{\partial_\sigma g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} \delta x^\sigma + 2g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{d(\delta x^\nu)}{d\lambda}}{g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} \right) d\lambda \\
&= \int \sqrt{g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} d\lambda + \int \frac{-\frac{1}{2} \partial_\sigma g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} \delta x^\sigma - g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{d(\delta x^\nu)}{d\lambda}}{\sqrt{g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}}} d\lambda.
\end{aligned} \tag{657}$$

D'où :

$$\delta\tau = \int \frac{-\frac{1}{2} \partial_\sigma g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} \delta x^\sigma - g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{d(\delta x^\nu)}{d\lambda}}{\sqrt{g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}}} d\lambda. \tag{658}$$

Or, de l'expression de  $d\tau^2 = ds^2$ , on déduit que :

$$d\tau = \int \sqrt{g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}} d\lambda, \tag{659}$$

donc, en substituant dans l'expression de  $\delta\tau$ , on obtient :

$$\delta\tau = \int \left( -\frac{1}{2} \partial_\sigma g_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} \delta x^\sigma - g_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{d(\delta x^\nu)}{d\tau} \right) d\tau. \tag{660}$$

Soit, en intégrant par parties :

$$\delta\tau = \int \left( -\frac{1}{2} \partial_\sigma g_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} + \frac{d}{d\tau} \left( g_{\mu\sigma} \frac{dx^\mu}{d\tau} \right) \right) \delta x^\sigma d\tau. \tag{661}$$

On cherche les chemins pour lesquels  $\delta\tau$  est nulle, donc cela implique que pour toute variation  $\delta x^\sigma$ , on ait :

$$-\frac{1}{2} \partial_\sigma g_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} + \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} \partial_\nu g_{\mu\sigma} + g_{\mu\sigma} \frac{d^2 x^\mu}{d\tau^2} = 0, \tag{662}$$

puisque :

$$\frac{dg_{\mu\sigma}}{d\tau} = \frac{dx^\nu}{d\tau} \partial_\nu g_{\mu\sigma}. \tag{663}$$

On obtient donc finalement, en remaniant les indices :

$$g_{\mu\sigma} \frac{d^2 x^\mu}{d\tau^2} + \frac{1}{2} (-\partial_\sigma g_{\mu\nu} + \partial_\nu g_{\mu\sigma} + \partial_\mu g_{\nu\sigma}) \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = 0. \tag{664}$$

Ce qui s'écrit finalement :

$$\frac{d^2 x^\mu}{d\tau^2} + \Gamma_{\mu\nu}^\sigma \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} = 0. \tag{665}$$

EURISTIDE : On retrouve bien l'équation géodésique que nous avons déduite tout à l'heure. Ceci achève de nous convaincre que les géodésiques sont à la fois les plus courts chemins d'un point à un autre sur un espace curviligne, et les courbes qui transportent parallèlement leurs tangentes d'un point à un autre.

BEATRIX : Si on prend l'exemple de la sphère (toujours elle!), la courbe géodésique joignant  $A$  et  $B$  est un arc de grand cercle passant par les deux points :

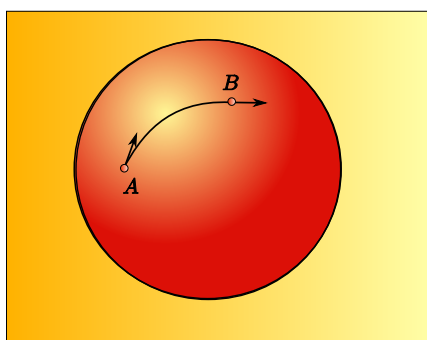


Fig. 125 - Géodésique d'une sphère : arc de grand cercle

Les méridiens et l'équateur sont des exemples de courbes géodésiques, mais d'une manière plus générale, tout arc situé sur un grand cercle de la sphère est une géodésique. Pour rappel, un grand cercle sur une sphère est un cercle à la surface de la sphère ayant pour rayon le rayon de la sphère ; son centre est également le centre de la sphère.

## 77 Scène XI-5 : Le tenseur de Riemann

EURISTIDE : Nous allons maintenant avoir besoin d'un outil pour représenter explicitement et mathématiquement la courbure de l'espace. Lorsque l'espace est plat, le transport parallèle d'un vecteur le long d'une courbe fermée résulte en un vecteur identique au vecteur d'origine :

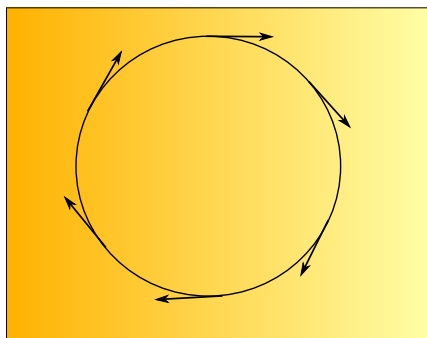


Fig. 126 - Transport parallèle le long d'une courbe plate fermée

En revanche, comme nous l'avons vu, lorsque nous considérons un espace curviligne, le transport parallèle d'un vecteur  $\vec{u}$  le long d'une courbe fermée va résulter en un vecteur  $\vec{u}_3$  en général différent du vecteur d'origine.

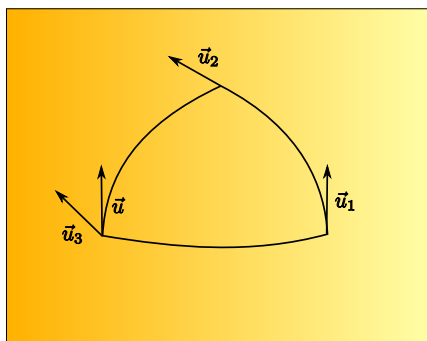


Fig. 127 - Transport parallèle sur une courbe curviligne fermée

Pour décrire la courbure, nous allons donc avoir besoin d'un tenseur qui avale le vecteur  $\vec{u}$  et recrache le vecteur  $\vec{u}_3$ . Mais ce tenseur devra également prendre en compte en entrée, la forme de la courbe fermée utilisée pour le transport parallèle, puisque nous savons que la modification apportée au vecteur dans son transport parallèle dépend de la forme de la courbe fermée.

BEATRIX : Alors, comment allons-nous représenter une courbe fermée avec des vecteurs ?

EURISTIDE : C'est simple. Nous allons considérer deux vecteurs infiniment petits  $\vec{a}$  et  $\vec{b}$ , et considérer la courbe fermée suivante, constituée du parallélogramme issu de ces deux vecteurs :

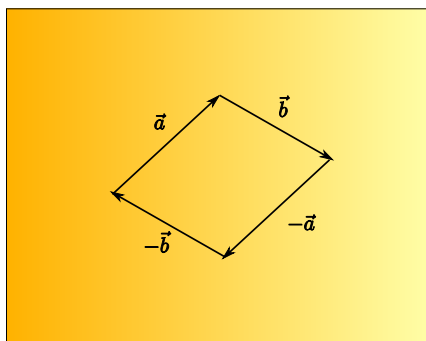


Fig. 128 - Modèle de la courbe fermée au moyen de 2 vecteurs

La courbe fermée sera donc caractérisée par deux vecteurs  $\vec{a}$  et  $\vec{b}$ . Par conséquent, le tenseur caractérisant la courbe de l'espace sera un tenseur qui mange trois vecteurs  $\vec{u}$ ,  $\vec{a}$  et  $\vec{b}$  en entrée, et recrache un vecteur  $\vec{u}_3$ . Ce sera donc un tenseur 3 fois covariant et 1 fois contravariant. Ce tenseur s'appelle le **tenseur de Riemann** et est noté :

$$R^{\rho}_{\sigma\mu\nu}. \quad (666)$$

MATHINE : Pour calculer les composantes de ce tenseur de Riemann, nous allons calculer la variation du vecteur  $\vec{u}$  lorsqu'il est transporté parallèlement le long du vecteur  $\vec{a}$ , puis le long du vecteur  $\vec{b}$  ; et nous allons comparer cette variation à celle du vecteur  $\vec{u}$  lorsqu'il est transporté parallèlement le long du vecteur  $\vec{b}$ , puis le long du vecteur  $\vec{a}$ .

Cette notion de transport par les vecteurs  $\vec{a}$  et  $\vec{b}$  a un sens dans l'espace curviligne parce que nous choisissons deux vecteurs infiniment petits, et que par conséquent l'espace autour de l'origine du vecteur  $\vec{u}$  peut être assimilé à un espace plat, conformément à ce que nous avons vu dans le cadre des variétés différentiables.

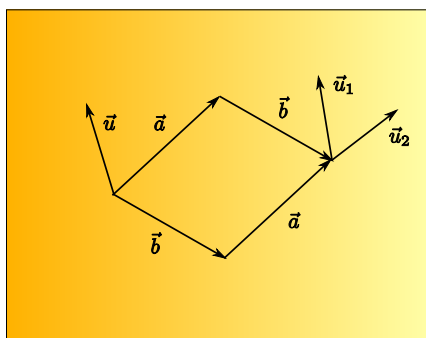


Fig. 129 - Transport du vecteur u

Si nous considérons maintenant que les vecteurs  $\vec{a}$  et  $\vec{b}$  sont assimilés aux dérivées partielles suivant les coordonnées respectivement par rapport à  $x^\mu$  et  $x^\nu$ , c'est-à-dire, en assimilant comme nous l'avons vu précédemment en étudiant les variétés différentiables, un opérateur dérivé à un vecteur :

$$\vec{a} = \frac{\partial}{\partial x^\mu}, \quad (667)$$

et :

$$\vec{b} = \frac{\partial}{\partial x^\nu}, \quad (668)$$

alors, nous pouvons dire que le déplacement de  $\vec{u}$  le long de  $\vec{a}$  est égal à  $\vec{u}$  plus la dérivée covariante de  $\vec{u}$  par rapport à  $x^\mu$  et le déplacement de  $\vec{u}$  le long de  $\vec{b}$  est égal à  $\vec{u}$  plus la dérivée covariante de  $\vec{u}$  par rapport à  $x^\nu$ .

Donc, la différence  $\vec{u}_1 - \vec{u}_2$  recherchée est en fait :

$$\nabla_\mu \nabla_\nu \vec{u} - \nabla_\nu \nabla_\mu \vec{u}, \quad (669)$$

soit, en s'intéressant aux coordonnées du vecteur  $\vec{u}$  :

$$\nabla_\mu \nabla_\nu u^\rho - \nabla_\nu \nabla_\mu u^\rho. \quad (670)$$

En utilisant l'expression de la dérivée covariante :

$$\nabla_\nu u^\rho = \frac{\partial u^\rho}{\partial x^\nu} + \Gamma_{\nu\lambda}^\rho u^\lambda, \quad (671)$$

on trouve :

$$\begin{aligned} \nabla_\mu \nabla_\nu u^\rho - \nabla_\nu \nabla_\mu u^\rho &= \nabla_\mu \left( \frac{\partial u^\rho}{\partial x^\nu} + \Gamma_{\nu\lambda}^\rho u^\lambda \right) - \nabla_\nu \left( \frac{\partial u^\rho}{\partial x^\mu} + \Gamma_{\mu\lambda}^\rho u^\lambda \right) \\ &= \left( \partial_\mu \Gamma_{\nu\sigma}^\rho - \partial_\nu \Gamma_{\mu\sigma}^\rho + \Gamma_{\mu\lambda}^\rho \Gamma_{\nu\sigma}^\lambda - \Gamma_{\nu\lambda}^\rho \Gamma_{\mu\sigma}^\lambda \right) u^\sigma. \end{aligned} \quad (672)$$

On obtient cette expression, en utilisant le fait que les symboles de Christoffel sont symétriques par rapport à leurs indices bas.

Et par conséquent, le tenseur de Riemann, qui est tel que :

$$\nabla_\mu \nabla_\nu u^\rho - \nabla_\nu \nabla_\mu u^\rho = R_{\sigma\mu\nu}^\rho u^\sigma, \quad (673)$$

est identifié à :

$$R_{\sigma\mu\nu}^{\rho} = \partial_{\mu}\Gamma_{\nu\sigma}^{\rho} - \partial_{\nu}\Gamma_{\mu\sigma}^{\rho} + \Gamma_{\mu\lambda}^{\rho}\Gamma_{\nu\sigma}^{\lambda} - \Gamma_{\nu\lambda}^{\rho}\Gamma_{\mu\sigma}^{\lambda}. \quad (674)$$

BEATRIX : Nous avons donc, contenue dans cette expression, toute la caractéristique de la courbure de l'espace qui fait tourner un vecteur lorsqu'il est transporté parallèlement le long d'une courbe fermée.

MATHINE : On peut définir, en utilisant la descente d'indice au moyen du tenseur métrique, un tenseur de Riemann avec tous les indices bas :

$$R_{\rho\sigma\mu\nu} = g_{\rho\lambda}R_{\sigma\mu\nu}^{\lambda}. \quad (675)$$

EURISTIDE : Nous allons maintenant considérer les coordonnées normales riemaniennes locales en un point de l'espace. Tu te souviens, ce sont les coordonnées qui en chaque point de l'espace permettent d'obtenir une expression de la métrique sous forme de matrice diagonale, dont les dérivées premières des composantes sont nulles (mais pas les dérivées d'ordre supérieur.) Rappelle-toi aussi que ces coordonnées sont les coordonnées cartésiennes de l'espace comme s'il était considéré plat localement.

Nous savons que les  $\Gamma_{ki}^j$  sont exprimés en fonction des dérivées premières des composantes du tenseur métrique :

$$\Gamma_{ki}^j = \frac{1}{2}g^{jl} \left( \frac{\partial g_{il}}{\partial y^k} + \frac{\partial g_{lk}}{\partial y^i} - \frac{\partial g_{ki}}{\partial y^l} \right). \quad (676)$$

MATHINE : Donc, dans un système de coordonnées locales normales riemaniennes, les  $\Gamma_{ki}^j$  qui sont exprimés sans être dérivés par rapport à une composante de coordonnées s'annulent.

Par conséquent, dans un tel système de coordonnées normales riemaniennes, le tenseur de Riemann s'écrit :

$$R_{\rho\sigma\mu\nu} = g_{\rho\lambda}(\partial_{\mu}\Gamma_{\nu\sigma}^{\lambda} - \partial_{\nu}\Gamma_{\mu\sigma}^{\lambda}), \quad (677)$$

d'où :

$$R_{\rho\sigma\mu\nu} = \frac{1}{2}(\partial_{\mu}\partial_{\sigma}g_{\rho\nu} - \partial_{\mu}\partial_{\rho}g_{\nu\sigma} - \partial_{\nu}\partial_{\sigma}g_{\rho\mu} + \partial_{\nu}\partial_{\rho}g_{\mu\sigma}). \quad (678)$$

On voit différentes propriétés de ce tenseur de Riemann : il est antisymétrique sur les indices  $\mu$  et  $\nu$ , c'est-à-dire que si on inverse les indices  $\mu$  et  $\nu$ , il change de signe :

$$R_{\rho\sigma\mu\nu} = -R_{\rho\sigma\nu\mu}. \quad (679)$$

De plus :

$$R_{\rho\sigma\mu\nu} + R_{\rho\mu\nu\sigma} + R_{\rho\nu\sigma\mu} = 0. \quad (680)$$

En outre, considérons la dérivée covariante du tenseur de Riemann, toujours exprimé en coordonnées normales riemaniennes :

$$\begin{aligned} \nabla_{\lambda}R_{\rho\sigma\mu\nu} &= \partial_{\lambda}R_{\rho\sigma\mu\nu} \\ &= \frac{1}{2}\partial_{\lambda}(\partial_{\mu}\partial_{\sigma}g_{\rho\nu} - \partial_{\mu}\partial_{\rho}g_{\nu\sigma} - \partial_{\nu}\partial_{\sigma}g_{\rho\mu} + \partial_{\nu}\partial_{\rho}g_{\mu\sigma}). \end{aligned} \quad (681)$$

Par un calcul simple, mais un peu fastidieux en substituant les différents indices de la formule précédente, on arrive au résultat suivant :

$$\nabla_{\lambda}R_{\rho\sigma\mu\nu} + \nabla_{\mu}R_{\rho\sigma\nu\lambda} + \nabla_{\nu}R_{\rho\sigma\lambda\mu} = 0. \quad (682)$$

Ce résultat est connu sous le nom de **seconde identité de Bianchi**.

EURISTIDE : Ce résultat exprime que si on prend les dérivées covariantes de 3 vecteurs de coordonnées déviés suivant une courbe fermée le long de trois axes de coordonnées distincts, leur somme est nulle.

MATHINE : Continuons à présenter les développements mathématiques autour du tenseur de Riemann. Nous pouvons définir un nouveau tenseur, dit **tenseur de Ricci**, en contractant deux indices du tenseur de Riemann :

$$R_{\mu\nu} = R_{\mu\lambda\nu}^{\lambda}. \quad (683)$$

Ce tenseur est symétrique :

$$R_{\mu\nu} = R_{\nu\mu}. \quad (684)$$

Nous pouvons monter un des indices du tenseur de Ricci, pour écrire :

$$R_{\mu}^{\lambda} = g^{\lambda\nu} R_{\mu\nu}. \quad (685)$$

Puis, en contractant une nouvelle fois ce dernier tenseur, nous obtenons un nombre (on dit un scalaire) qui s'appelle le **scalaire de Ricci** :

$$R = R_{\mu}^{\mu}. \quad (686)$$

Nous allons pouvoir utiliser ce scalaire et ce tenseur de Ricci pour exprimer sous une autre forme l'identité de Bianchi. Partons de l'identité de Bianchi, et effectuons sur les tenseurs de Riemann, une double contraction au moyen du tenseur métrique :

$$\nabla_{\lambda} R_{\rho\sigma\mu\nu} + \nabla_{\rho} R_{\sigma\lambda\mu\nu} + \nabla_{\sigma} R_{\lambda\rho\mu\nu} = 0. \quad (687)$$

D'où :

$$g^{\nu\sigma} g^{\mu\lambda} (\nabla_{\lambda} R_{\rho\sigma\mu\nu} + \nabla_{\rho} R_{\sigma\lambda\mu\nu} + \nabla_{\sigma} R_{\lambda\rho\mu\nu}) = 0. \quad (688)$$

Soit :

$$\nabla^{\mu} R_{\rho\mu} - \nabla_{\rho} R + \nabla^{\nu} R_{\rho\nu} = 0. \quad (689)$$

Ce qui peut s'écrire :

$$\nabla^{\mu} R_{\rho\mu} = \frac{1}{2} \nabla_{\rho} R. \quad (690)$$

Cette dernière expression est donc équivalente à l'identité de Bianchi.

Pour représenter la courbure, nous définissons également le **tenseur d'Einstein** par :

$$\boxed{G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu}}. \quad (691)$$

On constate que :

$$\nabla^{\mu} G_{\mu\nu} = \nabla^{\mu} R_{\mu\nu} - \frac{1}{2} \nabla^{\mu} R g_{\mu\nu} - \frac{1}{2} R \nabla^{\mu} (g_{\mu\nu}). \quad (692)$$

Or la dérivée covariante du tenseur métrique est nulle, puisque les symboles de Christoffel sont compatibles avec la métrique. Donc :

$$\begin{aligned} \nabla^{\mu} G_{\mu\nu} &= \nabla^{\mu} R_{\mu\nu} - \frac{1}{2} \nabla^{\mu} R g_{\mu\nu} \\ &= \nabla^{\mu} R_{\mu\nu} - \frac{1}{2} \nabla_{\nu} R = 0. \end{aligned} \quad (693)$$

Donc, l'identité de Bianchi est équivalente à l'équation :

$$\nabla^\mu G_{\mu\nu} = 0. \quad (694)$$

BEATRIX : Le tenseur d'Einstein est une représentation de la courbure plus synthétique. Mais il ne "code" pas toute l'information de la courbure, puisqu'il est le résultat de la contraction par deux fois du tenseur de Riemann, n'est-ce pas ?

## 78 Scène XI-6 : Courbure intrinsèque et extrinsèque

EURISTIDE : Oui, c'est vrai, le tenseur d'Einstein ne comprend qu'une partie de l'information sur la courbure de l'espace. On utilise parfois un autre tenseur, **le tenseur de Weyl**, qui fournit le complément du tenseur de Riemann qui ne se trouve pas dans le tenseur d'Einstein. Mais, je ne donnerai pas ici d'expression du tenseur de Weyl, qui fait appel à des notions mathématiques que nous n'avons pas abordées. Nous avons fait beaucoup de calculs. Il est temps maintenant de prendre un peu de recul, et de regarder ce que nous avons construit. Nous allons prendre l'exemple de notre sphère  $S^2$  favorite. Supposons que la sphère ait pour rayon  $a$ , et considérons la métrique de la surface de la sphère exprimée dans les coordonnées sphériques  $(\theta, \phi)$  :

$$ds^2 = a^2(d\theta^2 + \sin^2\theta d\phi^2). \quad (695)$$

A partir de là, on peut écrire les symboles de Christoffel de la sphère ; ceux qui ne sont pas nuls sont :

$$\begin{aligned} \Gamma_{\phi\phi}^\theta &= -\sin\theta \cos\theta \\ \Gamma_{\theta\phi}^\phi = \Gamma_{\phi\theta}^\phi &= \cot\theta. \end{aligned} \quad (696)$$

Par conséquent, toutes les composantes du tenseur de Riemann sont déduites de celle-ci :

$$\begin{aligned} R_{\phi\theta\theta\phi}^\theta &= \partial_\theta \Gamma_{\phi\phi}^\theta - \partial_\phi \Gamma_{\theta\phi}^\theta + \Gamma_{\theta\lambda}^\theta \Gamma_{\phi\phi}^\lambda - \Gamma_{\phi\lambda}^\theta \Gamma_{\theta\phi}^\lambda \\ &= (\sin^2\theta - \cos^2\theta) + \sin\theta \cos\theta \cot\theta \\ &= \sin^2\theta. \end{aligned} \quad (697)$$

Et :

$$\begin{aligned} R_{\theta\theta} &= g^{\phi\phi} R_{\phi\theta\theta\phi} = 1 \\ R_{\theta\phi} &= R_{\phi\theta} = 0 \\ R_{\phi\phi} &= g^{\theta\theta} R_{\theta\phi\phi\theta} = \sin^2\theta. \end{aligned} \quad (698)$$

Et pour finir, le scalaire de Ricci s'exprime par :

$$R = g^{\theta\theta} R_{\theta\theta} + g^{\phi\phi} R_{\phi\phi} = \frac{2}{a^2}. \quad (699)$$

On constate deux choses : le scalaire de Ricci est une constante. Cela signifie que la variété possède une courbure constante en tous points de la surface. C'est vrai que nous savons que la sphère, par symétrie, possède une courbure constante en tous points. Elle possède donc une forme de symétrie, vis à vis de sa courbure. On dit que la variété est **maximalement symétrique**, c'est-à-dire qu'elle possède le maximum de propriétés de symétrie, du fait de cette courbure constante. De plus, ce scalaire est positif, ce qui a également une signification physique significative.



Le signe du scalaire de Ricci exprime la forme de la courbure de l'espace. Prenons l'exemple d'une surface à 2 dimensions. Si le scalaire de Ricci est positif, l'espace est courbé comme une sphère ou une boule. Si le scalaire de Ricci est négatif, l'espace est courbé comme une selle de cheval, c'est-à-dire que dans une certaine direction, l'espace s'incurve vers le bas, et dans une autre direction, l'espace s'incurve vers le haut.

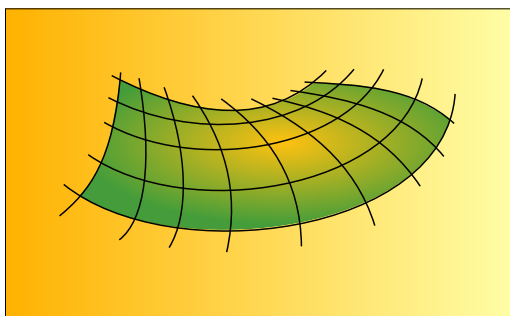


Fig. 130 - La courbure négative

On dit que l'espace dont la courbure est en forme de boule a une courbure positive et que l'espace dont la courbure est en forme de selle de cheval a une courbure négative. Un espace dont la courbure est nulle est tout simplement plat.

Ces notions seront utiles lorsque nous aborderons la cosmologie relativiste, pour nous permettre d'établir si le modèle d'univers possède une courbure nulle, positive ou négative.

## Treizième partie

# Acte XII - La gravitation

## 79 Scène XII-1 : Le principe d'équivalence

BEATRIX : Bon, je crois que les choses sont claires maintenant. La courbure d'un espace, propriété intrinsèquement codée dans la métrique de l'espace, peut être calculée comme la déviation d'un vecteur dans son transport parallèle le long d'un chemin fermé. C'est le tenseur de Riemann  $R_{\sigma\mu\nu}^{\rho}$  qui permet de modéliser la courbure suivant ce principe, ainsi que ses tenseurs déduits : le tenseur de Ricci  $R_{\mu\nu}$ , et le tenseur d'Einstein  $G_{\mu\nu}$ .

En contractant ce tenseur sur la totalité de ses indices, on obtient un scalaire appelé scalaire de Ricci. Ce scalaire donne en particulier la forme de la courbure. Aux endroits où ce scalaire est positif, l'espace curviligne possède une courbure positive, comme un boule; aux endroits où ce scalaire est négatif, l'espace curviligne possède une courbure négative, comme une selle de cheval; si le scalaire de Ricci est nul, alors l'espace est plat.

Enfin, on peut dire que les équivalents des lignes droites sur un espace curviligne sont les géodésiques. Les géodésiques sont des courbes qui minimisent la distance parcourue entre deux points dans un espace curviligne, et on peut aussi les définir comme des courbes qui transportent parallèlement leurs propres tangentes.

EURISTIDE : Voilà, c'est un bon résumé de la notion de courbure. Et c'est important de bien comprendre ces notions, puisqu'elles sont utilisées intensivement en relativité générale.

Nous avons d'ailleurs maintenant acquis avec fierté toutes les notions mathématiques nécessaires au développement de la relativité générale. Après ces nombreux pages de mathématiques, nous allons pouvoir refaire un peu de physique.

Après avoir mis en oeuvre la théorie de la relativité restreinte, Einstein s'est aperçu que la masse n'avait pas d'existence propre, mais qu'elle s'avérait être une simple manifestation de l'énergie et de l'impulsion. En effet, une masse au repos devenait en fait un manifestation de l'énergie, sous la forme :

$$E = mc^2. \tag{700}$$

Ceci a conduit Einstein à imaginer que la gravitation, qui était à l'origine pour Newton, une manifestation (sous forme de force) de la masse, devenait en fait une manifestation équivalente à l'accélération, et qu'il n'y avait en fait aucun moyen pour un physicien de distinguer entre une accélération et un champ gravitationnel. C'est ce qu'on appelle **le principe d'équivalence d'Einstein**.

Tout le génie d'Einstein a consisté à faire un pas supplémentaire et à interpréter cette dualité entre accélération et champ gravitationnel en considérant que l'énergie et l'impulsion (provoquées par exemple par un corps massif immobile ou en mouvement) provoquent une déformation de l'espace-temps; cette déformation de l'espace-temps le conduit à devenir curviligne; alors les trajectoires libres habituellement rectilignes dans l'espace-temps plat vont se mettre à suivre des géodésiques de l'espace-temps. C'est cette déformation curviligne de l'espace-temps et la transformation de ces lignes droites en géodésiques qui constitue les mouvements accélérés que nous observons et connaissons en présence d'un champ gravitationnel.

BEATRIX : En quelque sorte, on peut considérer que la masse ou l'énergie ou l'impulsion des corps

massifs provoque une déformation de l'espace-temps, et donne aux corps leur accélération ou les phénomènes que l'on observe en présence d'un champ gravitationnel.

EURISTIDE : Oui, c'est cela. Et la physique de la relativité générale va donc se traduire en un problème de géométrie dans une variété curviligne. C'est la raison pour laquelle nous avons dû réaliser une si longue introduction mathématique concernant les propriétés et les formulations du concept de variété en géométrie. Mais revenons à la notion d'équivalence entre accélération et gravitation. L'intuition d'Einstein provenait de ce que la masse utilisée pour la gravitation (masse pesante) et la masse utilisée dans les mouvements et donc pour l'accélération (masse inertielle) sont égales. Ces deux masses ont été mesurées très précisément par différentes expériences, et toutes concourent à démontrer que ces masses sont extrêmement proches, et par conséquent égales aux précisions de mesure près. Ces deux masses étant des choses fondamentalement différentes, le fait qu'elles soient égales devait avoir une signification physique. Et cette signification physique, c'était qu'il devait y avoir un lien fort entre gravitation et accélération.

C'est ce que le principe d'équivalence d'Einstein dit :

**Il n'y a aucun moyen de distinguer entre une accélération et un champ gravitationnel dans un région suffisamment petite de l'espace.**

Mais ce principe peut s'écrire autrement :

**Dans des régions suffisamment petites de l'espace-temps, les lois de la physique se réduisent à celles de la relativité restreinte; il est impossible de détecter l'existence d'un champ gravitationnel.**

BEATRIX : En d'autres termes, le principe d'équivalence d'Einstein dit que dans un région suffisamment petite, l'espace-temps est plat, puisque les lois y sont celles de la relativité restreinte. Donc l'espace-temps est une variété différentielle, puisque dans des petites régions, il peut être assimilé à un espace-temps plat. C'est ce que nous avons vu à propos des variétés différentielles, non ?

MATHINE : C'est exactement cela. L'espace-temps est une variété différentielle dont la courbure est une conséquence de l'énergie et de l'impulsion des corps massifs qui s'y trouvent. Ceci constitue les fondations de la relativité générale.

EURISTIDE : Pour étudier les lois de la relativité restreinte dans une petite région de l'espace, nous avons besoin d'un référentiel inertiel. Mais ce référentiel inertiel n'est applicable qu'à une petite région de l'espace-temps qui est plate, et ne peut pas s'étendre à l'ensemble de l'espace-temps qui est une variété, donc un espace-temps curviligne a priori. Le référentiel inertiel limité à une petite région est appelé **référentiel inertiel local** ou **référentiel localement inertiel** en un point.

MATHINE : Ce référentiel localement inertiel correspond aux coordonnées normales riemaniennes ou au référentiel de Lorentz local de la variété différentielle en un point.

EURISTIDE : Ce point est très important. Les lois de la physique que nous devons déterminer doivent être des lois compatibles avec la structure de variété différentielle. Comme nous l'avons vu, de telles lois doivent être des relations entre des tenseurs, afin qu'elles restent valables dans l'espace-temps curviligne quel que soit le point où elles sont appliquées.

En revanche, ces lois doivent se ramener aux lois de la relativité restreinte ou aux lois de Newton lorsque la courbure de l'espace-temps est nulle ou lorsque les vitesses considérées sont faibles.

BEATRIX : Nous avons donc un moyen de vérifier que les lois de la relativité générale que nous établissons sont justes : il suffira d'annuler la courbure de l'espace-temps pour vérifier qu'elles se ramènent à la relativité restreinte, ou de considérer des vitesses faibles pour vérifier qu'elles se ramènent à la mécanique de Newton.

EURISTIDE : On peut employer une technique simple pour passer des lois de l'espace-temps plat aux lois de l'espace-temps curviligne. Il suffit de remplacer les dérivées classiques par des dérivées covariantes. Par exemple, la formule classique de la conservation de l'énergie s'écrit :

$$\partial_{\mu} T^{\mu\nu} = 0, \quad (701)$$

où  $T^{\mu\nu}$  est le tenseur d'énergie-impulsion. Le tenseur représente la densité d'énergie et d'impulsion de corps massifs dans l'espace, mais nous l'étudierons plus en détail tout à l'heure. En particulier, le terme  $T^{00}$  représente l'énergie de repos du corps considéré.

Pour transformer cette équation et l'adapter à un espace-temps curviligne, nous écrivons tout simplement :

$$\nabla_{\mu} T^{\mu\nu} = 0. \quad (702)$$

Il s'agit de la loi de conservation de l'énergie-impulsion en présence d'un champ gravitationnel. Cette technique est appelée **principe de covariance**.

BEATRIX : Avons-nous un moyen de vérifier que le principe d'équivalence d'Einstein est vrai ? Cela me semble important, puisque c'est le principe fondateur de la relativité générale.

EURISTIDE : Comme tout principe physique, nous ne pouvons pas vérifier sa légitimité directement par la théorie. Mais ce que nous pouvons vérifier, c'est que ses conséquences ne sont pas contredites par l'expérience. Il en va ainsi de toute théorie physique : une théorie physique est considérée comme valide tant qu'elle permet de faire des prédictions qui ne sont pas contredites par l'expérience.

Une des prédictions célèbres du principe d'équivalence d'Einstein est le décalage vers le rouge gravitationnel.

BEATRIX : De quoi s'agit-il ?

EURISTIDE : Tu connais certainement l'effet Doppler. Lorsqu'une voiture s'approche de toi à grande vitesse, le bruit qu'elle fait semble plus aigu que dans la réalité. Lorsque la voiture passe à ta hauteur et puis s'éloigne, le bruit de la voiture devient brusquement plus grave.

BEATRIX : Oui, je connais cet effet Doppler. La voiture émet des ondes sonores qui viennent à ma rencontre, sous forme de vagues que nous appellerons **fronts d'ondes**, espacés d'une distance égale à la longueur d'onde  $\lambda$  du son de la voiture. Plus cette longueur d'onde est petite, plus le son est aigu.

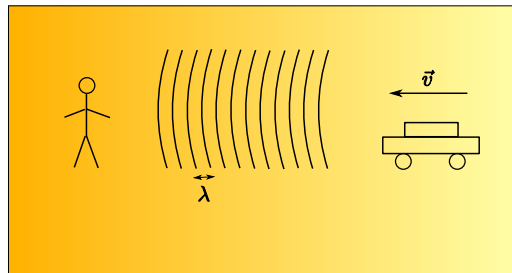


Fig. 131 - Propagation du son d'une voiture

A l'instant  $t_0$ , un front d'onde  $F_0$  est émis. Si  $T$  est la période du son de la voiture, le front d'onde suivant  $F_1$  est émis à l'instant  $t = t_0 + T$ . Mais pendant cette durée  $T$ , la voiture s'est approchée de moi d'une distance  $d = vT$ , où  $v$  est la vitesse de la voiture.

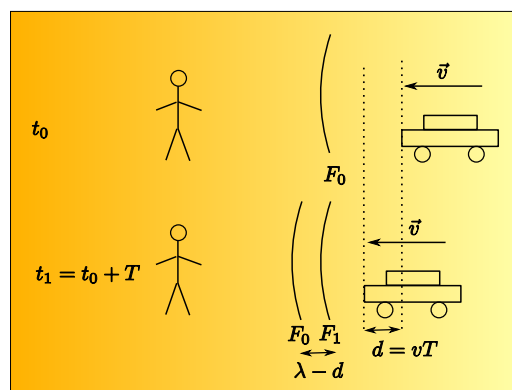


Fig. 132 - L'effet Doppler quand la voiture se rapproche

La distance entre les fronts d'onde  $F_0$  et  $F_1$  n'est pas conséquent plus  $\lambda$ , mais :

$$\lambda - d = \lambda - vT. \quad (703)$$

Il s'ensuit que le son est plus aigu lorsque la voiture roule dans ma direction.  
Supposons maintenant que la voiture s'éloigne de moi :

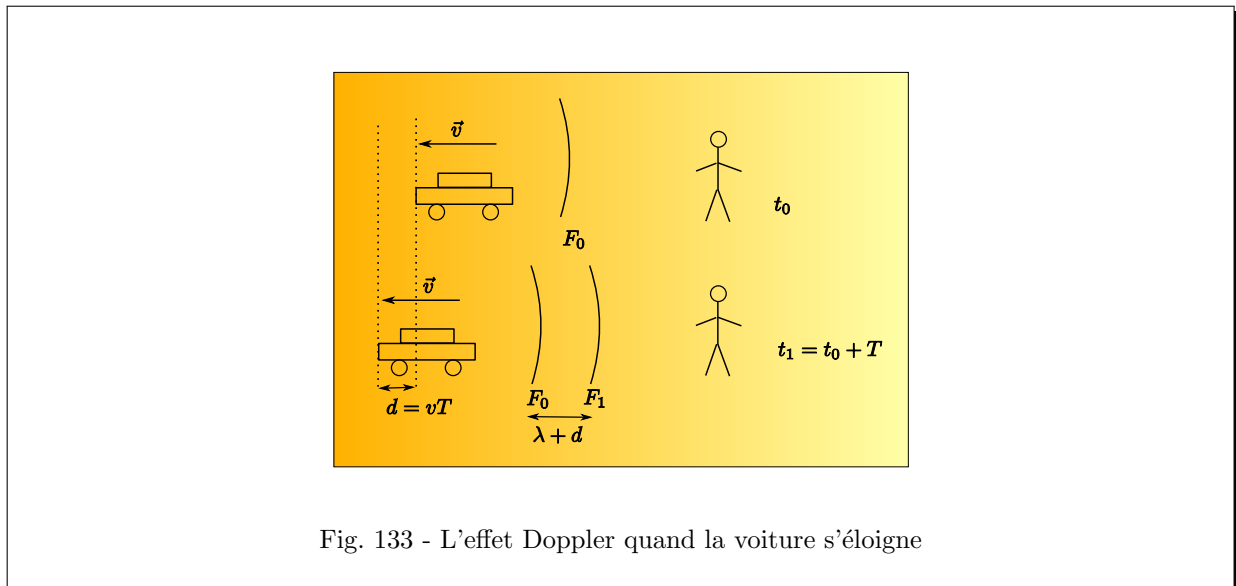


Fig. 133 - L'effet Doppler quand la voiture s'éloigne

A l'instant  $t_0$ , un front d'onde  $F_0$  est émis. Le front d'onde suivant  $F_1$  est émis à l'instant  $t_1 = t_0 + T$ . Mais pendant la durée  $T$ , la voiture s'est éloignée de moi d'une distance  $d = vT$ . La distance entre les fronts d'onde  $F_0$  et  $F_1$  n'est pas conséquent plus  $\lambda$ , mais :

$$\lambda + d = \lambda + vT. \quad (704)$$

Il s'ensuit que le son est plus grave lorsque la voiture roule en s'éloignant de moi.

**EURISTIDE** : Excellente démonstration. Eh bien, il se trouve que la gravitation est capable de provoquer également, d'après le principe d'équivalence d'Einstein, un "effet Doppler" particulier, appelé **décalage vers le rouge gravitationnel**. C'est une des célèbres prédictions de la théorie d'Einstein.

**MATHINE** : Considérons deux lampes émettrices de photons,  $A$  et  $B$ , distantes d'une distance  $d$ , et possédant une accélération  $a$ .

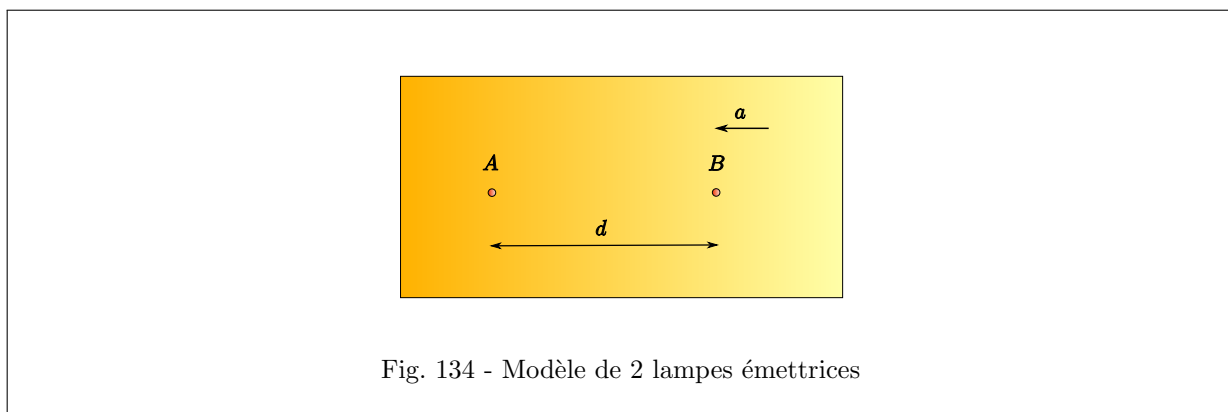


Fig. 134 - Modèle de 2 lampes émettrices

A l'instant  $t_0$ , on suppose que  $B$  envoie un photon vers  $A$  de longueur d'onde  $\lambda_0$ . Le photon atteint le point  $A$ , dans le référentiel des points  $A$  et  $B$ , après un instant :

$$\Delta t = \frac{d}{c}. \quad (705)$$

Pendant ce temps, les lampes ont pris une vitesse supplémentaire :

$$\Delta v = a\Delta t = \frac{ad}{c}. \quad (706)$$

Par conséquent, la longueur d'onde semble s'allonger pour devenir :

$$\lambda = \lambda_0 + \Delta\lambda = \lambda_0 + \Delta v T_0 = \lambda_0 + \Delta v \frac{\lambda_0}{c}, \quad (707)$$

$T_0$  étant la période de la lumière.

Soit :

$$\Delta\lambda = \Delta v \frac{\lambda_0}{c} = \frac{ad\lambda_0}{c^2}. \quad (708)$$

Cette expression donne la valeur du décalage de la longueur d'onde de la lumière dans ces conditions d'accélération. La longueur d'onde étant allongée, il s'ensuit un décalage de la lumière vers le rouge.

Considérons maintenant que nous nous trouvons à la hauteur  $d$  sur la surface d'une planète, dans un champ gravitationnel  $a_g$ . D'après le principe d'équivalence d'Einstein, les lois de la physique sont les mêmes dans un champ gravitationnel et dans un référentiel accéléré de même accélération.

Par conséquent, l'effet de décalage vers le rouge doit aussi être observé pour un corps placé à l'altitude  $d$  dans un champ gravitationnel de valeur  $a_g$  :

$$\Delta\lambda = \frac{a_g d \lambda_0}{c^2}. \quad (709)$$

**EURISTIDE** : C'est ce qu'on appelle le **décalage vers le rouge gravitationnel**. C'est une prédiction du principe d'équivalence d'Einstein qui a été vérifiée expérimentalement par Pound et Rebka en 1960, en mesurant la modification de fréquence de rayons Gamma voyageant du sol au sommet des laboratoires Jefferson à Harvard.

## 80 Scène XII-2 : Les équations du champ d'Einstein

BEATRIX : Mais je suppose que le décalage vers le rouge dû à la gravitation de la Terre est infime. En tout cas, nous ne percevons pas ce décalage.

EURISTIDE : C'est exact : il est infime. La formule que nous avons calculée montre qu'il est de l'ordre de grandeur du rapport entre la longueur d'onde de la lumière (exprimée en  $\mu m$ ), et le carré de la vitesse de la lumière (exprimée en millions de mètres par seconde!)

Nous allons maintenant passer au vif du sujet : nous allons pouvoir établir les équations du champ d'Einstein. Ces équations permettent d'établir le comportement des systèmes physiques dans un champ gravitationnel, et en particulier, elles remplaceront les équations de Newton.

BEATRIX : C'est-à-dire qu'en fait, les équations de Newton deviendront un cas particulier des équations d'Einstein, non ? Enfin, plus précisément un cas particulier lorsque le champ et la vitesse sont faibles.

EURISTIDE : Oui, nous verrons plus tard comment ces équations se ramènent aux équations de Newton pour un champ faible.

En attendant, pour établir les équations du champ d'Einstein, nous allons procéder de deux façons différentes. D'abord, nous allons raisonner intuitivement en partant du champ gravitationnel newtonien : en généralisant l'équation dite **de Poisson** pour le potentiel gravitationnel newtonien, nous allons déduire les équations du champ d'Einstein. Puis nous procéderons de façon plus rigoureuse en déduisant les équations d'Einstein du principe de moindre action.

MATHINE : Nous allons donc commencer par la démonstration intuitive. Nous partons en fait de l'équation qui décrit le comportement du potentiel de gravitation newtonien. Cette équation s'écrit :

$$\boxed{\nabla^2 \Phi = 4\pi G\rho.} \quad (710)$$

EURISTIDE : Cette équation nous dit ce qu'est le potentiel gravitationnel et la façon dont il se comporte. Avant de commenter cette équation, nous allons commencer par expliquer ce qu'est le potentiel gravitationnel. Le potentiel gravitationnel, c'est un champ qui détermine l'action que possède la gravitation sur les corps. Plus ce champ varie rapidement d'un point à un autre, plus l'action de la gravitation sur les corps est importante.

Dans la théorie de Newton, l'accélération  $\vec{a}$  appliquée aux corps du fait de la présence d'un champ gravitationnel, appelé potentiel gravitationnel, va être d'autant plus grande que la variation du potentiel gravitationnel est grande. Cette loi s'écrit :

$$\vec{a} = -\vec{\nabla}\Phi. \quad (711)$$

L'expression  $\vec{\nabla}$  représente ce qu'on appelle le gradient, c'est-à-dire la vitesse de variation du champ  $\Phi$  dans les différentes directions de l'espace. Cette expression signifie que l'accélération  $a$  est de sens opposé au gradient du potentiel gravitationnel et de même intensité.

Le gradient  $\vec{\nabla}\Phi$  du potentiel gravitationnel est :

$$\vec{\nabla}\Phi = \overrightarrow{\partial_i}\Phi, \quad (712)$$



autrement dit, c'est le vecteur dont les coordonnées sont les dérivées partielles du champ par rapport à chacune des coordonnées spatiales, ce qui exprime effectivement bien la vitesse de variation du champ dans les différentes directions de l'espace.

Interprétons cette loi : dessinons le champ sous forme de lignes isopotentielles, c'est-à-dire sous la forme de lignes dont les points possèdent tous le même potentiel ; et supposons que nous nous trouvons dans un cas de figure où le champ augmente lorsqu'on se rapproche d'un point  $P$  (c'est le cas d'une planète par exemple, pour laquelle le champ gravitationnel augmente au fur et à mesure qu'on se rapproche du centre de la planète) :

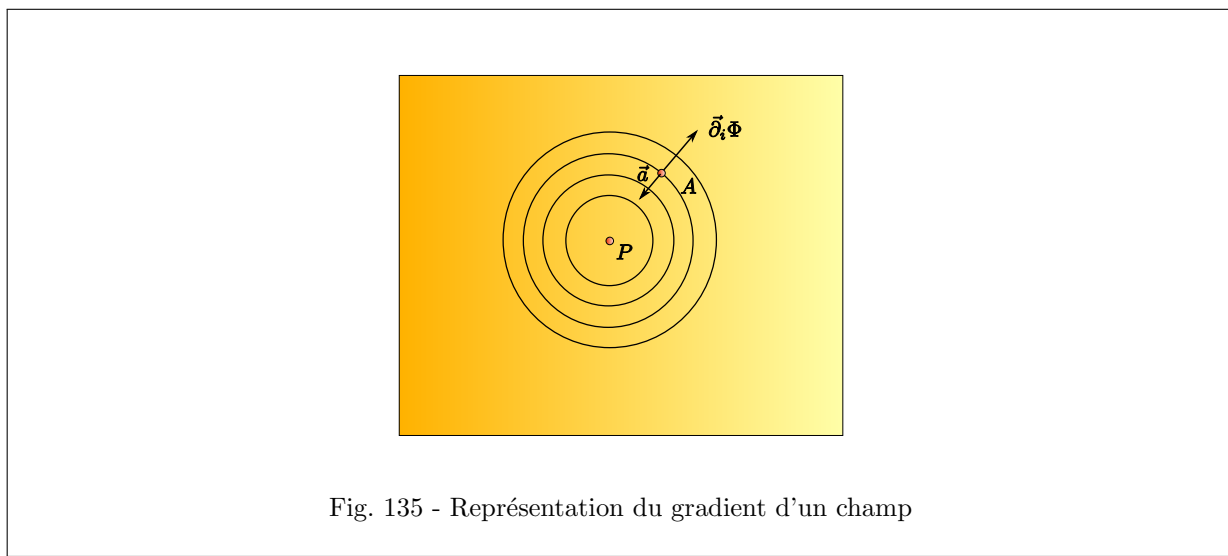


Fig. 135 - Représentation du gradient d'un champ

En un point  $A$  au voisinage du point  $P$ , on constate que le vecteur :

$$\vec{\nabla} \Phi = \vec{\partial}_i \Phi, \quad (713)$$

est un vecteur excentrique par rapport au point  $P$ .

Par conséquent, le vecteur accélération est opposé en sens à  $\vec{\nabla} \Phi$ .

Revenons maintenant à l'équation de Poisson :  $\nabla^2 \Phi = 4\pi G\rho$ . L'expression  $\nabla^2 \Phi$  représente ce qu'on appelle le Laplacien du potentiel gravitationnel dans les différentes directions de l'espace.  $G$  représente la valeur de la gravitation universelle de la Terre. Cette équation nous dit que le laplacien du potentiel gravitationnel est proportionnel à la densité de matière notée  $\rho$ . Le laplacien du potentiel gravitationnel est le module au carré du vecteur gradient du potentiel gravitationnel, et s'écrit :

$$\nabla^2 \Phi = \partial_i \partial_j \Phi \delta^{ij}. \quad (714)$$

C'est donc la somme des dérivées secondes du potentiel gravitationnel.

Par conséquent, la somme des dérivées secondes du potentiel gravitationnel est proportionnelle à la densité de matière. En d'autres termes, plus la densité de matière est importante, plus le champ potentiel s'accroît rapidement. Plus précisément, l'accélération avec laquelle le potentiel gravitationnel s'accroît est proportionnelle à la densité de matière.

BEATRIX : C'est somme toute très intuitif. Plus il y a de matière, plus la gravitation s'accroît rapidement lorsqu'on s'approche des corps.

MATHINE : Voilà. Ceci étant mis au point, nous avons donc une expression, dans la théorie de Newton, de l'évolution du potentiel gravitationnel :

$$\nabla^2\Phi = 4\pi G\rho. \quad (715)$$

Nous cherchons maintenant l'équivalent de cette loi en relativité générale. Nous allons donc chercher une loi qui reste valable même lorsque l'espace est curviligne. L'outil mathématique que nous connaissons et qui répond à cette exigence est le tenseur. Il nous faut donc trouver une équation tensorielle à partir de cette équation de Poisson, qui est une équation entre scalaires.

BEATRIX : Comment faire ? Il faut trouver l'équivalent tensoriel de la densité de matière et l'équivalent tensoriel du potentiel newtonien, je suppose.

MATHINE : Oui, c'est ce que nous devons chercher.

EURISTIDE : La généralisation de la densité de matière utilisée pour la relativité générale est le tenseur d'énergie-impulsion  $T^{\mu\nu}$ . C'est un tenseur symétrique de valence (0,2) qui nous dit tout ce que nous voulons savoir sur les aspects énergétiques d'un système : la densité d'énergie, la pression, la tension. Ce tenseur exprime en fait le flux de la 4-impulsion sur une surface ayant  $x^\nu$  constant, c'est-à-dire la quantité de 4-impulsion qui traverse cette surface par unité de temps. Comme nous le savons, d'après Einstein, la masse n'est qu'une forme particulière d'énergie. Dans ce tenseur, la composante  $T^{00}$  représente la densité d'énergie  $\rho$  du système. Si le système est constitué de matière uniquement, alors  $T^{00}$  est bien la densité de matière :

$$T^{00} = \rho. \quad (716)$$

Dans le cas d'un fluide parfait, qui est isotrope dans son référentiel de repos, les autres composantes diagonales  $T^{11}$ ,  $T^{22}$  et  $T^{33}$  représentent la pression du système.

Pour un fluide qui n'est pas parfait, et qui n'est pas isotrope, les composantes du tenseur vont représenter les différentes caractéristiques du système en termes de densité d'énergie, de pression et de tension. Dans la relativité générale, nous nous intéresserons en général à des systèmes simples, où ce tenseur d'énergie-impulsion n'a que des composantes diagonales.

BEATRIX : Nous avons donc trouvé l'équivalent tensoriel de la densité de matière ; c'est le tenseur d'énergie-impulsion. Qu'en est-il du potentiel gravitationnel ?

EURISTIDE : En fait, nous recherchons l'équivalent tensoriel du laplacien du potentiel gravitationnel. Le laplacien, comme nous l'avons vu, prend les dérivées partielles secondes des composantes du potentiel gravitationnel. Par ailleurs, l'équivalent du potentiel gravitationnel devrait être une quantité qui représente la gravitation dans le système. Or nous avons vu que la gravitation n'était qu'une courbure de l'espace-temps. Et nous avons vu aussi que la courbure de l'espace-temps est entièrement définie dans le tenseur métrique  $g^{\mu\nu}$ .

BEATRIX : Ah oui ! Je comprends. Nous allons chercher un tenseur dont les composantes sont des dérivées secondes du tenseur métrique. En faisant cela, nous prenons en compte l'hypothèse d'Einstein qui veut que la gravitation (ou son potentiel) soit une courbure de l'espace-temps.

EURISTIDE : Oui, c'est bien ce que signifie physiquement ce choix. Et quels tenseurs connais-tu qui soient de valence  $(0, 2)$  et qui aient des dérivées secondes du tenseur métrique dans leurs composantes ?

BEATRIX : Nous avons vu le tenseur de Ricci...

EURISTIDE : Le tenseur de Ricci est un bon exemple, mais ce n'est pas celui que nous allons utiliser. Nous allons prendre le tenseur d'Einstein :

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}, \quad (717)$$

où  $R_{\mu\nu}$  est le tenseur de Ricci.

BEATRIX : Pourquoi le tenseur d'Einstein plutôt que le tenseur de Ricci ?

EURISTIDE : Parce que nous avons une contrainte à prendre en compte. Nous savons que le tenseur d'énergie-impulsion doit être conservé, autrement dit ses dérivées covariantes doivent être nulles :

$$\nabla^\mu T_{\mu\nu} = 0. \quad (718)$$

C'est la loi de la conservation de l'énergie ("rien ne se crée, rien ne se perd"). Comme nous cherchons un tenseur représentant le potentiel gravitationnel et qui soit proportionnel à ce tenseur d'énergie-impulsion, il est donc indispensable que le tenseur que nous choisissons ait également une dérivée covariante nulle.

Ce n'est pas le cas du tenseur de Ricci, puisque nous connaissons l'identité de Ricci qui dit que :

$$\nabla^\mu R_{\mu\nu} = \frac{1}{2}\nabla_\nu R. \quad (719)$$

En revanche, nous avons vu que le tenseur d'Einstein (qui a été créé pour cela) répond bien à la contrainte :

$$\nabla^\mu G_{\mu\nu} = 0. \quad (720)$$

MATHINE : Et voilà. Nous avons donc notre équation tensorielle, qui généralise celle de Poisson :

$$G_{\mu\nu} = \alpha T_{\mu\nu}, \quad (721)$$

où  $\alpha$  est une constante que nous devons maintenant déterminer.

BEATRIX : Je suppose que nous allons adopter la méthode habituelle de l'approximation non relativiste...

EURISTIDE : Nous allons bien entendu analyser ce que donne cette équation tensorielle dans le cas d'un système non relativiste, dont le champ gravitationnel est faible et indépendant du temps.

MATHINE : Dans ces conditions, nous pouvons considérer que l'énergie de repos sera beaucoup plus grande que toutes les autres composantes du tenseur  $T^{\mu\nu}$ .

Par conséquent, nous pouvons limiter l'analyse à la composante  $T^{00}$  qui sera la seule composante significative du tenseur.

En contractant l'équation :

$$G_{\mu\nu} = \alpha T_{\mu\nu}, \quad (722)$$

en considérant que :

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}, \quad (723)$$

on obtient :

$$G = R - 2R = \alpha T, \quad (724)$$

où  $T$  est la contraction de  $T_{\mu\nu}$ . Donc :

$$R = -\alpha T. \quad (725)$$

Cette dernière expression peut être réinjectée dans l'équation d'origine, avec le tenseur d'Einstein :

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = R_{\mu\nu} + \frac{1}{2}\alpha Tg_{\mu\nu} = \alpha T_{\mu\nu}. \quad (726)$$

D'où :

$$R_{\mu\nu} = \alpha \left( T_{\mu\nu} - \frac{1}{2}Tg_{\mu\nu} \right). \quad (727)$$

Comme le champ gravitationnel est faible, la métrique peut être écrite sous la forme :

$$g_{\mu\nu} \approx \eta_{\mu\nu} + h_{\mu\nu}, \quad (728)$$

où  $h_{\mu\nu}$  est une petite perturbation, c'est-à-dire une valeur très petite par rapport à 1, et où  $\eta_{\mu\nu}$  est la métrique de Minkowski normalisée, sous sa forme canonique dans un référentiel lorentzien, telle que  $\eta_{00} = -1$ ,  $\eta_{ii} = 1$  ( $i=1$  à  $3$ ), et toutes les autres composantes soient nulles.

Par conséquent :

$$g_{00} \approx -1 + h_{00}, \quad (729)$$

et :

$$g^{00} \approx -1 - h_{00}. \quad (730)$$

Et par suite :

$$T \approx g^{00}T_{00} = -T_{00}. \quad (731)$$

Nous en déduisons que :

$$R_{00} = \alpha \left( T_{00} + \frac{1}{2}T_{00}(-1 + h_{00}) \right), \quad (732)$$

soit :

$$R_{00} \approx \frac{1}{2}\alpha T_{00}. \quad (733)$$

Par ailleurs, nous savons que :

$$R_{\sigma\mu\nu}^{\rho} = \partial_{\mu}\Gamma_{\nu\sigma}^{\rho} - \partial_{\nu}\Gamma_{\mu\sigma}^{\rho} + \Gamma_{\mu\lambda}^{\rho}\gamma_{\nu\sigma}^{\lambda} - \Gamma_{\nu\lambda}^{\rho}\Gamma_{\mu\sigma}^{\lambda}, \quad (734)$$

d'où :

$$R_{0j0}^i = \partial_j\Gamma_{00}^i - \partial_0\Gamma_{j0}^i + \Gamma_{j\lambda}^i\Gamma_{00}^{\lambda} - \Gamma_{0\lambda}^i\Gamma_{j0}^{\lambda}. \quad (735)$$

$\partial_0 \Gamma_{j0}^i$  est une dérivée par rapport au temps, qui s'annule pour un champ qui est statique par hypothèse.

Les termes :

$$\Gamma_{j\lambda}^i \Gamma_{00}^\lambda \quad (736)$$

et :

$$\Gamma_{0\lambda}^i \Gamma_{j0}^\lambda \quad (737)$$

contribuent au deuxième ordre de dérivée de la métrique, ils sont donc négligeables par rapport aux termes du premier ordre de dérivée de la métrique.

Par conséquent, nous pouvons faire l'approximation suivante :

$$R_{0j0}^i \approx \partial_j \Gamma_{00}^i. \quad (738)$$

On en déduit que :

$$\begin{aligned} R_{00} &= R_{0i0}^i = \partial_i \Gamma_{00}^i \\ &= \partial_i \left( \frac{1}{2} g^{i\lambda} (\partial_0 g_{\lambda 0} + \partial_0 g_{0\lambda} - \partial_\lambda g_{00}) \right) \end{aligned} \quad (739)$$

Le champ est statique, donc les dérivées par rapport au temps  $\partial_0$  du tenseur métrique sont nulles.

Donc :

$$\Gamma_{00}^i = -\frac{1}{2} g^{i\lambda} \partial_\lambda g_{00}, \quad (740)$$

ce qui, en utilisant l'approximation de la métrique de Minkowski, donne :

$$\begin{aligned} R_{00} &= -\frac{1}{2} \eta^{ij} \partial_i \partial_j h_{00} \\ &= -\frac{1}{2} \nabla^2 h_{00}. \end{aligned} \quad (741)$$

Donc, l'équation d'Einstein s'écrit :

$$R_{00} = -\frac{1}{2} \nabla^2 h_{00} = -\alpha T_{00}. \quad (742)$$

**BEATRIX** : Nous approchons du résultat. Cette équation ressemble à l'équation de Poisson. Il nous reste à trouver la relation entre  $h_{00}$  et le potentiel gravitationnel.

**EURISTIDE** : Oui, et pour cela, nous allons utiliser le fait que la généralisation du déplacement en ligne droite des particules libres, en déplacement le long d'une géodésique dans un espace-temps curviligne pour la relativité générale. En effet, les particules libres se déplacent en relativité générale le long des géodésiques de l'espace-temps.

**MATHINE** : Notre introduction mathématique nous permet de savoir maintenant que l'équation de la géodésique est :

$$\frac{d^2 x^\mu}{d\tau^2} + \Gamma_{\rho\sigma}^\mu \frac{dx^\rho}{d\tau} \frac{dx^\sigma}{d\tau} = 0 \quad (743)$$

où  $\tau$  est le temps propre.

Nous sommes dans une hypothèse non relativiste, et par conséquent les composantes :

$$\frac{dx^\rho}{d\tau} \quad (744)$$

où  $\rho \neq 0$  peuvent être négligées.

On obtient par conséquent :

$$\frac{d^2 x^\mu}{d\tau^2} + \Gamma_{00}^\mu \left( \frac{dt}{d\tau} \right)^2 = 0. \quad (745)$$

Par ailleurs, nous avons calculé tout-à-l'heure que :

$$\Gamma_{00}^i = -\frac{1}{2} \eta^{ij} \partial_j h_{00}. \quad (746)$$

Donc, l'équation de la géodésique devient :

$$\frac{d^2 x^\mu}{d\tau^2} = \frac{1}{2} \eta^{\mu\nu} \partial_\lambda h_{00} \left( \frac{dt}{d\tau} \right)^2. \quad (747)$$

Calculons la composante temporelle (donc pour  $\mu = 0$ ) :

$$\frac{d^2 t}{d\tau^2} = \frac{1}{2} \eta^{0\lambda} \partial_\lambda h_{00} \left( \frac{dt}{d\tau} \right)^2. \quad (748)$$

Or, dans le référentiel de Lorentz, les  $\eta^{0\lambda}$  sont tous nuls, sauf  $\eta^{00}$ , donc :

$$\frac{d^2 t}{d\tau^2} = \frac{1}{2} \eta^{00} \partial_0 h_{00} \left( \frac{dt}{d\tau} \right)^2. \quad (749)$$

Or le champ est statique, donc  $\partial_0 h_{00} = 0$ , d'où :

$$\frac{d^2 t}{d\tau^2} = 0. \quad (750)$$

Donc,  $\frac{dt}{d\tau}$  est constante.

Par ailleurs, en considérant uniquement les composantes spatiales des équations géodésiques :

$$\frac{d^2 x^i}{d\tau^2} = \frac{1}{2} \eta^{ij} \partial_j h_{00} \left( \frac{dt}{d\tau} \right)^2. \quad (751)$$

Or  $\eta^{ij}$  est la matrice identité :

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (752)$$

dans le référentiel lorentzien, donc :

$$\frac{d^2 x^i}{d\tau^2} = \frac{1}{2} \partial_i h_{00} \left( \frac{dt}{d\tau} \right)^2, \quad (753)$$

ou encore, en divisant les deux membres par le carré de  $\frac{dt}{d\tau}$  :

$$\frac{d^2 x^i}{d\tau^2} = \frac{1}{2} \partial_i h_{00}. \quad (754)$$

**EURISTIDE** : Cette équation ressemble à l'équation exprimant l'accélération en fonction du potentiel gravitationnel :

$$a = -\nabla\Phi, \quad (755)$$

à condition de poser :

$$\Phi = -\frac{1}{2}h_{00}. \quad (756)$$

MATHINE : Par conséquent, nous pouvons maintenant écrire l'équation d'Einstein sous la forme :

$$\nabla^2 h_{00} = -2\nabla^2 \Phi = -\alpha T_{00}. \quad (757)$$

Or, nous savons que, d'une part :

$$T_{00} = \rho, \quad (758)$$

où  $\rho$  est la densité de matière.

Et d'autre part, nous savons que :

$$\nabla^2 \Phi = 4\pi G\rho, \quad (759)$$

d'après l'équation de Poisson.

Par identification, on obtient donc :

$$-8\pi G\rho = -\alpha\rho, \quad (760)$$

d'où :

$$\alpha = 8\pi G. \quad (761)$$

Ce qui nous donne finalement l'équation d'Einstein pour la relativité générale :

$$\boxed{R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi GT_{\mu\nu}.} \quad (762)$$

BEATRIX : Ouf! Ca y est, nous les tenons, ces fameuses équations. Que nous disent-elles donc ?

EURISTIDE : Elles nous expliquent comment la courbure de l'espace-temps, représentée par le tenseur d'Einstein, ou les tenseurs  $R_{\mu\nu}$  et  $g_{\mu\nu}$  et par le scalaire de Ricci, est influencée par l'énergie et l'impulsion dans l'espace. Autrement dit, dès que l'espace contient de l'énergie, de la matière, une pression ou une tension, représentées par le tenseur d'énergie-impulsion  $T_{\mu\nu}$ , alors le tenseur d'Einstein n'est pas nul, ce qui signifie que l'espace-temps possède une courbure non nulle.

Dans le vide absolu, où il n'y a ni énergie, ni pression, ni tension, le tenseur d'énergie-impulsion est nul, et par conséquent :

$$R_{\mu\nu} = 0. \quad (763)$$

Cette dernière équation est l'équation d'Einstein dans le vide.

MATHINE : Nous avons donc construit les équations d'Einstein pour la relativité générale par une première méthode non formelle consistant à généraliser les équations de Newton et la notion de trajectoire droite de l'espace-temps plat.

Il nous faut maintenant employer une méthode plus rigoureuse pour retrouver ces équations. Cela consiste à utiliser le principe de moindre action.

La formulation de l'action a été proposée par Hilbert sous la forme :

$$S_H = \int \mathcal{L}_H d^n x, \quad (764)$$

c'est-à-dire sous la forme d'une intégrale sur la totalité de l'espace-temps d'une expression  $\mathcal{L}_H$  appelée densité de Lagrange.

EURISTIDE : Cette densité de Lagrange est ce qu'on appelle une densité tensorielle, c'est-à-dire que ce doit être une expression scalaire, multipliée par  $\sqrt{-g}$ , soit :

$$\mathcal{L} = \sqrt{-g}X, \quad (765)$$

où  $X$  est un scalaire faisant intervenir la métrique, et où  $g$  est le déterminant du tenseur métrique.

Or nous connaissons bien un scalaire faisant intervenir la métrique de l'espace-temps, il s'agit du scalaire de Ricci  $R$ . Hilbert a démontré que c'était le scalaire le plus simple répondant aux besoins de la formulation de l'action. C'est donc celui qui a été adopté :

$$\mathcal{L} = \sqrt{-g}R. \quad (766)$$

MATHINE : Nous allons maintenant calculer la variation de l'action, en faisant varier la métrique :

$$\delta S_H = \delta \left( \int \sqrt{-g} R d^n x \right). \quad (767)$$

Or, nous savons que :

$$R = g^{\mu\nu} R_{\mu\nu}, \quad (768)$$

donc :

$$\delta S_H = \int (\sqrt{-g} g^{\mu\nu} \delta R_{\mu\nu} + \sqrt{-g} R_{\mu\nu} \delta g^{\mu\nu} + R \delta \sqrt{-g}) d^n x. \quad (769)$$

Etudions d'abord le premier terme sous l'intégrale dans les parenthèses :

$$\sqrt{-g} g^{\mu\nu} \delta R_{\mu\nu}. \quad (770)$$

Nous devons calculer la variation du tenseur de Ricci. Or nous connaissons l'expression du tenseur de Riemann :

$$R_{\mu\lambda\nu}^{\rho} = \partial_{\lambda} \Gamma_{\nu\mu}^{\rho} + \Gamma_{\lambda\sigma}^{\rho} \Gamma_{\nu\mu}^{\sigma} - \partial_{\nu} \Gamma_{\lambda\mu}^{\rho} - \Gamma_{\nu\sigma}^{\rho} \Gamma_{\lambda\mu}^{\sigma}. \quad (771)$$

Par ailleurs, comme  $\delta \Gamma_{\nu\mu}^{\rho}$  est un tenseur deux fois covariant et une fois contravariant, on a :

$$\begin{aligned} \nabla_{\lambda} (\delta \Gamma_{\nu\mu}^{\rho}) &= \partial_{\lambda} (\delta \Gamma_{\nu\mu}^{\rho}) + \Gamma_{\lambda\sigma}^{\rho} (\delta \Gamma_{\nu\mu}^{\sigma}) \\ &\quad - \Gamma_{\lambda\nu}^{\sigma} (\delta \Gamma_{\sigma\mu}^{\rho}) - \Gamma_{\lambda\mu}^{\sigma} (\delta \Gamma_{\nu\sigma}^{\rho}). \end{aligned} \quad (772)$$

Il s'ensuit que :

$$\begin{aligned} \delta (R_{\mu\lambda\nu}^{\rho}) &= \partial_{\lambda} (\delta \Gamma_{\nu\mu}^{\rho}) + (\delta \Gamma_{\lambda\sigma}^{\rho}) \Gamma_{\nu\mu}^{\sigma} + \Gamma_{\lambda\sigma}^{\rho} (\delta \Gamma_{\nu\mu}^{\sigma}) - \partial_{\nu} (\delta \Gamma_{\lambda\mu}^{\rho}) \\ &\quad - (\delta \Gamma_{\nu\sigma}^{\rho}) \Gamma_{\lambda\mu}^{\sigma} - \Gamma_{\nu\sigma}^{\rho} (\delta \Gamma_{\lambda\mu}^{\sigma}) \\ &= \partial_{\lambda} (\delta \Gamma_{\nu\mu}^{\rho}) + (\delta \Gamma_{\lambda\sigma}^{\rho}) \Gamma_{\nu\mu}^{\sigma} - \Gamma_{\lambda\mu}^{\sigma} (\delta \Gamma_{\nu\sigma}^{\rho}) - \Gamma_{\lambda\nu}^{\sigma} (\delta \Gamma_{\sigma\mu}^{\rho}) \\ &\quad - \partial_{\nu} (\delta \Gamma_{\lambda\mu}^{\rho}) - \Gamma_{\nu\sigma}^{\rho} (\delta \Gamma_{\lambda\mu}^{\sigma}) + \Gamma_{\nu\mu}^{\sigma} (\delta \Gamma_{\lambda\sigma}^{\rho}) + \Gamma_{\lambda\nu}^{\sigma} (\delta \Gamma_{\sigma\mu}^{\rho}) \\ &= \nabla_{\lambda} (\delta \Gamma_{\nu\mu}^{\rho}) - \nabla_{\nu} (\delta \Gamma_{\lambda\mu}^{\rho}). \end{aligned} \quad (773)$$

Donc, le premier terme de l'intégrale s'écrit :

$$\begin{aligned} \sqrt{-g} g^{\mu\nu} \delta R_{\mu\nu} &= \sqrt{-g} g^{\mu\nu} (\nabla_{\lambda} (\delta \Gamma_{\nu\mu}^{\lambda}) - \nabla_{\nu} (\delta \Gamma_{\lambda\mu}^{\lambda})) \\ &= \sqrt{-g} \nabla_{\sigma} (g^{\mu\nu} (\delta \Gamma_{\mu\nu}^{\sigma}) - g^{\mu\sigma} (\delta \Gamma_{\lambda\mu}^{\lambda})). \end{aligned} \quad (774)$$



Le premier terme de l'intégrale s'écrit donc :

$$\int \sqrt{-g} \nabla_{\sigma} (g^{\mu\nu} (\delta\Gamma_{\mu\nu}^{\sigma}) - g^{\mu\sigma} (\delta\Gamma_{\lambda\mu}^{\lambda})) d^n x. \quad (775)$$

Il s'agit donc de l'intégrale de la divergence d'un vecteur par rapport à l'élément de volume sur la totalité de l'espace. Un important théorème d'analyse mathématique, le théorème d'Ostrogradski, montre qu'une telle intégrale sur le volume peut être ramenée à l'intégrale du même vecteur sur la frontière du volume (y compris à l'infini si le volume est ouvert), qui est donc une surface.

Or, par définition de la variation de l'action en dehors des frontières de l'intégration, les variations des  $g^{ij}$  s'annulent sur cette surface frontière. En conséquence, la variation de l'action s'annule sur la surface. Et l'intégrale sur la surface est nulle. Donc le premier terme de l'intégrale est nul.

Analysons maintenant le troisième terme :

$$R\delta\sqrt{-g}. \quad (776)$$

On démontre par le calcul matriciel que :

$$\delta\sqrt{-g} = -\frac{1}{2}\sqrt{-g}g_{\mu\nu}\delta g^{\mu\nu}. \quad (777)$$

Donc le troisième terme de l'intégrale s'écrit :

$$\int R\delta\sqrt{-g}d^n x = \int -\frac{1}{2}Rg_{\mu\nu}\sqrt{-g}\delta g^{\mu\nu}d^n x. \quad (778)$$

En transcrivant les deuxième et troisième termes dans l'équation, on obtient :

$$\delta S_H = \int \sqrt{-g}(R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu})\delta g^{\mu\nu}d^n x. \quad (779)$$

Cette expression doit être nulle pour toute variation de la métrique, donc le terme entre parenthèses dans l'intégrale est nul, c'est-à-dire :

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 0 = \frac{1}{\sqrt{-g}} \frac{\delta S}{\delta g^{\mu\nu}}. \quad (780)$$

Cette équation constitue l'équation du champ dans le vide.

Il s'agit maintenant d'analyser la situation en présence d'énergie ou de matière. L'action doit alors avoir la forme suivante :

$$S = S_H + S_M, \quad (781)$$

où  $S_M$  est l'action de la matière et où  $S_H$  doit avoir pour facteur constant  $-1/(8\pi G)$ .

On obtient :

$$\frac{1}{\sqrt{-g}} \frac{\delta S}{\delta g^{\mu\nu}} = \frac{1}{8\pi G}(R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu}) + \frac{1}{\sqrt{-g}} \frac{\delta S_M}{\delta g^{\mu\nu}} = 0. \quad (782)$$

En posant :

$$T_{\mu\nu} = -\frac{1}{\sqrt{-g}} \frac{\delta S_M}{\delta g^{\mu\nu}}, \quad (783)$$

on obtient l'équation d'Einstein :

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi GT_{\mu\nu}. \quad (784)$$

EURISTIDE : Nous en profitons pour avoir une définition du tenseur d'énergie-impulsion sous la forme :

$$T_{\mu\nu} = -\frac{1}{\sqrt{-g}} \frac{\delta S_M}{\delta g^{\mu\nu}}. \quad (785)$$

Cette expression montre que le tenseur d'énergie-impulsion est la variation de l'action de la matière par rapport au tenseur métrique. C'est-à-dire qu'il représente la variation de l'action de la matière et de l'énergie sur le système en fonction des variations de la courbure de l'espace-temps.

Ce tenseur d'énergie-impulsion peut être écrit sous forme matricielle :

$$T_{\mu\nu} = \begin{pmatrix} \rho & S_x/c & S_y/c & S_z/c \\ S_x/c & \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ S_y/c & \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ S_z/c & \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix}, \quad (786)$$

où le sens des composantes est le suivant :

- $\rho$  est la densité d'énergie dans l'espace-temps.
- $S_x$ ,  $S_y$  et  $S_z$  correspondent à la quantité d'énergie traversant l'unité de surface, par unité de temps dans la direction  $x$ ,  $y$  et  $z$  respectivement.
- $\sigma_{xx}$ ,  $\sigma_{xy}$ , etc. correspondent à la quantité d'impulsion traversant l'unité de surface par unité de temps, dans chacune des directions  $x$ ,  $y$  et  $z$ . Ce tenseur à trois dimensions est appelé **tenseur des contraintes**. Les termes diagonaux de ce tenseur correspondent à la pression lorsque la matière peut être considérée comme un fluide.

BEATRIX : Finalement, les équations d'Einstein constituent un modèle de tout système physique en établissant les relations entre d'une part la densité d'énergie, les flux d'énergie et les flux de quantité de mouvement de la matière, et d'autre part, la courbure de l'espace-temps. Si l'espace-temps est plat, c'est qu'il n'y a pas d'énergie, ni flux d'énergie, ni flux de quantité de mouvement. Si l'espace-temps est curviligne, c'est qu'il contient de l'énergie, un flux d'énergie ou de la matière en mouvement.

EURISTIDE : Oui, la signification de ces équations est aussi simple que cela. Mais quelle puissance de déduction et quelles curiosités physiques nous permettent de découvrir ces équations, aussi synthétiques soient elles !

BEATRIX : Oui, j'imagine que nous allons pouvoir maintenant en étudier les solutions ?

## Quatorzième partie

# Acte XIII - Les applications de la relativité générale

## 81 Scène XIII-1 : Le champ de gravitation faible

EURISTIDE : Oui, nous allons voir tour à tour les différentes applications et cas particuliers de ces équations, et nous allons pouvoir découvrir peu à peu les prédictions que cette théorie a pu apporter.

Et pour commencer, nous allons étudier le comportement et les conséquences de ces équations dans le cas d'un champ de gravitation faible. C'est évidemment le cas qui nous intéresse lorsque nous nous trouvons aux alentours d'une planète ou d'une étoile relativement peu massive.

Le fait que le champ de gravitation est faible va nous permettre de simplifier les équations en faisant des approximations, que nous aurons le droit de faire puisque certaines quantités sont faibles par rapport à d'autres, et par conséquent peuvent être négligées.

Nous allons donc considérer que le champ gravitationnel est faible. Mais il ne sera pas forcément statique, et les vitesses des particules ne seront pas forcément faibles. Si ces deux dernières hypothèses avaient été vérifiées, nous nous trouverions dans le cas de la limite newtonienne. Ce ne sera pas le cas ici.

MATHINE : Comme le champ gravitationnel est faible, la métrique peut être décomposée de la façon suivante :

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \quad (787)$$

où  $\eta_{\mu\nu}$  est la métrique de Minkowski décrivant l'espace-temps relativiste plat, et où  $h_{\mu\nu}$  est une petite perturbation de cette métrique, telle que le déterminant de la matrice des composantes de  $h_{\mu\nu}$  est très petit par rapport à 1 :

$$|h_{\mu\nu}| \lll 1. \quad (788)$$

Nous allons considérer que nous décrivons l'espace-temps dans un système de coordonnées où la métrique de Minkowski a pour composantes :

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & +1 \end{pmatrix}. \quad (789)$$

Nous avons, par définition :

$$g^{\mu\nu} g_{\nu\rho} = \delta_{\rho}^{\mu}. \quad (790)$$

En posant :

$$g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu}, \quad (791)$$

nous vérifions immédiatement que :

$$\begin{aligned} g^{\mu\nu} g_{\mu\nu} &= (\eta^{\mu\nu} - h^{\mu\nu})(\eta_{\mu\nu} + h_{\mu\nu}) \\ &= \eta^{\mu\nu} \eta_{\mu\nu} - h^{\mu\nu} \eta_{\mu\nu} + \eta^{\mu\nu} h_{\mu\nu} - h^{\mu\nu} h_{\mu\nu}. \end{aligned} \quad (792)$$

Les quantités  $-h^{\mu\nu}\eta_{\mu\nu}$  et  $+h^{\mu\nu}\eta_{\mu\nu}$  s'annulent, et la quantité  $h^{\mu\nu}h_{\mu\nu}$  est négligeable, puisqu'elle est du deuxième ordre par rapport à la quantité  $h_{\mu\nu}$  qui est considérée petite par hypothèse. Donc, on en déduit que :

$$g^{\mu\nu}g_{\mu\nu} \approx 1, \quad (793)$$

ce qui permet de vérifier qu'on a bien :

$$g^{\mu\nu} = \eta^{\mu\nu} - h^{\mu\nu}. \quad (794)$$

Par ailleurs, pour monter ou descendre les indices, il suffira d'employer  $\eta^{\mu\nu}$  et  $\eta_{\mu\nu}$ , puisque les corrections  $h_{\mu\nu}$  et  $h^{\mu\nu}$  introduisent des termes infinitésimaux du deuxième ordre que l'on peut négliger.

**EURISTIDE** : Pour interpréter cette expression du tenseur métrique, nous pouvons considérer que l'espace-temps est décrit par un tenseur perturbateur  $h_{\mu\nu}$  se propageant au sein d'un espace-temps plat, décrit par le tenseur de Minkowski  $\eta_{\mu\nu}$ . Ceci nous montre la nature physique de l'espace-temps que nous sommes en train de décrire.

**MATHINE** : Nous allons maintenant calculer les équations d'Einstein avec ces hypothèses. Pour commencer, nous aurons besoin de l'expression des symboles de Christoffel :

$$\begin{aligned} \Gamma_{\mu\nu}^{\rho} &= \frac{1}{2}g^{\rho\lambda}(\partial_{\mu}g_{\nu\lambda} + \partial_{\nu}g_{\lambda\mu} - \partial_{\lambda}g_{\mu\nu}) \\ &= \frac{1}{2}g^{\rho\lambda}(\partial_{\mu}(\eta_{\nu\lambda} + h_{\nu\lambda}) + \partial_{\nu}(\eta_{\lambda\mu} + h_{\lambda\mu}) - \partial_{\lambda}(\eta_{\mu\nu} + h_{\mu\nu})) \\ &\approx \frac{1}{2}\eta^{\rho\lambda}(\partial_{\mu}h_{\nu\lambda} + \partial_{\nu}h_{\lambda\mu} - \partial_{\lambda}h_{\mu\nu}), \end{aligned} \quad (795)$$

puisque les  $\eta_{\alpha\beta}$  sont constants et puisque nous pouvons remplacer le facteur  $g^{\rho\lambda}$  par  $\eta^{\rho\lambda}$ , étant donné que la perturbation s'y retrouverait au deuxième ordre, donc en quantité négligeable.

Calculons maintenant le tenseur de Riemann :

$$R_{\mu\nu\rho\sigma} = g_{\mu\lambda}(\partial_{\rho}\Gamma_{\sigma\nu}^{\lambda} - \partial_{\sigma}\Gamma_{\rho\nu}^{\lambda}) + g_{\mu\lambda}\Gamma_{\rho\lambda}^{\mu}\Gamma_{\sigma\nu}^{\lambda} - g_{\mu\lambda}\Gamma_{\sigma\lambda}^{\mu}\Gamma_{\rho\nu}^{\lambda}. \quad (796)$$

Or, les termes comportant le produit de deux symboles de Christoffel donnent des termes du deuxième ordre en  $h_{\alpha\beta}$ , donc ils sont négligeables.

Par ailleurs, en négligeant toujours les termes du deuxième ordre en  $h_{\alpha\beta}$ , nous pouvons remplacer le terme en facteur  $g_{\mu\lambda}$  par  $\eta_{\mu\lambda}$ . Ce qui donne :

$$\begin{aligned} R_{\mu\nu\rho\sigma} &= \eta_{\mu\lambda}\partial_{\rho}\Gamma_{\nu\sigma}^{\lambda} - \eta_{\mu\lambda}\partial_{\sigma}\Gamma_{\nu\rho}^{\lambda} \\ &= \frac{1}{2}\eta_{\mu\lambda}\eta^{\lambda\tau}(\partial_{\rho}\partial_{\nu}h_{\sigma\tau} + \partial_{\rho}\partial_{\sigma}h_{\tau\nu} - \partial_{\rho}\partial_{\tau}h_{\nu\sigma}) \\ &\quad - \frac{1}{2}\eta_{\mu\lambda}\eta^{\lambda\tau}(\partial_{\sigma}\partial_{\nu}h_{\rho\tau} + \partial_{\sigma}\partial_{\rho}h_{\tau\nu} - \partial_{\sigma}\partial_{\tau}h_{\nu\rho}) \\ &= \frac{1}{2}\delta_{\mu}^{\tau}(\partial_{\rho}\partial_{\nu}h_{\sigma\tau} + \partial_{\rho}\partial_{\sigma}h_{\tau\nu} - \partial_{\sigma}\partial_{\nu}h_{\rho\tau} - \partial_{\sigma}\partial_{\rho}h_{\tau\nu} \\ &\quad - \partial_{\rho}\partial_{\tau}h_{\nu\sigma} + \partial_{\sigma}\partial_{\tau}h_{\nu\rho}) \\ &= \frac{1}{2}(\partial_{\rho}\partial_{\nu}h_{\sigma\mu} + \partial_{\rho}\partial_{\sigma}h_{\mu\nu} - \partial_{\sigma}\partial_{\nu}h_{\rho\mu} - \partial_{\sigma}\partial_{\rho}h_{\mu\nu} \\ &\quad - \partial_{\rho}\partial_{\mu}h_{\nu\sigma} + \partial_{\sigma}\partial_{\mu}h_{\nu\rho}) \\ &= \frac{1}{2}(\partial_{\rho}\partial_{\nu}h_{\sigma\mu} + \partial_{\sigma}\partial_{\mu}h_{\nu\rho} - \partial_{\sigma}\partial_{\nu}h_{\rho\mu} - \partial_{\rho}\partial_{\mu}h_{\nu\sigma}). \end{aligned} \quad (797)$$

Contractons maintenant par rapport aux indices  $\rho\mu$ , on obtient :

$$R_{\mu\nu} = \frac{1}{2}(\partial_\sigma\partial_\nu h_\mu^\sigma + \partial_\sigma\partial_\mu h_\nu^\sigma - \partial_\mu\partial_\nu h_\mu^\mu - \square h_{\mu\nu}), \quad (798)$$

où :

$$\square = -\partial_t^2 + \partial_x^2 + \partial_y^2 + \partial_z^2. \quad (799)$$

Par ailleurs, en notant  $h = h_\mu^\mu$ , on obtient :

$$R_{\mu\nu} = \frac{1}{2}(\partial_\sigma\partial_\nu h_\mu^\sigma + \partial_\sigma\partial_\mu h_\nu^\sigma - \partial_\mu\partial_\nu h - \square h_{\mu\nu}). \quad (800)$$

D'où, en contractant une nouvelle fois en  $\mu\nu$  :

$$R = \partial_\mu\partial_\nu h^{\mu\nu} - \square h. \quad (801)$$

Par conséquent, le tenseur d'Einstein s'écrit :

$$\begin{aligned} G_{\mu\nu} &= R_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}R \\ &= \frac{1}{2}(\partial_\sigma\partial_\nu h_\mu^\sigma + \partial_\sigma\partial_\mu h_\nu^\sigma - \partial_\mu\partial_\nu h - \square h_{\mu\nu} - \eta_{\mu\nu}\partial_\mu\partial_\nu h^{\mu\nu} + \eta_{\mu\nu}\square h). \end{aligned} \quad (802)$$

L'équation du champ simplifiée (nous dirons **équation du champ linéarisée**) est donc :

$$\begin{aligned} &\frac{1}{2}(\partial_\sigma\partial_\nu h_\mu^\sigma + \partial_\sigma\partial_\mu h_\nu^\sigma - \partial_\mu\partial_\nu h - \square h_{\mu\nu} - \eta_{\mu\nu}\partial_\mu\partial_\nu h^{\mu\nu} + \eta_{\mu\nu}\square h) \\ &= 8\pi GT_{\mu\nu}, \end{aligned} \quad (803)$$

où  $T_{\mu\nu}$  est le tenseur d'énergie-impulsion, calculé à l'ordre 0 par rapport à  $h_{\mu\nu}$ .

**EURISTIDE** : Nous avons pratiquement déterminé les équations du champ linéarisées. Mais nous avons un problème à résoudre.

**BEATRIX** : Lequel ?

**EURISTIDE** : Dans les calculs précédents, nous avons considéré que l'espace-temps était la superposition d'un espace-temps plat et d'une perturbation petite curviligne, et nous avons pu constituer cette décomposition en choisissant un système de coordonnées adéquat. Mais cette décomposition espace-temps plat + perturbation n'est pas unique et dépend du système de coordonnées choisi.

Pour résoudre ce problème, et obtenir une détermination unique de la perturbation, il nous faut appliquer une contrainte au système de coordonnées. Cela se fait en choisissant ce qu'on appelle une **jaugé**. Une transformation de coordonnées est en fait appelée une **transformation de jaugé**, et le fait de choisir une jaugé permet de restreindre notre capacité à faire des transformations de jaugé.

**BEATRIX** : Si je comprends bien, ce choix d'une jaugé revient en fait à imposer une contrainte qui fixe le système de coordonnées.

**EURISTIDE** : Oui. On utilise habituellement la **jaugé harmonique**, qui applique la contrainte suivante :

$$\square x^\mu = 0, \quad (804)$$

où  $\square$  est le d'alembertien covariant, défini par :

$$\square = \nabla^\mu \nabla_\mu. \quad (805)$$

MATHINE : Cette jauge harmonique s'écrit :

$$\begin{aligned} 0 &= \square x^\mu \\ &= \nabla^\rho (\nabla_\rho x^\mu) \\ &= \nabla^\rho (\partial_\rho x^\mu + \Gamma_{\rho\sigma}^\mu x^\sigma) \\ &= g^{\rho\sigma} \nabla_\sigma (\partial_\rho x^\mu + \Gamma_{\rho\sigma}^\mu x^\sigma) \\ &= g^{\rho\sigma} (\partial_\rho \partial_\sigma x^\mu - \Gamma_{\rho\sigma}^\lambda \partial_\lambda x^\mu \\ &\quad + \partial_\sigma \Gamma_{\rho\sigma}^\mu x^\rho - \Gamma_{\rho\sigma}^\tau \Gamma_{\tau\sigma}^\mu x^\rho). \end{aligned} \quad (806)$$

On néglige toutes les expressions d'ordre 2, donc les produits des symboles de Christoffel entre eux et leurs dérivées. Il reste :

$$0 = g^{\sigma\rho} \partial_\sigma \partial_\rho x^\mu - g^{\sigma\rho} \Gamma_{\sigma\rho}^\lambda \partial_\lambda x^\mu. \quad (807)$$

Le premier terme est négligé, puisqu'il constitue une expression d'ordre 2. Le deuxième terme est réduit à l'expression où  $\lambda = \mu$ , d'où :

$$g^{\sigma\rho} \Gamma_{\sigma\rho}^\mu = 0. \quad (808)$$

Cette expression est la contrainte liée à la jauge harmonique. Cette jauge est également appelée **jauge de Lorentz** ou **jauge d'Einstein**.

Dans la limite du champ faible, cette expression s'écrit :

$$\frac{1}{2} \eta^{\sigma\rho} \eta^{\lambda\mu} (\partial_\sigma h_{\rho\lambda} + \partial_\rho h_{\lambda\sigma} - \partial_\lambda h_{\sigma\rho}) = 0, \quad (809)$$

ce qui s'écrit :

$$\frac{1}{2} \eta^{\lambda\mu} \partial_\rho h_\lambda^\rho + \frac{1}{2} \eta^{\lambda\mu} \partial_\rho h_\lambda^\rho - \frac{1}{2} \eta^{\lambda\mu} \partial_\lambda h = 0, \quad (810)$$

soit :

$$\eta^{\lambda\mu} \partial_\rho h_\lambda^\rho - \frac{1}{2} \eta^{\lambda\mu} \partial_\lambda h = 0, \quad (811)$$

ou encore :

$$\partial_\rho h_\lambda^\rho - \frac{1}{2} \partial_\lambda h = 0. \quad (812)$$

Revenons maintenant, en utilisant ce résultat, les équations linéarisées du champ :

$$\begin{aligned} &\partial_\sigma \partial_\nu h_\mu^\sigma + \partial_\sigma \partial_\mu h_\nu^\sigma - \partial_\mu \partial_\nu h - \square h_{\mu\nu} - \eta_{\mu\nu} \partial_\mu \partial_\nu h^{\mu\nu} + \eta_{\mu\nu} \square h \\ &= \frac{1}{2} \partial_\nu \partial_\mu h + \frac{1}{2} \partial_\mu \partial_\nu h - \partial_\mu \partial_\nu h - \square h_{\mu\nu} - \eta_{\mu\nu} \partial_\mu \partial_\nu h^{\mu\nu} + \eta_{\mu\nu} \square h \\ &= -\square h_{\mu\nu} - \eta_{\mu\nu} \partial_\mu \partial_\nu h^{\mu\nu} + \eta_{\mu\nu} \square h. \end{aligned} \quad (813)$$

D'où l'expression simplifiée des équations linéarisées du champ :

$$\square h_{\mu\nu} + \eta_{\mu\nu} \partial_\mu \partial_\nu h^{\mu\nu} - \eta_{\mu\nu} \square h = -16\pi G T_{\mu\nu}. \quad (814)$$

Dans le vide, on a :

$$R_{\mu\nu} = 0, \quad (815)$$

ce qui s'écrit :

$$\partial_\sigma \partial_\nu h_\mu^\sigma + \partial_\sigma h_\mu h_\nu^\sigma - \partial_\mu \partial_\nu h - \square h_{\mu\nu} = 0, \quad (816)$$

ou encore, en utilisant la jauge d'Einstein, sur la formule de contrainte de laquelle on applique une dérivée partielle, on obtient :

$$\square h_{\mu\nu} = 0. \quad (817)$$

BEATRIX : Les équations d'Einstein linéarisées ont maintenant une allure assez synthétique et élégante, n'est-ce pas ?

MATHINE : Nous allons pouvoir simplifier encore. Pour cela, nous allons considérer la perturbation suivante :

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2} \eta_{\mu\nu} h. \quad (818)$$

Alors, si nous réécrivons la condition de jauge d'Einstein :

$$\partial_\mu h_\lambda^\mu - \frac{1}{2} \partial_\lambda h = 0, \quad (819)$$

nous pouvons l'écrire maintenant sous la forme :

$$\partial_\mu (\bar{h}_\lambda^\mu + \frac{1}{2} \eta^{\nu\mu} \eta_{\lambda\nu} h) - \frac{1}{2} \partial_\lambda h = 0, \quad (820)$$

ou encore :

$$\partial_\mu \bar{h}_\lambda^\mu + \frac{1}{2} \partial_\mu h - \frac{1}{2} \partial_\mu h = 0, \quad (821)$$

soit finalement :

$$\partial_\mu \bar{h}_\lambda^\mu = 0. \quad (822)$$

Et les équations du champ d'Einstein linéarisées s'écrivent directement :

$$\square \bar{h}_{\mu\nu} = -16\pi G T_{\mu\nu}, \quad (823)$$

et enfin, les équations du vide s'écrivent :

$$\square \bar{h}_{\mu\nu} = 0. \quad (824)$$

EURISTIDE : Interprétons ces équations. Elles signifient que, pour un champ faible, nous pouvons remplacer le tenseur d'Einstein par le d'alembertien covariant de la perturbation de la métrique dans les équations du champ. C'est-à-dire que la courbure de l'espace-temps, dans cette approximation du champ faible, est déterminée uniquement par la perturbation de la métrique. Cette perturbation agit par son d'alembertien covariant, qui représente la somme des dérivées secondes covariantes selon chacun des axes du système de coordonnées.

Maintenant que nous avons ce résultat, nous allons considérer le cas particulier d'une planète ou d'une étoile. Pour un tel système, la composante  $T_{00}$  du tenseur d'énergie-impulsion représente la densité d'énergie de la planète ou de l'étoile et est beaucoup plus grande que les autres composantes qui représentent les flux d'énergie et les flux d'impulsion des corps massifs.

Si toutes les composantes  $T_{\mu\nu}$  différentes de  $T_{00} = \rho$  (densité d'énergie) sont négligeables, alors l'équation

du champ indique qu'il doit en être de même pour les composantes  $\bar{h}_{\mu\nu}$ , par rapport à la composante  $\bar{h}_{00}$ . Or, nous avons vu tout-à-l'heure que :

$$h_{00} = -2\Phi, \quad (825)$$

où  $\Phi$  est le potentiel gravitationnel de Newton :

$$\Phi = -\frac{GM}{r}. \quad (826)$$

Par ailleurs,  $\bar{h}_{00}$  est beaucoup plus grand que les autres composantes, donc :

$$h = -\bar{h} = -\eta^{\mu\nu}\bar{h}_{\mu\nu} = \bar{h}_{00}. \quad (827)$$

Il s'ensuit que :

$$\bar{h}_{00} = 2h_{00} = -4\Phi. \quad (828)$$

D'où :

$$h_{i0} = \bar{h}_{i0} + \frac{1}{2}\eta_{i0}\bar{h} \approx 0, \quad (829)$$

et :

$$h_{ij} = \bar{h}_{ij} - \frac{1}{2}\eta_{ij}\bar{h} \approx -2\Phi\delta_{ij}. \quad (830)$$

Par conséquent, la métrique, pour une étoile ou une planète dans la limite du champ faible est constituée de la somme de la métrique plate et de la perturbation que nous venons de calculer :

$$ds^2 = -(1 + 2\Phi)dt^2 + (1 - 2\Phi)(dx^2 + dy^2 + dz^2). \quad (831)$$

**BEATRIX** : Evidemment, lorsqu'on fait  $\Phi = 0$ , c'est-à-dire lorsque le potentiel gravitationnel est nul, on retrouve la métrique de l'espace-temps plat. Un univers sans gravitation est plat.

**EURISTIDE** : Injectons dans cette expression de la métrique, la valeur du potentiel gravitationnel :

$$\Phi = -\frac{Gm}{r}. \quad (832)$$

Alors, la métrique s'écrit :

$$ds^2 = -\left(1 - \frac{2Gm}{r}\right)dt^2 - \left(1 + \frac{2Gm}{r}\right)(dx^2 + dy^2 + dz^2). \quad (833)$$

Le rayon :

$$R_S = 2Gm, \quad (834)$$

est appelé **rayon de Schwarzschild**. Nous verrons la signification et les conséquences de ce rayon plus généralement dans l'étude des solutions de Schwarzschild pour le trou noir.

## 82 Scène XIII-2 : Les ondes gravitationnelles

**BEATRIX** : Il apparaît que pour une distance égale au rayon de Schwarzschild, la métrique ne comporte plus de terme temporel en  $dt$ . Qu'est-ce que cela signifie ?



EURISTIDE : Cela signifie qu'en ces points particuliers (qui constituent géométriquement un sphère de rayon  $R_S$  autour du centre du champ gravitationnel), l'espace-temps présente une particularité. On appelle cela une **singularité**. Nous étudierons plus tard ces singularités. Et en particulier, nous verrons que certaines singularités de l'espace-temps sont réelles et correspondent à des phénomènes physiques, et nous verrons aussi que certaines singularités sont factices, et ne sont créées que par le choix du système de coordonnées que nous avons fait. On peut alors faire disparaître ces singularités factices par un simple changement de coordonnées. C'est le phénomène géométrique que nous avons étudié lorsque nous avons vu les variétés différentielles. Le choix de la carte peut conduire à une singularité, par exemple les pôles de la Terre lorsqu'on fait une projection de Mercator de la surface du globe Terrestre. En effet, les pôles ne sont pas représentés dans cette projection. C'est donc une singularité, qui disparaît lorsque nous choisissons une autre carte.

Nous allons maintenant étudier le phénomène particulier que sont les ondes gravitationnelles.

BEATRIX : La gravitation est une onde ?

EURISTIDE : On peut considérer dans certains cas que la gravitation se comporte comme une onde. Ce sont les équations d'Einstein linéarisées qui le suggèrent, car on peut leur trouver une solution sous la forme d'une fonction ondulatoire.

MATHINE : En effet, dans un espace-temps vide, nous avons vu que les équations du champ s'écrivent :

$$\square \bar{h}_{\mu\nu} = 0. \quad (835)$$

Cette équation faisant intervenir le d'Alembertien, présente une analogie immédiate, pour les mathématiciens et les physiciens, avec les équations d'onde de l'électromagnétisme. C'est pourquoi, parmi les solutions possibles à ces équations, nous savons qu'elle admettent pour solution une onde plane du type :

$$\bar{h}_{\mu\nu} = A_{\mu\nu} \cos(k_\alpha x^\alpha), \quad (836)$$

où  $A_{\mu\nu}$  est un tenseur symétrique constant de rang 2 et  $k_\alpha$  est un vecteur d'onde constant.

EURISTIDE : Le vecteur d'onde donne la direction dans laquelle l'onde se propage.

MATHINE : Pour que les équations du champ soient vérifiées, il faut que :

$$\begin{aligned} 0 &= \square \bar{h}_{\mu\nu} & (837) \\ &= \eta^{\rho\sigma} \partial_\rho \partial_\sigma \bar{h}_{\mu\nu} \\ &= -\eta^{\rho\sigma} \partial_\rho (k_\sigma A_{\mu\nu} \sin(k_\alpha x^\alpha)) \\ &= -\eta^{\rho\sigma} k_\sigma k_\rho A_{\mu\nu} \cos(k_\alpha x^\alpha) \\ &= -k_\sigma k^\sigma \bar{h}_{\mu\nu}. \end{aligned}$$

Or, les composantes  $\bar{h}_{\mu\nu}$  ne sont a priori pas nulles partout, donc on a nécessairement :

$$k_\sigma k^\sigma = 0. \quad (838)$$

Donc, l'onde plane est une solution des équations du champ, à condition que le vecteur d'onde soit de genre lumière (de module nul).

EURISTIDE : Ceci, concrètement, signifie que l'onde de gravitation se propage à la vitesse de la lumière.

BEATRIX : Comme les ondes électromagnétiques, alors ! Justement, celles-ci produisent des effets physiques sur la matière. En est-il de même pour les ondes gravitationnelles ? Je suppose que oui.

EURISTIDE : En effet. Pour l'établir, nous allons considérer des particules soumises à des ondes gravitationnelles et calculer leur déplacement. Nous allons analyser le mouvement relatif de particules voisines, en travaillant avec l'équation de déviation géodésique, qui décrit la séparation de deux particules.

MATHINE : Cela donne :

$$\frac{d^2}{d\tau^2} S^\mu = R^\mu_{\nu\rho\sigma} U^\nu U^\rho S^\sigma, \quad (839)$$

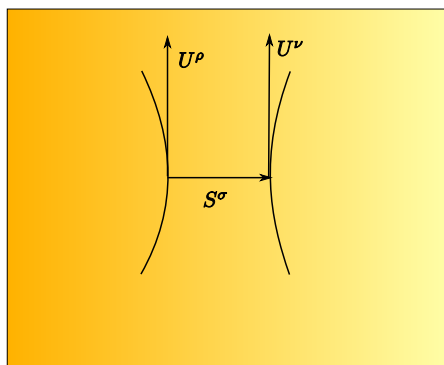


Fig. 136 - La déviation géodésique

où  $\tau$  est le temps propre de la particule, où  $S^\mu$  est le vecteur de séparation des deux particules, c'est-à-dire le vecteur passant de la particule 1 à la particule 2, où  $U^\rho$  est un vecteur unitaire tangent à la trajectoire de la particule 1, et où  $U^\nu$  est un vecteur unitaire tangent à la trajectoire de la particule 2. En effet, on a, par définition du tenseur de Riemann :

$$R^\mu_{\nu\rho\sigma} U^\nu = \nabla_\rho \nabla_\sigma U^\mu - \nabla_\sigma \nabla_\rho U^\mu, \quad (840)$$

soit :

$$\begin{aligned} R^\mu_{\nu\rho\sigma} U^\nu U^\rho U^\sigma &= \nabla_\rho \nabla_\sigma U^\mu U^\rho S^\sigma - \nabla_\sigma \nabla_\rho U^\mu U^\rho S^\sigma \\ &= \frac{d^2}{d\tau^2} S^\mu. \end{aligned} \quad (841)$$

Or, les particules se déplacent lentement, donc le vecteur  $U^\nu$  peut être considéré comme uniquement temporel, soit :

$$U^\nu \approx (1, 0, 0, 0). \quad (842)$$

Donc, seul le terme  $R_{00\sigma}^\mu$  est à calculer, puisque les autres termes sont nuls.

Or, nous savons que :

$$R_{\mu\nu\rho\sigma} = \frac{1}{2}(\partial_\rho\partial_\nu h_{\mu\sigma} + \partial_\sigma\partial_\mu h_{\nu\rho} - \partial_\sigma\partial_\nu h_{\mu\rho} - \partial_\rho\partial_\mu h_{\nu\sigma}), \quad (843)$$

d'où :

$$R_{\mu 00\sigma} = \frac{1}{2}(\partial_0\partial_0 h_{\mu\sigma} + \partial_\sigma\partial_\mu h_{00} - \partial_\sigma\partial_0 h_{\mu 0} - \partial_0\partial_\mu h_{0\sigma}). \quad (844)$$

Or, nous savons que les  $h_{\mu 0}$  et  $h_{0\mu}$  sont nuls, donc :

$$R_{\mu 00\sigma} = \frac{1}{2}\partial_0\partial_0 h_{\mu\sigma} + \frac{1}{2}\partial_\sigma\partial_\mu h_{00}. \quad (845)$$

Le terme en  $\partial_\sigma\partial_{\mu\mu} h_{00}$  est nul.

Donc, finalement :

$$R_{00\sigma}^\mu = \frac{1}{2}\partial_0\partial_0 h_\sigma^\mu. \quad (846)$$

Nous pouvons considérer, dans l'approximation de particules qui se déplacent lentement, que :

$$\tau = t. \quad (847)$$

Alors, l'équation de déviation géodésique devient :

$$\frac{\partial^2}{\partial t^2} S^\mu = \frac{1}{2} S^\sigma \frac{\partial^2}{\partial t^2} h_\sigma^\mu. \quad (848)$$

Par conséquent, pour une onde se déplaçant dans la direction  $x^3$ , par exemple, seuls les termes  $S^1$  et  $S^2$  seront non nuls.

On montre, par des calculs assez longs, que dans ces conditions, en choisissant une jauge adéquate, les coefficients de l'équation d'onde sont réduits à la matrice suivante :

$$A_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & A_{11} & A_{12} & 0 \\ 0 & A_{12} & -A_{11} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (849)$$

Par conséquent, l'onde est caractérisée par les deux nombres  $A_{11}$  et  $A_{12}$ . Supposons d'abord que  $A_{12} = 0$ , alors :

$$\frac{\partial^2}{\partial t^2} S^1 = \frac{1}{2} S^1 \frac{\partial^2}{\partial t^2} (A_{11} \cos(k_\alpha x^\alpha)), \quad (850)$$

et :

$$\frac{\partial^2}{\partial t^2} S^2 = -\frac{1}{2} S^2 \frac{\partial^2}{\partial t^2} (A_{11} \cos(k_\alpha x^\alpha)), \quad (851)$$

d'où l'on déduit, par intégration, que :

$$\begin{aligned} S^1 &= \left(1 + \frac{1}{2} A_{11} \cos(k_\alpha x^\alpha)\right) S^1(0) \\ S^2 &= \left(1 - \frac{1}{2} A_{11} \cos(k_\alpha x^\alpha)\right) S^2(0) \end{aligned} \quad (852)$$

EURISTIDE : Dans ce cas de figure, deux particules initialement séparées dans la direction  $x^1$  se mettent à danser en avant et en arrière dans la direction  $x^1$ .

De même, deux particules initialement séparées dans la direction  $x^2$  se mettent à danser en avant et en arrière dans la direction  $x^2$ .

On peut illustrer physiquement ce que cela donnera en imaginant des particules disposées sur un cercle. Le passage de l'onde gravitationnelle provoque l'oscillation de ce cercle, qui se déforme en ellipse contractée tour à tour dans les directions  $x^1$  et  $x^2$  :

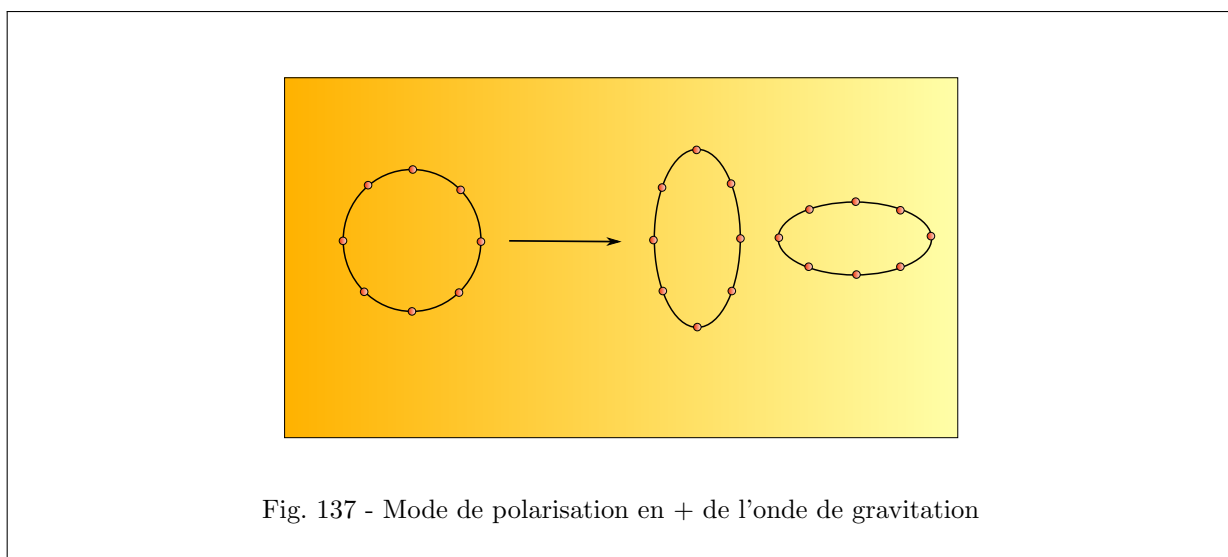


Fig. 137 - Mode de polarisation en + de l'onde de gravitation

BEATRIX : Les particules se mettent à danser en dessinant un signe plus !

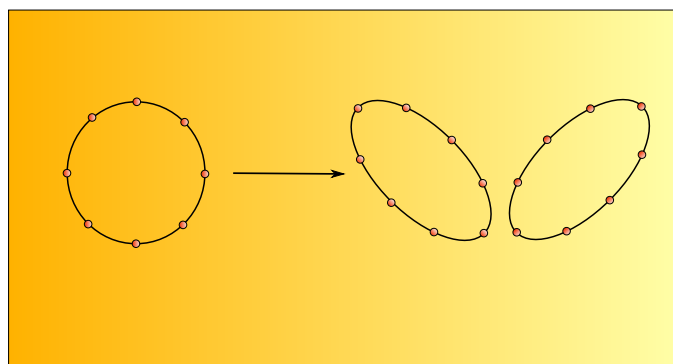
MATHINE : Supposons maintenant que  $A_{11} = 0$ . Alors :

$$\begin{aligned}\frac{\partial^2}{\partial t^2} S^1 &= \frac{1}{2} S^2 \frac{\partial^2}{\partial t^2} (A_{12} \cos(k_\alpha x^\alpha)) \\ \frac{\partial^2}{\partial t^2} S^2 &= \frac{1}{2} S^1 \frac{\partial^2}{\partial t^2} (A_{12} \cos(k_\alpha x^\alpha)),\end{aligned}\quad (853)$$

d'où l'on déduit, par intégration, que :

$$\begin{aligned}S^1 &= S^1(0) + \frac{1}{2} A_{12} \cos(k_\alpha x^\alpha) S^2(0) \\ S^2 &= S^2(0) + \frac{1}{2} A_{12} \cos(k_\alpha x^\alpha) S^1(0)\end{aligned}\quad (854)$$

EURISTIDE : Dans ce cas de figure, les particules disposées sur un cercle se déforment dans les directions diagonales :

Fig. 138 - Mode de polarisation en  $\times$  de l'onde de gravitation

BEATRIX : Les particules se mettent à danser en dessinant un signe multiplié!

EURISTIDE : Nous avons donc déterminé qu'une onde gravitationnelle possède deux modes de polarisation en  $+$  et en  $\times$ .

BEATRIX : A-t-on déjà observé de tels effets dus à l'onde gravitationnelle ?

EURISTIDE : Nous allons étudier maintenant le rayonnement gravitationnel émis par une source. C'est dans ce cadre que nous pourrons retrouver une prédiction connue relative au rayonnement gravitationnel émis par une étoile binaire.

MATHINE : Il s'agit donc maintenant d'établir la relation qui existe entre le rayonnement gravitationnel, représenté par le tenseur  $\bar{h}_{\mu\nu}$  et sa source, représentée par le tenseur d'énergie-impulsion  $T_{\mu\nu}$ . Par similitude avec l'électromagnétisme, on peut écrire la solution de l'équation du champ, dans l'hypothèse où les vitesses sont faibles, sous la forme :

$$\bar{h}_{\mu\nu}(t, \vec{r}) = \frac{4G}{r} \int T_{\mu\nu}(t - r, \vec{r}') dV, \quad (855)$$

où  $\vec{r}$  est le vecteur distance (en 3 dimensions) jusqu'au centre de la source de rayonnement gravitationnel, et où  $r$  est son module. Nous admettrons ce résultat qui peut être retrouvé dans les développements de la théorie des champs électromagnétiques. Le rayonnement gravitationnel est considéré dans une zone éloignée de la source. Par conséquent, le rayonnement est assimilable à une onde plane, et seule la partie spatiale de

$\bar{h}_{\mu\nu}$  est non nulle.

Il s'ensuit que nous ne considérerons que le terme spatial de l'intégrale.

Le principe de conservation de l'énergie-impulsion s'écrit :

$$\nabla_\nu T^{\mu\nu} = 0, \quad (856)$$

ce qui s'écrit, en décomposant les équations en partie temporelle et parties spatiales :

$$\begin{aligned} \frac{\partial T^{00}}{\partial x^0} + \frac{\partial T^{0k}}{\partial x^k} &= 0 \\ \frac{\partial T^{i0}}{\partial x^0} + \frac{\partial T^{ik}}{\partial x^k} &= 0. \end{aligned} \quad (857)$$

Par ailleurs, nous pouvons écrire, en intégrant par parties :

$$\int \frac{\partial}{\partial x^k} (T^{ik} x^j) dV = \int \frac{\partial}{\partial x^k} T^{ik} x^j dV + \int T^{ij} dV. \quad (858)$$

Or le premier membre de cette égalité est l'intégrale d'une divergence d'un tenseur sur tout le volume de l'espace. D'après le théorème de Gauss, cette intégrale est égale à une intégrale de surface, qui s'annule :

$$\int \frac{\partial}{\partial x^k} (T^{ik} x^j) dV = 0. \quad (859)$$

Par conséquent :

$$\int T^{ij} dV = - \int \frac{\partial}{\partial x^k} T^{ik} x^j dV \quad (860)$$

Or nous avons vu plus haut que :

$$\frac{\partial T^{ik}}{\partial x^k} + \frac{\partial T^{i0}}{\partial x^0} = 0. \quad (861)$$

Donc :

$$\int T^{ij} dV = \int \frac{\partial T^{i0}}{\partial x^0} x^j dV = \frac{d}{dt} \int T^{i0} x^j dV. \quad (862)$$

En interchangeant les indices  $i$  et  $j$ , et en utilisant la symétrie de  $T^{ij}$ , on a :

$$\int T^{ij} dV = \frac{d}{dt} \int T^{j0} x^i dV. \quad (863)$$

D'où, par sommation de l'équation d'origine et de son symétrique et par division par 2, nous obtenons :

$$\int T^{ij} dV = \frac{1}{2} \frac{d}{dt} \int (T^{i0} x^j + T^{j0} x^i) dV. \quad (864)$$

Mais, par ailleurs, nous savons que, au moyen d'une intégration par parties :

$$\int \frac{\partial}{\partial x^k} (T^{0k} x^i x^j) dV = \int \frac{\partial}{\partial x^k} T^{0k} x^i x^j + \int (T^{0i} x^j + T^{0j} x^i) dV. \quad (865)$$

En appliquant de nouveau le théorème de Gauss, nous pouvons exprimer le premier membre de l'égalité, qui est l'intégrale d'une divergence sur l'espace tout entier, sous la forme d'une intégrale de surface qui peut s'annuler :

$$\int \frac{\partial}{\partial x^k} (T^{0k} x^i x^j) dV = 0. \quad (866)$$

Donc :

$$\int (T^{0i} x^j + T^{0j} x^i) dV = - \int \frac{\partial}{\partial x^k} T^{0k} x^i x^j dV. \quad (867)$$

Or, de nouveau :

$$\frac{\partial}{\partial x^k} T^{0k} = -\frac{\partial}{\partial x^0} T^{00}, \quad (868)$$

donc :

$$\int (T^{0i} x^j + T^{0j} x^i) dV = \frac{d}{dt} \int T^{00} x^i x^j dV. \quad (869)$$

Par hypothèse, les particules se déplacent lentement, donc :

$$T^{00} \approx \rho, \quad (870)$$

où  $\rho$  est la densité de matière propre.

Donc, en intégrant cette approximation, et en revenant à l'expression :

$$\int T^{ij} dV = \frac{1}{2} \frac{d}{dt} \int (T^{i0} x^j + T^{j0} x^i) dV, \quad (871)$$

on obtient :

$$\int T^{ij} dV = \frac{1}{2} \frac{d^2}{dt^2} \left( \int \rho x^i x^j dV \right), \quad (872)$$

d'où :

$$\bar{h}^{ij} = \frac{2G}{r} \frac{d^2}{dt^2} \left( \int \rho x^i x^j dV \right)_{t'=t-r}. \quad (873)$$

On définit l'**impulsion quadripôle** de la source par l'expression :

$$\bar{q}^{ij} = \int \rho x^i x^j dV. \quad (874)$$

Par conséquent, la solution au rayonnement gravitationnel émis par une source d'impulsion quadripôle  $\bar{q}^{ij}$  est :

$$\bar{h}_{ij}(t, \vec{r}') = \frac{2G}{r} \frac{d^2}{dt^2} \bar{q}_{ij}. \quad (875)$$

**EURISTIDE** : Il faut comprendre l'impulsion quadripôle, comme la somme sur tout le volume spatial, des produits des masses de chaque particule (masses représentées par la densité de matière  $\rho$ ) par tous les couples de produits de vecteurs de coordonnées spatiales possibles. Ayant obtenu cette somme, qui est une masse multipliée par le carré d'une longueur, on en prend la dérivée seconde temporelle. Cela revient à calculer l'équivalent d'une impulsion comme on l'entend dans la théorie newtonienne, qui est le produit d'une masse avec le carré d'une vitesse. C'est donc cette "impulsion volumique" qui représente l'impact de la matière dans la génération du rayonnement gravitationnel représenté par le  $\bar{h}_{ij}(t, \vec{r}')$ .

**BEATRIX** : On retrouve la loi générale de la matière, ou plutôt ici de son impulsion, qui provoque une courbure de l'espace-temps, qui en l'occurrence est une onde de courbure. Avons-nous, comme vous l'aviez suggéré tout à l'heure, un exemple réel de tels effets ?

**EURISTIDE** : Oui, nous allons pouvoir appliquer cette formule au cas d'une étoile binaire. Il existe dans le ciel des objets qui sont des couples d'étoiles en rotation l'une par rapport à de l'autre. Par conséquent, nous allons modéliser une étoile binaire comme étant un ensemble de deux corps de masse  $M$  tournant sur une orbite circulaire de rayon  $R$  dans le plan  $xy$ . Et nous allons étudier le rayonnement gravitationnel de cet ensemble en un point  $P$  à une distance  $r$  de leur centre de rotation commun.

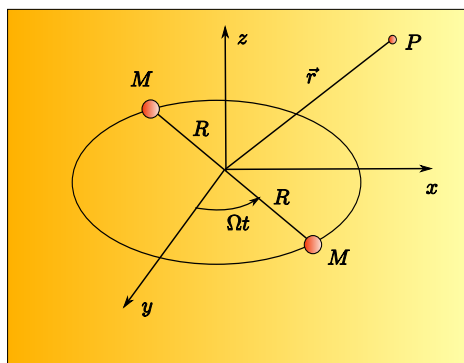


Fig. 139 - Modèle de l'étoile binaire

Comme nous pouvons considérer que ces deux étoiles en rotation ont un mouvement lent, donc non relativiste, et comme les masses sont faibles, nous pouvons appliquer les lois de la mécanique newtonienne. Par conséquent, les deux étoiles subissent une force d'attraction gravitationnelle d'intensité :

$$F = \frac{GM^2}{(2R)^2}. \quad (876)$$

Et par ailleurs, ces deux étoiles sont soumises à la force centrifuge d'intensité :

$$F' = \frac{Mv^2}{R}, \quad (877)$$

où  $v$  est la vitesse linéaire des étoiles sur leur orbite.

Ces deux forces sont égales, d'où :

$$\frac{GM^2}{(2R)^2} = \frac{Mv^2}{R}, \quad (878)$$

donc :

$$v = \sqrt{\frac{GM}{4R}}. \quad (879)$$

Par ailleurs, la durée d'une orbite complète est :

$$T = \frac{2\pi R}{v}, \quad (880)$$

et par conséquent, la vitesse angulaire étant :

$$\Omega = \frac{2\pi}{T}, \quad (881)$$

nous avons :

$$\Omega = \sqrt{\frac{GM}{4R^3}}. \quad (882)$$



Nous allons maintenant calculer la densité de masse du système constitué de ces deux étoiles en orbite. Pour calculer cette densité de masse, nous allons introduire la fonction  $\delta$  définie par :

$$\delta(x) = \begin{cases} 1 & \text{si } x = 0, \\ 0 & \text{sinon.} \end{cases} \quad (883)$$

Avec cette définition, la densité de masse est  $M$  lorsque nous la relevons au point  $\vec{r}$  et à l'instant  $t$  tels que, d'une part :

$$z = 0, \quad (884)$$

et, d'autre part (pour la première étoile) :

$$\begin{cases} x = R \cos \Omega t \\ y = R \sin \Omega t, \end{cases} \quad (885)$$

ou bien (pour la deuxième étoile) :

$$\begin{cases} x = -R \cos \Omega t \\ y = -R \sin \Omega t. \end{cases} \quad (886)$$

Par conséquent, la densité de masse peut être écrite :

$$\begin{aligned} \rho(t, \vec{r}) &= M\delta(z) \times (\delta(x - R \cos \Omega t)\delta(y - R \sin \Omega t) \\ &\quad + \delta(x + R \cos \Omega t)\delta(y + R \sin \Omega t)). \end{aligned} \quad (887)$$

Le moment du quadripôle s'écrit :

$$\bar{q}^{ij} = \int \rho dx^i dx^j dV. \quad (888)$$

Les composantes en  $z$ , c'est-à-dire pour lesquelles  $z \neq 0$ , sont évidemment nulles, puisque  $\delta(z) = 0$  pour  $z \neq 0$ .

seules les composantes  $\bar{q}_{xx}$ ,  $\bar{q}_{yy}$ ,  $\bar{q}_{xy}$  et  $\bar{q}_{yx}$  sont non nulles.

Pour la composante  $\bar{q}_{xx}$ , seuls les termes en :

$$x = R \cos \Omega t, \quad (889)$$

ou en :

$$x = -R \cos \Omega t, \quad (890)$$

sont non nuls, et pour ces termes  $\delta(z) = 1$ , et :

$$\delta(x - R \cos \Omega t)\delta(y - R \sin \Omega t) = 1 \quad (891)$$

ou :

$$\delta(x + R \cos \Omega t)\delta(y + R \sin \Omega t) = 1 \quad (892)$$

respectivement. Donc :

$$\bar{q}_{xx} = 2MR^2 \cos^2 \Omega t. \quad (893)$$

Soit :

$$\bar{q}_{xx} = MR^2(1 + \cos 2\Omega t). \quad (894)$$

De même, suivant un raisonnement analogue :

$$\bar{q}_{yy} = MR^2(1 - \cos 2\Omega t). \quad (895)$$

Et enfin, toujours en raisonnant de façon similaire :

$$\bar{q}_{xy} = \bar{q}_{yx} = 2MR^2 \cos \Omega t \sin \Omega t, \quad (896)$$

soit :

$$\bar{q}_{xy} = \bar{q}_{yx} = MR^2 \sin 2\Omega t. \quad (897)$$

Injectons ces résultats dans l'équation de l'onde gravitationnelle :

$$\bar{h}_{ij}(t, \vec{r}) = \frac{2G}{r} \frac{d^2}{dt^2} \begin{pmatrix} MR^2(1 + \cos 2\Omega t) & MR^2 \sin 2\Omega t & 0 \\ MR^2 \sin 2\Omega t & MR^2(1 - \cos 2\Omega t) & 0 \\ 0 & 0 & 0 \end{pmatrix}_{t-r}, \quad (898)$$

soit :

$$\bar{h}_{ij}(t, \vec{r}) = \frac{2G}{r} \frac{d}{dt} \begin{pmatrix} -MR^2 \times 2\Omega \sin 2\Omega t & MR^2 \times 2\Omega \cos 2\Omega t & 0 \\ MR^2 \times 2\Omega \cos 2\Omega t & MR^2 \times 2\Omega \sin 2\Omega t & 0 \\ 0 & 0 & 0 \end{pmatrix}_{t-r}, \quad (899)$$

ou encore :

$$\bar{h}_{ij}(t, \vec{r}) = \frac{2G}{r} \begin{pmatrix} -4MR^2\Omega^2 \cos 2\Omega t & -4MR^2\Omega^2 \sin 2\Omega t & 0 \\ -4MR^2\Omega^2 \sin 2\Omega t & 4MR^2\Omega^2 \cos 2\Omega t & 0 \\ 0 & 0 & 0 \end{pmatrix}_{t-r}. \quad (900)$$

Soit finalement :

$$\bar{h}_{ij}(t, \vec{r}) = \frac{8GM\Omega^2 R^2}{r} \begin{pmatrix} -\cos(2\Omega(t-r)) & -\sin(2\Omega(t-r)) & 0 \\ -\sin(2\Omega(t-r)) & \cos(2\Omega(t-r)) & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (901)$$

**EURISTIDE** : Ainsi, on voit que la fréquence de l'onde gravitationnelle est égale au double de la fréquence orbitale de l'étoile double.

Nous allons maintenant évaluer la puissance du rayonnement gravitationnel.

**MATHINE** : Les calculs sont complexes et font appels à des notions que nous n'avons pas abordées. Nous allons donc admettre que la puissance totale rayonnée est :

$$P(t, r) = \frac{G}{5} \left( \frac{d^3}{dt^3} Q_{ij} \frac{d^3}{dt^3} Q^{ij} \right)_{t-r}, \quad (902)$$

où  $Q_{ij}$  est l'impulsion quadripôle, donnée par :

$$Q_{ij} = \int \rho(x_i x_j - \frac{1}{3} \delta_{ij} r^2) = q_{ij} - \frac{1}{3} \delta_{ij} q_k^k. \quad (903)$$

En appliquant ces formules au cas de l'étoile double, nous avons :

$$Q_{ij} = \frac{1}{3} MR^2 \begin{pmatrix} 1 + 3 \cos 2\Omega t & 3 \sin 2\Omega t & 0 \\ 3 \sin 2\Omega t & 1 - 3 \cos 2\Omega t & 0 \\ 0 & 0 & -2 \end{pmatrix}, \quad (904)$$

et :

$$\frac{d^3 Q^{ij}}{dt^3} = 8MR^2 \Omega^3 \begin{pmatrix} \sin 2\Omega t & -\cos 2\Omega t & 0 \\ -\cos 2\Omega t & -\sin 2\Omega t & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (905)$$

Alors, la puissance rayonnée par l'étoile double est :

$$\begin{aligned} P &= \frac{128}{5} GM^2 R^4 \Omega^6 \\ &= \frac{2G^4 M^5}{5R^5} \\ &= \frac{128}{5} 4^{\frac{1}{3}} \frac{1}{G} \left( \frac{\pi GM}{T} \right)^{\frac{10}{3}}. \end{aligned} \quad (906)$$

EURISTIDE : L'effet du rayonnement gravitationnel émis par une étoile double a été observé pour un pulsar binaire PSR B1913+16. Le rayonnement émis extrait de l'énergie du système, et par conséquent diminue sa période de rotation et donc celle du rayonnement gravitationnel.

Nous allons maintenant calculer ce taux de décroissance de la période.

MATHINE : Pour cela, considérons l'énergie du système :

$$E = 2 \left( \frac{1}{2} Mv^2 \right) - \frac{GM^2}{2R}. \quad (907)$$

Or, nous savons que :

$$v = \sqrt{\frac{GM}{4R}} \quad (908)$$

et que :

$$\frac{2\pi}{T} = \sqrt{\frac{GM}{4R^3}}. \quad (909)$$

Donc :

$$E = -\frac{GM^2}{4R} = -\frac{GM}{4} \left( \frac{4\pi M}{TG^{1/2}} \right)^{\frac{2}{3}}. \quad (910)$$

Nous pouvons maintenant dériver E par rapport à t, et assimiler :

$$\frac{dE}{dt} = -P = -\frac{128}{5} 4^{\frac{1}{3}} \frac{1}{G} \left( \frac{\pi GM}{T} \right)^{\frac{10}{3}}. \quad (911)$$

Il s'ensuit que :

$$\frac{dT}{dt} = -\frac{96}{5} \pi 4^{\frac{1}{3}} \left( \frac{2\pi GM}{T} \right)^{\frac{5}{3}}. \quad (912)$$

EURISTIDE : En appliquant les valeurs numériques pour le pulsar considéré PSR B1913+16, on trouve que le taux de diminution de la période est de l'ordre de 10  $\mu s$  par an.

Ce ralentissement de la période du rayonnement gravitationnel a été détecté effectivement sur le pulsar PSR B1913+16.

## 83 Scène XIII-3 : Le champ à symétrie centrale

BEATRIX : Donc nous avons là une belle prédiction de la relativité générale, et plus précisément de la notion de rayonnement gravitationnel. L'avantage de l'étoile double, c'est qu'elle nous donne un moyen de vérifier l'impact de la perte d'énergie du système par rayonnement.

EURISTIDE : Nous allons maintenant étudier le cas plus général des champs à symétrie sphérique dans le vide. Ce peut être notamment le cas lorsque nous nous trouvons assez loin d'une étoile ou d'un trou noir. Cette étude ne prendra pas d'hypothèse sur la faiblesse du champ ou sur la lenteur des particules en mouvement.

BEATRIX : Comme nous sommes dans le vide, cela signifie que nous utilisons les équations d'Einstein dans le vide :

$$R_{\mu\nu} = 0. \quad (913)$$

EURISTIDE : C'est bien cela. Et les solutions de ces équations doivent comporter une symétrie sphérique, c'est-à-dire la symétrie d'une sphère  $S^2$  dans l'espace de dimension 3. Donc la métrique d'une telle solution doit avoir cette symétrie sphérique. Te rappelles-tu comment nous pouvons caractériser la symétrie d'une métrique ?

BEATRIX : Euh... Ce sont les champs de vecteurs de Killing, non ?

EURISTIDE : Bravo! Pour une sphère  $S^2$ , nous pouvons étudier le nombre de champs de vecteurs de Killing. Comme l'espace est à symétrie sphérique, cela signifie que nous pouvons nous déplacer sur la sphère arbitrairement, sans que la métrique ne change. Les déplacements qui ne changent pas la métrique, nous l'avons vu, sont appelés isométries. Il y a trois isométries possibles sur une sphère  $S^2$  : la translation le long des parallèles, la translation le long des méridiens, et la rotation qui transforme la coordonnée méridienne en parallèle, et vice-versa.

BEATRIX : Il y a donc 3 champs de vecteurs de Killing pour notre métrique solution des équations d'Einstein pour un champ à symétrie.

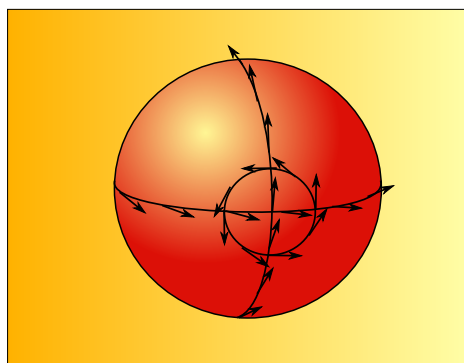


Fig. 140 - Les 3 champs de vecteurs de Killing d'un champ à symétrie sphérique

EURISTIDE : Exact. Nous allons appeler ces trois champs de vecteurs de Killing respectivement :

$$\vec{V}_{(1)}, \vec{V}_{(2)}, \vec{V}_{(3)}. \quad (914)$$

MATHINE : Il se trouve que nous pouvons choisir ces 3 champs de vecteurs de telle manière qu'ils vérifient les relations suivantes :

$$\begin{aligned} [\vec{V}_{(1)}, \vec{V}_{(2)}] &= \vec{V}_{(3)} \\ [\vec{V}_{(2)}, \vec{V}_{(3)}] &= \vec{V}_{(1)} \\ [\vec{V}_{(3)}, \vec{V}_{(1)}] &= \vec{V}_{(2)}, \end{aligned} \quad (915)$$

où nous notons :

$$[\vec{X}, \vec{Y}] = X^\lambda \partial_\lambda Y^\mu - Y^\lambda \partial_\lambda X^\mu. \quad (916)$$

EURISTIDE : Cette expression revient à comparer les transports d'un champ le long d'un autre champ vectoriel. Le champ résultant indique intuitivement la direction du décalage résultant du transport d'un champ le long du deuxième par rapport au transport du deuxième champ le long du premier. Cette direction détermine le troisième champ de vecteurs de Killing quand on effectue l'opération sur les deux premiers champs de vecteurs de Killing. Cela signifie que ces 3 champs de vecteurs de Killing choisis adéquatement ont une relation d'orthogonalité particulière liée au transport le long de leurs lignes de champ.

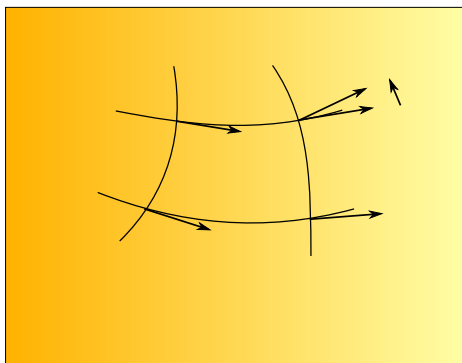


Fig. 141 - Relation d'orthogonalité des vecteurs de Killing

MATHINE : Il existe un théorème, appelé Théorème de Frobenius, qui nous dit que les champs de vecteurs constituent des ensembles de courbes intégrales qui chacun déterminent une sous-variété différentiable. Ces sous-variétés sont elles-mêmes des sphères  $S^2$ . Tout point de l'espace appartient donc à l'une de ces sphères. Cette technique permet de découper l'espace comme un mille-feuilles constitué de feuilles en forme de sphères  $S^2$ . On dit que **l'espace est follié** par les sphères  $S^2$ .

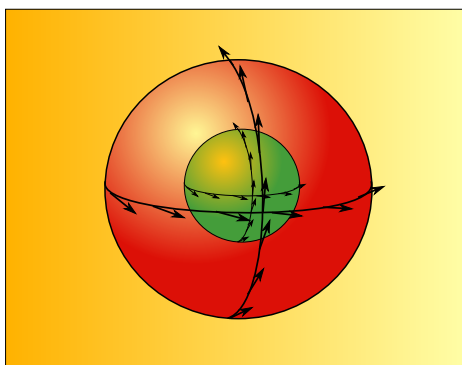


Fig. 142 - Foliation de l'espace

EURISTIDE : Naturellement, nous allons choisir un système de coordonnées cohérent avec cette foliation de l'espace. Par exemple, pour un espace de dimension 3, nous allons choisir, sur chaque sous-variété  $S^2$ , un système de 2 coordonnées  $u^i$  qui positionne un point sur la sous-variété (donc  $i$  varie de 0 à 1), et nous allons choisir un système d'une coordonnées  $v^I$  qui désigne la sous-variété sur laquelle nous travaillons (donc ici  $I$  est toujours égal à 0).

Pour l'espace-temps de dimension 4, nous choisirons le même système de coordonnées  $u^i$  dans les sous-variétés  $S^2$ , et nous devrions choisir un système de 2 coordonnées  $v^I$ , pour désigner la sous-variété (donc  $I$  variant de 0 à 1).

MATHINE : Il existe un théorème (que nous allons admettre) indiquant que si on a défini ainsi la foliation d'un espace, il est toujours possible de choisir le système de coordonnées  $u^i$  pour que la métrique sur la variété complète folliée soit sous la forme :

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu = g_{IJ}(v) dv^I dv^J + f(v) \gamma_{ij}(u) du^i du^j, \quad (917)$$

où  $\gamma_{ij}(u)$  est la métrique sur la sous-variété.

EURISTIDE : On voit que ce théorème nous indique que les  $g_{IJ}$  ne dépendent que des fonctions  $v^I$  seules et sont indépendants des  $u^i$ , et on voit également qu'il n'y a pas de termes croisés en  $du^i$  et  $dv^I$ . Ces deux propriétés sont tout à fait intuitives : par symétrie,  $g_{IJ}$  ne peut pas dépendre du choix de la sous-variété (donc des  $v^I$ ) puisque sinon, la symétrie sphérique de la foliation ne serait plus assurée ; de même, l'absence de termes croisés provient du fait qu'il est possible de choisir des coordonnées  $v^I$  orthogonales aux coordonnées  $u^i$  (en alignant de la même façon la foliation sur tout l'espace.)

MATHINE : En prenant en compte ce théorème, nous pouvons maintenant l'appliquer en prenant la métrique naturelle d'une sphère  $S^2$  :

$$dS^2 = d\theta^2 + \sin^2 \theta d\phi^2, \quad (918)$$

en prenant le système de coordonnées  $(\theta, \phi)$  où  $\theta$  est l'angle de rotation le long des parallèles, et  $\phi$  l'angle de rotation le long des méridiens :

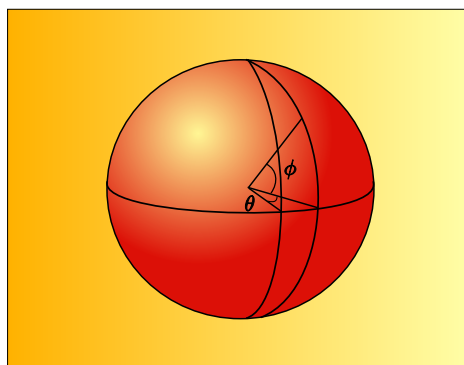


Fig. 143 - Coordonnées polaires sur la sphère

Comme l'espace-temps possède 4 dimensions, nous allons utiliser deux coordonnées supplémentaires pour désigner la sphère  $S^2$ . Dans ce cas, d'après le théorème précédent, la métrique doit s'écrire :

$$ds^2 = g_{aa}(a, b)da^2 + g_{ab}(a, b)(dad b + dbda) + g_{bb}(a, b)db^2 + r^2(a, b)dS^2. \quad (919)$$

Nous allons maintenant chercher une fonction de coordonnées  $t(a, r)$  telle que dans le système de coordonnées  $(t, r)$ , la métrique n'ait pas de terme croisé  $dt dr$ .

On a :

$$dt = \frac{\partial t}{\partial r} da + \frac{\partial t}{\partial a} dr, \quad (920)$$

d'où :

$$dt^2 = \left(\frac{\partial t}{\partial a}\right)^2 da^2 + 2\frac{\partial t}{\partial a}\frac{\partial t}{\partial r} dadr + \left(\frac{\partial t}{\partial r}\right)^2 dr^2. \quad (921)$$

Pour éliminer les termes croisés, nous voulons écrire  $ds^2$  sous la forme :

$$ds^2 = m dt^2 + n dr^2 + r^2 dS^2, \quad (922)$$

donc, par identification, cela donne :

$$\begin{aligned} g_{aa} &= m \left(\frac{\partial t}{\partial a}\right)^2 \\ g_{rr} &= n + m \left(\frac{\partial t}{\partial r}\right)^2 \\ g_{ar} &= m \left(\frac{\partial t}{\partial a}\right) \left(\frac{\partial t}{\partial r}\right) \end{aligned} \quad (923)$$

Ces trois équations permettent de déterminer théoriquement les 3 inconnues  $m$ ,  $n$ , et  $t$ .

Nous savons que l'espace-temps est lorentzien, donc  $m(t, r)$  ou  $n(t, r)$  est négatif. Supposons que ce soit  $m(t, r)$ . On peut arbitrairement représenter les fonctions  $m$  et  $n$  par des exponentielles :

$$\begin{aligned} m(t, r) &= -e^{2\alpha(t, r)} \\ n(t, r) &= e^{2\beta(t, r)} \end{aligned} \quad (924)$$

Alors, la métrique s'écrit :

$$ds^2 = -e^{2\alpha(t, r)} dt^2 + e^{2\beta(t, r)} dr^2 + r^2 dS^2. \quad (925)$$

Pour déterminer les fonctions  $\alpha(t, r)$  et  $\beta(t, r)$ , nous devons maintenant prendre en compte les équations du champ d'Einstein, et donc nous devons commencer par calculer les symboles de Christoffel pour en déduire le tenseur de Riemann, puis le tenseur de Ricci.

La coordonnée  $t$  est notée par l'indice 0,  $r$  par 1,  $\theta$  par 2, et  $\phi$  par 3.

Nous savons que :

$$\Gamma_{ik}^j = \frac{1}{2} g^{jl} (\partial_k g_{il} + \partial_i g_{lk} - \partial_l g_{ki}). \quad (926)$$

Le tenseur  $g_{\mu\nu}$  s'écrit :

$$g_{\mu\nu} = \begin{pmatrix} -e^{2\alpha} & 0 & 0 & 0 \\ 0 & r^2 e^{2\beta} & 0 & 0 \\ 0 & 0 & r^2 & 0 \\ 0 & 0 & 0 & \sin^2 \theta \end{pmatrix}. \quad (927)$$



D'où :

$$\Gamma_{00}^0 = \frac{1}{2}g^{0l}(\partial_0 g_{0l} + \partial_0 g_{l0} - \partial_l g_{00}). \quad (928)$$

Or, seule la composante  $g^{00}$  est non nulle, pour  $l = 0$ , donc :

$$\Gamma_{00}^0 = \frac{1}{2} \frac{1}{g_{00}} \partial_0 g_{00} = \frac{1}{2} \times (-e^{-2\alpha}) \times 2\partial_0 \alpha \times (-e^{2\alpha}) = \partial_0 \alpha. \quad (929)$$

De même, par des calculs équivalents, on trouve :

$$\begin{aligned} \Gamma_{01}^0 &= \partial_1 \alpha \\ \Gamma_{11}^0 &= e^{2(\alpha-\beta)} \partial_0 \beta \\ \Gamma_{00}^1 &= e^{2(\beta-\alpha)} \partial_1 \alpha \\ \Gamma_{01}^1 &= \partial_0 \beta \\ \Gamma_{11}^1 &= \partial_1 \beta \\ \Gamma_{12}^2 &= \frac{1}{r} \\ \Gamma_{22}^1 &= -re^{-2\beta} \\ \Gamma_{13}^3 &= \frac{1}{r} \\ \Gamma_{33}^1 &= -re^{-2\beta} \sin^2 \theta \\ \Gamma_{33}^2 &= -\sin \theta \cos \theta \\ \Gamma_{23}^3 &= \frac{\cos \theta}{\sin \theta}. \end{aligned} \quad (930)$$

D'autres symboles de Christoffel sont déduits par symétrie :

$$\begin{aligned} \Gamma_{10}^0 &= \Gamma_{01}^0 = \partial_1 \alpha \\ \Gamma_{10}^1 &= \Gamma_{01}^1 = \partial_0 \beta \\ \Gamma_{21}^2 &= \Gamma_{12}^2 = \frac{1}{r} \\ \Gamma_{31}^3 &= \Gamma_{13}^3 = \frac{1}{r} \\ \Gamma_{32}^3 &= \Gamma_{23}^3 = \frac{\cos \theta}{\sin \theta} \end{aligned} \quad (931)$$

Les autres symboles de Christoffel sont nuls.

Nous pouvons alors déduire les composantes du tenseur de Riemann par la relation :

$$R_{\sigma\mu\nu}^{\rho} = \partial_{\mu} \Gamma_{\nu\sigma}^{\rho} - \partial_{\nu} \Gamma_{\mu\sigma}^{\rho} + \Gamma_{\mu\lambda}^{\rho} \Gamma_{\nu\sigma}^{\lambda} - \Gamma_{\nu\lambda}^{\rho} \Gamma_{\mu\sigma}^{\lambda}. \quad (932)$$

Cela donne :

$$\begin{aligned}
R_{101}^0 &= e^{2(\beta-\alpha)}(\partial_0^2\beta + (\partial_0\beta)^2 - \partial_0\alpha\partial_0\beta) \\
&+ (\partial_1\alpha\partial_1\beta - \partial_1^2\alpha - (\partial_1\alpha)^2) \\
R_{202}^0 &= -re^{-2\beta}\partial_1\alpha \\
R_{303}^0 &= -re^{-2\beta}\sin^2\theta\partial_1\alpha \\
R_{212}^0 &= -re^{-2\alpha}\partial_0\beta \\
R_{313}^0 &= -re^{-2\alpha}\sin^2\theta\partial_0\beta \\
R_{101}^1 &= e^{2(\alpha-\beta)}[(\partial_1\alpha)^2 + \partial_1^2\alpha + \partial_1\beta\partial_1\alpha] + [\partial_0\beta\partial_0\alpha - \partial_0^2\beta - (\partial_0\beta)^2] \\
R_{212}^1 &= re^{-2\beta}\partial_1\beta \\
R_{313}^1 &= re^{-2\beta}\sin^2\theta\partial_1\beta \\
R_{020}^2 &= \frac{1}{r}e^{-2(\alpha-\beta)}\partial_1\alpha \\
R_{021}^2 &= \frac{1}{r}\partial_0\beta \\
R_{121}^2 &= \frac{1}{r}\partial_1\beta \\
R_{323}^2 &= (1 - e^{-2\beta})\sin^2\theta \\
R_{031}^2 &= \frac{1}{r}\partial_0\beta \\
R_{232}^2 &= 1 - e^{-2\beta}.
\end{aligned} \tag{933}$$

Les autres composantes du tenseur de Riemann sont nulles.

Calculons maintenant les composantes du tenseur de Ricci par contraction des deux indices :

$$\begin{aligned}
R_{00} &= -(\partial_0^2\beta + (\partial_0\beta)^2 - \partial_0\alpha\partial_0\beta) \\
&+ e^{2(\alpha-\beta)}(\partial_1^2\alpha + (\partial_1\alpha)^2 - \partial_1\alpha\partial_1\beta \\
&+ \frac{2}{r}\partial_1\alpha) \\
R_{11} &= -(\partial_1^2\alpha + (\partial_1\alpha)^2 - \partial_1\alpha\partial_1\beta - \frac{2}{r}\partial_1\beta) \\
&+ e^{2(\beta-\alpha)}(\partial_0^2\beta + (\partial_0\beta)^2 - \partial_0\alpha\partial_0\beta) \\
R_{01} &= \frac{2}{r}\partial_0\beta \\
R_{22} &= e^{-2\beta}(r(\partial_1\beta - \partial_1\alpha) - 1) + 1 \\
R_{33} &= (e^{-2\beta}(r(\partial_1\beta - \partial_1\alpha) - 1) + 1)\sin^2\theta.
\end{aligned} \tag{934}$$

Les équations d'Einstein dans le vide :

$$R_{\mu\nu} = 0, \tag{935}$$

permettent de déduire les fonctions  $\alpha$  et  $\beta$ .

Nous avons :

$$R_{01} = 0, \tag{936}$$

donc :

$$\partial_0\beta = 0, \tag{937}$$

donc  $\beta$  est une fonction de  $r$  uniquement :

$$\beta = \beta(r). \tag{938}$$

Nous avons :

$$R_{22} = 0, \quad (939)$$

donc :

$$\partial_0 R_{22} = 0 + e^{-2\beta} r (\partial_0 \partial_1 \beta - \partial_0 \partial_1 \alpha) = 0, \quad (940)$$

d'où :

$$\partial_0 \partial_1 \alpha = 0. \quad (941)$$

Donc :

$$\alpha = f(r) + g(t). \quad (942)$$

Donc, la métrique s'écrit :

$$ds^2 = -e^{2f(r)} e^{2g(t)} dt^2 + e^{2\beta(r)} dr^2 + r^2 dS^2. \quad (943)$$

Nous pouvons faire un changement de coordonnées en remplaçant  $dt^2$  par  $e^{2g(t)} dt^2$ . Dans ce cas, nous avons  $\alpha(r) = f(r)$ , d'où :

$$ds^2 = -e^{2\alpha(r)} dt^2 + e^{2\beta(r)} dr^2 + r^2 dS^2. \quad (944)$$

**EURISTIDE** : Il ressort à ce niveau que toutes les composantes de la métrique sont indépendantes du temps. Ce résultat est très important : cela signifie que toute métrique dans le vide avec une symétrie sphérique possède un champ de vecteurs de Killing de genre temps. On dit qu'une telle métrique est **stationnaire**.

Par ailleurs, il se trouve qu'il y a pas de termes en  $drdt$ , donc le champ de vecteurs de Killing  $\partial_0$  est orthogonal aux surfaces à  $t$  constant. On dit que la métrique est **statique**.

**BEATRIX** : Si je comprends bien, une métrique stationnaire décrit un espace-temps avec des déplacements à symétrie sphérique. Et une métrique statique décrit un espace-temps où rien ne bouge.

**EURISTIDE** : La métrique stationnaire permet de décrire des étoiles ou des trous noirs en rotation. La métrique statique permet de décrire des étoiles ou des trous noirs immobiles.

**MATHINE** : Poursuivons nos déductions. Nous savons que :

$$\begin{aligned} R_{00} &= 0 \\ R_{11} &= 0. \end{aligned} \quad (945)$$

Il s'ensuit que :

$$e^{2(\beta-\alpha)} R_{00} + R_{11} = \frac{2}{r} (\partial_1 \alpha + \partial_1 \beta) = 0. \quad (946)$$

Donc :

$$\alpha = -\beta + \text{Constante}. \quad (947)$$

En effectuant une translation du système de coordonnées, nous pouvons obtenir simplement :

$$\alpha = -\beta. \quad (948)$$

Nous avons également :

$$R_{22} = 0, \quad (949)$$

d'où :

$$e^{2\alpha}(2r\partial_1\alpha + 1) = 1. \quad (950)$$

Cette expression s'écrit également :

$$\partial_1(re^{2\alpha}) = 1. \quad (951)$$

Cette dernière équation a pour solution :

$$re^{2\alpha} = r + \mu, \quad (952)$$

où  $\mu$  est une constante quelconque. Par conséquent :

$$e^{2\alpha} = 1 + \frac{\mu}{r}. \quad (953)$$

Donc, la métrique s'écrit :

$$ds^2 = -\left(1 + \frac{\mu}{r}\right) dt^2 + \left(1 + \frac{\mu}{r}\right)^{-1} dr^2 + r^2 dS^2. \quad (954)$$

**EURISTIDE** : Nous allons maintenant déterminer la constante  $\mu$  en utilisant une technique dorénavant classique pour nous : en recherchant la limite newtonienne de cette métrique.

Lorsque  $r$  tend vers l'infini, l'expression ci-dessus permet de déduire que :

$$\begin{aligned} g_{00} &= -\left(1 + \frac{\mu}{r}\right) \\ g_{rr} &\approx 1 - \frac{\mu}{r}. \end{aligned} \quad (955)$$

Par ailleurs, nous avons vu dans l'étude du champ faible que :

$$\begin{aligned} g_{00} &= -\left(1 - \frac{2GM}{r}\right) \\ g_{rr} &= 1 + \frac{2GM}{r}, \end{aligned} \quad (956)$$

où  $G$  est la constante de gravitation universelle, et  $M$  la masse de l'étoile ou du trou noir.

Par identification, nous trouvons que :

$$\mu = -2GM, \quad (957)$$

d'où le résultat finalement, pour la métrique d'un système à symétrie sphérique et dans le vide, appelée **métrique de Schwarzschild** :

$$ds^2 = -\left(1 - \frac{2GM}{r}\right) dt^2 + \left(1 - \frac{2GM}{r}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2). \quad (958)$$

**BEATRIX** : On voit tout de suite deux propriétés intéressantes. Pour une masse  $M$  nulle, l'espace-temps devient plat. C'était prévisible, puisque la métrique d'un espace-temps vide et sans gravitation est la métrique de l'espace-temps plat, c'est-à-dire la métrique de Minkowski :

$$ds^2 = -dt^2 + dr^2 + r^2 d\theta^2 + r^2 d\phi^2 \sin^2\theta. \quad (959)$$

Par ailleurs, plus  $r$  est grand, plus les termes  $\frac{2GM}{r}$  deviennent négligeables, et donc plus la métrique se rapproche de celle de Minkowski. C'est également prévisible, puisque lorsque nous nous éloignons à l'infini du corps massif (étoile ou trou noir), ses effets deviennent négligeables.

EURISTIDE : Cette dernière propriété est appelée **platitude asymptotique** de l'espace-temps.

MATHINE : La solution que nous avons trouvée a été déduite des données initiales du problème, sans hypothèse supplémentaire. Le théorème de Birkhoff, que nous admettrons ici, dit que cette solution est la seule solution possible pour les conditions données d'un espace-temps à symétrie sphérique dans le vide.

EURISTIDE : En regardant de plus près, nous voyons que pour deux valeurs de  $r$ ,  $r = 0$  et  $r = 2GM$ , la métrique possède une valeur infinie. Elle n'est donc pas définie en ces points, qui correspondent d'une part au centre du corps massif, et d'autre part à une sphère centrée sur le corps massif, de rayon  $r = 2GM$ . Ces lieux de l'espace-temps sont appelés **singularités de l'espace-temps**.

BEATRIX : Qu'est-ce que cela signifie ? Cela veut-il dire que la relativité générale prédit des effets extraordinaires en ces points ? Ou bien cela signifie-t-il que la théorie que nous avons développée n'est pas complètement valide, et en particulier ne s'applique pas en ces points ?

EURISTIDE : Nous devons en fait distinguer deux cas de figure : certaines de ces singularités peuvent constituer de véritables **singularités physiques**, pour lesquelles le comportement physique de l'espace présente effectivement une singularité en ces lieux. Mais certaines de ces singularités sont en quelque sorte factices et résultent du choix du système de coordonnées, et non pas de quelque chose qui est intrinsèque au monde physique. De telles singularités sont appelées **singularités de coordonnées**. Elles peuvent être en général neutralisées par un changement adéquat de système de coordonnées.

BEATRIX : Oui, je comprends ce que cela veut dire. Par exemple, la métrique d'un espace plat à deux dimensions s'écrit en coordonnées polaires :

$$ds^2 = dr^2 + r^2 d\theta^2. \quad (960)$$

Le centre  $O$  de ce système de coordonnées provoque une dégénérescence de la métrique, puisque l'angle  $\theta$  n'a plus de signification. Il constitue donc une singularité de la métrique. Mais, ce n'est pas une singularité physique, et nous le savons à l'évidence, puisque si je choisis le système de coordonnées cartésiennes, la métrique devient :

$$ds^2 = dx^2 + dy^2, \quad (961)$$

et le point  $O$  origine du système de coordonnées n'est absolument plus une singularité.

EURISTIDE : Tu as parfaitement compris le principe. La question maintenant est de savoir détecter les singularités de coordonnées, sans devoir tatonner en testant différents systèmes de coordonnées.

MATHINE : Pour cela, nous allons utiliser la courbure. Nous pouvons calculer pour la métrique considérée un scalaire caractérisant la courbure, tel que le scalaire de Ricci :

$$R = g^{\mu\nu} R_{\mu\nu}, \quad (962)$$

ou d'autres scalaires tels que  $R^{\mu\nu}R_{\mu\nu}$ ,  $R^{\mu\nu\rho\sigma}R_{\mu\nu\rho\sigma}$ , etc. S'il s'avère que l'un de ces scalaires devient infini en un point, cela signifie que ce point est bien une singularité physique. Nous avons donc un moyen de vérifier qu'un point est une singularité physique.

Mais il faut que nous soyons bien conscients du fait que la réciproque n'est pas vraie : ce n'est pas parce que nous n'avons pas trouvé de scalaire infini caractérisant la courbure qu'un point n'est pas une singularité physique.

EURISTIDE : Appliquons ce principe à la métrique de Schwarzschild :

$$\begin{aligned} g_{00} &= -\left(1 - \frac{2GM}{r}\right) \\ g_{11} &= \left(1 - \frac{2GM}{r}\right)^{-1} \\ g_{22} &= r^2 \\ g_{33} &= r^2 \sin^2 \theta. \end{aligned} \quad (963)$$

Or, nous savons que :

$$R_{\rho\sigma\mu\nu} = \frac{1}{2}(\partial_\rho\partial_\sigma g_{\mu\nu} + \partial_\mu\partial_\rho g_{\nu\sigma} - \partial_\nu\partial_\sigma g_{\rho\mu} - \partial_\nu\partial_\rho g_{\mu\sigma}). \quad (964)$$

Donc :

$$\begin{aligned} R_{\rho\sigma\mu\nu} &= \frac{1}{2}(\partial_\rho\partial_\sigma g_{00} + \partial_\rho\partial_\sigma g_{11} + \partial_\rho\partial_\sigma g_{22} + \partial_\rho\partial_\sigma g_{33} \\ &\quad + \partial_\mu\partial_\rho g_{00} + \partial_\mu\partial_\rho g_{11} + \partial_\mu\partial_\rho g_{22} + \partial_\mu\partial_\rho g_{33} \\ &\quad - \partial_\nu\partial_\sigma g_{00} - \partial_\nu\partial_\sigma g_{11} - \partial_\nu\partial_\sigma g_{22} - \partial_\nu\partial_\sigma g_{33} \\ &\quad - \partial_\nu\partial_\rho g_{00} - \partial_\nu\partial_\rho g_{11} - \partial_\nu\partial_\rho g_{22} - \partial_\nu\partial_\rho g_{33}) \\ &= \frac{1}{2}(\partial_r\partial_r g_{00} + \partial_r\partial_r g_{11} + \partial_r\partial_r g_{22} + 2\partial_r\partial_\theta g_{33} + \partial_r\partial_r g_{33} + \partial_\theta\partial_\theta g_{33} \\ &\quad + \partial_r\partial_r g_{00} + \partial_r\partial_r g_{11} + \partial_r\partial_r g_{22} + 2\partial_r\partial_\theta g_{33} + \partial_r\partial_r g_{33} + \partial_\theta\partial_\theta g_{33} \\ &\quad - \partial_r\partial_r g_{00} - \partial_r\partial_r g_{11} - \partial_r\partial_r g_{22} - 2\partial_r\partial_\theta g_{33} - \partial_r\partial_r g_{33} - \partial_\theta\partial_\theta g_{33} \\ &\quad - \partial_r\partial_r - \partial_r\partial_r g_{11} - \partial_r\partial_r g_{22} - 2\partial_r\partial_\theta g_{33} - \partial_r\partial_r g_{33} - \partial_\theta\partial_\theta g_{33}). \end{aligned} \quad (965)$$

D'où, après tous calculs faits :

$$R^{\mu\nu\rho\sigma}R_{\mu\nu\rho\sigma} = \frac{12G^2M^2}{r^6}. \quad (966)$$

Cette expression montre immédiatement que  $r = 0$  représente une singularité physique réelle, puisque la courbure est infinie en ce point.

On peut vérifier (mais on ne va pas le faire ici), qu'aucun scalaire caractéristique de la courbure ne diverge pour  $r = 2GM$ . On peut donc penser que cet ensemble de points ne correspond pas à une singularité physique, mais nous n'en sommes pas certains.

Pour le vérifier, il faudra trouver un nouveau système de coordonnées qui neutralise cette singularité. Nous verrons cela plus tard.

Dans l'immédiat, nous allons étudier le comportement des particules dans l'espace-temps, considéré en présence d'un champ à symétrie sphérique. Nous allons pour cela considérer que  $r$  est plus grand que  $2GM$ . En effet, dans le cas des étoiles habituelles, le rayon de l'étoile est largement supérieur au rayon  $2GM$ . C'est pourquoi les mouvements des particules devant être considérés dans le vide, il est indispensable que  $r$  soit supérieur au rayon de l'étoile et par conséquent au moins à  $2GM$ .

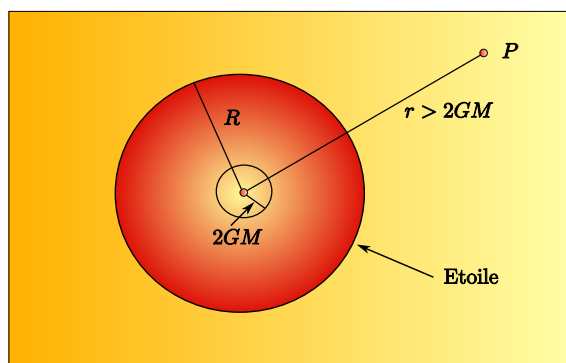


Fig. 144 - Une étoile selon Schwarzschild

**MATHINE** : Le mouvement des particules dans l'espace-temps que nous avons décrit, dont la métrique est la métrique de Schwarzschild, est décrit par les géodésiques, puisque nous savons que les particules libres se déplacent le long de géodésiques de l'espace-temps.

Pour écrire les équations géodésiques, nous avons besoin des symboles de Christoffel, que nous allons calculer avec la formule habituelle :

$$\Gamma_{jk}^i = \frac{1}{2} g^{il} (\partial_k g_{jl} + \partial_j g_{kl} - \partial_l g_{kj}). \quad (967)$$

Nous obtenons :

$$\begin{aligned} \Gamma_{01}^0 &= \frac{GM}{r(r-2GM)} \\ \Gamma_{00}^1 &= \frac{GM}{r^3} (r-2GM) \\ \Gamma_{11}^1 &= \frac{-GM}{r(r-2GM)} \\ \Gamma_{22}^1 &= -(r-2GM) \\ \Gamma_{33}^1 &= -(r-2GM) \sin^2 \theta \\ \Gamma_{12}^2 &= \frac{1}{r} \\ \Gamma_{33}^2 &= -\sin \theta \cos \theta \\ \Gamma_{13}^3 &= \frac{1}{r} \\ \Gamma_{23}^3 &= \frac{\cos \theta}{\sin \theta}. \end{aligned} \quad (968)$$

Puis les équations décrivant les géodésiques sont :

$$\frac{d^2 x^\mu}{d\lambda^2} + \Gamma_{\rho\sigma}^\mu \frac{dx^\sigma}{d\lambda} \frac{dx^\rho}{d\lambda} = 0. \quad (969)$$

D'où, en utilisant les symboles de Christoffel et en développant pour chacune des 4 coordonnées de l'espace-temps :

$$\frac{d^2 t}{d\lambda^2} + \frac{2GM}{r(r-2GM)} \frac{dr}{d\lambda} \frac{dt}{d\lambda} = 0 \quad (970)$$

$$\frac{d^2 r}{d\lambda^2} + \frac{GM}{r^3} (r-2GM) \left(\frac{dt}{d\lambda}\right)^2 - \frac{GM}{r(r-2GM)} \left(\frac{dr}{d\lambda}\right)^2 \quad (971)$$

$$-(r-2GM) \left( \left(\frac{d\theta}{d\lambda}\right)^2 + \sin^2 \theta \left(\frac{d\phi}{d\lambda}\right)^2 \right) = 0$$

$$\frac{d^2 \theta}{d\lambda^2} + \frac{2}{r} \frac{d\theta}{d\lambda} \frac{dr}{d\lambda} - \sin \theta \cos \theta \left(\frac{d\phi}{d\lambda}\right)^2 = 0$$

$$\frac{d^2 \phi}{d\lambda^2} + \frac{2}{r} \frac{d\phi}{d\lambda} \frac{dr}{d\lambda} + 2 \frac{\cos \theta}{\sin \theta} \frac{d\theta}{d\lambda} \frac{d\phi}{d\lambda} = 0.$$

Pour résoudre ces équations, nous avons besoin d'utiliser les hypothèses de symétrie. Nous savons que la symétrie sphérique implique que l'espace-temps possède 3 champs de vecteurs de Killing. Nous savons également que l'espace-temps est statique, donc l'invariance par translation du temps nous fournit un 4-ième champ de vecteurs de Killing.

Le choix des coordonnées que nous avons fait en cohérence avec les champs de vecteurs de Killing permet d'écrire :

$$K_\mu \frac{\partial x^\mu}{\partial \lambda} = \text{Constante}, \quad (972)$$

où  $K_\mu$  est un champ de vecteurs de Killing. Par ailleurs, la compatibilité de la métrique fait que :

$$-g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} = \text{Constante} = C. \quad (973)$$

BEATRIX : Quelle est la signification physique des champs de vecteurs de Killing ?

EURISTIDE : L'invariance dans toute translation temporelle a pour signification physique la conservation de l'énergie. L'invariance dans les rotations spatiales correspond à la conservation du moment angulaire, en l'absence de force extérieure au système.

Le moment angulaire peut être représenté comme un vecteur à 3 composantes, 2 pour la direction, 1 pour l'intensité du moment angulaire.

Deux champs de vecteurs de Killing  $\partial_r$  et  $\partial_\theta$  correspondent à la conservation de la direction du moment angulaire. Donc les particules se déplacent dans un plan, que nous pouvons choisir équatorial arbitrairement, sans nuire à la généralité, puisqu'il suffit de faire un changement de coordonnées adéquat.

Par conséquent, la symétrie sphérique permet de fixer  $\theta$ , en l'occurrence :

$$\theta = \frac{\pi}{2}. \quad (974)$$

MATHINE : Le champ de vecteurs de Killing correspondant à la conservation de l'énergie est  $\partial_t$ , et a pour composantes :

$$K_\mu = \left( - \left( 1 - \frac{2GM}{r} \right), 0, 0, 0 \right). \quad (975)$$



Le champ de vecteurs de Killing correspondant à la conservation de l'intensité du moment angulaire est  $\partial_\phi$  et a pour composantes :

$$L_\mu = (0, 0, 0, r^2 \sin^2 \theta). \quad (976)$$

On en déduit les deux quantités conservées :

$$\begin{aligned} E &= \left(1 - \frac{2GM}{r}\right) \frac{dt}{d\lambda} \\ L &= r^2 \sin^2 \theta \frac{d\phi}{d\lambda}. \end{aligned} \quad (977)$$

Or, nous avons :

$$-g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} = \text{Constante}. \quad (978)$$

Nous pouvons choisir un système d'unités afin que cette constante soit 1.

D'où en développant :

$$-\left(1 - \frac{2GM}{r}\right) \left(\frac{dt}{d\lambda}\right)^2 + \left(1 - \frac{2GM}{r}\right)^{-1} \left(\frac{dr}{d\lambda}\right)^2 + r^2 \left(\frac{d\phi}{d\lambda}\right)^2 = -1. \quad (979)$$

Soit, en intégrant les expressions de  $E$  et de  $L$  :

$$-E^2 + \left(\frac{dr}{d\lambda}\right)^2 + \left(1 - \frac{2GM}{r}\right) \left(\frac{L^2}{r^2} + 1\right) = 0. \quad (980)$$

Posons :

$$V(r) = \frac{1}{2} - \frac{GM}{r} + \frac{L^2}{2r^2} - \frac{GML^2}{r^3}, \quad (981)$$

alors :

$$\frac{1}{2} \left(\frac{dr}{d\lambda}\right)^2 + V(r) = \frac{1}{2} E^2. \quad (982)$$

**EURISTIDE** : Cette dernière équation est l'équation du mouvement radial d'une particule dans un potentiel de champ  $V(r)$ , la particule ayant pour énergie  $\frac{1}{2}E^2$ .

Ce potentiel est décrit par l'équation :

$$V(r) = \frac{1}{2} - \frac{GM}{r} + \frac{L^2}{2r^2} - \frac{GML^2}{r^3}, \quad (983)$$

Le premier terme est une constante. Le second terme représente le potentiel newtonien classique. Le troisième terme est la contribution du moment angulaire, qui est également le même dans la théorie newtonienne. Le quatrième terme est spécifique à la relativité générale. C'est lui qui va déterminer les conséquences de la théorie d'Einstein sur le mouvement des particules.

**MATHINE** : L'équation du mouvement de la particule peut être réécrit sous la forme :

$$\frac{1}{2} \left(\frac{dr}{d\lambda}\right)^2 = \frac{1}{2} E^2 - V(r). \quad (984)$$

EURISTIDE : En général, pour  $R > 2GM$ , la particule se déplacera dans le potentiel jusqu'à un point de pivot où :

$$V(r) = \frac{1}{2}E^2. \quad (985)$$

A ce moment, la dérivée seconde de  $r$  par rapport à  $\lambda$  s'annulera, et la particule repartira dans l'autre direction.

En revanche, pour  $V(r) = \text{Constante}$ , on aura une trajectoire circulaire.

MATHINE : Pour le cas où  $V(r)$  est une constante, on peut différencier l'expression de  $V(r)$  :

$$\frac{dV(r)}{dr} = 0 = +\frac{GM}{r^2} - \frac{2L^2}{2r^3} + \frac{3L^2GM}{r^4}. \quad (986)$$

Soit en multipliant l'expression par  $r^4$  :

$$GMr^2 - L^2r + 3GML^2 = 0. \quad (987)$$

EURISTIDE : C'est une équation du second degré. La trajectoire peut donc avoir deux valeurs de rayon :

$$r_c = \frac{L^2 \pm \sqrt{L^4 - 12G^2M^2L^2}}{2GM}, \quad (988)$$

soit :

$$r_c = \frac{L^2}{2GM} \left( 1 \pm \sqrt{1 - \frac{12G^2M^2}{L^2}} \right). \quad (989)$$

Le rayon le plus intérieur  $r_-$  correspond à un potentiel maximal, c'est donc une trajectoire instable : la moindre perturbation provoquera une chute de la particule vers la singularité  $r = 0$ , ou vers l'infini de l'espace.

Le rayon le moins intérieur  $r_+$  correspond à un potentiel minimal, c'est donc une trajectoire stable circulaire.

Il faut noter que dans le cas newtonien, le terme :

$$\frac{3L^2GM}{r^4} \quad (990)$$

n'est plus présent. Donc l'équation donnant le rayon de la trajectoire circulaire se réduit à :

$$GMr^2 - L^2r = 0, \quad (991)$$

soit :

$$r = \frac{L^2}{GM}. \quad (992)$$

Il n'y a donc dans ce cas qu'une seule trajectoire circulaire.

MATHINE : Pour le cas où  $r = 2GM$ , le potentiel  $V(r)$  est toujours nul.

EURISTIDE : Nous étudierons plus loin le cas où  $r \leq 2GM$  qui correspond au cas du trou noir.

En attendant, examinons le comportement des solutions de trajectoires circulaires en fonction des valeurs

du moment angulaire  $L$ .

Si  $L$  est grand et tend vers l'infini, alors les deux solutions deviennent approximativement :

$$\begin{aligned} r_- &= 3GM \\ r_+ &= \frac{L^2}{GM}. \end{aligned} \quad (993)$$

Par conséquent, l'orbite circulaire stable devient de plus en plus lointaine, et l'orbite circulaire instable s'approche de  $3GM$ .

Lorsque  $L$  diminue, les orbites circulaires se rapprochent progressivement, pour se confondre en une seule orbite circulaire pour :

$$L = \sqrt{12}GM, \quad (994)$$

et pour laquelle :

$$r = 6GM. \quad (995)$$

Pour les valeurs de  $L$  inférieures à  $\sqrt{12}GM$ , il n'y a plus d'orbite circulaire.

**BEATRIX** : En résumé, pour une étoile à symétrie sphérique, les particules sont en mouvement sur une orbite circulaire instable entre les rayons  $3GM$  et  $6GM$ . Au dessus de  $6GM$ , les particules sont sur une orbite circulaire stable. En dessous de  $3GM$ , il n'y a plus d'orbite circulaire.

**EURISTIDE** : Nous allons considérer maintenant deux applications concrètes et vérifiées expérimentalement du modèle d'étoile à symétrie sphérique de la métrique de Schwarzschild : il s'agit de la déflexion de la lumière, et de la précession du périhélie de Mercure.

Commençons par la déflexion de la lumière. La lumière, tout comme la matière ordinaire, est soumise à l'effet de la gravitation. En effet, la gravitation provoquant une courbure de l'espace-temps, cela a pour effet d'incurver toutes les géodésiques de l'espace-temps, et par conséquent d'incurver également les trajectoires des rayons lumineux qui s'y déplacent.

**MATHINE** : Reprenons l'équation du mouvement, en introduisant la variable :

$$u = \frac{1}{r}. \quad (996)$$

Alors, en considérant que le paramètre  $\lambda$  est la coordonnée  $\phi$ , nous avons :

$$\frac{d^2u}{d\phi^2} + u = \frac{GM}{L^2} + 3GMu^2. \quad (997)$$

Pour la trajectoire de la lumière, nous savons que l'expression :

$$g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} = \text{Constante}, \quad (998)$$

est en fait nulle.

Donc, dans l'équation ci-dessus, le terme en  $\frac{1}{L^2}$  est absent pour la trajectoire de la lumière et l'équation de la trajectoire de la lumière devient :

$$\frac{d^2u}{d\phi^2} + u = 3GMu^2. \quad (999)$$

Pour résoudre cette équation, nous considérons que la solution est une fonction du type :

$$u = u_0 + p \quad (1000)$$

où  $p$  est une fonction de perturbation, et  $u_0$  la solution de l'équation au plus petit ordre pour  $u$  petit (donc  $r$  grand) :

$$\frac{d^2 u_0}{d\phi^2} + u_0 = 0. \quad (1001)$$

La solution de cette équation est de la forme :

$$u_0 = \frac{1}{b} \cos \phi. \quad (1002)$$

Le paramètre  $b$  porte le nom de **paramètre d'impact**, et correspond à la distance la plus proche entre la trajectoire de la lumière et l'étoile. Pour calculer la perturbation, nous considérons qu'elle est telle que :

$$u = \frac{1}{b} (\cos \phi + B + A \sin^2 \phi). \quad (1003)$$

En réinjectant  $u$  dans l'équation d'origine, nous obtenons l'approximation :

$$\frac{1}{b} (B + 2A - 3A \sin^2 \phi) = \frac{3GM}{b^2} (1 - \sin^2 \phi). \quad (1004)$$

Ce qui signifie que :

$$B = A = \frac{GM}{b}. \quad (1005)$$

Donc, finalement, la trajectoire de la lumière est telle que :

$$u = \frac{1}{b} (\cos \phi + \frac{GM}{b} (1 + \sin^2 \phi)). \quad (1006)$$

EURISTIDE : Rappelons que la fonction  $u$  est en fait la fonction :

$$u(\phi) = \frac{1}{r(\phi)}. \quad (1007)$$

Donc, la trajectoire de la lumière à l'infini est telle que :

$$u(\phi) = 0. \quad (1008)$$

Supposons que la lumière passe près d'une étoile en arrivant sous un angle :

$$\frac{\delta\phi}{2} \quad (1009)$$

par rapport à l'axe :

$$\phi = \frac{\pi}{2}, \quad (1010)$$

et supposons que cette même lumière s'éloigne de l'étoile en repartant sous un angle opposé :

$$\frac{\delta\phi}{2}. \quad (1011)$$

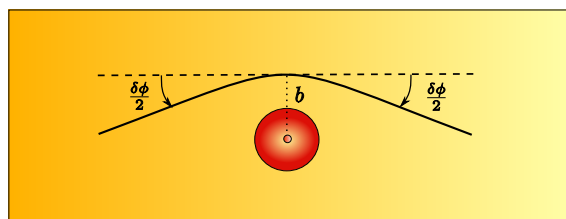


Fig. 145 - Déflexion de la lumière au voisinage du Soleil

Pour déterminer  $\frac{\delta\phi}{2}$ , il nous reste à résoudre l'équation :

$$u\left(\frac{\pi}{2} + \frac{\delta\phi}{2}\right) = 0, \quad (1012)$$

soit :

$$\frac{1}{b} \left( \cos\left(\frac{\pi}{2} + \frac{\delta\phi}{2}\right) + \frac{GM}{b} \left(1 + \sin^2\left(\frac{\pi}{2} + \frac{\delta\phi}{2}\right)\right) \right) = 0. \quad (1013)$$

Soit encore, en développant ces fonctions trigonométriques selon un développement limité par rapport à  $\pi/2$  :

$$\begin{aligned} 0 &= \frac{1}{b} \left( -\frac{\delta\phi}{2} + \frac{GM}{b} \left(1 + 1 + \frac{\delta\phi^2}{4}\right) \right) \\ &= -\frac{\delta\phi}{2b} + \frac{2GM}{b^2} + \frac{GM\delta\phi^2}{4b^2}. \end{aligned} \quad (1014)$$

En négligeant  $\delta\phi^2$  par rapport à  $\delta\phi$ , puisque l'angle de déflexion est infiniment petit, on obtient par approximation :

$$\delta\phi = \frac{4GM}{b}. \quad (1015)$$

Ce qui nous donne la formule pour calculer l'angle de déflexion d'un rayon lumineux passant à la distance  $b$  d'une étoile de masse  $M$ .

En appliquant cette formule à un rayon lumineux rasant la surface du soleil, on obtient :

$$\delta\phi = 1,75'', \quad (1016)$$

donc un angle très petit de 1,75 secondes d'arc. En 1919, cet angle a pu être mesuré précisément lors d'une éclipse solaire, en observant le décalage apparent des étoiles dans le ciel lorsqu'elles étaient voisines du Soleil, résultant de la déflexion de leurs rayons lumineux passant au voisinage du Soleil.

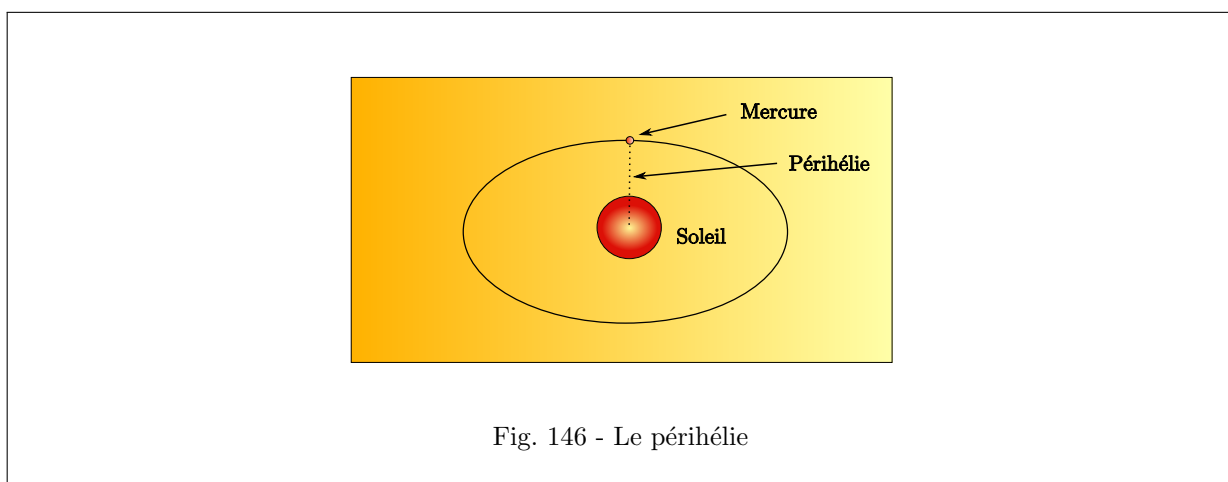
Cette observation a permis de confirmer la justesse des prédictions de la relativité générale, et a été considérée justement comme la preuve définitive que la théorie d'Einstein n'était pas seulement une curiosité mathématique, mais constituait réellement une théorie physique expliquant les lois de l'Univers.

**BEATRIX** : C'est fantastique. On a donc pu vérifier directement que les rayons lumineux ne suivaient pas toujours des trajectoires rectilignes, et qu'ils pouvaient être incurvés par l'effet de la gravitation d'un corps massif à proximité.

EURISTIDE : La théorie de la relativité a permis d'expliquer un autre phénomène : la précession du périhélie de Mercure.

BEATRIX : Qu'est que cela signifie ?

EURISTIDE : Tu sais que les planètes tournent autour du Soleil en suivant des trajectoires elliptiques planes. Ce qu'on appelle le périhélie, c'est le point de la trajectoire de la planète qui est le plus proche du Soleil.



Il se trouve qu'avant la publication des résultats de la relativité générale, les astronomes avaient un problème. Ils observaient que l'axe principal de l'orbite elliptique de Mercure se décalait progressivement d'une amplitude de 500 secondes d'arc par siècle. C'est ce qu'on appelle la précession du périhélie de Mercure. Les physiciens avaient expliqué majoritairement cette précession comme étant due à la perturbation des planètes avoisinantes en utilisant la théorie de Newton, mais il restait une précession d'environ 40 secondes d'arc par siècle qui était inexpliquée.

MATHINE : Pour déterminer l'équation du mouvement d'une particule (la planète Mercure en l'occurrence), autour d'un champ gravitationnel à symétrie sphérique (le Soleil), nous allons utiliser comme précédemment l'équation classique du mouvement, de laquelle nous ôtons le terme contribuant à la relativité générale, et les solutions de cette équation classique seront modifiées pour y ajouter une fonction perturbation.

L'équation du mouvement de la relativité générale est :

$$\frac{d^2u}{d\phi^2} + u = \frac{GM}{L^2} + 3GMu^2. \quad (1017)$$

Dans un premier temps, nous négligeons le terme  $3GMu^2$  pour résoudre l'équation classique du mouvement :

$$\frac{d^2 u_0}{d\phi^2} + u_0 = \frac{GM}{L^2}. \quad (1018)$$

Cette équation a des solutions du type :

$$u_0 = \frac{GM}{L^2} (1 + e \cos \phi), \quad (1019)$$

où  $e$  est appelée **excentricité de l'orbite** de la planète. Comme nous étudions le mouvement d'une planète sur une ellipse autour du Soleil, nous allons considérer que  $e$  est compris entre 0 et 1 (sinon la trajectoire serait une parabole ( $e = 1$ ) ou une hyperbole ( $e > 1$ )). Donc la trajectoire est une ellipse plane, paramétrée par l'angle  $\phi$  autour du Soleil.

Comme cette trajectoire est elliptique, elle doit être écrite sous la forme suivante, qui est l'équation en coordonnées polaires d'une ellipse :

$$u_0 = \frac{1 + e \cos \theta}{a(1 - e^2)}, \quad (1020)$$

où  $e$  est l'excentricité de l'ellipse, et  $a$  son axe demi-majeur. Donc nous avons :

$$\frac{L^2}{GM} = a(1 - e^2), \quad (1021)$$

ou encore :

$$L^2 = GMa(1 - e^2). \quad (1022)$$

Nous allons maintenant considérer que la perturbation due à la relativité générale intervient sous la forme d'un facteur proche de 1, noté  $\omega$ , appliqué à l'angle de rotation de la planète sur l'ellipse :

$$u = \frac{1 + e \cos \omega \phi}{a(1 - e^2)}. \quad (1023)$$

En injectant cette fonction  $u$  dans l'équation du mouvement intégrant le terme de la relativité générale, nous obtenons :

$$\frac{d^2}{d\phi^2} \left( \frac{1 + e \cos \omega \phi}{a(1 - e^2)} \right) + \frac{1 + e \cos \omega \phi}{a(1 - e^2)} = \frac{1}{a(1 - e^2)} + 3M \frac{(1 + e \cos \omega \phi)^2}{a^2(1 - e^2)^2}. \quad (1024)$$

En multipliant l'équation par  $a(1 - e^2)$ , et en développant la dérivée seconde de  $\cos \omega \phi$ , on obtient :

$$-e\omega^2 \cos \omega \phi + 1 + e \cos \omega \phi = 1 + \frac{3M + 6Me \cos \omega \phi + 3Me^2 \cos^2 \omega \phi}{a(1 - e^2)}. \quad (1025)$$

En considérant que  $e$  est petit, et en effectuant les approximations qui en résultent, le calcul donne :

$$1 - \omega = \frac{3M}{a(1 - e^2)} = \delta\omega. \quad (1026)$$

Par ailleurs, nous avons l'expression de la vitesse de précession  $\Omega$  :

$$\Omega = \frac{2\pi\delta\omega}{T}. \quad (1027)$$

Et, compte tenu de la **troisième loi de Kepler**, nous avons :

$$4\pi^2 a^3 = GMT^2. \quad (1028)$$

En réinjectant ces valeurs dans l'expression de la vitesse de précession  $\Omega$ , on obtient :

$$\Omega = \frac{3(GM)^{3/2}}{c^2(1-e^2)a^{5/2}}. \quad (1029)$$

En appliquant les données numériques pour Mercure et le Soleil, nous obtenons une valeur de la vitesse de précession de 43 secondes d'arc.

BEATRIX : Et voilà ! La prédiction est conforme aux données expérimentales ; nous tenons une nouvelle vérification directe de la relativité générale.

## 84 Scène XIII-4 : Les trous noirs

EURISTIDE : Nous allons maintenant examiner le cas particulier des trous noirs décrits par la métrique de Schwarzschild.

On appelle **trous noirs** des objets stellaires tels qu'il est possible qu'un observateur soit dans le vide lorsqu'il se trouve à une distance inférieure à  $2GM$  par rapport à ce corps stellaire. Cela suppose que le corps stellaire soit extrêmement dense parce que la plupart des étoiles que nous connaissons ont un rayon très largement supérieur à  $2GM$ , donc il n'est pas possible de se trouver dans le vide à l'intérieur de ce rayon. Il s'agit en général d'étoiles très massives qui se sont écroulées sur elles-mêmes lorsque la gravitation devient trop forte.

Pour comprendre le comportement d'un observateur à proximité d'un tel trou noir, nous allons étudier des courbes radiales de genre lumière (donc les trajectoires de rayons lumineux) au voisinage d'un trou noir.

Une trajectoire radiale de genre lumière, dans la métrique de Schwarzschild, correspond à  $ds^2 = 0$ , pour  $\theta$  et  $\phi$  constants. Donc :

$$-\left(1 - \frac{2GM}{r}\right) dt^2 + \left(1 - \frac{2GM}{r}\right)^{-1} dr^2 = 0. \quad (1030)$$

Nous en déduisons que :

$$\frac{dt}{dr} = \pm \left(1 - \frac{2GM}{r}\right)^{-1}. \quad (1031)$$

Cela signifie que pour  $r$  s'approchant de  $2GM$ , l'évolution du temps par rapport au rayon  $r$  devient infinie.

BEATRIX : Cela veut dire que la lumière met un temps infini à se déplacer lorsqu'elle approche le rayon  $r = 2GM$ . Elle ne parvient donc jamais à l'atteindre.

EURISTIDE : En fait, c'est une illusion. N'oublions pas que nous sommes en train d'étudier la variation du temps universel ou de coordonnées  $dt$  par rapport au rayon  $r$ . Cela veut dire que nous analysons le comportement d'un rayon lumineux lorsqu'il est étudié par un observateur distant au repos. Pour un tel observateur, le rayon lumineux semble ne jamais atteindre le rayon  $r = 2GM$  ; il s'en approche en fait de plus en plus lentement.

BEATRIX : Maintenant, il serait intéressant de voir ce qu'il se passe dans le temps propre d'un observateur qui s'approcherait de ce rayon  $r = 2GM$ .



EURISTIDE : Pour faire cela, nous allons avoir besoin de changer de système de coordonnées pour adopter un système de coordonnées qui ne comporte pas de singularité en  $r = 2GM$ .

Pour cela, nous allons choisir le système de coordonnées  $u, r, \theta, \phi$  tel que :

$$u = t + r + 2GM \ln \left( \left| \frac{r}{2GM} - 1 \right| \right). \quad (1032)$$

C'est un peu compliqué. Mais c'est la bonne technique pour éliminer la singularité des coordonnées précédentes. En effet, dans ce système de coordonnées, la métrique s'écrit :

$$ds^2 = - \left( 1 - \frac{2GM}{r} \right) du^2 + (drdu + dudr) + r^2(d\theta^2 + \sin^2 \theta d\phi^2). \quad (1033)$$

Ce système de coordonnées est appelé **coordonnées d'Eddington-Finkelstein**.

Dans ce nouveau système de coordonnées, écrivons l'équation de la trajectoire de genre lumière d'un rayon lumineux, avec  $ds^2 = 0$  pour  $\theta$  et  $\phi$  constants :

$$- \left( 1 - \frac{2GM}{r} \right) du^2 + 2dudr = 0, \quad (1034)$$

d'où deux solutions :

$$\begin{aligned} \frac{du}{dr} &= 0 \\ \frac{du}{dr} &= 2 \left( 1 - \frac{2GM}{r} \right)^{-1} = \frac{2}{1 - \frac{2GM}{r}}. \end{aligned} \quad (1035)$$

La solution  $\frac{du}{dr} = 0$  correspond aux trajectoires entrantes, se dirigeant vers le centre du trou noir.

L'autre solution correspond aux trajectoires sortantes, partant du centre du trou noir pour se diriger vers l'extérieur de celui-ci. On voit qu'une particule provenant de l'intérieur de la sphère  $r < 2GM$  ne parvient pas à franchir la barrière  $r = 2GM$ , et reste inéluctablement à l'intérieur de la sphère de rayon  $r = 2GM$ .

D'ailleurs, pour  $r < 2GM$ ,  $\frac{du}{dr}$  est négatif, ce qui veut dire que le futur correspond systématiquement à des trajectoires pour lesquelles  $dr < 0$ , donc pour lesquelles la particule se rapproche du centre du trou noir.

BEATRIX : En bref, le trou noir ne laisse rien (ni particule, ni lumière) s'échapper, dès que cette particule ou ce rayon lumineux a franchi la barrière de la sphère de rayon  $r = 2GM$ .

EURISTIDE : C'est précisément pour cela qu'on appelle ces corps célestes des trous noirs : c'est parce qu'on ne peut pas en voir l'intérieur ou la surface, puisque rien, pas même la lumière, ne peut s'en échapper. La sphère de rayon  $r = 2GM$  est appelée **horizon des événements**. C'est l'horizon, à proximité du trou noir, en dessous duquel un observateur externe ne peut plus rien voir.

---

## Quinzième partie

# Acte XIV - La cosmologie relativiste

## 85 Scène XIV-1 : Les principes cosmologiques

BEATRIX : Le modèle de la métrique de Schwarzschild est très efficace. Nous parvenons à décrire le comportement des étoiles et des planètes, et nous pouvons prédire et modéliser le comportement des trous noirs. C'est à peu près tout ce que l'on peut trouver dans l'univers.

EURISTIDE : C'est un peu vrai. Mais en fait, les physiciens ont dû élaborer d'autres modèles pour s'adapter à des situations particulières, telles que des trous noirs chargés électriquement ou des trous noirs en rotation. Nous n'allons pas étudier ces modèles ici, mais il faut savoir qu'ils existent.

BEATRIX : Alors, nous savons tout sur la relativité générale, maintenant ?

EURISTIDE : Il reste une chose à voir : le modèle de l'Univers tout entier. La relativité générale permet aussi de comprendre le comportement et l'évolution de l'univers dans son ensemble. C'est l'objet de la science appelée **cosmologie**. Nous allons donc maintenant étudier la cosmologie relativiste.

Pour commencer, nous allons définir les principes qui régissent l'univers. Ces principes sont appelés **principes cosmologiques** et ils sont au nombre de deux : l'homogénéité et l'isotropie de l'univers.

L'univers est homogène : cela signifie qu'il n'existe, à grande échelle, aucun point particulier dans l'univers. Les galaxies sont distribuées de façon uniforme dans l'espace, si on considère ce dernier à grande échelle.

L'univers est isotrope : cela signifie qu'il n'y a pas de directions spatiales particulières dans l'univers. Les galaxies sont distribuées de façon uniforme dans toutes les directions angulaires si on considère l'espace à grande échelle.

BEATRIX : Cela me rappelle ce que disaient les hommes autrefois : ils considéraient que la Terre était le centre de l'univers. Les principes cosmologiques réfutent cela aussi...

## 86 Scène XIV-2 : Modèles Friedmann-Robertson-Walker

EURISTIDE : L'univers matériel en 3 dimensions est donc homogène et isotrope. Mais nous savons que l'univers est en expansion : il n'est donc pas correct de considérer que l'univers est également homogène et isotrope par rapport à sa dimension temporelle.

Nous sommes donc conduits à considérer que l'espace est découpé en foliations homogènes et isotropes à 3 dimensions, chaque foliation correspondant à un instant  $t$  de l'espace-temps :

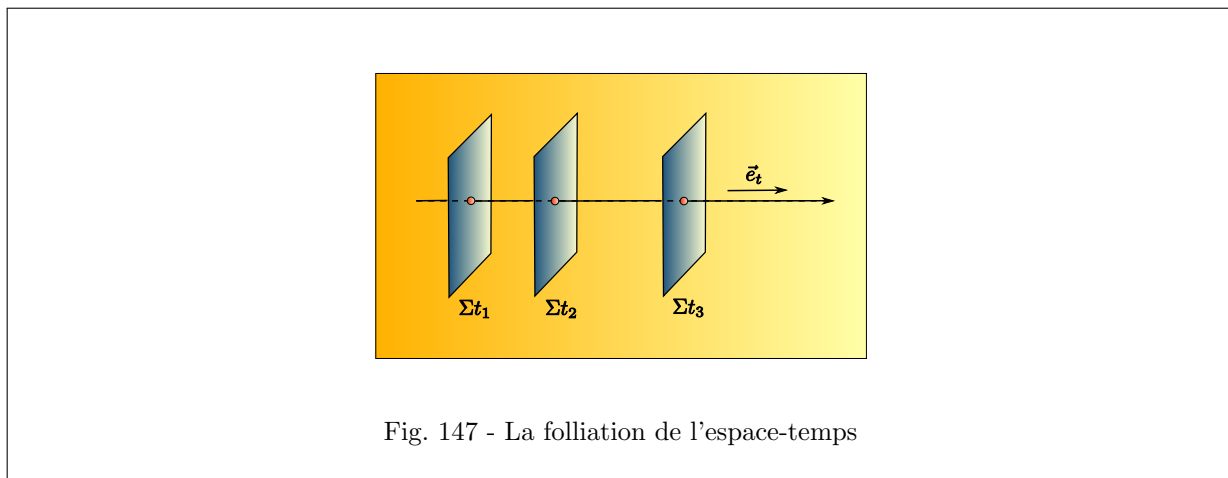


Fig. 147 - La foliation de l'espace-temps

Chaque foliation est un espace de dimension 3 dans l'espace-temps à 4 dimensions. Ces foliations sont notées  $\Sigma_t$  et sont des hypersurfaces. L'univers peut donc être représenté par :

$$\mathbb{R} \times \Sigma_t, \quad (1036)$$

où  $\mathbb{R}$  représente la droite temporelle, et  $\Sigma_t$  chaque hypersurface correspondant à l'espace matériel tel qu'il est à l'instant  $t$ .

La droite temporelle est représentée par le vecteur  $\vec{e}_t$ . Ce vecteur est obligatoirement orthogonal aux  $\Sigma_t$ , car sinon, sa projection sur  $\Sigma_t$  déterminerait une direction privilégiée sur  $\Sigma_t$ , ce qui serait en contradiction avec le principe d'isotropie de l'univers.

Par conséquent, la métrique de l'espace-temps peut s'écrire sans faire apparaître de termes croisés temporels et spatiaux. Nous pouvons donc écrire la métrique sous la forme :

$$ds^2 = -dt^2 + a(t)^2(d\chi^2 + r(\chi)^2(d\theta^2 + \sin^2 \theta d\phi^2)), \quad (1037)$$

où  $\chi$  est la coordonnée radiale,  $\theta$  et  $\phi$  sont les coordonnées polaires sphériques.

La fonction  $a(t)$  est appelée **facteur d'échelle**, puisqu'elle constitue le facteur multiplicatif de la coordonnée radiale.

**MATHINE** : Compte tenu des principes cosmologiques d'isotropie et d'homogénéité, nous pouvons considérer que la métrique spatiale  $\gamma_{ij}$  est maximale symétrique. Dans ce cas, il se trouve que le tenseur de Riemann vérifie :

$$R_{ijkl} = k(\gamma_{ik}\gamma_{jl} - \gamma_{il}\gamma_{jk}), \quad (1038)$$

où  $k$  est une constante. Nous allons admettre cette propriété, que l'on peut retrouver par un calcul direct un peu fastidieux.

Nous en déduisons le tenseur de Ricci :

$$R_{jl} = 2k\gamma_{jl}. \quad (1039)$$

Par ailleurs, l'espace ayant une symétrie sphérique, nous pouvons appliquer ici ce que nous avons vu dans le cadre de la métrique de Schwarzschild, en l'appliquant au cas tridimensionnel spatial :

$$d\sigma^2 = \gamma_{ij}du^i du^j = e^{2\beta(r)} dr^2 + r^2(d\theta^2 + \sin^2 \theta d\phi^2). \quad (1040)$$

Le calcul donne :

$$\begin{aligned} R_{11} &= \frac{2}{r} \partial_1 \beta \\ R_{22} &= e^{-2\beta} (r \partial_1 \beta - 1) + 1 \\ R_{33} &= (e^{-2\beta} (r \partial_1 \beta - 1) + 1) \sin^2 \theta. \end{aligned} \quad (1041)$$

Or, nous avons vu que les composantes du tenseur de Ricci étaient proportionnelles aux composantes de la métrique spatiale :

$$R_{jl} = 2k\gamma_{jl}, \quad (1042)$$

donc, par identification, on obtient :

$$\beta = -\frac{1}{2} \ln(1 - kr^2). \quad (1043)$$

D'où la métrique de l'espace-temps :

$$ds^2 = -dt^2 + a^2(t) \left( \frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin^2 \theta d\phi^2) \right). \quad (1044)$$

EURISTIDE : Cette métrique est appelée **métrique de Robertson-Walker**. Par un simple changement de coordonnées, on peut substituer à  $k$  la constante :

$$\kappa = \frac{k}{|k|}, \quad (1045)$$

d'où la forme suivante de la métrique :

$$ds^2 = -dt^2 + a^2(t) \left( \frac{dr^2}{1 - \kappa r^2} + r^2(d\theta^2 + \sin^2 \theta d\phi^2) \right), \quad (1046)$$

où  $\kappa = -1, +1$  ou  $0$ .

BEATRIX : Cela va nous donner trois types d'univers, en fonction du paramètre  $\kappa$ , n'est-ce pas ?

EURISTIDE : Oui, en effet. Pour  $\kappa = 0$ , les foliations spatiales sont plates. Le modèle d'univers correspondant est appelé **modèle d'univers plat**. Mais attention, le fait que l'univers spatial soit plat n'exclut pas que l'espace-temps soit curviligne !

Pour  $\kappa = -1$ , les foliations possèdent une courbure négative constante, donc elles sont en forme de selle de cheval. Le modèle d'univers correspondant est appelé **modèle d'univers ouvert**. L'espace s'étend indéfiniment.

Pour  $\kappa = +1$ , les foliations spatiales possèdent une courbure positive constante, donc elles sont repliées sur elles-mêmes et n'ont pas une étendue infinie. Le modèle d'univers correspondant est appelé **modèle d'univers clos**.

BEATRIX : Si je comprends bien, on peut imaginer l'univers sous 3 formes. Un monde fini, en forme de boule à 3 dimensions, ou bien un monde infini, en forme de selle de cheval à 3 dimensions, ou enfin, un monde infini tout plat dans ses 3 dimensions spatiales.

## 87 Scène XIV-3 : Cosmologies homogènes et isotropes

MATHINE : Comme nous allons en avoir besoin plus tard, voici les composantes du tenseur d'Einstein qui ne sont pas nulles :

$$\begin{aligned} G_{tt} &= 3\frac{\dot{a}^2 + k}{a^2} \\ G_{ii} &= -2\frac{\ddot{a}}{a} - \frac{\dot{a}^2 + k}{a^2}, \end{aligned} \quad (1047)$$

où nous notons :

$$\begin{aligned} \dot{a} &= \frac{da}{dt} \\ \ddot{a} &= \frac{d^2a}{dt^2}. \end{aligned} \quad (1048)$$

EURISTIDE : Nous allons avoir besoin également du tenseur d'énergie-impulsion pour écrire les équations du champ d'Einstein. Pour modéliser la matière et l'énergie de l'univers à grande échelle, nous allons considérer que c'est un fluide, dont la densité est  $\rho$ , et dont la pression est  $p$  dans toutes les directions spatiales. Dans ce cas, il existe un système de coordonnées dans lequel le tenseur d'énergie-impulsion est réduit à l'expression diagonale la plus simple :

$$T_{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix}. \quad (1049)$$

Les équations du champ d'Einstein s'écrivent :

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}. \quad (1050)$$

Cela donne :

$$\begin{aligned} &\begin{pmatrix} 3\frac{\dot{a}^2+k}{a^2} & 0 & 0 & 0 \\ 0 & -2\frac{\ddot{a}}{a} - \frac{\dot{a}^2+k}{a^2} & 0 & 0 \\ 0 & 0 & -2\frac{\ddot{a}}{a} - \frac{\dot{a}^2+k}{a^2} & 0 \\ 0 & 0 & 0 & -2\frac{\ddot{a}}{a} - \frac{\dot{a}^2+k}{a^2} \end{pmatrix} \\ &= \begin{pmatrix} 8\pi G\rho & 0 & 0 & 0 \\ 0 & 8\pi Gp & 0 & 0 \\ 0 & 0 & 8\pi Gp & 0 \\ 0 & 0 & 0 & 8\pi Gp \end{pmatrix}. \end{aligned} \quad (1051)$$

Ce qui se réduit à deux équations :

$$\begin{aligned} 3\frac{\dot{a}^2 + k}{a^2} &= 8\pi G\rho \\ -2\frac{\ddot{a}}{a} - \frac{\dot{a}^2 + k}{a^2} &= 8\pi Gp. \end{aligned} \quad (1052)$$

EURISTIDE : Ces équations sont appelées **équations de Friedmann**.

MATHINE : En combinant ces deux équations on obtient :

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p). \quad (1053)$$

Par ailleurs, nous savons que l'énergie est conservée, ce qui s'exprime par le fait que la divergence du tenseur d'énergie-impulsion est nulle :

$$\nabla^\mu T_{\mu\nu} = 0. \quad (1054)$$

Le calcul donne en conséquence de l'équation de la conservation de l'énergie :

$$\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + p) = 0. \quad (1055)$$

EURISTIDE : Par ailleurs, comme nous considérons que la matière et l'énergie de l'univers constituent, à grande échelle, un gaz parfait, nous pouvons écrire l'équation d'état de ce gaz parfait qui donne l'expression de sa pression  $p$  :

$$p = \omega\rho, \quad (1056)$$

où  $\omega$  est une constante indépendante du temps.

MATHINE : Il s'ensuit que l'équation de conservation de l'énergie-impulsion s'écrit maintenant :

$$\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + \omega\rho) = 0, \quad (1057)$$

soit :

$$\frac{d}{dt}(\rho a^3) + \omega\rho \frac{d}{dt}a^3 = 0. \quad (1058)$$

Sous cette forme, on peut trouver que les solutions de cette équation sont :

$$\rho a^{3(\omega+1)} = \rho_0, \quad (1059)$$

où  $\rho_0$  est la densité de l'énergie et de la matière de l'univers à l'instant présent.

EURISTIDE : Nous pouvons distinguer trois types de fluides et donc d'univers : pour  $\omega = -1$ , le fluide est vide, et donc l'univers est vide. Pour  $\omega = 0$ , le fluide est appelé **poussière**. Pour  $\omega = \frac{1}{3}$ , le fluide est appelé **univers de rayonnement**.

Dans ce modèle, l'équation de Friedmann qui décrit le comportement de l'Univers à grande échelle, s'écrit :

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3} \frac{\rho_0}{a^{3(\omega+1)}} - \frac{k}{a^2}. \quad (1060)$$

Le taux d'expansion de l'Univers :

$$H = \frac{\dot{a}}{a}, \quad (1061)$$

est appelé **paramètre de Hubble**.

Il est intéressant également d'isoler l'expression suivante :

$$\Omega = \frac{8\pi G}{3H^2}\rho, \quad (1062)$$

qui est appelé **paramètre de densité**.

En posant pour expression de la **densité critique** :

$$\rho_C = \frac{3H^2}{8\pi G}, \quad (1063)$$

on peut écrire :

$$\Omega = \frac{\rho}{\rho_C}, \quad (1064)$$

et par suite, en transférant ces données dans l'équation de Friedmann :

$$\frac{\rho}{\rho_C} - 1 = \frac{k}{H^2 a^2}. \quad (1065)$$

Arrêtons-nous un instant sur cette équation.

Si la densité  $\rho$  est plus grande que la densité critique  $\rho_C$ , alors  $k$  est positif, et nous savons alors que l'univers est fermé. Si la densité  $\rho$  est plus petite que  $\rho_C$ , alors  $k$  est négatif et l'univers est ouvert. Si la densité  $\rho$  est égale à la densité  $\rho_C$ , alors  $k$  est nul et l'univers est plat.

**BEATRIX** : Autrement dit, le paramètre de densité détermine la topologie de notre univers.

**EURISTIDE** : Oui, et c'est essentiel. C'est la raison pour laquelle la détermination de ce paramètre de densité (et donc la densité de matière et de l'énergie) fait l'objet de recherches intenses.

Voyons maintenant comment interpréter le comportement des modèles en fonction de leurs paramètres.

Supposons pour commencer que l'Univers est rempli de fluide avec une énergie positive ( $\rho > 0$ ) et une pression positive ou nulle ( $p \geq 0$ ). Alors, d'après l'équation :

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p), \quad (1066)$$

nous déduisons que le premier membre est obligatoirement négatif, donc :

$$\ddot{a} < 0. \quad (1067)$$

Par ailleurs, nous savons que l'univers est en expansion, puisqu'il a pu être observé que les galaxies distantes s'éloignaient de nous.

Donc  $\ddot{a} < 0$  signifie que l'expansion de l'Univers est en train de décélérer. Cela signifie que l'Univers se trouve dans une phase d'expansion dont la vitesse d'expansion ralentit en permanence. Cela signifie que l'Univers a dû avoir une expansion plus rapide dans le passé. En remontant le temps, nous pouvons déduire qu'à un instant donné, l'univers a eu un facteur d'échelle égal à 0 :

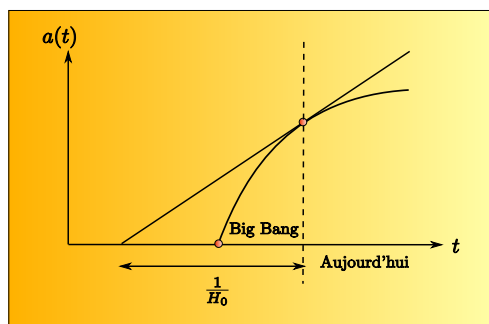


Fig. 148 - Le Big Bang

Cet événement a dû se produire après l'événement situé à un instant dans le passé de  $1/H_0$  par rapport à aujourd'hui, qui se détermine comme la rencontre de la tangente à la courbe  $a(t)$  avec l'axe  $a = 0$ . En effet, puisque  $\ddot{a} < 0$ , la courbe  $a(t)$  a sa convexité tournée vers le haut, et donc elle rencontre l'axe  $a = 0$  après l'événement situé à un instant  $1/H_0$  par rapport à aujourd'hui.

Cet événement, qui est une singularité de l'espace-temps, est appelé **Big Bang**.

BEATRIX : Le Big Bang, c'est donc la création de l'Univers ?

EURISTIDE : Oui, c'est la création de l'Univers à partir de la singularité de l'espace-temps. Mais contrairement à son nom, il ne faut pas l'imaginer comme une grosse explosion de quelque chose d'antérieur. Il s'agit seulement de la création de l'univers à partir de rien.

Considérons maintenant l'évolution de l'Univers dans le futur. Nous avons l'équation de Friedmann :

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2}. \quad (1068)$$

En multipliant les deux membres de l'équation par  $a^2$ , on obtient :

$$\dot{a}^2 = \frac{8\pi G}{3}\rho a^2 - k. \quad (1069)$$

Considérons les univers ouvert et plat, pour lesquels  $k \leq 0$ , et considérons toujours des univers pour lesquels la densité d'énergie est positive. Alors, le membre de droite de l'équation est forcément strictement positif. Donc  $\dot{a}$  ne s'annule jamais. Physiquement, cela veut dire que l'Univers poursuivra son expansion indéfiniment.

BEATRIX : Par conséquent, les univers ouvert ou plat dont la densité d'énergie est positive sont en expansion perpétuelle. Qu'en est-il des univers clos ?



EURISTIDE : Pour un univers clos, nous savons que  $k = +1$ , donc l'équation de Friedmann s'écrit :

$$\dot{a}^2 = \frac{8\pi G}{3}\rho a^2 - 1. \quad (1070)$$

Considérons la quantité  $\rho a^3$  :

$$\frac{d}{dt}(\rho a^3) = a^3(\dot{\rho} + 3\rho\frac{\dot{a}}{a}). \quad (1071)$$

Or, la loi de conservation de l'énergie nous a permis tout à l'heure de montrer que :

$$\dot{\rho} = -3\frac{\dot{a}}{a}(\rho + p). \quad (1072)$$

Donc :

$$\begin{aligned} \frac{d}{dt}(\rho a^3) &= a^3(3\rho\frac{\dot{a}}{a} - 3\rho\frac{\dot{a}}{a} - 3p\frac{\dot{a}}{a}) \\ &= -3pa^2\dot{a}. \end{aligned} \quad (1073)$$

L'univers est en expansion, donc  $\dot{a} > 0$ , et la pression du fluide que constitue l'Univers est supposée positive ou nulle. Donc :

$$\frac{d}{dt}(\rho a^3) \leq 0. \quad (1074)$$

Donc, la quantité  $\rho a^3$  doit diminuer lorsque le temps s'écoule, et donc lorsque  $a$  croît vers l'infini.

Par conséquent, il en va de même pour  $\rho a^2$ , qui doit tendre vers 0. Mais cela conduirait à une incohérence, puisque l'équation de Friedmann donnerait :

$$\dot{a}^2 \rightarrow -1, \quad (1075)$$

ce qui est impossible.

Donc, il existe un instant où  $a(t)$  ne croît plus. L'Univers clos n'est pas en expansion indéfiniment.  $a(t)$  atteint une valeur maximale que nous noterons  $a_{max}$ .

Reprenons l'équation de Friedmann :

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p). \quad (1076)$$

Alors, pour  $a$  tendant vers sa valeur maximale  $a_{max}$ , la dérivée seconde de  $a$  tend vers :

$$\ddot{a} \rightarrow -\frac{4\pi G}{3}(\rho + 3p)a_{max}. \quad (1077)$$

Cela signifie qu'au voisinage de  $a_{max}$ ,  $\ddot{a}$  est négative. Par conséquent, la fonction  $a(t)$  atteint un maximum et recommence à décroître à partir de cette valeur.

BEATRIX : Jusqu'où décroît-elle ?

EURISTIDE : Elle continue de décroître sans contrainte jusqu'à la valeur 0. Par conséquent, pour un espace clos dont la densité d'énergie et la pression sont positives, l'Univers est en expansion depuis le Big Bang jusqu'à une valeur limite de son expansion, puis commence à se contracter jusqu'à zéro, pour atteindre une nouvelle singularité appelée **Big Crunch**.

## 88 Scène XIV-4 : La loi de Hubble

BEATRIX : Pour résumer, je retiens que dans l'hypothèse d'un univers à densité positive et à pression positive, les modèles d'univers plat et ouvert sont en expansion perpétuelle, tandis que le modèle d'univers clos termine son expansion dans le futur pour se contracter à nouveau. Cela découle d'un peu d'intuition, puisqu'on peut s'imaginer que l'univers clos possède une densité plus importante, telle que les effets de gravitation parviennent à contrecarrer l'expansion.

EURISTIDE : L'idée est bonne. Et c'est vrai que la densité a un rôle à jouer dans le comportement de l'univers. Nous avons vu que le signe de  $k$  dépendait du rapport de la densité à la densité critique. C'est bien donc que l'univers clos correspond à un univers de densité critique plus importante que les univers plat ou ouvert.

Nous allons maintenant analyser la vitesse d'expansion de l'univers et ses conséquences.

Considérons une galaxie lointaine située à une distance  $\chi$  sur l'axe radial par rapport à un observateur situé en  $r = 0$ .

La distance propre de la galaxie est donnée par :

$$D = a(t)\chi. \quad (1078)$$

La vitesse de la galaxie pour l'observateur immobile est :

$$v = \frac{dD}{dt} = \frac{\dot{a}}{a}a(t)\chi = HD, \quad (1079)$$

où  $H$  est le paramètre de Hubble que nous avons défini précédemment.

La valeur actuelle du paramètre de Hubble est notée  $H_0$  et est approximativement égale à :

$$H_0 = 21 \text{ km/s par million d'années lumière.} \quad (1080)$$

La relation :

$$v = HD, \quad (1081)$$

est appelée **loi de Hubble** et exprime que la vitesse d'une galaxie est proportionnelle à sa distance par rapport à l'observateur.

Hubble avait déduit cette loi de l'observation en 1929 et ce résultat a constitué alors une preuve de l'expansion de l'Univers, avant même que la relativité générale puisse le prédire. D'ailleurs Einstein lui-même pensait que l'Univers était statique.

Nous allons maintenant regarder comment l'expansion cosmologique produit un décalage vers le rouge de la lumière.

La lumière voyage le long de géodésiques, et les rayons lumineux qui s'approchent de nous se déplacent suivant les géodésiques d'équation :

$$ds^2 = 0, \quad (1082)$$

avec :

$$\begin{aligned} d\theta &= 0 \\ d\phi &= 0. \end{aligned} \quad (1083)$$

Par conséquent, d'après la métrique de l'Univers :

$$dt = -a(t)d\chi. \quad (1084)$$

Considérons la période  $\Delta t_e$  de l'événement lumineux émis. Et considérons la période  $\Delta t_0$  de l'événement reçu par l'observateur :

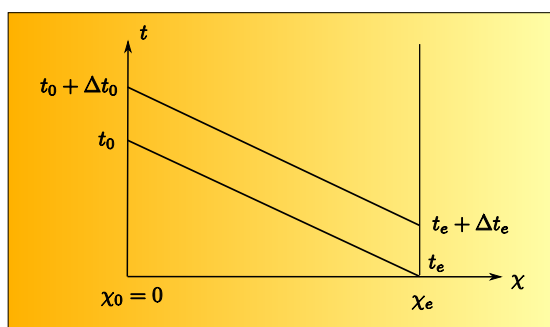


Fig. 149 - Emission de deux signaux lumineux

Par conséquent, aux instants  $t_e$  et  $t_e + \Delta t_e$ , sont émis deux signaux lumineux qui sont reçus aux instants respectifs  $t_0$  et  $t_0 + \Delta t_0$ .

La distance radiale  $\chi_e$  peut être calculée de deux façons différentes :

$$\chi_e = \int_{t_2}^{t_0} \frac{dt}{a(t)} \quad (1085)$$

$$\chi_e = \int_{t_2 + \Delta t_e}^{t_0 + \Delta t_0} \frac{dt}{a(t)}.$$

Par conséquent, en retranchant membre à membre les deux équations, on obtient :

$$\int_{t_e + \Delta t_e}^{t_0 + \Delta t_0} \frac{dt}{a(t)} - \int_{t_e}^{t_0} \frac{dt}{a(t)} = 0, \quad (1086)$$

ce qui peut s'écrire encore :

$$\int_{t_0}^{t_0 + \Delta t_0} \frac{dt}{a(t)} = \int_{t_e}^{t_e + \Delta t_e} \frac{dt}{a(t)}. \quad (1087)$$

On peut considérer par approximation que  $a(t)$  est constant dans les intervalles de temps petits  $[t_0, t_0 + \Delta t_0]$  et  $[t_e, t_e + \Delta t_e]$ .

Soient  $a(t_0)$  et  $a(t_e)$  les valeurs respectives du facteur d'expansion. Alors, nous pouvons écrire :

$$\frac{\Delta t_e}{a(t_e)} = \frac{\Delta t_0}{a(t_0)}. \quad (1088)$$

En considérant que la longueur d'onde s'écrit :

$$\lambda = c\Delta t, \quad (1089)$$

nous obtenons :

$$\frac{\lambda_0}{\lambda_e} = \frac{a(t_0)}{a(t_e)}. \quad (1090)$$

Cette équation montre que les ondes lumineuses sont étirées par l'expansion de l'espace. Par conséquent, la lumière est décalée vers le spectre rouge.

**BEATRIX** : C'est ce qu'on appelle le **décalage vers le rouge cosmologique**.

**EURISTIDE** : Voilà. Je crois que nous avons fini d'étudier les principaux aspects de la théorie de la relativité générale et de ses conséquences. Nous n'avons pas abordé certains points plus complexes, notamment l'analyse des modèles d'univers spécifiques, tels que ceux à rayonnement dominant, ou à matière dominante, ou à énergie du vide dominante. Ce sont des développements plus complexes, mais tu possèdes maintenant les bases mathématiques pour les aborder et en comprendre le sens physique.

Je te souhaite donc bon courage et plein de plaisir et d'émerveillement pour la suite de ton parcours.

**BEATRIX** : Merci à tous les deux. C'était un beau voyage, plein de surprises et de rebondissements!

## Index

- écliptique, 111
- élan, 87
- énergie, 90
- énergie cinétique, 93
- énergie de repos, 166
- énergie potentielle, 95
- énergie potentielle de déformation, 93
- équation d'Einstein, 239
- équation d'Einstein dans le vide, 239
- équation d'Euler-Lagrange, 118
- équation de Lagrange, 120
- équation de Poisson, 232
- équation différentielle, 49
- équation du transport parallèle, 214
- équation géodésique, 216
- équations de Friedmann, 286
- équations linéarisées d'Eistein, 247
- équilibre d'un corps, 76
- équilibre sur un axe, 105
- étoile binaire, 256
- événement de genre espace, 148
- événement de genre lumière, 148
- événement de genre temps, 148
  
- aberration de la lumière, 158
- accélération, 66
- action, 79
- addition de matrices, 61
- addition de tenseurs, 136
- addition de vecteurs, 57
- aphélie, 111
- atlas, 172
- axes, 17
  
- balle ouverte, 169
- Big Bang, 288
- Big Crunch, 289
- bras de levier, 105
  
- cône de lumière, 148
- calcul matriciel, 60
- calcul tensoriel, 122
- carte, 169
- carte injective, 171
- champ vectoriel, 197
- champ vectoriel de Killing, 199
- changement de coordonnées, 29
- changement de coordonnées curviligne, 130
  
- cinématique, 63
- composantes covariantes d'un vecteur, 140
- conservation de l'énergie, 96
- conservation de la quantité de mouvement, 86
- contraction de Lorentz, 157
- contraction de tenseurs, 137
- contravariant, 133
- convention d'Einstein, 122
- coordonnées, 16
- coordonnées d'Eddington-Finkelstein, 281
- coordonnées de Gauss, 128
- coordonnées normales riemaniennes, 186
- coordonnées polaires, 31
- corps solide, 101
- cosmologie, 282
- courbes intégrales, 197
- courbure concave, 26
- courbure convexe, 26
- courbure d'un espace, 23
- covariant, 132
  
- Décalage vers le rouge cosmologique, 292
- décalage vers le rouge gravitationnel, 230
- Déflexion de la lumière, 277
- dérivée covariante, 205
- dérivée d'un produit, 43
- dérivée d'une fonction, 40
- dérivée d'une somme, 43
- dérivée de Lie, 198
- dérivée du produit d'une fonction par un nombre, 43
- dérivée partielle, 45
- déterminant d'une matrice, 62
- développement de Taylor, 38
- déviatoin géodésique, 250
- degré de liberté, 15
- densité critique, 287
- densité de Lagrange, 240
- deuxième loi de Kepler, 111
- difféomorphisme, 194
- différentielle d'une fonction, 44
- différentielle d'une fonctionnelle, 114
- dimension, 15
- distance entre points, 23
- dual de l'espace tangent, 180
- dynamique, 75
  
- ensemble ouvert, 170
- espace cotangent, 179

- espace curviligne, 24  
espace de Minkowski, 162  
espace tangent, 28, 177  
espace vectoriel, 59  
espace-temps, 19, 143  
extrémale d'une fonctionnelle, 115
- facteur d'échelle, 283  
foliation de l'espace, 262  
foliation de l'espace-temps, 283  
fonction, 32  
fonction négligeable devant une autre, 38  
fonctionnelle, 113  
fonctions équivalentes, 38  
force, 75  
forme quadratique fondamentale, 139  
front d'onde, 228
- générateur, 197  
géodésique, 215  
gradient du potentiel gravitationnel, 232  
gravitation terrestre, 80
- horizon cosmologique, 149  
horizon des événements, 281
- impulsion d'une force, 87  
impulsion quadripôle, 255  
intégrale d'une fonction, 47  
intégrale triple, 48  
intervalle entre 2 événements, 146  
isométrie, 199
- jauge, 245  
jauge d'Einstein, 246  
jauge de Lorentz, 246  
jauge harmonique, 245
- laplacien du potentiel gravitationnel, 233  
ligne d'univers, 143  
limite d'une fonction, 34  
limite en  $+\infty$  d'une fonction, 36  
limite en 0 d'une fonction, 35  
loi de Hubble, 290
- méridiens, 129  
métrique de Minkowski normalisée, 236  
métrique de Robertson-Walker, 284  
métrique de Schwarzschild, 268  
métrique lorentzienne, 183  
métrique riemannienne, 183  
métrique stationnaire, 267
- métrique statique, 267  
matrice, 60  
modèle d'univers clos, 284  
modèle d'univers ouvert, 284  
modèle d'univers plat, 284  
module d'un vecteur, 54  
moment d'inertie, 108  
moment par rapport à un axe, 105  
moment rotatif, 107  
mouvement lié, 84  
multiplication d'un vecteur par un nombre, 59  
multiplication de matrices, 61  
multiplication de tenseurs, 136
- onde gravitationnelle, 249  
origine, 16
- périhélie, 111  
parallèles, 129  
paramètre de Hubble, 287  
plan tangent, 177  
plan tangent à une surface, 28  
platitude asymptotique, 269  
plongement, 174  
points d'univers, 143  
potentiel gravitationnel, 232  
précession du périhélie de Mercure, 280  
première loi de Kepler, 110  
première loi de Newton, 77  
principe d'équivalence d'Einstein, 226  
principe de covariance, 228  
principe de moindre action de Hamilton, 121  
principe de relativité, 142  
principe de relativité d'Einstein, 142  
principes cosmologiques, 282  
produit scalaire de vecteurs, 91  
produit vectoriel, 103  
pullback, 188  
pulsar, 259  
pushforward, 190
- quadri-impulsion, 167  
quadrivitesse, 162  
quantité de mouvement, 78, 84
- réaction, 79  
référentiel galiléen, 97  
référentiel inertiel local, 227  
référentiel localement inertiel, 227  
référentiel non galiléen, 98  
référentiel en rotation, 99

- rayon de Schwarzschild, 248  
rayonnement gravitationnel, 255  
relativité restreinte, 142  
rotation, 101
- scalaire, 59  
scalaire de Ricci, 223  
seconde identité de Bianchi, 223  
seconde loi de Newton, 78  
singularité de coordonnées, 269  
singularité de l'espace-temps, 269  
singularité physique, 269  
symétrie, 199  
symbole d'antisymétrie, 123  
symbole de Kronecker, 123  
symbole de Levi-Civita, 186  
symboles de Christoffel, 206  
système de coordonnées, 16
- tangente à une courbe, 27  
temps propre, 150  
tenseur, 133  
tenseur contravariant, 133  
tenseur covariant, 133  
tenseur d'énergie-impulsion, 228  
tenseur d'Einstein, 223  
tenseur de Ricci, 223  
tenseur de Riemann, 220  
tenseur de Weyl, 224  
tenseur des contraintes, 242  
tenseur euclidien fondamental, 139  
tenseur métrique, 180  
théorème d'Ostrogradski, 241  
théorème de Birkhoff, 269  
théorème de Frobenius, 262  
théorème de Ricci, 209  
torision, 209  
trajectoire, 68  
trajectoire d'un point, 63  
transformation de jauge, 245  
transformation de l'énergie, 90  
transformations de Lorentz, 156  
translation, 101  
transmission de la quantité de mouvement, 87  
transport parallèle, 211  
travail, 90  
travail d'une force, 90  
troisième loi de Kepler, 112  
troisième loi de Newton, 79  
trou noir, 280
- univers de poussière, 286  
univers vide, 286
- valence d'un tenseur, 136  
variété différentiable, 173  
variété maximale symétrique, 224  
vecteur, 54  
vecteur d'onde, 249  
vecteur de Killing, 199  
vecteur position, 69  
vitesse angulaire, 103  
vitesse d'un point, 64  
voisinage, 37