Introduction au calcul tensoriel

Applications à la physique

Claude Semay • Bernard Silvestre-Brac

Introduction au calcul tensoriel

Applications à la physique

3e édition

DUNOD

Graphisme de couverture : Elizabeth Riba

11, rue Paul Bert 92240 Malakoff
www.dunod.com

ISBN 978-2-10-088576-3

Table des matières

AVANT-PROPOS vii

Chapitre 1 Rappels et conventions d'écriture 1
1.1 Convention d'Einstein 1
1.2 Symbole de Kronecker 4
1.3 Symbole de Levi-Civita 5
1.4 Éléments de calcul matriciel 6
1.5 Dérivées et différentielles 9
Exercices 11

Chapitre 2 Espaces vectoriels 14
2.1 Vecteurs de l'espace ordinaire 14
2.2 Définitions et premières propriétés 16
2.3 Sous-espaces vectoriels 19
2.4 Base et dimension 23
2.5 Changement de base 27
2.6 Notation matricielle 30
Exercices 32

Chapitre 3 Dualité 35
3.1 Formes linéaires 35
3.2 Espace dual 37
3.3 Base duale 38
3.4 Changement de base 39
3.5 Bidualité 40
Exercices 40

Chapitre 4 Algèbre tensorielle 44
4.1 Produit tensoriel de 2 espaces vectoriels 44
4.2 Tenseurs attachés à un espace vectoriel 50
4.3 Opérations sur les tenseurs 59
4.4 Tenseurs symétriques et antisymétriques 64
Exercices 66

Chapitre 5 Produit scalaire 71
5.1 Définitions 71
5.2 Orthogonalité 74
5.3 Base orthonormée 78
5.4 Changement de variance 82
5.5 Base réciproque 88
5.6 Applications linéaires et problèmes aux valeurs propres 91
Exercices 96

Chapitre 6 Éléments d'algèbre extérieure 105
6.1 Algèbre extérieure d'ordre 2 105
6.2 Tenseurs complètement antisymétriques 109
6.3 Produit vectoriel 115
Exercices 117

Chapitre 7 Espaces ponctuels 121
7.1 Espace ponctuel affine 121
7.2 Repères dans un espace ponctuel 122
7.3 Coordonnées rectilignes 125
7.4 Espace ponctuel euclidien 125
7.5 Volume d'un parallélépipède 127
7.6 Espaces $\mathcal{E}_3$ dextrogyre et lévogyre 128
7.7 Champs tensoriels 131
7.8 Dimension d'une grandeur physique 133
7.9 Repère orthonormé 134
Exercices 134

Chapitre 8 Coordonnées curvilignes dans un espace euclidien 137
8.1 Repère naturel 138
8.2 Changement de coordonnées curvilignes 144
8.3 Dérivée covariante d'un champ de vecteurs 148
8.4 Dérivée covariante d'un tenseur 158
8.5 Opérateurs différentiels 161
8.6 Relations particulières pour l'espace $\mathcal{E}_3$ 167
8.7 Espace de Riemann 172
Exercices 176

Chapitre 9 Intégration des champs tensoriels 183
9.1 Intégration d'un champ tensoriel sur une variété 183
9.2 Formules pour l'espace à trois dimensions 194
9.3 Interprétation physique des opérateurs différentiels 197
Exercices 201

Chapitre 10 Applications à la physique 207
10.1 Expression des lois physiques 207
10.2 Tenseur d'inertie 209
10.3 Théorie de l'élasticité 212
10.4 Réseau réciproque 218
10.5 Relativité restreinte 220
Exercices 232

SOLUTIONS DES EXERCICES 241

BIBLIOGRAPHIE 259

INDEX 261

Avant-propos

On rapporte, à propos d'une peinture (sans doute abstraite), le dialogue suivant entre un homme et Groucho Marx [1] :

L'HOMME – Un enfant de cinq ans comprendrait ça.
GROUCHO – Vite, allez me chercher un enfant de cinq ans !

Il ne fait aucun doute que cet échange verbal pourrait avoir lieu entre un expert du calcul tensoriel et un novice en la matière – ayant le sens de la répartie. Cette théorie a en effet mauvaise réputation auprès des étudiants – et pas seulement auprès d'eux. Elle passe pour très abstraite et fait peur, et cela depuis longtemps. En 1920, Alfred North Whitehead, un spécialiste de la théorie de la relativité, écrivait [2] :

« Ce n'est pas aller trop loin de dire que l'annonce que les physiciens auraient à étudier dans le futur la théorie des tenseurs créa une véritable panique parmi eux lorsque fut annoncée pour la première fois la vérification des prédictions d'Einstein. »

Le but de cet ouvrage est de montrer que, loin d'être une théorie abstraite, le calcul tensoriel est en fait un outil puissant – et bien plus encore – qui ne demande qu'à être maîtrisé – avec un minimum d'efforts – pour vous emmener plus loin sur les chemins de la connaissance. Mais commençons par le commencement. Qu'est-ce que le calcul tensoriel et à quoi peut-il bien servir ?

Au-delà de son intérêt mathématique, la nécessité d'utiliser le calcul tensoriel en physique est une conséquence directe du postulat fondamental suivant :

Une grandeur physique possède une existence intrinsèque, indépendante de tout système de référence.

1. Les Marx Brothers, *Pensées, répliques et anecdotes*, Éditions J'ai lu, 1999.
2. Alfred North Whitehead, *The Concept of Nature*, Cambridge University Press, 1920, p. 182. Cité en p. 190 dans : Jeffrey Crelinsten, « Physicists receive relativity: Revolution and reaction », *The Physics Teacher*, vol. 18, mars 1980, p. 187-193.

La validité de cette prescription semble difficilement contestable si on veut pouvoir donner un sens aux phénomènes physiques et être capable de faire des prédictions. Cependant, la manipulation, voire la définition, d'une grandeur physique requièrent généralement l'utilisation d'un repère (ou référentiel) particulier muni d'un système de coordonnées adéquat. Le choix de ce système de référence est souvent dicté par la nature même du problème. Ainsi, l'étude du mouvement d'une particule test dans un champ de force central se fera plus aisément dans un système de coordonnées sphériques dont l'origine est située sur la source du champ de force, que dans un repère cartésien. Les conclusions auxquelles on peut aboutir dans ces deux référentiels, ou dans n'importe quel autre repère, sont pourtant identiques. On peut donc traduire le postulat ci-dessus par l'énoncé équivalent :

Aucun système de référence n'est privilégié.

Évidemment, la valeur d'une quantité physique dépend du système de référence dans lequel elle est mesurée. Pensons à la position d'un point qui dépend du choix des axes du repère, à la vitesse d'un objet qui dépend de l'état de mouvement de l'observateur par rapport à cet objet, etc. Il convient donc d'établir avec soin les lois de transformation des grandeurs considérées lorsque l'on passe d'un référentiel à un autre. La traduction mathématique d'une loi physique – supposée universelle – se réduisant souvent à une égalité, on arrive donc à énoncer ce principe fondamental :

Une équation ne peut être l'expression d'une loi physique que si elle se ramène à l'égalité de deux grandeurs qui se transforment de la même manière lors d'un changement de référentiel. [3]

On peut de ce fait classer les quantités physiques suivant leur loi de transformation. Les nombres (ou scalaires) se différencient ainsi nettement des vecteurs. Mais on est amené à introduire d'autres grandeurs plus complexes, moins familières, les tenseurs, dont le nom provient de l'étude des forces de pression ou des tensions internes qui s'exercent dans les milieux continus. L'art de manipuler correctement ces tenseurs – dont scalaires et vecteurs sont des représentants particuliers – constitue ce qu'on appelle le calcul tensoriel.

Le calcul tensoriel est donc le formalisme mathématique qui permet d'exprimer les lois de la physique sous une forme indépendante du système de référence choisi. Bien que les prescriptions énoncées plus haut semblent assez naturelles, ce n'est que récemment dans l'histoire des sciences, avec l'avènement des théories relativistes,

3. Il faut également que les grandeurs, présentes de part et d'autre du signe égal, aient le même contenu dimensionnel (voir la section 7.8). Pour effectuer un calcul numérique correct, il faudra de plus s'assurer que le système d'unités choisi est cohérent. Enfin, des contraintes supplémentaires, liées à la structure même de notre univers, sont également à prendre en compte (voir le chapitre 10).

qu'elles ont été érigées en principes fondamentaux contraignants pour toute théorie physique. C'est l'existence préalable du calcul tensoriel qui a permis le développement de la théorie de la relativité restreinte et sa généralisation en une théorie englobant la gravitation, la relativité générale. En retour, cette dernière a été le moteur d'un essor important du calcul tensoriel. Ce formalisme est rapidement devenu un instrument essentiel de la physique théorique moderne et un outil puissant pour tout scientifique, chercheur, ingénieur, enseignant et étudiant.

Bien souvent, un calcul avec des grandeurs vectorielles se résume à manipuler des composantes, et il en va de même avec les grandeurs tensorielles. Il semble ainsi, à première vue, que l'on puisse se contenter d'un ensemble de recettes pratiques pour aboutir au résultat voulu. Il faut cependant aller bien au-delà pour réellement tirer profit de la théorie. Nous ne proposons pas non plus d'adopter ici une démarche purement mathématique qui consisterait à n'introduire que des quantités abstraites et à définir des opérations sur celles-ci. Nous préférons tenter le pari d'un juste milieu en introduisant les grandeurs de manière intrinsèque (tenseurs) dans un formalisme mathématique rigoureux et en les traduisant ensuite en termes de quantités dépendant d'une représentation donnée (composantes) mais aussi plus aisément manipulables.

Ce livre est divisé en deux parties. Dans la première, nous construirons les grandeurs tensorielles, en toute généralité, en nous appuyant sur le concept d'espace vectoriel (chapitres 1 à 6). Dans la deuxième partie, nous utiliserons ces notions pour bâtir des systèmes de coordonnées dans les espaces de points servant de cadre aux phénomènes physiques et donnerons des exemples d'application du calcul tensoriel (chapitres 7 à 10). Dans un souci de simplification, avec pour but de donner un premier aperçu élémentaire de cette théorie générale, nous nous limiterons aux espaces dits « plats » ou « sans courbure » : les espaces euclidiens, comme celui de la géométrie ordinaire ou de la mécanique classique, et les espaces pseudo-euclidiens, comme l'espace-temps de la théorie de la relativité restreinte. Les espaces non euclidiens, comme ceux nécessaires à l'élaboration de la théorie de la relativité générale, ne seront qu'évoqués dans cet ouvrage.

Nous proposons ici un petit guide de lecture. Certains mots importants, souvent propres au calcul tensoriel, sont écrits en italique (dans un texte en romain et vice versa) lorsqu'ils apparaissent pour la première fois dans le texte. Nous insistons sur un mot ou un bout de phrase en l'écrivant en gras. Adopter une bonne notation, c'est souvent faciliter la compréhension du lecteur et limiter les risques d'erreurs lors de la manipulation des formules. Nous avons donc apporté un soin tout particulier aux conventions d'écriture. De plus, le texte est truffé, de-ci de-là, d'encarts portant un titre, « Remarque concernant les notations », qui suffit à préciser leur utilité. Des exemples seront également insérés en encart pour illustrer des notions théoriques qui viennent d'être exposées.

Une bonne compréhension d'une théorie se traduit par la capacité à résoudre des problèmes spécifiques. À la fin de chaque chapitre, des exercices sont donc proposés au lecteur. Ils sont classés, de manière approximative et subjective, par ordre de difficulté croissante. Les exercices sont par nature une œuvre collective : certains de ceux que nous proposons sont inspirés de problèmes parus dans la littérature ; d'autres sont extraits d'examens dont nous avons eu connaissance ; quelques-uns sont de notre cru. Des notes sur ces exercices, donnant les solutions – quand elles ne sont pas contenues dans l'énoncé du problème ou dans le texte –, sont rassemblées à la fin de l'ouvrage.

Nos plus vifs remerciements vont à nos professeurs, qui nous ont fait découvrir cette belle théorie, et à toutes les personnes – étudiants, enseignants et amis –, qui nous ont prodigué remarques et conseils avisés. En particulier, nous exprimons notre gratitude à Fabien Buisseret, Damien Galant et Vincent Mathieu pour leur lecture attentive du manuscrit.

Chapitre 1

Rappels et conventions d'écriture

La matière abordée dans cet ouvrage repose en grande partie sur l'algèbre linéaire et sur les propriétés élémentaires des fonctions à plusieurs variables. Ces notions sont supposées connues et ce chapitre n'en est qu'un bref rappel qui servira surtout à fixer nos conventions. De plus, les entités mathématiques présentées peuvent se révéler très lourdes à manipuler si un symbolisme adéquat n'est pas utilisé. Ce chapitre est destiné à clarifier toutes ces notations.

1.1 CONVENTION D'EINSTEIN

En physique, les phénomènes sont décrits à l'aide de variables, qui peuvent être dynamiques ou de simples paramètres. Lorsqu'un grand nombre de variables entrent en jeu, ou lorsque ces variables se réfèrent à des quantités présentant des similitudes, il n'est pas judicieux de les nommer par des lettres différentes. On préfère dans ce cas utiliser la même lettre et distinguer chaque variable par un *indice*. Ainsi pour définir un point de l'espace on utilisera plutôt la notation x^1, x^2, x^3 au lieu de x, y, z. De même, la vitesse $\vec{v}$ sera décomposée en ses composantes v^1, v^2, v^3.

Définition 1.1

On appelle suite ***contravariante*** *la donnée de n quantités étiquetées par la même lettre et qui diffèrent par un indice placé en* ***position supérieure*** *:* $\{a^1, a^2, \ldots, a^n\}$.

On appelle suite ***covariante*** *la donnée de n quantités étiquetées par la même lettre et qui diffèrent par un indice placé en* ***position inférieure*** *:* $\{b_1, b_2, \ldots, b_n\}$.

On peut déjà faire quelques remarques :

– Les indices sont toujours des nombres entiers ; le plus souvent ils sont représentés par des lettres latines $i, j, \dots, k$, ou par des lettres elles-mêmes indicées (toujours en bas) $i_1, i_2, \dots, i_k$.
– Les indices supérieurs d'une suite indicée ne doivent pas être confondus avec une puissance de la quantité : x^3 indiquera généralement la variable nommée « x^3 » et non le cube de la variable x ; l'indication d'une puissance pour une quantité indicée se fera en la mettant entre parenthèse, ainsi la notation $\left(x^4\right)^2$ signifie le carré de la variable x^4.
– Ce n'est pas pour le plaisir de compliquer que l'on utilise un formalisme à base d'indices et que l'on fait un distinguo entre indices supérieurs (contravariants) et inférieurs (covariants). En algèbre tensorielle, c'est une nécessité capitale. La différence entre indices contravariants et indices covariants sera développée plus loin.

On peut généraliser cette notion de suite au cas où plusieurs indices de type contravariant ou covariant apparaissent. Par exemple, dans le cas de deux indices, on peut poser :

Définition 1.2

*On appelle suite **deux fois contravariante** la donnée de n^2 quantités étiquetées par la même lettre $\{a^{11}, \dots, a^{1n}, a^{21}, \dots, a^{2n}, \dots, a^{n1}, \dots, a^{nn}\}$ et qui diffèrent par 2 indices placés en **position supérieure**. Le terme générique est noté a^{ij}.*

*On appelle suite **deux fois covariante** la donnée de n^2 quantités étiquetées par la même lettre $\{a_{11}, \dots, a_{1n}, a_{21}, \dots, a_{2n}, \dots, a_{n1}, \dots, a_{nn}\}$ et qui diffèrent par 2 indices placés en **position inférieure**. Le terme générique est noté a_{ij}.*

*On appelle suite mixte **une fois contravariante et une fois covariante** la donnée de n^2 quantités étiquetées par la même lettre $\{a_1^1, \dots, a_1^n, a_2^1, \dots, a_2^n, \dots, a_n^1, \dots, a_n^n\}$ et qui diffèrent par 1 indice placé en **position supérieure** et 1 indice placé en **position inférieure**. Le terme générique est noté a_i^j.*

Nous verrons plus loin comment généraliser ces notions à une *variance* quelconque, avec un nombre arbitraire d'indices contravariants et covariants. Nous établirons également une distinction entre les éléments notés a_j^i, $a^i{}_j$ et $a_i{}^j$, avec des indices haut et bas en 1re ou en 2e position.

L'utilisation d'indices se révèle très pratique lorsqu'il s'agit de faire des sommes de quantités ; on se sert alors de la lettre grecque Σ. Ainsi, on note

$$a_1 + a_2 + \cdots + a_n = \Sigma_{i=1}^n a_i. \tag{1.1}$$

Une telle écriture se lit « somme sur i allant de 1 à n des quantités a_i ». La représentation symbolique Σ est réservée à l'opération d'addition, notée par le symbole +. Très souvent les quantités a_i sont des nombres réels et la loi binaire + est l'addition traditionnelle. Mais il arrive aussi souvent que les a_i soient des éléments d'un autre ensemble. La notation précédente n'a de sens que si on a défini, au préalable, dans cet ensemble une opération + nommée addition, qui possède la propriété d'être associative et commutative.

En algèbre tensorielle, nous serons très souvent amenés à effectuer des sommes portant sur des monômes (cela suppose que l'on ait aussi défini une « multiplication » agissant sur les quantités, multiplication indiquée habituellement par une absence de symbole) formés de quantités indicées. Parfois la sommation peut porter sur plusieurs indices. Par exemple, on peut avoir

$$S^i = a_1^i b_1^1 c^1 + a_1^i b_2^1 c^2 + a_2^i b_1^2 c^1 + a_2^i b_2^2 c^2 = \Sigma_{j=1}^2 \left(\Sigma_{k=1}^2 a_j^i b_k^j c^k \right) = \Sigma_{k=1}^2 \left(\Sigma_{j=1}^2 a_j^i b_k^j c^k \right). \tag{1.2}$$

Puisque l'addition est commutative et que les sommes sont finies, l'ordre dans lequel on effectue les opérations est sans importance, ce qui justifie les deux écritures en termes de Σ. On peut profiter de cette propriété pour écrire l'égalité précédente de façon encore plus concise

$$S^i = \Sigma_{j,k=1}^2 a_j^i b_k^j c^k, \tag{1.3}$$

où on a regroupé sous le même symbole unique Σ une somme double portant sur les deux indices j et k.

Pour manipuler les quantités apparaissant en algèbre tensorielle, même cette notation concise s'avère encore trop lourde. Einstein a imaginé une notation réduite à sa plus simple expression en proposant d'écrire la double somme précédente sous la forme

$$S^i = a_j^i b_k^j c^k. \tag{1.4}$$

Tout symbole de sommation a désormais disparu ; ne restent que le monôme générique et les indices concernés. Cette façon d'écrire les sommations s'appelle la *convention d'Einstein*, qui peut être énoncée de la manière générale suivante :

Définition 1.3 *Dans le cadre de la convention d'Einstein, toute expression de type monôme indicé contenant un indice contravariant et un indice covariant identiques doit être comprise comme une sommation sur cet indice. m sommations doivent être effectuées s'il y a m paires différentes d'indices identiques.*

On peut faire quelques commentaires utiles :

- Du fait de son caractère ultra-simplifié, la convention d'Einstein a fait perdre des informations : les valeurs sur lesquelles courent les indices de sommation ne figurent nulle part. Ce n'est pas un réel handicap car cette information doit être évidente d'après le contexte et est donc toujours sous-entendue ; en pratique tous les indices courent sur les mêmes valeurs (en général de 1 à n).
- La sommation court sur **toutes** les valeurs affectées aux indices ; si elle n'affecte qu'une partie des indices, le symbole $\sum$ est réintroduit.

– **Les indices concernés par la sommation doivent être l'un en position supérieure, l'autre en position inférieure**. Les raisons de cette contrainte apparaîtront plus tard. Ainsi $a_i b^i$ signifie $\Sigma_{i=1}^n a_i b^i$, mais l'expression $a^i b^i$ désigne simplement un monôme unique. Si l'on veut effectivement effectuer une sommation sur ce type de monôme, on utilisera explicitement le symbole de sommation $\Sigma_{i=1}^n a^i b^i$.
– Dans un monôme générique, il existe en général des paires d'indices répétés et des indices non répétés. Les indices répétés, appelés *indices muets*, doivent être compris comme des indices de sommation. En tant que tels on peut leur affecter une lettre arbitraire (puisqu'au bout du compte toutes les valeurs disponibles vont être utilisées), à condition de prendre **une lettre différente** pour des sommations courant sur des indices différents. Par opposition, les indices non répétés, appelés *indices libres*, possèdent une valeur bien précise qui est celle représentée par la lettre qui leur est affectée. Ces valeurs sont en quelque sorte figées et l'expression finale dépend de ces indices, qu'on ne peut donc pas changer de façon arbitraire. Ainsi l'expression $a_{ij} b_k^j c^{kl}$, qui s'écrit explicitement $\Sigma_{j,k=1}^n a_{ij} b_k^j c^{kl}$, possède deux paires d'indices muets j, k et deux indices libres i, l. Le résultat final dépend de i et l et peut s'écrire par exemple S_i^l. Il est important dans le résultat final de garder le caractère de variance (soit contravariant, soit covariant) à chaque indice libre. L'expression originale peut tout aussi bien s'écrire $a_{im} b_n^m c^{nl}$ (changement de notation des indices muets), mais en aucun cas $a_{ij} b_j^j c^{jl}$ (même notation pour deux indices muets qui conduit à une ambiguïté pour les sommations) ou $a_{mj} b_k^j c^{kn}$ (changement de notation pour les indices libres qui donne un résultat S_m^n différent de celui attendu S_i^l).

La convention d'Einstein, par sa grande simplicité, est utilisée de façon systématique dans les calculs concernant l'algèbre tensorielle. Elle sera la règle dans toute la suite de cet ouvrage.

1.2 SYMBOLE DE KRONECKER

Très fréquemment, il arrive qu'une quantité étiquetée par deux indices soit nulle si les indices sont différents, et possède la valeur 1 lorsqu'ils sont identiques. Cette quantité est généralement représentée par la lettre grecque δ et est appelée *symbole de Kronecker* [1]. Ce symbole est d'un emploi très pratique et nous en ferons un usage fréquent.

1. En l'honneur du mathématicien allemand Léopold Kronecker (1823-1891).

Définition 1.4 *Le symbole de Kronecker, qui concerne deux indices i et j, et qui est noté indifféremment $\delta_{ij} = \delta^{ij} = \delta^i_j = \delta^j_i$, prend la valeur 0 si $i \neq j$ et la valeur 1 si $i = j$.*

> ***Remarque concernant les notations :*** Le symbole de Kronecker doit être considéré comme une convention d'écriture. Plus loin dans l'ouvrage, nous définirons un objet (tenseur) très particulier dont les composantes sont numériquement égales à celles du symbole de Kronecker, mais dont la signification va au-delà d'une simple convention d'écriture (voir la section 5.4).

Le symbole de Kronecker montre tout son intérêt lorsqu'il apparaît dans une sommation. En pratique, quand le symbole δ^j_i apparaît dans une somme multiple qui court entre autres sur les indices i et j, on peut faire disparaître une des sommes – par exemple la somme sur i – et remplacer partout dans le terme générique l'indice i par l'indice j. Par exemple, la somme $a_{ik}\,\delta^i_j$ se transforme en l'élément unique a_{jk}. Les indices j et k étant libres dans ce cas, on les retrouve bien dans l'expression finale.

1.3 SYMBOLE DE LEVI-CIVITA

Une autre facilité d'écriture fait intervenir le *symbole de Levi-Civita*[2], ne prenant que 3 valeurs possibles, mais qui peut dépendre d'un nombre arbitraire d'indices. Avant de le définir proprement, disons un mot sur les permutations de n entiers.

Soit $N = \{1, 2, \ldots, n\}$ l'ensemble des n premiers nombres entiers. On appelle permutation P sur N toute bijection de N sur lui-même. Autrement dit, P fait correspondre à N l'ensemble des n nombres $\{p_1, p_2, \ldots, p_n\}$ qui reproduit les n nombres $\{1, 2, \ldots, n\}$ dans un ordre différent. Il est bien connu que le nombre de permutations différentes sur N vaut $n! = 1 \times 2 \times \ldots \times n$. Il est parfois utile d'introduire des permutations spéciales, appelées *permutations circulaires*, qui reprennent la permutation originale en décalant la totalité des nombres d'une valeur donnée en conservant le même ordre, par exemple $\{p_1, \ldots, p_k, p_{k+1}, \ldots, p_n\} \to \{p_{k+1}, \ldots, p_n, p_1, \ldots, p_k\}$.

De plus l'ensemble des permutations sur N, muni de la loi de composition traditionnelle des applications, possède la structure de groupe, la permutation identité étant simplement $\{1, 2, \ldots, n\}$. Une caractéristique fondamentale d'une permutation est sa *signature*. Il existe plusieurs définitions équivalentes de cette notion. Nous adoptons la définition suivante :

Définition 1.5 *Soit P une permutation définie par $\{p_1, p_2, \ldots, p_n\}$. Considérons toutes les paires (p_i, p_j) possibles et comptons le nombre r de celles-ci qui inversent l'ordre naturel (c'est-à-dire telles que si $i < j$ alors $p_i > p_j$). La signature de P,*

2. En hommage au mathématicien italien Tullio Levi-Civita (1873-1941).

notée σ_P, est le nombre $(-1)^r$. Une permutation de signature +1 *est dite paire, tandis qu'une permutation de signature* -1 *est dite impaire.*

> **Exemple 1.1 :** *(Signature des permutations d'ordre 6)* Avec $n = 6$, la permutation $P = \{2, 5, 3, 1, 6, 4\}$ présente 6 inversions (sur les 15 paires testées), à savoir (2, 1), (5, 3), (5, 1), (5, 4), (3, 1), (6, 4), et sa signature est $\sigma_P = (-1)^6 = +1$; la permutation est paire. La permutation $Q = \{1, 3, 4, 5, 2, 6\}$ montre, elle, 3 inversions, à savoir (3, 2), (4, 2), (5, 2), et sa signature est $\sigma_Q = (-1)^3 = -1$; la permutation est impaire.

Une propriété remarquable (que nous ne démontrons pas) des signatures peut être résumée dans l'équation $\sigma_{PQ} = \sigma_P \sigma_Q$: la signature du produit de deux permutations est égal au produit de leurs signatures. Cela implique en particulier qu'il y a autant de permutations paires que de permutations impaires dans N, à savoir $n!/2$. De plus, une permutation paire effectuée sur une permutation donnée ne change pas la signature de cette permutation.

Nous sommes à présent armés pour définir le symbole de Levi-Civita.

Définition 1.6 *Soit un ensemble de n indices p_i, chaque indice pouvant varier de* 1 *à n. On appelle symbole de Levi-Civita le nombre dépendant des n indices et noté indifféremment*[3] $\epsilon_{p_1 p_2 \ldots p_n}$ *ou* $\epsilon^{p_1 p_2 \cdots p_n}$ *dont la valeur vaut :*

- 0 *si deux indices p_l et p_m sont égaux ;*
- +1 *si la permutation $\{p_1, p_2, \ldots, p_n\}$ est paire ;*
- -1 *si la permutation $\{p_1, p_2, \ldots, p_n\}$ est impaire.*

1.4 ÉLÉMENTS DE CALCUL MATRICIEL

Le calcul matriciel est l'outil de base pour toutes les propriétés de l'algèbre linéaire. Nous le supposerons connu et ne ferons ici que des rappels utiles à notre propos, en précisant les notations en usage dans cet ouvrage.

1.4.1 Opérations sur les matrices

Définition 1.7 *Une matrice A de type $m \times n$ est un ensemble de mn nombres $(A)_{ij}$ (ou A_{ij} si aucune ambiguïté n'est permise) arrangés en un tableau rectangulaire de m lignes et de n colonnes. Le nombre ou élément $(A)_{ij}$ se trouve placé à la croisée de la ligne n^o i et de la colonne n^o j.*

3. Dans l'exercice 10.14, nous verrons que, dans un espace pseudo-euclidien, il faut veiller à imposer une bonne cohérence de signe aux tenseurs bâtis sur le symbole de Levi-Civita.

La matrice A se représente sous la forme

$$A = \begin{pmatrix} A_{11} & \dots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{m1} & \dots & A_{mn} \end{pmatrix}. \tag{1.5}$$

Les éléments A_{ij} peuvent être des nombres réels ou des nombres complexes ; on dit dans ce cas que la matrice est réelle ou complexe. Dans cet ouvrage, nous traiterons exclusivement des problèmes faisant intervenir des matrices réelles. Des matrices particulières sont : les *matrices lignes* constituées d'une seule ligne ($m = 1$) ; les *matrices colonnes* constituées d'une seule colonne ($n = 1$). Quand le nombre de lignes d'une matrice est égal au nombre de colonnes $m = n$, on dit qu'il s'agit d'une *matrice carrée* $n \times n$ ou d'ordre n.

> ***Remarque concernant les notations :*** Les indices i et j accolés à l'élément de matrice A_{ij} ou $(A)_{ij}$ n'ont pas de caractère de variance particulier. Ils sont toujours placés en bas et indiquent simplement la position de l'élément dans la matrice.

Dans l'ensemble des matrices $m \times n$, il est possible de définir une addition, $C = A + B$, telle que $(C)_{ij} = (A)_{ij} + (B)_{ij}$, et une multiplication par un scalaire, $B = \lambda\, A$, telle que $(B)_{ij} = \lambda\, (A)_{ij}$. On peut également définir le produit de deux matrices : la multiplication d'une matrice A de type $m \times n$ par une matrice B de type $n \times p$ est une matrice $C = AB$ de type $m \times p$, telle que $(C)_{ij} = \sum_{k=1}^{n} (A)_{ik}\, (B)_{kj}$. La matrice carrée I d'éléments $(I)_{kl} = \delta_{kl}$ possède la propriété $IA = AI = A$, $\forall\, A$ carrée. On l'appelle la matrice identité ou la matrice unité ; c'est l'élément neutre pour la multiplication des matrices carrées.

À partir d'une matrice donnée A, on peut construire d'autres matrices associées. Dans la suite nous emploierons essentiellement la notion de matrice transposée.

Définition 1.8 *Soit une matrice A de type $m \times n$. La matrice transposée de A, notée $\widetilde{A}$ et de type $n \times m$, est la matrice obtenue en permutant les lignes et les colonnes, c'est-à-dire dont les éléments sont $(\widetilde{A})_{ij} = (A)_{ji}$.*

On démontre facilement la propriété

$$(\widetilde{AB \dots Q}) = \widetilde{Q} \dots \widetilde{B}\widetilde{A}. \tag{1.6}$$

La matrice transposée d'une matrice colonne est une matrice ligne, et réciproquement.

1.4.2 Trace et déterminant

Dans le reste de cette section, on se bornera à considérer des matrices carrées d'ordre n. Pour deux matrices quelconques A et B, on a $AB \neq BA$ en général (bien que A et B soient de même type). Un élément du type $(A)_{ii}$ est dit *diagonal* et l'ensemble de ces éléments forme la *diagonale* de la matrice A. On va définir deux applications importantes de ces matrices dans $\mathbb{R}$: la trace et le déterminant.

Définition 1.9 *La trace d'une matrice A est la somme de tous ses éléments diagonaux :* $\mathrm{tr}\,(A) = \Sigma_{i=1}^{n}(A)_{ii}$.

On a de façon évidente l'égalité $\mathrm{tr}\,(\widetilde{A}) = \mathrm{tr}\,(A)$. Il est aussi facile de montrer pour une **permutation circulaire** (mais pas pour une permutation quelconque !)

$$\mathrm{tr}\,(AB \ldots FGH \ldots Q) = \mathrm{tr}\,(GH \ldots QAB \ldots F). \tag{1.7}$$

Le déterminant est également une notion fondamentale d'algèbre linéaire.

Définition 1.10 *Le déterminant d'une matrice A, noté $|A|$, est le nombre réel défini par*

$$|A| = \epsilon^{p_1 p_2 \cdots p_n}(A)_{1p_1}(A)_{2p_2} \ldots (A)_{np_n}. \tag{1.8}$$

> ***Remarque concernant les notations :*** La notation $|A|$ est utilisée pour désigner le déterminant de la matrice A. Si la grandeur entre les deux barres verticales est un nombre a, l'expression $|a|$ signifie alors la valeur absolue de ce nombre. Le contexte sera toujours suffisamment explicite pour éviter tout risque de confusion. Quant à la notation $|\boldsymbol{u}|$, elle désigne la norme du vecteur $\boldsymbol{u}$.

Une définition de grande importance est celle de *cofacteur*.

Définition 1.11 *Soit une matrice A d'éléments $(A)_{ij}$. On appelle mineur associé à l'élément $(A)_{ij}$, noté M_{ij}, le déterminant de la matrice obtenue de A en supprimant la ligne et la colonne contenant cet élément (autrement dit la ligne n^o i et la colonne n^o j). On appelle cofacteur ou complément algébrique de l'élément $(A)_{ij}$ le nombre $\alpha_{ij} = (-1)^{i+j} M_{ij}$.*

Le calcul d'un déterminant est souvent assez pénible (car il contient $n!$ monômes). La méthode systématique consiste à le développer par ligne ou par colonne en se servant des cofacteurs. On montre que l'on peut écrire

$$|A| = \Sigma_{k=1}^{n}(A)_{ik}\alpha_{ik} = \Sigma_{k=1}^{n}(A)_{ki}\alpha_{ki} \quad \forall\, i. \tag{1.9}$$

On a la propriété suivante :

$$|AB \ldots Q| = |A|\,|B| \ldots |Q|\,. \tag{1.10}$$

Une permutation paire des colonnes (ou des lignes) laisse le déterminant invariant, alors qu'une permutation impaire des colonnes (ou des lignes) change son signe. En utilisant le symbole de Levi-Civita et la formule (1.8), cette propriété peut s'écrire

$$\epsilon_{k_1 k_2 \ldots k_n}\,|A| = \epsilon^{p_1 p_2 \cdots p_n}(A)_{k_1 p_1}(A)_{k_2 p_2} \ldots (A)_{k_n p_n}. \tag{1.11}$$

Étant donné une matrice A telle que $|A| \neq 0$ (matrice *non singulière* ou *régulière*), il existe une matrice inverse **unique**, notée A^{-1}, qui vérifie

$$AA^{-1} = I = A^{-1}A. \tag{1.12}$$

Comme en algèbre élémentaire, on note $A^n = \underbrace{AA\ldots A}_{n \text{ fois}}$ la puissance n de la matrice A et $A^{-n} = (A^n)^{-1} = (A^{-1})^n$ la matrice inverse de A^n. Dans tous les cas, on a la propriété $A^m\, A^n = A^{m+n}$ et, en particulier, $A^0 = I$.

À chaque élément $(A)_{ij}$ de la matrice A, on peut associer un nombre qui est le cofacteur correspondant α_{ij}. Ces nombres, comme les éléments originaux, peuvent être arrangés dans une matrice α, appelée *matrice des cofacteurs* ou *comatrice*. Une propriété remarquable est que la matrice inverse s'exprime simplement à l'aide de la matrice des cofacteurs par la formule

$$A^{-1} = \frac{1}{|A|}\widetilde{\alpha}. \tag{1.13}$$

Notez que la matrice inverse est proportionnelle à la transposée de la comatrice. On vérifie facilement la propriété

$$(ABC\ldots FGH)^{-1} = H^{-1}G^{-1}F^{-1}\ldots C^{-1}B^{-1}A^{-1}. \tag{1.14}$$

1.5 DÉRIVÉES ET DIFFÉRENTIELLES

Tout au long du livre, nous aurons l'occasion d'utiliser abondamment les fonctions de variables réelles et leurs dérivées successives.

1.5.1 Fonction à une variable

Commençons par une fonction à une variable, qui est une application [4] de $\mathbb{R}$ sur $\mathbb{R}$: on la note $f(x)$, où x est la variable. Pour la valeur x la fonction vaut $f(x)$; lorsque la variable est modifiée et prend la valeur $x + \Delta x$, la fonction est aussi modifiée et devient $f(x + \Delta x)$. La variation de la fonction est notée $\Delta f = f(x + \Delta x) - f(x)$. Si la fonction est continue, lorsque $\Delta x \to 0$ on a aussi $\Delta f \to 0$. Pourtant, la plupart du temps, le rapport $\Delta f/\Delta x$ tend vers une valeur finie qui dépend de la fonction f et de la valeur de la variable x. Cette valeur est la *dérivée* de la fonction pour la variable x (ou au point x)

$$f'(x) = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x}. \tag{1.15}$$

4. Plus exactement le domaine de définition est un sous-ensemble de $\mathbb{R}$ et l'ensemble image aussi un sous-ensemble de $\mathbb{R}$. Bien que ces notions soient importantes, nous ne les aborderons pas ici.

La dérivée $f'(x)$ représente la pente de la tangente à la fonction $f(x)$ au point x. La fonction dérivée est généralement dérivable elle-même et on peut définir la dérivée seconde $f''(x)$ comme la dérivée de la fonction dérivée puis, par processus itératif, les *dérivées d'ordre supérieur*.

On appelle *différentielle* df au point x la variation de la fonction f pour une variation infinitésimale dx de la variable

$$df = f'(x)\,dx. \tag{1.16}$$

En utilisant la notation (1.16), on écrit très souvent les dérivées d'ordres supérieurs de la fonction f sous la forme

$$f'(x) = f^{(1)}(x) = \frac{df}{dx} \quad ; \quad f^{(n)}(x) = \frac{d^n f}{dx^n} = \frac{d}{dx} f^{(n-1)}(x). \tag{1.17}$$

On suppose connues les règles de dérivation des différentes opérations sur les fonctions (somme, produit, quotient, composition, etc.).

1.5.2 Fonctions à plusieurs variables

Très souvent, nous aurons à manipuler des fonctions dépendant de plusieurs variables (en nombre fini). Nous notons $f(x^1, x^2, \ldots, x^n)$ une fonction dépendant de n variables x^i ; c'est une application de $\mathbb{R}^n$ dans $\mathbb{R}$. Notons que les variables sont ici indiquées sous forme d'une suite contravariante.

La notion de dérivée se généralise sans peine à condition de préciser par rapport à quelle variable on prend la dérivée ; cela entraîne la définition de *dérivée partielle*

$$f'_{x^i}(x^1, x^2, \ldots, x^n) = \frac{\partial f(x^1, x^2, \ldots, x^n)}{\partial x^i} = \partial_i f(x^1, x^2, \ldots, x^n), \tag{1.18}$$

qui se lit « dérivée partielle de f par rapport à x^i » ; on la calcule en prenant la dérivée habituelle par rapport à la variable x^i, en considérant toutes les autres variables comme des paramètres fixés. La dernière notation pour la dérivée partielle, $\partial_i f$, est très condensée et très pratique ; l'indice covariant va nous permettre d'appliquer facilement la règle de convention d'Einstein. Cette notation sera pleinement justifiée plus loin (voir le chapitre 8).

Il est possible de définir les *dérivées partielles d'ordre supérieur*, mais à priori il faut être très prudent dans l'ordre des dérivations. Le miracle est que, pour des conditions sur f suffisamment générales pour qu'elles soient en pratique toujours satisfaites en physique, le résultat est indépendant de l'ordre. On dit que les **dérivations partielles commutent**. On note ces dérivées du second ordre

$$f''_{x^i x^j} = \frac{\partial^2 f}{\partial x^i \partial x^j} = \partial^2_{ij} f = \partial_i(\partial_j f) = \partial_j(\partial_i f). \tag{1.19}$$

La *différentielle totale* d'une fonction f s'écrit (utilité de la convention d'Einstein)

$$df = \frac{\partial f}{\partial x^1} dx^1 + \ldots + \frac{\partial f}{\partial x^n} dx^n = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i} dx^i = (\partial_i f) dx^i. \qquad (1.20)$$

Elle représente la variation de la fonction f au point $(x^1, \ldots, x^n)$ lorsque chaque variable x^i est augmentée de la valeur infinitésimale dx^i.

Définition 1.12 *Soit une suite contravariante quelconque $(x^1, x^2, \ldots, x^n)$, une suite contravariante infinitésimale (dx^1, dx^2,..., dx^n) et une suite covariante de fonctions $(P_1, P_2, \ldots, P_n)$, chaque fonction P_i étant une fonction des n variables x^k : $P_i(x^1, x^2, \ldots, x^n)$. On appelle forme différentielle linéaire la quantité*

$$\delta\omega = P_i dx^i. \qquad (1.21)$$

La différentielle totale d'une fonction df est bien une forme différentielle (voir la relation (1.20)) ; mais c'est une forme bien particulière puisque les fonctions P_i sont les dérivées partielles par rapport à x^i **d'une seule et même fonction** : $P_i = \partial_i f$. On peut se poser la question réciproque : sous quelles conditions une forme différentielle est-elle, **en tout point**, la différentielle totale d'une fonction : $\delta\omega = df$? La réponse est fournie par le théorème suivant, que nous énonçons sans démonstration :

Théorème 1.13 *Une condition nécessaire et suffisante pour qu'une forme différentielle $\delta\omega = P_i dx^i$ soit, en tout point, la différentielle totale d'une fonction df est que les fonctions P_i satisfassent en tout point les conditions suivantes :*

$$\partial_i P_j - \partial_j P_i = 0 \quad \forall i, j. \qquad (1.22)$$

EXERCICES

Les réponses à ces exercices sont données à la page 241.

Exercice 1.1 Sommation

Écrire explicitement les sommations suivantes dans le cas $n = 3$: $a^{ij} b_j$, $a^{kl} b_k c_l$, $a^i b_{ijk} c^j$.

Exercice 1.2 Propriétés du symbole de Kronecker

Calculer $\delta_i^j \delta_j^i$.

Exercice 1.3 Une propriété générale des sommations

Démontrer l'égalité $(a_{ij} - a_{ji})(s^{ij} + s^{ji}) = 0$.

Exercice 1.4 Déterminant et symbole de Levi-Civita

Montrer que les expressions connues des déterminants d'une matrice A d'ordre 2 et 3 peuvent s'écrire respectivement $\epsilon^{ij}(A)_{1i}(A)_{2j}$ et $\epsilon^{ijk}(A)_{1i}(A)_{2j}(A)_{3k}$ (règle de Sarrus), en accord avec la formulation générale (1.8).

Exercice 1.5 Symboles de Levi-Civita et de Kronecker

Démontrer la relation $\epsilon^{i_1 i_2 \ldots i_n}\epsilon_{i_1 i_2 \ldots i_n} = n!$. Dans le cas $n = 3$: démontrer les relations $\epsilon^{ijk}\epsilon_{ijm} = 2\delta^k_m$ et $\epsilon^{ijk}\epsilon_{lmk} = \delta^i_l\delta^j_m - \delta^i_m\delta^j_l$; au moyen de ces deux dernières égalités, vérifier que $\epsilon^{ijk}\epsilon_{ijk} = 6$.

Exercice 1.6 Jeux sur les matrices

Soit A une matrice ligne à n éléments A_i, B une matrice colonne à n éléments B_j et C une matrice carrée d'ordre n d'éléments C_{ij}. Donner le type des matrices ABC, ACB, BAC, BCA, CAB, CBA, ainsi que leurs éléments de matrice.

Exercice 1.7 Puissance d'une matrice particulière

On donne les deux matrices 3×3 suivantes

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \quad ; \quad B = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

1. Vérifier que $A = B + I$, où I est la matrice identité.
2. Calculer B^2 puis B^3.
3. En déduire la valeur de A^n.
4. Vérifier ce résultat au moyen d'une démonstration par récurrence.

Exercice 1.8 Commutateurs de matrices

Étant donné deux matrices carrées régulières A et B, on appelle commutateur de A et B la matrice $[A, B] = AB - BA$. Lorsque deux matrices commutent, leur commutateur est identiquement nul. Montrer les propriétés suivantes (A, B, C sont des matrices carrées régulières et λ est un nombre réel)

1. $[A, A] = 0$.
2. $[B, A] = -[A, B]$.
3. $[A, B + C] = [A, B] + [A, C]$.
4. $[A, \lambda B] = \lambda[A, B]$.
5. $[A, BC] = [A, B]C + B[A, C]$.
6. $[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$ (Identité de Jacobi).
7. Montrer que les trois matrices

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad , \quad B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad , \quad C = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

   satisfont les relations de commutation $[A, B] = C$, $[A, C] = 0$, $[B, C] = 0$. Vérifier sur elles l'identité de Jacobi.

Exercice 1.9 Propriétés des cofacteurs

Soit A une matrice régulière et α la matrice des cofacteurs. Vérifier les propriétés :

1. $\sum_k (A)_{ki}(\alpha)_{ki} = |A| = \sum_k (A)_{ik}(\alpha)_{ik}$ (développement en colonne et ligne).
2. $\sum_k (A)_{ki}(\alpha)_{kj} = 0 = \sum_k (A)_{ik}(\alpha)_{jk}$ si $i \neq j$. On pourra, par exemple, remarquer que cette expression correspond à un déterminant contenant deux lignes identiques.
3. En se basant sur le résultat des questions précédentes, en déduire l'expression (1.13) pour l'inverse de la matrice A.

Chapitre 2

Espaces vectoriels

Pour décrire les phénomènes naturels qui se déroulent dans l'espace ordinaire – ou l'espace-temps en relativité – le physicien utilise des quantités dynamiques qui peuvent être des grandeurs définies par un seul nombre et qu'on appelle des *scalaires* comme la température, la pression, la masse volumique, etc., ou des grandeurs caractérisées par plusieurs nombres et qu'on appelle des *vecteurs* comme la vitesse, l'accélération, la force, etc.

Les mathématiciens ont exploité la notion de vecteur introduite par les physiciens pour la généraliser et définir la structure *d'espace vectoriel* qui est à la base de toute la matière de cet ouvrage. Il n'est pas déraisonnable de lui consacrer un chapitre qui sera essentiellement un rappel de propriétés que nous supposons connues du lecteur.

2.1 VECTEURS DE L'ESPACE ORDINAIRE

Vous êtes sans doute familiers avec les vecteurs de l'espace ordinaire. Un vecteur de cet espace est habituellement noté par un symbole surmonté d'une flèche, par exemple $\vec{A}$. Ces vecteurs peuvent avoir des significations physiques (et souvent des unités) différentes : position, vitesse, accélération, force, champ électrique ou magnétique, etc. Ils sont caractérisés par une direction et un sens, ainsi que par un module. On représente graphiquement un vecteur par un segment de droite entre deux points M et N, orienté à l'aide d'une flèche partant de *l'origine* M et allant vers *l'extrémité* N. La direction du vecteur correspond à la droite MN, appelée le *support* du vecteur. Le sens va de M à N. Le module est la longueur du segment par rapport à une longueur choisie comme unité. Un tel vecteur est noté $\vec{A} = \overrightarrow{MN}$ et son module est noté $A = |\vec{A}| = |MN|$.

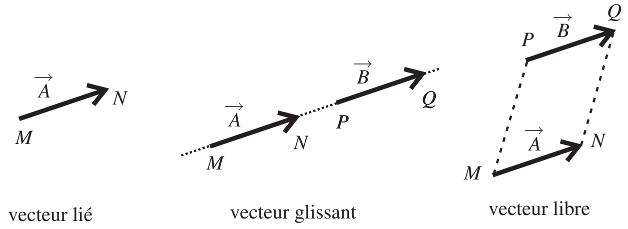

Figure 2.1 *Vecteurs lié, glissant et libre.*

Dans l'espace ordinaire, on distingue habituellement trois types de vecteurs : les *vecteurs liés* qui possèdent une origine et une extrémité fixes, les *vecteurs glissants* qui ont même direction, même sens et même module, mais dont l'origine a la liberté d'être n'importe où sur le support, et les *vecteurs équipollents* ou *vecteurs libres* qui ont même direction (ils sont sur des supports parallèles), même sens et même module mais dont l'origine peut être n'importe où dans l'espace (voir la figure 2.1). L'équipollence de deux vecteurs est une relation d'équivalence pour les vecteurs liés ; en général on désigne une classe d'équivalence comme un vecteur unique [1]. C'est cette dernière interprétation que nous utiliserons par la suite. Dans ce cadre, deux vecteurs $\vec{A}$ et $\vec{B}$ sont dits égaux $\vec{A} = \vec{B}$ s'ils ont des directions, des sens et des modules identiques. Géométriquement, l'égalité entre les vecteurs $\vec{A} = \overrightarrow{MN}$ et $\vec{B} = \overrightarrow{PQ}$ signifie qu'ils sont représentés par les deux côtés opposés d'un parallélogramme $MNQP$ (voir le dessin de droite sur la figure 2.1).

Par définition le vecteur nul, noté $\vec{0}$, possède un module égal à 0 ; autrement dit son extrémité coïncide avec son origine. Le vecteur opposé à $\vec{A}$ est noté $-\vec{A}$. Il est défini par rapport au vecteur $\vec{A}$ comme le vecteur dont on a changé le sens ou, de façon équivalente, dont on a interverti origine et extrémité $-\overrightarrow{MN} = \overrightarrow{NM}$.

On définit pour les vecteurs de l'espace ordinaire deux types d'opérations : l'addition et la multiplication par un scalaire.

– l'addition de deux vecteurs $\vec{A}$ et $\vec{B}$, notée $\vec{C} = \vec{A} + \vec{B}$, est définie par la *règle du parallélogramme* (voir la figure 2.2). De façon explicite, on choisit un vecteur $\overrightarrow{MN}$ équipollent à $\vec{A}$ d'origine M et d'extrémité N, puis un vecteur $\overrightarrow{NP}$ équipollent à $\vec{B}$ d'origine N et d'extrémité P. Alors le vecteur somme $\vec{C}$ est équipollent au vecteur $\overrightarrow{MP}$. L'égalité $\overrightarrow{MP} = \overrightarrow{MN} + \overrightarrow{NP}$ est appelée la *relation de Chasles* sur les vecteurs.
  L'addition de deux vecteurs est commutative, $\vec{A} + \vec{B} = \vec{B} + \vec{A}$, est associative, $\vec{A} + (\vec{B} + \vec{C}) = (\vec{A} + \vec{B}) + \vec{C}$, possède un élément neutre $\vec{0}$, $\vec{A} + \vec{0} = \vec{A}$, et affecte un vecteur opposé $-\vec{A}$ à chaque vecteur $\vec{A}$, $\vec{A} + (-\vec{A}) = \vec{0}$.

1. Cette propriété fait le lien entre espace vectoriel et espace de points, et sera développée en détail dans le chapitre 7.

L'addition au vecteur $\vec{B}$ de l'opposé du vecteur $\vec{A}$, s'appelle la soustraction et on note volontiers $\vec{B} - \vec{A} = \vec{B} + (-\vec{A})$. Sur le parallélogramme $MNPQ$, la première diagonale correspond à la somme $\vec{A} + \vec{B} = \overrightarrow{MP}$, tandis que la seconde diagonale correspond à la différence $\vec{B} - \vec{A} = \overrightarrow{NQ}$.

– À chaque vecteur $\vec{A}$, on peut affecter un vecteur unique $\vec{B} = a\vec{A}$, multiplication du vecteur $\vec{A}$ par le scalaire (nombre réel) a de la façon suivante. Le vecteur $\vec{B}$ a même support et même sens que $\vec{A}$ si $a > 0$, le sens opposé si $a < 0$, et son module a été multiplié par la valeur absolue $|a|$: $|a\vec{A}| = |a|\,|\vec{A}|$. Il est facile de vérifier la propriété $(ab)\vec{A} = a(b\vec{A}) = b(a\vec{A})$. De plus, la multiplication par un scalaire et l'addition sont respectivement distributives l'une par rapport à l'autre. Cela se traduit par les égalités $(a + b)\vec{A} = a\vec{A} + b\vec{A}$ et $a(\vec{A} + \vec{B}) = a\vec{A} + a\vec{B}$.

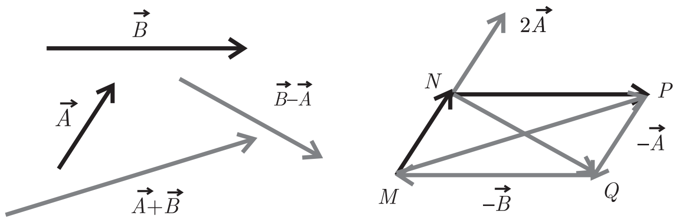

Figure 2.2 *Somme et différence de vecteurs.*

Choisissons trois vecteurs de l'espace ordinaire $\vec{e}_1, \vec{e}_2, \vec{e}_3$ tels que l'un quelconque ne soit pas contenu dans le plan des deux autres. La propriété d'un vecteur $\vec{X}$ quelconque est qu'il s'exprime de **façon unique** comme combinaison linéaire de ces trois vecteurs : $\vec{X} = x^1\vec{e}_1 + x^2\vec{e}_2 + x^3\vec{e}_3$. Les nombres x^i s'appellent les *composantes* ou les *projections parallèles* de $\vec{X}$ sur les vecteurs $\vec{e}_i$. La décomposition sur ces vecteurs étant unique, on dit que l'espace ordinaire est de dimension 3 ; il est équivalent de raisonner sur la quantité vectorielle $\vec{X}$ ou sur le triplet de nombres réels (x^1, x^2, x^3). L'utilisation des composantes est d'usage très courant car souvent plus pratique. Ainsi l'addition vectorielle, qui est définie géométriquement de façon un peu compliquée par la règle du parallélogramme, est définie de façon équivalente et beaucoup plus simple par l'égalité entre composantes $(x^1, x^2, x^3) + (y^1, y^2, y^3) = (x^1 + y^1, x^2 + y^2, x^3 + y^3)$; de même la multiplication par un scalaire se traduit par l'égalité $a(x^1, x^2, x^3) = (ax^1, ax^2, ax^3)$.

Dans le reste de ce chapitre nous généralisons cette notion intuitive de vecteur à d'autres espaces avec des dimensions arbitraires.

2.2 DÉFINITIONS ET PREMIÈRES PROPRIÉTÉS

Définition 2.1 *Soit V un ensemble non vide – dont les éléments sont notés par un caractère gras* $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$, *etc. – muni de deux lois : une loi de composition interne,* *l'addition, notée +, qui associe à deux éléments $\boldsymbol{u}$ et $\boldsymbol{v}$ de V un troisième élément $\boldsymbol{w} = \boldsymbol{u} + \boldsymbol{v}$ de V, et une loi de composition externe, la multiplication par un scalaire (nombre réel) indiquée par une absence de symbole qui associe à un élément a de $\mathbb{R}$ et un élément $\boldsymbol{u}$ de V un élément unique $a\boldsymbol{u}$ de V.*

V est appelé espace vectoriel sur $\mathbb{R}$ – et ses éléments des vecteurs – si les axiomes suivants sont vérifiés (plusieurs présentations sont possibles) :

Addition vectorielle

A_1 – *Commutativité :* $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$, $\forall\, \boldsymbol{u}, \boldsymbol{v} \in V$.

A_2 – *Associativité :* $\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$, $\forall\, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$.

A_3 – *Existence d'un élément neutre* $\mathbf{0}$ *:* $\boldsymbol{u} + \mathbf{0} = \mathbf{0} + \boldsymbol{u} = \boldsymbol{u}$, $\forall\, \boldsymbol{u} \in V$.

A_4 – *Chaque élément* $\boldsymbol{u}$ *possède un élément inverse ou opposé noté* $-\boldsymbol{u}$ *:* $\boldsymbol{u} + (-\boldsymbol{u}) = (-\boldsymbol{u}) + \boldsymbol{u} = \mathbf{0}$, $\forall\, \boldsymbol{u} \in V$.

Multiplication par un scalaire

B_1 – *Associativité :* $(ab)\boldsymbol{u} = a(b\boldsymbol{u})$, $\forall\, a, b \in \mathbb{R}$, $\forall\, \boldsymbol{u} \in V$.

B_2 – *Distributivité par rapport à l'addition vectorielle :* $a(\boldsymbol{u} + \boldsymbol{v}) = a\boldsymbol{u} + a\boldsymbol{v}$, $\forall\, a \in \mathbb{R}$, $\forall\, \boldsymbol{u}, \boldsymbol{v} \in V$.

B_3 – *Distributivité par rapport à l'addition scalaire :* $(a + b)\boldsymbol{u} = a\boldsymbol{u} + b\boldsymbol{u}$, $\forall\, a, b \in \mathbb{R}$, $\forall\, \boldsymbol{u} \in V$.

B_4 – *Rôle du scalaire* 1 *:* $1\boldsymbol{u} = \boldsymbol{u}$, $\forall\, \boldsymbol{u} \in V$.

Notons que les 4 premiers axiomes A_1-A_4 confèrent à V muni de l'addition vectorielle une structure de groupe commutatif. À strictement parler, il faudrait faire une distinction entre le signe typographique représentant l'addition de deux nombres réels et celui représentant l'addition de deux vecteurs ; de même il faudrait distinguer la multiplication de réels et la multiplication de vecteurs par des réels. L'utilisation du même signe typographique pour désigner des lois différentes est de pratique courante ; elle ne pose pas de problème et devient tellement naturelle qu'on a tendance à oublier ce point.

Nous avons défini un espace vectoriel sur le corps des nombres réels $\mathbb{R}$; il est possible d'être plus général et de le définir sur le corps des nombres complexes $\mathbb{C}$, ou même sur un corps commutatif quelconque K. Une telle généralisation ne sera pas nécessaire dans cet ouvrage et on se contentera d'utiliser le corps des nombres réels.

À cause du caractère associatif de l'addition, il n'est pas utile de mettre des parenthèses dans une somme multiple ; à cause du caractère commutatif, l'ordre dans une telle somme est sans importance. Il est d'usage de parler de *soustraction* lorsqu'on ajoute l'opposé d'un vecteur et on utilise la notation simplifiée suivante $\boldsymbol{u} + (-\boldsymbol{v}) = \boldsymbol{u} - \boldsymbol{v}$.